2000 Basic Energy Sciences Metals and Ceramics Sciences Journal Publications

Alloy Theory

Atomic Mechanisms in Interface Sciences and; Semiconductor Nanocrystals for CO_2 Fixation

Radiation Effects

Laves Phase Intermetallics

Science of Non-Equilibrium Processing of Materials

High-Temperature Alloy Design

Microstructural Design of Advanced Ceramics

KC 02 01 01 0
ERKCM01 (Alloy Theory)

Refereed Journal Publications

*Becher, P. F., G. S. Painter, E. Y. Sun, C. H. Hsueh, and M. J. Lance, "The Importance of Amorphous Intergranular Films in Self-Reinforced Si_3N_4 Ceramics," *Acta Mater.* **48**, 4493-99 (2000).

†Butler, W. H., X.-G. Zhang, and J. M. MacLaren, "Semiclassical Theory of Spin-dependent Transport in Magnetic Multilayers," *Journal of Superconductivity and Novel Magnetism* **13**, 221-38 (2000).

Butler, W. H., X.-G. Zhang, and J. M. MacLaren, "Solution to the Boltzmann Equation for Layered Systems for Current Perpendicular to the Planes," *J. Appl Phys*. **87**, 5173 (2000).

Fu, C. L. and Xindong Wang, "Thermal Expansion Coefficients of Mo-Si Compounds by First-Principles Calculations," *Phil. Mag. Lett.* **80**, 683-90 (2000).

*Hong, S.-L., C. L. Fu, and M. H. Yoo, "First-Principles Calculation of Stacking Fault and Twin Boundary Energies of Cr_2Nb," *Phil. Mag. A* **80**, 871-80 (2000).

Kowalewski, M.,W. H. Butler, N. Moghadam, G. M. Stocks, T. C. Schulthess, K. J. Song, J. R. Thompson, A. S. Arrott, T. Zhu, J. Drewes, R. R. Katti, M. T. McClure, and O. Escorcia, "The Effect of Ta on the Magnetic Thickness of Permalloy ($Ni_{81}Fe_{19}$) Films," *J. Appl. Phys.* **87 (Part 2)**, 5732-34 (2000).

Moghadam, N. Y., G. M. Stocks, B. Ujfalussy, W. A. Shelton, A. Gonis, and J. S. Faulkner, "Application of the Screened Korringa-Kohn-Rostoker Method," *J. Phys. Condens. Matter.* **11**, 5505-15 (1999).

*Nicholson, D. M. C., P. A. Sterne, J. C. Swihart, J. Tran, and Y. Wang, "Electron and Positron States in Amorphous $Ni_{0.4}Pd_{0.4}P_{0.2}$," *Model. Simul. Mater. Sci. Eng*. **8**, 261 (2000).

Orgassa, D., H., Fujiwara, T. C. Schulthess, and W. H. Butler, "Disorder Dependence of the Magnetic Moment of the Half-Metallic Ferromagnet NiMnSb From First Principles," *J. Appl. Phys*. **87**, 5870 (2000).

Petit L., S. V. Beiden, W. M. Temmerman, Z. Szotek, G. M. Stocks, and G. A. Gehring, "Screened Real-Space Korringa-Kohn-Rostoker Description of the Relativistic and Magnetic Properties of Transition Metals," *J. Phys. Condens Mat.* **12**(39), 8439-54 (2000).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

Schulthess, T. C. and W. H. Butler, "Magnetostatic Coupling in Spin Valves: Revisiting Neel's Formula," *J. Appl. Phys*. **87**, 5759 (2000).

†Stocks, G.M., V. F. Los, S. P. Repetsky, and T. D. Shatnii, "Short-Range and Long-Range Magnetic Order and Phase Diagram of Systems with Strong Electron Correlation," *Metallofiz. Nov. Tekh.* **22**(4), 3-12 (2000).

Ujfalussy, B., J. S. Faulkner, N. Y. Moghadam, G. M. Stocks, and Y. Wang, "Calculating Properties With the Polymorphous Coherent Potential Approximation," *Phys. Rev. B* **61**, 12005 (2000).

William, J. J., Y. Y. Ye, M. J. Kramer, K. M. Ho, L. Hong, C. L. Fu, and S. K. Malik, "Theoretical Calculations and Experimental Measurements of the Structure of Ti_5Si_3 and $Ti_5Si_3Z_{0.5}$ (Z=B, C, N or O)," *Intermetallics* **8**, 917-43 (2000).

*Zhang, X.-G. and W. H. Butler, "A Simple Treatment of the 'Scattering-In' Term of the Boltzmann Equation for Multilayers," *J. Appl. Phys.* **87**, 5176 (2000).

*Zhang, X.-G. and D. M. C. Nicholson, "A New Density Functional for Spherical Potentials," *Phys. Rev. B* **60**, 4551 (1999).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

Pergamon

Acta mater. 48 (2000) 4493–4499

Acta MATERIALIA

www.elsevier.com/locate/actamat

THE IMPORTANCE OF AMORPHOUS INTERGRANULAR FILMS IN SELF-REINFORCED Si_3N_4 CERAMICS

P. F. BECHER*, G. S. PAINTER, E. Y. SUN‡, C. H. HSUEH and M. J. LANCE

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Abstract—High-fracture-strength and high-toughness β-Si_3N_4 ceramics can be obtained by tailoring the size and number of the elongated bridging grains. However, these bridging mechanisms rely on debonding of the reinforcing grains from the matrix to increase toughness. Interfacial debonding is shown to be influenced by sintering aids incorporated in the amorphous intergranular films. In one case, the interface strength between the intergranular glass and the reinforcing grains increases with the aluminum and oxygen content of an interfacial epitaxial β-SiAlON layer. In another, the incorporation of fluorine in the intergranular film allows the crack to circumvent the grains. Atomic cluster calculations reveal that these two debonding processes are related to (1) strong Si–O and Al–O bonding across the glass/crystalline interface with an epitaxial SiAlON layer and (2) a weakening of the amorphous network of the intergranular film when difluorine substitutes for bridging oxygen. *© 2000 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.*

Keywords: Ceramics; Interfaces; Microstructure; Toughness; Cluster calculations

1. BACKGROUND

There have been a number of observations of increases in the mechanical performance of silicon nitride ceramics associated with the formation of elongated grains; for example, the early studies by Himsolt *et al*. [1] and Lange [2]. Kawashima and co-workers [3] and Mitomo [4] provided evidence to show that the fracture toughness increased with the diameter of these elongated grains. More recent studies by Hirao and co-workers [5, 6] and Emoto and Mitomo [7] have shown that beta seeds can be used to control the size and volume fraction of elongated grains in these so-called self-reinforced ceramics. The toughening effect arises from deflection of the crack tip by these elongated grains and by subsequent generation of bridging grains in the wake of the crack tip, similar to what occurs in both whisker-reinforced [8–11] and fiber-reinforced [12] ceramics. The crack-bridging reinforcements dissipate considerable strain energy through friction generated by their motion against the matrix during elastic stretching along the debonded interface and pull-out. Thus, increases in the number, diameter and debonded length of such bridging reinforcements should enhance the toughening effects [8–11].

In the self-reinforced silicon nitride ceramics, this involves the controlled incorporation of large (>1 μm) diameter elongated grains [13, 14]. For instance, a bimodal (~0.5 and ~2 μm) diameter distribution of the elongated grains yielded strengths >1 GPa and toughness values >10 MPa $m^{1/2}$, while a uniform ~0.5 μm equiaxed grain size resulted in a strength of ~600 MPa and a toughness of ~3.5 MPa $m^{1/2}$ [14]. Recent work has shown that a distinct bimodal diameter distribution of the elongated grains may not be required to achieve similar toughening effects [15]. In this work, a silicon nitride with a monomodal diameter distribution (average ~1 μm) and one with a bimodal diameter distribution (~0.6 and ~2 μm) exhibited toughness values of ~10 MPa $m^{1/2}$, each with similar area fraction of reinforcing grains. The high toughness achieved with the smaller (~1 μm) reinforcing grains can probably be attributed to a greater degree of crystallographic texture of the reinforcing grains, since when the textures of the two microstructures are similar, the toughness is lower in the material with the smaller diameter reinforcing grains. This points out that orienting the reinforcing grains so they actively participate in the crack-bridging process can contribute to the level of toughness obtained. The effects of the diameter and texture of the reinforcing grains on the toughness are

* To whom all correspondence should be addressed. *E-mail address:* Becherpf@ornl.gov (P.F. Becher)

‡ Now with United Technologies Research Center, East Hartford, CT, USA.

1359-6454/00/$20.00 © 2000 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.
PII: S1359-6454(00)00236-6

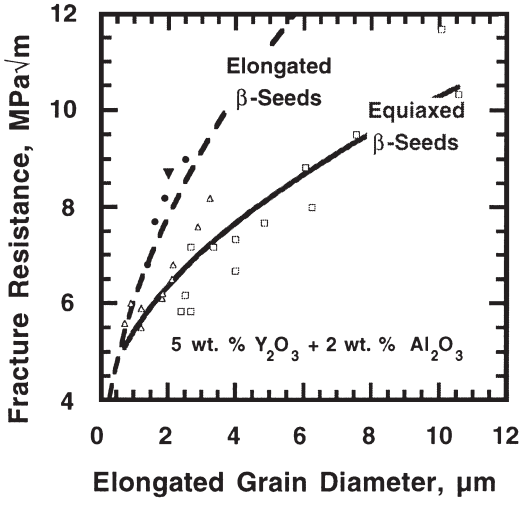


Fig. 1. The toughness of silicon nitride can be scaled with the diameter of the elongated reinforcing grains. The toughening effect in ceramics with randomly oriented reinforcements (equiaxed seeds, [3, 4]) can be increased by enhancing the alignment of these grains (elongated seeds, [6, 14]).

illustrated in Fig. 1. In summary, both toughening models and experiments reveal that tailoring the microstructure is a versatile method for developing substantially enhanced fracture properties, but only if attention is paid to the microstructural parameters affecting the bridging contributions.

Against this background, we must emphasize that, for these toughening mechanisms to be operative, it is imperative that the interface between the matrix and the reinforcing grains debond (or separate), Fig. 2. This debonding process must occur so that the crack tip is deflected along the grain face, rather than cutting through the elongated grain, while leaving intact elongated grains to bridge the crack in its wake. In the case of silicon nitride ceramics the microstructure consists of an amorphous silicate-based intergranular film (IGF) approximately 1 to 1.5 nm thick that surrounds all Si_3N_4 grains, including the elongated grains. The question then is how to tailor the structure or chemistry of the interfacial region to optimize the debonding process and, hence, the fracture toughness of these ceramics.

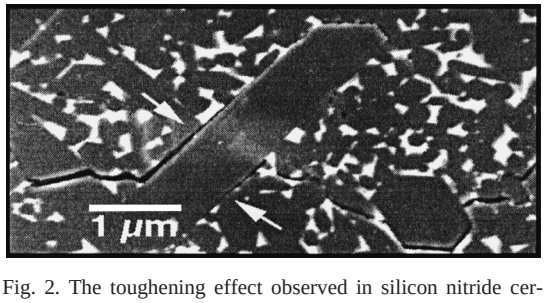


Fig. 2. The toughening effect observed in silicon nitride ceramics is associated with elongated grains that bridge the crack in the tip wake. Bridging by the grain occurs when the interface between it and the matrix separates or debonds (arrows). Each silicon nitride grain is surrounded by an amorphous intergranular film (white).

2. INTERFACIAL DEBONDING PROCESSES

Based on recent observations there appear to be at least two interfacial debonding processes in silicon nitride ceramics. In one, the introduction of appropriate anions generates crack-like defects within the amorphous intergranular film (IGF) (i.e., weakening the intergranular film by "crack formation in the IGF") [16]. In the other process, the strength of the interface between the amorphous film and the elongated grains is found to depend on the growth of an epitaxial SiAlON layer at the interface—"interfacial strengthening" [17, 18]. The intent here is to focus discussion on the experimental and theoretical aspects of these two chemical/structural effects.

It is recognized that either of these debonding processes may be affected by the generation of thermal expansion mismatch stresses between the intergranular phase(s) and the silicon nitride. These stresses arise because the coefficients of thermal expansion of the amorphous IGF films are at least two times greater than those of the anisotropic beta grains. As a result, the interface along prismatic grain surfaces will be subjected to radial compression and the film will be subjected to tension. These radial compressive stresses could inhibit interfacial debonding along these critical interfaces. Earlier investigations showed that the radial compressive stresses are, at best, a secondary factor in the presence of SiAlON epitaxial layers [17]. When SiAlON is not present at the interface, the extent of debonding at the Si_3N_4/glass interface is diminished by increases in the radial compressive stress [18]. In addition, the local tensile stresses, including those due to thermal expansion mismatch, are seen to promote the growth of crack-like defects introduced in the amorphous films containing fluorine [16]. The local residual stresses would then be a key component in the generation and growth of cracks within the intergranular film and debonding at the interface with the reinforcing phase.

Clearly, work to understand the effects of additives on crack formation in the IGF and interfacial strengthening (as well as on the high-temperature properties of the ceramic system, such as viscosity) has entered the regime of nanoscale chemistry, focusing predominantly on atomic-level phenomena that manifest themselves at the macroscopic scale in often dramatic ways. This evolution to the nanoscale in efforts to better understand IGF behavior is evidenced in the number of detailed characterization studies, such as those involving high-resolution transmission electron microscopy and electron energy-loss spectroscopy (e.g., [19]). Atomistic theoretical studies with first-principles [20, 21] and molecular dynamics [22, 23] modeling are, then, natural complements in the interpretation of experiments at this scale, and

much is being learned about atomic-level behavior both at the interface and within the IGFs.

2.1. Crack formation in the IGF

The process of IGF crack formation involves the formation *of weak bonds within the amorphous intergranular film*, for example by the addition of an anion such as fluorine, as discussed by Pezzotti *et al.* [16]. When the local density of these weakened bonds is sufficiently high, a crack-like feature forms, which grows in the presence of either an applied stress, a residual thermal expansion mismatch stress, or a combination of these stresses. According to the description of Pezzotti *et al.* [16], insertion of fluorine in the amorphous phase results in a pair of fluorine ions (difluorine), required to maintain the local charge balance, substituting in the position formerly taken by the oxygen linking the SiO_4 tetrahedra in the intergranular film (see Fig. 3). The fluorine pair is considered to move the tetrahedra apart due to the Coulomb repulsion between the two F^- ions, thus dilating the film, consistent with high-resolution electron microscopy (HREM) observations [19]. As a result of fluorine doping, oxygen-bridge linkages between tetrahedra are lost with each being replaced by a defect (i.e., missing bond) whose number density increases with fluorine content.

For the present study, the following first-principles atomic cluster calculations were made to provide an atomic-level description of the effects of substitution of difluorine for oxygen, from which two important features emerge: (1) strong Si–F bond formation (as expected on the basis of valency considerations) and (2) very weak interaction between the tetrahedra joined by the F–F pair. Here, the partial-wave self-consistent-field method [25] was used to solve the local density equations [26] precisely within an atomic cluster chosen to capture the essentials of the problem of interest: here, the energetics and structural effects of replacing a bridging oxygen atom linking two SiO_4 units by a pair of fluorine atoms. The chemical reaction steps involved in the interaction of CF_2 with the oxygen site in amorphous SiO_2, with removal of the bridging oxygen atom and replacement with two fluorine atoms, is the subject of another study [27]. Here we compare the energy and structure for a "lattice fragment" of the SiO_2 glass made up of two oxygen-bridged SiO_4 tetrahedra with results for the difluorine-doped cluster.

Electronic structure and total energy calculations were carried out self-consistently, allowing charge density rearrangements, for the IGF reference SiO_2 cluster made up of two SiO_4 tetrahedra corner-linked through a bridging oxygen. Atoms were allowed to relax according to the calculated energy and force field, evaluated for each atom in the cluster. The fully relaxed fluorine-free cluster displays bond lengths in good (within 2%) agreement with experiment [28], justifying its use as a reference for determining dopant effects in the atomic cluster model. The total binding energy for this $O_3Si–O–SiO_3$ fragment (42.317 eV) reflects its high stability. Calculations of the energy of removal of the bridging oxygen atom show that its binding is strong (~6.40 eV) compared with the average binding energy of 6.05 eV per oxygen atom in the SiO_2 fragment. Similarly, for a *single* fluorine atom occupying the bridge site, the binding energy of fluorine is found to be ~1.40 eV less than that of oxygen. This tells us that the attack of silicon by fluorine alone in the glass would not disrupt the SiO_2 network.

Results for the cluster in which two fluorine ions are centered about the origin previously occupied by the bridging oxygen atom show the greatest stability with a binding energy of 47.432 eV. However, each of the very strong Si–F bonds is localized between fluorine and its nearest neighbor silicon, such that the two O_3SiF tetrahedra are only weakly interacting. Most of the binding energy originates within each O_3SiF unit. Most significant with regard to the IGF fracture behavior, these units become essentially non-interacting: only a very small force (≤0.85 eV/nm) is needed to separate these two tetrahedra beyond the local binding energy minimum of ~0.04 eV at an F–F distance of ~0.230 nm, Fig. 4. In contrast, the

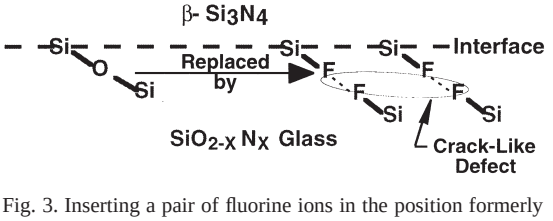


Fig. 3. Inserting a pair of fluorine ions in the position formerly taken by the oxygen linking the SiO_4 tetrahedra in the intergranular film results in the loss of the very strong oxygen link between tetrahedra. The very weak bond between the two fluorines is comparable to introducing a defect in the glass network.

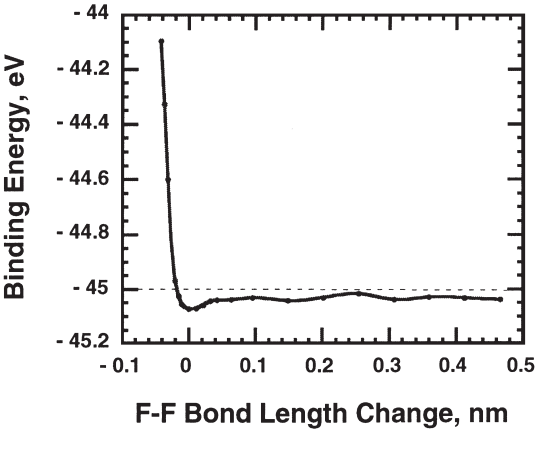


Fig. 4. The binding energy curve for the D_{3d} symmetry SiO_3–F–F–SiO_3 cluster as a function of separation of the tetrahedral units along the F–F axis. (The symmetry constraint raises the energy by ~2.3 eV relative to symmetry-relaxed values given in text). The zero of energy is for free atom constituents, and the separation is relative to equilibrium (F–F bond = 0.1193 nm).

tetrahedra of the oxygen-bridged reference cluster are bound so strongly to the joining oxygen atom that very large restoring forces (~41 eV/nm) act to return any displaced unit to equilibrium. In this analysis, IGF crack formation is a result of the fluorine pair substitution creating an extremely weak F–F bond and not a missing or dangling bond (plus a charge imbalance) that enhances crack growth in the amorphous phase. This enhanced crack formation and growth due to the fluorine reaction with the network is similar to that occurring when a water molecule interacts with the SiO_2 bonds at a crack tip in glasses [24].

In the difluorine structure, there is a strong repulsion when the F–F separation decreases below 0.185 nm and the fluorine densities overlap significantly (penetration beyond ~30% of the nominal ion radius). This repulsive force resists compression of the F–F bonds between the tetrahedra due to the van der Waals dispersion forces attempting to pull adjacent grains together [29]. The fact that the IGF is amorphous poses some problem for directly applying the calculated change in structure of the linked tetrahedra to the experimentally reported dilation of the IGF of 0.1 nm (e.g., see [19]). The increase in dimension is almost entirely in the F–F spacing, which becomes 0.243 nm. At first sight this seems too large, but if one incorporates the interactive forces between adjacent grains that define the IGF, film thicknesses are accurately given to within experimental precision (±0.1 nm) for both the SiO_2 glass and the fluorine-doped film (e.g., [19]). Taken together, the studies of anion dopant effects in IGFs show that weakening of the amorphous network by anion additions results from interactions at the atomic level in the strictest sense: the interaction range involved in the replacement of oxygen by difluorine is distinctly localized to the region of the former oxygen-bridge bond. This suggests the possibility of debonding control with selectivity at the atomic level, since the effect of dopants is specific to a particular type of site in the IGF and the results of the doping are known.

2.2. Interfacial strengthening due to epitaxial SiAlON

The second process that controls debonding is associated with the strength of bonding across the glass/crystalline interface. In the experimental approach used here, an increase in the interface strength is reflected by a decrease in the critical angle between the plane of the propagating crack and the prism plane at which the crack starts to deflect up the interface [17, 18]. High-resolution, high-magnification scanning electron microscopy images of the whiskers embedded in the various glasses indicate that the debond crack does run along the crystalline/glass interface. Studies of both β-Si_3N_4 whiskers embedded in oxynitride glasses and self-reinforced silicon nitride ceramics reveal that the critical debond angle is modified by the formation of an epitaxial β - $Si_{6-z}Al_zO_zN_{8-z}$ layer between the β-Si_3N_4 grain and the amorphous IGF, Table 1.

Comparison of the critical debond angle with analysis of the radial compressive stresses imposed on the prismatic interfaces indicates that the radial stress is not a major factor when an epitaxial SiAlON layer is present, as mentioned earlier. This is clearly shown for the 55Si10Al35Y20N80O glass system by comparing the critical debond with and without an epitaxial SiAlON layer present, Table 1. In addition, reductions in the nitrogen content of the glasses, while having only a minor effect on the radial stress levels, exclude the formation of an epitaxial layer and interfacial debonding occurs quite readily then.

The observations summarized in Table 1 reveal that the introduction of aluminum and oxygen into this epitaxial layer increases the strength of the interface [17, 18]. At the same time, the fracture toughness of self-reinforced Si_3N_4 ceramics with identical microstructures is found to decrease with increases in aluminum and oxygen level in the epitaxial SiAlON layers [30]. As a result, the potential for differences in the types of bond that formed across the interface, seen schematically in Fig. 5, was examined as a source of the interface strength. In the presence of an SiAlON layer, both Si–O and Al–O bonds, as well as Si–N bonds, could form, and thus silicon, aluminum, oxygen and nitrogen atoms would participate in the bonding from both sides of the interface with the Si–Al–RE–O–N amorphous intergranular films. This would increase the density of bonding sites along the interface and suggests a mechanism for the rise in strength of the interface. While Al–N bonds are involved in the SiAlON crystal structure, there is little evidence for them in the oxynitride glasses [31, 32], and thus they were not included in the interface bonding considerations. On the other hand, the formation of Si–O, Al–O and Si–N bonds is well documented in the oxynitride glasses.

Earlier electronic structure calculations showed that the substitution of aluminum and oxygen into the lattice of β-Si_3N_4 reduces the total overlap population (a measure of electron bond density) to which the substituted atoms contribute, in keeping with a reduction of the mechanical properties of β-Si_3N_4 versus β - $Si_{6-z}Al_zO_zN_{8-z}$ [33, 34]. However, more recent studies that include charge redistribution indicate that while the substitution of aluminum for a silicon in an Si_3N_4 unit results in a weaker Al–N bond, the strengths of other aluminum- and silicon-containing bonds in the structure actually increase [35]. In addition, crystal-orbital-based calculations have assessed the bonding of aluminum and oxygen to the β-Si_3N_4 surfaces and suggested stronger bonding of these elements to the prism faces, but only as compared with the basal surface [36].

The observed influence of the SiAlON composition on the interface strength motivated the current studies using first-principles atomic cluster calculations to address the following. (1) Why is the oxynitride

Table 1. Summary of the critical interfacial debond angles for the β-Si_3N_4/oxynitride glass

Sample	Critical debond angle (°)	Compressive stress on interface (MPa)[a]	$Si_{6-z}Al_zO_zN_{8-z}$	β-SiAlON predicted from phase equilibria
β-Si_3N_4 whiskers in oxynitride glass matrix				
Composition (eq%)				
55Si25Al20Y10N90O	~70	~350	0	No
55Si25Al20Y20N80O	~55	~320	1.0	Yes
55Si10Al35Y10N90O	~70	~515	0	No
55Si10Al35Y20N80O				
without SiAlON layer	~72	~530	0	Yes
with SiAlON layer[b]	~50	~530	0.15	
57Si43La20N80O	~68	~590	0	No
46Si27Al27La27N73O	~50	~555	1.6–2.0	Yes
41Si30Al29Yb23N77O	~55	~455	1.6–2.0	Yes
Self-reinforced β-Si_3N_4 with amorphous intergranular film				
Additives (wt%) yttria/alumina				
6.25/1.0	~75		0.01	Yes
5.0/2.0	~70		0.03	Yes
4.0/2.8	~60		0.06	Yes

[a] Calculated stress based on measured values of thermal expansion coefficients, softening temperatures and Young's modulus of the glasses and silicon nitride (e.g., [18]).
[b] Sample after additional nitrogen (5 MPa) anneal at 1680°C for 30 min to develop epitaxial layer.

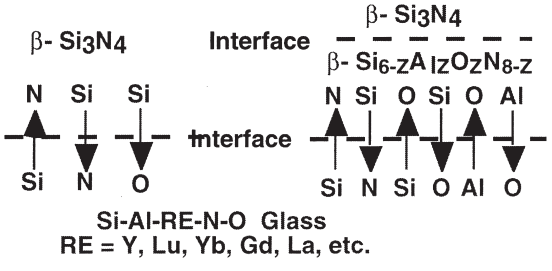


Fig. 5. As shown schematically here, only Si–N bonds are established in both directions across the interface in the case of the Si_3N_4/glass interface. When an epitaxial SiAlON layer forms, Si–O and Al–O bonds, as well as Si–N bonds, form in both directions across the glass/crystalline interface.

glass/β-Si_3N_4 interface weak compared with the β - $Si_{6-z}Al_zO_zN_{8-z}$/glass interface? (2) What is the strengthening mechanism when $Si_{6-z}Al_zO_zN_{8-z}$ forms at the interface? (3) What is the influence of composition on the interfacial strength in terms of atomic-level interactions?

The model structure used for this study was the basic tetrahedron, which is a common structural unit of both the crystalline phases (i.e., β-Si_3N_4 and β - $Si_{6-z}Al_zO_zN_{8-z}$) and the oxynitride glasses. The interfacial debonding of most interest here involves that parallel to the prism planes of the elongated reinforcing grains. A slice normal to these planes, shown schematically in Fig. 6, reveals that the atoms in the grain and the glass involved in bonding across the prism interfaces can also, indeed, be defined in tetrahedral units. Thus, one can use the simple tetrahedral unit building blocks to begin to describe the interfacial bonding and the effects of composition by altering the atomic species within the body-centered silicon- and aluminum-based tetrahedral clusters.

The partial-wave self-consistent cluster method [25] was used to solve local density equations for such tetrahedral cluster models. This allows one to determine the energy and force field changes that would occur by a variety of atomic substitutions to gain a first approximation of the influence of composition on the interface cohesion. To achieve an atomic-level understanding of compositional effects, the electronic structures, total energies and atomic force fields were calculated to high accuracy using only the atomic numbers and parameters describing the local exchange correlation approximation (plus gradient corrections) as inputs.

Then the source was sought for the increase in interface strength with the formation of an epitaxial $Si_{6-z}Al_zO_zN_{8-z}$ layer at the interface and the rapid rise in strength over a compositional range of $z = 0$ to ~0.2 (i.e., increasing substitution of aluminum for silicon and oxygen for nitrogen). The influence of aluminum and oxygen can be addressed by determining how the binding energy of the basic body-centered silicon (and aluminum) tetrahedra are affected by variations in the nitrogen-to-oxygen ratio. As seen in Fig. 7, the trend is for the binding energy to increase with substitution of oxygen for nitrogen in both the silicon- and aluminum-based tetrahedra. At the same N/O ratio, the silicon-based tetrahedra do exhibit higher binding energies than the aluminum-based tetrahedra. However, the AlO_3N and AlO_4 tetrahedra exhibit greater binding energies than do the SiN_4 and $SiON_3$ tetrahedra. In fact, the binding energy of the AlO_4 unit closely rivals that of the SiO_4 unit. On the other hand, the AlN_4 and SiN_4 tetrahedra are the most weakly bound units.

The above trends of increasing binding energies within both silicon- and aluminum-based tetrahedra with increasing substitution of oxygen for nitrogen are confirmed by additional calculations based on two and three corner-connected tetrahedra with oxygen substituted for nitrogen and aluminum for silicon.

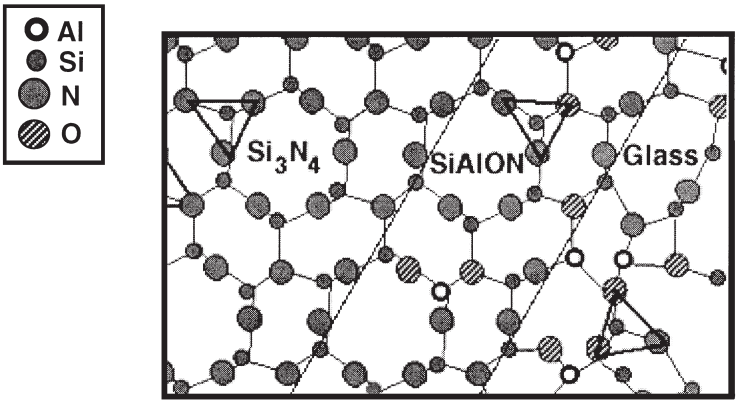


Fig. 6. The atomic structure involved in bonding across the prism planes is shown in the cross-section parallel to the basal plane of Si_3N_4 that cuts across the Si_3N_4/SiAlON and SiAlON/glass interfaces. Basic to all three components are the simple silicon-based and aluminum-based body-centered tetrahedral unit cells, which are used in the atomic cluster calculations to describe bonding at the interface.

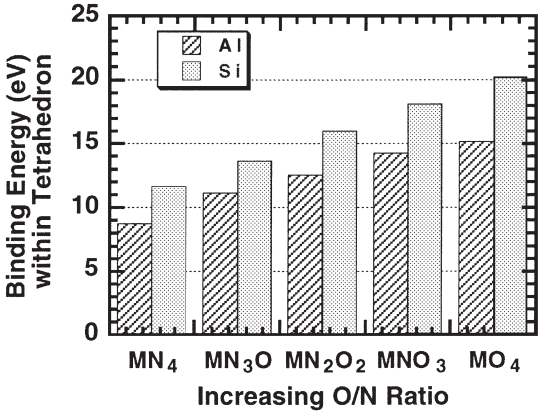


Fig. 7. First-principles calculations of the binding energy reveal a continuous increase with substitution of oxygen for nitrogen in both the silicon- and aluminum-based tetrahedra. While the silicon-based tetrahedra exhibit greater binding energies than do their aluminum-based counterparts, the oxygen-rich AlO_3N and AlO_4 tetrahedra exhibit greater binding energies than do the oxygen-deficient SiN_4 and $SiON_3$ tetrahedra.

These, then, further support the strengthening of the interface bonding by the addition of Si–O and Al–O bonding across the interface as first suggested in Fig. 5. At the β-Si_3N_4/glass interfaces, silicon on the Si_3N_4 surface can bond to oxygen in the glass, while Si–N bonds can form in both directions across the interface. With an epitaxial $Si_{6-z}Al_zO_zN_{8-z}$ layer at the interface, Si–O bonds, the strongest bonds, can form across the interface in both directions. In addition, these are supplemented by Al–O bonds, the second strongest bonds in the series, which can be formed in either direction across the interface. Both the greater strengths of these oxide bonds and the fact that they can involve silicon and aluminum atoms present on both sides of the interface contribute to the increase in interfacial strength and resistance to debonding observed with the increasing substitution of aluminum and oxygen in the interfacial $Si_{6-z}Al_zO_zN_{8-z}$ layer.

3. SUMMARY

Interfacial debonding, which is a key step in crack-bridging mechanisms, appears to involve at least two processes. When fluorine is incorporated into the amorphous intergranular film, the glass network is weakened, which allows a propagating crack to bypass the reinforcing grains without cutting through them. In the presence of SiYAl oxynitride intergranular glass films, reductions in the Al/Y ratio decrease the aluminum and oxygen content of the epitaxial $Si_{6-z}Al_zO_zN_{8-z}$ layer on the Si_3N_4 grains and debonding is enhanced.

Atomic cluster calculations show that the substitution of two fluorine atoms in the place of the oxygen linking two SiO_4 tetrahedra results in (1) fluorines that are strongly bonded within each tetrahedra and (2) weak bonding between the two fluorine atoms, which are energetically favored to rest on the axis between the silicon atoms. As such, the calculated structure for the oxygen-linked and the fluorine-pair-linked tetrahedra yields intergranular film thicknesses in agreement with observations. The (at best) weak bond between the fluorine pair then acts like a defect in the glass network. The calculations show that the restoring force resisting the separation of the fluorine pair is extremely low. As a result, increases in the local density of such defects with fluorine additions will substantially weaken the intergranular glass network and, thus, allow cracks within the IGF to propagate around, rather than through, the reinforcing grains.

In the case of SiYAl oxynitride glass intergranular films, an increase in resistance to interfacial debonding occurs with the formation of the layer at the interface. Atomic cluster studies reveal that the binding energies of the SiN_4 tetrahedra increase as oxygen replaces the nitrogen, with SiO_4 exhibiting very high binding energies. The formation of $Si_{6-z}Al_zO_zN_{8-z}$ requires the incorporation of alumninum in place of silicon, and the binding energies of the aluminum-

based tetrahedra show the same increase with oxygen substitution for nitrogen. In fact, the AlO_4 tetrahedron has only slightly lower binding energy than does the SiO_4 tetrahedron. This increase in binding energies with oxygen substitution has an important impact on the interface strength.

When the glass is in contact with the Si_3N_4 grain, Si–N bonds can form in both directions across the interface; however, Si–O bonds will only occur between the silicon on the grain surface and oxygen in the glass. Thus opportunity to form the strongest bond (i.e., Si–O) is fairly limited. On the other hand, the growth of an $Si_{6-z}Al_zO_zN_{8-z}$ layer on the Si_3N_4 grains allows Si–O, as well as Si–N, bonds to form in both directions across the interface. The strong Si–O bonding will also be supplemented by the addition of nearly as strong Al–O bonding in both directions across the interface. Thus, the atomic cluster studies have established a relationship between composition and bond strength as a function of O/N ratio that provides insight into the role of composition on the debonding processes in the toughened silicon nitride ceramics.

Acknowledgements—This research was sponsored by the US Department of Energy, Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, under contract DE-AC05-00OR22725 with UT-Battelle, LLC.

REFERENCES

1. Himsolt, G., Knoch, H., Huebner, H. and Kleinlein, F. W. *J. Am. Ceram. Soc.*, 1979, **62**(1), 29.
2. Lange, F. F. *J. Am. Ceram. Soc.*, 1979, **62**(7-8), 428.
3. Kawashima, T., Okamoto, H., Yamamoto, H. and Kitamura, A. *J. Ceram. Soc. Jpn.*, 1991, **99**, 1.
4. Mitomo, M. in *Proceedings: Science of Engineering Ceramics*, ed. S. Kimura and K. Niihara, Ceramic Society of Japan, Tokyo, 1991, p. 101.
5. Hirao, K., Nagaoka, T., Brito, M. E. and Kanzaki, S. *J. Am. Ceram. Soc.*, 1994, **77**(7), 1857.
6. Hirao, K., Nagaoka, T., Brito, M. E. and Kanzaki, S. *J. Ceram. Soc. Jpn.*, 1996, **104**, 55.
7. Emoto, H. and Mitomo, M. *J. Eur. Ceram. Soc.*, 1997, **17**(4), 797.
8. Becher, P. F., Hsueh, C. H., Angelini, P. and Tiegs, T. N. *J. Am. Ceram. Soc.*, 1988, **71**(12), 1050.
9. Becher, P. F. *J. Am. Ceram. Soc.*, 1991, **74**(2), 255.
10. Becher, P. F., Hsueh, C. H., Alexander, K. B. and Sun, E. Y. *J. Am. Ceram. Soc.*, 1996, **79**(2), 298.
11. Sun, E. Y., Hsueh, C. H. and Becher, P. F. in *Fracture-Instability Dynamics, Scaling and Ductile/Brittle Behavior*, ed. R. B. Selinger, J. J. Mecholsky, A. E. Carlsson and E. R. Fuller, *MRS Proceedings*, Vol. 409 Materials Research Society, Boston, MA, 1996, p. 223.
12. Evans, A. G. and McMeeking, R. M. *Acta metall.*, 1986, **34**(12), 2435.
13. Li, C. W. and Yamanis, J. *Ceram. Eng. Sci. Proc.*, 1989, **10**(7-8), 632.
14. Becher, P. F., Sun, E. Y., Plucknett, K. P., Alexander, K. B., Hsueh, C. -H., Lin, H. -T., Waters, S. B., Westmoreland, C. G., Kang, E. -S., Hirao, K. and Brito, M. *J. Am. Ceram. Soc.*, 1998, **81**(11), 2821.
15. Hirao, K., Imamura, H., Watari, K., Brito, M. E., Toriyama, M. and Kanzaki, S. in *Science of Engineering Ceramics II*, ed. K. Niihara, T. Sekino, E. Yasuda and T. Sasa, Ceramic Society of Japan, Tokyo, 1999, p. 469.
16. Pezzotti, G., Nishida, T., Kleebe, H. -J., Ota, K., Muraki, N. and Sergo, V. *J. Mater. Sci.*, 1999, **34**(7), 1667.
17. Becher, P. F., Sun, E. Y., Hsueh, C. H., Alexander, K. B., Hwang, S. L., Waters, S. B. and Westmoreland, C. G. *Acta metall.*, 1996, **44**(10), 3881.
18. Sun, E. Y., Becher, P. F., Hsueh, C. H., Painter, G. S., Waters, S. B., Hwang, S. L. and Hoffmann, M. J. *Act. mater.*, 1999, **47**(9), 2777.
19. Kleebe, H. -J. *J. Ceram. Soc. Jpn.*, 1997, **105**(6), 453.
20. Ching, W. -Y., Xu, Y. -N., Gale, J. D. and Rühle, M. *J. Am. Ceram. Soc.*, 1998, **81**(12), 3189.
21. Kaneko, K., Yoshiya, M., Tanaka, I. and Tsurekawa, S. *Acta. mater*, 1999, **47**(4), 1281.
22. Blonski, S. and Garofalini, S. H. *J. Am. Ceram. Soc.*, 1997, **80**(8), 1997.
23. Litton, D. A. and Garofalini, S. H. *J. Mater. Res.*, 1999, **14**(4), 1418.
24. Michalske, T. A. and Freiman, S. W. *J. Am. Ceram. Soc.*, 1983, **66**(4), 284.
25. Averill, F. W. and Painter, G. S. *Phys. Rev. B*, 1994, **50**(11), 7262.
26. Jones, R. O. and Gunnarsson, O. *Rev. Mod. Phys.*, 1989, **61**(3), 689.
27. Painter, G. S., unpublished.
28. Jin, J., Yoko, T., Miyaji, F., Sakka, S., Fukunaga, T. and Misawa, M. *Phil. Mag. B*, 1994, **70**(2), 191.
29. Clarke, D. R. *J. Am. Ceram. Soc.*, 1987, **70**(1), 15.
30. Sun, E. Y., Becher, P. F., Hsueh, C. -H., Waters, S. B., Plucknett, K. P., Hirao, K. and Brito, M. *J. Am. Ceram. Soc.*, 1998, **81**(11), 2831.
31. Rouxel, T., Besson, J. L., Rzepka, E. and Goursat, P. *J. Non-Cryst. Solids*, 1990, **122**, 298.
32. McMillan, P. F., Sato, R. K. and Poe, B. T. *J. Non-Cryst. Solids*, 1998, **224**, 267.
33. Tanaka, I., Niihara, K., Nasu, S. and Adachi, H. *J. Am. Ceram. Soc.*, 1993, **76**(11), 2833.
34. Tanaka, I., Nasu, S., Adachi, H., Miyamoto, Y. and Niihara, K. *Acta metall. mater.*, 1992, **40**(9), 1995.
35. Ching, W.-Y., Huang, M.-Z. and Mo, S.-D., *J. Am. Ceram. Soc.*, 2000, **83**(4), 780.
36. Dudesek, P. and Benco, L. *J. Am. Ceram. Soc.*, 1998, **81**(5), 1248.

Solution to the Boltzmann equation for layered systems for current perpendicular to the planes

W. H. Butler and X.-G. Zhang
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

J. M. MacLaren
Tulane University, New Orleans, Louisiana 70018

Present theories of giant magnetoresistance (GMR) for current perpendicular to the planes (CPP) are based on an extremely restricted solution to the Boltzmann equation that assumes a single free electron band structure for all layers and all spin channels. Within this model only the scattering rate changes from one layer to the next. This model leads to the remarkable result that the resistance of a layered material is simply the sum of the resistances of each layer. We present a solution to the Boltzmann equation for CPP for the case in which the electronic structure can be different for different layers. The problem of matching boundary conditions between layers is much more complicated than in the current in the planes (CIP) geometry because it is necessary to include the scattering-in term of the Boltzmann equation even for the case of isotropic scattering. This term couples different values of the momentum parallel to the planes. When the electronic structure is different in different layers there is an interface resistance even in the absence of intermixing of the layers. The size of this interface resistance is affected by the electronic structure, scattering rates, and thicknesses of nearby layers. For Co–Cu, the calculated interface resistance and its spin asymmetry is comparable to that measured at low temperature in sputtered samples. © *2000 American Institute of Physics.* [S0021-8979(00)71908-6]

I. INTRODUCTION

Although most studies of GMR are for the case of applied electric fields and hence net current in the plane of the film (CIP geometry), there has been considerable experimental and theoretical work on the current perpendicular to the plane (CPP) geometry. Gijs and Bauer have provided a useful review[1] of recent work on CPP GMR. Camblong, Zhang, and Levy[2] presented a simple model for the CPP conductance for a single spin channel in the case of identical free electron band structure in all layers, with the only difference between layers being possible differences in the scattering rates. Valet and Fert, in a classic paper,[3] generalized this model to include scattering processes that connect the two spin channels. The Michigan State group and co-workers[4,5] measured GMR for numerous systems for the CPP geometry, and found that many of their results could be fit with the Camblong–Levy or Valet–Fert models.

In this paper, we present a solution to the Boltzmann equation for the case in which the electronic structure can be different for different layers and different spin channels. We compare our results to the ''standard'' theory.[2,3]

II. BOLTZMANN EQUATION FOR CPP TRANSPORT

Although it is not necessary in our approach, in order to make contact with the standard approach, we shall assume that the scattering probability is isotropic. Then, within a single layer in which we can assume that $v_z^s(\mathbf{k})$ is independent of z, the Boltzmann equation can be written[3] in terms of the anisotropic function, $g_s(z,\mathbf{k})$,

$$v_z^s \frac{\partial g_s(z,\mathbf{k})}{\partial z} + \frac{g_s(z,\mathbf{k})}{\tau_s} = v_z^s \frac{\partial \bar{\mu}_s(z)}{\partial z}, \tag{1}$$

where we have defined $\bar{\mu}_s(z) = \mu(z) - eV(z)$. The ''standard'' treatment of CPP transport[2,3] is obtained by using this equation for all layers, thereby implicitly assuming that there is no change in electronic structure from layer to layer. The reason for this assumption, of course, is the difficulty in solving the CPP Boltzmann equation for an inhomogeneous system. This difficulty arises from the requirement that $g_s(z,\mathbf{k})$ be anisotropic, i.e., that

$$\sum_{\mathbf{k}} \frac{\partial f_0(E_{\mathbf{k}}^s - \mu_0)}{\partial E_{\mathbf{k}}^s} g_s(z,\mathbf{k}) \equiv 0. \tag{2}$$

This requirement arises from the definition of $g_s(z,\mathbf{k})$,

$$g_s(z,\mathbf{k}) = \mu_s(z) - h_s(z,\mathbf{k}), \tag{3}$$

where the chemical potential $\mu(z)$ is defined as the average of the distribution function, $\mu_s(z) = \langle h_s(z,\mathbf{k}) \rangle$, and the angular brackets indicate a Fermi surface average. The requirement that $g_s(z,\mathbf{k})$ satisfy Eq. (2) is equivalent to including the scattering-in term of the usual form of the Boltzmann equation and is necessary for current conservation.

In Eq. (1), the scattering rate, $1/\tau_s$, can be allowed to vary with z without significantly complicating the solution. Thus it is easy to show that for homogeneous electronic structure but varying scattering rate, the CPP resistance depends only on the averaged scattering rate,

$$R_s A = \frac{\int dz \frac{1}{\tau_s(z)}}{e^2 n_s \langle (v_z^s)^2 \rangle} = \int dz \rho_s(z), \tag{4}$$

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

where A is the area of the film perpendicular to the direction of the current, n_s is the Fermi energy density of electronic states, $\rho_s(z)$ is a local resistivity of free electrons and the angular brackets denote an average over the Fermi surface. The second expression in Eq. (4) is often assumed to be the proper generalization of the free electron formulas to general electron dispersion. In the following we shall refer to this result as the ''standard'' expression for CPP resistance.

In order to treat a system in which the electronic structure is different in successive layers, it is necessary to match the solutions to the Boltzmann equation obtained for each layer across the interfaces between the layers. The procedure for performing this matching is relatively straightforward,[6] but involves the full distribution function, $h_s(z,\mathbf{k})$, not just the anisotropic part,

$$h_{i+1}^{+,j}(z_i^+,\mathbf{k}_\parallel)=\sum_{j',\mathbf{k}'_\parallel}^{N_R} T_i^{+-}(j\mathbf{k}_\parallel,j'\mathbf{k}'_\parallel)h_{i+1}^{-,j'}(z_i^+,\mathbf{k}'_\parallel)+\sum_{j',\mathbf{k}'_\parallel}^{N_L} T_i^{++}(j\mathbf{k}_\parallel,j'\mathbf{k}'_\parallel)h_i^{+,j'}(z_i^-,\mathbf{k}'_\parallel)$$

$$h_i^{-,j}(z_i^-,\mathbf{k}_\parallel)=\sum_{j',\mathbf{k}'_\parallel}^{N_L} T_i^{-+}(j\mathbf{k}_\parallel,j'\mathbf{k}'_\parallel)h_i^{+,j'}(z_i^-,\mathbf{k}'_\parallel)+\sum_{j',\mathbf{k}'_\parallel}^{N_R} T_i^{--}(j\mathbf{k}_\parallel,j'\mathbf{k}'_\parallel)h_{i+1}^{-,j'}(z_i^+,\mathbf{k}'_\parallel). \quad (5)$$

Here N_L and N_R denote the number of states on the left or right of the interface, respectively, for a given value of $k'_\parallel$.

In order to properly match the distribution functions at the boundaries of the layers, it is necessary to utilize the general solution to Eq. (1) and to admit exponentially varying solutions that could be omitted for the homogeneous case. Thus, it can be verified that the anisotropic distribution function for spin s in layer i can be written as

$$g_s(z,\mathbf{k})=\frac{J_s}{\sigma_{si}}ev_z^{si}(\mathbf{k})\tau_{si}-F_{si}(\mathbf{k})e^{(-z/v_z^{si}(\mathbf{k})\tau_{si})}+\left\langle \frac{F_{si}(\mathbf{k}')e^{(-z/v_z^{si}(\mathbf{k}')\tau_{si})}}{v_z^{si}(\mathbf{k})-v_z^{si}(\mathbf{k}')}\frac{v_z^{si}(\mathbf{k})}{\left\langle\frac{v_z^{si}(\mathbf{k})}{v_z^{si}(\mathbf{k})-v_z^{si}(\mathbf{k}')}\right\rangle_{\mathbf{k}}}\right\rangle_{\mathbf{k}'}, \quad (6)$$

$$\overline{\mu}(z)=\alpha_{is}J_s-\frac{J_s}{\sigma_{si}}ez+\left\langle\frac{F_{si}(\mathbf{k}')e^{(-z/v_z^{si}(\mathbf{k}')\tau_{si})}}{\left\langle\frac{v_z^{si}(\mathbf{k})}{v_z^{si}(\mathbf{k})-v_z^{si}(\mathbf{k}')}\right\rangle_{\mathbf{k}}}\right\rangle_{\mathbf{k}'}. \quad (7)$$

Here, J_s is the current density for spin-channel s and σ_{si} is the bulk conductivity for spin s of the material in layer i,

$$\sigma_{si}=-\frac{e^2}{V}\sum_{\mathbf{k}}(v_z^{si}(\mathbf{k}))^2\tau_{si}\delta(E_{\mathbf{k}}-E_F), \quad (8)$$

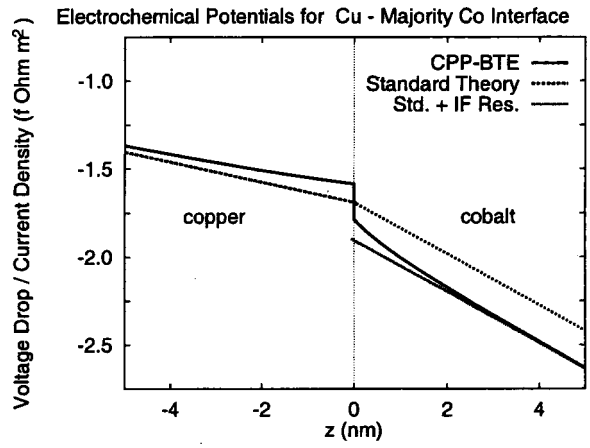


FIG. 1. Chemical potential divided by current density for copper–cobalt majority interface.

and α_{is} and $F_{si}(\mathbf{k})$ are parameters determined by the matching conditions at the boundaries, Eq. (5).

If the distribution function, Eq. (6), is used to calculate the current density, the first term yields J_s and the contributions of the other two terms cancel. The second term would have been expected from the general solution to the CIP case. The third term is made necessary by the requirement that $g(z,\mathbf{k})$ be purely anisotropic.

III. TRANSPORT FOR CURRENT PERPENDICULAR TO THE PLANES

The layer Korringa–Kohn–Rostoker approach[7] was used to calculate the self-consistent electronic structures of cobalt, copper, and cobalt–copper interfaces. These were used to evaluate the transmission (T^{++},T^{--}) and reflection probabilities (T^{+-} and T^{-+}) for Bloch electrons impinging on the interfaces. The Boltzmann equation, including the boundary matching equations, was then solved using an iterative procedure. In the calculations presented here, it was assumed that the interfaces were epitaxial and that there was no additional disorder in the vicinity of the interface.

Figures 1 and 2 show the calculated electrochemical potential, $\overline{\mu}$, for the majority and minority channels in the vi-

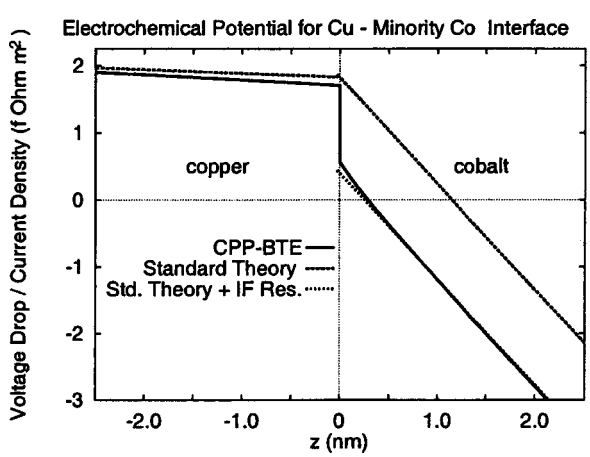


FIG. 2. Chemical potential divided by current density for copper–cobalt minority interface.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

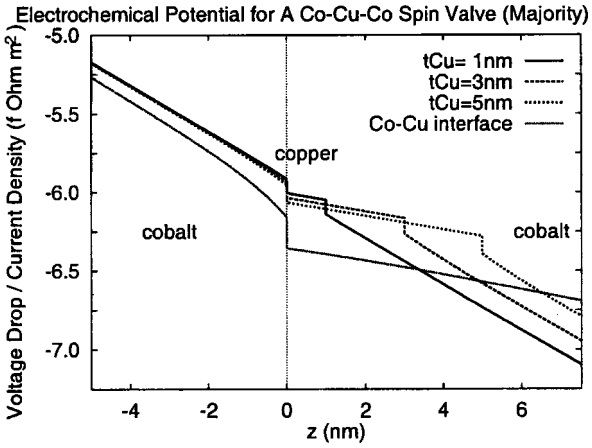


FIG. 3. Chemical potential divided by current density for the majority channel of a cobalt–copper–cobalt spin valve for various thicknesses of the copper layer.

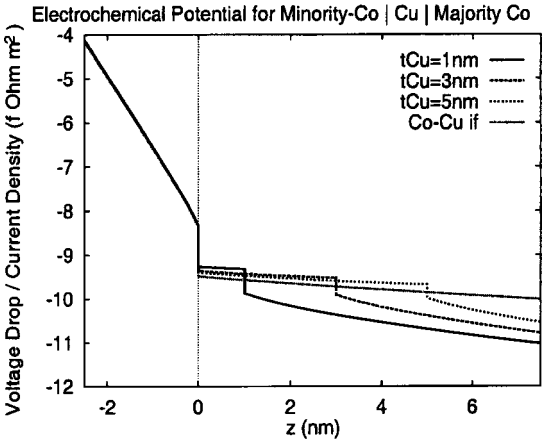


FIG. 4. Chemical potential divided by current density for a cobalt–copper spin valve with cobalt moments aligned antiparallel.

cinity of a copper–cobalt interface. The electrochemical potential has been divided by the current density so that the plots yield $\int dz \rho_s(z)$. For the standard theory this is simply a straight line for each layer. The actual solution to the Boltzmann equation differs in two ways: (1) There is a discontinuity in the chemical potential at the interface which is equivalent to an interfacial resistance. This interfacial resistance is not due to intermixing or additional scattering at the interface (although this effect can be included in the model if desired) but to the mismatch of the bands across the interface which causes some of the electrons incident on the interface to be reflected. (2) There are exponential terms in the electrochemical potential in the vicinity of the interface that decay at a rate comparable to the component of the mean free path perpendicular to the layers. The effect of these terms can be included as an additional interfacial resistance that is added to the discontinuous contribution just described, as is indicated in the dotted lines of Figs. 1 and 2. If this is done, however, it must be taken into consideration that this additional contribution depends on the environment of the interface, e.g., the proximity of other interfaces and the bulk scattering rates. The calculated interfacial resistances can be read off the figures (0.21 and 1.2 $f\Omega m^2$ for majority and minority, respectively) and are comparable to those deduced from low-temperature measurements[5] on cobalt–copper multilayers. This indicates that much of the interfacial resistance previously attributed to disorder at the interface may equally well arise from reflection.

Figures 3 and 4 show that both the discontinuous interfacial resistance and the interfacial resistance arising from the exponential terms can depend on the type and thicknesses of neighboring layers. In Fig. 3, for example, it can be seen that the discontinuous interfacial resistance on both sides of the Co–Cu–Co (majority) spin valve increase with the thickness of the copper spacer layer. The exponential contributions also increase with copper layer thickness. For the case of antiparallel alignment shown in Fig. 4, however, the discontinuous contribution increases slightly with spacer layer thickness on the minority side but *decreases* with thickness on the majority side.

Recently, Chien *et al.*[8] used a tight-binding based quantum model to calculate CPP transport through Co–Cu multilayers. They pointed out the necessity of including vertex corrections which are equivalent to the scattering-in term of the Boltzmann equation in the semiclassical limit. Very recently, Penn and Stiles[9] reported calculations of the interface resistance for free electrons incident on a model interface.

ACKNOWLEDGMENTS

Helpful conversations and communications with P. Bruno, S. Hershfield, and especially M. Stiles, who helped us to find the form of Eq. (6), are gratefully acknowledged. Work at Oak Ridge was sponsored by the Defense Advanced Research Projects Agency and by the Division of Materials Sciences Office of Basic Energy Sciences of the U.S. DOE, under Contract No. DE-AC05-96OR22464 with Lockheed Marietta Energy Research Corp. J.M.M. acknowledges support from the Oak Ridge Institute for Science and Education and DARPA Grant No. MDA 972-97-1-003.

[1] M. A. M. Gijs and G. E. W. Bauer, Adv. Phys. **46**, 285 (1997).
[2] H. E. Camblong and P. M. Levy, Phys. Rev. Lett. **69**, 2835 (1992); J. Appl. Phys. **73**, 5533 (1993); H. E. E. Camblong, S. Zhang, and P. M. Levy, Phys. Rev. B **47**, 4735 (1993); H. E. E. Camblong, P. M. Levy, and S. Zhang, *ibid.* **51**, 16052 (1995).
[3] T. Valet and A. Fert, Phys. Rev. B **48**, 7099 (1993).
[4] W. P. Pratt, Jr. *et al.*, Phys. Rev. B **46**, 548 (1992); A. C. Reilly *et al.*, J. Magn. Magn. Mater. **195**, L269 (1999); W. Park *et al.*, J. Appl. Phys. **85**, 4542 (1999); J. A. Borchers *et al.*, Phys. Rev. Lett. **82**, 2796 (1999); P. Holody *et al.*, Phys. Rev. B **58**, 12230 (1998).
[5] W. P. Pratt, Jr. *et al.*, J. Magn. Magn. Mater. **126**, 406 (1993); Q. Yang *et al.*, Phys. Rev. B **51**, 3226 (1995).
[6] W. H. Butler, X.-G. Zhang, and J. M. MacLaren, IEEE Trans. Magn. **34**, 927 (1998).
[7] J. M. MacLaren, X.-G. Zhang, W. H. Butler, and X. Wang, Phys. Rev. B **59**, 5470 (1999).
[8] J. Chien, T.-S. Choy, and S. Hershfield, J. Appl. Phys. **85**, 4551 (1999).
[9] D. R. Penn and M. D. Stiles, Phys. Rev. B **59**, 13338 (1999).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

PHILOSOPHICAL MAGAZINE LETTERS, 2000, VOL. 80, NO. 10, 683–690

Thermal expansion coefficients of Mo–Si compounds by first-principles calculations

C. L. FU† and XINDONG WANG‡

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6114, USA

[*Received 16 February 2000 and accepted in revised form 19 June 2000*]

ABSTRACT

Taking Mo–Si compounds as model systems, we show that the coefficients of thermal expansion (CTEs) of complex structures can be calculated precisely from first principles by incorporating the Debye model for acoustic response. Specifically, we obtain a nearly isotropic CTE in $MoSi_2$ but a highly anisotropic CTE in Mo_5Si_3. The CTE anisotropy in Mo_5Si_3 is due to an elastically more rigid basal plane and a higher anharmonicity along the *c* axis. As the structure of 5–3 compounds is modified from $D8_m$ to $D8_l$ by boron substitutions (Mo_5SiB_2), we predict a significant decrease in the CTE anisotropy, which is confirmed by experiments.

§ 1. INTRODUCTION

In the synthesis and processing of non-cubic high-temperature materials, microcracks are often induced owing to the anisotropy in the coefficients of thermal expansion (CTEs). To minimize the thermal stresses in polycrystalline materials, a more isotropic CTE is desirable. The degree of the anisotropy, however, is not predictable by any empirical rules and can vary substantially within any specific material type. For example, among the transition-metal silicides, the CTEs are highly anisotropic for 5–3 silicides, but become nearly isotropic for disilicides (Shah *et al.* 1992). No *ab initio* theory has been developed to correlate the degree of the anisotropy with the nature of crystal and electronic structures. In this paper, taking the tetragonal Mo–Si alloys ($MoSi_2$, Mo_5Si_3 and Mo_5SiB_2) as model systems, we show that *ab initio* theory can precisely predict and elucidate the theromoelastic properties of these alloys.

That the CTEs are highly anisotropic in 5–3 transition-metal silicides is well documented. There is no exception for Mo_5Si_3. Measurements by Chu *et al.* (1999) show that the CTE in the [001] direction is more than twice the CTE in the [100] direction. We shall calculate the CTEs and address the origin of anisotropy in this class of compounds. One prominent feature of 5–3 Mo–Si is that the substitution by boron (Mo_5SiB_2) substantially improves the oxidation resistance (Meyer *et al.* 1996) but, at the same time, changes the crystal structure from $D8_m$ (T1 phase) to

† Email: fucl@ornl.gov

‡ Present address: New Century Applied Science, Inc., 7501 Forbes Boulevard, Suite 104, Seabrook, Maryland 20706, USA.

Philosophical Magazine Letters ISSN 0950–0839 print/ISSN 1362–3036 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk /journals

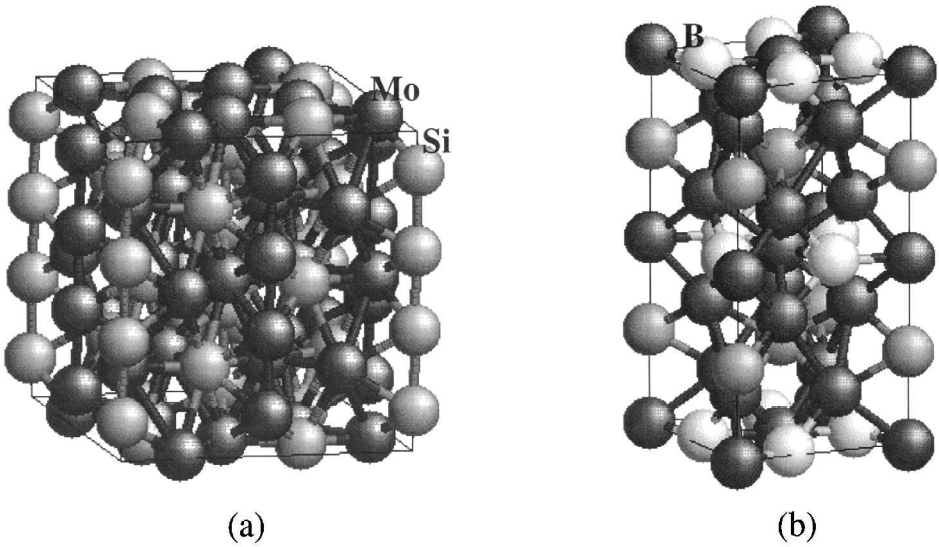


Figure 1. The crystal structures of the (*a*) $D8_m$ and (*b*) $D8_l$ structures.

$D8_l$ (T2 phase). An important basic question is whether or not the CTE anisotropy of the T1 phase is retained in the T2 phase. The answer provided by our investigation is that the CTEs of the T2 phase become nearly isotropic. This prediction is further confirmed by experiments. The structures of $D8_m$ and $D8_l$ are depicted in figure 1. Note the existence of prominent chain structures along the *c* axis shown at the cell boundaries in figure 1 (*a*).

§ 2. The theoretical model

Although thermal expansion is a classic subject, the calculation of CTEs by first principles is still largely limited to simple cubic elements (for example Fleszar and Gonze (1990) and Quong and Liu (1997)). This is the situation because rigorous calculations of the entire phonon spectra become lengthy and difficult for complex multicomponent systems. The problem can be simplified by approximating the phonon contribution through the elastic acoustic response (i.e. the Debye model). The semiempirical approach based on this model has had quantitative success in describing the CTEs of monatomic cubic metals (Moruzzi *et al.* 1988). Generally, the use of a Debye model for thermal expansion tends to overestimate the hydrostatic pressure for multicomponent systems. For the present systems, however, we find that the use of a Debye model for the lattice anharmonicity in our formalism (see below) is adequate enough to give CTE values to within about 10% of the available measured values in most cases. On physical grounds, this is not surprising. The materials investigated here display strong covalent bonding characteristics (Fu *et al.* 1999), which imply high optical phonon energies. Because of their higher activation energies, optical phonons are not expected to modify significantly the results presented here. The good agreement between our results and the measurements on the prototype system $MoSi_2$ (see below) further confirms that optical phonons do not play significant roles.

To calculate the CTE of a tetragonal structure, we consider the free energy $F(T)$ in the presence of elastic strain components (uniform expansions or contractions) u_i as

$$F(T) = E_0 + \frac{1}{2} C_{\mu\nu} u_\mu u_\nu + \frac{1}{3!} C_{\mu\nu\sigma} u_\mu u_\nu u_\sigma + F_{\text{phonon}}(T), \tag{1}$$

where the Greek subscripts take values from 1 to 3 and the Einstein summation convention is implied. F_{phonon} is the phonon free energy (including zero-point motion), and E_0 is the total energy at 0 K. C_{ij} and C_{ijk} are second-order and third-order elastic constants respectively. The sound velocities are determined from the Christoffel equation (Ashcroft and Mermin 1976). A minimization of the free energy with respect to the strains ($\delta a/a$ and $\delta c/c$ in [100] and [001] directions respectively) yields the following equations for the change in lattice parameters at temperature T:

$$\frac{\delta a}{a} = \frac{2\tilde{C}_{13}^{(2)} \Gamma_c(T) - \tilde{C}_{33} \Gamma_a(T)}{2(\tilde{C}_{11} + \tilde{C}_{12})\tilde{C}_{33} - 4\tilde{C}_{13}^{(1)} \tilde{C}_{13}^{(2)}}, \tag{2}$$

$$\frac{\delta c}{c} = \frac{2\tilde{C}_{13}^{(2)} \Gamma_a(T) - 2(\tilde{C}_{11} + \tilde{C}_{12}) \Gamma_c(T)}{2(\tilde{C}_{11} + \tilde{C}_{12})\tilde{C}_{33} - 4\tilde{C}_{13}^{(1)} \tilde{C}_{13}^{(2)}}. \tag{3}$$

Here, $\Gamma(T)$ and $\Gamma_c(T)$ are the derivatives of F_{phonon} with respect to $\delta a/a$ and $\delta c/c$, and the $\tilde{C}$ values are given by

$$\tilde{C}_{11} + \tilde{C}_{12} = C_{11} + C_{12} + (C_{113} + C_{123}) \frac{\delta c}{c} + \frac{1}{2}(C_{111} + 3C_{112}) \frac{\delta a}{a}, \tag{4}$$

$$\tilde{C}_{33} = C_{33} + \frac{1}{2} C_{333} \frac{\delta c}{c} + 2C_{133} \frac{\delta a}{a}, \tag{5}$$

$$\tilde{C}_{13}^{(1)} = C_{13} + \frac{1}{2}(C_{113} + C_{123}) \frac{\delta a}{a}, \tag{6}$$

$$\tilde{C}_{13}^{(2)} = C_{13} + \frac{1}{2} C_{133} \frac{\delta c}{c}. \tag{7}$$

The algebraic equations (2)–(7) are solved iteratively once $\Gamma_a(T)$ and $\Gamma_c(T)$ are obtained from first principles. Clearly, thermal expansion depends on both crystal elasticity and lattice anharmonicity. If the system is isotropic in lattice anharmonicity (i.e. $\Gamma_a = 2\Gamma_c$), the ratio of thermal expansions becomes inversely proportional to the ratio of elastic tensile stiffness in the basal plane and along the c axis.

The calculation of the CTEs for these complex structures imposes a tremendous challenge, since the CTE depends not only on the second-order elastic constants but also on higher-order elastic constants (dominated by third-order terms). For a tetragonal structure, there are six independent second-order elastic constants and ten independent third-order elastic constants. While the calculation of second-order elastic constants has become more common recently, the calculation of higher-order elastic constants remains challenging. This is because that the calculation of higher-order terms involves the *difference* in the elastic strain energies at a small change in lattice parameter. The higher-order elastic constants contribute to both crystal anharmonicity (through the dependence of the Debye temperature on the lattice parameter in our model) and the temperature dependence of elastic constants.

The elastic constants were calculated by the full-potential linearized augmented plane-wave (FLAPW) method (Wimmer *et al.* 1981) within the local-density-functional approach. The calculational details of the second-order elastic constants of a tetragonal structure have been described by Fu *et al.* (1999). The third-order elastic constants were determined through the derivative of second-order elastic constants with respect to lattice strains (i.e. $\delta a/a$ and $\delta c/c$). Within a range of 3–4% change in the lattice parameters, the dependence of the second-order elastic constants on lattice parameters is found to be linear, indicating that the third-order terms are dominant in the higher-order elastic constants. In calculations of the strained structure systems, the internal coordinates of the atoms within the unit cell were relaxed using the calculated FLAPW atomic forces.

§ 3. Results and discussion

The calculated second- and third-order elastic constants are listed in table 1 and table 2 respectively. We find excellent agreement between theory and experiments (Nakamura *et al.* 1990; Chu *et al.* 1999) in the second-order elastic constants for $MoSi_2$ and Mo_5Si_3. The measured second-order elastic constants are $C_{11} = 446$ GPa, $C_{12} = 174$ GPa, $C_{33} = 390$ GPa, $C_{13} = 140$ GPa, $C_{44} = 110$ GPa and $C_{66} = 140$ GPa for Mo_5Si_3, and $C_{11} = 417$ GPa, $C_{12} = 104.2$ GPa, $C_{33} = 514.5$ GPa, $C_{13} = 83.8$ GPa, $C_{44} = 204.2$ GPa and $C_{66} = 193.6$ GPa for $MoSi_2$. However, we are not aware of any measurement of higher-order terms. For Mo_5SiB_2, there is no experimental measurement to compare with, since single-crystal T2 phase is not currently available.

To examine the validity of our theory, we first consider $MoSi_2$. It has the simplest structure studied here. There are two interesting features in the elastic properties of $MoSi_2$.

(1) $C_{11} + C_{12} \approx C_{33}$ (i.e. the same degree of elastic rigidity in the basal plane and along the c-axis).

Table 1. Theoretical second-order elastic constants of $MoSi_2$, Mo_5Si_3 and Mo_5SiB_2.

	C_{11} (GPa)	C_{12} (GPa)	C_{33} (GPa)	C_{13} (GPa)	C_{44} (GPa)	C_{66} (GPa)
$MoSi_2$[a]	404	109	508	87	198	195
Mo_5Si_3[a]	438	162	371	136	106	143
Mo_5SiB_2	483	154	419	188	179	127

[a] At experimental lattice parameters.

Table 2. Theoretical third-order elastic constants of $MoSi_2$, Mo_5Si_3 and Mo_5SiB_2. Here $C_{111} + 3C_{112}$ and $C_{113} + C_{123}$ measure the 'softening rate' of the tensile modulus $C_{11} + C_{12}$ with [100] and [001] expansions respectively.

	$C_{111} + 3C_{112}$ (GPa)	$C_{113} + C_{123}$ (GPa)	C_{333} (GPa)	C_{133} (GPa)	C_{144} (GPa)	C_{344} (GPa)	C_{166} (GPa)	C_{366} (GPa)
$MoSi_2$	−4700	−1450	−4560	−950	−700	−600	−900	−750
Mo_5Si_3	−6520	−1540	−2830	−780	−370	−665	−820	−370
Mo_5SiB_2	−5700	−600	−2600	−200	−600	−770	−650	−200

(2) $C_{166} \approx C_{366}$ and $C_{144} \approx C_{344}$ (i.e. the response of shear elastic constants with respect to lattice expansions in different directions are approximately equal), indicating a small difference between the lattice anharmonicities in the [100] and [001] directions.

Indeed, the calculated CTEs of $MoSi_2$ shown in figure 2 (*a*) are nearly isotropic with magnitudes of about 8–10 ppm K^{-1} at high temperatures. The calculated CTEs are in good agreement with experiment (Thomas *et al.* 1985). The measured CTEs

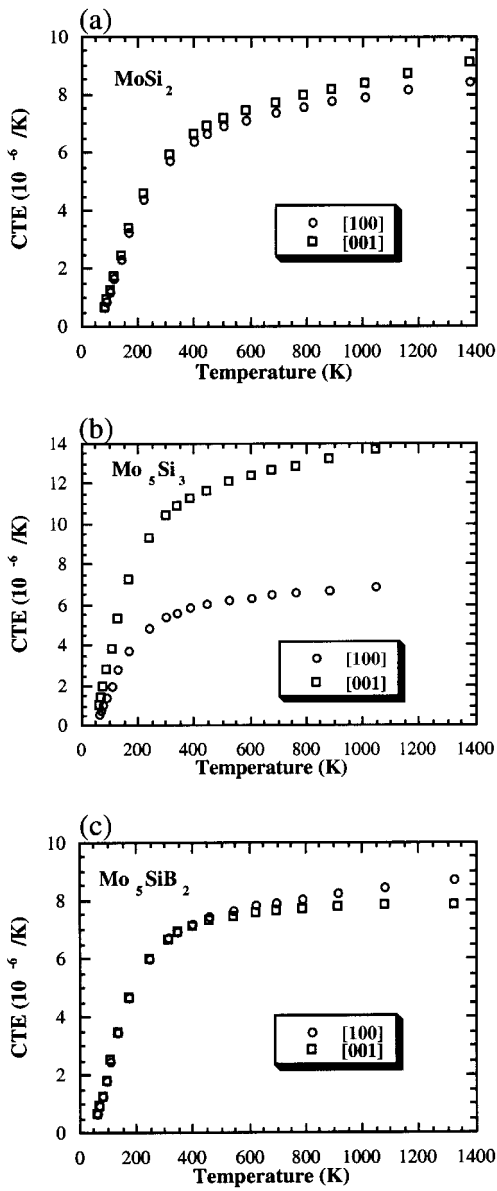


Figure 2. The calculated CTEs in the [100] and [001] directions for (*a*) $MoSi_2$, (*b*) Mo_5Si_3 and (*c*) Mo_5SiB_2.

are 8.2 and 9.4 ppm K^{-1} in the [100] and [001] directions respectively. This good agreement confirms that our model is valid for these strong covalent Mo–Si systems. In the following, we report our theoretical predictions for the 5–3 Mo–Si compounds.

The calculation for the CTE of Mo_5Si_3 becomes far more demanding owing to the complexity of its structure. The bonding in Mo_5Si_3 is shown to have pronounced multicentred covalent components (Fu *et al.* 1999), characterized by the planar Mo–Si–Mo bonding units in the basal plane and by the unusually short Mo–Mo bonds along the *c* axis (shorter by about 10% relative to the Mo—Mo distance in the basal plane). While the basal plane is found to be elastically more rigid than the *c* axis ($C_{11} + C_{12} > C_{33}$), the Mo–Mo bonds along the *c* axis will be shown to play a more important role in determining the anisotropy of lattice anharmonicity. Anisotropy is found in the elastic response to lattice expansions, which is evidenced in the marked difference in the magnitudes of calculated third-order terms: $C_{344} > C_{144}$ and $C_{366} < C_{166}$ for [100](001) and [100](010) shears respectively.

The anharmonicity can be described by examining the volume dependence of Debye temperature θ_D, that is the Grüneisen constant ($\gamma = -\partial(\ln \theta_D)/\partial(\ln V)$). We find that γ is higher for the lattice expanded in the [001] direction ($\gamma_{[001]} = 2.55$) than in the [100] direction ($\gamma_{[100]} = 2.17$). In other words, the lattice vibration energy decreases more rapidly by [001] expansions than by [100] expansions. The calculated CTEs of Mo_5Si_3 shown in figure 2 (*b*) are consistent with the experimental measurement by Chu *et al.* (1999). The CTE along the [001] direction is about twice that in the [100] direction. The measurements were made in the range from 300 to 700 K with fitted CTEs of 5.2 and 11.5 ppm K^{-1} in the [100] and [001] directions respectively. It should be noted, however, that the calculated CTEs still increase linearly with increasing temperature at higher temperatures (in particular, in the [001] direction).

An examination of the third-order elastic constants of $MoSi_2$ and Mo_5Si_3 shows the difference between their elastic responses to lattice expansions. In $MoSi_2$, the response of either C_{44} or C_{66} to lattice expansions is nearly isotropic. By contrast, the corresponding shear elastic response in Mo_5Si_3 is anisotropic. Because of a planar covalent bonding in the basal plane in Mo_5Si_3, the finding that $C_{366} < C_{166}$ for (001) *intralayer* shear is not entirely surprising (i.e. C_{66} decreases more rapidly by [100] expansions than by [001] expansions). The large difference (by a factor of two) between C_{344} and C_{144} for (001) *interlayer* shear, however, is unexpected, since an increase in lattice spacing in either the [001] or the [100] direction has the same effect to decrease the [100](001) shear elastic strength.

Analysis shows that, as the lattice spacing is varied, the change in Debye temperature (i.e. lattice anharmonicity) is particularly sensitive to the variation in shear elastic constants. In the case of Mo_5Si_3, the major contribution to the anisotropy in anharmonicity comes from the anisotropy in C_{344} and C_{144} (i.e. C_{44} decreases more rapidly by [001] expansions than by [100] expansions). (Note that the response of [100](001) shear (i.e. C_{44}) is about twice as important as the response of [100](010) shear (i.e. C_{66}), since C_{44} and C_{55} are degenerate for tetragonal systems.) The physical origin for this anisotropy lies in the dominant role of the unusually short Mo–Mo [001] covalent bonds in coupling the (001) layers. These covalent Mo–Mo bonds, which characterize the [100](001) shear, are weakened more by [001] expansions than by [100] expansions. Increasing the [001] spacing reduces the Mo–Mo [001] bond strength and decreases the lattice vibrational energy (giving

higher lattice anharmonicity) more rapidly than in-plane expansions. This directional dependence of the Mo–Mo bond strength (in coupling the (001) layers) contributes significantly to the difference in the softening behaviour of the [100](001) shear characteristics.

Thus, for Mo_5Si_3, there are two physical sources contributing to a higher CTE along the [001] direction than along the [100] direction: a stronger bonding in the basal plane ($C_{11} + C_{12} > C_{33}$) and a higher lattice anharmonicity along the c axis (dominated by $C_{344} > C_{144}$). Both of these physical mechanisms can be understood in terms of electronic structure: the existence of planar multicentred Mo–Si–Mo covalent bonds in the basal plane, and the dominant role of directional Mo–Mo bonds along the c axis in the (001) interlayer coupling. In fact, we believe that the same mechanisms are also responsible for the observed CTE anisotropy in other 5–3 transition-metal silicides. For example, although Ti_5Si_3 has a different crystal structure (hexagonal $D8_8$ structure) from the tetragonal $D8_m$ structure discussed above, the interlayer coupling between basal planes is still dominated by atomic chains along the c axis with unusually short interatomic distance. As a result, the calculated CTEs of Ti_5Si_3 are also highly anisotropic. It follows that it is possible to reduce the CTE anisotropy if the interlayer coupling between basal planes is no longer dominated solely by the [001] bonding component. This can probably be achieved either by interstitial alloying additions (to modify bonding direction) or by alloying substitutions (to increase the interatomic distance along the chains). These suggestions are currently examined by experiments.

The substitution of boron in Mo–Si changes the crystal structure from $D8_m$ (T1 phase) to $D8_l$ (T2 phase). One notable feature of the T2 phase is the absence of transition-metal atomic chains along the c axis in this structure. The calculated lattice parameters are 6.027 and 10.97 Å for a and c respectively.

For Mo_5SiB_2, the averaged elastic moduli are higher than those of Mo_5Si_3. This increase is partly attributed to the formation of Mo–B covalent bonds. While the calculated second-order elastic constants for the T2 phase still indicate that the basal plane is elastically more rigid than the c axis ($C_{11} + C_{12} > C_{33}$), the effect of this difference on the CTE is balanced by a substantial increase in the elastic coupling (C_{13}) between the basal plane and c axis, presumably owing to the effect of boron at interstitial sites (cf. figure 1). Most significantly, the lattice anharmonicity for the T2 phase is found to be nearly isotropic in the [100] and [001] directions (with Grüneisen constants of 2.05 and 1.98 respectively). We identify this near isotropy in lattice anharmonicity as being due to the absence of a directionally bonded [001] chain structure characteristic of the T1 phase. Indeed, in this case, the difference between C_{344} and C_{144} becomes smaller and is balanced by a relatively larger difference between C_{166} and C_{366}.

As a result of decreased anisotropy in both the elastic (static) contribution and the lattice anharmonicity compared with the T1 phase, the CTEs for the T2 phase become nearly isotropic in [100] and [001] directions with the CTE in the [100] direction being slightly higher. The calculated CTEs for the T2 phase are presented in figure 2 (c). Experimentally, it was observed that, in the processing of these alloys, grain-boundary cracking problems (characteristic of the T1 phase) are virtually eliminated in the T2 phase, indicating that the CTE anisotropy in the T2 phase is much less than that of the T1 phase, in agreement with theoretical prediction.

More recently, the CTEs of Mo_5SiB_2 were measured by neutron powder diffraction (Rawn *et al.* 2000) and by synchrotron X-ray diffraction (Kramer 2000) from

room temperature to 1400°C. The experiments confirm the theoretical prediction not only of the CTE values but also of a nearly isotropic CTE for Mo_5SiB_2.

In summary, we have shown that the CTEs of complex systems can be predicted by first-principles calculation. The high CTE anisotropy in 5–3 silicides is understood in terms of the anisotropy in lattice anharmonicity. The removal of atom chains (characteristic of the T1 phase) along the *c* axis by boron substitutions substantially decreases the CTE anisotropy in the T2 phase.

Acknowledgements

Research was sponsored by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy under contract DE-AC05-96OR 22464 with Lockheed Martin Energy Research Corporation.

References

Ashcroft, N. W., and Mermin, N. D., 1976, *Solid State Physics* (New York: Holt, Rinehart and Winston), p. 443.

Chu, F., Thoma, D. J., McClellan, K., Peralta, P., and He, Y., 1999, *Intermetallics*, **7**, 611.

Fleszar, A., and Gonze, X., 1990, *Phys. Rev. Lett.*, **64**, 2961.

Fu, C. L., Wang, X., Ye, Y. Y., and Ho, K. M., 1999, *Intermetallics*, **7**, 179.

Kramer, M. J., 2000 (to be published).

Meyer, M. K., Kramer, M|. J., and Akinc, M., 1996, *Intermetallics*, **4**, 273.

Moruzzi, V. L., Janak, J. F., and Schwarz, K., 1988, *Phys. Rev.* B, **37**, 790.

Nakamura, M., Matsumoto, S., and Hirano, T., 1990, *J. Mater. Sci.*, **25**, 3309.

Quong, A. A., and Liu, A. Y., 1997, *Phys. Rev.* B, **56**, 7767.

Rawn, C. J., Hoffmann, C. M., Schneibel, J. H., and Hubbard, C. R., 2000, *Intermetallics* (to be published).

Shah, D. M., Berczik, D., Anton, D. L., and Hecht, R., 1992, *Mater. Sci. Engng*, **A155**, 45.

Thomas, O., Senateur, J. P., Madar, R., Laborde, O., and Rosencher, E., 1985, *Solid St. Commun.*, **55**, 629.

Wimmer, E., Krakauer, H., Weinert, M., and Freeman, A. J., 1981, *Phys. Rev.* B, **24**, 864.

PHILOSOPHICAL MAGAZINE A, 2000, VOL. 80, NO. 4, 871–880

First-principles calculation of stacking fault and twin boundary energies of Cr_2Nb

SUKLYUN HONG†, C. L. FU and M. H. YOO

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6114, USA

[*Received 15 July 1998 and accepted 7 June 1999*]

ABSTRACT

The stacking fault and twin boundary energies of C15 Cr_2Nb are calculated by the first-principles local-density-functional approach. It is found that the intrinsic and extrinsic stacking fault energies are 116 and 94 $mJ\,m^{-2}$, respectively, and the twin boundary energy is 39 $mJ\,m^{-2}$. The lower extrinsic stacking fault energy is consistent with the fact that the C36 structure has a lower energy than the C14 structure. The calculated stacking fault energies at 0 K are larger than the experimental values available in the literature. The equilibrium separations between Shockley partials based on the calculated elastic constants and stacking fault energies are also calculated.

§ 1. INTRODUCTION

For a wide variety of structural applications, Laves phases have some unique properties such as high melting temperature, low density, and high oxidation resistance. Unfortunately, this potential has not been well exploited, largely because of low temperature brittleness due to the lack of plastic deformation. Cr_2Nb, one of the most studied Laves phase compounds, has either the cubic C15 or hexagonal C14 (or C36) structure, which is topologically akin to the face-centred cubic (fcc) (A1) or the hexagonal close-packed (hcp) (A3) structure, or any one of the polytypic phases. The topologically close-packed (TCP) plane of the C15 structure is of the $\{111\}$ type, and the potential modes of plastic deformation in Cr_2Nb are twelve $\{111\}\langle 110\rangle$ slip systems and twelve $\{111\}\langle 112\rangle$ twin systems. Since each of the TCP units consists of the quadruple atomic layers, the slip, twinning, or stress-induced polytypic transformation will require a coordinated process of atomic motions, such as synchroshear, in order to effect the motion of Shockley partial dislocations.

Among other quantities, the knowledge of stacking fault energy (SFE), γ_{SF}, is necessary to understand the deformation mechanism, since SFE will play an important role in processes such as dislocation dissociation, cross-slip and twinning. First-principles calculations are useful to estimate the energetics of stacking faults (SFs) and understand the interaction between Shockley partials.

Using the linear muffin-tin orbital (LMTO) method, Chu *et al.* (1995a) estimated the intrinsic stacking fault energy of Cr_2Nb from the structural energy difference between the C15 and C14 structures, and obtained $\gamma_{SF} = 90\,mJ\,m^{-2}$. On the other

† Present address: Department of Physics, Sejong University, Seoul 143-747, Korea.

Philosophical Magazine A ISSN 0141–8610 print/ISSN 1460-6992 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journ als/tf/01418610.html

hand, Yoshida *et al.* (1995) obtained a value of $8\,\mathrm{mJ\,m^{-2}}$ for SFE from transmission electron microscopy (TEM) observations of extended dislocation nodes at 1623°C. More recently, Kazantzis *et al.* (1996) obtained a value of $25\,\mathrm{mJ\,m^{-2}}$ from TEM observations of extended nodes in Cr_2Nb at 1400 and 1500°C.

In a previous paper (Hong and Fu 1999), we investigated the phase stability of three Laves phases (C15, C14 and C36) of Cr_2Nb. It was found that the C15 phase is the ground-state structure with the lowest energy and the C36 phase is an intermediate state between C15 and C14. These three phases, however, are very close in energy, i.e. within a range of about 60 meV/formula unit (Hong and Fu 1999) indicating the possibility of low stacking fault energies in this system. In this paper, we report the calculation of SFE using supercell geometry, and an evaluation of the equilibrium separation between Shockley partials using the anisotropic elastic theory.

§ 2. Laves structures

In describing the structures of Laves phases and their stacking faults and twin boundary, we follow the notations of Hazzledine (1994).

Laves phases have ideal chemical compositions S_2L; they contain smaller atoms S and larger atoms L in alternate sheets parallel to the TCP planes (i.e. (111) plane for the C15 structure and (0001) plane for the C14 and C36 structures). The main geometric characteristic of Laves phases S_2L is that they consist of two types of atomic stacking sequence, $\alpha A\alpha$ ($\beta B\beta$ and $\gamma C\gamma$) and $\alpha c\beta$ ($\beta a\gamma$ and $\gamma b\alpha$). Here, the Greek letters (α, β, γ) denote the L atoms, while lower case Latin letters (a, b, c) and capital letters (A, B, C) denote the type-1 and type-2 S atoms, respectively. Note that the $\alpha c\beta$-type stackings are more closely spaced (in terms of the interlayer spacings). For Cr_2Nb, Latin and Greek letters represent Cr and Nb, respectively.

Figure 1 represents an atomic layer of A atoms based on the hexagonal unit cell on the TCP plane (a Kagome net) which is determined by two lattice vectors

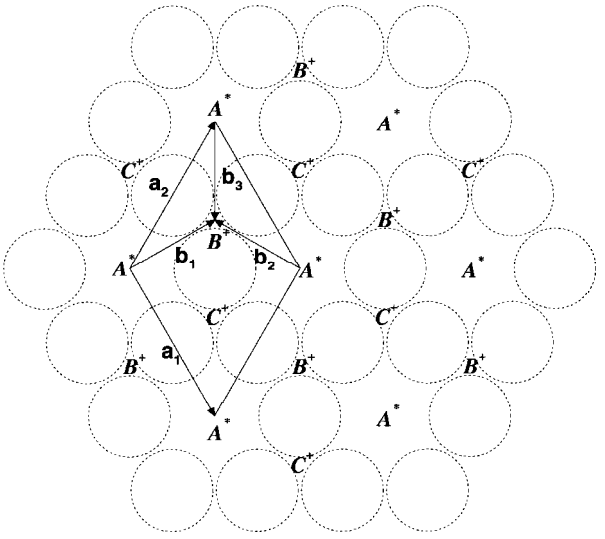


Figure 1. The hexagonal unit cell on the basal planes which is determined by two lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. See the text for details.

$\mathbf{a}_1 = \frac{1}{2}[0\bar{1}1]$ and $\mathbf{a}_2 = \frac{1}{2}[\bar{1}10]$. The vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are Shockley partial Burgers vectors of the type $\frac{1}{6}\langle\bar{1}12\rangle(\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 0)$. The dotted circles in figure 1 denote A atoms, one of the type-2 S atoms, projected on the (111) plane. Let the centres of hexagonal rings consisting of A atoms be denoted by A^*, and thus the A atoms are placed midway between the lattice points A^*. Similarly, the lattice points B^* and C^*, which determine B and C atoms respectively, can be obtained by the displacement of A^* by $\mathbf{b}_1$ and $-\mathbf{b}_1$, respectively, on higher or lower planes. Also, we denote the projected points of B^* and C^*onto the plane given in figure 1 as B^+ and C^+, respectively, which are the saddle points for larger α atoms. Then, the projected atomic positions of α (or a) onto the given plane are at the sites of A^* in figure 1, and the projected positions of β (or b) and γ (or c) are at those of B^+ and C^+, respectively. Consequently, the TCP C15 structure is defined by repeated XYZ stacking:

$$\text{C15}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{1}$$

where the TCP plane unit X (Y or Z) consists of one single layer of A (B or C) atoms and one triple layer of αcβ (βaγ or γbα). A synchroshear mechanism was first introduced by Kronberg (1957), and was used to explain the deformation twinning process by several groups (Livingston and Hall 1990, Chu and Pope 1993, Hazzledine 1994). Synchroshearing of the X unit (Aαcβ) creates the X′ unit (Aαbγ), and similarly for the Y′ (Bβcα) and Z′ (Cγaβ) units. The C14 and C36 Laves phases have repeated X′Z and XYZ′Y′ stackings, respectively, as follows:

$$\text{C14}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{b}\gamma}^{\mathbf{X'}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{2}$$

$$\text{C36}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{a}\beta}^{\mathbf{Z'}}\,\overbrace{\mathrm{B}\beta\mathrm{c}\alpha}^{\mathbf{Y'}}\cdots. \tag{3}$$

It can be seen that the primed X′, Y′, Z′ units are introduced by synchroshear in order to make sure that TCP structure is maintained in C14 and C36. For example, c and β of the X unit in C15 are synchrosheared by $-\mathbf{b}_2$ and $-\mathbf{b}_1$, respectively, to become b and γ of X′ unit. This displacement preserves a close packing between X and Z units in C14.

The stacking sequences for an intrinsic stacking fault (ISF) and an extrinsic stacking fault (ESF) and a twin boundary are given in table 1 in terms of X, Y, Z units. Note that the ISF contains a local C14-like structure in C15, while the ESF contains a local C36-like structure in C15. Figure 2 shows schematic illustrations of an ISF and an ESF obtained from C15 through the synchroshear mechanism: (*a*) C15 → ISF, and (*b*) C15 → ESF. The C15 stacking has a repeated XYZ sequence (no stacking fault). For the ISF, one unit, namely Y, is missing from the C15 XYZ sequence, and X becomes X′ by synchroshear (shifts by $-\mathbf{b}_2$ and $-\mathbf{b}_1$ shown in figure 2 (*a*)), while for an ESF a unit Y′ is added into the original C15 sequence. For the ESF, two successive synchroshears (two synchro-shifts shown in figure 2 (*b*)) are operative in the αcβ type to maintain a TCP structure by introducing two primed units, Z′ and Y′. Twinned structure is equivalent to a mirror reflection about the X′ unit (see table 1).

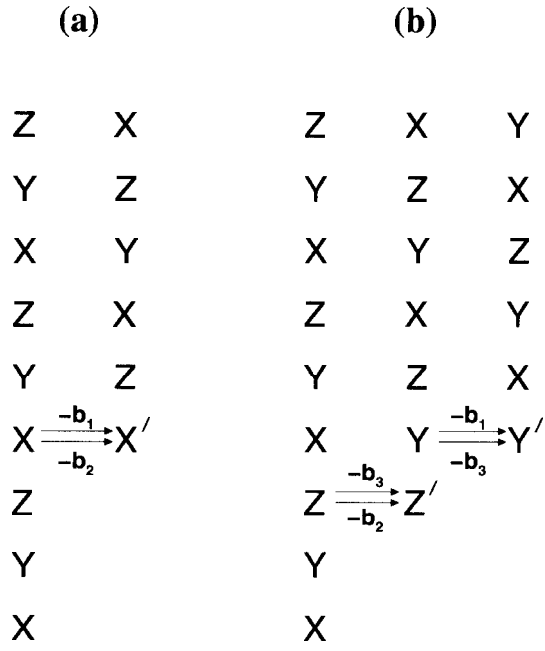


Figure 2. Schematic illustrations of an ISF and an ESF obtained from C15 → ISF, and (*b*) C15 → ESF.

Table 1. Stacking sequences of the C15 structure (no stacking fault), intrinsic stacking fault, extrinsic stacking fault, and twin (C15 + C15T) structure.

	Sequence of faults
C15	···XYZXYZXYZ···
Intrinsic stacking fault	···XYZX′ZXYZ···
Extrinsic stacking fault	···XYZXYZ′Y′XYZ···
Twin (C15 + C15T)	···XYZX′Z′Y′X′Z′Y′···

§ 3. Stacking fault and twin boundary energies

Total-energy calculations for the stacking fault energies of C15 Cr_2Nb are performed using the full-potential linearized augmented plane-wave (FLAPW) method (Wimmer *et al.* 1981) within the local-density approximation. The FLAPW method solves the local-density-functional equations without any shape approximation to the potential or charge density. The atomic positions are relaxed by calculating Hellmann–Feynman forces acting on the atoms.

The supercell geometry is used to obtain the energies of ISF, ESF and twin boundary. In the supercell calculation, we use the experimental lattice constant (6.991 Å) of C15 Cr_2Nb and spacings along $\mathbf{a}_3$ axis corresponding to the ideal hcp c/a ratio, which gives ~4. 04 Å for the thickness of each X, Y, Z unit. We consider a supercell containing XYZX′Z for an ISF and XYZXYZ′Y′ for an ESF. In these supercells, the separations between ISF and ESF planes are about 20 Å and 28 Å,

respectively. Also, a supercell containing XYZX′Z′Y′ (~24 Å) is considered for twin boundary. The supercell for twinning contains two twin boundaries in it, thus the separation between twin boundaries is about 12 Å.

The internal coordinates of each structure are fully relaxed from the ideal positions (defined by the atomic positions of the C15 lattice) by calculating Hellmann–Feynman forces acting on the atoms. The relaxation energies of ISF, ESF, and twin boundary are 41, 20, and 17 mJ m^{-2}, respectively.

The results for the stacking fault and twin boundary energies are given in table 2. It is found that the fault energies are γ_{ISF} = 116 and γ_{ESF} = 94 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is γ_T = 39 mJ m^{-2}. Since ESFs and ISFs contain local C36-like and C14-like structures, respectively, a lower ESF energy (γ_{ESF}) compared to the ISF energy (γ_{ISF}) is consistent with the fact that the C36 structure has a lower energy than the C14 structure (Hong and Fu 1999).

For comparison, we also calculated the stacking fault and twin boundary energies using the theoretical lattice constant (6.822 Å). The relaxation energies of ISF, ESF, and twin boundary are 48, 40, and 21 mJ m^{-2}, respectively. It is also found that the fault energies are 140 and 108 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is 52 mJ m^{-2}. The results using the theoretical lattice constant are slightly larger than those using the experimental one.

Although the intrinsic faults are expected to prevail in fcc crystals rather than the extrinsic faults (Hirth and Lothe 1982), extrinsic stacking faults were observed to be dominant in Laves phase Co_2Ti (Allen *et al.* 1972), and Nb-doped HfV_2 (Chu *et al.* 1998). Since the calculated γ_{ESF} is smaller than γ_{ISF}, it is likely that the observed SFs in Cr_2Nb are also of extrinsic type.

We compare our results with others. Chu *et al.* (1995a) obtained γ_{ISF} = 90 mJ m^{-2} by an estimation from the structural energy difference between the bulk C14 and C15 structures. This value is close to our value of 116 mJ m^{-2}. On the other hand, Yoshida *et al.* (1995) observed extended dislocation nodes in C15 Cr_2Nb deformed at 1623 K, in which the SFs are bounded by three Shockley partials (of the type $\frac{1}{6}\langle 112\rangle$) with a radius of curvature of R. They obtained a smaller value of 8 mJ m^{-2}. Note that Chu *et al.* (1995a) re-estimated SFE from the data of Yoshida *et al.* (1995) to obtain γ_{SF} = 15–60 mJ m^{-2}. More recently, Kazantzis *et al.* (1996) obtained a value of 25 mJ m^{-2} from TEM observations of extended triple-junction nodes in Cr_2Nb at 1400 and 1500°C. These results are tabulated in table 2.

Our results for SFEs are higher than the experimental results. Certainly, SFE obtained at high temperatures can be expected to be lower than the calculated value at 0 K, which has also been suggested by Kazantzis *et al.* (1996). The unknown image

Table 2. Stacking fault and twin boundary energies for C15 Cr_2Nb, using the experimental lattice constant.

Method	γ_{SF} (mJ m^{-2})		γ_T (mJ m^{-2})
This work	116 (intrinsic)	94 (extrinsic)	39
LMTO[a]	90 (intrinsic)		—
Exp[b]	25		—
Exp[c]	>8		—

[a] Chu *et al.* (1995a): estimated from energy difference between C14 and C15.
[b] Kazantzis *et al.* (1996): estimated from dislocation triple-junction.
[c] Yoshida *et al.* (1995): estimated from dislocation triple-junction.

shift of Shockley partial dislocations can also give an uncertainty in the SFE determination; for example, for SFE in silicon, $\gamma_{SF} = 30\,\mathrm{erg\,cm^{-2}}$ has an uncertainty in the range $15 < \gamma_{SF} < 75\,\mathrm{erg\,cm^{-2}}$, after taking into account the image shift (Aerts *et al.* 1962). From these points of view, the temperature and image shift effects may be responsible for the difference between our calculated value and experimental results.

Our theoretical calculations on Cr_2Nb were performed for the stoichiometric composition. In experiments, however, some localized variations in chemical composition are possible in polycrystalline compounds. For instance, the two Cr_2Nb alloys investigated by Yoshida *et al.* (1995) were Cr–32. 2% Nb and Cr-34. 0% Nb, in which the second phase particles observed are Cr solid solution in the former and Nb solid solution in the latter. Also, polycrystals used in experiments may have compositions that deviate from the ideal stoichiometry ratio, since the phase region of the C15 Cr_2Nb is relatively large. For example, as mentioned by Hong and Fu (1999), the calculated elastic moduli for the stoichiometry alloy at the experimental lattice constant are very different from the experimental values (Chu *et al.* 1995b) obtained from polycrystals. Therefore, this composition effect can also be partially responsible for the discrepancy between theory and experiment.

The calculated twin boundary energy, $\gamma_T = 39\,\mathrm{mJ\,m^{-2}}$, at the experimental lattice constant for Cr_2Nb is relatively low, for instance, in comparison to $\gamma_T = 60\,\mathrm{mJ\,m^{-2}}$ for TiAl of the Ll_0 structure (Fu and Yoo 1990). In view of the energetics of twin nucleation, such a low twin boundary energy suggests a high propensity of twinning in Cr_2Nb. This is consistent with the experimental observations of twinned microstructures in Cr_2Nb, formed due to plastic deformation at elevated temperatures (Yoshida *et al.* 1995) and to the internal stresses resulting from the C14–C15 transformation as well as the thermal contraction differences between Cr solid solution and Cr_2Nb in the two-phase alloy (Kumar and Liu 1997). However, the absence of deformation twinning in Cr_2Nb at low temperatures is not understood. Kinetic aspects of the motion of synchro-Shockley partials need to be elucidated in order to better understand twin formation in the C15 Laves phase.

§ 4. Interaction between Shockley partials

From the anisotropic elasticity theory (Stroh 1958, Hirth and Lothe 1982) we calculate the equilibrium separation between Shockley partials, using the calculated elastic constants (Hong and Fu 1999) and stacking fault energies.

First, we consider the following case for the ISF:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \mathrm{ISF} + \frac{1}{6}[2\bar{1}\bar{1}], \tag{4}$$

where $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$, $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$ correspond to $-\mathbf{a}_2$, $\mathbf{b}_3$, and $-\mathbf{b}_1$ in figure 1, respectively. This expression for the formation of an ISF is approximate since the Shockley partial $\mathbf{b}_3 = \frac{1}{6}[121]$ is the sum of two synchro-Shockley partial vectors $-\mathbf{b}_2$ and $-\mathbf{b}_1$, lying in the two successive atomic planes, shown in figure 2 (*a*). The equilibrium separation between Shockley partials can be obtained through the balance of the attractive force (the surface tension due to the ISF) and the repulsive force (due to elastic interaction between the partials).

The radial and tangential components of the interaction force (per unit length) between two parallel dislocations can be calculated by

$$F_r = f_r/(2\pi r), \tag{5}$$

$$F_\theta = f_\theta/(2\pi r). \tag{6}$$

Here, r is the separation between the two partials of Burgers vectors $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$, and f_r and f_θ are the radial (in the $\{111\}$ plane) and tangential (out-of-the plane) components of the interaction force constants, which are determined from the anisotropic elasticity theory of dislocations (Stroh 1958). For the dissociation of equation (4), the interaction force constants f_r and f_θ are obtained numerically (Yoo 1987), by (i) letting $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$, (ii) using the experimental lattice constant $a_0 = 6.991$ Å for C15, and (iii) the calculated elastic constants and γ_{ISF} at experimental lattice constant.

As shown in figure 3, the radial component f_r increases monotonically from $\phi = 0$ (screw) to $\phi = 90°$ (edge), where ϕ is the angle between a dislocation line and the Burgers vector **B**. The tangential component f_θ is maximum in magnitude at $\phi = 31°$ and zero at the edge orientation. At $\phi \sim 30°$, the two parallel Shockley partials are inclined at about 60° and 0° to their respective Burgers vectors. This implies that while the repulsive F_r balances the surface tension γ_{SF} (see equation (7) below), the out-of-plane force ($F_\theta = 0.5F_r$ at $\phi = 30°$) promotes cross-slip of the screw Shockley partial and climb of the 60° Shockley partial, most likely onto the $(\bar{1}01)$ plane. This Shockley partial dissociation configuration of the cross-slip and climb combination may lead to a possible mechanism for the thickening process of a twin embryo originating from a mixed ($\phi = 30°$) $\frac{1}{2}[1\bar{1}0]$ dislocation.

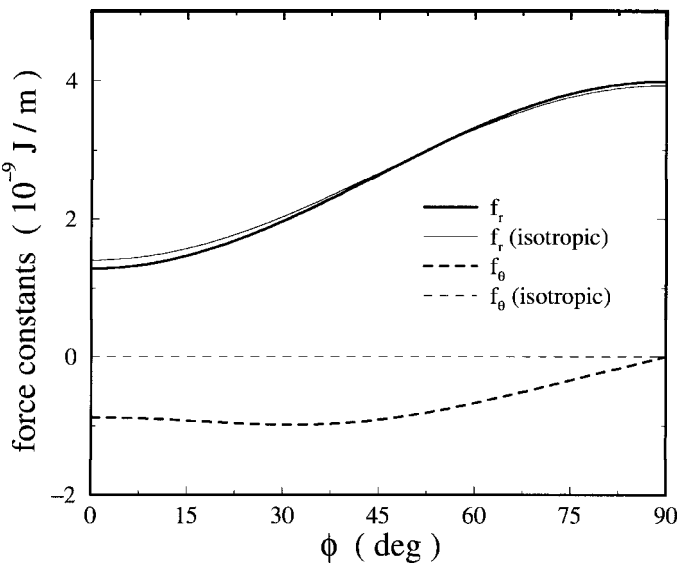


Figure 3. Orientation dependence of the interaction force constants in Cr_2Nb for the ISF-type dissociation, using the results obtained at the experimental lattice constant. ϕ is an angle between a dislocation line and the Burgers vector $\frac{1}{2}[1\bar{1}0]$.

The width of the equilibrium separation, w, can be obtained from

$$\gamma_{\mathrm{SF}} = F_r = f_r/(2\pi w), \tag{7}$$

$$w = f_r/(2\pi\gamma_{\mathrm{SF}}). \tag{8}$$

The separation is $w_{\mathrm{s}} = 17.6$ Å for the $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$ screw dislocation, and $w_{\mathrm{e}} = 54.7$ Å for the edge dislocation. These results are shown in table 3, together with those in the isotropic case (discussed below). Also, the results using γ_{ESF} and the experimental result are tabulated.

For the isotropic case, the equilibrium separation w is given by a simple formula (Hirth and Lothe 1982):

$$w = \frac{Gb^2}{8\pi\gamma_{\mathrm{SF}}}\frac{2-\nu}{1-\nu}\left(1 - \frac{2\nu\cos 2\phi}{2-\nu}\right). \tag{9}$$

Here, $b = |\mathbf{b}^{(1)}| = |\mathbf{b}^{(2)}| = a_0/6^{1/2}$. The Hill's average values of shear modulus and the Poisson's ratio are $G = 50.0$ GPa and $\nu = 0.383$, respectively. Using γ_{ISF}, it is found that $w_{\mathrm{s}} = 19.3$ Å and $w_{\mathrm{e}} = 53.9$ Å. As given in table 3, the results in anisotropic and isotropic cases are rather close, because the shear anisotropy of C15 Cr_2Nb is moderate, $A = 1.45$ (Hong and Fu 1999), compared to the isotropic case ($A = 1$).

The radial component f_r of the isotropic case is very close to that of the anisotropic one, whereas its tangential component f_θ is exactly zero (figure 3). While the anisotropic corrections to f_r and hence to w in equation (9) are very small, the anisotropic tangential component of the elastic interaction force is quite large; for instance, $f_\theta/f_r \sim 0.7$ at $\phi = 0$. This indicates that cross-slip of $\frac{1}{2}[1\bar{1}0]$ dislocation is difficult in Cr_2Nb because of a large constriction energy for the Shockley partials that include the non-radial component of the interaction energy.

Next, let us consider the ESF case:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \frac{1}{6}[11\bar{2}] + \mathrm{ESF} + \frac{1}{6}[2\bar{1}\bar{1}] + \frac{1}{6}[\bar{1}\bar{1}2] \tag{10}$$

$$\approx \frac{1}{6}[2\bar{1}\bar{1}] + \mathrm{ESF} + \frac{1}{6}[1\bar{2}1]\,. \tag{11}$$

To calculate the width of the ESF, we approximate two Shockley partials on two successive TCP units on either side of the ESF by a single Shockley partial (as in the ISF case), and use the same formulas for the ISF. In other words, the only difference

Table 3. Equilibrium separation between Shockley particles bounding an ISF and an ESF, using the results obtained at experimental lattice constant. w_{s} and w_{e} are the width for screw and edge dislocations, respectively. The numbers are in angstroms.

		w_{s} (screw)		w_{e} (edge)	
		ISF	ESF	ISF	ESF
Theory ($T = 0$ K)	Anisotropic	18	22	55	68
	Isotropic	19	24	54	67
Exp[a] ($T = 1400$°C)		99		—	

[a] Kazantzis *et al.* (1996): measured from the ribbon after the correction for image shift and projection effects.

between ISF and ESF in the calculation of separation w is that the stacking fault energy used in the formulas is of intrinsic or extrinsic type. Using γ_{ESF}, it is found that $w_s = 21.7$ Å and $w_e = 67.5$ Å for the anisotropic case (table 3). For the isotropic case, they are $w_s = 24.4$ and $w_e = 66.5$ Å. As mentioned, the results of both anisotropic and isotropic cases are very close. Since γ_{ESF} is not much different from γ_{ISF}, the force constants f_r and f_θ using γ_{ESF} are expected to be similar to those in figure 3 using γ_{ISF}.

On the other hand, we consider the separation between partials by using the theoretical lattice constant. For the ISF, the separation is $w_s = 17.1$ Å and $w_e = 53.5$ Å for the anisotropic case, while $w_s = 18.7$ Å and $w_e = 52.8$ Å for the isotropic case. For the ESF, the separation is $w_s = 22.2$ Å and $w_e = 69.4$ Å for the anisotropic case, while $w_s = 24.2$Å and $w_e = 68.4$ Å for the isotropic case. It can be seen that the separations are very close when using both experimental and theoretical lattice constants.

Kazantzis *et al.* (1996) measured, at $T = 1400°C$, the width of the ribbon (separation between two parallel partial dislocations) as 99 Å after correction for image shift and projection effects. Note that they also obtained a different value of $w_s = 82$ Å for the width using equation (9) with $\gamma_{SF} = 25\,mJ\,m^{-2}$, which was estimated from the measured curvature of R in extended dislocation nodes and the temperature-corrected shear modulus. Compared with the experimental result at high temperatures, our results for the separation between partials are very small.

§ 5. Summary

We performed first-principles total-energy calculations to obtain stacking fault and twin boundary energies. The intrinsic and extrinsic stacking fault energies were calculated to be 116 and 94 mJ m^{-2}, respectively, and the twin boundary energy was 39 mJ m^{-2}. The calculated stacking fault energies are larger than the available experimental data, measured at high temperatures. We also calculated the equilibrium separations between Shockley partials using the calculated elastic constants and stacking fault energies. Our results of the equilibrium separations are very small compared with the experimental results reported at high temperatures. This discrepancy may be due to temperature and/or composition and image shift effects.

Acknowledgements

The research was sponsored by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

References

Aerts, E., Delavignette, P., Siems, R., and Amelinckx, S., 1962, *J. appl. Phys.*, **33**, 3078.
Allen, C. W., Delavignette, P., and Amelinckx, S., 1972, *Phys. status solidi* (a), **9**, 237.
Chu, F., Ormeci, A. H., Mitchell, T. E., Wills, J. M., Thoma, D. J., Albers, R. C., and Chen, S. P., 1995a, *Phil. Mag. Lett.*, **72**, 147.
Chu, F., He, Y., Thoma, D. J., and Mitchell, T. E., 1995b, *Scripta metall. mater.*, **33**, 1295.
Chu, F., Lu, Y.-C., Kotula, P. G., Mitchell, T. E., and Thoma, D. J., 1998, *Phil. Mag. A*, **77**, 941.

CHU, F., and POPE, D. P., 1993, *High-Temperature Ordered Intermetallic Alloys V*, *Materials Research Society Symposium Proceedings*, Vol. 288, edited by I. Baker, R. Darolia, J. D. Whittenberger and M. H. Yoo (Materials Research Society), p. 561.
FU, C. L., and YOO, M. H., 1990, *Phil. Mag. Lett.*, **62**, 159.
HAZZLEDINE, P. M., 1994, *Twinning in Advanced Materials*, edited by M. H. Yoo and M. Wuttig (The Minerals, Metals & Materials Society), p. 403.
HAZZLEDINE, P. M., and PIROUZ, P., 1993, *Scripta metall. mater.,* **28**, 1277.
HIRTH, J. P., and LOTHE, J., 1982, *Theory of Dislocations* (New York: Wiley).
HONG, S., and FU, C. L., 1999, *Intermetallics*, **7**, 5.
KAZANTZIS, A. V., AINDOW, M., and JONES, I. P., 1996, *Phil. Mag. Lett.*, **74**, 129.
KRONBERG, M. L., 1957, *Acta metall.,* **5**, 507.
KUMAR, K. S., and LIU, C. T., 1997, *Acta mater.,* **45**, 3671.
LIVINGSTON, J. D., and HALL, E. L., 1990, *J. mater. Res.,* **5**, 5.
STROH, A. N., 1958, *Phil. Mag.,* **3**, 625.
WIMMER, E., KRAKAUER, H., WEINERT, M., and FREEMAN, A. J., 1981, *Phys. Rev.* B, **24**, 864.
YOO, M. H., 1987, *Acta metall.,* **35**, 1559.
YOSHIDA, M., TAKASUGI, T., and HANADA, S., 1995, *High-Temperature Ordered Intermetallic Alloys VI*, *Materials Research Society Symposium Proceedings*, Vol. 364, edited by J. Horton, I. Baker, S. Hanada, R. D. Noebe and D. S. Schwartz (Materials Research Society), p. 1395.

The effect of Ta on the magnetic thickness of permalloy ($Ni_{81}Fe_{19}$) films

M. Kowalewski, W. H. Butler, N. Moghadam, G. M. Stocks, and T. C. Schulthess
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

K. J. Song and J. R. Thompson
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

A. S. Arrott, T. Zhu, J. Drewes, and R. R. Katti
Honeywell, Inc., MN14-3C15, 12001 State Highway 55, Plymouth, Minnesota 55441

M. T. McClure and O. Escorcia
Commonwealth Scientific Corporation, 500 Pendleton Street, Alexandria, Virginia 22314

The effect of Ta and Ta/Cu seed layers, and Ta and Cu cap layers on the effective magnetic thickness of ultrathin permalloy ($Ni_{81}Fe_{19}$) was investigated for MRAM applications. The films were deposited by Ion Beam Deposition. The magnetic moment of each as-deposited permalloy film was measured using a B-H looper and a SQUID magnetometer. The films were further annealed at either 525 K for 1/2 h or 600 K for 1 h to study the effect of thermally driven interdiffusion on the magnetic moment of the permalloy film. Our theoretical calculations showed that the presence of 12% intermixing at the interface reduces the Ni moments to zero. Experimentally, it was shown that the tantalum rather than the copper interfaces are primarily responsible for the magnetically dead layers. The Ta seed layer interface produces a loss of moment equivalent to a magnetically dead layer of thickness 0.6 ± 0.2 nm. The Ta metal in the cap layer results in a loss of moment equivalent to a dead layer of thickness 1.0 ± 0.2 nm. Upon annealing, thermally driven interdiffusion is concluded to have a strong effect on the Ta(seed)/ $Ni_{81}Fe_{19}$ as-deposited interface, based on the doubling of the magnetically dead layer to 1.2 ± 0.2 nm. The $Ni_{81}Fe_{19}$/Ta(cap) as-deposited interface slightly increases its equivalent magnetically dead layer upon annealing to 1.2 ± 0.2 nm. As-deposited interfaces of Ta(seed)/permalloy and permalloy/Ta(cap) are not chemically equivalent and result in different magnetically dead layers, whereas after annealing to 600 K both interfaces attain comparable intermixing and magnetically dead layers. It was also shown that a half-hour anneal at the lower 525 K annealing temperature, which is closer to the actual processing temperature, results in only slight increase of the magnetically dead layer at both interfaces. © *2000 American Institute of Physics.* [S0021-8979(00)73808-4]

In this work, the effect of tantalum seed and cap layers on the magnetic thickness of permalloy is investigated in detail for MRAM (Magnetic Random Access Memory) applications. The MRAMs employ a PSV (Pseudo-Spin-Valve) structure which consists of two magnetic layers of unequal thickness separated by a nonmagnetic spacer. PSVs exhibit the well known GMR (Giant Magneto-Resistance) effect: The resistance across the trilayer film depends on the relative orientation of the top and bottom magnetic films. The effective magnetic thickness of the bottom (storage) and top (read) magnetic layers determines the switching fields of the read and write events. Thinner permalloy layers are desirable to lower switching fields and therefore currents needed to address a particular bit for reading or writing. Both the storage and the read layers typically have permalloy ($Ni_{81}Fe_{19}$) in direct contact with tantalum. Therefore, it is of technological interest to study the effects of Ta seed layers and Ta cap layers on the effective magnetic thickness of ultrathin permalloy ($Ni_{81}Fe_{19}$).

Experiments were designed to study the effect of the Ta seed and cap layers separately, as shown in Fig. 1. The films were deposited using a Commonwealth Scientific Ion Beam Deposition system, and the following structures, with permalloy thickness ranging from $d=2$ to 10 nm, were grown:

(i) 5 nm Ta(seed)/d nm $Ni_{81}Fe_{19}$/10 nm Ta(cap) and

(ii) 5 nm Ta/5 nm Cu/d nm $Ni_{81}Fe_{19}$/10 nm Cu to separate the influence of Ta and Cu layers on the permalloy and to exclude any possible effects of the permalloy/Cu and Cu/permalloy interfaces, and both

(iii) 10 nm Ta(seed)/d nm $Ni_{81}Fe_{19}$/10 nm Cu, to study the influence of the Ta seed layer;

(iv) 10 nm Ta(seed)/5 nm Cu/d nm $Ni_{81}Fe_{19}$/10 nm Ta(cap) to study the influence of the Ta cap layer.

Physical thicknesses were calibrated based on deposition rates. The magnetic moment of each permalloy film was measured using SHB109 B-H looper on a whole 6 in. wafer. The magnetic moments for films in series (iii) and (iv) were also measured on a Quantum Design SQUID magnetometer for 5 mm by 6 mm pieces cut from the same wafers. These films were further annealed at 600 K for 1 h to study the effect of thermally driven interdiffusion on the magnetic moment of the permalloy film.

As shown in Fig. 2, our theoretical studies of the magnetic structure of Ni-rich Ni-Ta alloys using the KKR-CPA method[1,2] show that, in the ideal random alloy limit, the average magnetic moment vanishes at 12% Ta concentration. Experimentally, as shown in Fig. 3(a) and 3(b), by comparing series (i) and (ii), it is evident that the Ta interfaces are

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

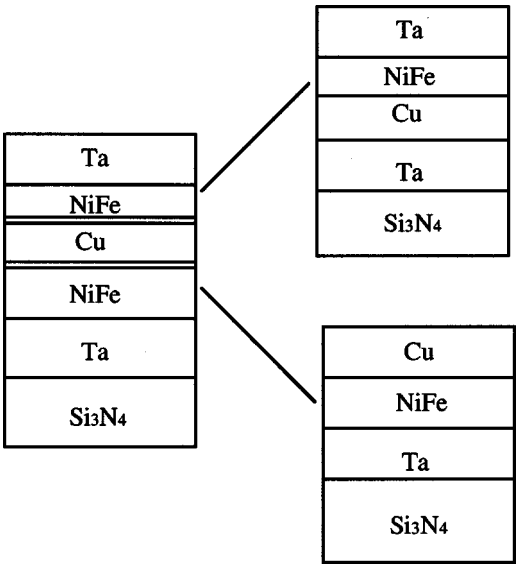


FIG. 1. Design of experiments to investigate separate effects of tantalum seed and cap layers on the magnetic moment of permalloy films used in the PSV. There are two types of Ta/permalloy interfaces: the bottom permalloy layer is deposited on top of a Ta seed layer (denoted as $Ta/Ni_{81}Fe_{19}$), whereas the top permalloy layer is covered by a deposited Ta cap layer (denoted as $Ni_{81}Fe_{19}/Ta$).

primarily responsible for the magnetically dead layers. The amount of missing moment in terms of equivalent magnetically dead layer was obtained from the horizontal axis intercept of the straight line dependence of magnetization as a function of the permalloy thickness. Permalloy layers surrounded by copper films exhibit virtually no loss of magnetic moment within the experimental error of 0.2 nm, whereas permalloy layers surrounded by tantalum films exhibit a 1.5 $\pm$0.2 nm magnetically dead layer. Series (ii) clearly shows that the moment of $Ni_{81}Fe_{19}$ is minimally reduced due to the

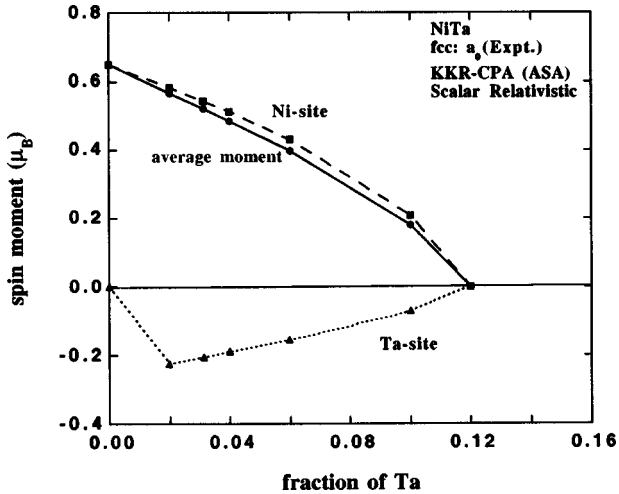


FIG. 2. Spin moment versus tantalum concentration in Ni-Ta alloy. Calculations were done at $l_{max}=3$ using scalar relativistic KKR-CPA and the atomic sphere approximation (ASA) method. Solid line represents alloy's average moment. The dashed and dotted lines show magnetic moment of Ni and Ta sites, respectively. The average moment in the alloy vanishes at 12% Ta concentration.

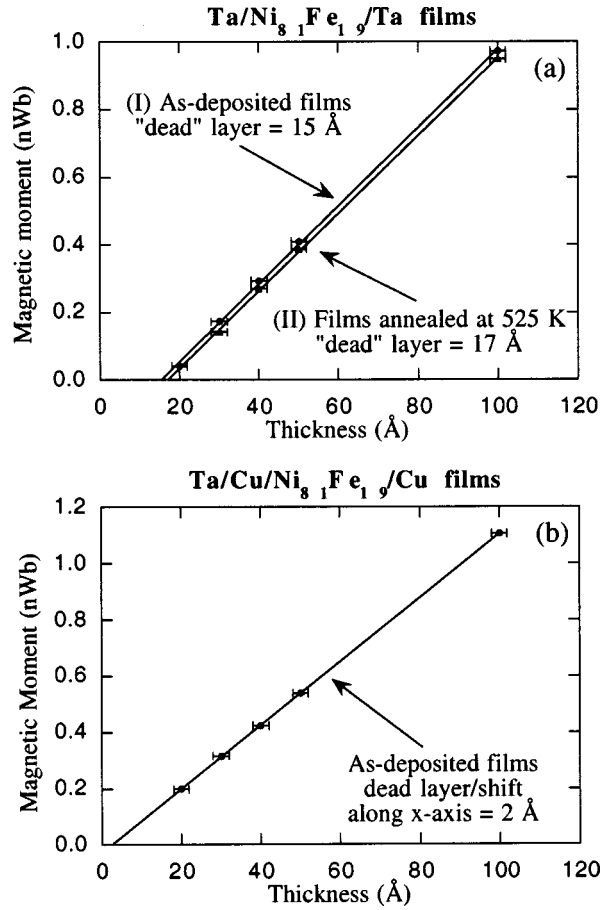


FIG. 3. (a) Total magnetic moment of $Ni_{81}Fe_{19}$ films surrounded by 5 nm Ta seed and 10 nm Ta cap layers (denoted as $Ta/Ni_{81}Fe_{19}/Ta$) as a function of the permalloy thickness. Measurements were performed on a whole 6 in. wafer using a B-H looper for (I) as-deposited films and (II) the same films annealed at 525 K for 1/2 h. 525 K temperature for this set of films was chosen to investigate the annealing temperatures close to actual processing temperatures. The data were fitted with a linear dependence. If there were no losses of moment at the interfaces, the linear dependence of the moment as a function of thickness would start at the origin (0,0) of the graph. From an intercept of the straight line with the thickness axis, the loss of moment can be expressed in terms of an equivalent magnetic dead layer of thickness (I) 1.5 ± 0.2 nm for as-deposited and (II) 1.7 ± 0.2 nm for annealed films. Notice that upon annealing at 525 K the moment decreases by approximately 20 μWb which corresponds to a 0.2 nm increase in terms of a magnetic dead layer. Both Ta interfaces contribute to the magnetic dead layer, see text. (b) The total magnetic moment of $Ni_{81}Fe_{19}$ film surrounded by 5 nm Ta/5 nm Cu seed and 10 nm Cu cap layers (denoted as $Ta/Cu/Ni_{81}Fe_{19}/Cu$) as a function of the permalloy thickness measured on whole 6 in. wafers using a B-H looper. Ta is still used as a seed layer, but the permalloy is surrounded by Cu only. The 0.2 nm shift of the linear dependence measured along the thickness axis is within the measurement error. Error bars in magnetic moment are smaller than symbols.

interfaces with Cu and therefore there is no significant paramagnetic alloy at those interfaces.

Further, as shown in Fig. 4(a), by studying the as-deposited films in series (i), it was found that the Ta seed layer interface with the $Ni_{81}Fe_{19}$ produces a 0.6 ± 0.2 nm thick magnetic dead layer. The Ta metal in the cap series (ii), on the other hand, intermixes more readily in the as-deposited films, producing a 1.0 ± 0.2 nm thick magnetic dead layer, see Fig. 4(b). Since Ta surface free energy (3.15 J/m^2) is higher than that of Ni (2.45 J/m^3),[3] it is not surpris-

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

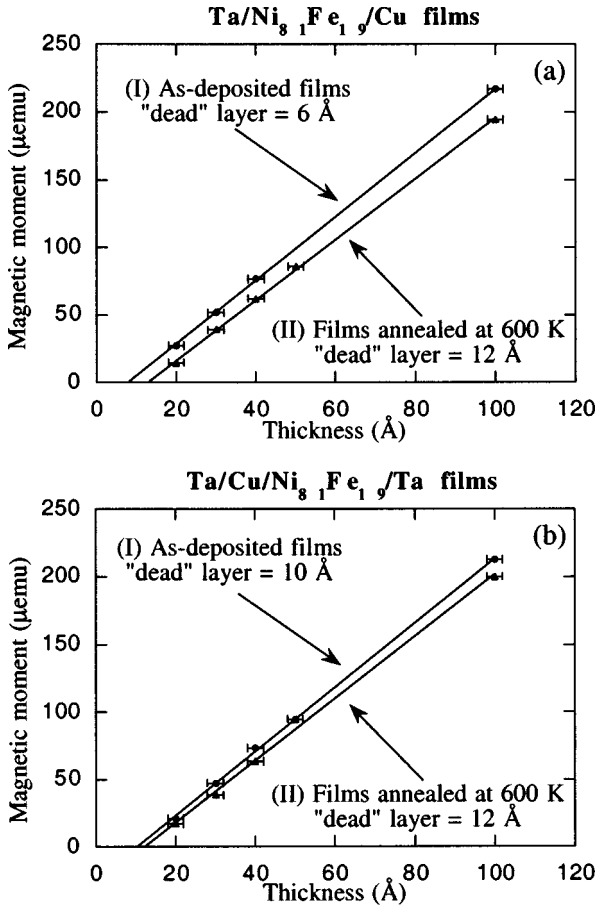


FIG. 4. (a) The total magnetic moment of $Ni_{81}Fe_{19}$ film surrounded by 10 nm Ta seed and 10 nm Cu cap layers (denoted as Ta/$Ni_{81}Fe_{19}$/Cu) as a function of the permalloy thickness measured on a 5 mm by 6 mm piece of a wafer using a SQUID magnetometer for (I) as-deposited films and (II) the same films annealed at 600 K for 1 h. The data were fitted with a linear dependence. The loss of moment can be expressed in terms of an equivalent magnetically dead layer of thickness (I) 0.6 ± 0.2 nm for as-deposited and (II) 1.2 ± 0.2 nm for annealed films. Upon annealing to 600 K the moment decreases by approximately 11 μemu which corresponds to a 0.6 nm increase of a magnetically dead layer. Only the Ta seed layer interface contributes significantly to the magnetically dead layer. (b) The total magnetic moment of $Ni_{81}Fe_{19}$ film surrounded by 10 nm Ta/5 nm Cu seed and 10 nm Ta cap layers (denoted as Ta/Cu/$Ni_{81}Fe_{19}$/Ta) as a function of the permalloy thickness measured on a 5 mm by 6 mm piece of a wafer using a SQUID magnetometer for (I) as-deposited films and (II) the same films annealed at 600 K for 1 h. The data were fitted with a linear dependence. The loss of moment can be expressed in terms of an equivalent magnetically dead layer of (I) 1.0 ± 0.2 nm for as-deposited and (II) 1.2 ± 0.2 nm for annealed films. Notice that upon annealing to 600 K the moment decreases by approximately 3.3 μemu which corresponds to 0.2 nm increase in terms of a magnetic dead layer. Only the Ta cap layer interface contributes significantly to the magnetic dead layer. Error bars in magnetic moment are smaller than symbols.

ing that the Ta atoms in the cap layer interdiffuse into the permalloy during deposition more easily and create larger magnetically dead layer than Ta atoms which are in the seed layer. Ni atoms deposited on the seed layer do not intermix as readily with Ta atoms of higher surface free energy. Therefore the two interfaces, Ta/permalloy, with 0.6 nm dead layer, and permalloy/Ta, with 1.0 nm dead layer, are not equivalent in as-deposited films. Note that, as expected, the 1.6 ± 0.2 nm total sum of thickness from series (iii) and (iv) agrees well within experimental error with the magnetic dead layer thickness for Ta/permalloy/Ta structures of series (i) of 1.5 ± 0.2 nm for as-deposited films.

Our results for as-deposited films agree very well with the previously published studies of magnetically dead layers at the permalloy interfaces with Cu (0.4 ± 0.016 nm for both interfaces in Cu/NiFe/Cu structure[4]), Ta (1.28 ± 0.2 nm for both interfaces in Ta/NiFe/Ta and 0.78 ± 0.2 nm for Ta/NiFe/Cu structures[4]) in sputtered films. However, the authors in that paper were not able to differentiate between the effects of bottom and top permalloy interfaces on the magnetically dead layers.

As shown in Fig. 4(a), upon annealing to 600 K, thermally driven interdiffusion has a much stronger effect on the originally chemically smooth Ta(seed)/ $Ni_{81}Fe_{19}$ as-deposited interface than on chemically rough $Ni_{81}Fe_{19}$/Ta(cap) as-deposited interface, and produces a doubling of the magnetically dead layer to 1.2 ± 0.2 nm. As shown in Fig. 4(b), the chemically rougher $Ni_{81}Fe_{19}$/Ta(cap) as-deposited interface only slightly increases its magnetic dead layer thickness upon annealing to 1.2 ± 0.2 nm. Therefore, the chemically and magnetically different as-deposited interfaces of Ta(seed)/permalloy and permalloy/Ta(cap) reach comparable intermixing and magnetic dead layers upon annealing to 600 K.

Furthermore, series (i) was annealed for 1/2 h at 525 K to study the effect of a more realistic processing temperature on the Ta/permalloy and permalloy/Ta interfaces. As shown in Fig. 3(a), it was found that intermixing increases slightly from 1.5 to 1.7 ± 0.2 nm for both the seed and the cap Ta interfaces. Therefore, it is concluded that up to 525 K annealing temperature the two interfaces still maintain their difference in the level of chemical intermixing.

Work supported through a CRADA with Honeywell SSEC and Nonvolatile Electronics, Inc. sponsored by the Laboratory Technology Research Program, Office of Science, USDOE, under Contract No. DE-AC05-96-OR22464 with Lockheed Martin Energy Research Corp.

[1] H. Winter and G. M. Stocks, Phys. Rev. B **26**, 882 (1983).
[2] B. L. Gyorffy, D. D. Johnson, F. J. Pinski, D. M. Nicholson, and G. M. Stocks, *Alloy Phase Stability*, NATO ASI Series E: Applied Sciences, Vol. 163, edited by G. M. Stocks and A. Gonis (Dordrecht, Kluwer, 1989).
[3] F. R. de Boer, R. Boom, W. C. M. Mattens, A. R. Miedema, and A. K. Niessen, *Cohesion in Metals* (North-Holland, Amsterdam, 1988).
[4] V. S. Speriosu, J. P. Nozieres, B. A. Gourney, B. Dieny, T. C. Huang, and H. Lefakis, Phys. Rev. B **47**, 11579 (1993).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

J. Phys.: Condens. Matter **11** (1999) 5505–5515. Printed in the UK PII: S0953-8984(99)98608-3

Application of the screened Korringa–Kohn–Rostoker method

Nassrin Y Moghadam†, G M Stocks†, B Ujfalussy†, W A Shelton†, A Gonis‡ and J S Faulkner§

† Oak Ridge National Laboratory, Oak Ridge, TN 37831-6114, USA
‡ Lawrence Livermore Laboratory, Livermore, CA 94550, USA
§ Florida Atlantic University, Boca Raton, FL 33431, USA

Received 21 October 1998, in final form 10 May 1999

Abstract. In the familiar Korringa–Kohn–Rostoker (KKR) or Green function method, wave propagation between scattering sites is described by the so-called structure constants of the lattice (KKR structure constants), where these quantities are treated as having infinite spatial extent. In a recent development, it has been shown that the KKR method can be formulated in terms of screened structure constants that are of finite range (the screened KKR method). Here, we present an alternative formulation of the screened KKR method that is derived from simple manipulations of the multiple-scattering equations. We carry out density-of-states and total-energy calculations for spin-polarized and non-spin-polarized materials using the method. We point out possible unphysical features of the method.

1. Introduction

The Korringa–Kohn–Rostoker (KKR) or, as it is alternatively referred to, the Green function method has been used extensively in the study of materials properties connected to the electronic structure. Among its most notable successes, the calculation of the electronic structure of ordered elemental solids (the standard KKR method) [1, 2], substitutionally disordered alloys, the KKR coherent potential approximation (KKR-CPA) [3, 4], and impurities [5] can be mentioned. Based on the multiple-scattering theory (MST) [6], the KKR method calculates the system Green function which leads directly to the calculation of observable quantities such as the electron density and the ground-state energy of the system. Application of MST relies heavily on the construction of the structure constants which describe free-particle propagation between scattering sites. Even though these structure constants can be calculated straightforwardly using Ewald's method [7], they remain computationally cumbersome.

Therefore, it is sensible to search for the development of MST within a framework that would keep intact its advantages while also reducing the difficulties associated with the infinite extent of the structure constants of the lattice. Such a framework has been proposed recently based on ideas originally put forth by Braspenning and Lodder [8, 9] for developing multiple-scattering theory in the presence of a reference medium rather than free space (see [10, 11] and references therein). The formalism leads to the construction of screened structure constants whose extent often does not reach beyond a few nearest neighbours (nn) in a lattice. This formalism has been successfully applied not only to bulk materials but to surfaces as well [10, 12].

In this paper, we present an alternative formulation of the screened KKR method that is more transparent than previous ones. We use the method to calculate self-consistently the

 © 1999 IOP Publishing Ltd

electronic structure and the density of states (DOS) for the ordered paramagnetic materials, fcc copper and bcc molybdenum, and the ferromagnetic bcc iron and fcc nickel. We compare results with those obtained from the standard ($\boldsymbol{k}$-space) Green function method and with calculations from reference [12]. We also point out unphysical behaviour that may arise in such calculations. The remainder of the paper takes the following form. In section 2, we present the derivation of the screened KKR equations. In section 3, we illustrate the formal expressions derived in section 2 by means of numerical examples. Spin-polarized calculations are presented in section 4. A final discussion containing our conclusions is given in section 5.

2. Formulation of the screened KKR method

The formal equations leading to effective, short-ranged structure constants for use within the Green function method can be derived from simple manipulations of the multiple-scattering equations.

The site-diagonal elements of the Green function can be written as [13]

$$G(\boldsymbol{r},\boldsymbol{r}') = \sum_{L,L'} Z^n_L(\boldsymbol{r})\tau^{nn}_{LL'} Z^n_{L'}(\boldsymbol{r}') - \sum Z^n_{L'}(\boldsymbol{r}) J^n_{L'}(\boldsymbol{r}') \tag{1}$$

where the functions $Z^n_L(\boldsymbol{r})$ and $J^n_L(\boldsymbol{r})$ are solutions of the Schrödinger equation in the nth Wigner–Seitz cell. The points $\boldsymbol{r}$ and $\boldsymbol{r}'$ are in that cell. The superscripts on the elements of the scattering path operator [14], $\tau^{ij}_{LL'}$, refer to the cells centred at the lattice sites $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$, and the subscripts are the angular momentum indices. The $\tau^{ij}_{LL'}$ depend on the energy, and are obtained from the inverse of the matrix

$$\mathbf{M} = \mathbf{m} - \mathbf{g}. \tag{2}$$

The non-zero elements of $\mathbf{m}$ are blocks $\mathbf{m}^i$ corresponding to $i = j$. The matrix $\mathbf{m}^i$ is the inverse of the scattering t-matrix for the atom on site i, and, for the special case of muffin-tin potentials, is given by

$$m^i_{LL'} = (t^i_{LL'})^{-1} = (-\kappa \cot \eta^i_l + \mathrm{i}\kappa)\delta_{LL'} \tag{3}$$

where $\kappa = \sqrt{\epsilon}$ and the η^i_l are the scattering phase shifts. The elements of $\mathbf{g}$ are the propagators for electrons in free space, which are

$$g^{ij}_{LL'} = -4\pi\kappa \mathrm{i}^{l-l'+1} \sum \mathrm{i}^{l''} C^{L''}_{LL'} h^+(\kappa|\boldsymbol{R}_{ij}|) Y_{L''}(\boldsymbol{R}_{ij}) \tag{4}$$

where $\boldsymbol{R}_{ij} = \boldsymbol{R}_j - \boldsymbol{R}_i$, and the $C^{L''}_{LL'}$ are Gaunt factors. Note that the matrix elements $g^{ij}_{LL'}$ are defined to be zero for $i = j$. It is standard to truncate the angular momentum expansions at some l_{max}. If there are N atoms in the crystal, the dimensions of $\mathbf{M}$ are $N(l_{max}+1)^2 \times N(l_{max}+1)^2$. The matrix elements $\tau^{ij}_{LL'}$ are the ij, LL' elements of $\mathbf{M}^{-1}$, i.e., $\tau = \mathbf{M}^{-1}$. Since the number of atoms in the crystal is infinite, $\mathbf{M}$ is an infinite matrix and taking the inverse is not a well-defined operation.

For the special case in which all of the atoms are the same, $\mathbf{m}$ has a set of identical $(l_{max}+1)^2 \times (l_{max}+1)^2$ scattering matrices on the diagonal. Using the matrix $\mathbf{U}$, with elements

$$U^{ij}_{LL'} = (1/\sqrt{N}) \exp(-\mathrm{i}\boldsymbol{R}_i \cdot \boldsymbol{k}_j)\delta_{LL'} \tag{5}$$

$\mathbf{g}$ can be transformed into block-diagonal form:

$$[\mathbf{U}^\dagger \mathbf{g} \mathbf{U}]^{ij}_{LL'} = g_{LL'}(\boldsymbol{k})\delta_{ij} = \sum_{j=1}^{N} \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{R}_{0j}) g^{0j}_{LL'} \tag{6}$$

which defines the $(l_{max}+1)^2 \times (l_{max}+1)^2$ matrices $\mathbf{g}(\boldsymbol{k})$. It is easy to take the inverse of the transformed matrix one block at a time, and the result, after allowing N to approach infinity, is the standard one [13]:

$$\tau^{ij} = \frac{\Omega}{(2\pi)^3} \int \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{ij})[\mathbf{m} - \mathbf{g}(\boldsymbol{k})]^{-1}\,\mathrm{d}\boldsymbol{k} \tag{7}$$

where Ω is the volume of the unit cell.

When the atoms in the crystal are not all the same, the block matrices on the diagonal of $\mathbf{m}$, $\mathbf{m}^i$, are different and finding $\mathbf{M}^{-1}$ is not as straightforward. A method for dealing with this problem is to write

$$\mathbf{m} = \mathbf{m}^s + (\mathbf{m} - \mathbf{m}^s). \tag{8}$$

The matrix $\mathbf{m}^s$ is chosen to have identical blocks on the diagonal, as was the case with the ordered crystal. It follows that

$$\boldsymbol{\tau} = \boldsymbol{\tau}^s - \boldsymbol{\tau}^s(\mathbf{m} - \mathbf{m}^s)\boldsymbol{\tau} = [1 + \boldsymbol{\tau}^s(\mathbf{m} - \mathbf{m}^s)]^{-1}\boldsymbol{\tau}^s \tag{9}$$

with

$$\boldsymbol{\tau}^s = [\mathbf{M}^s]^{-1} = (\mathbf{m}^s - \mathbf{g})^{-1}. \tag{10}$$

One way to calculate $\boldsymbol{\tau}^{s,ij}$ is to use (7) with $\mathbf{m}^s$ replacing $\mathbf{m}$.

This method for calculating the scattering path operator has been used frequently in multiple-scattering theory. For one impurity atom embedded in a perfect lattice, the blocks $\mathbf{m}^s$ describe the scattering from the host atoms, $\mathbf{m} - \mathbf{m}^s$ has only one non-zero block, and (9) leads to a simple formula for $\boldsymbol{\tau}^{ij}$. For a cluster of n impurities, n of the blocks in $\mathbf{m} - \mathbf{m}^s$ are non-zero, and it is only slightly more complicated to calculate $\boldsymbol{\tau}^{ij}$ [15].

For an arbitrary collection of atoms in the crystal, it is useful to define the scattering matrices $\mathbf{m}^s$ such that the scattering path operators $\boldsymbol{\tau}^{s,ij}$ are essentially zero when $|\boldsymbol{R}_j - \boldsymbol{R}_i| \geqslant R_{max}$. For instance, calculating the $\boldsymbol{\tau}^{nn}$ needed for the evaluation of the Green's function in (1), truncated matrices $\boldsymbol{\tau}^s$, $\mathbf{m}$, and $\mathbf{m}^s$ are obtained from the ones defined above by setting the blocks corresponding to sites for which $|\boldsymbol{R}_j - \boldsymbol{R}_i| \geqslant R_{max}$ equal to zero. Then $\boldsymbol{\tau}^{nn}$ is the nn block of the finite matrix

$$\boldsymbol{\tau} = [1 + \boldsymbol{\tau}^s(\mathbf{m} - \mathbf{m}^s)]^{-1}\tau^s. \tag{11}$$

Note that the advantage in transforming the scattering path operator matrix instead of the structure constants, as stated in equation (9) in reference [11], is that the present derivation leads directly to the scattering path operator of the real system from which the physical properties are calculated as will be shown in the next section. Furthermore, it is easy to fit τ^s as a function of energy as opposed to the screened structure constants in reference [11].

The experience gained in the research on the screened structure constants described in references [10–12] suggests a convenient scattering potential to use in the calculation of the $\mathbf{m}^s$. The potential is a positive constant v^s within the muffin-tin sphere of each cell in the crystal, and zero in the interstitial region. The phase shifts can be calculated quite easily for this potential, and $\mathbf{m}^s$ is obtained from (3). The matrix $\boldsymbol{\tau}^s$ is calculated from a truncated version of the matrix defined in (10), rather than using the $\boldsymbol{k}$-space formula in (7). A complete real-space calculation of the total energy can be carried out using the $\boldsymbol{\tau}^{nn}$ from (11). The ramifications of this approach will be considered elsewhere. In the present paper, we calculate the scattering path operator $\boldsymbol{\tau}^s$ in real space and then use it to obtain the total energy and DOS from $\boldsymbol{k}$-space integration. The details of these calculations are given in the next section.

3. Calculations

We perform total-energy and DOS calculations for real materials employing the screened KKR method. We compare results of these calculations with those that are obtained using the usual KKR method in which the KKR structure constants are obtained using the Ewald procedure [7]. For the screened KKR method we calculate the screened scattering path matrix ($\boldsymbol{\tau}^s$) in real space, whilst calculation of the scattering path matrix for the system is carried out in $\boldsymbol{k}$-space.

The screening medium is constructed by placing the same constant repulsive muffin-tin potential of height v^s on all of the sites of the underlying lattice. An N-site cluster, comprised of the central site (denoted by 0) and its $N-1$ neighbouring sites, is used to approximate scattering processes within this medium. The calculation of $\boldsymbol{\tau}^s$ is performed by inverting the matrix $\mathbf{M}^s$ for the cluster using (10), where $\mathbf{m}^s$ is the inverse of the t-matrix corresponding to the screening potential v^s.

The $\boldsymbol{\tau}^{00}$ block of the scattering path matrix for the system is obtained using

$$\boldsymbol{\tau}^{00}(\epsilon) = (1/\Omega_{BZ}) \int_{\Omega_{BZ}} [1 + \boldsymbol{\tau}^s(\boldsymbol{k})(\mathbf{m} - \mathbf{m}^s)]^{-1} \boldsymbol{\tau}^s(\boldsymbol{k})\, \mathrm{d}\boldsymbol{k} \tag{12}$$

and the Green function is obtained using (1).

In (12), $\boldsymbol{\tau}^s(\boldsymbol{k})$ is the lattice Fourier transform of $\boldsymbol{\tau}^{s,ij}$ and is given by

$$\boldsymbol{\tau}^s(\boldsymbol{k}) = \sum_{j=0}^{N} \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{R}_{0j}) \boldsymbol{\tau}^{s,0j}. \tag{13}$$

It should be noted that, unlike the corresponding equations for the free-particle propagator (6), this lattice Fourier transform contains a contribution from the origin since $\boldsymbol{\tau}^{s,00} \neq 0$.

Because we have obtained the elements $\boldsymbol{\tau}^{s,ij}$ using a finite cluster in real space, we must choose the ones to use in the lattice Fourier transform (13). The point is that the use of a finite cluster breaks the translational invariance that exists between elements $\tau^{s,ij}$ and $\tau^{s,mn}$ when the pair of sites i, j and m, n have the same positional relationship to one another. We use the matrix elements $\tau^{s,0j}$ ($j = 1, N$) connecting the central site to its neighbours on the basis that they are better approximation to those of the infinite array of v^ss than are general matrix elements $\tau^{s,ij}$ ($i \neq 0$) connecting sites on the periphery of the cluster.

Once the Green function is determined, the charge density, $\rho(\boldsymbol{r})$, and the DOS, $n(\boldsymbol{r})$, are calculated using

$$\rho(\boldsymbol{r}) = -(1/\pi)\, \mathrm{Im} \int_{-\infty}^{\epsilon_F} G(\boldsymbol{r}, \boldsymbol{r}, \epsilon)\, \mathrm{d}\epsilon \tag{14}$$

and

$$n(\epsilon) = -(1/\pi)\, \mathrm{Im} \int_{\Omega_{ws}} G(\boldsymbol{r}, \boldsymbol{r}, \epsilon)\, \mathrm{d}\boldsymbol{r} \tag{15}$$

respectively. In the above, ϵ_F is the Fermi energy.

It should be apparent that the parameters controlling the convergence of the screened KKR method are the usual truncation of the angular momentum at l_{max} used to calculate the Green function from (1), the number of sites in the cluster N, the angular momentum cut-off, l_{scr} ($l_{max} \leqslant l_{scr}$), retained in the cluster used to calculate $\tau^{s,ij}$, and the height of the screening potential, v^s.

In carrying out calculations, we make extensive use of the complex-energy plane. When performing total-energy calculations, integrations over energy are carried out in the complex-energy plane using a semi-circular contour. In calculations of the density of states, used for display purposes, the energy contour is parallel to the real-energy axis with a small imaginary

part (typically $\mathrm{Im}\,\epsilon \sim 0.001$ Ryd). The self-consistent-field (SCF) total-energy calculations use the local density approximation for the exchange–correlation potential and are carried out at zero temperature. We use the muffin-tin approximation for magnetic Fe and Ni, and the atomic sphere approximation for Cu and Mo. This is to test the method for the most commonly used approximations for the atomic potentials. The BZ integration required in (12) is performed using the direction (prism) method [16].

In figure 1 we show the DOS obtained with the screened KKR for fcc Cu with lattice parameter $a = 6.83$ Bohr radii. The dashed curve corresponds to SCF potentials obtained using the screened KKR, and the solid curve to the standard KKR method. Figure 1 also shows the error in the DOS (in Ryd) that is obtained by taking the difference between the DOS calculated at each energy point using the two methods. Figure 2 shows similar results obtained for bcc Mo with lattice constant $a = 5.80$ Bohr radii.

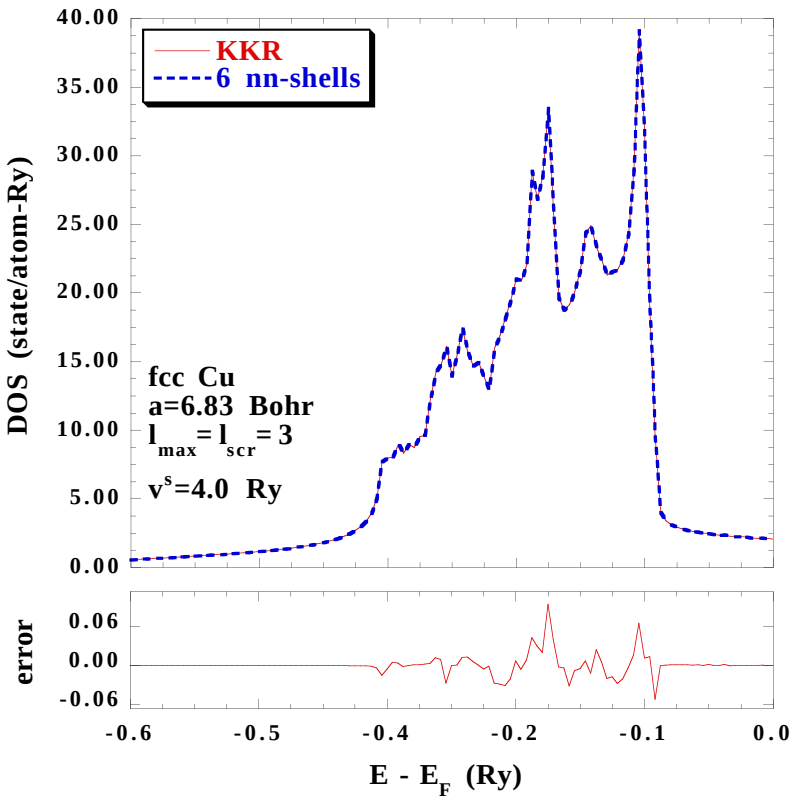


Figure 1. Comparison of KKR (solid curve) and screened KKR (dashed curve) DOS for fcc Cu. The error in the DOS with respect to the standard KKR is also shown.

It can be seen from figures 1 and 2 that for both Cu and Mo the densities of states obtained using the different techniques are essentially identical. The corresponding total energies differ by 0.03 and 0.006 mRyd respectively. In the calculation of the screened structure constants, we used $l_{scr} = 3$ and $v^s = 4.0$ Ryd for both structures. For the fcc structure we used $N = 87$ (six nn shells), while for bcc we used $N = 89$ (seven nn shells). In both cases, $l_{max} = 3$ was used in the KKR calculations. Clearly, for these parameters the screened structure constants are converged.

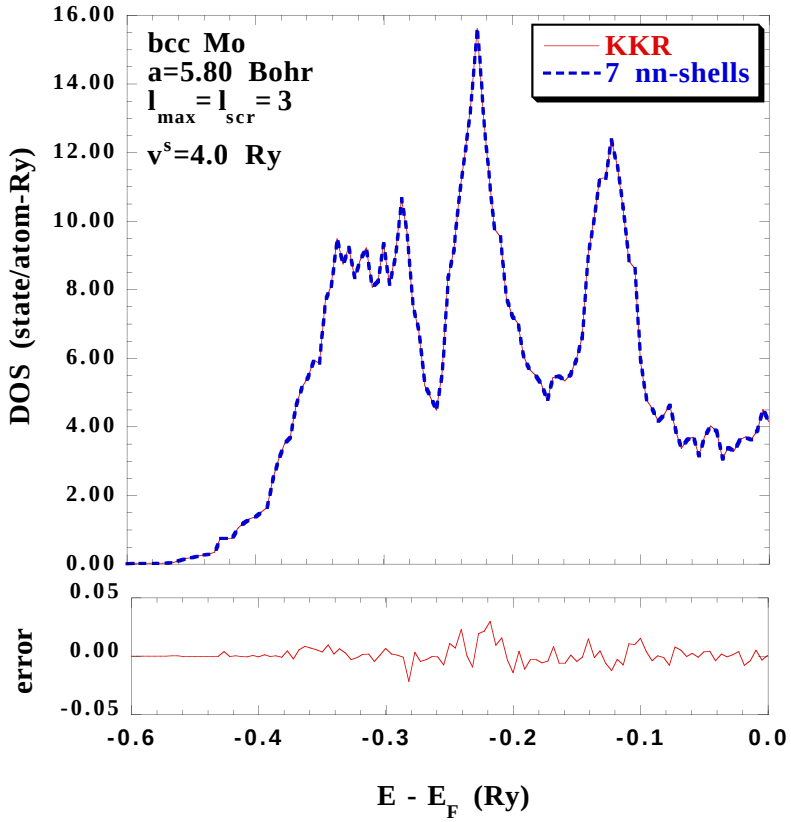


Figure 2. A comparison of the DOS of bcc Mo. The KKR DOS is represented by the solid curve and the screened KKR DOS is shown by the dashed curve. The error in the DOS with respect to the standard KKR is also shown.

In table 1 we show the effect of truncation of the cluster size on the total energy. The screened structure constants were calculated with fixed values of $l_{scr} = 3$ and $v^s = 4.0$ Ryd. Again, $l_{max} = 3$ was used in the KKR calculations. Columns 2 and 3 show the errors in the total energy using the screened KKR method with clusters containing various numbers of

Table 1. Calculated errors in total energy at different numbers of nn shells used in the calculation of the screened structure constants. The table is for fcc Cu and bcc Mo at the respective lattice parameters $a = 6.83$, $a = 5.80$ Bohr radii, $l_{scr} = l_{max} = 3$, and $v^s = 4.0$ Ryd.

	Cu error in energy (mRyd)	Mo error in energy (mRyd)
1 nn shell	−2.95	*
2 nn shells	0.75	*
3 nn shells	0.26	−0.400
4 nn shells	−0.02	*
5 nn shells	0.03	−0.100
6 nn shells	0.03	−0.002
7 nn shells		−0.006

nearest-neighbour shells. The errors are measured with respect to total energies obtained using the standard KKR method. For Cu, an accuracy of ∼0.03 mRyd is obtained for five nn shells. For bcc Mo, an accuracy of 0.002 is attained for six nn shells. An asterisk in the table indicates that no SCF solution was obtained using the above parameters. Adjusting the height of the screened potential to $v^s = 1.0$ Ryd, we were able to obtain a SCF total energy for Mo with one nn shell that differs from the KKR result by 8 mRyd. For two nn shells, adjustment of the screening potential to $v^s = 8.0$ Ryd resulted in a SCF total energy with an error of 0.5 mRyd. Adjustment of the v^s did not yield a SCF total energy in the case of four nn shells. It is not clear why, in some instances, adjustment of the height of v^s led to self-consistency and in other instances it did not. It is not surprising that an approximation to the Green function using screened structure constants that are not fully converged may result in a non-analytic Green function.

There are differences between the calculations in this paper and those reported by Zeller in reference [12]. For example, he uses direct sampling for the BZ integrations, the exchange–correlation potential of Alder and Ceperley, and the full-potential MST. The calculated errors in the total energy for fcc Cu listed in table 1 above show the same trend as those listed in table I of reference [12]. The reason is that the differences in the total energies cancel out. The primary difference between the calculations described here and those in reference [12] is that Zeller calculates the total energy and the DOS at finite electron temperature. This enables him to avoid problems that arise in zero-temperature calculations, which will be discussed.

Next, we show the effect that the height of the screened potential v^s has on the total energy as a function of the number of shells used in the calculation of the screened structure constants. In figure 3, we plot the error in total energy of fcc Cu with lattice parameter $a = 6.83$ versus the number of nearest-neighbour shells. The screened KKR calculations were done at two different screening potentials $v^s = 4$ and 8 Ryd, with $l_{scr} = l_{max} = 3$.

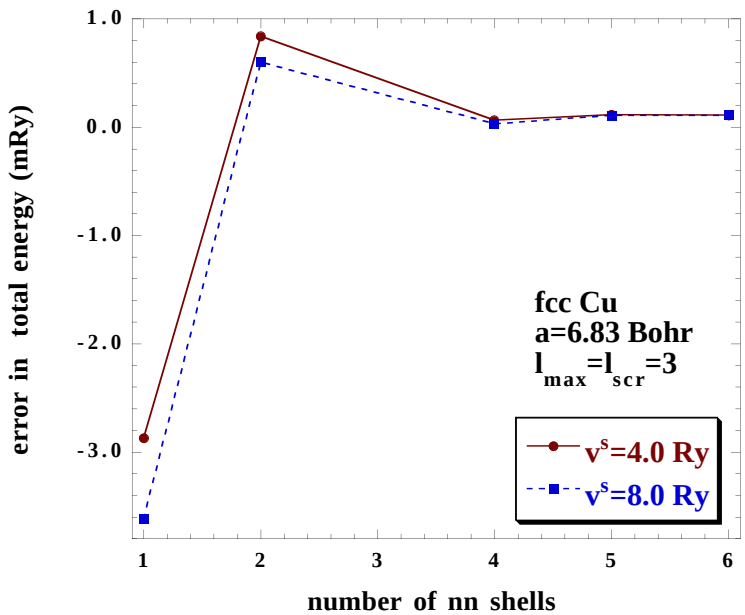


Figure 3. The errors in total energy calculated at two values of v^s are plotted as functions of the number of nearest-neighbour shells used to calculate the screened structure constants. The calculations were done at $l_{scr} = l_{max} = 3$.

From figure 3, it is evident that for four and more nn shells, the strength of the screened potential has essentially no effect on the total energy. Empty-lattice DOS calculations using the screened KKR method are reported in reference [12]. These, however, do not address the problems that arise in DOS calculations for real materials.

In figure 4 we show the calculated DOS of Cu at $l_{scr} = l_{max} = 2$ and $v^s = 2.0$ Ryd where only one nn shell has been used in the calculation of the screened structure constants.

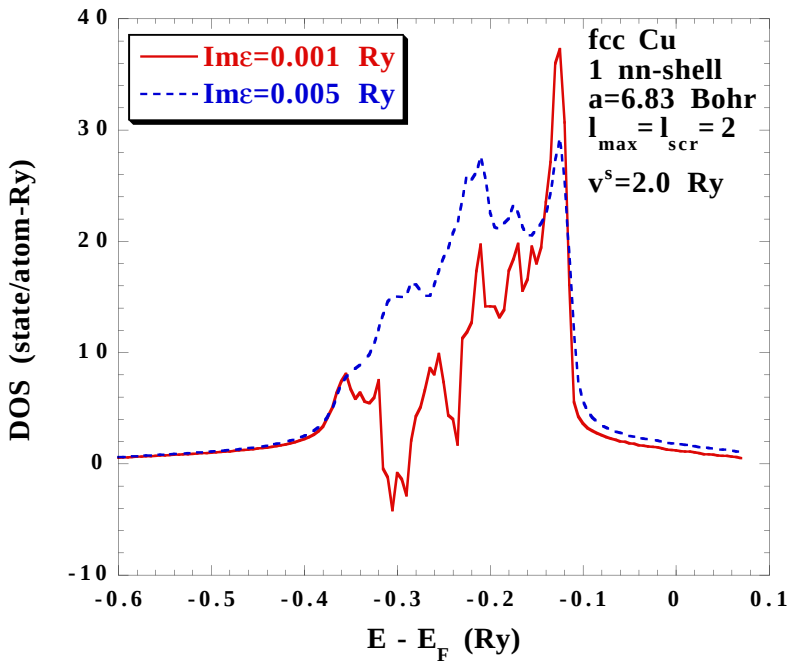


Figure 4. The screened KKR DOS for fcc Cu for one nn shell. The calculations were done at $l_{scr} = l_{max} = 2$, and $v^s = 2.0$ Ryd. The solid curve is the DOS calculated along a contour parallel to the real axis that is 0.001 Ryd off in the complex plane. The dashed curve represents the DOS calculated along a similar contour 0.005 Ryd off in the complex plane.

It can be seen that, for the above parameters, the DOS (solid curve) is negative in the middle of the d bands. This is indicative of a non-analytic Green function. Recall that when we perform SCF calculations we determine ϵ_F and calculate the charge density by integrating over a contour in the complex-energy plane. Thus, any non-analytic behaviour in the Green function will give rise to spurious results. The solid curve represents the DOS calculated along a contour parallel to the real axis that is 0.001 Ryd off in the complex plane ($\mathrm{Im}\,\epsilon = 0.001$ Ryd). The negative DOS are indicative of a spurious pole (or cut) in the Green function that is further off in the complex plane than this. The density of states calculated along a contour for which $\mathrm{Im}\,\epsilon = 0.005$ Ryd (dashed curve) does not exhibit unphysical behaviour. Thus the effect of this non-analytic behaviour will not be seen if calculations are performed at finite electron temperature, $T > 700$ K.

From the above studies it is clear that the screened KKR produces results that are essentially identical to the standard KKR ones if six (seven) nn shells are used for fcc (bcc) structures, but truncation of the inversion to include $\leqslant 4$ nn shells may lead to unreliable DOS, and, in some instances, can lead to non-analytic behaviour.

4. Spin-polarized calculations

It is interesting to apply the screened KKR method to magnetic materials and investigate the validity of the method in such calculations. We present total-energy calculations for bcc Fe and fcc Ni in the ferromagnetic state using the screened KKR method. In the standard KKR calculations used as a reference, the structure constants are obtained from a polynomial fit rather than the Ewald method. In these calculations, we obtain the screened structure constants at $l_{scr} = 3$ and $v^s = 4.0$ Ryd. As before, the KKR calculations were done with $l_{max} = 3$. Figures 5 and 6 show total energies as functions of lattice parameters. The screened KKR calculations for Ni were done with $N = 13$ (one nn shell), and those for Fe with $N = 27$ (three nn shells).

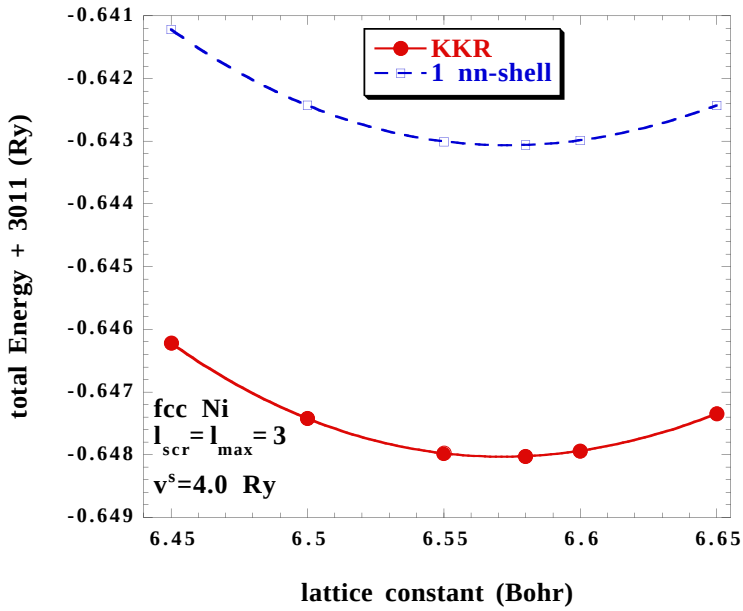


Figure 5. The total energy of fcc Ni as a function of the lattice parameter is shown. The solid curve represents the standard KKR and the dashed curve the screened KKR total energies. The calculations were done at $l_{scr} = l_{max} = 3$, and $v^s = 4.0$ Ryd. Only one nn shell was included in the cluster to calculate the screened structure constants.

The total energies were fitted to a third-degree polynomial to obtain the equilibrium lattice constant a_0, total energy E_0, and bulk modulus B_0. Table 2 shows the results.

Table 2. The equilibrium lattice constant a_0, total energy E_0, and bulk modulus B_0 were obtained for nickel and iron at $l_{scr} = l_{max} = 3$ and $v^s = 4.0$ Ryd. $N = 13$ for nickel and 27 for iron. The corresponding standard KKR values are also listed. The zero of energy for Ni is −3011 and that of Fe is −2522 Ryd.

	fcc Ni		bcc Fe	
	Standard KKR	Screened KKR ($N = 13$)	Standard KKR	Screened KKR ($N = 27$)
a_0 (Bohr radii)	6.572	6.573	5.281	5.280
E_0 (Ryd)	−0.64803	−0.64307	−0.82884	−0.82924
B_0 (Mbar)	2.20	2.29	2.21	2.16

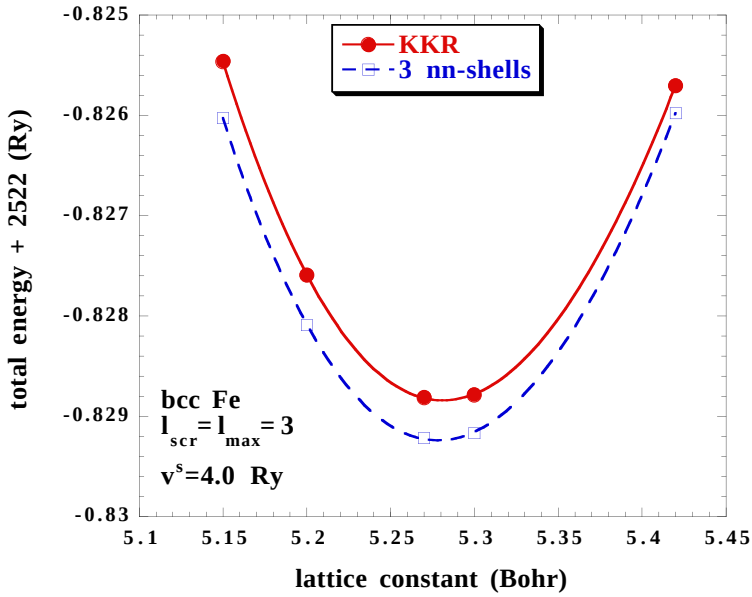


Figure 6. The total energy of bcc Fe is shown for five lattice parameters. The screened KKR calculations were done at $N = 27$, $l_{scr} = l_{max} = 3$, and $v^s = 4.0$ Ryd.

From table 2, the equilibrium lattice constants and bulk moduli agree well with the corresponding KKR values. The equilibrium total energies differ from the standard KKR values by about -5.0 mRyd for Ni and 0.4 mRyd for Fe. Also, magnetic moments calculated for the two systems using the screened KKR method yield values that are different from those obtained from the standard KKR calculations by -0.002 Bohr magneton (μ_B) for Ni and 0.005 μ_B for Fe. Errors in the total energy and magnetic moment reduce considerably when the number of sites N used in the calculation of the screened structure constants is increased. For nickel at $a = 6.65$ Bohr radii, $N = 55$ (four nn shells for fcc structure) and iron at $a = 5.27$ Bohr radii, $N = 65$ (six nn shells for bcc structure), the errors in the total energy are about 0.004 and 0.03 mRyd respectively. The corresponding errors in the magnetic moment are -0.0003 and 0.000 07 μ_B.

5. Conclusions

We have presented an alternative formulation of the screened KKR method that is derived from simple manipulations of the multiple-scattering equations. The derivation focuses on transformation of the scattering τ-matrix rather than the Green function, and leads to an expression for the τ-matrix for the system that is more transparent than previous ones. To illustrate how the formalism works in practice, we performed total-energy and DOS calculations for copper and nickel in fcc phases and molybdenum and iron in bcc phases. It was found that the method yields total energies that are within tens of μRyd of the standard KKR results. The high degree of accuracy is obtained only when the screened structure constants are calculated with $\geqslant 5$ nearest-neighbour shells of repulsive scatterers. The effect of the height of the repulsive potentials on the total energy was shown to be minimal using converged screened structure constants. However, when the screened structure constants are not converged, the total energy is more sensitive to variations in the strength of the screening

potential. We have demonstrated unphysical features (negative DOS) that may arise in the calculations when the maximum angular momentum, the number of neighbouring shells, and the height of repulsive potentials are not chosen properly. The negative DOS, we believe, are mainly due to spurious poles in the Green function. The poles appear at energies close to the real axis ($\mathrm{Im}\,\epsilon \sim 0.001$ Ryd), when the system's Green function is obtained from the screened structure constants that are not converged. We also pointed out that non-analytic behaviour is not seen when calculations are done at finite electron temperature, $T > 700$ K.

The equilibrium total energies calculated for Fe and Ni with $\leqslant 3$ nn shells were shown to be in reasonable agreement with their standard KKR counterparts (a few mRyd). For these systems the magnetic moments differ by only a few thousandths of μ_{B} from the standard KKR results. We pointed out that the total energy and the magnetic moment improve considerably as more sites are used in the calculation of the screened structure constants. The method may be used for tight-binding and near-tight-binding purposes to obtain accurate magnetic moments, and good equilibrium lattice constants and bulk moduli.

Acknowledgments

This work was supported by BES–DMS, and OCTR–MICS, US DOE, under subcontract DEAC05–960R22464 with Lockheed–Martin Energy Research Corporation.

References

[1] Korringa J 1947 *Physica* **13** 392
[2] Kohn W and Rostoker N 1954 *Phys. Rev.* **94** 1111
[3] Winter H and Stocks G M 1983 *Phys. Rev.* B **27** 882
[4] Gyorffy B L, Johnson D D, Pinski F J, Nicholson D M and Stocks G M 1989 *Alloy Phase Stability (NATO ASI Series E: Applied Sciences, vol 163)* ed G M Stocks and A Gonis (Dordrecht: Kluwer)
[5] Zeller R and Dederichs P H 1979 *Phys. Rev. Lett.* **42** 1713
[6] Strutt J W and Rayleigh Baron 1892 *Phil. Mag.* **34** 4816
[7] Ham F S and Segall B 1961 *Phys. Rev.* **124** 1786
[8] Braspenning P J and Lodder A 1994 *Phys. Rev.* B **49** 10 222
[9] Lodder A and Braspenning P J 1994 *Phys. Rev.* B **49** 10 215
[10] Szunyogh L, Újfalussy B, Weinberger P and Kollár J 1994 *Phys. Rev.* B **49** 2721
[11] Zeller R, Dederichs P H, Újfalussy B, Szunyogh L and Weinberger P 1995 *Phys. Rev.* B **52** 8807
[12] Zeller R 1997 *Phys. Rev.* **55** 9400
[13] Faulkner J S and Stocks G M 1980 *Phys. Rev.* B **21** 3222
[14] Gyorffy B L and Stott M J 1973 *Band Structure Spectroscopy of Metals and Alloys* ed D J Fabian and L M Watson (New York: Academic) p 385
[15] Gonis A, Butler W H and Winter H 1984 *Phys. Rev.* B **29** 555
[16] Stocks G M, Temmerman W M and Gyorffy B L 1978 *Electrons in Disordered Metals and at Metallic Surfaces* ed P Phariseau, B L Gyorffy and L Scheire (London: Plenum) p 193

Modelling Simul. Mater. Sci. Eng. **8** (2000) 261–268. Printed in the UK PII: S0965-0393(00)11953-9

Electron and positron states in amorphous $Ni_{0.4}Pd_{0.4}P_{0.2}$

D M C Nicholson†, P A Sterne‡, J C Swihart§, Jeffry Tran‖ and Yang Wang¶

† Oak Ridge National Laboratory, Oak Ridge, TN 37831-6114, USA
‡ Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
§ Physics Department, Indiana University, Bloomington, IN 47405, USA
‖ Physics Department, Luther College, Decora, IA 52101, USA
¶ Pittsburgh Supercomputing Center, Pittsburgh, PA 15213, USA

Received 16 September 1999, accepted for publication 10 February 2000

Abstract. We performed local density calculations of the electronic and positron states for a 300-atom model of bulk amorphous $Ni_{0.4}Pd_{0.4}P_{0.2}$. The procedure for constructing the model and the resulting distribution of bond angles and free volume are described. Comparisons are made to experiment and to models of amorphous $Ni_{0.8}P_{0.2}$.

0. Introduction

Bulk amorphous metals are an interesting class of new materials possessing unique properties that offer exciting possibilities for applications to a broad range of technologies. In contrast to the previous generation of amorphous metals, bulk amorphous metals can be produced in *bulk* form at cooling rates as low as $\sim$1 K s^{-1}. The understanding of their structure, is important to the explanation of their low cooling rate. One of the simplest and most studied bulk amorphous metal is $Ni_{0.4}Pd_{0.4}P_{0.2}$ [1]. We can benefit from the earlier work of Weber and Stillinger [2] who developed interatomic potentials for the conventional metallic glass $Ni_{0.8}P_{0.2}$ that were constructed specifically to reproduce the measured partial pair distribution functions (PDFs) [3]. Structures generated using these potentials have been used as the basis for density functional calculations of the electronic conductivity, density of electronic states, atomic density, and optical reflectivity [4]. The agreement of these calculations with experiment validates the atomic models and the interatomic potentials used in their construction. Recent measurements of the PDF and the Pd distribution function of $Ni_{0.4}Pd_{0.4}P_{0.2}$ [5] provide a very useful guide for extending the potentials of Weber and Stillinger [2] to the ternary alloy $Ni_{0.4}Pd_{0.4}P_{0.2}$. We describe the use of these potentials to generate a 300-atom unit cell model of amorphous $Ni_{0.4}Pd_{0.4}P_{0.2}$. We have confidence in our model because it reproduces the measured partial distribution functions, and electronic and positron states calculated with this model agree with photo-emission and positron lifetime measurements. The calculations of the electron and positron states were performed using the first-principles, order-N, locally self-consistent multiple scattering (LSMS) method [6]. We analysed the structure to determine the distribution of bonds, free volume, electronic states, and positron states. Comparisons were made to models of the conventional amorphous metal, $Ni_{0.8}P_{0.2}$. The most striking contrast is in the amount and distribution of free volume. Because free volume is so closely related to diffusion and the kinetics of glass stability it may be the key to the glass forming ability of $Ni_{0.4}Pd_{0.4}P_{0.2}$.

 © 2000 IOP Publishing Ltd

1. Computational technique: the LSMS method

The electronic structure of systems with a number of atoms, N in the unit cell as large as 300 cannot be obtained by conventional band structure methods. The LSMS method is an $O(N)$ multiple scattering method that is specifically designed for massive parallel processors [6]. It relies on several stationary properties of density functional theory to ensure accurate free energy determination based on an electron density that is determined self-consistently at each site [7]. The LSMS method makes the simplifying assumption that, for the purpose of solving the Schrödinger equation, the effective potential beyond a cluster of neighbouring atoms (the local interaction zone (LIZ)) surrounding each atom can be approximated by a constant. Each atom is then taken to be at the centre of its own LIZ. In the calculations presented here the atomic potentials are taken to be spherical and are obtained self-consistently within the atomic-sphere-muffin-tin approximation [8]. The LIZ is taken to be a sphere of radius 10 Bohr radii.

2. Atomic model

The calculations were performed for a periodically reproduced cubic box containing an amorphous network consisting of 300 atoms for which the atomic positions had first been relaxed to a local energy minimum via two-body interactions [4].

The starting point for generating the amorphous configurations was a previously published [4] structure for amorphous $Ni_{0.8}P_{0.2}$. This configuration was generated by random packing, followed by interchanges to eliminate P–P neighbours, and finally relaxation via the pair potentials of Weber and Stillinger. The $Ni_{0.8}P_{0.2}$ model structure has PDFs that agree well with experiment. We have considerable confidence in this structure because it has been used as the basis for many calculations that agree with experiment. However the algorithm that eliminated P–P nearest neighbours only frees the sample of P–P pairs that are closer than the smallest Ni–Ni distance. To remedy this weakness, without completely abandoning this model that has served us well, we scaled 'r' by a factor of 0.5 in the P–P potential in order to force greater P–P separation. We also made a very small adjustment to the Ni–P potential. This left the model substantially intact, but pushed those few closely-spaced phosphorus pairs to positions clearly identifiable as second nearest neighbours. The Pd-containing structure was then generated by randomly replacing half of the Ni by Pd. We then relaxed the structure using Ni–Pd, Pd–Pd, and Pd–P potentials introduced to augment the already defined Ni–Ni, Ni–P, and P–P potentials. The newly-introduced potentials were adjusted to optimize agreement with the measured partial distribution functions. The potentials are of the form

$$V_{ij}(R) = C_{ij}\, \mathrm{e}^{(\alpha_{ij}R-1.652\,194)^{-1}}[(\alpha_{ij}R)^{-12} - 1)\Theta(1.652\,194 - \alpha_{ij}R) \qquad (1)$$

with the parameters as given in table 1. $\Theta(x)$ is the heaviside function. The PDFs from the model are compared to the measured values [5] in figure 1.

Table 1. Parameters for potentials. Units: α is in units of inverse fcc Ni–Ni distance of 2.49 Å and C is in 10^{-12} ergs.

	Ni–Ni	Ni–Pd	Ni–P	Pd–Pd	Pd–P	P–P
C_{ij}	1.134	1.172	1.701	1.211	1.7011	0.567
α_{ij}	1.000	0.900	1.080	0.850	0.950	0.755

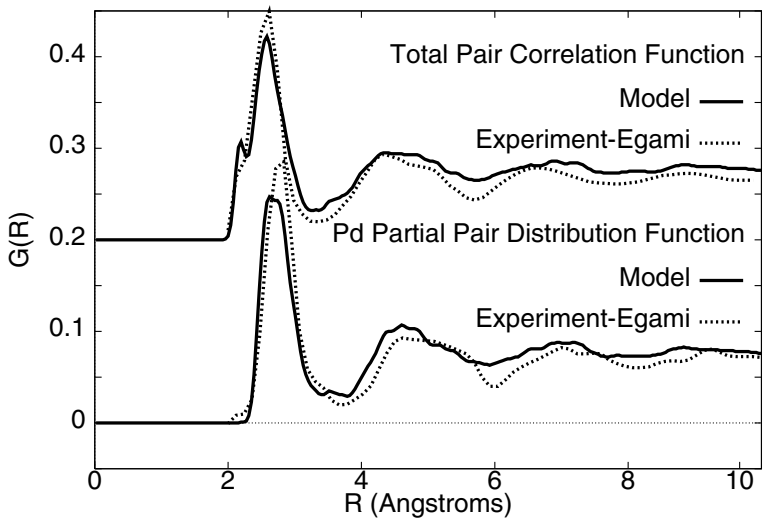


Figure 1. Comparison of model (full curves) amorphous pair distribution functions to measurements by Egami *et al* [5] (broken curves).

3. Stability and diffusion

Like many metallic glasses $(Ni_xPd_{1-x})_{0.8}P_{0.2}$ alloys are associated with a deep eutectic. A reasonable assumption is that at compositions where the liquid state is stable at unusually low temperatures the amorphous phase, which is essentially a frozen liquid, has a free energy competitive with alternative crystalline phases. If long diffusion paths are required to reach energetically-favoured crystalline structures, or if diffusion barriers are high, the conditions for glass formation are enhanced.

In crystals, the diffusion rate is controlled by the density of vacancies and the hopping time. The hopping time depends on the height of the barrier to vacancy hopping. In glasses the definition of a vacancy is not clear cut. Sietsma and Thijsse [9] have made progress toward a workable definition of a vacancy in glass. They characterize the unoccupied regions in the glass system according to their volumes and number of surrounding atoms. They observe that annealing eliminates large, local, free-volume elements that have greater than nine surrounding atoms. These large volumes, which tend to anneal out, are considered to play the role of vacancies, while those that remain are thought to be similar to interstitial sites or constitutional vacancies in crystals. The large and small volumes are referred to as holes and voids, respectively. The distribution of free volume, because it dominates diffusion is likely to play an important role in the kinetic stability of the glass.

We attempt to extend the understanding of the free volume in two ways. First, because our model is based on matching the measured PDFs, there is the hope that it also reproduces characteristics of the distribution of free volume found in the experimental sample. Hence, we study the distribution of free volume in our model. Although the PDFs do not uniquely determine the atomic arrangement, they do constrain it considerably and the distribution of free volume in our model may be indicative of its distribution in a real sample. Second, we study positron wavefunctions. Because positrons are repelled by the nuclei, injected positrons spend most of their time meandering through the interstitial regions and vacancies until they annihilate. If defects that trap the positrons are plentiful, the majority of positrons

will diffuse through the system for only a short time before being trapped. Once trapped they find themselves in a region of low electron density and have an extended lifetime characteristic of the type of defect (i.e. in a crystal typical traps would be monovacancies and divacancies). We take the initial steps necessary for interpreting the information available in the annihilation photons; we calculate the positron density and lifetime.

4. Atomic structure

We can further analyse the structure by tabulating distributions of bond angles. We find that the distribution of angles (figure 2) in $Ni_{0.4}Pd_{0.4}P_{0.2}$ is very similar to the distribution in $Ni_{0.8}P_{0.2}$. They both peak near $\theta = 60^\circ$ and 108° characteristic of icosahedral packing and differ only in small details. Fcc packing would have had peaks at 60°, 90°, and 120°. The population of bond angle can be subdivided according to whether the vertex atom is phosphorus or transition metal and the number of phosphorus atoms among the remaining two atoms that form the angle. The angle populations involving only transition metal atoms at the two compositions are very similar and are very icosahedral in nature. The angles with a P vertex have very similar distributions at the two compositions; the angles are predominately around 70°, consistent with a lower coordination around P. The case of a transition metal vertex and a P and transition metal atom completing the angle shows the greatest change with composition. This change is because Pd and Ni sit at different distances from the phosphorus due to their different sizes. The angles formed by a transition metal atom at the vertex and two phosphorus atoms are distributed in a broad peak (not shown) centred at 100° with no angles below about 80° because there are no P–P nearest neighbours.

Another way of describing the bonds is to look at the atoms surrounding each bond. In a perfectly icosahedral system each bond would have five surrounding atoms that are common

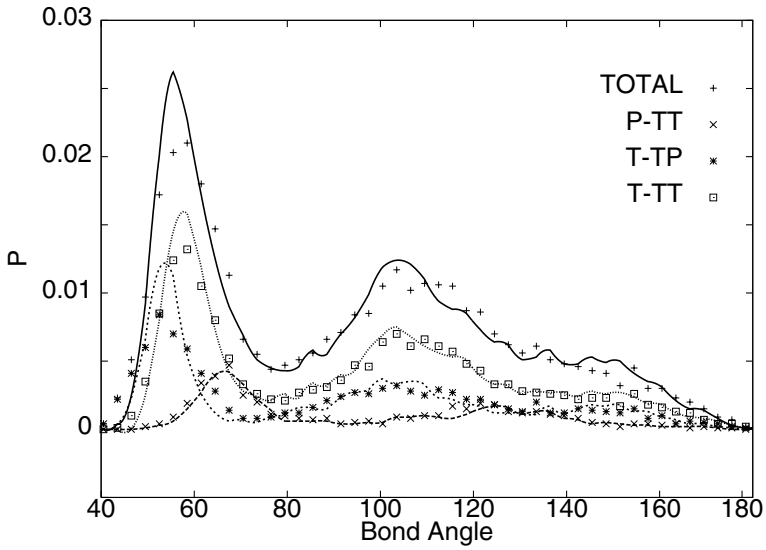


Figure 2. The distribution of bond angles: the curves denote results for $Ni_{0.8}P_{0.2}$; the solid curve is for all bond angles, the long-dash curve is for the P-vertex bonding to two Ni, the short-dash curve is for the Ni-vertex bonding to a P and a Ni, the dotted curve is for all Ni. The symbols apply to $Ni_{0.4}Pd_{0.4}P_{0.2}$ and are labelled in the key, T stands for Ni or Pd. For example, P–TT stands for a phosphorus-vertex bonding with two transition metal atoms.

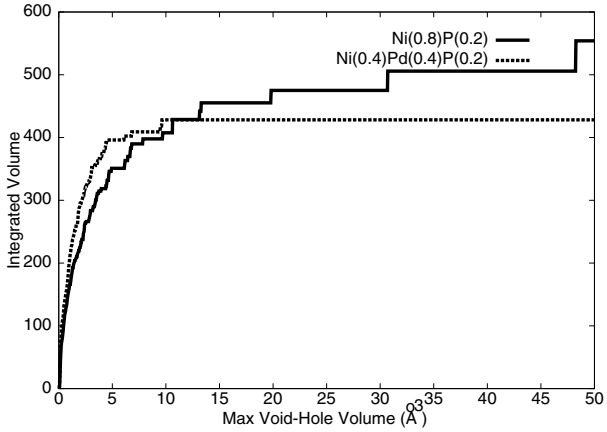


Figure 3. The integrated distribution of free volumes: $Ni_{0.8}P_{0.2}$ is denoted by a full curve and $Ni_{0.4}Pd_{0.4}P_{0.2}$ is denoted by a broken curve.

nearest neighbours to the two atoms that form the bond. Fcc would have four bond neighbours. In our models we find that the most commonly occurring number of bond neighbours is five. About half the bonds have five bond neighbours, with the remaining bonds having numbers of neighbours predominately in the ranging from four to seven. For transition metal bonds about four of the surrounding atoms are transition metal atoms and one is P. The bonds with more than five neighbours usually have an additional P. Again, there is no clear-cut distinction between the two systems. From the point of view of bonds the structure behaves substitutionally, with local dilation to account for the larger size of Pd, but on average preserving the bond angles. This is consistent with the observation made by Egami *et al* [5]: that the PDFs are roughly concentration independent; and the observation of Alamgir *et al* [10]: that the core levels are roughly concentration independent. It is doubtful that these small differences in bond statistics can be held accountable for the much greater glass forming ability of $Ni_{0.4}Pd_{0.4}P_{0.2}$.

Another way to characterize the atomic structure is to look where the atoms are not. We construct Voronoi polyhedra [11] according to the radical plane construction [12] using the radii 2.293, 2.66, and 1.963 au for Ni, Pd and P, respectively. We then shift our attention to the points farthest from the atoms, the vertices of the Voronoi polyhedra. These are the points associated with the free volume [9]. For each vertex we find the largest sphere that does not overlap any atomic sphere. We then group those vertex-centred spheres that overlap each other. The space defined by the overlapping, vertex-centred spheres defines a cell of free volume. The distribution of the volumes of these cells is shown in figure 3. The figure shows the integrated free volume (i.e. the sum of all free volume below a limiting void or hole size). There are two major differences between the two compositions. The Pd-containing composition has less free volume and less of the free volume is associated with large-volume cells. This could be a manifestation of the higher packing fraction possible with a distribution of atomic sizes.

In bulk amorphous alloys, as proposed by Sietsma and Thijsse, diffusion is probably controlled by the large free volumes, holes, that are smaller than vacancies but larger than the interstitial volumes in ordered materials. Reduction of the number of these holes may be the mechanism by which bulk amorphous alloys are prohibited from diffusive transformation to the crystalline ground state even when cooled slowly. It would be valuable to have measured values for the positron lifetimes at these two compositions. The validity of our model free-volume distributions would be supported if the lifetime was longer for $Ni_{0.8}P_{0.2}$, indicating that the more plentiful holes trapped the positrons.

5. Positron states

Positrons are trapped by defects, particularly vacancies. When they annihilate they provide information about the electron density at the defect as well as the atomic species of surrounding atoms. In a simple system, such as well annealed Cu with a few well dispersed monovacancies, two lifetimes are observed. One is a short lifetime associated with itinerant positrons in the bulk, and the other is a longer lifetime associated with positrons trapped at the vacancies. The density of vacancies and the electron density at the vacancy site can be determined. In more complex systems there may be one lifetime, or several lifetimes. For example, amorphous metals typically have one broadened lifetime, the interpretation of which is not clear. It is still debatable whether all annihilation is of the bulk type or from defects. In these more complex systems models are needed to extract useful information from the positron experiments.

The two photons created when a positron annihilates with an electron carry away the momentum and energy of the pair. Because the positrons have (shortly after injection) only thermal energies, a high-momentum photon pair indicates annihilation with a high-momentum core electron, rather than a low-momentum valence electron. The high momentum part of the core annihilation profile as a function of momentum is characteristic of the element and can be obtained by experiments on the pure element and by calculation [13]. In multicomponent systems, the high momentum profile of positrons trapped in vacancies can be compared to profiles from the pure elements to indicate the atomic number of the atom whose core electron participated in the annihilation [13]. The simplest assumption is that the atoms participating in the annihilation are nearest neighbours of the trapping defect.

At this stage we can calculate the positron wavefunctions and determine which nuclei have the greatest overlap with the positron density. We did this by calculating the electrostatic potential of our self-consistent electron density. The local approximation to the electron–positron correlation potential [14] was added to the electrostatic part to provide a one-particle Schrödinger equation for the positrons. We used the LSMS to solve for the positron wavefunctions. The site decomposed density of states shows that the low-energy positrons are predominately on P sites. The product of the electron and positron densities adjusted for correlation and multiplied by the annihilation cross section gives the positron annihilation rate [14]. We have not yet incorporated the matrix elements that couple the positron wavefunctions to the electronic core levels. Therefore, we cannot predict the characteristic patterns that will be seen in high-momentum positron annihilation experiments on $Ni_{0.4}Pd_{0.4}P_{0.2}$ when they are carried out. We have observed that the positrons have their greatest overlap with the phosphorus core so we can anticipate that the high-momentum annihilation will indicate annihilation at phosphorus cores.

We can compare our calculated positron lifetime to some interesting preliminary measurements by Somieski [15] of the positron lifetime in $Ni_{0.4}Pd_{0.4}P_{0.2}$. Their measurements give lifetimes in an as-quenched specimen distributed in the range 140–180 ps [15]. After annealing for 24 h at 200 °C a second group of lifetimes appears in the range 100–120 ps. Our calculated average lifetime for the lowest eight positron states is 114 ps. One possible explanation of these results is that the as-quenched sample has large defects similar to quenched-in thermal vacancies in rapidly-cooled crystals. These defects in the amorphous structure may have long lifetimes in the 140–180 ps range. These defects could have large trapping cross sections that could account for nearly all annihilation events. In the computer-generated model these large vacancies are probably precluded by the model construction procedure. When the experimental sample is annealed, the number of large vacancies is reduced and eventually the smaller-free-volume cells similar to those seen in the model begin to trap positrons with sufficient probability to be observed in the experiments in the range 100–120 ps.

6. Electronic states

The electronic density of states is modified considerably by the addition of Pd (figure 4). Substitution of Pd for Ni greatly increases the overall d-band width of $Ni_{0.4}Pd_{0.4}P_{0.2}$ relative to $Ni_{0.8}P_{0.2}$ [4]. This occurs because of the broader d-band width of elemental Pd and because the centres of the Ni and Pd d-bands are displaced from each other. The Ni states occupy the upper part of the band and the Pd states dominate in the lower part of the band. This separation is seen in the behaviour of the x-ray photo-emission spectrum (XPS) as a function of Pd concentration [10].

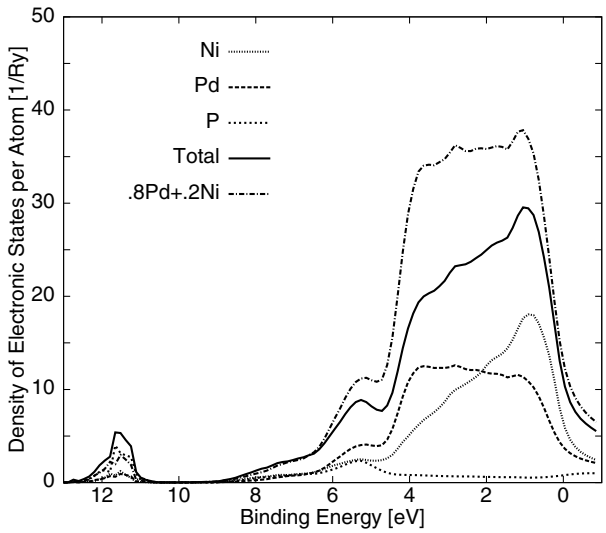


Figure 4. Density of states as function of electron energy for amorphous $Ni_{0.4}Pd_{0.4}P_{0.2}$; the zero of the energy scale is the Fermi energy. The dash–dot curve is the XPS calculated using a rough approximation to the inclusion of matrix elements (arbitrary units).

There is a long tail of states below the d-band that are contributed equally from Ni, Pd, and P states. From our experience with Ni–P we speculate that these tail states are somewhat localized due to the disorder. The fact that all species contribute equally to these tails would tend to make them independent of composition. This can be seen in the lack of composition dependence of the XPS data between 5 and 8 eV.

We find low-lying phosphorus states split-off from the d-bands. These states may be responsible for the barely perceptible bulge in the XPS data at a binding energy of 12.5 eV.

Our calculated density of states at the Fermi energy is higher than that seen in the XPS. Furthermore, we did not see a minimum in the density of states near the Fermi energy that, if present, would have indicated the effects of the stabilizing mechanism suggested by Nagel and Tauc [16]. Our density of states has a dip at 4.5 eV that is not seen in the XPS. A further difference between the density of states and the XPS is that the XPS d-band is almost symmetric about the band centre, while the calculation shows a strong Ni peak at the upper d-band edge. In order to show how inclusion of matrix elements affect the agreement, we plotted the weighted sum of the Ni and Pd densities of states. We weighted the Pd states by a factor of four relative to the Ni states. This is based on the statement of Almagir *et al* [10] that the P matrix element is very small and the matrix element for Pd is four times as large as for

Ni. This rough approximation to a true calculation [17] of the XPS gives excellent agreement with the measured spectrum.

7. Conclusions

We have used the PDFs recently measured by Egami *et al* [5] to develop interatomic potentials suitable for Ni–Pd–P amorphous alloys. We do not make any claim for the usefulness of the energy comparisons based on these potentials. Combined with the procedure of randomly packing hard spheres they should be viewed as part of an interpolation or extrapolation procedure that can use the PDFs of $Ni_{0.8}P_{0.2}$ and $Ni_{0.4}Pd_{0.4}P_{0.2}$ to construct models at other compositions. The geometric properties of the model are discussed along with the resulting electron and positron states. The comparisons of the electronic density of states and positron lifetimes with measurements are promising. These comparisons are not straight forward and there are additional questions to be addressed; still, it is fair to say at this stage that there are no contradictions that discredit the validity of our model structure.

Acknowledgments

Electronic and positron state calculations were performed at the Center for Computational Science at ORNL and the Pittsburgh Supercomputing Center. Work sponsored by ORNL Laboratory Directed Research and Development Fund, BES–DMS, and OCTR–MICS US Department of Energy under subcontract DEAC05-96OR22464 with Lockheed-Martin Energy Research Corporation. This work was also performed under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under contract W-7405-ENG-48.

References

[1] He Y, Schwarz R B and Archuleta J I 1996 *Appl. Phys. Lett.* **69** 1861
Schwarz R B and He Y 1997 *Proc. Int. Symp. on Metastable, Mechanically Alloyed and Nanocrystalline Materials* (Los Alamos, NM: Los Alamos National Laboratory) Publication LA-UR-96-1703
[2] Weber T W and Stillinger F H 1985 *Phys. Rev.* B **31** 1954–63
[3] Lamparter P and Steeb S 1985 *Rapidly Quenched Metals* ed S Steeb and H Warlimont (Amsterdam: Elsevier) pp 459–68
[4] Yang H, Swihart J C, Nicholson D M and Brown R H 1993 *Phys. Rev.* B **47** 107–14
Yang H, Swihart J C, Nicholson D M and Brown R H 1993 *Materials Theory and Modeling (Mater. Res. Soc. Symp. Proc. vol 291)* ed J Broughton, P D Bristowe and J M Newsam, pp 419–24
[5] Egami T, Dmowski W, He Y and Schwarz R B 1998 *Met. Mater. Trans.* A **29** 1805
[6] Wang Y, Stocks G M, Shelton W A, Nicholson D M C, Szotek Z and Temmermann W M 1995 *Phys. Rev. Lett.* **75** 2867–70
[7] Nicholson D M C, Stocks G M, Wang Y, Shelton W A, Szotek Z and Temmerman W M 1994 *Phys. Rev.* B **50** 14 686–9
[8] Zhang X-G and Nicholson D M C 1999 *Phys. Rev.* B **60** 4551
[9] Sietsma J and Thijsse B J 1995 *Phys. Rev.* B **52** 3248
[10] Alamgir F M, Jain H, Miller A C, Williams D B and Schwarz R B 1999 *Phil. Mag.* B **79** 239
[11] Yang Wang and Stocks G M 1994 *Phys. Rev.* B **49** 5028
[12] Watson R E and Bennett L H 1991 *Phys. Rev.* B **43** 11 642
[13] Asokakumar P *et al* 1996 *Phys. Rev. Lett.* **77** 2097
[14] Sterne P A and Kaiser J A 1991 *Phys. Rev.* B **43** 13 892
[15] Bertram Somieski 1998 Oak Ridge National Laboratory, unpublished results
[16] Nagel S R and Tauc J 1975 *Phys. Rev. Lett.* **35** 380
[17] Winter H, Durham P J, Temmerman W M and Stocks G M 1986 *Phys. Rev.* B **33** 2370

Disorder dependence of the magnetic moment of the half-metallic ferromagnet NiMnSb from first principles

D. Orgassa[a)] and H. Fujiwara
Center for Materials for Information Technology (MINT), The University of Alabama, Tuscaloosa, Alabama 35487-0209

T. C. Schulthess and W. H. Butler
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6114

Using half-metallic ferromagnets in spin-dependent devices, like spin valves and ferromagnetic tunnel junctions, is expected to increase the device performance. However, using the half-metallic ferromagnet NiMnSb in such devices led to much less than ideal results. One of the possible sources for this behavior is atomic disorder. First-principles calculations of the influence of atomic disorder on the electronic structure of NiMnSb underline the sensitivity of half-metallic properties in NiMnSb to atomic disorder. In this article, we report on the disorder dependence of the total magnetic moment calculated by applying the layer Korringa–Kohn–Rostoker method in conjunction with the coherent potential approximation. We consider the following types of disorder: (1) intermixing of Ni and Mn, (2) partial occupancy of a normally vacant lattice site by Ni and Mn, and (3) partial occupancy of this site by Mn and Sb. In all cases the composition is kept stoichiometric. All three types of disorder decrease the moment monotonically with increasing disorder levels. For the experimentally seen disorder of 5% Mn and 5% Sb on the normally vacant lattice site, the total moment is decreased by 4.1%. The results suggest that precise measurement of the saturation magnetization of NiMnSb thin films can give information on the disorder. © *2000 American Institute of Physics.* [S0021-8979(00)50008-5]

INTRODUCTION

Spin-dependent devices increasingly gain importance in many applications. Spin valves are being used in hard-disk read heads and in many sensors. Both spin valves and ferromagnetic tunnel junctions are considered as building blocks for nonvolatile memory devices. Spin-dependent devices gain efficiency if the spin polarization in their ferromagnetic thin films is increased. An ideal 100% spin polarization at the Fermi energy is found in half-metallic ferromagnets. The semi-Heusler alloy NiMnSb is such a ferromagnet[1,2] and it has been attempted to use NiMnSb in spin-dependent devices.[3–6] However, the results are not consistent with a high spin polarization. It has been pointed out that atomic disorder in the related half-metallic ferromagnet PtMnSb can significantly influence the half-metallic property of this material.[7] First-principles electronic structure calculations for NiMnSb show indeed that a disorder level of only a few percent is sufficient to create minority-spin states at the Fermi energy and significantly reduce the spin polarization.[8] Such disorder levels are realistic in thin films[9] and thus indeed could be the explanations for the present failure to produce half-metallic NiMnSb films in spin-dependent devices. In this article we will show that a second consequence of atomic disorder is the reduction of the magnetic moment in NiMnSb.

[a)]Electronic mail: dorgassa@mint.ua.edu.

CALCULATION

NiMnSb crystallizes in the $C1_b$ structure. It consists of four interpenetrating fcc lattices labeled A, B, C, and D. Sites A, B, and D are occupied by Ni, Mn, and Sb, respectively, while site C is unoccupied. We apply the layer Korringa–Kohn–Rostoker method in conjunction with the coherent potential approximation (LKKR–CPA)[10] to calculate the total magnetic moment of NiMnSb with atomic disorder. The types of disorder we consider are summarized in Table I. Type A–B has intermixing of Ni and Mn. In the case of disorder type C (C′) the normally empty site C is occupied by Ni and Mn (Mn and Sb). In all cases the composition is kept stoichiometric. X-ray diffraction measurements by Kautzky *et al.*[11] are consistent with disorder type C′ and a disorder level of 5%. The experimental lattice constant[12] of 5.927 Å is used throughout. We use the atomic sphere approximation with an empty sphere for unoccupied sites. In our calculations we use the layer doubling algorithm,[10] repeating (111) layers with four sites in the layer

TABLE I. Types of disorder considered in this article. The second column shows the sites with interchange of atoms (or vacancies). Occupancies of the four lattice sites are shown for the three types of disorder.

Disorder		Occupancies			
Type	Scheme	Site A	Site B	Site C	Site D
A–B	A↔B	$Ni_{1-x}Mn_x$	Ni_xMn_{1-x}	⋯	Sb
C	AB↔C	Ni_{1-x}	Mn_{1-x}	Ni_xMn_x	Sb
C′	BD↔C	Ni	Mn_{1-x}	Mn_xSb_x	Sb_{1-x}

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

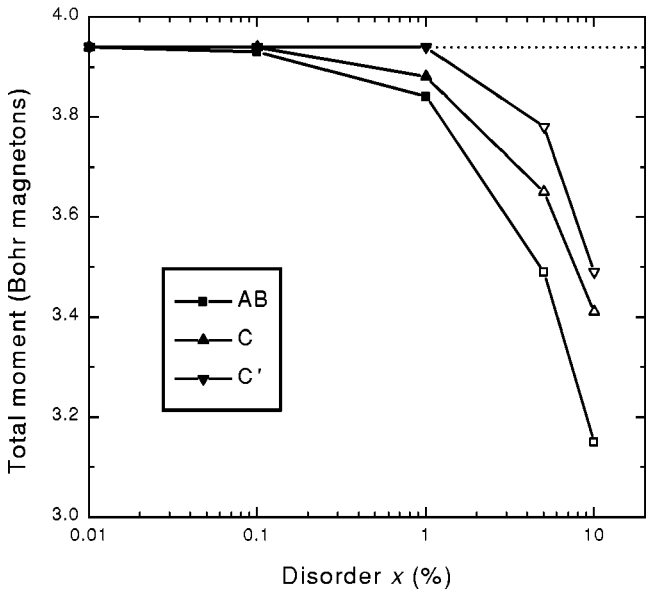


FIG. 1. Calculated total magnetic moment as a function of the disorder level for A–B-type disorder (squares), C-type disorder (up triangles), and C'-type disorder (down triangles). The dotted line shows the moment calculated for ordered NiMnSb. Open symbols indicate a spin polarization below 100%. The lines between the data points are to guide the eyes.

unit cell. All numerical results are based on fully self-consistent calculations performed with a basis set that includes *s*, *p*, *d*, and *f* partial waves on each site and 75 two-dimensional plane waves. The Brillouin zone sums are performed over 64 special $\vec{k}_{\parallel}$-points in the Brillouin zone. For the energy integrals we use 16 points on a semicircle in the complex plane.

RESULTS

We find a total magnetic moment of 3.94 μ_B for ordered NiMnSb. The small deviation from the expected (integer) value of 4.00 μ_B is related to approximations made in determining charge densities and moments in the calculations.[13] For all three types of disorder the total moment decreases monotonically with an increasing level of disorder. This disorder dependence is shown in Fig. 1. Also indicated are cases with minority-spin states at the Fermi energy, i.e., cases with a spin polarization smaller than 100%, as reported in Ref. 8. In these cases, the moment is always smaller than 3.8 μ_B (a reduction by 3.6%), while the half-metallic cases have a moment above this value. For the experimentally observed[11] disorder of type C′ at a disorder level of 5% the calculated moment is 3.78 μ_B or a reduction by 4.1%.

DISCUSSION AND CONCLUSION

The reduction of the moment is caused by disorder-induced minority-spin states which also cause the reduction of the spin polarization for disorder levels above 1%. If we scale our calculated moments by the accepted moment of 4.00 μ_B for ordered NiMnSb we obtain a moment of 3.86 μ_B below which the spin-polarization is smaller than 100%.

Experimental data for bulk NiMnSb is very well in agreement with a moment of 4.00 μ_B or only slightly higher. Hordequin *et al.*[14] measure 4.03 ± 0.02 and 4.025 ± 0.02 μ_B in different directions of single-crystal samples and 4.01 ± 0.02 μ_B in a polycrystalline sample. Kabani *et al.*[3] report a moment of 3.97 ± 0.08 μ_B for single-crystal material. The same group reports a moment of 3.9 ± 0.2 μ_B for thin-film samples obtained by coevaporation of Ni, Mn, and Sb. In this case, the samples were annealed at 500 °C for 2 h. The apparent reduction of the moment should be more pronounced for NiMnSb thin films used in spin valves and tunnel junctions which require lower annealing temperatures and thus higher levels of disorder can be expected.

In order to better understand the effect of disorder on the half-metallic property of NiMnSb and the implication for spin-dependent devices it will be helpful to obtain precise structural and magnetic measurements on thin-film samples. A precise measurement of the magnetization in thin films for which both the atomic disorder and the spin-polarizations are known would in our opinion lead to an independent check of predicted changes of the electronic structure due to atomic disorder and thus help validate the argument that failure to produce half-metallic NiMnSb films is due to atomic disorder.

In conclusion, we reported on the dependence of the total magnetic moment of NiMnSb on atomic disorder. The moment decreases with increasing disorder. A calculated moment below 3.8 μ_B was seen in those cases with minority-spin states at the Fermi energy. While more data on atomic disorder in NiMnSb thin films is needed it will also be interesting to precisely measure the low-temperature saturation magnetization.

ACKNOWLEDGMENTS

This work has been supported by the United States Department of Energy, Office of Basic Energy Sciences through Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research, Inc. and the United States Department of Defense under Grant No. DAAH04-96-1-0316.

[1] R. A. de Groot, F. M. Mueller, P. G. van Engen, and K. H. J. Buschow, Phys. Rev. Lett. **50**, 2024 (1983).
[2] S. J. Youn and B. I. Min, Phys. Rev. B **51**, 10436 (1995).
[3] R. Kabani, M. Terada, A. Roshko, and J. S. Moodera, J. Appl. Phys. **67**, 4898 (1990).
[4] C. T. Tanaka, J. Nowak, and J. S. Moodera, J. Appl. Phys. **81**, 5515 (1997).
[5] C. Hordequin, J. P. Nozières, and J. Pierre, J. Magn. Magn. Mater. **183**, 225 (1998).
[6] J. A. Caballero, Y. D. Park, J. R. Childress, J. Bass, W.-C. Chiang, A. C. Reilly, W. P. Pratt, Jr., and F. Petroff, J. Vac. Sci. Technol. A **16**, 1801 (1998).
[7] H. Ebert and G. Schütz, J. Appl. Phys. **69**, 4627 (1991).
[8] D. Orgassa, H. Fujiwara, T. C. Schulthess, and W. H. Bulter, Phys. Rev. B **60**, 13237 (1999).
[9] M. C. Kautzky, F. B. Mancoff, J.-F. Bobo, P. R. Johnson, R. L. White, and B. M. Clemens, J. Appl. Phys. **81**, 4026 (1997).
[10] J. M. MacLaren, S. Crampin, D. D. Vvedensky, and J. B. Pendry, Phys. Rev. B **40**, 12164 (1989).
[11] M. C. Kautzky, F. B. Mancoff, J.-F. Bobo, P. R. Johnson, R. L. White, and B. M. Clemens (unpublished).
[12] M. J. Otto, R. A. M. van Woerden, P. J. van der Valk, J. Wijngaard, C. F. van Bruggen, C. Haas, and K. H. J. Buschow, J. Phys.: Condens. Matter **1**, 2341 (1989).
[13] In the LKKR–CPA method the charge density is determined from the Green function and the total charge and moments are determined from numerical integration over the atomic sphere. Together with the truncation of the basis sets this may lead, in the present case, to small deviation of the moment from integer values.
[14] C. Hordequin, J. Pierre, and R. Currat, J. Magn. Magn. Mater. **162**, 75 (1996).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

J. Phys.: Condens. Matter **12** (2000) 8439–8454. Printed in the UK PII: S0953-8984(00)14811-8

Screened real-space Korringa–Kohn–Rostoker description of the relativistic and magnetic properties of transition metals

L Petit†‡, S V Beiden†, W M Temmerman§, Z Szotek§, G M Stocks‖ and G A Gehring†

† Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, UK
‡ Institut de Physique et de Chimie des Matériaux de Strasbourg, 28 rue du Loess, 67037 Strasbourg Cédex, France
§ Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK
‖ Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Received 19 June 2000, in final form 4 August 2000

Abstract. The use of the relativistic and spin-polarized real-space Korringa–Kohn–Rostoker (KKR) method is limited to small systems (less than 100 atoms). This is due to the prohibitively large CPU times needed for the inversion of the KKR matrix. To study systems of more than a thousand atoms, we have implemented the concept of a screened reference medium, within the fully relativistic spin-polarized version of the real-space locally self-consistent multiple-scattering method (LSMS). The LSMS method makes use of a local interaction zone (LIZ) for solving the quantum mechanical problem, while the Poisson equation is solved in the whole space. The screened reference medium gives rise to sparse KKR matrices and using state-of-the-art sparse matrix technology a substantial reduction in the CPU times is obtained, enabling applications of the method to systems whose LIZ consists of more than a thousand atoms. The method is benchmarked by application to the elemental transition metals, the fcc (face-centred-cubic) Co and Ni, and the bcc (body-centred cubic) Fe, and compared to the results of the conventional $\boldsymbol{k}$-space methods. The convergence in real space of the magnetic moments, the magnetocrystalline anisotropy energy and the orbital moment anisotropy is discussed in detail.

1. Introduction

First-principles quantum mechanical calculations of the magnetic properties of periodic solids are usually performed by means of $\boldsymbol{k}$-space methods. However, not for all problems of interest can one exploit translational invariance. These include compositional inhomogeneities and strains at surfaces, interfaces and grain boundaries, nanostructures such as nanowires and nanoclusters with potential applications as giant-magnetoresistance (GMR) systems [1, 2]. Very often, the lowering or loss of symmetry determines the technological importance of materials. The magnetocrystalline anisotropy energy (MAE), for example, tends to be much larger at the surface than in the bulk, and its direction in thin films is determined by the deposited material and the layer thickness [3, 4].

In $\boldsymbol{k}$-space, surfaces can be treated as layered structures with the two-dimensional periodicity in the plane, and one real-space dimension in the perpendicular direction. If planar disorder occurs, one uses large two-dimensional unit cells, which are periodically repeated. Inside these cells, the real-space problem has to be addressed. At interfaces and grain boundaries, the periodicity is eventually completely lost and the problem needs to be treated entirely in real space.

 © 2000 IOP Publishing Ltd

In this paper, we present the screened real-space fully relativistic spin-polarized Korringa–Kohn–Rostoker (KKR) method. Here real space is implemented along the lines of the locally self-consistent multiple-scattering (LSMS) formalism [5, 6], where it is assumed, for each site of the solid, that the quantum mechanics is determined by the surrounding atoms, while the Poisson equation is solved in the whole space. Thus, each site is considered as the centre of a local interaction zone (LIZ), and the electron density on this site is calculated by considering only the multiple-scattering processes from the neighbouring sites in the LIZ.

Incorporating relativistic effects in the unscreened version of the LSMS method, Beiden *et al* [7, 8] have explored the convergence, as a function of the LIZ's size, of the spin and orbital magnetic moments and the magnetocrystalline anisotropy energy of elemental transition metals. Fast convergence in real space has been observed for the bcc Fe and the fcc Co. For the fcc Ni, for less than four shells of atoms in the LIZ, no magnetic moment has been obtained. Moreover, no convergence for the magnetic moment and the MAE has been observed within the LIZs of up to 135 atoms (seven shells).

In the unscreened formalism, it is very difficult to deal with the LIZs of more than a hundred atoms, because the inversion of the KKR matrices scales as $O((2 \times (l_{max} + 1)^2 \times M)^3)$ with the number M of atoms in the LIZ. Here l_{max} defines the angular momentum cut-off, and the overall multiplication by 2 is due to the relativistic quantum mechanics. The required CPU times become quickly prohibitive. However, on introducing the screened reference medium [9], a concept originating from Andersen's tight-binding formalism [10, 11], and implemented in the KKR method by Zeller *et al* [12], the structure constant matrices become sparse and, as has been previously shown [13], the matrix inversion times can be substantially reduced by using state-of-the-art sparse matrix technology. Consequently, as will be demonstrated in the present paper, much larger LIZs become accessible to numerical investigation.

The screening technique has so far been applied to situations where real space enters either as zero or one dimension. It is either used as a means to generate $\boldsymbol{k}$-dependent structure constants by a lattice Fourier transform, without the need to use the Ewald summation [14], or in applications to layered systems where a two-dimensional Fourier transform over the layer is performed whilst the remaining dimension is dealt with in real space [16]. Here we present a study with screening without any lattice Fourier transform. Such a study has not been done before and the motivation was to determine the convergence of the method and the usefulness of this approach for calculating the magnetocrystalline anisotropy energy for complex systems. An important aspect of our approach is the use of state-of-the-art sparse matrix technology [17] to push our method to the largest possible system sizes with the present generation of workstations. The physics being targeted by this methodology is the study of the MAE and the orbital moment anisotropy in circumstances of low symmetry which occur at relaxed surfaces and interfaces, steps at surfaces, impurities and surface alloying. These are situations that cannot easily be studied by $\boldsymbol{k}$-space methods.

The remainder of the paper is organized as follows. In the next section, we present the formalism of the screened real-space KKR method, and discuss difficulties associated with the non-periodic reference medium. In section 3, we describe how to construct the screened structure constant matrix, and introduce the approximations enabling us to study systems with significantly more than a thousand atoms. Here, to benchmark the present method against the existing $\boldsymbol{k}$-space methods, all calculations have been performed for the elemental fcc Co and Ni, and the elemental bcc Fe. In section 4, we concentrate on the convergence of the magnetic moments, and discuss the validity and accuracy of the approximations employed. We compare the present results for the spin and orbital moments, and the MAE and the orbital moment anisotropy (OMA), for Fe, Co and Ni, to those obtained with the standard $\boldsymbol{k}$-space methods. Finally, in section 5, we draw our conclusions.

2. Theory

For the sake of clarity and comparison with the $\boldsymbol{k}$-space methods, in this paper we concentrate on bulk materials, with all atoms equivalent. In the LSMS formalism such systems are approximated by a single LIZ consisting of M atoms of the material: LIZ $= \{p\}$ where the sites p are such that $|\boldsymbol{R}_p - \boldsymbol{R}_0| < \boldsymbol{R}_{LIZ}$, i.e. the atoms are contained within a sphere of radius $\boldsymbol{R}_{LIZ}$ around the central site $\boldsymbol{R}_0$. Beyond the $\boldsymbol{R}_{LIZ}$, the boundary of the LIZ, the potential is set to zero, i.e. the muffin-tin zero. The Poisson equation is solved for the infinite real system, whereas for the quantum mechanical problem the LIZ is used. The scattering-path matrix for the cluster of M atoms is obtained as

$$\tilde{\tau}_M(E) = [\tilde{t}_M^{-1}(E) - \tilde{G}_M^0]^{-1}. \tag{1}$$

The LIZ t-matrix, $\tilde{t}_M$, has M non-zero sub-blocks on the diagonal, each corresponding to a specific single-site scattering matrix t_j, where j is one of the M atoms in the local interaction zone. The free-space structure constant matrix, $\tilde{G}^0$, consists of the site- and angular-momentum-decomposed matrix elements of the free-space propagator:

$$\tilde{G}_M^0 = \{\tilde{G}^{0;pn}(E)\} \qquad \tilde{G}^{0;pn}(E) = \{G_{LL'}^{0;pn}(E)\} \tag{2}$$

where p and n run over all the sites in the LIZ.

The electronic and magnetic properties are derived from the multiple-scattering Green's function, $G(\boldsymbol{r}, \boldsymbol{r}'; E)$, of the system. In the vicinity of the central site $\boldsymbol{R}_0$, it can be written as [18, 19]

$$G(\boldsymbol{r}, \boldsymbol{r}'; E) = \sum_{\Lambda,\Lambda'} Z_\Lambda(\boldsymbol{r}_0)\tau_{\Lambda\Lambda'}^{00}(E)Z_{\Lambda'}^{\times}(\boldsymbol{r}_0') - \sum_{\Lambda} Z_\Lambda(\boldsymbol{r}_<)J_\Lambda^{\times}(\boldsymbol{r}_>). \tag{3}$$

Here, $\boldsymbol{r}_0 = \boldsymbol{r} - \boldsymbol{R}_0$ and $\boldsymbol{r}_0' = \boldsymbol{r}' - \boldsymbol{R}_0$, and $\boldsymbol{r}_<$ ($\boldsymbol{r}_>$) means the smaller (bigger) of $(\boldsymbol{r}, \boldsymbol{r}')$. The relativistic Green's function, $G(\boldsymbol{r}, \boldsymbol{r}'; E)$, is a 4×4 matrix, Λ stands for the pair of relativistic quantum numbers (κ, μ), and Z_Λ and J_Λ represent respectively the normalized regular and irregular scattering solutions to the single-site Dirac equation. $\times$ refers to taking the complex conjugate of only the spherical harmonic, and not the radial part of these solutions. The site-diagonal matrix elements of the scattering-path operator, $\tau_{\Lambda,\Lambda'}^{00}$, give the scattered wave at site 0, due to an incident wave at 0, taking into account all possible scattering processes in between. It is obtained as the 00 site-diagonal sub-block of $\tilde{\tau}_M$ of equation (1) and is actually the only part of $\tilde{\tau}_M$ which is needed for the Green's function of equation (3).

In the KKR method, knowledge of the Green's function, $G(\boldsymbol{r}, \boldsymbol{r}'; E)$, allows one to calculate such quantities as the charge density $\rho(\boldsymbol{r})$, spin magnetic moment $\boldsymbol{m}_{spin}$ and orbital magnetic moment $\boldsymbol{m}_{orb}$.

The force theorem [20], which follows from the variational character of the total energy, defines the change in the total energy between two states with similar charge densities to be equal to the change in the sum of the one-electron energies E_{sum}. Consequently, the magnetocrystalline anisotropy energy, being the difference in the total energies corresponding to two different magnetization directions, is defined as

$$E_{MCA} = \sum_i^{occ} E_i^{[\hat{M}_1]} - \sum_i^{occ} E_i^{[\hat{M}_2]} = \int^{E_{F_1}} \mathrm{d}E\, n_{[\hat{M}_1]}(E)E - \int^{E_{F_2}} \mathrm{d}E\, n_{[\hat{M}_2]}(E)E \tag{4}$$

where $n_{[\hat{M}_1]}(E)$ and $n_{[\hat{M}_2]}(E)$ are the single-particle densities of states when the magnetization is in the directions $[\hat{M}_1]$ and $[\hat{M}_2]$, respectively. Here E_{F_1} and E_{F_2} are the Fermi energies corresponding to the single-particle densities of states $n_{[\hat{M}_1]}(E)$ and $n_{[\hat{M}_2]}(E)$.

The major computational effort of the real-space method is associated with the determination of $\tau_{\Lambda,\Lambda'}^{00}$ through the matrix inversion in equation (1). The free-space structure

constants are long ranged, i.e., all the sites in the LIZ are connected to all the other sites in the LIZ, and the matrix inversion times scale as $O(N^3)$ with the matrix dimension N. Following Andersen *et al* [10, 11], Zeller *et al* [12] have proposed a new reference medium where the structure constants decay exponentially in real space. This reference medium consists of spherically symmetric constant repulsive potentials at the non-overlapping muffin-tin sites. In the interstitial space, the potential remains zero. The exponential decay of the corresponding so-called 'screened' structure constants is due to the fact that, in the energy range of the valence bands, and with repulsive potentials in excess of 2 Ryd, there exist no eigensolutions for the new reference system.

The scattering with respect to the new reference medium is described by the scattering matrix

$$\Delta t_j = t_j - t_j^{sc} \tag{5}$$

where t_j and t_j^{sc} are respectively the t-matrices of the muffin-tin potential and the repulsive potential. The structural Green's function for the screened reference medium is obtained from the Dyson equation

$$\tilde{G}_M^{sc} = \tilde{G}_M^0(\tilde{I} - \tilde{t}_M^{sc}\tilde{G}_M^0)^{-1}. \tag{6}$$

Similarly to equation (3), the Green's function for the LIZ can be written in terms of the new scattering-path operator:

$$\tilde{\tau}_{\Delta,M}(E) = [(\Delta\tilde{t})_M^{-1}(E) - \tilde{G}_M^{sc}]^{-1} \tag{7}$$

and the single-site solutions are now defined with respect to the difference scattering potential.

The change of reference medium is a purely mathematical concept, and completely identical Green's functions are obtained when the matrices in equation (6) involve the entire LIZ. However, it has been shown by Zeller [14] that, because of the exponential decay in real space, the screened structure constants connecting sites at $\boldsymbol{R}_n$ and $\boldsymbol{R}_p$, can be ignored if $|\boldsymbol{R}_n - \boldsymbol{R}_p| > R_{sc}$, and that only those with $|\boldsymbol{R}_n - \boldsymbol{R}_p| < R_{sc}$ need be considered. Here R_{sc} denotes a certain distance, called the screening radius, beyond which the screened structure constants are essentially zero. Consequently, the screened structure constants are still calculated through the Dyson equation (6), but by inverting clusters of much smaller size than the LIZ.

In periodic systems, all the sites of the infinite reference medium are equivalent, and the $\boldsymbol{k}$-space screened structure constants, $G_{LL'}^{sc}(\boldsymbol{k}, E)$, can be calculated by lattice Fourier transformation. Because of the exponential decay, the summation converges fast and does not require the Ewald procedure [14]. When applied to layered systems, the screening leads to band diagonal matrices [15, 16]. In real space, the important feature associated with the new reference medium is the sparsity of the screened structure constants, which results from the exponential decay of its matrix elements in real space. In fact, as has been shown in reference [13], by applying a specific sparse matrix solver [17], and depending on the degree of sparsity, the inversion scales essentially as $O(N)$ with matrix dimension N. This means that we can expect considerable gains in the execution times of our code when exploiting the freedom of choice of the reference medium.

There is however a drawback. The structure constants for the free-space reference medium are known analytically, whereas the screened structure constants have to be calculated numerically (equation (6)). As can be seen from figure 1, our reference system for the LIZ is not translationally invariant. The repulsive potentials are placed only on those sites that correspond to atoms in the LIZ. Outside this zone, the reference medium is the free space. Therefore, every site is surrounded by a different scattering neighbourhood, and the screened structure constant matrix for the LIZ cannot be reconstructed by translations of the screened

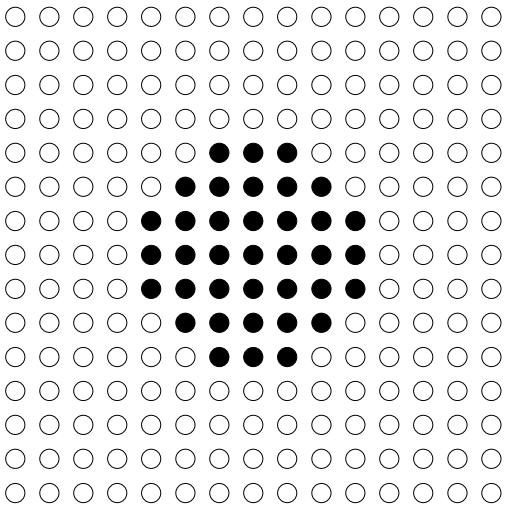


Figure 1. The reference medium for the screened real-space method. In the area covered by the LIZ, constant repulsive potentials at the muffin-tin sites are used (dark spheres). Outside this area the reference medium is free space (white spheres).

structure constants, calculated for one limited small cluster contained within a screening radius. To find the screened structure constant matrix we can either invert the matrix for the entire LIZ, which is clearly not an economical solution, or we can calculate the screened structure constants for each site of the LIZ by inverting a small matrix corresponding to the sites within its screening radius. This means that some of the gain in CPU time that results from dealing with sparse matrices will be lost in the construction of these matrices.

The speeding up in the matrix inversions, with respect to ordinary LU decomposition, allows one to study much larger cluster sizes than was previously possible. In the next section we will discuss how this gain in the CPU time can be maximized, and what approximations can be used to further increase the size of the LIZs.

3. Matrix construction and approximations

The execution time of the sparse matrix inversion is to a large extent determined by the degree of sparsity, i.e., the number of matrix elements that are zero. Therefore, the aim is to make the screened structure constants as short ranged as possible. However, a cutting off of the screened structure constants after only a few shells, for example a screening radius of one or two shells, can only be justified if the height of the repulsive potentials causes the screened structure constants to decay quickly enough that they can be neglected beyond that distance. Moreover, the height of the repulsive potentials has to be chosen according to the energy window of applicability. It has been previously noted [14] that the energy range of applicability of the screened KKR method is larger for higher potentials, and that accurate densities of states (DOS) for energies of up to 3 Ryd can be obtained with a repulsive potential of 8 Ryd. In the present work we have used a screening potential of 4 Ryd to reproduce reliably the DOS at the energies corresponding to the valence bands and the angular momenta up to $l_{max} = 3$.

Concerning the screening radius, we find that three shells of screening are sufficient to reproduce the DOS of bcc Fe. For Co and Ni, which crystallize in the more close packed fcc structure, two shells of screening are sufficient. Compared to the unscreened results, for these

respective screening radii, the total energies are reproduced with the precision of 10^{-3} Ryd, whereas the spin magnetic moments have been calculated with the accuracy of 10^{-1} μ_B. A larger screening radius implies more accurate computed values, meaning that for example the spin magnetic moment of Fe can be determined with the accuracy of 10^{-3} μ_B, when a screening radius of five shells is used. However, in this case, the sparsity of the screened structure constant matrix is considerably reduced, and the matrix inversion time increases accordingly.

In the absence of structural disorder in the cluster, in the central region of the cluster the translational invariance is fulfilled, and only the sites at the surface are surrounded by a different number of sites within the screening radius. Moving to ever-increasing cluster sizes, the relative importance of the surface with respect to the central area decreases. Therefore, one might expect that for large LIZs, it would be a good approximation to assume translationally invariant structure constants on all the sites. We have tested this assumption by calculating the spin magnetic moment on the central site of the Co cluster as a function of the number of shells in the LIZ. As can be seen in figure 2, even with 29 shells (1055 atoms) in the LIZ, the magnetic moment shows no sign of convergence. Since, however, it has been shown in the unscreened calculations [7] that the magnetic moment has been converged for smaller LIZs, we have to conclude that the use of the translationally invariant screened structure constants is not justified. Assuming translationally invariant screened structure constants is equivalent to having a reference medium with the repulsive potentials both inside and outside the LIZ. The physical picture that it corresponds to is that of a muffin-tin potential within the LIZ, embedded in a repulsive potential, giving rise to a quantum well with standing waves inside the chosen cluster [21]. In calculations for layered systems, with only one real-space dimension, this problem, though noticeable, is less serious [16].

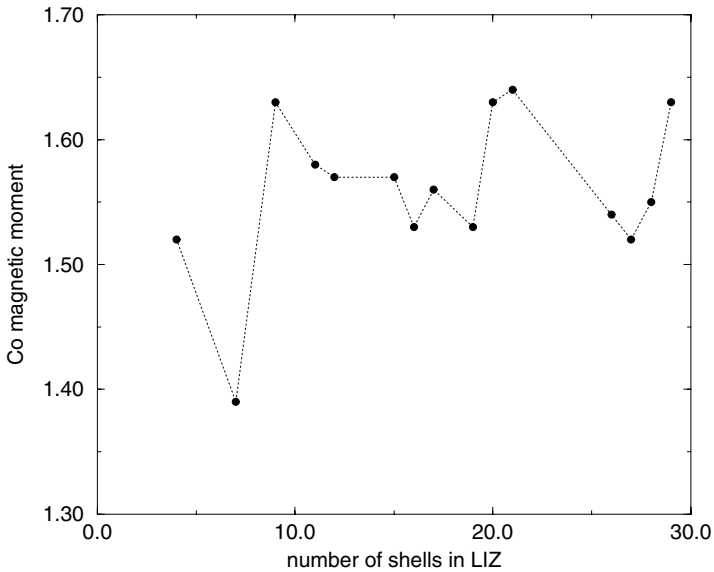


Figure 2. The magnetic moment of Co as a function of the number of shells in the LIZ. The reference medium is assumed to be translationally invariant.

The matrix dimension N is determined by the size of the LIZ, and the angular momentum expansion of the structure constants, i.e., $N = 2(l_{max} + 1)^2 \times M$, where l_{max} is a maximum value of the on-site angular momentum cut-off. It has been suggested earlier [7] that it is

sufficient to use large values for l_{max} only on the first few shells of the cluster, and to reduce l_{max} on more distant shells of the LIZ cluster. This of course reduces the matrix dimension considerably, as we now have

$$N = \sum_{i=1}^{M} 2(l_{max,i} + 1)^2$$

with i the atom index for the M-atom LIZ. In a tight-binding picture of the electronic structure, this approximation can be justified by noting the fact that for example the d electrons are more localized than the s electrons, and therefore the d levels on a remote atom will only have a negligible effect on the electronic properties of the central site. If this approximation were justified, it would be sufficient to use only $l_{max} = 0$ on all the outer sites of the cluster. This would imply further that one could consider extremely large LIZ sizes, and at the same time reduce the time needed for the reconstruction of the screened structure constant matrix. We shall explore this possibility further in the next section.

4. Results and discussion

In comparison with the standard KKR $\boldsymbol{k}$-space calculations, the present method relies on three major approximations: the size of the LIZ, the angular momentum cut-off and the screening radius. In this section we shall discuss our results for the convergence of the magnetic properties on the central site of the LIZ as a function of the number of shells in the LIZ. Furthermore, we shall analyse the dependence of the results on the angular momentum cut-off, i.e., the second approximation. The third approximation, associated with the value of the screening radius, is considered when comparing our results to those of the equivalent $\boldsymbol{k}$-space methods.

4.1. Convergence

In figure 3 we show the convergence of the spin and orbital magnetic moments in Co, as a function of the number of shells in the LIZ. The angular momentum configuration in the LIZ used in these calculations is referred to as $l_{max} = 3^3 2^6 1^n$, meaning that on the central site and the nearest- and second-nearest-neighbour shells we have included angular momenta up to $l_{max} = 3$, then on the atoms belonging to the six subsequent shells we have set $l_{max} = 2$ and for the rest of the sites we have used $l_{max} = 1$. For Fe, shown in figure 4, additionally to the angular configuration used for Co, we have also performed calculations for the $l_{max} = 3^3 2^m$ configuration and the corresponding results are plotted in the relevant panel for comparison. As can be seen in figure 5, the latter configuration has also been studied for Ni, but instead of the $l_{max} = 3^3 2^6 1^n$ angular momentum configuration, we have performed calculations for the $l_{max} = 3^3 2^6 0^n$ configuration. Our test studies have shown that, although the converged value depends on the choice of the screening radius, the convergence pattern is independent of it. As a consequence, the observed convergence behaviour has been obtained by using screening radii of respectively one shell for Co and Ni, and two shells for Fe.

As can be seen in figure 3, in the case of Co for LIZs larger than 17 shells, i.e., 459 atoms, all the calculated properties for the central site are converged. The magnetic moments of Fe are seen to converge at around 17 shells in the LIZ. The converged value is 2.05 μ_B. This value differs by approximately 0.05 μ_B from the result obtained with three shells of screening, i.e., a change of 2.5%. The results obtained with three shells of screening are generally in better agreement with the results obtained with the free-space reference medium.

For Ni the situation is different. It has been noticed in the unscreened calculations that one has to include at least five shells in the LIZ to obtain a magnetic moment at all [7], and that

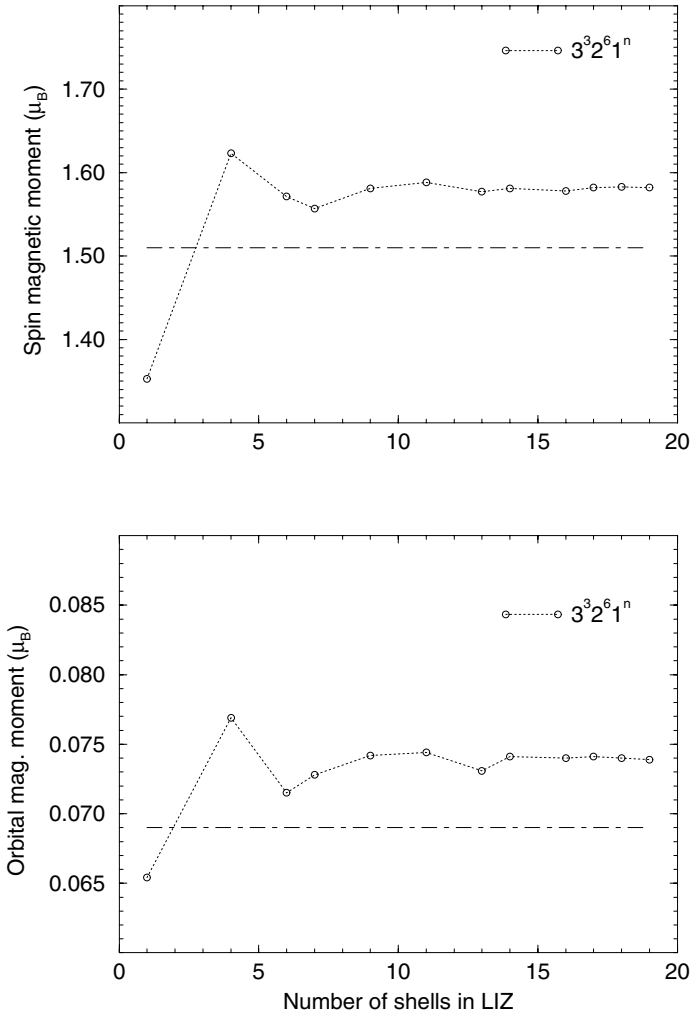


Figure 3. Convergence of the spin and orbital magnetic moments of fcc Co as a function of shells in the LIZ. The angular momentum configuration used is $l_{max} = 3^3 2^6 1^n$ (dotted line). The screening radius consists of one shell. The dashed line corresponds to the KKR $\boldsymbol{k}$-space result [24].

the moments showed no sign of convergence even for seven shells in the LIZ. Larger cluster sizes could not be investigated with the unscreened real-space code, as the CPU times became prohibitive. With the screened version of the method one can explore much larger cluster sizes, and we find that the magnetic moments only start to converge for LIZs containing more than 1200 atoms, i.e., at least 32 shells. For the LIZ of 1253 atoms (33 shells, $l_{max} = 3^3 2^6 0^{25}$), we find the converged value of the spin magnetic moment to be 0.598 μ_B. Note that in our calculations, $l_{max} = 3$ has been implemented only on the central site and the two surrounding shells. Including also the $l = 3$ contributions from further shells changes the moments on the central site only marginally. Equivalently, the inclusion of the $l = 1$ scattering channel from all the sites in the LIZ leads only to minor changes in the moments, which is why the $l = 1$ contributions beyond the eight shells are ignored in the case of Ni, where we have to investigate very large cluster sizes. The fact that the $l = 3$ and $l = 1$ contributions do not

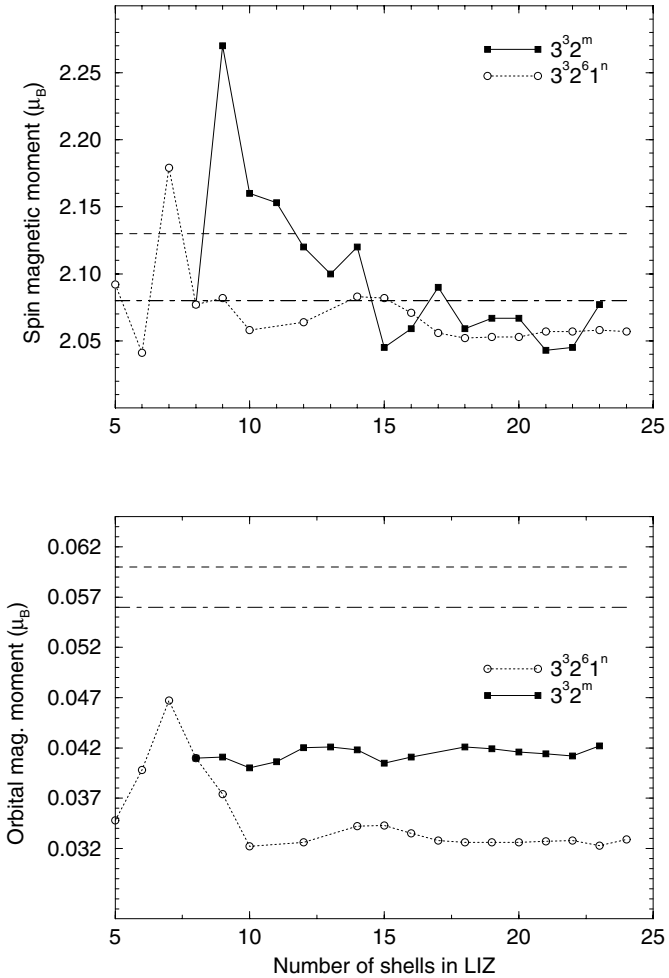


Figure 4. Convergence of the spin and orbital magnetic moments of bcc Fe as a function of the number of shells in the LIZ. The angular momentum configurations used are $l_{max} = 3^32^61^n$ (dotted line) and $l_{max} = 3^32^m$ (solid line). A screening radius of two shells is used. The dashed and dashed–dotted lines correspond to the KKR $\boldsymbol{k}$-space results given by Strange *et al* [25] and Ebert *et al* [24].

significantly influence the convergence of the magnetic moments might, in the tight-binding picture, be understood as being related to the fact that f or p electrons of these constituent atoms hardly contribute to the valence band on the central site.

The execution times of the self-consistency cycles increase significantly as a function of the number of sites with $l_{max} = 2$. Therefore in the case of Ni for example, when going beyond the eighth shell, both the $l = 2$ and $l = 1$ scattering channels are no longer included, which means that in the present approximation (dotted line in figure 5), the p, d and f contributions from sites beyond the central area have been removed. The substantial long-range effects that we see in Ni, implying that the multiple scattering from sites situated on shells far away still influences the electronic structure on the central site, seem to originate from the s scattering channel. Similarly for Fe and Co, those sites situated beyond the eighth shell are assumed to contribute only to the s and p scattering.

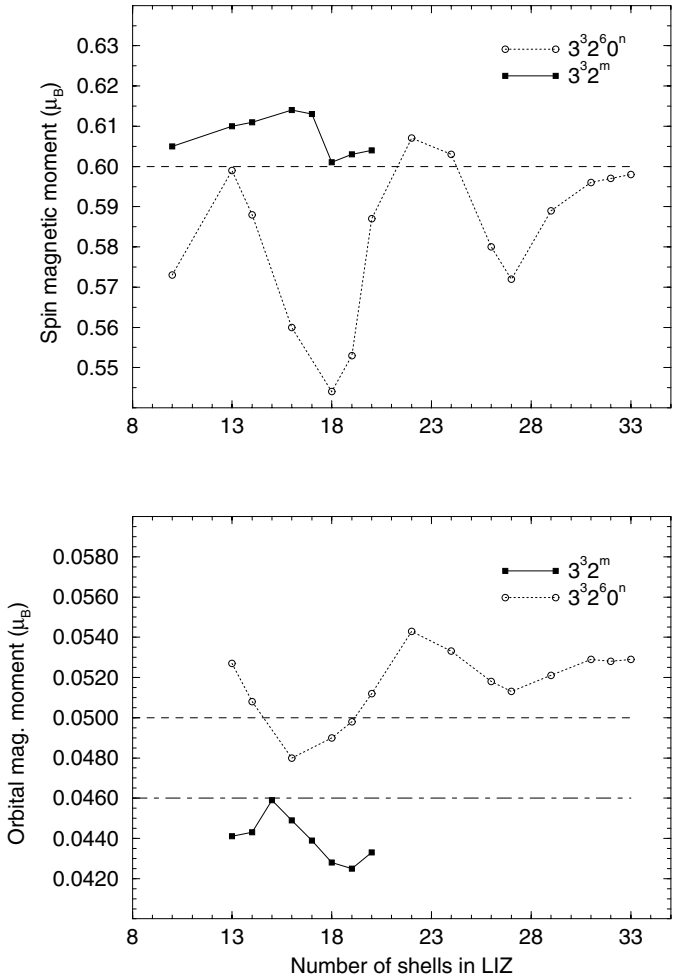


Figure 5. Convergence of the spin and orbital magnetic moments of fcc Ni as a function of the number of shells in the LIZ. The angular momentum configurations used are $l_{max} = 3^3 2^6 0^n$ (dotted line) and $l_{max} = 3^3 2^m$ (solid line). A screening radius of one shell is used. The dashed and dashed–dotted lines correspond to the KKR $\boldsymbol{k}$-space results given by Strange *et al* [25] and Ebert *et al* [24].

In the tight-binding picture, the $l = 2$ contributions are due to the overlap of the atomic d levels. The solid line in figure 4 shows the convergence of the magnetic moment of Fe, and for the angular momentum configurations $l_{max} = 3^3 2^m$, i.e. the $l = 2$ contribution from all the sites in the LIZ are taken into account here. It can be seen that, although there are considerable differences for small LIZs, the converged value seems less dependent on whether the $l = 2$ scattering channel is included. This implies that it is mainly the itinerant s contribution, through s–d hybridization, that determines the magnetic moment on the central site of the LIZ. As expected, the orbital moment is more sensitive to the angular momentum convergence. In particular, we note from figure 4 a 20% variation in the orbital moment between the angular momentum configuration $l_{max} = 3^3 2^m$ and $l_{max} = 3^3 2^6 1^n$.

A similar picture emerges for Ni (figure 5), although we have not been able to investigate the convergence for LIZs larger than 19 shells, i.e. $l_{max} = 3^3 2^{17}$. The results seem to

indicate that the converged magnetic moment will not differ considerably from the value obtained with the configuration $l_{max} = 3^3 2^6 0^{25}$. The values for the spin and orbital magnetic moments that are given in table 1 correspond to the largest possible LIZ that we were able to investigate, i.e. 1253 atoms with the angular momentum configurations $l_{max} = 3^3 2^{15} 0^{16}$ and $l_{max} = 3^3 2^{11} 0^{16}$ for respectively one and two shells of screening. For small LIZs, as can be seen from figure 5, the variation in the magnetic moment is much less pronounced. This might be due to the fact that for close-packed fcc Ni the d contributions are screened out faster than in the more open bcc structure (compare to the result for Fe). Thus, the angular momentum cut-off, justified for p and f contributions, needs to be employed with care when dealing with the d contributions.

Table 1. Screened real-space results for the spin and orbital moment of Fe, Co and Ni, compared to corresponding results obtained by the $\boldsymbol{k}$-space KKR method and the unscreened LSMS method.

	Fe (bcc)		Co (fcc)		Ni (fcc)	
Method	$m_s(\mu_B)$	$m_o(\mu_B)$	$m_s(\mu_B)$	$m_o(\mu_B)$	$m_s(\mu_B)$	$m_o(\mu_B)$
This work						
Two-shell screened (LIZ = 609 atoms)	2.057	0.033				
Three-shell screened (LIZ = 531 atoms)	2.109	0.035				
One-shell screened (LIZ = 555 atoms)			1.58	0.074		
One-shell screened (LIZ = 1253 atoms)					0.62	0.044
Two-shell screened (LIZ = 1253 atoms)					0.60	0.045
SPR-LSMS [7]						
(LIZ = 65 atoms)	2.08	0.041				
(LIZ = 87 atoms)			1.59	0.074		
(LIZ = 135 atoms)					0.546	0.044
KKR [23]	2.15		1.56		0.59	
SPRKKR [24]	2.08	0.056	1.51	0.069	0.60	0.046
SPRKKR [25]	2.13	0.06			0.60	0.05
Experiment [22]	2.13	0.09			0.57	0.05

4.2. Real space versus $\boldsymbol{k}$-space

In the limit of the 'exact' real-space implementation, i.e. for large screening radii and large LIZs, our results should converge to the 'exact' $\boldsymbol{k}$-space KKR results. In table 1, we compare our results for the spin and orbital magnetic moments to the experiment [22] and results obtained with the unscreened real-space method and the $\boldsymbol{k}$-space techniques [23–25]. The results given by Moruzzi *et al* [23] are non-relativistic, which explains why there is no value for the orbital moment. The first fully relativistic KKR calculations, with spin polarization and spin–orbit coupling being treated on an equal footing, are those given by Ebert *et al* [24] and Strange *et al* [25], marked respectively by the dashed–dotted and dashed lines in figures 4 and 5.

For Fe, we see that we have very good agreement for the spin magnetic moment with both the $\boldsymbol{k}$-space KKR method and the experiment. For the orbital moment of Fe the agreement is not as satisfactory as for the spin moment. However, the orbital moment is not well reproduced by any band-structure method, except maybe for the calculation by Trygg *et al* [26], who used the full-potential LMTO, both with and without orbital polarization (OP), finding a much improved agreement with experiment when the orbital polarization has been taken into account. The spin moments are insensitive to whether OP is included or not. This may

indicate the importance of going beyond the local spin density (LSD) and considering the effect of orbital currents.

The screened real-space results for Ni were obtained by using a LIZ of 1253 atoms (33 shells) and the respective angular momentum configurations were $l_{max} = 3^3 2^{11} 0^{20}$ with two shells of screening and $l_{max} = 3^3 2^{15} 0^{16}$ with one shell of screening. Again the agreement with the $\boldsymbol{k}$-space methods has been very good, in particular as regards the value for the orbital moment. These results show that our real-space approach has converged for the magnetic moment within the accuracy of the present state-of-the-art calculations. The spin magnetic moment is in general a more robust quantity to converge whilst the orbital magnetic moment is sensitive to the angular momentum cut-off in real space.

4.3. Magnetocrystalline anisotropy energy

In table 2, we compare our results for the magnetocrystalline anisotropy energy for the transition metals, as a function of the LIZ size and screening radius, to the results of the $\boldsymbol{k}$-space methods [27, 28]. For the bulk, these energies are extremely small, i.e. of the order of μeV. Therefore, to get meaningful results, the band energies had to be converged to the accuracy of at least 10^{-8} eV. To accomplish this, up to a hundred iterations in the self-consistency cycle have been needed for the modest sizes of the LIZs and screening radii. Note that the $\boldsymbol{k}$-space methods, which we compare our results to, also suffer due to the latter and additionally have problems with the Brillouin zone integrations, where huge numbers of $\boldsymbol{k}$-points are needed to determine the tiny MAEs. Here the MAE is calculated as the one-electron sum energy difference for the magnetizations pointing along the [001] and [111] directions, i.e. MAE $= E_{sum}^{[001]} - E_{sum}^{[111]}$.

Table 2. Results for the magnetocrystalline anisotropy energy, $E_{sum}^{[001]} - E_{sum}^{[111]}$, for fcc Co, fcc Ni and bcc Fe. The screened real-space method results are compared to the unscreened real-space method and the $\boldsymbol{k}$-space method.

Method	MAE (μeV/atom) fcc Co	MAE (μeV/atom) fcc Ni	MAE (μeV/atom) bcc Fe
This work			
One-shell screened (LIZ = 555 atoms)	0.544		
One-shell screened (LIZ = 135 atoms)		1.8	
One-shell screened (LIZ = 1175 atoms)		2.2	
Two-shell screened (LIZ = 135 atoms)		−0.31	
Two-shell screened (LIZ = 1175 atoms)		−0.34	
Two-shell screened (LIZ = 65 atoms)			−0.73
Two-shell screened (LIZ = 339 atoms)			−0.66
SPR-LSMS [7]			
(LIZ = 87 atoms)	1.050		
(LIZ = 135 atoms)		−0.434	
(LIZ = 65 atoms)			−0.78
SPRKKR [27]		10.5 ± 7.0	
SPRKKR [28]	0.86	0.11	−0.95
Experiment [22]	1.52	2.7	−1.3

For Co and Fe, the correct easy axis of the magnetization is found, i.e., the sign agrees with the one observed experimentally. When compared to experiment, our values for the MAE are too small by a factor of 3 for Co and a factor of 2 for Fe, but fit well within the rather broad

range of values found by $\boldsymbol{k}$-space KKR and LMTO [26] calculations. Our results are smaller than the corresponding KKR $\boldsymbol{k}$-space values and the unscreened real-space values. The reason for this is most probably insufficient convergence with respect to the screening radius.

For Ni, we find the wrong easy axis of magnetization. This result is in agreement with the LMTO calculations [26, 29, 30], whereas Razee *et al* [28], using the $\boldsymbol{k}$-space KKR method, have obtained a positive value. In Ni, d electrons are slightly more localized, which is indicated by a narrower d band, and one may expect the orbital moment to play a more prominent role in this system. Trygg *et al* [26] have taken this into account by including the orbital polarization effects within the full-potential LMTO method (FP-LMTO-OP). Whereas this usually gives a better agreement with experimental results as regards the magnitude, it does not have any effect on the sign of the magnetic anisotropy, i.e., they also find the wrong easy axis for Ni. In their calculations the MAE has been obtained from the total energies, i.e., the force theorem has not been used. Since the LSD has been the only approximation used in their calculation, one is inclined to conclude that the LSD cannot describe Ni, because the orbital motion of the electrons cannot be well represented by the free-electron gas [31].

Further to our results, we should mention that although the MAE is rather insensitive to the size of the LIZ, it is very dependent on the size of the screening radius, which needs to be large enough for good convergence. Also, as can be seen in table 2, the sign of the magnetic anisotropy energy of Ni can change when a screening radius of two shells is used instead of the screening radius of only one shell.

What we have not been able to investigate in this paper with respect to the MAE, because of the huge matrix sizes it gives rise to, is the influence of the angular momentum cut-off, when going beyond the eighth shell in the LIZ. We have seen that the magnetic moment has been very sensitive to this, and it would have been interesting to see the effect, if any, on the MAE. Since it is possible to approximately determine the MAE using relatively small clusters, it seems that one is dealing here with a phenomenon that is rather localized in space. Overall, our results have been quite satisfactory, indicating that within the validity of the LSD, the MAE can be calculated in real space to an acceptable degree of accuracy.

4.4. Orbital moment anisotropy

In transition metals, both the MAE and the orbital moment are due to the spin–orbit coupling [32]. It has been shown by Bruno [33], when treating the spin–orbit interaction in first-order perturbation theory and assuming the band splitting to be much larger than the crystal-field splitting, that the magnetocrystalline anisotropy energy, with respect to the magnetization directions $\hat{M}_1$ and $\hat{M}_2$, is related to the corresponding anisotropy in the orbital moment through

$$E_{sum}^{[\hat{M}_1]} - E_{sum}^{[\hat{M}_2]} = \frac{\xi}{4\mu_B}(m_{orb}^{[\hat{M}_2]} - m_{orb}^{[\hat{M}_1]}) \tag{8}$$

where ξ is the spin–orbit coupling parameter. From this equation it follows that it should be possible to identify the easy axis from the magnetization direction which has the largest orbital moment [3, 4]. If the magnetization lies along the hard axis, the spin–orbit coupling will force the orbital moment away from its preferred direction, giving rise to a small component of the orbital moment along the hard axis.

For all systems studied, we have calculated the components of the orbital moment for the magnetization along both the [001] and the [111] directions. Defining the anisotropy in the orbital moment as OMA $= L_z^{[001]} - L_z^{[111]}$, we have found for Co and Ni, respectively, the values of $-52 \times 10^{-6}\ \mu_B$ and $78 \times 10^{-6}\ \mu_B$. Since for the fcc structure the easy axis is along the [111] direction, one can see that our real-space calculations (based on formula (8))

give the wrong easy axis of magnetization for Ni, but the correct easy axis is obtained for Co, where the orbital moment is largest for the magnetization along [111].

The value of the orbital moment is very sensitive to the assumed angular momentum configuration in the LIZ. It has been noticed in the case of the bcc Fe that the LIZs of the same size, but with a different number of shells with the $l_{max} = 2$ configuration on the sites, can actually lead to different signs for the OMA. Namely, for the LIZ of 339 atoms, and the angular momentum configurations of respectively $l_{max} = 3^3 2^6 0^8$ and $l_{max} = 3^3 2^7 0^7$, we have calculated the OMAs to be $-2.98 \times 10^{-6}\ \mu_B$ and $28 \times 10^{-6}\ \mu_B$ whilst the sign of the MAE remains unaltered. The force theorem has been used to derive both the MAEs and the OMAs. Although there is no proof for the variational nature of the OMA, it has been reassuring that a fully self-consistent calculation for both magnetization directions for the $l_{max} = 3^3 2^6 0^8$ configuration has also given OMA $= -2.98 \times 10^{-6}\ \mu_B$, thus justifying the use of the force theorem.

The easy axis for bcc Fe is along [001], and our results show that to get the correct easy axis, the $l = 2$ contributions are of paramount importance. Thus, the MAE is less sensitive to the $l = 2$ contributions than the OMA, which might be due to the fact that the orbital moment points along a specific direction and its calculation requires a higher angular momentum expansion. Note also that for 65 atoms in the LIZ, and with all the sites having at least $l_{max} = 2$, the OMA is $17.3 \times 10^{-6}\ \mu_B$, i.e. it has the correct sign. The actual angular momentum configuration for this cluster is $l_{max} = 3^3 2^3$, i.e. there are fewer shells with $l_{max} = 2$ than in either of the two configurations for the cluster of 339 atoms, discussed above. What these results seem to imply is that as long as the convergence with respect to the $l = 2$ contributions has not been reached, including as many outer shells as possible with the $l_{max} = 0$ configuration gives worse accuracy than when simply restricting the size of the LIZ cluster, but having the $l_{max} = 2$ configuration on all the sites.

5. Conclusions

We have shown that the screened relativistic and spin-polarized real-space KKR band-structure method can be useful for studying magnetic properties of solids. Owing to the exponentially decaying screened structure constants and the use of the sparse matrix techniques, the inversion of the corresponding KKR matrix scales nearly as $O(N)$. The subsequent gains in the execution times allow one to study systems with local interaction zones of more than a thousand atoms. However, it is important to find the right balance between the size of the LIZ and the angular momentum configuration. The magnetic moments on the central site can be obtained to a good accuracy with restricted angular momentum configurations, but may need substantial LIZs as we demonstrated for Ni. On the other hand, directional properties, such as the MAE and the OMA, require the inclusion of the $l_{max} = 2$ configuration on all the sites of the LIZ. In this case the size of the LIZ is not so much the determining quantity. Of course, a drawback of the method is that the screened structure constants need to be calculated numerically. Since the dimension of the screened structure constant matrix and its sparsity are the determining factors in the matrix inversion, approximations with respect to the screening radius and the on-site angular momentum configuration have to be implemented. Their implications have been thoroughly examined, and it has been found out that the d scattering channel has been most sensitive to those approximations. The present application to the elemental transition metals, the fcc Co and Ni, and the bcc Fe, to study their magnetic properties, has been motivated by the need for benchmarking the method. While the spin and orbital magnetic moments of Fe and Co have turned out to converge fast in real space, the LIZs of about 1200 atoms have been necessary to reach convergence for Ni. It has been encouraging, however, that the converged values for

these quantities compare well with the results of the corresponding $\boldsymbol{k}$-space methods. The fact that we have succeeded in calculating the magnetic moments and the magnetocrystalline anisotropy energies within a real-space approach proves that these properties are determined by the local environment, and that the method can provide new insight into the physics at the origin of magnetism in these materials. Unfortunately, the computational effort, for very large LIZs, can be quite substantial. The specific two-shell screening result for Ni in table 1 took a week to complete on a 500 MHz DEC-Alpha workstation, whereas the same result could be easily obtained in minutes with the $\boldsymbol{k}$-space KKR. However, it is not the elemental transition metals that the method is aimed at. For periodic systems, where $\boldsymbol{k}$-space methods can be applied, the present real-space method is uncompetitive, but not many real systems are periodic. At interfaces and grain boundaries, for steps on surfaces and for nanowires, for example, the $\boldsymbol{k}$-space methods can no longer be applied, and it is here that the advantage of the present method can be appreciated and moreover can be implemented with varying degrees of accuracy depending on the amount of computational effort one wishes to use.

Acknowledgments

Useful discussions with Professor Peter H Dederichs and Dr Rudi Zeller are highly appreciated. Moreover, one of us (LP) would like to acknowledge the financial support of the Psi-k European Network, and the TMR 'Interface Magnetism' Network (contract FMRX-CT96-0089).

References

[1] Berkowitz A E, Mitchell J R, Carey M J, Young A P, Zhang S, Spada F E, Parker F T, Hutten A and Thomas G 1992 *Phys. Rev. Lett.* **68** 3745
[2] Xiao J Q, Jiang J S and Chien C L 1992 *Phys. Rev. Lett.* **68** 3749
[3] Dürr H A, Guo G Y, van der Laan G, Lee J, Lauhoff J A C and Bland J A C 1997 *Science* **277** 213
[4] Hjortstam, O, Baberschke, K, Wills J M, Johansson B and Eriksson O 1997 *Phys. Rev.* B **55** 15 026
[5] Nicholson D M C, Stocks G M, Wang Y, Shelton W A, Szotek Z and Temmerman W M 1994 *Phys. Rev.* B **50** 14 686
[6] Wang Y, Stocks G M, Shelton W A, Nicholson D M C, Szotek Z and Temmerman W M 1995 *Phys. Rev. Lett.* **75** 2867
[7] Beiden S V, Temmerman W M, Szotek Z, Gehring G A, Stocks G M, Wang Yang, Nicholson D M C, Shelton W A and Ebert H 1998 *Phys. Rev.* B **57** 14 247
[8] Beiden S V, Guo G Y, Temmerman W M, Szotek Z, Gehring G A, Wang Y, Stocks G M, Nicholson D M C, Shelton W and Ebert H 1996 *Materials Theory, Simulations and Parallel Algorithms (MRS Proc. No 408)* ed E Kaxiras, J Joannopoulos, P Vashishta and R K Kalia (Pittsburgh, PA: Materials Research Society)
[9] Braspenning B J and Lodder A 1994 *Phys. Rev.* B **49** 10 222
[10] Andersen O K and Jepsen O 1984 *Phys. Rev. Lett.* **53** 2571
[11] Andersen O K, Postnikov A V and Savrasov S Y 1992 *Application of Multiple Scattering Theory to Materials Science (MRS Proc. No 253)* ed W H Butler, P H Dederichs, A Gonis and R L Weaver (Pittsburgh, PA: Materials Research Society)
[12] Zeller R, Dederichs P H, Ujfalussy B, Szunyogh L and Weinberger P 1995 *Phys. Rev.* B **52** 8807
[13] Petit L, Beiden S V, Temmerman W M, Szotek Z, Stocks G M and Gehring G A 1998 *Phil. Mag.* B **78** 449
[14] Zeller R 1997 *Phys. Rev.* B **55** 9400
[15] Szunyogh L, Ujfalussy B, Weinberger P and Kollar J 1994 *Phys. Rev.* B **49** 2721
[16] Wildberger K, Zeller R and Dederichs P H 1997 *Phys. Rev.* B **55** 10 074
[17] Duff I S 1996 *Rutherford Appleton Laboratory Report No* RAL-TR-047
[18] Strange P, Ebert H, Staunton J B and Gyorffy B L 1989 *J. Phys.: Condens. Matter* **1** 2959
[19] Strange P, Staunton J and Gyorffy B L 1984 *J. Phys. C: Solid State Phys.* **17** 3355
[20] Mackintosh A R and Andersen O K 1980 *Electrons at the Fermi Surface* ed M Springford (Cambridge: Cambridge University Press)
[21] Zeller R, private communication
[22] *Landolt–Börnstein New Series* 1988 Group III, vol 19a, ed H P J Wein (Berlin: Springer)

[23] Moruzzi V L, Janak J F and Williams A R 1978 *Calculated Electronic Properties of Metals* (New York: Pergamon)
[24] Ebert H, Strange P and Gyorffy B L 1988 *J. Phys. F: Met. Phys.* **18** L135
[25] Strange P, Staunton J B, Gyorffy B L and Ebert H 1991 *Physica* B **172** 51
[26] Trygg J, Johansson B, Eriksson O and Wills J M 1995 *Phys. Rev. Lett.* **75** 2871
[27] Strange P, Ebert H, Staunton J B and Gyorffy B L 1989 *J. Phys.: Condens. Matter* **1** 3947
[28] Razee S S A, Staunton J B and Pinski F J 1997 *Phys. Rev.* B **56** 8082
[29] Daalderop G H O, Kelly P J and Schuurmans M F H 1990 *Phys. Rev.* B **41** 11 919
[30] Guo G Y, Temmerman W M and Ebert H 1991 *Physica* B **172** 61
[31] Jansen H J F 1995 *Phys. Today* (4) 50
[32] Brooks H 1940 *Phys. Rev.* **58** 909
[33] Bruno P 1989 *Phys. Rev.* B **39** 865

Magnetostatic coupling in spin valves: Revisiting Néel's formula

T. C. Schulthess[a)] and W. H. Butler
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6114

We use a numerical, atomistic approach to calculate the magnetostatic coupling in spin valves. In addition to the numerical treatment, the coupling energy is evaluated analytically and it is shown that Néel's formula is accurate to first order in the ratio of roughness amplitude to grain size. We also generalize the formula so that it can be applied to systems such as Py/Co/Cu/Co/Py spin valves that have complex ferromagnetic layers. © *2000 American Institute of Physics.*
[S0021-8979(00)45208-4]

I. INTRODUCTION

The two ferromagnetic (FM) layers in spin valve type trilayers (magnetic heterostructures consisting of two ferromagnetic films separated by a nonmagnetic spacer) tend to couple ferromagnetically. In many cases this coupling is attributed to the roughness of the interfaces on both sides of the nonmagnetic spacer and the magnetic charges that form as a consequence of this roughness.[1–3] Since this roughness is often correlated, especially when film growth is by columnar grain growth of the spin valve stack (Fig. 1), the magnetostatic interactions between the charges favor parallel alignment of the respective magnetizations of the two FM layers. This coupling is known as orange peel coupling (OPC) or Néel coupling. In the early 1960s Néel,[4] following a suggestion by Metfessel,[5] calculated the magnetostatic energy of two semi-infinite ferromagnetic layers A and A' that separated by a vacuum region B such that the interfaces $A-B$ and $B-A'$ are assumed to have a two dimensional waviness of the form

$$z(x,y) = -h \sin\frac{2\pi x}{\lambda} \sin\frac{2\pi y}{\lambda}. \tag{1}$$

In his calculation, Néel projected the magnetic charge distribution that results from the waviness onto the xy plane and solved for the magnetostatic energy (per unit area) of two FM layers with charged interface planes. For the case in which the magnetization in both FM layers is rigid,[6] the solution for parallel alignment of the FM layers is

$$E_{\uparrow\uparrow} = \frac{\pi^2}{\sqrt{2}} M_s M_s' h h' e^{-2\pi d\sqrt{2}/\lambda}, \tag{2}$$

where M_s and M_s' represent the saturation magnetizations of the FM layers A and A', respectively. The magnetostatic energy for antiparallel alignment of the FM layers is $E_{\uparrow\downarrow} = -E_{\uparrow\uparrow}$.

The assumption that the FM magnetization is rigid is justified since the grain sizes in typical spin valves are not much larger than the exchange length of the FM materials involved. However, it is not clear how valid the assumption is of a flat distribution of charges for the case in which the amplitude of the roughness h is of the same order of magnitude as the thickness d of the nonmagnetic spacer (typical values in spin valve applications $h = 0.5-1$ nm and $d = 2$ nm).

In the present work we take an atomistic approach to calculate the OPC which does not make any assumption about the topology of the interfacial charge distribution. We derive expressions for the magnetostatic energy that can be evaluated numerically and we use an approximate analytic evaluation of these expressions to derive Néel's formula. The direct comparison of numerical and analytical results gives a clear understanding of the range of validity of Néel's formula and indicates how OPC can be included in micromagnetic models of spin valves. This last point is particularly important because present micromagnetics models neglect interfacial roughness even though the OPC related coupling field can be as large as 10 Oe,[1,2] which is comparable to other characteristic fields that appear in spin valves.[7]

II. NUMERICAL APPROACH

For the numerical treatment we assume that the FM is composed of localized atomic moments $\mathbf{m}_i$ that are situated on atomic sites i with position vectors $\mathbf{R}_i$. Furthermore we choose models for the interface roughness in such a way that the film is two dimensional (2D) periodic in the xy plane. The magnetization can thus be written as

$$\mathbf{M}(\mathbf{r}) = \sum_i \frac{\mathbf{m}_i}{V_i} \sum_{\mathbf{R}} \delta[\mathbf{r} - (\mathbf{R}_i - \mathbf{R})], \tag{3}$$

where $\mathbf{R}$ are the 2D lattice vectors, V_i represents the volume of the ith atomic cell, and the i summation runs over all sites in the unit cell. The magnetostatic energy per unit area is then given by

$$E = -\frac{1}{2\Omega} \sum_{i \neq j} \mathbf{m}_i \mathbf{D}_{ij} \mathbf{m}_j, \tag{4}$$

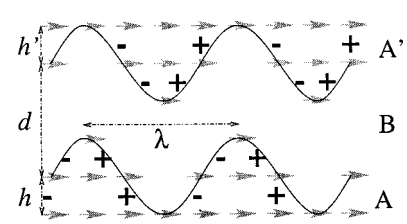


FIG. 1. Schematic of two FM layers with correlated wavy surfaces. Arrows indicate positions of atomic moments.

[a)]Author to whom correspondence should be addressed; electronic mail: schulthesstc@ornl.gov

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

where Ω is the area of the unit cell, and the dipolar interaction matrix is given by

$$(\mathbf{D}_{ij})_{\mu\nu}=\sum_{\mathbf{R}} \frac{3(\hat{\mathbf{n}}_i)_\mu(\hat{\mathbf{n}}_j)_\nu-\delta_{\mu\nu}}{|\mathbf{R}_i-\mathbf{R}_j-\mathbf{R}|^3} \tag{5}$$

with $\hat{\mathbf{n}}_i=\mathbf{R}_i/|\mathbf{R}_i|$ and $\mu(\nu)$ labeling Cartesian coordinates.

The OPC is usually extracted from a positive offset, H_o in the magnetization loop of the free layer f of the spin valve.[1] In our atomistic approach, when all moments in a layer are parallel,

$$H_o=\frac{J}{\left(\sum_{i\in f}\frac{m_i}{V_i}\right)t_f}, \tag{6}$$

where t_f is the thickness of the free layer and

$$J=(E_{\uparrow\downarrow}-E_{\uparrow\uparrow}) \tag{7}$$

is the coupling energy between the layers. Since the sum in Eq. (5) is conditionally convergent a few more transformations are required before Eqs. (4), (5), and (7) can be evaluated numerically.

In the present work we assume that $d>2h$ so that $|\mathbf{R}_i-\mathbf{R}_j|\neq 0$, when the sites i and j are not in the same FM layer. Since, in the case for which the magnetization of an individual layer FM is rigid, the only contributions to the magnetostatic energy [Eq. (4)] that do not cancel in Eq. (7) are those for which i and j are in different layers, we can calculate the sums in Eq. (5) with Fourier transforms. Assuming that the magnetization points along the x axis we get

$$J=\frac{1}{\Omega}\sum_{i\in A, j\in A'} m_i(D_{ij})_{xx}m_j, \tag{8}$$

where

$$(D_{ij})_{xx}=-\frac{2\pi}{\Omega}\sum_{G\neq 0} e^{-G|R_{i,z}-R_{j,z}|}\frac{G_x^2}{G}\cos[\mathbf{G}(\mathbf{R}_i-\mathbf{R}_j)]. \tag{9}$$

Evidently, this last expression converges whenever $|R_{i,z}-R_{j,z}|>0$. For the case of a distribution of moments on a square lattice, the reader can verify that this expression reduces to the result obtained by Tsymbal.[8]

III. APPROXIMATE ANALYTIC TREATMENT

To derive Néel's OPC result from the expressions given in Sec. II, we begin by calculating the field at some site $\mathbf{r}$ (with $r_z>h$) due to magnetic moments which are situated below the surface

$$z(x,y)=h\left(\sin^2\frac{\pi x}{L}\cos^2\frac{\pi y}{L}-\cos^2\frac{\pi x}{L}\sin^2\frac{\pi y}{L}\right). \tag{10}$$

Note that Eqs. (1) and (10) are equivalent when $\lambda=\sqrt{2}L$. In accordance with the assumptions that lead to Eq. (2), we can assume here that all the moments vectors are identical and point along the x axis, i.e., $\mathbf{m}_i=m\hat{\mathbf{x}}$ and $M_s=m/V$. Using Eq. (9) we have for the x component of the magnetostatic field

$$H_x=-\frac{2\pi}{\Omega}M_s\sum_{G\neq 0}\frac{G_x^2}{G}\sum_i e^{-G|r_z-R_{i,z}|}\cos[\mathbf{G}(\mathbf{r}-\mathbf{R}_i)]. \tag{11}$$

The first approximation we make is to retain only the first shell in the reciprocal lattice sum and replace the summation over sites by integrals. This converts Eq. (11) to

$$H_x=-\frac{8\pi^2 e^{-r_z 2\pi/L}}{\Omega L}M_s\int_0^L dx\,dy\cos\left(\frac{2\pi}{L}(r_x-x)\right)\times\int_{-h}^{z(x,y)}dz'\,e^{z'2\pi/L}. \tag{12}$$

The z integral yields an exponential function which we evaluate only to first order in (h/L), i.e.,

$$\int_{-h}^{z(x,y)}dz'\,e^{z'2\pi/L}=\frac{L}{2\pi}(e^{z(x,y)2\pi/L}-e^{-h2\pi/L})=[z(x,y)+h]+o(h/L)^2. \tag{13}$$

Note that this is the second approximation which we made. The remaining integrals in Eq. (12) are now straightforward to evaluated and we find

$$H_x=2\pi^2 M_s\frac{h}{L}\cos\frac{2\pi R_x}{L}e^{-R_z 2\pi/L}. \tag{14}$$

An analogous expression can be obtained for the field at a site with $R_z<d-h$ due to moments situated on sites above $z(x,y)+d$.

Using this last result to calculate the magnetostatic self energy $E=-1/(2\Omega)\int MHd^3r$ (where we again replace the summation over moment by an integral) for the parallel and antiparallel states of the spin valve, we get, after exploiting all possible symmetries,

$$J=\frac{4\pi^2}{\Omega}M_s^2\frac{h}{L}\int_0^L dx\,dy\cos\frac{2\pi x}{L}\int_{d+z(x,y)}^{\infty}dz\,e^{-z2\pi/L}+o(h/L)^2. \tag{15}$$

Repeating the steps which lead from Eq. (12) to Eq. (14) yields the final result for the coupling energy

$$J=\sqrt{2}\pi^2 M_s^2\frac{h^2}{\lambda}e^{-2\pi\sqrt{2}d/\lambda}+o(h/\lambda)^2, \tag{16}$$

where we have replaced L by $\lambda=L/\sqrt{2}$. This formula is equivalent to the combination of Néel's formula, Eq. (2), and Eq. (4).[9]

IV. DISCUSSION

The results of Eqs. (8) and (16) are compared in Fig. 2. For the numerical treatment we have assumed two 8 nm thick Permalloy films for which, for simplicity, we assume that atomic moments of $1\mu_B$ occupy a simple cubic lattice with lattice constant $a=2.25$ Å and that interface planes are parallel to (001). The interfaces with the spacer layer have a sinusoidal waviness of the from discussed in Sec. II and the outer surfaces are flat. This is a very accurate model of two semi-infinite FM slabs, because the contributions of perfect

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

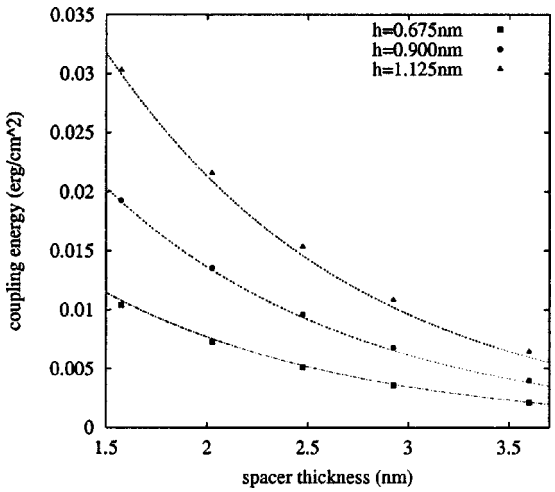


FIG. 2. Numerical evaluation of coupling energy (symbols) compared with results of Eq. (16) as a function of spacer thickness for different roughness amplitudes. Results are for $L=\lambda/\sqrt{2}=7.875$ nm.

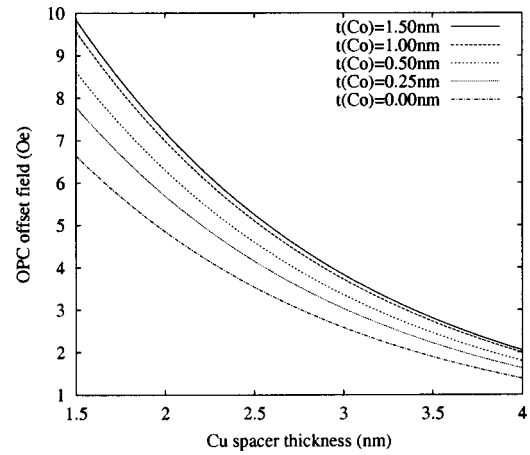


FIG. 3. OPC offset field for a Py(3 nm)Co(t_{Co})/Cu/Co(t_{Co})/Py(3 nm) spin value with a double sinusoidal waviness of amplitude $h=0.3$ nm on all interfaces. Results are for $L=\lambda/\sqrt{2}=10$ nm.

layers to the energy fall off with a factor $-2\pi|z|/a$ instead of $-2\pi|z|/L$ in the exponential of Eq. (9). The results in Fig. 2 confirm our expectations from the derivation of Néels formula in Sec. III: Independent of the spacer thickness, the agreement is best for the smallest values of h but the deviations are still small for values $h\simeq 1$ nm since the error is of order $(h/L)^2$.

The formulas of Sec. II are valid for any distribution of magnetic moments in the unit cell and are thus applicable to inhomogeneous films. It is also straightforward to generalize the analytic derivation of Sec. III to the case of multiple FM layers. For the important case of spin-valve applications, where Co is introduced as a diffusion barrier between Py and the Cu on both sides of the spacer, one has to consider a Py/Co/Cu/Co/Py multilayer. For this case the coupling energy is (to first order in h/L)

$$
\begin{aligned}
J=\frac{\pi^2}{L}h^2e^{-(2\pi d/L)}(M_{\mathrm{Py}})^2\Bigg\{&\left(\frac{M_{\mathrm{Co}}}{M_{\mathrm{Py}}}\right)^2[1-e^{-(2\pi/L)t_2}\\
&-e^{-(2\pi/L)t_3}+e^{-(2\pi/L)(t_2+t_3)}]+\frac{M_{\mathrm{Co}}}{M_{\mathrm{Py}}}e^{-(2\pi/L)t_2}\\
&\times[1-e^{-(2\pi/L)t_1}-e^{-(2\pi/L)t_3}+e^{-(2\pi/L)(t_1+t_3)}]\\
&+\frac{M_{\mathrm{Co}}}{M_{\mathrm{Py}}}e^{-(2\pi/L)t_3}[1-e^{-(2\pi/L)t_2}-e^{-(2\pi/L)t_4}\\
&+e^{-(2\pi/L)(t_2+t_4)}]+e^{-(2\pi/L)(t_2+t_3)}[1-e^{-(2\pi/L)t_1}\\
&-e^{-(2\pi/L)t_4}+e^{-(2\pi/L)(t_1+t_4)}]\Bigg\},
\end{aligned}
$$

where t_i represent the respective thicknesses of the FM layers in the stack Py/Co/Cu/Co/Py. For simplicity we have assumed that all interfaces have the same roughness amplitude. Results for the corresponding offset field are illustrated in Fig. 3, from which one concludes that, in order to reduce the OPC related offset, the Co diffusion barrier should be kept as thin as possible. This is because the magnetic charges at the Co/Py interfaces are anticorrelated with the charges at the Co/Cu interface on the other side of the spacer and thus cancel the OPC. A similar result was found for the case of very thin homogeneous FM layers by Zhang and White[10] numerically and by Kools *et al.*[11] in terms of Néel's model.

In summary, we have investigated the magnetostatic coupling due to correlated roughness in spin valves both numerically and analytically. The derivation and the numeric examples show that Néel's formula is correct to first order in h/λ. A generalization for Py/Co/Cu/Co/Py spin valves has been given, which shows that in order to reduce the OPC offset field, the Co diffusion barrier should be kept a thin as possible.

ACKNOWLEDGMENT

This work is supported by Defense Advanced Research Projects Agency Office of Naval Research.

[1] J. C. S. Kools, J. Appl. Phys. **77**, 2993 (1995).
[2] R. D. K. Misra, T. Ha, Y. Kadmon, C. J. Powell, M. D. Stiles, R. D. McMichael, and W. F. Eglehoff Jr., Mater. Res. Soc. Symp. Proc. **384**, 373 (1995).
[3] Th. G. S. Rijks, R. F. O. Reneerkens, R. Coehoorn, J. C. S. Kools, M. F. Gillies, J. N. Chapman, and W. J. M. de Jonge, J. Appl. Phys. **82**, 3442 (1997).
[4] L. Néel, C. R. Acad. Sci., Ser. R **255**, 1545 (1962); **255**, 1676 (1962).
[5] See R. Montmory and A. Yelon, in *Basic Problems in Thin Film Physics*, edited by R. Niedermayer and H. Mayer (Vandenhoeck & Ruprecht, Göttingen, 1966), p. 470.
[6] Néel in Ref. 4 also treated the case where the magnetization within a layer relaxes. For reasons given below we will not considerer this case here.
[7] A. S. Arrott, Y. Lu, R. Katti, S. Li, J. Sather, D. Schipper, T. Zhu, M. Kowalevski, W. H. Butler, and T. C. Schulthess, J. Appl. Phys. (these proceedings).
[8] E. Tsymbal, J. Magn. Magn. Mater. **130**, L6 (1994).
[9] Note that the corresponding expression in S.I. units is $J=(\pi/\sqrt{8})M_s^2\mu_0(h^2/\lambda)\exp(-2\pi\sqrt{2}d/\lambda)$, which differs by a factor of 2π from the expression used by Kools in Ref. 1 to evaluate the experimental data. This equation gives values for the OPC related fields that are in good quantitative agreement with experiment (Fig. 7 in Ref. 1) when the measured values for the roughness parameter (Fig. 2 in Ref. 1) are used.
[10] J. Zhang and R. M. White, IEEE Trans. Magn. **32**, 4630 (1996).
[11] J. C. S. Kools, W. Kula, D. Mauri, and T. Lin, J. Appl. Phys. **85**, 4466 (1999).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

Calculating properties with the polymorphous coherent-potential approximation

B. Ujfalussy
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

J. S. Faulkner
Alloy Research Center and Department of Physics, Florida Atlantic University, Boca Raton, Florida 33431

N. Y. Moghadam and G. M. Stocks
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

Yang Wang
Pittsburgh Supercomputing Center, Pittsburgh, Pennsylvania 15213
(Received 21 May 1999; revised manuscript received 10 February 2000)

The formulas for calculating properties of an alloy such as the density of states, the charge density, and the Bloch spectral density function are derived from multiple-scattering theory for the polymorphous coherent-potential approximation (PCPA). The chemical shifts obtained for three alloy systems using the PCPA, the Korringa-Kohn-Rostoker CPA, and the locally self-consistent multiple-scattering method are compared with experiment. A significant improvement in the treatment of Coulomb effects is achieved using the PCPA with only a little more computational effort than for the older isomorphous CPA's.

I. INTRODUCTION

The one-electron method based on the density-functional theory and the local-density approximation[1] (DFT-LDA) is used routinely to calculate the energetics of ordered crystalline solids. These band-theory methods cannot be used to treat disordered solids, even such conceptually simple ones as substitutional solid-solution alloys, because of the lack of long-range order. Faced with this situation, many researchers turn to completely different approaches in their quest to explain the properties of alloy systems. Some introduce heuristic models with parameters that are obtained from fitting to observations.[2] Others carry out DFT-LDA calculations on intermetallic compounds having relatively small numbers of atoms in the unit cells with the purpose of interpolating the energetics of the infinite disordered system from those of the ordered solids.[3] Today, the electronic structure for models of disordered solids can be calculated using supercells that contain thousands of atoms. This has been made possible by the development of order-N methods based on plane-wave expansions[4] or multiple-scattering theory.[5] Still, it is useful for certain applications, and more satisfying philosophically, to have a simple approximate theory that will describe the important features of the electronic structure of alloys. Exact calculations on one-dimensional and three-dimensional models of alloys, as well as more mathematical considerations,[6] convinced theorists that the coherent-potential approximation[7] (CPA) provides such a simple approximate theory.

The CPA calculations on solvable models of alloys give qualitative guidance for the interpretation of experiments on real alloys, but to make the predictions quantitative it was necessary to merge the CPA with the Korringa-Kohn-Rostoker[8] (KKR) band theory method.[9] The KKR-CPA was later extended to produce self-consistent one-electron potentials.[10] Care was taken to insure that they satisfy the requirement of the DFT-LDA that they are obtained as functional derivatives of the potential energy with respect to the local charge densities. The addition of a self-consistency step, however, brings in an aspect of the CPA theory that had never been considered before.

In order to understand the difficulties involved with including charge self-consistency in a CPA, it is necessary to be aware that all of the model calculations and mathematical studies of the CPA made use of isomorphous models of alloys. An isomorphous model is one in which, for a binary alloy, the A atoms are all assumed to have identical charge densities $\rho_A(\mathbf{r})$ and hence potential functions $\nu_A(\mathbf{r})$, and the B atoms all have identical $\rho_B(\mathbf{r})$ and $\nu_B(\mathbf{r})$. In the CPA, a scattering matrix $\hat{\mathbf{t}}_A(\mathbf{t}_c)$ is calculated for an A atom embedded in a lattice with the effective t matrix $\mathbf{t}_c$ on all the other sites. The scattering matrix for an A atom embedded in a vacuum, which appears in $\hat{\mathbf{t}}_A(\mathbf{t}_c)$, is calculated in the usual way from the potential $\nu_A(\mathbf{r})$. The scattering matrix $\hat{\mathbf{t}}_B(\mathbf{t}_c)$ is calculated analogously. The desired t matrix, $\mathbf{t}_c$, is obtained from the null scattering requirement, $c_A\hat{\mathbf{t}}_A(\mathbf{t}_c)+c_B\hat{\mathbf{t}}_B(\mathbf{t}_c)=0$. The assumption of isomorphous models was so ubiquitous that theorists made it without thinking.

It is known from band-theory calculations on ordered intermetallic compounds that there is a charge transfer between the different species of atoms, and this leads to a Madelung contribution to the self-consistent potentials. The net charges on the A and B sites are $q_A=\int\rho_A(\mathbf{r})d\mathbf{r}-Z_A$ and $q_B=\int\rho_B(\mathbf{r})d\mathbf{r}-Z_B$, where the integrals are over the unit cells and the Z's are the atomic numbers. The charge self-consistent KKR-CPA method also predicts nonzero net charges, but it contains a curious inconsistency in that there is no Madelung potential in it. A careful analysis shows that the derivation of self-consistent potentials in a disordered

 ©2000 The American Physical Society

alloy subject to the condition that the resulting model must be isomorphic leads inevitably to the conclusion that the Madelung potentials must be zero. In addition to the arguments in the original derivations,[10] a way of understanding this paradoxical result is to note that any distribution of charged atoms will give a different Madelung potential at each site, irrespective of the kind of atom at the site. The only way for all the A atoms to have the same potential and all the B atoms to have another, as required by the isomorphous model, is to approximate the Madelung contributions with zero. It was understood at the time that the isomorphous picture was only an approximation. However, the technology to test such ideas had not yet been developed, and it was hoped that the fluctuations about the average charges would be small.

The isomorphous KKR-CPA was criticized because the Madelung potentials were set equal to zero,[11] and efforts were made to deal with this criticism. These led to two very similar methods, the screened-impurity model[12] CPA (SIM-CPA) and the screened CPA (S-CPA).[13] Both of these models are isomorphous, and the Madelung potential is calculated by placing a shell of charge around each atom at a radius R_{eff}. The total charge on the shell is equal to the charge on the atom. Using calculations on the screening of single impurities in an otherwise perfect crystal as a guide,[14] the effective radius is usually chosen to be equal to radius of the nearest-neighbor shell R_1. The resulting self-consistent potentials contain a Madelung term, and there is a Coulomb contribution to the total energy. The charge transfers q_A and q_B predicted by the SIM-CPA and S-CPA are different from the ones obtained from the KKR-CPA, and they improve the agreement with many experiments. If one accepts the argument that the mathematically correct value for the Madelung potential in an isomorphous model of an alloy is zero, as derived in Ref. 10, the derivations of the isomorphous SIM-CPA and S-CPA must contain inconsistencies. There is a sense in which the SIM-CPA and S-CPA can be justified, and that will be explained below.

The environment for developing theories of alloys changed dramatically with the advent of the order-N calculations mentioned above.[4,5] Using a technique called the locally self-consistent multiple-scattering method (LSMS),[15] first-principles DFT-LDA calculations on models of alloys using supercells that contain hundreds or even thousands of atoms have been carried out.[16–18] Figure 1 of Ref. 18 shows the distribution of atomic charges q^i for a 50% copper-zinc alloy on a fcc Bravais lattice calculated with a supercell containing 500 atoms. It is seen that there is quite a broad distribution of charges on the sites, and the numerical output shows that the q^i are different for every site i. Of course, q falls within one range of values if there is an A atom on site i, and another range if there is a B atom there. It follows that an alloy is more properly described by a polymorphous model in which the charge density $\rho^i(\mathbf{r})$ on every site is unique. This result was anticipated to some extent in Ref. 11, although those authors deduced from their calculations on supercells containing 4–12 atoms that the number of possible charges that a given kind of atom can have is small and depends only on the occupation of the sites in the nearest-neighbor shell. It is shown in Fig. 2 of Ref. 17 that the Madelung potential at any lattice site is not screened at the nearest-neighbor shell radius. Instead, the contributions to this potential from succeeding shells diminish very slowly. It is demonstrated in that same reference that a part of the Coulomb energy has a contribution u_{C1} that depends only on the average charge transfer, as would be anticipated from an isomorphous CPA, but it is also demonstrated that there is a contribution u_{C2} that is not determined by the average charge transfer but depends on the distribution of the charges about the average.

Isomorphous CPA models have been remarkably successful in explaining many interesting properties of alloys,[6] and some of the reasons for that are given in Ref. 18. Even though the Madelung potentials are set equal to zero in the KKR-CPA, that approximation gives surprisingly good values for the free energy of mixing. Another curve in Fig. 2 of Ref. 17 shows that the average of the Madelung potentials for all the A sites or B sites in the alloy is short range as assumed in the SIM-CPA and S-CPA. The parameter R_{eff} turns out to be approximately equal to R_1, and this explains the successes that these isomorphous CPA's have had. On the other hand, locating the screening charge on a shell is clearly an approximation, and defining R_{eff} to be R_1 makes the SIM-CPA and S-CPA theories precise but removes the possibility for improving them by treating R_{eff} as an adjustable parameter. In addition, it has been shown that the form for the total Coulomb energy that arises naturally in the models is unsatisfactory. In the SIM-CPA, the expression is multiplied by an adjustable parameter β that cannot be obtained from within the theory.[19] It is even more disturbing that this parameter multiplies the Coulomb energy, but it does not occur in the one-electron potential. This violates the requirement of the DFT-LDA that the one-electron potential is the functional derivative of the energy with respect to the charge density. The parameter β can be understood as an effort to emulate the contribution u_{C2} to the Coulomb energy within the limitations of an isomorphous CPA.[20]

Because of the theoretical objections to isomorphous models, it was suggested in Ref. 18 that Coulomb effects could be included better at the level of the coherent-potential approximation with a polymorphous CPA (PCPA) than with any isomorphous CPA. The PCPA will generate charge densities $\rho_{A,i}(\mathbf{r})$ and $\rho_{B,i}(\mathbf{r})$ and hence potential functions $\nu_{Ai}(\mathbf{r})$ and $\nu_{Bi}(\mathbf{r})$ that are different for every site in the alloy, as found from the first-principles LSMS calculations. The theory of the PCPA was deduced from a careful study of the results of order-N calculations, so it is necessary to describe some of the inner workings of the order-N methods in more detail.

The LSMS makes use of the principle of near-sightedness that has been espoused by Kohn[21] insofar as the continuity of the wave functions is concerned. Infinitely many atoms are included in the calculation, with supercells containing N atoms being reproduced periodically to fill all space. The multiple-scattering equations are solved completely for all the atoms in a local interaction zone (LIZ) surrounding a given atom. The t matrices for the sites outside the LIZ are set equal to zero. This process is repeated for LIZ's centered on each atom in the supercell, which makes the calculation order-N. The principle of near-sightedness is not used in the Coulomb part of the calculation. The Madelung potential for each site is calculated exactly with the contributions from the

full infinity of atomic charges, not just the ones in the LIZ, being included. It could be said that the multiple-scattering part of the calculation is order-N, but the Coulomb part is not. The Madelung potentials for solids with periodic boundary conditions are easy to calculate with the Ewald method,[22] so that does not slow down the overall procedure.

A locally self-consistent Greens function (LSGF) method has been suggested,[23] which was influenced by the success of the LSMS. In the LSGF, the LIZ's are made smaller by putting an effective scatterer on the sites outside the LIZ, rather than zero as in the LSMS. A reasonable choice, although not the only choice, for this scattering matrix is a CPA t matrix $\mathbf{t}_c$.[24] As is the case in the LSMS the Madelung potentials are calculated exactly, the sums including the infinity of atomic charges. Placing the effective scatterers on the sites of a Bravais lattice outside the LIZ limits the LSGF to systems for which the atomic sites are periodic, while the LSMS can be applied to systems in which the atoms have arbitrary atomic positions, such as bulk amorphous solids.[25] The LSMS must also be used for systems for which a homogeneous alloy is a poor reference medium, such as magnetic multilayers and interfaces,[26] and magnetic noncollinearity.[27] The LSGF is clearly the best method for the study of disordered substitutional alloys, and it predicts the same electronic structure as the LSMS if both methods are converged. In particular, the model of an alloy produced by a LSGF calculation is also polymorphous, and this holds for LSGF calculations in which the LIZ is chosen to contain only one atom. As pointed out in Ref. 18, it is the authors position that the particular kind of LSGF calculation in which the effective scatterer is determined by the CPA condition and the LIZ contains one atom has the shape that is needed for a PCPA. The originators of the LSGF did not have the construction of a PCPA as one of their goals because, among other things, they are also the originators of the SIM-CPA.[23,24,12]

It is reasonable for an expert in the CPA to worry that, while the PCPA should lead to an improved treatment of Coulomb effects and give a physically more correct picture of the alloy, all of the other desirable features of the CPA that have been so useful in applications over many years will be lost. The main purpose of this paper is to develop a mathematical formalism which leads to the PCPA and from which the site-diagonal and non-site-diagonal average Green's functions are obtained. These Green's functions are different from the ones derived for the isomorphous CPA because charge correlations are built into them, but they can be used equally well to calculate the properties of alloys. A rather subtle point that comes from the formalism is that full potentials (i.e., not muffin-tin) can be used in PCPA calculations, while they cannot in isomorphous CPA calculations.

The formulas for calculating properties of an alloy such as the density of states, the charge density, and the Bloch spectral density function with an isomorphous CPA were derived from multiple-scattering theory in a paper that will be referred to as FS.[28] In the following section, the analogous formulas will be derived for the PCPA. The differences between the two sets of formulas for the non-site-diagonal Green's function and the Bloch spectral density function $A^B(E,\mathbf{k})$ are particularly interesting. It is demonstrated that the manipulation of the DFT-LDA Green's function for an alloy under the class of restrictions that define a CPA leads to the PCPA. If the additional restriction that the resulting model must be isomorphous is invoked, the result is the KKR-CPA. The S-CPA and SIM-CPA do not fit into this chain of approximations, although, as mentioned above, they are useful approximations.

Calculations of various alloy parameters have been carried out with the LSGF within the tight-binding linear muffin-tin orbital (TB-LMTO) method, and the LIZ was chosen to contain one atom in some of these. Calculations based on the multiple-scattering theory, which is the natural language of the CPA method, will be shown in this paper. The use of this method makes it easier to compare with LSMS and KKR-CPA results. In Sec. III, some of the details of the computational methods are described. In a previous paper,[29] it was shown that the chemical shifts in alloys are a particularly sensitive measure of the Coulomb effects. The chemical shifts obtained from the KKR-CPA, PCPA, and LSMS are compared with experiment in Sec. IV. The atomic charges predicted by the three theories in real alloy systems are also compared with each other. In the last section, the argument is made that the PCPA gives useful results with only a little more effort than the older isomorphous CPA's, and other aspects of the theory are discussed.

II. CALCULATING PROPERTIES WITH THE PCPA

A. The Green's function for alloys

The DFT-LDA Schrödinger equation for a collection of N atoms can be written as

$$\left[-\nabla^2+\sum_{i=1}^{N}\nu_i(\mathbf{r})\right]\psi(\mathbf{r})=E\psi(\mathbf{r}), \tag{1}$$

where the one-electron potentials will be assumed to have the muffin-tin form. Multiple-scattering theory can deal with more general potentials, such as atomic-sphere approximation (ASA) potentials or even full potentials, but the notation becomes complex. The atomic potential $\nu_n(\mathbf{r}_n)$, where $\mathbf{r}_n=\mathbf{r}-\mathbf{R}_n$, is spherically symmetric when $\mathbf{r}$ is within a sphere centered on the lattice site $\mathbf{R}_n$, and is zero otherwise. One way to write the corresponding Green's function was shown in FS to be

$$G(E,\mathbf{r},\mathbf{r}')=\sum_{L,L'}Z_L^n(E,\mathbf{r}_n)\tau_{LL'}^{nn}Z_{L'}^n(E,\mathbf{r}_n')-\sum_{L}Z_L^n(E,\mathbf{r}_n)J_L^n(E,\mathbf{r}_n'). \tag{2}$$

In this equation, $Z_L^n(E,\mathbf{r})$ is the solution of

$$[-\nabla^2+\nu_n(\mathbf{r})-E]Z_L^n(\mathbf{r})=0, \tag{3}$$

that is regular at the origin and equals

$$Z_L^n(E,\mathbf{r})=Y_L(\mathbf{r})j_l(\kappa r)m_l^n(E)-i\kappa Y_L(\mathbf{r})h_l(\kappa r), \tag{4}$$

when r is greater than the radius of the nth muffin-tin sphere. It is assumed in Eq. (2) that $r_n'>r_{\mathbf{n}}$. The matrix $\mathbf{m}^n(E)$ is the inverse of the t matrix $\mathbf{t}^n$ that describes the scattering from the potential $\nu_n(\mathbf{r})$. Since $\mathbf{t}^n$ is diagonal for muffin-tin poten-

tials, its inverse is as well, and m_l^n is a diagonal element of that matrix. The functions $j_l(\kappa r)$ and $h_l(\kappa r)$ are Bessel functions, and κ is the square root of the energy. Following FS, the spherical harmonics $Y_L(\mathbf{r})$ are chosen to be real. The function $J_L^n(E,\mathbf{r})$ is the solution of Eq. (1) that is not regular at the origin and approaches $Y_L(\mathbf{r})j_l(\kappa r)$ when r is greater than the radius of the nth muffin-tin sphere. The coefficients $\tau_{LL'}^{nm}$ are elements of the scattering-path matrix defined below. This form of the Green's function is valid when $\mathbf{r}$ and $\mathbf{r}'$ are inside the nth muffin-tin sphere or in the interstitial region between the muffin-tin spheres, where the potential is zero.

The same Green's function can be written in a different way when $\mathbf{r}$ is in the nth muffin-tin sphere and $\mathbf{r}'$ is in the mth sphere, or they are in the interstitial region. For this case, the Green's function is

$$G(E,\mathbf{r},\mathbf{r}')=\sum_{L,L'} Z_L^n(E,\mathbf{r}_n)\tau_{LL'}^{nm}Z_{L'}^m(E,\mathbf{r}'_m). \tag{5}$$

The expressions for the Green's functions in Eqs. (2) and (5) have been used in a wide range of calculations, and are reliable. Clearly, the convergence of the sums becomes a problem if the magnitudes of the position vectors are too large.

The elements of the scattering-path matrix, $\tau_{LL'}^{nm}(E)$, are most easily obtained by taking the inverse of the matrix

$$M_{LL'}^{nm}=m_l^n\delta_{LL'}\delta_{nm}-g_{LL'}^{nm}. \tag{6}$$

That is to say,

$$\tau_{LL'}^{nm}=[\mathbf{M}^{-1}]_{LL'}^{nm}. \tag{7}$$

The functions $g_{LL'}^{nm}$ are components of the free-electron Green's functions that describe propagation from lattice sites $\mathbf{R}_n$ to $\mathbf{R}_m$, and are zero when $n=m$. The elements of the inverse of the t matrix on site $\mathbf{R}_n$, m_l^n, were defined above.

B. The averaging process

A theory of the electronic states in a disordered alloy must have a statistical as well as a quantum-mechanical aspect because the knowledge about the structure of such a system will, of necessity, be incomplete. In the ideal random alloy, which is considered here, the atoms are distributed on the sites of a Bravais lattice. The probability of an A or B atom occupying a site is c_A or c_B. In more realistic models, the Warren-Cowley short-range order coefficients measure the deviation from a random distribution.

A major difference between the isomorphous CPA and the PCPA is the nature of the averaging process used in the statistical stage. It was natural in FS to use an ensemble-averaging process, the ensemble being the set of $(N_A+N_B)!/N_A!N_B!$ alloys that can be formed by distributing N_AA atoms and N_BB atoms on the lattice sites, and then passing to the limit that N_A and N_B approach infinity. In the PCPA, the average is over the sites of one infinitely large sample. The reason for this type of averaging is that the only reliable way to calculate the Madelung potential at a site is to include the contributions from the charges on all the other sites. This is the lesson that has been learned from the LSMS, the LSGF, and other order-N techniques, as was described in the previous section. This point will be made even more clearly in Sec. II D, where the calculation of the self-consistent potential in the PCPA is discussed in more detail. The difference between the various A and B atoms is due to their spatial correlation with all of the other atoms, and information about this spatial correlation is lost in the kind of ensemble-averaging process used in FS. The site-averaged Green's function will be seen to be periodic, as it should be.

Since the properties of interest are self-averaging, the ensemble- and site-averaging processes lead to the same results when applied to the same model. The concept of self-averaging is used frequently in modern discussions of statistical physics, but it first appeared in alloy theory in the writings of Lifshitz.[30] The argument was made in FS that the ensemble-averaged quantities should represent the properties of a single real crystal because of self-averaging. For the PCPA, it is necessary to argue that average values for properties can be found from one large sample because of self-averaging.[31] It will be shown in the following that formulas derived with the site-averaging process become identical to those derived with ensemble averaging when the model is taken to be isomorphous.

C. Averaging the Green's function for the site-diagonal case

Using the site-averaging process, the average of the Green's function defined in Eq. (2) over all the N sites in a volume V of the large sample is

$$\langle G(E,\mathbf{r},\mathbf{r}')\rangle=\frac{1}{N}\sum_{\substack{i=1\\ i\subset V}}^{N}\left[\sum_{L,L'} Z_L^i(E,\mathbf{r})\tau_{LL'}^{ii}Z_{L'}^i(E,\mathbf{r}')-\sum_L Z_L^i(E,\mathbf{r})J_L^i(E,\mathbf{r}')\right]. \tag{8}$$

In each term in the sum over i, the origin of the coordinate system is moved to the lattice position $\mathbf{R}_i$. The limit of this process is reached as N and hence V increase without bound. It should be clear that this averaged Green's function is periodic, $\langle G(E,\mathbf{r},\mathbf{r}')\rangle=\langle G(E,\mathbf{r}+\mathbf{R}_n,\mathbf{r}'+\mathbf{R}_n)\rangle$, because the sum is unchanged.

Approximating this averaged Green's function using the philosophy of the coherent-potential approximation entails simplifying the scattering-path matrix elements in a specific way. As was discussed in FS, the single-site approximation to the scattering-path matrix leads to a redefinition of the matrix elements $\tau_{LL'}^{ii}$ so that they are given by the inverse of a matrix $\mathbf{M}_c^i$ whose elements are given by

$$\mathbf{M}_c^{i,ii}=\mathbf{m}^i \quad \text{for } n=i,\ m=i,$$

$$\mathbf{M}_c^{i,nm}=\mathbf{m}^c\delta_{nm}-\mathbf{g}^{nm} \quad \text{for } n\neq i, \text{ or } m\neq i, \tag{9}$$

in a block-matrix notation that eliminates the angular momentum indices. The matrix $\mathbf{m}^i$ is the inverse of the t matrix that defines scattering from the potential $\nu_i(\mathbf{r})$, and the matrix $\mathbf{m}^c$ is the inverse of the effective scatterer $\mathbf{t}_c$. Another way to write this scattering-path operator is

$$\tau^{ii}\rightarrow\tau_c^{i,00}=\mathbf{D}^i\tau_c^{00}=\tau_c^{00}\tilde{\mathbf{D}}^i, \tag{10}$$

where

$$\mathbf{D}^i=[I+\tau_c^{00}(\mathbf{m}^n-\mathbf{m}^c)]^{-1}. \tag{11}$$

As discussed in Sec. I, the CPA averaged Green's function describes a periodic system with the scattering matrix $\mathbf{t}_c$ on every site. The matrix τ_c^{nm} is the scattering-path matrix for this system, and is obtained from the inverse of the matrix $\mathbf{M}_c$ with elements

$$\mathbf{M}_c^{nm}=\mathbf{m}^c\boldsymbol{\delta}_{nm}-\mathbf{g}^{nm} \quad \text{for all } n \text{ and } m. \tag{12}$$

Since this system is periodic, $\tau_c^{ii}=\tau_c^{jj}=\tau_c^{00}$.

The site-diagonal average Green's function in the single-site approximation is then

$$G^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r}')=\frac{1}{N}\sum_{\substack{i=1\\ i\subset V}}^{N}\left[\sum_{L,L'} Z_L^i(E,\mathbf{r})\tau_{c,LL'}^{i,00}Z_{L'}^i(E,\mathbf{r}')-\sum_L Z_L^i(E,\mathbf{r})J_L^i(E,\mathbf{r}')\right]. \tag{13}$$

The average density of states per site for the alloy is an example of a site-diagonal property, and it can be found from this Green's function by

$$\langle n(E)\rangle=-\frac{1}{\pi}\mathrm{Im}\int_\Omega G^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r})dv, \tag{14}$$

where Ω is the volume of the central unit cell. Clearly, the density of states associated with any site in the effective crystal is the average of that quantity for the individual atoms in V

$$\langle n(E)\rangle=\frac{1}{N}\sum_{i=1}^{N} n^i(E), \tag{15}$$

where

$$n^i(E)=-\frac{1}{\pi}\mathrm{Im}\left[\sum_{L,L'}\int_\Omega Z_L^i(E,\mathbf{r})Z_{L'}^i(E,\mathbf{r})d\mathbf{r}\tau_{LL'}^{i,00}\right]. \tag{16}$$

The charge density on a site in the effective crystal is similarly

$$\langle\rho(\mathbf{r})\rangle=\frac{1}{N}\sum_{i=1}^{N}\rho^i(\mathbf{r}), \tag{17}$$

where

$$\rho^i(\mathbf{r})=-\frac{1}{\pi}\int_{-\infty}^{E_F}\mathrm{Im}\left[\sum_{L,L'} Z_L^i(E,\mathbf{r})Z_{L'}^l(E,\mathbf{r})\tau_{LL'}^{i,00}(E)\right]dE. \tag{18}$$

The term in Eq. (13) that includes the singular solution $J_L^i(E,\mathbf{r})$ normally does not appear in formulas for properties because it is real.

As stated above, Eq. (13) should become identical with the corresponding KKR-CPA equation when it is applied to an isomorphous model. For that case, the functions of position $Z_L^i(E,\mathbf{r})$ and $J_L^i(E,\mathbf{r})$ are all $Z_L^A(E,\mathbf{r})$ and $J_L^A(E,\mathbf{r})$ when there is an A atom on the ith site or $Z_L^B(E,\mathbf{r})$ and $J_L^B(E,\mathbf{r})$ when a B atom is there. This is equivalent to the assumption that the potentials on all the A or B sites are $\nu_A(\mathbf{r})$ or $\nu_B(\mathbf{r})$, and hence the t matrices are $\mathbf{t}^A$ or $\mathbf{t}^B$. The resulting equation is

$$\begin{aligned}G_{\mathrm{iso}}^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r}')=&\sum_{L,L'}[c_A Z_L^A(E,\mathbf{r})\tau_{LL'}^{A,00}Z_{L'}^A(E,\mathbf{r}')\\&+c_B Z_L^B(E,\mathbf{r})\tau_{LL'}^{B,00}Z_{L'}^B(E,\mathbf{r}')]\\&-\sum_L[c_A Z_L^A(E,\mathbf{r})J_L^A(E,\mathbf{r}')\\&+c_B Z_L^B(E,\mathbf{r})J_L^B(E,\mathbf{r}')],\end{aligned} \tag{19}$$

which is identical with Eq. (2.33) of FS. The most obvious difference between the Eqs. (13) and (19) is that in the PCPA all the atoms are assumed to be unique, and hence their concentrations are just $1/N$. The formulas that are the analog of Eqs. (15) and (17) are

$$\langle n(E)\rangle=c_A n^A(E)+c_B n^B(E) \tag{20}$$

and

$$\langle\rho(\mathbf{r})\rangle=c_A\rho^A(\mathbf{r})+c_B\rho^B(\mathbf{r}). \tag{21}$$

The Green's functions defined in Eqs. (13) and (19) are periodic, as they should be, $G_{\mathrm{iso}}^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r}')=G_{\mathrm{iso}}^{\mathrm{SD}}(E,\mathbf{r}+\mathbf{R}_n,\mathbf{r}'+\mathbf{R}_n)$.

D. The PCPA condition and self-consistency

To this point, the single-site approximation has been used, but nothing has been said about the definition of the effective scattering matrix $\mathbf{t}_c$. The relation that defines $\mathbf{t}_c$ in the isomorphous CPA appears in Eq. (5.24) of FS as

$$c_A\tau_c^{A,00}+c_B\tau_c^{B,00}=\tau_c^{00}, \tag{22}$$

where $\tau^{A,00}$ is obtained from Eqs. (9) by putting the inverse of the scattering matrix for the potential $\nu_A(\mathbf{r})$ on the central site and $\tau^{B,00}$ is obtained similarly. This is just the conversion of the original definition of the CPA (Ref. 7) into the language of the multiple-scattering theory.[32] The extension of this relation that defines the effective scattering matrix in the PCPA is

$$\frac{1}{N}\sum_{i=1}^{N}\tau_c^{i,00}=\tau_c^{00}, \tag{23}$$

where the scattering-path matrices are defined in Eq. (10).

The information necessary for calculating self-consistent potentials and total energies in the PCPA is contained in the site-diagonal Green's functions in Eq. (13). As in any other system in which there is charge transfer, the local part of the DFT-LDA potential $\nu_i(\mathbf{r})$ is calculated from the charge density $\rho^i(\mathbf{r})$ defined in Eq. (18). The Madelung contribution is defined using the net charges $q^j=\int\rho^j(\mathbf{r})d\mathbf{r}-Z_j$ for all $j\neq i$, where $Z_j=Z_A$ if there is an A atom on the jth site, and $Z_j=Z_B$ if there is a B atom on the site. This is the same way that $\nu_i(\mathbf{r})$ is calculated in the LSMS or the LSGF.

In practical applications of the PCPA, the charge densities $\rho^i(\mathbf{r})$ and potentials $\nu_i(\mathbf{r})$ cannot be calculated for N equal to infinity, so it is necessary to make a supercell approximation.

In this approximation, the configuration of atoms in the central supercell is reproduced periodically to fill all space. This same approximation is made in order-N methods, such as the LSMS and LSGF. The use of supercells does not interfere with the periodicity of the Green's functions demonstrated in the previous paragraph. The calculation of the Madelung potential for all the sites in the supercell takes very little time when the Ewald (Ref. 22) method is used. As a practical matter, calculations of interesting properties using supercells of various sizes have shown that the supercell approximation is not serious as long as the cells contain some hundreds of atoms. One reason for this is that the Madelung sums are not seriously affected if the actual contents of the adjoining supercells in the large sample are replaced with replicas of the central cell. Another is that the properties of interest in the calculations are self-averaging.

It may seem contradictory to talk about an effective crystal that is periodic, and at the same time talk about net charges on the sites. A similar contradiction occurs in the isomorphous CPA, because one uses $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$ to calculate the DFT-LDA potential in spite of the fact that the effective scatterers on all the sites are the same. Ensuring this does not lead to a contradiction is the purpose of the CPA condition. The major point of FS is that this is a necessary part of the formalism. A quantity like the total density of states does not require knowledge of the wave functions and can be calculated from the periodic effective crystal with $\mathbf{t}_c$ on every site using Lloyd's formula.[33] Information about individual sites must be used in the calculation of a property that requires wave functions, such as the potentials, because there is no average wave function in an alloy, only an average Green's function.

E. Averaging the Green's function for the non-site-diagonal case

The most interesting non-site-diagonal properties are related to the Bloch spectral density function. The PCPA, like the isomorphous CPA, leads to an effective Green's function that is periodic. This means that Bloch vectors $\mathbf{k}$ should play a role in the theory, even though they are not good quantum numbers. The E vs $\mathbf{k}$ relation of ordinary band theory is replaced in alloy theory with the Bloch spectral density, which is the density of states in $\mathbf{k}$ space. The formulas for the Bloch spectral density function derived in FS and displayed in Eqs. (4.8) and (4.9) of that paper are

$$A^B(E,\mathbf{k}) = -(1/\pi)\,\mathrm{Im}\,G(E,\mathbf{k},\mathbf{k}), \tag{24}$$

where

$$G(E,\mathbf{k},\mathbf{k}) = \sum_n e^{i\mathbf{k}\cdot\mathbf{R}_n}\int_\Omega G(E,\mathbf{r},\mathbf{r}+\mathbf{R}_n)\,d\mathbf{r}. \tag{25}$$

In Eq. (25), the sum is over all Bravais lattice vectors $\mathbf{R}_n$, and the integral is over the central unit cell. It should be noted that $G(E,\mathbf{k},\mathbf{k})$ is not the Fourier transform of the Green's function, as is used, e.g., in the analysis of positron annihilation experiments.

The Green's function that must be used in $A^B(E,\mathbf{k})$ is obtained from Eq. (5) by means of the site-averaging process

$$\langle G(E,\mathbf{r},\mathbf{r}+\mathbf{R}_n)\rangle = \frac{1}{N}\sum_{i=1}^{N}\sum_{L,L'} Z_L^i(\mathbf{r})\,\tau_{LL'}^{ij} Z_{L'}^j(\mathbf{r}). \tag{26}$$

The sum is only over site i, because the site j is related to site i by $\mathbf{R}_j = \mathbf{R}_i + \mathbf{R}_n$. It is shown in Eq. (2.43) in FS that the single-site approximation to the scattering-path matrix in this equation is

$$\tau^{ij} = \mathbf{D}^i \tau_c^{ij} \tilde{\mathbf{D}}^j, \tag{27}$$

where τ_c^{ij} is the scattering path matrix for a periodic lattice that has the PCPA scattering matrix $\mathbf{t}_c$ on every site. This matrix is found from the inverse of the matrix $\mathbf{M}_c$ defined in Eq. (12), and depends only on the separation between sites i and j

$$\tau_c^{ij} = \tau_c(\mathbf{R}_{ij}) = [M_c^{-1}]^{ij}. \tag{28}$$

The matrix $\mathbf{D}^i$ is given in Eq. (11). Thus, the non-site-diagonal averaged Green's function is

$$G^{\mathrm{NSD}}(E,\mathbf{r},\mathbf{r}+\mathbf{R}_n) = \frac{1}{N}\sum_{i=1}^{N}\sum_{L_1,L_2}\sum_{L,L'} Z_{L_1}^i(\mathbf{r}) D_{L_1L}^i \tau_{c,LL'}^{ij} \tilde{D}_{L'L_2}^j Z_{L_2}^j(\mathbf{r}). \tag{29}$$

Inserting this in Eq. (25) leads to the expression for the averaged Green's function in the $\mathbf{k}$ representation

$$G(E,\mathbf{k},\mathbf{k}) = \int_\Omega G^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r})\,d\mathbf{r} + \sum_{\substack{n\\ R_n\neq 0}} e^{i\mathbf{k}\cdot\mathbf{R}_n}\sum_{LL'} F_{LL'}(\mathbf{R}_n)\,\tau_{c,LL'}(\mathbf{R}_n), \tag{30}$$

where $G^{\mathrm{SD}}(E,\mathbf{r},\mathbf{r}')$ is defined in Eq. (13). In Eq. (30),

$$F_{LL'}(\mathbf{R}_n) = \frac{1}{N}\sum_{i=1}^{N}\sum_{L_1,L_2}\int_\Omega Z_{L_1}^i(\mathbf{r}) Z_{L_2}^j(\mathbf{r})\,d\mathbf{r}\,D_{L_1L}^i D_{L_2L'}^j, \tag{31}$$

which depends on $\mathbf{R}_n$ because $\mathbf{R}_j = \mathbf{R}_i + \mathbf{R}_n$. The matrix $\mathbf{F}(\mathbf{R}_n)$ has weight one in the sense that it is the sum of N integrals, but divided by N.

Specializing to the isomophous case, for which $Z_L^i(\mathbf{r}) = Z_L^A(\mathbf{r})$ with probability c_A and $Z_L^i(\mathbf{r}) = Z_L^B(\mathbf{r})$ with probability c_B, $F_{LL'}(\mathbf{R}_n)$ becomes

$$\begin{aligned} F_{LL'}^{cc} = {} & c_A^2 \sum_{L_1,L_2}\int_\Omega Z_{L_1}^A(\mathbf{r}) Z_{L_2}^A(\mathbf{r})\,d\mathbf{r}\,D_{L_1L}^A D_{L_2L'}^A \\ & + c_A c_B \sum_{L_1,L_2}\int_\Omega Z_{L_1}^A(\mathbf{r}) Z_{L_2}^B(\mathbf{r})\,d\mathbf{r}\,D_{L_1L}^A D_{L_2L'}^B \\ & + c_B c_A \sum_{L_1,L_2}\int_\Omega Z_{L_1}^B(\mathbf{r}) Z_{L_2}^A(\mathbf{r})\,d\mathbf{r}\,D_{L_1L}^B D_{L_2L'}^A \\ & + c_B^2 \sum_{L_1,L_2}\int_\Omega Z_{L_1}^A(\mathbf{r}) Z_{L_2}^B(\mathbf{r})\,d\mathbf{r}\,D_{L_1L}^A D_{L_2L'}^B, \end{aligned} \tag{32}$$

which does not depend on $\mathbf{R}_n$. The resulting formula for $G(E,\mathbf{k},\mathbf{k})$ is identical to Eq. (4.10) of FS, which means that the Bloch spectral densities will be the same.

It is reassuring that the formulas for $F^{cc}_{LL'}$ and $A^B(E,\mathbf{k})$ obtained with the site-averaging process are identical with the ones obtained with the ensemble-averaging process when the former are applied to an isomorphous model. When ensemble averaging is used, the sum in Eq. (26) is over sites i and j independently. Since the function $F^{cc}_{LL'}$ in Eq. (32) does not depend on $\mathbf{R}_n$, it can be factored out of the integral in Eq. (30). This makes the expression for the Bloch spectral density for the isomorphous case, Eq. (4.15) of FS, considerably easier to deal with computationally.

F. Beyond the PCPA

There are three steps in the derivations given above; the site-averaging process, the single-site approximation, and defining the effective scattering matrix $\mathbf{t}_c$ as the one given by Eq. (23). It should be observed that these steps are independent, and that different levels of theory can be obtained by truncating the derivation after step one or making different choices in step 3. In particular, after a LSMS calculation on a supercell has been carried out, the results can be inserted into Eqs. (8) and (26) to get site-diagonal and non-site-diagonal averaged Green's functions. These can be used directly to obtain such quantities as the Bloch spectral densities without making a single-site or a CPA approximation. The process of averaging reduces the amount of information in the LSMS results, but there are situations in which this is desirable. Experience with order-N calculations like the LSMS has demonstrated that there is frequently too much information in the results. The averaging process may be helpful in resolving the more important physical effects from the mass of computed data. The averaged Green's functions could also provide a good standard by which approximations like the PCPA can be tested.

It should also be noted that the site-averaging process may be applied to supercells that have short-range order or displacements of the atoms from their average lattice sites. It will have to be seen if these capabilities are of any practical use. LSMS calculations on copper-zinc alloys with short-range order have been published,[5] and it would be interesting to see the degree to which these could be reproduced with a PCPA that contains such order. There has been a lot of interest in the inclusion of displacements in alloy theories recently, and this is a proposal for including them in a CPA level theory.

III. COMPUTATIONAL METHODS

The computer codes used for the PCPA calculations described in this paper are based on the ones that were developed for the implementation of the LSMS method.[5,15] The supercell is first generated for the underlying face centered cubic (fcc) or body centered cubic (bcc) Bravais lattice. For bcc alloys, the dimensions of the supercell are normally chosen to be $5\times5\times5$ lattice spacings, and it contains 250 atoms. For fcc alloys, the supercell dimensions are typically $4\times4\times4$, and contain 256 atoms. The next step is to assign atoms to the lattice sites using a random number generator, with the constraint that the alloy model must have the chosen concentration. The computer code can build in short-range order as measured by the Warren-Cowley coefficients, but that capability was not used in the present work.

Ignoring static displacements of the atoms from the sites of the Bravais lattice is an approximation, but there is no reason to expect it to be a serious one for the present case. From the sharpness of the Bragg peaks and the smallness of the static Debye-Waller factors observed in diffraction experiments on metallic alloys, it is clear that, to a first approximation, the atoms fall on the sites of the average lattice. This conclusion is corroborated by the most recent experimental studies.[34] Thermal displacements are also important. At room temperature, the rms average of the thermal displacements is 3–4 times larger than the rms average of the static displacements,[35] even for alloys with a large size mismatch.

Initial guesses are made for the atomic potentials, and the t matrices are calculated. The PCPA equations in Eq. (23) are solved iteratively, using a generalization of the programs that were originally developed for use in KKR-CPA calculations. The new charge densities for each site are found using Eq. (18), and the whole process is repeated until the total energy and potentials have converged. It has been demonstrated computationally[36] that the muffin-tin approximation introduces no significant error in calculations like the ones discussed here. Even for the most extreme case, the copper-palladium alloy system, the heats of mixing calculated with muffin-tin and non-muffin-tin methods are not significantly different. That is, calculations on copper-palladium in the $L1_2$ structure with the muffin-tin LSMS method give -6.6 mRy, while the non-muffin-tin Vienna *ab initio* simulation package[37] (VASP) gives -6.8 mRy. It has been reported in the literature[38] that the non-muffin-tin linear augmented plane-wave method gives -6.3 mRy. The advantage in using the muffin-tin approximation, as compared with an atomic-sphere approximation, is that the multiple-scattering equations are exact. Questions of convergence that must be addressed when full-potential methods are used are avoided in the present approach, although such methods will be incorporated in later calculations when they are deemed necessary. As emphasized earlier, the Madelung contribution is calculated without approximation.

It might be thought that the solution of the equation for the effective scattering matrix $\mathbf{t}_c$ for the N-atom PCPA case, Eq. (23), would be much more difficult than for the isomorphous CPA, Eq. (22). It turns out that, using the standard method described in Ref. 28, only twice as many iterations were needed to solve the equation for the PCPA than for the KKR-CPA. This is due to the fact that, although the charge is different on every atom, the scattering matrices for the atoms of a given species are quite similar. The Brillouin zone integrations required when calculating τ_c^{00} were performed using the prism method[39] with 36 directions in the irreducible wedge of the Brillouin zone, which insures milli-Rydberg accuracy. All the calculations reported here were performed with the Cray T3E-900 512-processor supercomputer operated by the National Energy Research Scientific Computing Center located at the Lawrence Berkeley National Laboratory. It is, of course, always desirable to use the most powerful computing facilities that are available. However, one of

TABLE I. Convergence test of the total energy, Fermi energy, and charge transfer for a 50% copper-zinc alloy.

Supercell size	16	54	128	250
Total energy (Rv)	−3 414.465 156	−3 414.465 272	−3 414.465 129	−3 414.465 064
Fermi energy (Rv)	0.660 9	0.660 9	0.660 9	0.660 9
Charge transfer (e)	0.103 26	0.105 36	0.102 64	0.101 43

the major thrusts of our future program development is to simplify the PCPA calculations so that they can be carried out on small workstations or microcomputers.

Tests were made on the dependence of the PCPA results on N, the number of atoms in the supercell. It was found that surprisingly small samples give quite acceptable predictions for such self-averaging quantities as the total energy per atom and the charge transfer. This is illustrated by the calculations on a bcc 50% CuZn alloy with a lattice constant of 5.5 Bohr radii that are shown in Table I. The total energies calculated with supercells containing 16 and 250 atoms differ by only 92 micro-Rydbergs, and the charge transfer changes by 0.0018 electron charges. All of these supercells used in these calculations were generated randomly. This is, at the same time, a test of the sensitivity of the energy to the arrangement of the atoms in the supercell, since these samples have nothing in common except for their concentration. The Warren-Cowley short-range order parameters, other than the first, are small for these samples, which indicates that they are random. The distributions of Warren-Cowley parameters for some larger supercells are shown in Fig. 15 of Ref. 17. It is expected that unwanted order is more likely to appear in a randomly generated supercell when N is small. Such small supercells will rarely be used because PCPA calculations on large supercells are not significantly more time consuming, since the calculation of Madelung potentials with the Ewald method takes very little time.

IV. ANALYSIS OF CALCULATIONS

In Table II, experimental chemical shifts in copper-palladium, copper-zinc, and silver-palladium alloys are compared with the results of calculations using the LSMS, PCPA, and KKR-CPA. The LSMS calculations appeared in a previous publication.[29] The charge transfers predicted by the three theories are also shown. In the following, the reasons for focusing on chemical shifts rather than the densities of states that can be measured with ultraviolet photoemission spectroscopy, as was conventional in earlier evaluations of alloy theories, will be made clear.

The chemical shifts are defined as follows. In the DFT-LDA, the binding energy of a core level $E_{n,l}^{A,i}(c)$ of an A atom on site i in an A-B alloy with a concentration c is the negative of the one-electron core-level energy measured relative to the Fermi energy. The binding energy for the A atoms $E_{n,l}^{A}(c)$ is the average of the $E_{n,l}^{A,i}(c)$ over all the A sites. The binding energy for the B atoms $E_{n,l}^{B}(c)$ is defined similarly. The chemical shifts for atoms in an alloy, $\delta E_{n,l}^{A}(c)$ or $\delta E_{n,l}^{B}(c)$, are the change in the binding energy relative to the pure A or B metal. It is well understood that there are significant differences between the Koopman approximation to the binding energy and the measured binding energy of a core electron, due to the relaxation of the electrons around the core hole.[40] It has been proposed that the many-body relaxation corrections are approximately independent of concentration,[41] so they are not expected to be very important in measurements of chemical shifts.

The LSMS and PCPA calculations for the CuPd alloys shown in Table II use supercells containing 256 atoms based on fcc Bravais lattices, the lattice constant for the 50% alloy being 6.9 atomic units (au) and for the 80% alloy being 7.1 a.u. The disordered 50% CuZn alloy is in the β phase (bcc) with a lattice constant of 5.5 a.u. The supercell used with the LSMS calculations contains 432 atoms, while the one for the PCPA calculations contains 256. The ordered β' phase of

TABLE II. Experimental and theoretical values for the chemical shifts in various alloy systems. Net atomic charges calculated with the indicated theories.

Alloy	a alloy Bohr radii	atom level	Exp shift (eV)	LSMS shift (eV)	PCPA shift (eV)	KKR-CPA shift (eV)	LSMS charge ($\|e\|$)	PCPA charge ($\|e\|$)	KKR-CPA charge ($\|e\|$)
50% CuPd	7.1000	$2p$ Cu	−0.700 00	−0.716 88	−0.856 03	−0.931 02	−0.176 09	−0.160 84	−0.043 34
		$3d$ Pd	0.260 00	0.337 97	0.232 02	−0.006 29	0.176 09	0.160 84	0.043 84
80% CuPd	6.9000	$2p$ Cu	−0.250 00	−0.268 17	−0.295 92	−0.323 44	−0.077 21	−0.074 53	−0.021 62
		$3d$ Pd	0.700 00	0.654 57	0.606 68	0.436 36	0.310 87	0.299 59	0.086 48
50% CuZn	5.5000	$2p$ Cu		0.339 58	0.225 99	0.189 42	−0.100 86	−0.101 43	−0.075 78
		$2p$ Zn		−0.008 71	−0.133 88	0.046 90	0.100 86	0.101 43	0.075 78
CuZn $B2$	5.5000	$2p$ Cu	0.350 00	0.568 69			−0.125 89		
		$2p$ Zn	−0.200 00	−0.285 57			0.125 89		
50% AgPd	7.6100	$3d$ Ag	−0.500 00	−0.477 15	−0.573 89	−0.535 90	0.049 56	0.048 63	0.057 24
		$3d$ Pd	0.000 00	−0.101 63	−0.193 47	−0.376 86	−0.049 56	−0.048 63	−0.057 24

CuZn is in the $B2$ (CsCl) structure with the same lattice constant as the β phase. The 50% AgPd alloy is modeled with a supercell containing 256 atoms and a fcc Bravais lattice with a lattice constant of 7.61 a.u. in both the LSMS and PCPA calculations. The same lattice constants are used in the KKR-CPA calculations on all these alloy systems, but, of course, supercells are not used in such calculations. The pure metals Cu, Pd, Ag, and Zn are all taken to be fcc with lattice constants 6.7677, 7.43, 7.2744, and 7.79 a.u. The experimental chemical shifts for CuPd are taken from Ref. 42, those for CuZn are from Ref. 43, and those for AgPd are from Ref. 44. The choice of core levels listed in Table II is dictated by the existence of well-defined peaks in the energy range that can be sampled using Al $K\alpha$ or Mg $K\alpha$ x rays.

The first observation that can be made from Table II is that the chemical shifts calculated with the LSMS agree with the experimental data to within the accuracy of the experiments, as pointed out in Ref. 29. It can be concluded from this agreement that the prediction of chemical shifts in alloys is trivial in the sense that an accurate DFT-LDA calculation will yield them. Overall, the chemical shifts predicted by the PCPA calculations are very good, particularly when it is taken into account that the computational effort expended in obtaining them is a very small fraction of that required for the LSMS numbers. The effort required to calculate chemical shifts with the KKR-CPA is even less, but the agreement with experiment is further reduced.

The agreement between the LSMS calculations and experiment for such a sensitive property as the chemical shift increases the confidence that can be placed in that theory. For this reason, it is disturbing that the KKR-CPA predictions for the average charges on the constituent atoms, shown in Table II, differ considerably from the predictions of the LSMS. This is particularly noticeable for the CuPd alloys. If one takes the conventional view that chemical shifts are a measure of charge transfer, it is surprising that the chemical shifts given by the KKR-CPA are as accurate as they are. The PCPA gives values for the average charges that agree with the LSMS to within a few percent. This is expected, because the PCPA includes Coulomb effects as well as they can be within the level of a single-site approximation.

In Fig. 1, calculations of the density of states for the 50% CuPd alloy are shown. The LSMS curve is obtained by averaging the densities of states calculated for the 256 sites in the supercell. The average density of states given by the KKR-CPA is clearly different from the LSMS prediction, but it would be difficult to see differences of this magnitude in experiments. One might have expected that the difference would have been greater, given the fact that the KKR-CPA prediction of the charge transfer is only one fourth of the LSMS value. There are 10 electrons in the d bands, and the charge transfer is only 2% of that. Thus, the difference between the positions of the Fermi energies does not appear to be very large on the scale of the drawing. This explains the early successes that the KKR-CPA had in predicting the results of photoemission spectroscopy experiments, and also the reason that charge transfer and chemical shifts provide a more sensitive test of the ability of an alloy theory to treat Coulomb effects correctly.

An interesting observation that was made on the basis of data from LSMS calculations[16] is that the DFT-LDA predicts

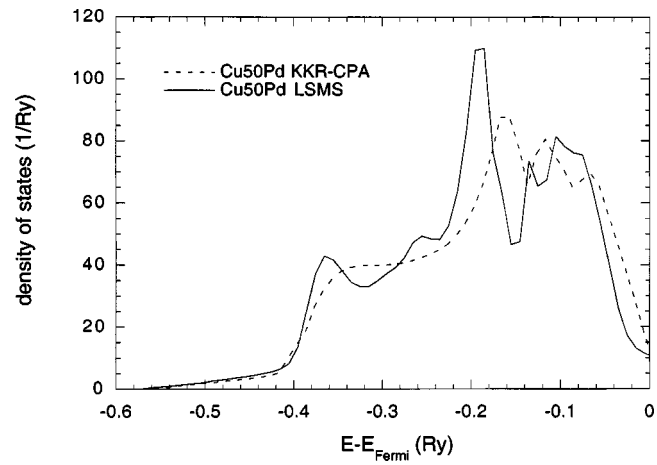


FIG. 1. The density of states of a 50% copper-palladium alloy as a function of the energy relative to the Fermi energy. The lattice constant of the fcc Bravais lattice is 6.9 atomic units. The solid line shows the average of the 256 densities of states calculated with the LSMS. The dotted line shows the density of states calculated with the standard KKR-CPA.

a linear relation between the Madelung potentials at the sites of an alloy V^i and the net charges on the sites q^i. This qV relation is not trivial because it is only true for the charges calculated in the final stage of a self-consistent calculation, while such simple conditions as charge neutrality hold at every iteration. It can be seen from Fig. 2 that the V^i and q^i from a PCPA calculation on the 50% CuPd alloy fall on lines that are as straight as the ones determined by the LSMS data. The slopes are not identical. For copper sites, the slope from the PCPA data is -1.3738 Ry/$|e|$, compared with -1.1955 from the LSMS data. For the Pd sites, the slopes are -1.2767 and -1.1826. Thus, the PCPA not only gives an accurate value for the average net charge on the atomic sites, as can be seen from Table II, but also it gives a very good description of the distribution of the charges on the sites.

Since the computer codes used for the present calculations are based on the LSMS codes, it is not difficult to extend

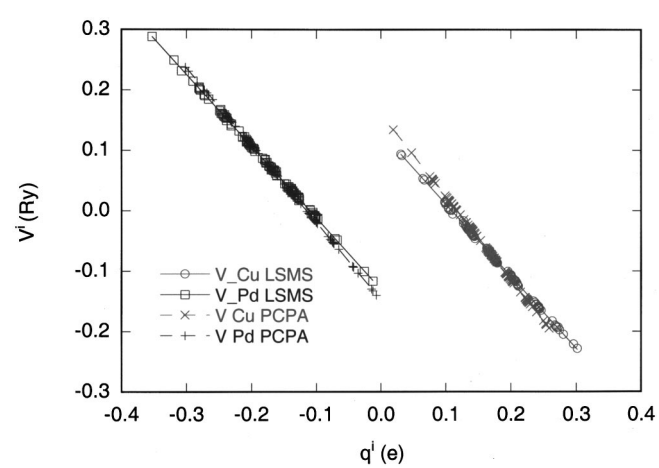


FIG. 2. The Coulomb potential at the lattice sites as a function of the net charge on the lattice sites for a 50% copper-palladium alloy. The lattice constant of the fcc Bravais lattice is 6.9 atomic units. The straight lines are the result of fits to the data. The sign convention for the charges is opposite to the one used in Table II.

them to treat LIZ's larger than one. Such LSGF calculations on the 50% CuPd alloy have been carried out with a LIZ containing 19 atoms or two nearest-neighbor shells. These calculations differ from the ones in Ref. 19 in that the multiple-scattering method is used. The improvement of the LSGF charges over the PCPA charges is not as great as might have been expected. The average net charge on the Cu atoms predicted by the calculations using $N_{\mathrm{LIZ}}=19$ is $-0.161\,53\ |e|$, only 0.4% larger than the one from the PCPA calculations. The qV relations for this alloy obtained from the LSMS and PCPA are shown in Fig. 2. The calculation using $N_{\mathrm{LIZ}}=19$ gives straight lines like the ones in the figure. The slope of the line for Cu is -1.3161 Ry/$|e|$, and for Pd it is -1.2420 Ry/$|e|$. These values are 3–4 % smaller than the ones predicted by the PCPA, and are about a third of the correction that is needed to obtain values that agree with the LSMS slopes.

V. DISCUSSION

It can be concluded that a CPA level theory that treats Coulomb effects without significant approximation exists. There are several advantages in the PCPA approach. The first is that it predicts a continuous distribution of charges about the mean for the A and B atoms, and this is in accordance with reality. The second is that there are no adjustable parameters in the theory, and this is an improvement over the isomorphous alloy theories that attempt to include nonzero Madelung potentials. The third advantage is that the inclusion of the exact form for the Madelung potentials is conceptually simpler than the approximations that have been proposed, and requires only slightly more computational effort. We emphasize that the changes to the predictions of the electronic structure obtained from other self-consistent CPA theories are quantitative but not qualitative. Those methods have had great success in explaining many properties of alloys. In hindsight, it is clear that the derivation of the Madelung potential in the KKR-CPA (Ref. 10) is logically and mathematically sound. The reason it led to a null result was the belief, generally held at that time, that the isomorphous alloy model is a necessary precondition for such derivations.

Site averaging with only the CPA conditions leads to the polymorphous PCPA. The addition of the condition that the potentials and wave functions correspond to the isomorphous model leads to the isomorphous KKR-CPA, in which the Madelung potential is zero. The isomorphous S-CPA and SIM-CPA do not fall within the algebraic sequence, although they are useful approximations. Of course, the order-N LSMS and LSGF theories give more accurate descriptions of the electronic states than any CPA, but many theoretical studies do not require so much detail.

The polymorphous nature of the PCPA leads to some interesting consequences, not all of which are explored in this paper. Not only does the PCPA give a more accurate description of the average charge transfer, it also includes fluctuations about the average. These charge fluctuations modify the density of states and the Bloch spectral density function. In particular the effects on the Bloch spectral function have the potential to change the width of peaks and, therefore, to alter electron lifetimes. This will affect calculations of such properties as the residual resistivity[45] and ordering temperatures for Fermi surface driven long period ordered structures.[46] A major problem with the isomorphous KKR-CPA is that there is no clear pathway to make the theory full potential because of difficulties with continuity of the charge density at the cell boundaries, even if one allows nonspherical corrections to the scattering. In the PCPA the fact that the Coulomb effects are treated using a supercell means that charge density is well defined throughout all space. Unfortunately, it will not be precisely continuous at the cell boundary since the multiple scattering equations are still solved in a single site approximation. How severe a restriction this imposes remains to be investigated numerically.

The development of the PCPA is only beginning, and there are many extensions that need to be carried out. A comparison of the binding energies of alloys calculated with the LSMS and KKR-CPA is given in Ref. 18. The PCPA binding energies should be added to this comparison. The formulas for $G(E,\mathbf{k},\mathbf{k})$ and hence the Bloch spectral density function $A^B(E,\mathbf{k})$ derived in Sec. II E should be utilized for calculations on some real alloys as soon as possible. The LSGF with LIZ's containing more than one atom also leads to an effective scattering matrix, $\mathbf{t}_c$. We do not consider that to be a PCPA because it is more difficult to fit it into the single-site picture, and it presupposes that one is prepared to do more time-consuming calculations. It would be interesting to compare the average Green's functions obtained using the LSGF and PCPA effective scattering matrices. The existing programs, which are based on the LSMS codes, should be simplified so that PCPA calculations can be carried out with readily available microcomputers and workstations.

ACKNOWLEDGMENTS

Work supported by Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, and Office of Computational and Technology Research, Mathematical Information and Computational Sciences Division, U. S. Department of Energy, under Subcontract No. DEAC05-960R22464 with Lockheed-Martin Energy Research Corporation and DEAC03-76SF00098. B.U. and N.Y.M. were supported by an appointment to the Oak Ridge National Laboratory Postdoctoral Research Associates Program administered jointly by the Oak Ridge National Laboratory and the Oak Ridge Institute for Science and Education.

[1] Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[2] L. Kaufman and H. Bernstein, *Computer Calculation of Phase Diagrams* (Academic, New York, 1970); M. Hillert and M. Schalin, J. Phase Equilib. **19**, 206 (1998).

[3] J. W. D. Conolly and A. R. Williams, Phys. Rev. B **27**, 5169 (1983); J. M. Sanchez, F. Ducastelle, and D. Gratias, Physica (Amsterdam) **128A**, 334 (1984); A. Zunger, in *Statics and Dynamics of Alloy Phase Transitions*, edited by P. E. A. Turchi and A. Gonis, Vol. 319 of *NATO Advanced Study Institute, Series B: Physics* (Kluwer, Dordrecht, 1993).

[4] R. Car and M. Parinello, Phys. Rev. Lett. **55**, 2471 (1985); M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992); V. Godlevsky, J. R. Chelikowsky, and N. Trouillier, Phys. Rev. B **52**, 13 281 (1995); J. R. Chelikowsky, X. Jing, K. Wu, and Y. Saad, *ibid.* **53**, 12 071 (1996).

[5] Yang Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, and W. M. Temmerman, Phys. Rev. Lett. **75**, 2867 (1995).

[6] J. S. Faulkner, in *Progress in Materials Science*, edited by J. W. Christian, P. Haasen, and T. B. Massalski (Pergamon, Oxford, England, 1982), Vol. 27.

[7] P. Soven, Phys. Rev. **156**, 809 (1967).

[8] J. Korringa, Physica (Amsterdam) **13**, 392 (1947); W. Kohn and N. Rostoker, Phys. Rev. **94**, 111 (1954).

[9] G. M. Stocks, W. M. Temmerman, and B. L. Gyorffy, Phys. Rev. Lett. **41**, 339 (1978).

[10] H. Winter and G. M. Stocks, Phys. Rev. B **27**, 882 (1983); D. D. Johnson, D. M. Nicholson, F. J. Pinski, B. L. Gyorffy, and G. M. Stocks, Phys. Rev. Lett. **56**, 2088 (1986); Phys. Rev. B **41**, 9701 (1990).

[11] R. Magri, S. H. Wei, and A. Zunger, Phys. Rev. B **42**, 11 388 (1990).

[12] I. A. Abrikosov, Yu. H. Vekilov, A. V. Ruban, Phys. Lett. A **154**, 407 (1991); I. A. Abrikosov, Yu. H. Vekilov, P. A. Korzhavyi, A. V. Ruban, and L. E. Shilkrot, Solid State Commun. **83**, 867 (1992).

[13] D. D. Johnson and F. J. Pinski, Phys. Rev. B **48**, 11 553 (1993).

[14] N. Papanikolao, R. Zeller, P. H. Dederichs, and N. Stefanou, Phys. Rev. B **55**, 4157 (1997).

[15] G. M. Stocks, D. M. C. Nicholson, Y. Wang, W. A. Shelton, Z. Szotek, and W. M. Temmermann, in *High Performance Computing Symposium; Grand Challenges in Computer Simulation, Proceedings of the 1994 Simulation Multiconference*, edited by A. M. Tentner (The Society for Computer Simulation, San Diego, 1994); D. M. C. Nicholson, G. M. Stocks, Y. Wang, W. A. Shelton, Z. Szotek, and W. M. Temmermann, Phys. Rev. B **50**, 14 686 (1994).

[16] J. S. Faulkner, Yang Wang, and G. M. Stocks, Phys. Rev. B **52**, 17 106 (1995); J. S. Faulkner, Yang Wang, and G. M. Stocks, in *Stability of Materials*, Vol. 355 of *NATO Advanced Study Institute, Series B: Physics*, edited by A. Gonis, P. E. A. Turchi, and J. Kudrnovsky (Plenum, New York, 1996); J. S. Faulkner, Yang Wang, Nassrin Moghadam, and G. M. Stocks, in *Properties of Complex Inorganic Solids*, edited by A. Gonis, A. Meike, and P. E. A. Turchi (Plenum, New York, 1997); J. S. Faulkner, Yang Wang, and G. M. Stocks, J. Phase Equilib. **18**, 499 (1997).

[17] J. S. Faulkner, Yang Wang, and G. M. Stocks, Phys. Rev. B **55**, 7492 (1997).

[18] J. S. Faulkner, Nassrin Moghadam, Yang Wang, and G. M. Stocks, Phys. Rev. B **57**, 7653 (1998).

[19] I. A. Abrikosov and B. Johansson, Phys. Rev. B **57**, 14 164 (1998).

[20] P. A. Korzhavyi, A. V. Ruban, I. A. Abrikosov, and H. L. Skriver, Phys. Rev. B **51**, 5773 (1995).

[21] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[22] P. P. Ewald, Ann. Phys. (Leipzig) **64**, 253 (1921).

[23] I. A. Abrikosov, A. M. N. Niklasson, S. I. Simak, B. Johansson, A. V. Ruban, and H. L. Skriver, Phys. Rev. Lett. **76**, 4203 (1996).

[24] I. A. Abrikosov, S. I. Simak, B. Johansson, A. V. Ruban, and H. L. Skriver, Phys. Rev. B **56**, 9319 (1997).

[25] D. M. C. Nicholson, G. M. Stocks, W. A. Shelton, Y. Wang, and J. C. Swihart, Metall. Mater. Trans. A **29**, 1845 (1998).

[26] A. B. Oparin, D. M. C. Nicholson, X.-G. Zhang, W. H. Butler, W. A. Shelton, G. M. Stocks, and Yang Wang, J. Appl. Phys. **85**, 4548 (1999); W. H. Butler, X.-G. Zhang, T. C. Schulthess, D. M. C. Nicholson, A. B. Oparin, J. M. MacLaren, *ibid.* **85**, 5834 (1999).

[27] G. M. Stocks, B. Ujfalussy, X. Wang, D. M. C. Nicholson, W. A. Shelton, Y. Wang, and B. L. Gyorffy, Philos. Mag. B **78**, 665 (1998).

[28] J. S. Faulkner and G. M. Stocks, Phys. Rev. B **21**, 3222 (1980).

[29] J. S. Faulkner, Yang Wang, and G. M. Stocks, Phys. Rev. Lett. **81**, 1905 (1998).

[30] I. M. Lifschitz, Adv. Phys. **13**, 483 (1964).

[31] The following equations were first put forward in J. S. Faulkner, B. Ujfalussy, Nassrin Y. Moghadam, G. M. Stocks, and Yang Wang, Bull. Am. Phys. Soc. **44**, 1035 (1999).

[32] B. L. Gyorffy, Phys. Rev. B **5**, 2382 (1972).

[33] P. Lloyd, Proc. Phys. Soc. London **90**, 207 (1967).

[34] X. Jiang, G. E. Ice, C. J. Sparks, L. Robertson, and P. Zschack, Phys. Rev. B **54**, 3211 (1996); J. S. Faulkner, *ibid.* **56**, 2299 (1997).

[35] F. H. Herbstein, B. S. Borie, Jr., and B. L. Averbach, Acta Crystallogr. **9**, 466 (1956).

[36] S. G. Louie, K. Ho, and M. L. Cohen, Phys. Rev. B **19**, 1774 (1979); J. Neve, B. Sundqvist, and O. Rapp, *ibid.* **28**, 629 (1983); A. R. Jani, N. E. Brenner, and J. Callaway, *ibid.* **38**, 9425 (1988); D. M. Nicholson and J. S. Faulkner, *ibid.* **39**, 8187 (1989).

[37] G. Kresse and J. Hafner, Phys. Rev. B **47**, R558 (1993); G. Kresse and J. Furthmueller, *ibid.* **54**, 11 169 (1996).

[38] Z. W. Lu, D. B. Laks, S.-H. Wei, and A. Zunger, Phys. Rev. B **50**, 6642 (1994).

[39] G. M. Stocks, W. M. Temmerman, and B. L. Gyorffy, in *Electrons in Disordered Metals and at Metallic Surfaces*, Vol. 42 of *NATO Advanced Study Institute, Series B: Physics*, edited by P. Phariseau, B. L. Gyorffy and L. Scheire (Kluwer, Dordrecht, 1979), p. 193.

[40] A. R. Williams and N. D. Lang, Phys. Rev. Lett. **40**, 954 (1978); Phys. Rev. B **16**, 2408 (1978).

[41] M. Weinert and R. E. Watson, Phys. Rev. B **51**, 17 168 (1995).

[42] R. J. Cole, N. J. Brooks, and P. Weightman, Phys. Rev. B **56**, 12 178 (1997).

[43] G. K. Wertheim, M. Campagna, and S. Hufner, J. Phys.: Condens. Matter **18**, 133 (1974).

[44] P. Steiner and S. Hufner, Acta Metall. **29**, 1885 (1981).

[45] G. M. Stocks, and W. H. Butler, Phys. Rev. Lett. **48**, 55 (1982); W. H. Butler, and G. M. Stocks, Phys. Rev. B **29**, 4217 (1984).
[46] B. L. Gyorffy and G. M. Stocks, Phys. Rev. Lett. **50**, 374 (1983); J. Wadsworth, B. L. Gyorffy, and G. M. Stocks, in *High Temperature Alloys: Theory and Design*, edited by J. O. Stiegler (Metallurgical Society AIME, Warrendale, PA, 1984); J. B. Staunton, D. D. Johnson, and F. J. Pinski, Phys. Rev. B **50**, 1450 (1994).

ELSEVIER

Intermetallics 8 (2000) 937–943

Intermetallics

www.elsevier.com/locate/intermet

Theoretical calculations and experimental measurements of the structure of Ti_5Si_3 with interstitial additions

J.J. Williams [a,b], Y.Y. Ye [c], M.J. Kramer [a,b,*], K.M. Ho [a], L. Hong [c], C.L. Fu [d], S.K. Malik [e,f]

[a] *Ames Laboratory, US Department of Energy, Ames, IA 50011, USA*
[b] *Department of Materials Science and Engineering, Iowa State University, Ames, IA 50011, USA*
[c] *Center of Analysis and Testing, Wuhan University, Wuhan, China 430072*
[d] *Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[e] *Tata Institute of Fundamental Research, Bombay India 400005*
[f] *University of Missouri Research Reactor, Columbia, MO 65211, USA*

Received 15 December 1999; accepted 28 April 2000

Abstract

The equilibrium structural parameters, enthalpies of formation and partial densities of state for Ti_5Si_3 and $Ti_5Si_3Z_{0.5}$ (Z = B, C, N or O) were calculated based on first-principle techniques. Enthalpy of formation calculations suggest that of the known structures for transition metal (TM) silicide compounds containing TM_5Si_3 ($D8_8$, $D8_l$ and $D8_m$) the $D8_8$ structure is the most stable form of Ti_5Si_3, and the stability of the structure increases as Z atoms are added. The theoretically determined structural trends as a function of interstitial element, Z, agreed well with experimentally determined values. Both indicate bonding between Ti and Z atoms based on contraction of Ti–Z separations. The calculated partial densities of state suggest that p(Si)–d(Ti) and d(Ti)-d(Ti) interactions are responsible for most of the bonding in pure Ti_5Si_3, which agrees with previous studies. As Z atoms are added, p(Z)–d(Ti) interactions become significant at the expense of weakening some of the d(Ti)–d(Ti) interactions. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Titanium silicides; B. Crystal chemistry of intermetallics; E. Electronic structure, calculation; F. Diffraction (electron, neutron and X-ray)

1. Introduction

Interest in M_5Si_3 intermetallics (M = transition metal of groups III–VI) which began in the 1950′s and continues today is primarily a result of their high melting points (> 2000°C), wide homogeneity ranges and large alloying potentials. Nowotny and coworkers performed much of the early characterization of these materials [1–3]. The crystal structures were determined to be either a hexagonal Mn_5Si_3-type ($D8_8$, M = Sc, Ti, Mn, Y), tetragonal Cr_5B_3-type ($D8_l$, M = Cr, Nb, La, Ta) or tetragonal W_5Si_3-type ($D8_m$, M = V, Mo, W). Additionally, Nowotny discovered that all of these M_5Si_3 compounds, except La_5Si_3, reverted to the hexagonal form in the presence of boron, carbon, nitrogen, or oxygen [1]. M_5X_3 compounds stabilized in the hexagonal structure by ternary additions have since become known as Nowotny phases [2]. Experimental observations also suggested that carbon was most efficient and oxygen least efficient in stabilizing the hexagonal structure. However, no direct experimental or theoretical evidence currently exists which explains why this stabilization occurs.

Although experimental evidence suggests that Ti_5Si_3 does not require ternary additions to form the hexagonal structure, recent work has shown that small additions of carbon have a significant effect on the crystal structure, thermal expansion and high temperature oxidation resistance [4,5]. In fact, by adding carbon to the structure, Ti_5Si_3 becomes a considerably more promising material for engineering applications [6,7] by reducing the anisotropy of thermal expansion and increasing oxidation resistance. However, little is known why carbon additions have such a striking effect on these properties. The goal of this study is to combine experimentally determined structural data with first-principle electronic calculations to understand bonding changes that occur with the addition of boron, carbon, nitrogen or oxygen to Ti_5Si_3. Determining these bonding changes will aid in

* Corresponding author. Tel.: +1-515-294-0276.

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(00)00064-9

understanding and predicting the changes that occur in the thermal and electronic properties.

No previous first-principle calculations have been attempted on ternary $Ti_5Si_3Z_x$ (Z = B,C,N or O); however, two studies do exist on binary Ti_5Si_3. The first study, by Long and Chong [8], used a semi-empirical tight binding energy band method with the extended Hückel approximation to calculate band structure and density of states (DOS). They concluded that bonding–antibonding in Ti_5Si_3 is primarily a result of d(Ti)–p(Si) interaction above and below the Fermi energy (E_f) as well as d(Ti)–d(Ti) interactions spanning energies around E_f. These orbital interactions are very typical of transition metal silicides in general and explain their good electrical conductivity [9,10]. The second study, by Ekman and Ozolins [11], made calculations based on the full potential version of the linear muffin tin orbital (LMTO) method. Their conclusions of d(Ti)–p(Si) and d(Ti)–d(Ti) hybridization were identical to those above. However, their electron density maps suggest that all Ti–Si interactions in Ti_5Si_3 are multi-centered bonds as opposed to simple two-atom covalent bonds. The study also calculated the equilibrium volume, bulk modulus and enthalpy of formation, all of which were only slightly lower than the experimental values.

This study used the LMTO method with the atomic sphere approximation (ASA) to calculate the angular-momentum decomposed electronic DOS, but unlike the previous studies, equilibrium lattice parameters and atomic positions were also calculated. These calculations were made for Ti_5Si_3, $Ti_5Si_3Z_{0.25}$ and $Ti_5Si_3Z_{0.5}$ (Z = B,C,N or O) and compared to experimentally determined values. Heats of formation were also calculated for most compositions including Ti_5Si_3 in the equilibrium Mn_5Si_3 structure ($D8_8$) as well as in the possible alternate structures of W_5Si_3 ($D8_m$) and Cr_5B_3 ($D8_l$).

2. Crystal structure of Ti_5Si_3

Fig. 1 gives the hexagonal structure of Ti_5Si_3. The unit cell contains two distinct titanium sites and one silicon site: Ti at 4d sites at (1/3, 2/3, 0); Ti at 6g sites at (x_{Ti}, 0, 1/4); Si at 6g sites at (x_{Si}, 0, 1/4).

Theoretical calculations were based on this unit cell, which contains two formula units of atoms (i.e. $Ti_{10}Si_6$). The Ti^{4d} atoms form a linear chain parallel to the c-axis, and the Ti^{6g} atoms form a chain of face-shared trigonal antiprisms along the c-axis. The silicon atoms form a chain of distorted face-shared trigonal antiprisms parallel to the c-axis such that one Ti^{4d} site is at the center of each antiprism. This structure leads to an ABAC stacking sequence along the c-direction. The B and C planes consist of Ti^{6g} and Si atoms, which form the shared faces of the antiprsims, and the A planes consist solely of Ti^{4d} atoms. The Z atoms are thought to occupy

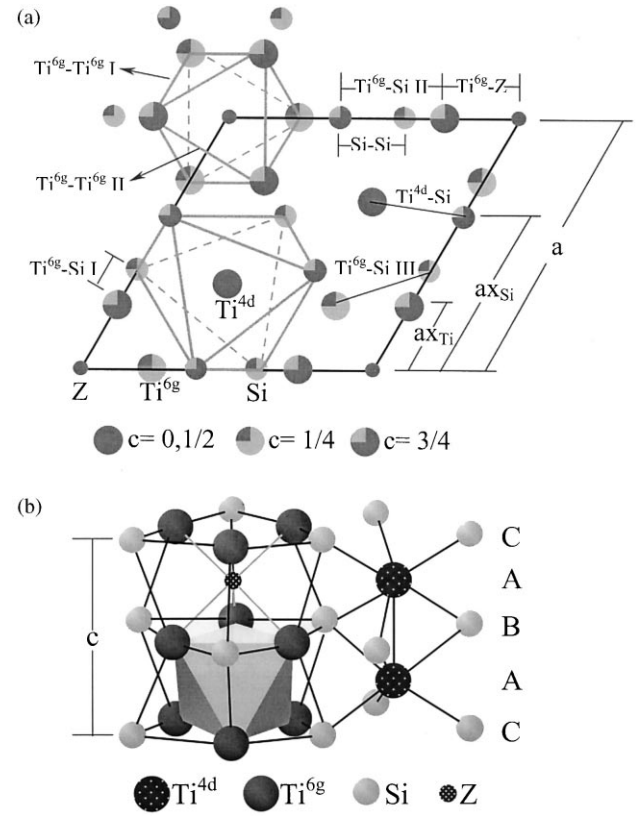


Fig. 1. $D8_8$ crystal structure of Ti_5Si_3: (a) 001 orthographic projection of lattice with highlighted trigonal antiprisms; (b) depiction of the face-sharing of the trigonal antiprisms along the c-axis. Z atoms sit at the center of the trigonal antiprisms formed by six surrounding Ti^{6g} atoms.

the interstitial region at the center of the antiprism formed by the Ti^{6g} atoms, and hence, would also lie in the A plane with the Ti^{4d} atoms.

Strong experimental evidence exists to support the assertion that Z atoms occupy this interstitial site. Neutron diffraction studies of Mo_5Si_3C by Parthe et al. [3] and of $Ti_5Si_3H_x$ by Kajitani et al. [12] as well as single crystal x-ray diffraction studies of $La_5Ge_3O_x$ by Guloy and Corbett [13] and of $Er_5Si_3C_x$ by Al-Shahery et al. [14] all agree that occupation of this antiprismatic interstice by Z atoms is most probable. Thus, the solubility of Z in Ti_5Si_3 should vary from zero to the stoichiometric limit of $Ti_5Si_3Z_1$ (there are only two antiprismatic interstices in the unit cell, which is equivalent to a maximum of one Z atom per formula unit). This is in full agreement with the actual solubility limits of C,N or O in Ti_5Si_3 as measured by diffusion couple experiments [15,16].

3. Theoretical approach

Calculations to determine the stable structure were made within the local density approximation [17] by using the Hedin–Lundqvist [18] form for the local exchange and correlation potential. Electronic Bloch

states [19] were expanded as a mixed basis set with norm conserving scalar-relativistic pseudopotentials [20] used for the constituent elements. Relaxation of atomic coordinates was facilitated by computing the Hellmann–Feynman forces [21] acting on the atoms. A Broyden algorithm for estimating and updating this force matrix was used to predict the new atomic coordinates during the relaxation process. The atomic positions and lattice parameters were fully relaxed to the true equilibrium structure. Using these calculated equilibrium lattice parameters, electronic densities of state (DOS) were calculated by the LMTO-ASA method. To improve efficiency and accuracy of the calculations, unoccupied octahedral interstices were filled with empty spheres. For Ti_5Si_3, the radius of these spheres was set to 0.7 times the radius of the Ti atom. For $Ti_5Si_3Z_x$, the radius was set to the radius of the Z atom, which ranged from 0.68 to 0.7 times the radius of the Ti atom. Heats of formation were also calculated (see Section 5.1).

4. Experimental approach

All $Ti_5Si_3Z_x$ samples were synthesized by arc-melting. The starting materials included sponge titanium (Timet, 99.7%), silicon pieces (Alfa Æsar, 99.9999%), spectrographic grade graphite electrodes for carbon, boron pieces (Alfa Æsars, 99.5%), titanium nitride for nitrogen (Johnson Matthey, 99.8%) and titanium dioxide for oxygen (Fischer Scientific, 99.8%). Arc melting was done in a ultra-high-purity argon atmosphere on a water-chilled copper hearth. Samples were melted at least three times via a non-consumable tungsten electrode. Total weight losses after arc-melting were typically much less than 0.5 wt.%.

Arc-melted samples were then ground to <20 μm and mixed with a silicon line position standard (NIST SRM 640b). Room temperature X-ray diffraction spectra were obtained from a Scintag diffractometer with solid state detector. Room temperature neutron diffraction spectra were obtained from the Missouri University Research Reactor (MURR) using a curved Ge monochromator and position sensitive detector. Rietveld analysis software (GSAS, Los Alamos National Laboratory, 1985) was used to calculate the lattice parameters and the two variable atomic coordinates, x_{Ti} and x_{Si}. Oxygen and nitrogen content were measured on a Leco TC-436 analyzer; carbon content was measured on a Horiba EMIA-520 analyzer.

5. Results and discussion

5.1. Enthalpy of formation

The enthalpies of formation were calculated from total energies, E, according to: $H + E_{Ti_5Si_3Z_x} - \Sigma_i(X_iE_i)$, where X_i is the concentration of the ith elemental component. The total energies of the elements, E_i, were calculated using their most stable structures: titanium ($P6_3/mmc$), silicon (Fd-3m), boron (R-3m) and graphite ($P6_3/mmc$), as well as O_2 and N_2 gas. For $Ti_5Si_3Z_x$, total energies, $E_{Ti_5Si_3Z_x}$, were calculated using the relaxed atomic positions and lattice parameters that were determined by the pseudopotential method described in Section 3.

Table 1 lists the results of these calculations. Of the three possible crystal structures in which M_5Si_3 compounds form, the $D8_8$ structure has the largest negative value. Thus, these calculations agree with experimental observations that suggest that the $D8_8$ structure is the most stable structure for Ti_5Si_3. In most other M_5Si_3 compounds, where M is heavier than Ti, the $D8_8$ structure only becomes favorable in the presence of interstitial atoms. As an example, Mo_5Si_3, which exists in the $D8_m$ structure, converts to the $D8_8$ structure when carbon is added. In a study by Fu et al. [22], calculations of enthalpies of formations did correctly suggest that in agreement with experimental evidence, Mo_5Si_3 should form in the $D8_m$ structure instead of the $D8_8$ and $D8_l$ structures. Thus, this study provides further support that theoretical calculations of the enthalpy of formation can be used to predict which crystal structure is most stable.

As seen in Table 1, the enthalpy of formation becomes more negative, as more carbon or boron is added to the lattice. This suggests that these interstitial atoms, in accordance with experimental observations, do increase the stability of the $D8_8$ structure. A comparison of the enthalpy of formation of Ti_5Si_3 with that of $Ti_5Si_3Z_{0.5}$ indicate $Ti_5Si_3C_{0.5}$ is 3% more negative, $Ti_5Si_3N_{0.5}$ is 7% more negative and $Ti_5Si_3O_{0.5}$ is 38% more negative than Ti_5Si_3. This trend agrees well with experimental values of the Gibbs energy of formation at 1100°C for Ti_5Si_3Z. As reported in Goldstein et al. [16], Ti_5Si_3C is 5% more negative, Ti_5Si_3N is 11% more negative and Ti_5Si_3O is 41% more negative than Ti_5Si_3.

5.2. Equilibrium structural parameters

Theoretical and experimental structural parameters are given, respectively, in Tables 2 and 3. The theoretical

Table 1
Calculated enthalpies of formation

Composition	Structure	Enthalpy, eV/f.u.
Ti_5Si_3	$D8_l$	−6.060
Ti_5Si_3	$D8_m$	−6.170
Ti_5Si_3	$D8_8$	−6.410
$Ti_5Si_3B_{0.25}$	$D8_8$	−6.604
$Ti_5Si_3B_{0.5}$	$D8_8$	−7.104
$Ti_5Si_3C_{0.25}$	$D8_8$	−6.625
$Ti_5Si_3C_{0.5}$	$D8_8$	−7.251
$Ti_5Si_3N_{0.5}$	$D8_8$	−6.866
$Ti_5Si_3O_{0.5}$	$D8_8$	−8.870

Table 2
Calculated structural parameters for $Ti_5Si_3Z_x$[a]

Z_X	a (Å)	c (Å)	x_{Ti}[a]	x_{Si}[a]
–	7.377	5.084	0.2473	0.6063
$B_{0.25}$	7.4027	5.1079		
$B_{0.5}$	7.4101	5.1130	0.2470	0.6023
$C_{0.25}$	7.3775	5.0892		
$C_{0.5}$	7.3925	5.0956	0.2400	0.6023
$N_{0.25}$				
$N_{0.5}$	7.3834	5.0746	0.2410	0.6025
$O_{0.25}$				
$O_{0.5}$	7.3730	5.0829	0.2440	0.6025

[a] Values for the x coordinates of the 6g Ti (x_{Ti}) and 6g Si (x_{Si}) sites are in fractional units.

Table 3
Measured structural parameters for $Ti_5Si_3Z_x$[a]

Z_x	a (Å)	c (Å)	x_{Ti}[a]	x_{Si}[a]
$O_{0.02}$	7.4601(1)	5.1510(1)	0.2507(3)	0.6067(2)
$B_{0.24}$	7.4670(2)	5.1722(2)	0.2495(3)	0.6076(5)
$B_{0.47}$	7.4781(1)	5.1788(1)	0.2478(3)	0.6053(5)
$C_{0.25}$	7.4497(1)	5.1596(1)	0.2446(3)	0.6032(2)
$C_{0.47}$	7.4415(1)	5.1687(1)	0.2391(3)	0.6004(2)
$N_{0.27}$	7.4387(1)	5.1453(1)	0.2439(2)	0.6049(3)
$N_{0.46}$	7.4273(1)	5.1543(1)	0.2379(3)	0.6025(5)
$O_{0.22}$	7.4469(2)	5.1410(1)	0.2454(2)	0.6041(2)
$O_{0.4}$	7.4342(1)	5.1334(1)	0.2419(2)	0.6015(2)

[a] Values for the x coordinates of the 6g Ti (x_{Ti}) and 6g Si (x_{Si}) sites are in fractional units.

lattice parameters underestimate the experimental lattice parameters by 0.6–1.6% as is common for this method of calculation. However, the normalized trends between experimental and theoretical values are in good qualitative agreement, particularly for the change in c-axis as a function of interstitial content. Boron and carbon, being the larger atoms, expand the "c"-lattice parameter; whereas, the smaller oxygen and nitrogen contract it. The agreement between theoretical and experimental trends in the "a"-lattice parameter is not quite as good. Based on experimental measurements, all but boron contract the "a"-lattice parameter; although theoretical calculations show all but oxygen expand the "a"-lattice parameter. The reason for this discrepancy may be due to the underestimation of the lattice by theoretical calculations, although additional research is necessary to confirm this assertion.

Based on the experimental and theoretical structural parameters, nearest-neighbor atomic separations were calculated in order to infer bonding changes as interstitial atoms are added to Ti_5Si_3. Fig. 2 shows the change in nearest-neighbor atomic separations as carbon is added to Ti_5Si_3. In general, all studied interstitial additions led to changes similar to those seen in Fig. 2. There is very good qualitative agreement between the experimental and theoretical changes in atomic separations as a function of interstitial content. The most striking effect of interstitial atoms is to contract the Ti^{6g}–Ti^{6g} and Ti^{6g}–Z separations and to expand the Ti^{6g}–Si separations. This suggests bonding between the

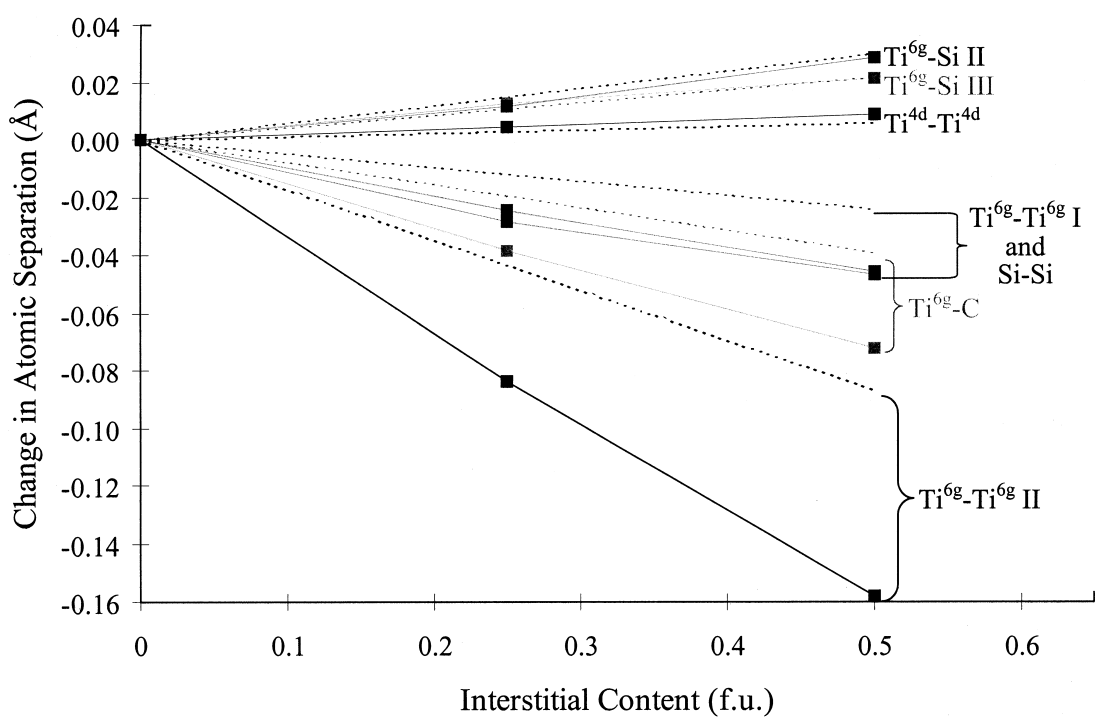


Fig. 2. Change in atomic separations as carbon is added to Ti_5Si_3. Dotted lines represent theoretical calculations, solid lines represent experimental data.

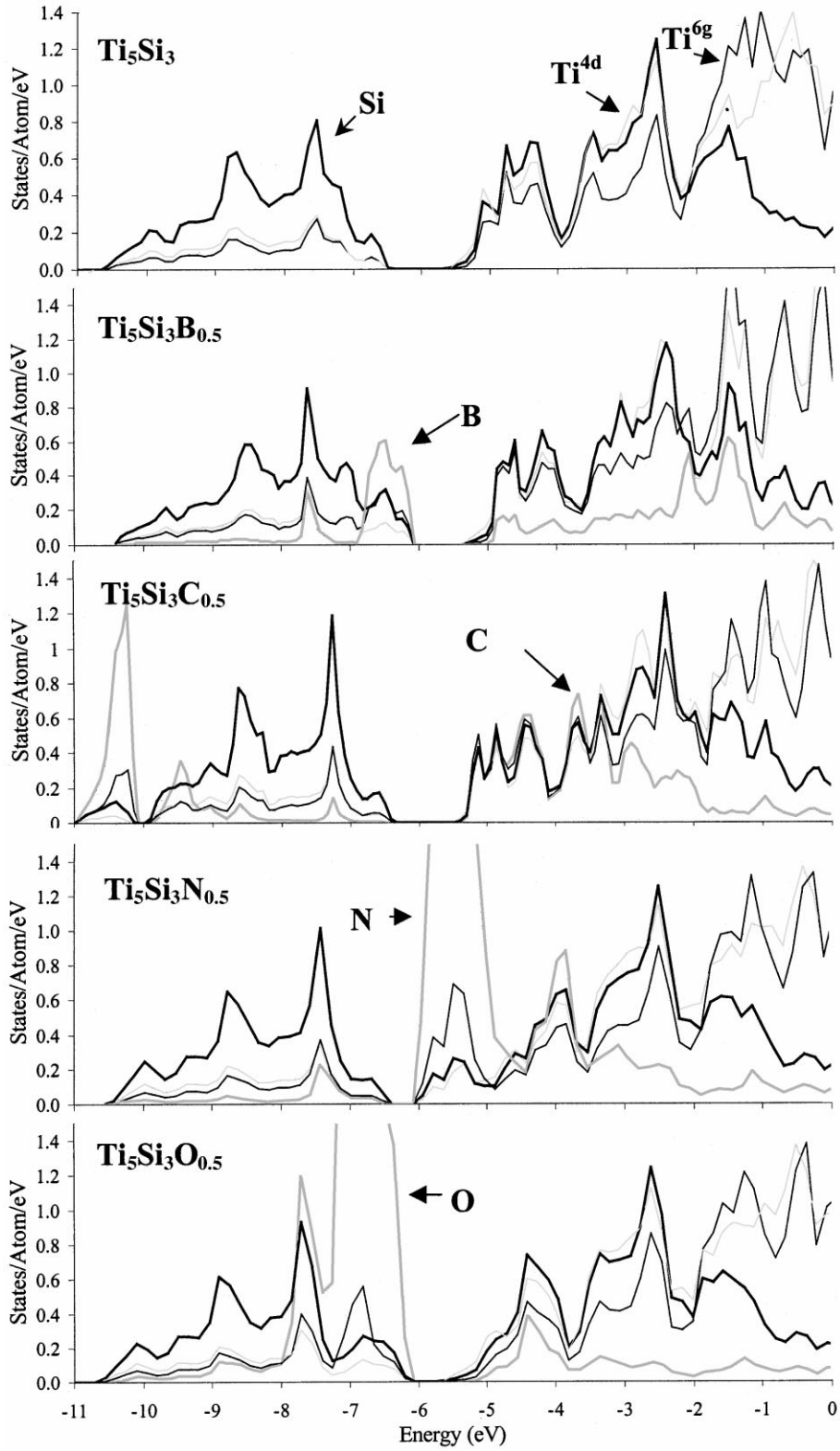


Fig. 3. PDOS's for Ti_5Si_3 and $Ti_5Si_3Z_{0.5}$. Thick, black lines represent Si states; thick, grey lines are Z states; thin, black lines are Ti^{6g} states; thin, grey lines are Ti^{4d} states.

Ti^{6g} and Z atoms and a possible weakening of bonding between the Ti^{6g} and Si atoms. Furthermore, the theoretical calculations suggest that these changes are strongest for carbon. Thus, carbon atoms may be more strongly bonded with the Ti than the other interstitial atoms.

5.3. Densities of state

Fig. 3 gives the partial densities of state (PDOS's) for each of the atoms in Ti_5Si_3 and $Ti_5Si_3Z_{0.5}$. The calculated densities of state for Ti_5Si_3 qualitatively agree with the previously mentioned studies by Long and Chong [8] and Ekman and Ozolins [11]; that is, the region from −2 to −5.5 eV is dominated by d(Ti)–p(Si) mixing, and the region at the Fermi level (0 eV) to −2 eV is dominated by d(Ti) states. These d(Ti) states most likely consist of both bonding and non-bonding electrons. Also, little mixing occurs with the s(Si) states from −6.5 to −10.5 eV.

As boron, carbon, nitrogen or oxygen is added to the lattice, mixing occurs primarily between the interstitial atom's p-state and the surrounding Ti^{6g} atom's d-state. This is shown in Fig. 4 for interstitial carbon and oxygen atoms, which is a plot of the cumulative area of the difference between the PDOS's of $Ti_5Si_3Z_{0.5}$ and Ti_5Si_3. A positive slope in Fig. 4 indicates an increase in the PDOS of an atom in $Ti_5Si_3Z_{0.5}$ compared to that same atom in Ti_5Si_3. Similarly, a negative slope indicates a decrease and a zero slope indicates no change in the PDOS's in $Ti_5Si_3Z_{0.5}$ relative to Ti_5Si_3. Thus, as seen in Fig. 4, addition of 0.5 formula units of oxygen leads to an increase of about 0.3 states for each Ti^{6g} atom at −6 to −7 eV which corresponds exactly to the position of

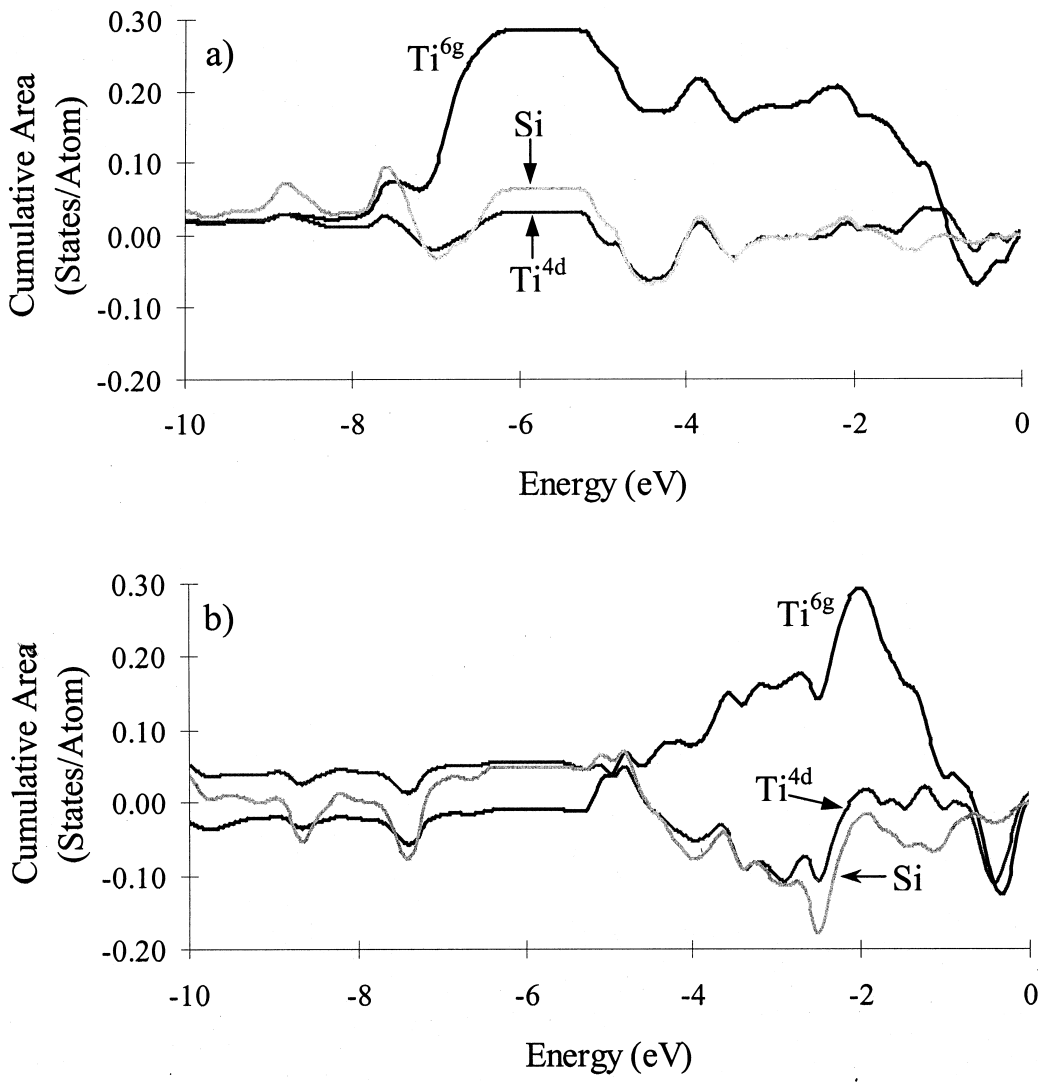


Fig. 4. (a) Cumulative area of the difference between the DOS of an atom of $Ti_5Si_3O_{0.5}$ and the DOS of that same atom in Ti_5Si_3. (b) Cumulative area of the difference between the DOS of an atom in $Ti_5Si_3C_{0.5}$ and the DOS of that same atom in Ti_5Si_3. Large, positive slopes in the Ti^{6g} curves correspond exactly in energy to the position of the interstitial atom's p-state. Smaller features in these figures are due to slight changes in the Fermi level as interstitial atoms are added to Ti_5Si_3.

the oxygen p-band. Also, addition of oxygen leads to a reduction of 0.1 states per Ti^{6g} atom at −4.5 to −5 eV and a reduction of 0.25 states per atom at −0.5 to −2 eV which corresponds to areas of d(Ti)–p(Si) and d(Ti)–d(Ti) interaction, respectively.

Furthermore, little change occurs between the PDOS's of Ti^{4d} and Si atoms in $Ti_5Si_3O_{0.5}$ relative to Ti_5Si_3. Based on these observations, addition of oxygen to Ti_5Si_3 leads to the formation of d(Ti^{6g})–p(O) bonds at the expense of d(Ti^{6g})–d(Ti^{6g}) interaction and to a lesser extent, d(Ti^{6g})–p(Si) interaction. However, d(Ti^{4d})–d(Ti^{4d}) and d(Ti^{4d})-p(Si) interactions remain relatively unaffected.

The effect of carbon addition is similar in that Ti^{6g} atoms show the most dramatic redistribution of electronic states. Also, the increase in states of the Ti^{6g} atoms coincides in energy with the carbon p-states (−2 to −5 eV) and the decrease in states of Ti^{6g} atoms coincides with states associated with d(Ti^{6g})–d(Ti^{6g}) interaction (−0.5 to −2 eV). However, unlike oxygen (and nitrogen), the carbon p-band is considerably broader and is located at similar energy levels as the Si p-band (−2 to −5 eV).

Although not shown, the effect of nitrogen additions is very similar to that of oxygen additions. Boron additions, however, leads to a redistribution of Ti^{6g} states around boron's s-band, which is located at −6 to −7 eV. The majority of boron's p-states are located at −1 to −2 eV, where d(Ti)–d(Ti) interactions predominate.

Finally, although d(Ti^{6g})–p(Z) bonding apparently forms at the expense of d(Ti^{6g})–d(Ti^{6g}) interactions, no significant change in DOS occurs at the Fermi level for all studied compositions. Also, although not obvious in Figs. 2 and 3, the overlap between Ti^{4d} and Si atoms increases slightly as interstitial atoms are added to Ti_5Si_3. This may suggest an increase in bonding between Ti^{4d} and Si atoms.

6. Conclusions

Based on enthalpy of formation calculations, the $D8_8$ structure appears to be the most stable structure for Ti_5Si_3 in accordance with experimental evidence. Also, interstitial additions appear to increase the stability of the $D8_8$ structure. This increase in stability is apparently a result of bonding between the interstitial atom's p-electrons and the Ti^{6g} atom's d-electrons, which results in a strong contraction in separation between these atoms. These bonds form at the expense of Ti^{6g} states located near the Fermi level. All other PDOS features remain relatively unaffected by the incorporation of interstitial atoms.

Acknowledgements

Research at Ames Laboratory was contracted under contract No. W-7405-ENG-82 with the US Department of Energy by Iowa State University. Research at ORNL was sponsored by the Division of Materials Sciences, Office of Basic Energy Research, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The work at MURR was supported by the US Department of Energy grant No. DE-FE02-90ER45427.

References

[1] Nowotny H, Lux B, Kudielka H. Mh Chem 1956;87:447.
[2] Nowotny H. Electronic structure and alloy chemistry of the transition elements. In: Beck, PA editor. New York: Plenum Press, 1963. p.179.
[3] Parthe E, Jeitschko W, Sadagopan V. Acta Crystall 1965;19:1031.
[4] Thom AJ, Akinc M. J Mater Sci Lett 1994;13:1657.
[5] Thom AJ, Akinc M. Advanced ceramics for structural and tribiological applications. In: Hawthorne HM, Troczynski T, editors. Vancouver (BC): The Minerals, Metals & Materials Society, 1994. p.143.
[6] Thom AJ, Young VG, Akinc M. J of Alloys and Compounds 2000;296:59–66.
[7] Williams JJ, Kramer MJ, Malik SK, Akinc M. J Mater Res (in press).
[8] Long X, Chong Z. Trans Nonferrous Meta Soc China 1994; 4(3):25.
[9] Speier W, Kumar L, Sarma DD, de Groot RA, Fuggle JC. J Phys: Condens Matter 1989;1:9117.
[10] Gelatt CD, Williams AR, Moruzzi VL. Phys Rev B 1983;27:2005.
[11] Ekman M, Ozolins V. Phys Rev B 1998;57:4419.
[12] Kajitani T, Kawase T, Yamada K, Hirabayashi M. Trans Jpn Inst of Met 1986;27(9):639.
[13] Guloy AM, Corbett JD. Inorg Chem 1996;35:4669.
[14] Al-Shahery GMY, Jones DW, McColm IJ, Steadman R. J Less Common Metals 1982;87:99.
[15] Wakelkamp W. Diffusion and phase relations in the systems Ti–Si–C and Ti–Si–N. Thesis, Technical University of Eindhoven 1991.
[16] Goldstein JI, Choi SK, Von Loo FJJ, Bastin GF, Metselarr R. J Am Ceram Soc 1995;78(2):313.
[17] Jones RO. Rev Mod Phys 1989;61:689.
[18] Hedin L, Lundqvist BI. J Phys C 1971;4:2064.
[19] Louie SG, Ho KM, Cohen ML. Phys Rev B 1979;19:1774.
[20] Hamann DR, Schluter M, Chiang C. Phys Rev B 1982;25:2103.
[21] Ho KM, Elsasser C, Chan CT, Fahnle M. J Phys: Condens Matter 1992;4:5189.
[22] Fu CL, Wang X, Ye YY, Ho KM. Intermetallics 1998;7:1.

A simple treatment of the "scattering-in" term of the Boltzmann equation for multilayers

X.-G. Zhang[a)]
Computational Physics and Engineering Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6114

W. H. Butler
Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6114

We present a simple approximation for treating anisotropic scattering within the semiclassical Boltzmann equation for current in plane geometry in magnetic multilayers. This approximation can be used to qualitatively account for the forward scattering that is neglected in the lifetime approximation, and requires only one additional parameter. For the case of a bulk material its effect is a simple renormalization of the scattering rate. The simplicity of this term has allowed quick and simple solution to the Boltzmann equation for magnetic multilayers using realistic band structures. When we use the band structures for Cu|Co multilayers obtained from first-principles calculations, we find an increase in the resistance of the multilayer, compared to the solution without the scattering-in term, due to the higher scattering rates needed to fit the same bulk conductivities. The giant-magnetoresistance ratio is also changed when the vertex corrections are included. © *2000 American Institute of Physics.* [S0021-8979(00)72008-1]

Most studies of giant magnetoresistance[1] (GMR) are for the current in plane (CIP) geometry. Typically, these have been based on the Fuchs–Sondheimer solution[2,3] to the Boltzmann equation for thin films and its generalization to multilayers.[4,5] However, the Fuchs–Sondheimer approach utilizes the lifetime approximation which assumes that the scattering probability between states $\mathbf{k}$ and $\mathbf{k}'$ in the Boltzmann equation, $P_{\mathbf{kk}'}$, is isotropic, i.e., is independent of the electron wave vectors before and after the scattering. Specifically, $P_{\mathbf{kk}'}$ is assumed to have the form,

$$P_{\mathbf{kk}'}=\frac{1}{N(E_{\mathbf{k}})\tau}\delta(E_{\mathbf{k}}-E_{\mathbf{k}'}), \tag{1}$$

where τ is the lifetime and $N(E_{\mathbf{k}})$ is the density of states at $E_{\mathbf{k}}$. This isotropic scattering probability allows one to neglect the scattering-in term of the Boltzmann equation because it does not contribute to the current for bulk or CIP.

This approximation usually works rather well for bulk materials, because the effect of the scattering-in term can be included simply by scaling the lifetime. In an inhomogeneous system, however, simply scaling the lifetime would lead to an incorrect spatial distribution of the current. Therefore, the scattering-in corrections need to be included at least approximately and we need to obtain at least a qualitative understanding of their effect on the resistance, GMR, and current distribution. The effects of the scattering-in terms are sometimes referred to as vertex corrections because of the manner in which they appear as a correction to the self-energy in a quantum mechanical calculation.

One approach to including the scattering-in corrections for substitutional impurity scattering is to evaluate the Kubo formula using the coherent potential approximation with vertex corrections.[6] Another approach would be to solve the Boltzmann equation with the scattering-in term using the scattering rate calculated from the multiple scattering solutions of dilute substitutional impurities in a first-principles calculation. For example, in Fig. 1 we show the scattering rate $P_{\mathbf{kk}'}$ due to Co impurities in Cu at a fixed $\mathbf{k}$. The scattering is highly anisotropic, with the maximum occurring at about $\mathbf{k}'=\mathbf{k}$, and a smaller peak at $\mathbf{k}'=-\mathbf{k}$. Thus, the scattering is dominated by forward scattering. For the same measured resistivity, this can greatly reduce the effective lifetime and significantly change the calculated current distribution. Inclusion of such a scattering term in the Boltzmann equation and its numerical solution is straightforward.

However, substitutional impurities represent only one possible source of scattering in a GMR material. Most other defects that cause electron scattering are difficult to model from the first principles. Therefore, it is desirable to have a simple phenomenological model that can account for all forms of defect scattering and can be easily incorporated into the Boltzmann equation. In this article we present such a model that needs only one additional parameter, captures most of the qualitative features of the vertex corrections, and allows a fast solution of the Boltzmann equation.

We use g to describe the deviation of the distribution function from equilibrium, i.e.,

$$f(z,\mathbf{k}_{\parallel})=f_0(E_{\mathbf{k}})+g(z,\mathbf{k}_{\parallel})\delta(E_{\mathbf{k}}-E_F), \tag{2}$$

where f_0 is the equilibrium distribution function and E_F is the Fermi energy. For the CIP geometry, we assume that the layers are stacked in the z direction, and the current flows in the x direction. The Boltzmann transport equation for multilayers is

[a)]Electronic mail: xgz@ornl.gov

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

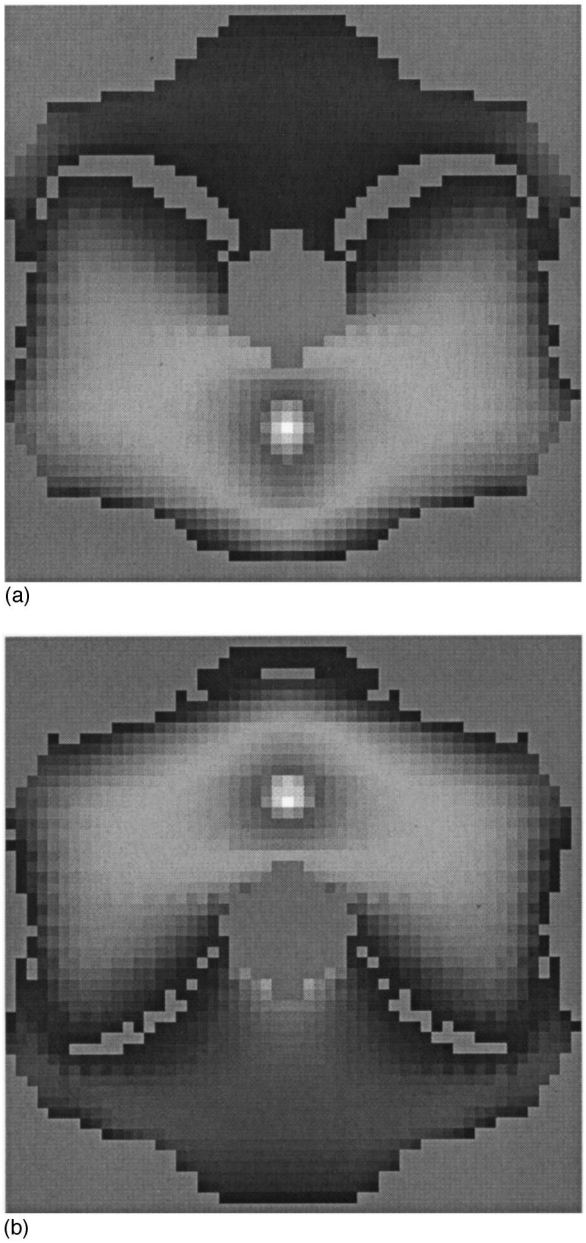


FIG. 1. Scattering rates for a Co impurity in Cu as a function of outgoing wave vector for an incoming wave vector of (0.326, −0.188, −0.561). The upper panel (a) is for negative outgoing wave vectors k_z' (forward scattering) where the highest scattering rate has the value of 0.922, and the lower panel (b) is for positive k_z' (backward scattering) where the highest scattering rate is 0.525. The calculation is for the majority spin channel with Co moment oriented parallel to spin quantization axis.

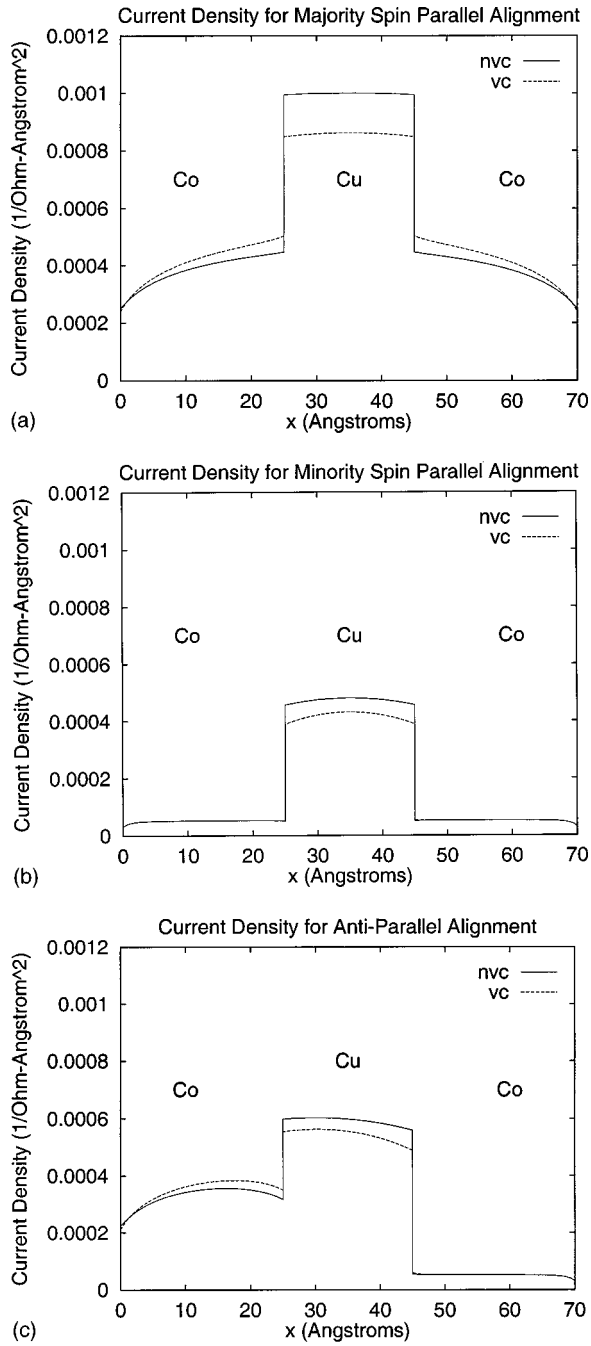


FIG. 2. Current distribution for a Co|Cu|Co trilayer with and without vertex corrections. (a) Majority spin channel for parallel alignment; (b) minority spin channel for parallel alignment; (c) one of the spin channels for antiparallel alignment. Solid lines are without the vertex corrections (nvc) and dotted lines are with the vertex corrections (vc).

$$v_z \frac{\partial g(z,\mathbf{k}_\parallel)}{\partial z} - \sum_{\mathbf{k}_\parallel'} P_{\mathbf{k}\mathbf{k}'}[g(z,\mathbf{k}_\parallel') - g(z,\mathbf{k}_\parallel)] = \frac{eE_x v_x}{m}, \quad (3)$$

where $P_{\mathbf{k}\mathbf{k}'}$ is the scattering rate between states $\mathbf{k}$ and $\mathbf{k}'$ and E_x is the external electric field. The lifetime approximation to the Boltzmann transport equation assumes that the scattering rate is isotropic, Eq. (1), so that

$$v_z \frac{\partial g(z,\mathbf{k}_\parallel)}{\partial z} + \frac{1}{\tau} g(z,\mathbf{k}_\parallel) = \frac{eE_x v_x}{m}, \quad (4)$$

using $\Sigma_{\mathbf{k}_\parallel} g(z,\mathbf{k}_\parallel) = 0$. This leads to the Fuchs–Sondheimer solution,[2,3]

$$g^{(0)}(z,\mathbf{k}_\parallel) = \frac{eE_x v_x \tau}{m}[1 + F^{(0)}(\mathbf{k}_\parallel) e^{-z/v_z \tau}], \quad (5)$$

where the coefficients $F(\mathbf{k}_\parallel)$ are determined by the continuity condition of the distribution function $g(z,\mathbf{k}_\parallel)$ at the boundaries of each layer, and the superscript (0) indicates a solution without the vertex corrections.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

The condition that the scattering rate $P_{\mathbf{kk}'}$ satisfies is $P_{\mathbf{kk}'}=P_{\mathbf{k}'\mathbf{k}}$. Therefore the lowest order correction to the isotropic approximation is

$$P_{\mathbf{kk}'}=\left[\frac{1}{N(E_{\mathbf{k}})\tau}+p_x v_x v_x'+p_y v_y v_y'+p_z v_z v_z'\right]\delta(E_{\mathbf{k}}-E_{\mathbf{k}}'), \tag{6}$$

where v_i are the components of the group velocity at $\mathbf{k}$, and v_i' are those at $\mathbf{k}'$, and p_i are the coefficients of expansion, and are presumably small. Substituting back into Eq. (3), we find

$$v_z\frac{\partial g(z,\mathbf{k}_\parallel)}{\partial z}+\frac{1}{\tau}g(z,\mathbf{k}_\parallel)=\frac{eE_x v_x}{m}+p_x j_x(z)v_x, \tag{7}$$

where $j_x(z)$ is the current density at z,

$$j_x(z)=\sum_{\mathbf{k}_\parallel} v_x g(z,\mathbf{k}_\parallel). \tag{8}$$

The second term on the right-hand side represents a simple approximation to the vertex corrections. For a homogeneous bulk system, there is no z dependence of j_x and g. Thus g can be trivially solved as

$$g(\mathbf{k}_\parallel)=\frac{eE_x v_x\overline{\tau}}{m}, \tag{9}$$

and

$$j_x=\frac{eE_x\overline{\tau}}{m}\sum_{\mathbf{k}_\parallel} v_x^2, \tag{10}$$

where

$$\overline{\tau}=\frac{\tau}{1-p_x\tau\Sigma_{\mathbf{k}_\parallel}v_x^2}. \tag{11}$$

This is identical to the solution without the vertex corrections, but with a renormalized lifetime.

The solution to Eq. (7) for a general multilayer system is

$$g(z,\mathbf{k}_\parallel)=g^{(0)}(z,\mathbf{k}_\parallel)+g^{(1)}(z,\mathbf{k}_\parallel), \tag{12}$$

where the extra term due to the vertex corrections can be calculated iteratively through

$$g^{(1)}(z,\mathbf{k}_\parallel)=e^{-z/v_z\tau}\left[g^{(1)}(z_0,\mathbf{k}_\parallel)+\frac{p_x v_x}{v_z}\int_{z_0}^{z}dz\ e^{z/v_z\tau}j_x(z)\right]. \tag{13}$$

The term $g^{(1)}(z_0,\mathbf{k}_\parallel)$ represents the boundary condition at $z=z_0$ which is determined from the reflection and the transmission coefficients of the interfaces, which are determined using the layer-Korringa–Koha–Rostoker approach.[7]

We calculated the effect of the vertex corrections for a Co|Cu|Co trilayer system. The current distribution in the trilayer is plotted for parallel and antiparallel alignments for each spin channel in Fig. 2. The effect of the vertex correction term is compared. The scattering rates (with and without the vertex corrections) in each layer are adjusted so that they always give the same bulk resistivity as measured experimentally. Consequently, the scattering rates when the vertex corrections are included are much higher than without the vertex corrections. The size of the vertex correction term is chosen so that the scattering rates for the Cu layer and for the Co majority spin channel are changed by a factor of 2. For the Co minority spin channel the scattering is assumed to be isotropic so that the scattering-in term vanishes. The calculation shows that the vertex corrections cause a large reduction in the current in the Cu layers, but have much smaller effect in the Co layers. The combined effect is a higher resistance for the film. We also observe a small reduction in the GMR ratio, from about 20% to about 17%.

In conclusion, a simple term is added to the Boltzmann equation to account for the ''scattering-in'' term. It is demonstrated that this term can change dramatically the current distribution in the GMR systems, but has a small effect on the GMR ratio. We note that this term captures qualitatively the anisotropic feature of defect scattering, but except for a maximum in the forward direction, it is rather unlike the anisotropic impurity scattering rates one would obtain from first principles. The biggest difference is that the impurity scattering rates have a small maximum in the backward direction, while Eq. (6) gives a minimum in the backward direction. This difference, however, should have minimal effect on the change of current distribution due to vertex corrections.

This work was supported by DARPA through NASA Grant No. NAG-1-1962.

[1] M. N. Baibich *et al.*, Phys. Rev. Lett. **61**, 2472 (1988).
[2] K. Fuchs, Proc. Cambridge Philos. Soc. **34**, 100 (1938).
[3] E. H. Sondheimer, Adv. Phys. **1**, 1 (1952).
[4] R. E. Camely and J. Barnaś, Phys. Rev. Lett. **63**, 664 (1989).
[5] R. Q. Hood, L. M. Falicov, and D. R. Penn, Phys. Rev. B **49**, 368 (1994).
[6] W. H. Butler, X.-G. Zhang, D. M. C. Nicholson, and J. M. MacLaren, J. Appl. Phys. **76**, 6808 (1994).
[7] J. M. MacLaren, X.-G. Zhang, and W. H. Butler, Phys. Rev. B **59**, 5470 (1999).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

Generalized local-density approximation for spherical potentials

X.-G. Zhang and D. M. C. Nicholson
Computational Physics and Engineering Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6114
(Received 29 May 1998; revised manuscript received 5 April 1999)

An alternative density functional for the spherical approximation of cell potentials is formulated. It relies on overlapping atomic spheres for the calculation of the kinetic energy, similar to the atomic sphere approximation (ASA), however, a shape correction is used that has the same form as the interstitial treatment in the nonoverlapping muffin-tin (MT) approach. The intersite Coulomb energy is evaluated using the Madelung energy as computed in the MT approach, while the on-site Coulomb energy is calculated using the ASA. The Kohn-Sham equations for the functional are then solved self-consistently. The ASA is known to give poor elastic constants and good point defect energies. Conversely the MT approach gives good elastic constants and poor point defect energies. The proposed new functional maintains the simplicity of the spherical potentials found in the ASA and MT approaches, but gives good values for both elastic constants and point defects. This solution avoids a problem, absent in the ASA but suffered by the MT approximation, of incorrect distribution of site charges when charge transfer is large. Relaxation of atomic positions is thus facilitated. Calculations confirm that the approach gives similar elastic constants to the MT approximation, and defect formation energies similar to those obtained with ASA. [S0163-1829(99)09631-9]

I. INTRODUCTION

Traditional band structure techniques based on the multiple scattering Green function approach[1] require the spherical approximation of the electronic cell potentials. Although great progress has been made in the development of full-potential multiple scattering theory,[2] calculations using spherical potentials still remain the norm because of their speed and simplicity. There are two ways that the spherical approximation is invoked. The first is the so-called ''muffin-tin'' (MT) approximation,[3] where space is divided into nonoverlapping spherical volumes centered at each nucleus within which the potential is spherical, and the interstitial region between these spheres, where the potential is constant. The second form is the ''atomic sphere approximation'' (ASA),[4] which approximates the space by a collection of spheres, centered at the nuclei, whose volume equals the volume of the corresponding atomic cell.

Both of these approximations have been quite successful in predicting a wide range of properties of metals and alloys. However, they suffer severe limitations. In particular, the ASA cannot give reliable energy differences between structures that differ in the shape of the atomic cell, thus it is unable to reliably predict the stable structure for materials. This is mainly due to the incorrect Coulomb energy contribution from the interstitial charge in the ASA. The interstitial charge plays an important role in determining structural properties, but the ASA is the poorest in describing this charge. On the other hand, the MT approximation does a reasonably good job of describing the interstitial charge for transition metals, as reflected in the good elastic constants obtained for several materials using the MT approximation.[5] However, the MT approach usually fails in systems where there is a large charge transfer. It cannot in general even predict the correct amount of charge transfer in these systems. This is because, in the reconstruction of the effective potential, the interstitial charge is averaged and then redistributed over all the atomic cells. A local change in the interstitial charge around a defect is transferred by the averaging process to the far reaches of the unit cell. This introduces an artificial charge transfer when the real electron density in the interstitial regions of different atomic cells is different. This problem becomes worse for large unit cells, such as those often used in connection with the locally self-consistent multiple scattering (LSMS) technique.[6] In these applications, even a small error in the interstitial charge due to the MT approximation can give an artificial charge transfer between very distant atoms in the unit cell. This can lead to a large error in the Coulomb energy because of the large size of the unit cell.

There has been work directed at overcoming these difficulties. One approach is to add a perturbative correction to the ASA density functional by replacing the Coulomb energy by a more accurate term using the MT charge density.[7] This has been improved to include corrections due to the full-cell charge density to the ASA total energy.[8] This approach allows one to calculate elastic constants accurately for transition metals. However the correction term is not self-consistent, thus its reliability cannot be affirmed. Alternatively one may employ a full potential approach. However, experience in the past decade has shown that a general, fully self-consistent full-potential approach for MST would be rather cumbersome.

In consideration of the fact that for some time to come a majority of MST calculations will use a spherical approximation, we propose a scheme that combines the strengths of each approach, overcomes their limitations, and maintains the simplicity of spherical approaches. We will present a new energy functional, from which a self-consistent procedure is derived. This procedure contains a bandstructure part that is based on MST within the ASA plus a shape correction, and an intersite Coulomb energy part that is analogous to the standard MT approach. We demonstrate this new approach

 ©1999 The American Physical Society

with calculations of elastic constants of a few transition metals, and the vacancy formation energy of copper.

II. THE ELECTRON DENSITY FUNCTIONAL

The total energy functional in a typical local density functional approach[9,10] is written as

$$E[\rho]=T[\rho]+U[\rho]+E_{\mathrm{xc}}[\rho], \tag{1}$$

where the three terms on the right are the kinetic energy, Coulomb energy, and exchange-correlation energy, respectively for the electron density $\rho(\mathbf{r})$. The kinetic energy is given by

$$T=\sum_{\epsilon<E_F}\epsilon-\int d^3\mathbf{r}\,\rho(\mathbf{r})V_{\mathrm{eff}}(r), \tag{2}$$

where $\Sigma\epsilon$ is the band energy sum obtained using one of the standard bandstructure approaches to the Shrödinger equation for the potential V_{eff}, and E_F is the Fermi energy for the electrons.

The total Coulomb energy, in atomic units, is,

$$U[\rho]=\int d^3\mathbf{r}\int d^3\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}+\sum_{ij,i\neq j}\frac{Z_iZ_j}{|\mathbf{R}_i-\mathbf{R}_j|}-2\sum_i\int d^3\mathbf{r}\frac{Z_i\rho(\mathbf{r})}{|\mathbf{R}_i-\mathbf{r}|}, \tag{3}$$

where Z_i are the atomic numbers and $\mathbf{R}_i$ are the positions of the nuclei. We separate U into site diagonal and site off-diagonal parts,

$$U[\rho]=U^{(\mathrm{o})}+\sum_i U_i^{(\mathrm{d})}[\rho_i,Z_i], \tag{4}$$

with

$$U_i^{(\mathrm{d})}[\rho_i,Z_i]=\int d^3\mathbf{r}\int d^3\mathbf{r}'\frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}-2\int d^3\mathbf{r}\frac{Z_i\rho_i(\mathbf{r})}{|\mathbf{r}|}. \tag{5}$$

The Kohn-Sham effective potential[10] V_{eff} is obtained from the requirement that the total energy functional is variational with respect to the charge density,

$$\frac{\delta E[\rho]}{\delta\rho(\mathbf{r})}=-V_{\mathrm{eff}}(r)+\frac{\delta W[\rho]}{\delta\rho(\mathbf{r})}=0, \tag{6}$$

where

$$W[\rho]=U[\rho]+E_{\mathrm{xc}}[\rho]. \tag{7}$$

In order to find a generalized approach to calculate the Coulomb energy which reduces to the standard Coulomb energy involving the Madelung term in the limit of MT approximation, we introduce a reference system composed of point charges $Q_i^{\text{ASA-MT}}$ at the lattice sites and a compensating uniform electron density,

$$\rho^{\mathrm{comp}}=\frac{\sum_i Q_i^{\text{ASA-MT}}}{\sum_i \Omega_i}. \tag{8}$$

We will see later that it is the different definitions for the charge $Q_i^{\text{ASA-MT}}$ that gives rise to the different energy expressions in the ASA, MT, and ASA-MT approximations. The electrostatic energy of the reference system is given by a Madelung sum and can also be separated into site diagonal and off-diagonal parts,

$$U_{\mathrm{Mad}}[\rho]=U_{\mathrm{Mad}}^{(\mathrm{o})}+\sum_i U_i^{(\mathrm{d})}[\rho^{\mathrm{comp}},Q_i^{\text{ASA-MT}}]=\frac{1}{2}\sum_{ij}Q_i^{\text{ASA-MT}}M_{ij}Q_j^{\text{ASA-MT}}, \tag{9}$$

where M_{ij} is the Madelung matrix. Upon adding and subtracting the Madelung energy of the reference system,

$$W[\rho]=\sum_i W_i[\rho]+U^{(\mathrm{o})}[\rho]+\frac{1}{2}\sum_{ij}Q_i^{\text{ASA-MT}}M_{ij}Q_j^{\text{ASA-MT}}-U_{\mathrm{Mad}}^{(\mathrm{o})}[\{Q_i^{\text{ASA-MT}}\}], \tag{10}$$

and

$$W_i[\rho]=U_i^{(\mathrm{d})}[\rho_i,Z_i]-U_i^{(\mathrm{d})}[\rho^{\mathrm{comp}},Q_i^{\text{ASA-MT}}]+E_{\mathrm{xc}}[\rho_i], \tag{11}$$

where $\rho_i=\rho(\mathbf{r})\sigma_i(\mathbf{r})$ and $\sigma_i(\mathbf{r})$ is the cell shape function for cell i. We now look for an approximation such that $U^{(\mathrm{o})}[\rho]=U_{\mathrm{Mad}}^{(\mathrm{o})}[\{Q_i^{\text{ASA-MT}}\}]$.

The spherical shape approximations such as MT or ASA can be thought of as approximations to the explicit form of the Coulomb and exchange-correlation energy $W[\rho]$ as a functional of the electron density $\rho(\mathbf{r})$. Approximate forms for $W[\rho]$ can be constructed by introducing auxiliary densities that in turn depend on $\rho(\mathbf{r})$. A general form for such a density that covers both the MT and ASA cases consists of a spherical part, $\rho_i^{\mathrm{S}}(r)$, that is defined over a sphere of radius $r_{i\mathrm{S}}$, an interstitial part, ρ^{int}, and an additional term containing ΔQ_i, a charge that will be defined later. The spherical part inside the MT sphere is easily defined as the angular average of the full charge density,

$$\rho_i^{\mathrm{S}}(r)=\frac{1}{4\pi}\int_i d\Omega\,\rho_i(\mathbf{r}),\quad r<r_{i\mathrm{MT}}. \tag{12}$$

If $r_{i\mathrm{S}}>r_{i\mathrm{MT}}$, then for r outside the MT radius but inside $r_{i\mathrm{S}}$, the definition of $\rho^{\mathrm{S}}(r)$ in terms of the full charge density is not clear. Specifically, the site electron density is zero in those volumes that are inside r_{S} but outside the WS cell. In this case the standard interpretation is that a space of equal volume outside r_{S} but inside the WS cell is mapped into this volume, and then an angular average of the charge is performed. Thus

$$\rho_i^{\mathrm{S}}(r)=\frac{1}{4\pi}\int_i d\Omega\,\rho_i(\mathbf{r}'[\mathbf{r}]),\quad r_{i\mathrm{S}}>r>r_{i\mathrm{MT}}, \tag{13}$$

where $\mathbf{r}'[\mathbf{r}]$ indicates the mapping. This mapping leaves an interstitial space of volume $\Omega_i - \Omega_i^{\mathrm{S}}$, that lies inside the ''ith'' WS cell. Where Ω_i is the cell volume, and

$$\Omega_i^{\mathrm{S}} = \frac{4\pi}{3} r_{i\mathrm{S}}^3 .$$

This volume vanishes when $r_{i\mathrm{S}} = r_{i\mathrm{WS}}$. We introduce the function $s_i(\mathbf{r}) = 1$ (0) for points, $\mathbf{r}$, inside (outside) this interstitial region. If the electron density is nearly flat in the outer portion of the cell or if $r_{i\mathrm{S}}$ is very close to $r_{i\mathrm{MT}}$ the exact nature of the mapping in Eq. (13) is inconsequential. We adhere to the ASA convention and will therefore leave the mapping unspecified with the understanding that it will typically have little impact on results. The interstitial charge density is determined by the charge in the entire interstitial volume, i.e., that outside the MT spheres,

$$\rho^{\mathrm{int}} = \frac{1}{\Omega^{\mathrm{int}}} \sum_i \left[Z_i - \int d^3\mathbf{r}\, \rho_i^{\mathrm{S}}(r) \Theta(r_{i\mathrm{MT}} - r) \right], \tag{14}$$

where $\Omega^{\mathrm{int}} = \Sigma_i (\Omega_i - 4\pi/3 r_{i\mathrm{MT}}^3)$. Note that this definition is independent of $r_{i\mathrm{S}}$, and in the limit of $r_{i\mathrm{S}} = r_{i\mathrm{WS}}$, ρ^{int} remains unchanged although the volume it occupies goes to zero. Finally, we add a number of electrons ΔQ_i to adjust the total site charge. As we will see later, this term is required if one includes a shape correction to the ASA. If $\Delta Q_i = 0$ the standard MT(ASA) densities are recovered when $r_{i\mathrm{S}} = r_{i\mathrm{MT}}(r_{i\mathrm{WS}})$. Combining all three terms, we have

$$\rho_i^{\text{ASA-MT}}(\mathbf{r}) = \rho_i^{\mathrm{S}}(r) + \Delta Q_i \delta(r - r_{i\mathrm{S}}) + \rho^{\mathrm{int}} s_i(\mathbf{r}). \tag{15}$$

The region occupied by ρ^{int} has zero volume when $r_{\mathrm{S}} = r_{\mathrm{WS}}$. When $r_{\mathrm{S}} = r_{\mathrm{MT}}$, s_i is the volume in the cell that is outside the ''muffin-tin.'' At values of $r_{i\mathrm{S}}$ between $r_{i\mathrm{MT}}$ and $r_{i\mathrm{WS}}$, the shape of this electron density becomes complicated because $s_i(\mathbf{r})$ is complicated. It is still, however, usable and leads to spherical potentials, because it only enters the energy as a prefactor to ρ^{int} which depends only on the spherical part of the electron density. Furthermore, this form provides a continuous link between the MT and ASA approaches. The calculations presented in this work are done with $r_{i\mathrm{S}} = r_{i\mathrm{WS}}$.

We follow the standard MT and ASA procedure of approximating W_i by replacing ρ_i in E_{xc} and $U_i^{(\mathrm{d})}$ by $\rho_i^{\text{ASA-MT}}$, and replacing ρ^{comp} by

$$\rho_{\mathrm{comp}}^{\text{ASA-MT}} = \rho^{\mathrm{comp}} [\Theta(r_{i\mathrm{S}} - r) + s_i(r)], \tag{16}$$

so that

$$\begin{aligned} W_i^{\text{ASA-MT}}[\rho] &= U_i^{(\mathrm{d})}[\rho_i^{\text{ASA-MT}}, Z_i] - U_i^{(\mathrm{d})}[\rho_{\mathrm{comp}}^{\text{ASA-MT}}, Q_i^{\text{ASA-MT}}] + E_{\mathrm{xc}}[\rho_i^{\mathrm{S}}] + \rho^{\mathrm{int}} \epsilon_{\mathrm{xc}}(\rho^{\mathrm{int}})(\Omega_i - \Omega_i^{\mathrm{S}}) + \mu_{\mathrm{xc}}(\rho_i^{\mathrm{S}}(r_{i\mathrm{S}}))\Delta Q_i \\ &= \int d^3\mathbf{r}\, d^3\mathbf{r}' \frac{\rho_i^{\text{ASA-MT}}(\mathbf{r}) \rho_i^{\text{ASA-MT}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - 2 \int d^3\mathbf{r} \frac{\rho_i^{\text{ASA-MT}}(\mathbf{r}) Z_i}{r} - \int d^3\mathbf{r}\, d^3\mathbf{r}' \frac{\rho_{\mathrm{comp}}^{\text{ASA-MT}}(\mathbf{r}) \rho_{\mathrm{comp}}^{\text{ASA-MT}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &\quad + 2 \int d^3\mathbf{r} \frac{\rho_{\mathrm{comp}}^{\text{ASA-MT}}(\mathbf{r}) Q_i^{\text{ASA-MT}}}{r} + E_{\mathrm{xc}}[\rho_i^{\mathrm{S}}] + \rho^{\mathrm{int}} \epsilon_{\mathrm{xc}}(\rho^{\mathrm{int}})(\Omega_i - \Omega_i^{\mathrm{S}}) + \mu_{\mathrm{xc}}(\rho_i^{\mathrm{S}}(r_{i\mathrm{S}}))\Delta Q_i , \end{aligned} \tag{17}$$

where μ_{xc} and ϵ_{xc} are the exchange-correlation potential and energy density, respectively. We have used a linear expansion to obtain the exchange-correlation energy which is valid for the case where ρ_i^{S} for $r > r_{\mathrm{MT}}$ is slowly varying and large compared to the electron density associated with ΔQ_i. We continue following the standard MT and ASA procedure and take $U^{(\mathrm{o})}$ to be equal to $U_{\mathrm{Mad}}^{(\mathrm{o})}$. They therefore cancel in Eq. (10) leaving the Madelung energy,[11] explicitly,

$$W^{\text{ASA-MT}}[\rho] = \sum_i W_i^{\text{ASA-MT}}[\rho] + \frac{1}{2} \sum_{ij} Q_i^{\text{ASA-MT}} M_{ij} Q_j^{\text{ASA-MT}}. \tag{18}$$

At this point we have a fairly general form for the density functional that depends on the electron density only through $\rho_i^{\text{ASA-MT}}$ and $Q_i^{\text{ASA-MT}}$. The density $\rho_i^{\text{ASA-MT}}$ depends on the spherical average of $\rho(\mathbf{r})$ and on ΔQ_i which has not, at this point, been specified. Likewise the dependence of $Q_i^{\text{ASA-MT}}$ on $\rho(\mathbf{r})$ has yet to be specified. It is convenient to define the cell charge,

$$Q_i^{\mathrm{cell}} = \int d^3\mathbf{r}\, \rho_i^{\text{ASA-MT}} = Q_i^{\mathrm{S}} + \Delta Q_i + \rho^{\mathrm{int}}(\Omega_i - \Omega_i^{\mathrm{S}}), \tag{19}$$

where

$$Q_i^{\mathrm{S}} = \int_0^{r_{i\mathrm{S}}} d^3\mathbf{r}\, \rho_i^{\mathrm{S}}(\mathbf{r}). \tag{20}$$

The sum over all sites of the cell charge gives the total number of electron is the system and is used to determine the Fermi level. The essence of the various spherical approximations, ASA or MT, is the choice of $r_{i\mathrm{S}}$ and $Q_i^{\text{ASA-MT}}$. The ASA, where $\Delta Q_i = 0$ and $r_{i\mathrm{S}} = r_{i\mathrm{WS}}$, is derived by taking

$$\lim_{\mathrm{ASA}} Q_i^{\text{ASA-MT}} = Q_i^{\mathrm{ASA}} = Z_i - Q_i^{\mathrm{cell}}. \tag{21}$$

It is easy to see that, in the ASA, $\rho^{\mathrm{comp}} = 0$. One arrives at the MT approximation, where $\Delta Q_i = 0$ and $r_{i\mathrm{S}} = r_{i\mathrm{MT}}$, by setting

$$\lim_{\mathrm{MT}} Q_i^{\text{ASA-MT}} = Q_i^{\mathrm{MT}} = Z_i - Q_i^{\mathrm{cell}} + \rho^{\mathrm{int}} \Omega_i . \tag{22}$$

TABLE I. Calculated equilibrium lattice constants a and bulk moduli B_0 for Cu, Ni, and Fe as compared with experimental data.

	a (atomic units)		B_0 (Mbar)	
	Expt.	ASA-MT	Expt.	ASA-MT
Cu	6.83	6.78	1.37	1.43
Ni	6.66	6.59	1.80	1.82
Fe	5.42	5.28	1.78	0.98

This leads to $\rho^{\text{comp}}=\rho^{\text{int}}$. In order to benefit from the more accurate description of the Madelung Coulomb energy in the MT approximation we follow the MT procedure and make the following definition:

$$Q_i^{\text{ASA-MT}}=Z_i-Q_i^{\text{cell}}+\rho^{\text{int}}\Omega_i. \tag{23}$$

This gives a Madelung energy that corresponds to $\rho^{\text{comp}}=\rho^{\text{int}}$ which is the same interstitial charge as in the MT case. However, the major difference between the MT approach and the ASA-MT approach is that the interstitial part of the on-site Coulomb energy is calculated using ρ^{int} in the MT approach, but using ρ_i^{S} in the ASA-MT approach.

We now turn our attention to the definition of ΔQ_i. This quantity is introduced as a means of assigning a physically reasonable amount of charge to the cell density $\rho^{\text{ASA-MT}}$. If $r_{\text{S}}=r_{\text{MT}}$ and all atoms are equivalent then ΔQ_i is zero because the total number of electrons in the cell must equal Z_i. However, if there are inequivalent atoms, ΔQ_i is given by the integrated difference over the interstitial volume of cell i between the electron density and ρ^{int}. In the MT approximation the electron densities in the interstitial regions are given by the plane wave solutions for a constant potential, V_{MT} that match to the MST solutions on the surfaces of the MT spheres. We extend this definition to larger values of r_{S} by defining,

$$\Delta Q_i=q_i-\rho^{\text{int}}(\Omega_i-\Omega_i^{\text{S}}), \tag{24}$$

where q_i is the integral, of the square of plane waves matched to the MST solutions at r_{S}, from the sphere of radius r_{S} to the WS cell boundary with the convention that contributions to the integral inside r_{S} are negative and those from outside r_{S} are positive. As r_{S} increases the magnitude of q_i is reduced because the volume of integration inside r_{S} approaches the volume outside. When $r_{\text{S}}=r_{\text{MT}}$, for a given self-consistent iteration, the Fermi energy is independent of whether or not ΔQ_i is set to zero because the sum of all ΔQ_i is zero by construction. However, for other choices of r_{S} the Fermi level and hence the eigenvalue sum will be affected at each iteration. Replacing ΔQ_i in Eqs. (23) and (19) by its definition according to Eq. (24) we obtain

$$Q_i^{\text{ASA-MT}}=Z_i-Q_i^{\text{S}}-q_i+\rho^{\text{int}}\Omega_i \tag{25}$$

and

$$Q_i^{\text{cell}}=Q_i^{\text{S}}+q_i. \tag{26}$$

Using the definitions of $\rho_i^{\text{ASA-MT}}$, ρ^{comp}, and $Q_i^{\text{ASA-MT}}$ in Eq. (17) we find,

$$W_i^{\text{ASA-MT}}[\rho_i]=\int d^3\mathbf{r}\,d^3\mathbf{r}'\frac{\rho_i^{\text{S}}(r)\rho_i^{\text{S}}(r')}{|\mathbf{r}-\mathbf{r}'|}-2\int d^3\mathbf{r}\frac{\rho_i^{\text{S}}(r)Z_i}{r}-\frac{6[\rho^{\text{int}}\Omega_i^{\text{S}}]^2}{5r_{i\text{S}}}+4\pi\rho^{\text{int}}Q_i^{\text{ASA-MT}}r_{i\text{S}}^2+E_{\text{xc}}[\rho_i^{\text{S}}]$$
$$+\rho^{\text{int}}\epsilon_{\text{xc}}(\rho^{\text{int}})(\Omega_i-\Omega_i^{\text{S}})+\mu_{\text{xc}}(\rho_i^{\text{S}}(r_{i\text{S}}))\Delta Q_i+\Delta Q_i\left(\frac{\Delta Q_i+2(Q_i^{\text{S}}-Z_i)}{r_{i\text{S}}}\right). \tag{27}$$

The requirement that the functional derivative of the energy with respect to the electron density be zero gives the one-electron effective potential. There are two contributions to the V_{eff}. The first is from the derivative with respect to the explicit spherical electron density, $\rho_i^{\text{S}}(r)$, which gives a contribution to the potential for all r inside $r_{i\text{S}}$. The second contribution comes from the derivative with respect to ρ^{int} which depends implicitly on $\rho_i^{\text{S}}(r)$ for $r<r_{\text{MT}}$ and therefore contributes only to the potential inside the MT radius. This introduces a discontinuity in the effective potential at r_{MT}, which must be dealt with explicitly in the solution of radial wave equations. Using Eqs. (25) and (14), we can take the

TABLE II. Shear moduli for Cu, Ni, and Fe calculated at the experimental lattice constants. All numbers are in Mbar.

	$C_{11}-C_{12}$				C_{44}			
	Exp	MT	ASA	ASA-MT	Exp	MT	ASA	ASA-MT
Cu	0.47	0.36	−0.48	0.36	0.75	1.60	−0.96	1.56
Ni	0.99	0.85	0.03	0.83	1.25	2.21	−0.65	2.11
Fe	0.96	1.08	0.63	1.04	1.12	1.63	−0.93	1.54

derivative with respect to $\rho(r)$ of the Coulomb parts of the potential energy and obtain the electrostatic part of the potential,

$$\phi_i^{\text{ASA-MT}}=\frac{\delta U^{\text{ASA-MT}}}{\delta\rho_i(r)}=\int d^3\mathbf{r}'\frac{2\rho_i^{\text{S}}(r)}{|\mathbf{r}-\mathbf{r}'|}-\frac{2Z_i}{r}-V_i^0 -V_1\Theta(r_{i\text{MT}}-r),\quad r<r_{i\text{S}}, \tag{28}$$

where V_i^0 and V_1 are constants,

$$V_i^0=\sum_j M_{ij}Q_j^{\text{ASA-MT}}+4\pi\rho^{\text{int}}r_{i\text{S}}^2-\frac{2\Delta Q_i}{r_{i\text{S}}}, \tag{29}$$

and

$$V_1=\sum_i V_{i1}$$

$$\Omega^{\text{int}}V_{i1}=\sum_j \Omega_i M_{ij}Q_j^{\text{ASA-MT}}+\frac{4\pi}{5}\rho^{\text{int}}\Omega_i^{\text{S}}r_{i\text{S}}^2+4\pi r_{i\text{S}}^2Q_i^{\text{ASA-MT}} +(\Omega_i-\Omega_i^{\text{S}})\left[\frac{4\pi}{3}\rho^{\text{int}}r_{i\text{S}}^2+\frac{2Q_i^{\text{ASA-MT}}}{r_{i\text{S}}}\right]. \tag{30}$$

V_i^0 is the potential at the origin from a lattice of point charges in a compensating uniform background but with the contribution from the uniform charge within $r_{i\text{S}}$ subtracted and the potential from ΔQ_i added. V_1 is from derivatives with respect to ρ^{int} and is a shift in the potential at the MT radius, similar to the MT step in the case of the MT approximation.

The derivative with respect to q_i involves the density in the interstitial region and is site dependent. It can potentially give contributions to the potential that are not spherical and are site dependent. Recall however, that the charge q_i is derived from a sum of free electron solutions in a constant potential, $\overline{V}_i$. The appropriate value of this constant is determined by minimizing the energy with respect to q_i. Noting that $\delta U^{\text{ASA-MT}}/\delta q_i=\int d^3\mathbf{r}\,\delta U^{\text{ASA-MT}}/\delta\rho_i^{\text{S}}(r)\,\delta(r-r_{i\text{S}})$, we obtain

$$\overline{V}_i=\phi_i^{\text{ASA-MT}}(r_{i\text{S}})+\mu_{\text{xc}}[\rho_i^{\text{S}}(r_{i\text{S}})]. \tag{31}$$

This implies that the MST solution is to be referenced from a different constant potential for each site. This is inconsistent with determination of the eigenvalue sum within the MST approach. However, we are at liberty to solve the MST equations with a site independent constant potential, $\overline{V}$, and afterwards correct the eigenvalue sum to any desired level of perturbation in the difference between $\overline{V}$ and $\overline{V}_i$. We chose to include only first order perturbation corrections, and further choose the value of $\overline{V}$ that renders these to be zero, namely,

$$\overline{V}=\frac{\sum_i q_i\overline{V}_i}{\sum_i q_i}. \tag{32}$$

Finally, collecting the pieces discussed above and setting the zero of energy for the potential to $\overline{V}$ we have the ASA-MT energy functional,

$$E[\rho]=T^{\text{ASA-MT}}[\rho]+U^{\text{ASA-MT}}[\rho]+E_{\text{xc}}^{\text{ASA-MT}}[\rho], \tag{33}$$

where

$$T^{\text{ASA-MT}}[\rho]=\sum_{\epsilon<E_F}\epsilon-\sum_i\int d^3\mathbf{r}\,\rho_i^{\text{S}}(r)V_{i\text{eff}}(r), \tag{34}$$

$$U^{\text{ASA-MT}}[\rho]=\sum_i\left\{-\int_0^{r_{i\text{S}}}d^3\mathbf{r}\frac{\rho_i^{\text{S}}(r)2Z_i}{r}+2\int d^3\mathbf{r}\,\rho_i^{\text{S}}(r)\int^r d^3\mathbf{r}'\frac{\rho_i^{\text{S}}(r')}{r'}\right\}+\sum_i\left\{-\frac{6}{5}\sum_i\frac{[\rho^{\text{comp}}\Omega_i^{\text{S}}]^2}{r_{i\text{S}}} +4\pi\rho^{\text{comp}}\sum_i Q_i^{\text{ASA-MT}}r_{i\text{S}}^2+\Delta Q_i\left(\frac{\Delta Q_i+2(Q_i^{\text{S}}-Z_i)}{r_{i\text{S}}}\right)\right\}+\frac{1}{2}\sum_{ij}Q_i^{\text{ASA-MT}}M_{ij}Q_j^{\text{ASA-MT}}, \tag{35}$$

and

$$E_{\text{xc}}^{\text{ASA-MT}}=\sum_i\{E_{\text{xc}}[\rho_i^{\text{S}}]+\rho^{\text{int}}\epsilon_{\text{xc}}(\rho^{\text{int}})(\Omega_i-\Omega_i^{\text{S}})+\mu_{\text{xc}}[\rho_i^{\text{S}}(r_{\text{S}})]\Delta Q_i\}. \tag{36}$$

The exchange-correlation potential, $\mu_{\text{xc}}^{\text{ASA-MT}}[\rho_i]=\delta E_{\text{xc}}^{\text{ASA-MT}}/\delta\rho_i$, is given by

$$\mu_{\text{xc}}^{\text{ASA-MT}}[\rho_i]=\mu_{\text{xc}}[\rho_i^{\text{S}}(r)]-\mu_{\text{xc}}[\rho^{\text{int}}]\Theta(r_{i\text{MT}}-r)\frac{\Omega_i-\Omega_i^{\text{S}}}{\Omega^{\text{int}}}+\mu'_{\text{xc}}[\rho_i^{\text{S}}(r)]\frac{\delta(r-r_{i\text{S}})}{4\pi r_{i\text{S}}^2}\Delta Q_i. \tag{37}$$

The self-consistent effective potential is then,

$$V_{i\mathrm{eff}}(r)=\frac{-2Z_i}{r}+\frac{2}{r}\int_0^r d^3\mathbf{r}'\rho^{\mathrm{S}}(r')+2\int_r^{r_{i\mathrm{S}}}d^3\mathbf{r}'\frac{\rho^{\mathrm{S}}(r')}{r'}-V_i^0-\left[V_1+\mu_{\mathrm{xc}}(\rho^{\mathrm{int}})\frac{(\Omega_i-\Omega_i^{\mathrm{S}})}{\Omega^{\mathrm{int}}}\right]\Theta(r_{i\mathrm{MT}}-r)+\mu_{\mathrm{xc}}[\rho_i^{\mathrm{S}}(r)]-\overline{V}$$
$$+\mu'_{\mathrm{xc}}[\rho_i^{\mathrm{S}}(r)]\frac{\delta(r-r_{i\mathrm{S}})}{4\pi r_{i\mathrm{S}}^2}\Delta Q_i, \tag{38}$$

where $\mu'_{\mathrm{xc}}(x)$ is the derivative of $\mu_{\mathrm{xc}}(x)$ with respect to x. This completes our derivation of the new energy functional. The last term in the effective potential involving a delta function at $r=r_{i\mathrm{S}}$ is neglected in our calculations. This term introduces a small change in the phase shifts of the corresponding spherical potentials.

In the calculations described below in which we used $r_{\mathrm{S}}=r_{\mathrm{WS}}$, we have $\Delta Q_i=q_i$, and the energy becomes

$$E[\rho]=\sum_{\epsilon<E_F}\epsilon-\sum_i\int d^3\mathbf{r}\,\rho_i^{\mathrm{S}}(r)V_{i\mathrm{eff}}(r)+\sum_i\left\{-\int_0^{r_{i\mathrm{WS}}}d^3\mathbf{r}\frac{\rho_i^{\mathrm{S}}(r)2Z_i}{r}-\frac{2Z_iq_i}{r_{i\mathrm{WS}}}+2\int d^3\mathbf{r}\,\rho_i^{\mathrm{S}}(r)\int^r d^3\mathbf{r}'\frac{\rho_i^{\mathrm{S}}(r')}{r'}\right\}$$
$$+\sum_i\left\{\frac{2q_iQ_i^{\mathrm{S}}+q_i^2}{r_{i\mathrm{WS}}}-\frac{6}{5}\frac{[\rho^{\mathrm{comp}}\Omega_i]^2}{r_{i\mathrm{WS}}}+4\pi\rho^{\mathrm{comp}}Q_i^{\mathrm{ASA\text{-}MT}}r_{i\mathrm{WS}}^2\right\}+\sum_i\{E_{\mathrm{xc}}[\rho_i^{\mathrm{S}}]+\mu_{\mathrm{xc}}(\rho_i^{\mathrm{S}}(r_{\mathrm{WS}}))q_i\}$$
$$+\frac{1}{2}\sum_{ij}Q_i^{\mathrm{ASA\text{-}MT}}M_{ij}Q_j^{\mathrm{ASA\text{-}MT}}, \tag{39}$$

where

$$V_{i\mathrm{eff}}(r)=\frac{2}{r}\int_0^r d^3\mathbf{r}'\rho^{\mathrm{S}}(r')+2\int_r^{r_{i\mathrm{WS}}}d^3\mathbf{r}'\frac{\rho^{\mathrm{S}}(r')}{r'}-\frac{2Z_i}{r}-V_i^0-V_1\Theta(r_{i\mathrm{MT}}-r)+\mu_{\mathrm{xc}}[\rho_i^{\mathrm{S}}(r)]-\overline{V} \tag{40}$$

and

$$Q_i^{\mathrm{ASA\text{-}MT}}=Q_i^{\mathrm{S}}+q_i+\rho^{\mathrm{int}}\Omega_i-Z_i. \tag{41}$$

We note that the ASA-MT approach actually represents a spectrum of approximations, according to the choice of r_{S}, since there is no restriction within the approach on the value of r_{S}. At one end of the spectrum, when $r_{\mathrm{S}}=r_{\mathrm{WS}}$, we can set $\Delta Q_i=0$ and have essentially a conventional ASA. In this case, the above derivation provides a variational energy functional for the ASA. It is interesting to note that from this derivation the zero scale of the energy in the ASA is not arbitrary, but is determined by Eq. (32), which in the ASA limit is simply the weighted average of the ASA potential at r_{WS} over all atoms. Only with this choice of the energy zero the variational property of the total energy is maintained.

III. APPLICATIONS

To test our theory, we calculated the second order elastic constants of Cu, Ni, and Fe. Although the ASA usually gives reasonable bulk moduli, it is known to give negative shear modulus for many systems. On the other hand, MT approximation typically gives fair elastic constants for transition metals.[5] We used the LSMS to calculate the equilibrium lattice constants, bulk modulus (at the experimental lattice constants), B_0, and the shear moduli at experimental lattice constants $C_{11}-C_{12}$, and C_{44}, for Cu, Ni, and Fe. The results are compared with the experimental values in Tables I, II, and III. As in a typical LDA calculation, the error in the bulk modulus varies from 10% to 50%. We have tabulated only the ASA-MT values of the bulk moduli because the ASA-MT agrees almost exactly with both the MT and ASA. The bulk modulus from ASA-MT for Fe appears to be surprisingly small. This is because at the experimental lattice constant the LDA gives a moment that is too large ($2.34\mu_B$ for the ASA-MT). At the ASA-MT equilibrium lattice constant, of 5.28 atomic units, the moment is $2.14\mu_B$ and the bulk modulus is 2.04 MBar. The lattice constant, moment, and bulk modulus agree very well with the MT results of Moruzzi *et al.*[12] The error in $C_{11}-C_{12}$ is reasonably small, while the calculated values for C_{44} are consistently off by almost a factor of two. The ASA-MT approach gives essentially the same values as the MT for $C_{11}-C_{12}$ and C_{44} for all materials, but the ASA gives mostly negative values, as expected. We speculate that the error in the ASA-MT values of C_{44} is due mostly to the nonspherical charge distribution, since a full potential method such as FLAPW usually gives much more accurate values for C_{44}.

Another test of the method is to calculate the vacancy formation energy. This is usually calculated using the ASA since the MT approximation does not treat the charge transfer associated with vacancies very well. We calculated the vacancy formation energy of Cu for an unrelaxed lattice,

TABLE III. Vacancy formation energy for Cu compared to the experimental value, and the calculated charge on the vacancy site. Note the calculated values are for unrelaxed lattices. The charge transfer in parentheses for the MT case is before the interstitial average is taken.

	Exp	MT	ASA	ASA-MT
Vacancy energy	1.3 eV	4.3 eV	2.8 eV	2.4 eV
Vacancy charge	—	−1.133 (−0.876)	−0.848	−0.857

using the MT approximation, the ASA, and the ASA-MT approach. The results are listed in Table III and are compared with the experimental value.[13] In this table we also compare the amount of charge on the vacancy site given by three methods. It is evident that the ASA and the ASA-MT results are very close in this case, and both are significantly better than the MT value. The remaining difference between the ASA-MT value and the experimental value is probably from nonspherical charge effects (especially the dipole terms) and energy due to lattice relaxation. The comparison of the vacancy charge highlights the fact that the ASA-MT approach yields a charge density that is significantly closer to the ASA charge density than the MT one.

ACKNOWLEDGMENTS

This work was supported by the Division of Materials Science Office of Basic Energy under Contract No. DE-AC05-96OR22464 with the Lockheed Martin Energy Research Corporation.

[1] Antonios Gonis, *Green Functions for Ordered and Disordered Systems* (Elsevier Science Publishers B. V., Amsterdam, 1992).

[2] For a useful discussion see W. H. Butler, in *Application of Multiple Scattering Theory to Materials Science*, edited by W. H. Butler, P. H. Dederichs, A. Gonis, and R. L. Weaver, MRS Symposia Proceedings No. 253 (Materials Research Society, Pittsburgh, 1992), p. 149.

[3] J.F. Janak, Phys. Rev. B **9**, 3985 (1974).

[4] O.K. Andersen, Phys. Rev. B **12**, 3060 (1975); H.L. Skriver, Phys. Rev. Lett. **39**, 1768 (1982); Phys. Rev. B **31**, 1909 (1985).

[5] G.Y. Guo and W.M. Temmerman, in *Application of Multiple Scattering Theory to Materials Science* (Ref. 2), p. 317.

[6] D.M.C. Nicholson *et al.*, Phys. Rev. B **50**, 14 686 (1994); Y. Wang *et al.*, Phys. Rev. Lett. **75**, 2867 (1995).

[7] S. Asano and J. Yamashita, J. Phys. Soc. Jpn. **30**, 667 (1971); N.E. Christensen and S. Satpathy, Phys. Rev. Lett. **55**, 600 (1985).

[8] L. Vitos, J. Kollár, and H.L. Skriver, Phys. Rev. B **55**, 13 521 (1997).

[9] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[10] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).

[11] E. Madelung, Phys. Z. **19**, 524 (1918).

[12] V.L. Moruzzi, J.F. Janak, and A.R. Williams, *Calculated Electronic Properties of Metals* (Pergamon Press, New York, 1978).

[13] F. Agullo-Lopez, C.R.A. Catlow, and P.D. Townsend, *Point Defects in Materials* (Academic Press, New York, 1988).

ERKCS30 (Atomic Mechanisms in Interface Sciences) and ERKCS41 (KC 02 01 03 0; Semiconductor Nanocrystals for CO_2 Fixation) - publications are not separated by FWP

Refereed Journal Publications

*Buczko, R., D. Duscher, S. J. Pennycook, and S. T. Pantelides," Excitonic Effects in Core Excitation Spectra," *Phys. Rev. Lett.* **85**, 2168 (2000).

*Buczko, R., S. J. Pennycook, and S. T. Pantelides, "Bonding Arrangements at the Si-SiO_2 and SiC-SiO_2 Interfaces and a Possible Origin of Their Contrasting Properties," *Phys. Rev. Lett.* **84**, 943 (2000).

†*Chen, C. L., Z. H. Zhang, H. H. Feng, G. P. Luo, S. Y. Chen, A. K. Heilman, W. K. Chu, C. W. Chu, H. J. Gao, B. Rafferty, S. J. Pennycook, Y. Liou, and F. A. Miranda, "Epitaxial Behavior and Interface Structures of Ferroelectric BSTO Thin Films," *Integr. Ferroelectr.* **29**, 237 (2000).

*Dai, S., M. C. Burleigh, Y. H. Ju, H.-J. Gao, J. S. Lin, S. J. Pennycook, C. E. Barnes, and Z. L. Xue, "Hierarchically Imprinted Sorbents for the Separation of Metal Ions," *J. Am. Chem. Soc.* **122**, 992 (2000).

†*Dai, S., Y. H. Ju, H.-J. Gao, J. S. Lin, S. J. Pennycook, and C. E. Barnes, "Preparation of Silica Aerogel Using Ionic Liquids as Solvents," *Chem. Commun.* **3**, 243 (2000).

*Dickey, E. C., X. Fan, and S. J. Pennycook, "Direct Atomic-Scale Imaging of Ceramic Interfaces," *Acta Mater.* **47**, 4061 (1999).

†*Duscher, G., J. P. Buban, N. D. Browning, M. F. Chisholm, and S. J. Pennycook, "The Electronic Structure of Pristine and Doped (100) Tilt Grain Boundaries in $SrTiO_3$," *Interface Sci.* **8**, 199 (2000).

*Fan, X., E. C. Dickey, P. C. Eklund, K. A. Williams, L. Grigorian, R. Buczko, S. T. Pantelides, and S. J. Pennycook, "Atomic Arrangement of Iodine Atoms Inside Single-Walled Carbon Nanotubes," *Phys. Rev. Lett.* **84**, 4621 (2000).

*Fan, X., E. C. Dickey, S. J. Pennycook, and M. K. Sunkara, "*Z*-Contrast Imaging and Electron Energy-Loss Spectroscopy Analysis of Chromium Doped Diamond-Like Carbon Films," *Appl. Phys. Lett.* **75**, 2740 (1999).

†*Gai, P. L., E. D. Boyes, C. B. Carter, D. J. H. Cockayne, L. D. Marks, and S. J. Pennycook, "Celebrations in Pioneering Electron Microscopy: A Symposium in Honor of Professor Archie Howie – Introduction," *Microscopy and Microanalysis* **6**, 281 (2000).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

*Gao, H. J., C. L. Chen, B. Rafferty, S. J. Pennycook, S. J. Luo, and C. W. Chu, "Atomic Structure of $Ba_{0.5}Sr_{0.5}TiO_3$ Thin Films on $LaA10_3$," *Appl. Phys. Lett.* **75**, 2542 (1999).

*Gao, H. J., G. Duscher, M. Kim, S. J. Pennycook, D. Kumar, K. G. Cho, and R. K. Singh, "Cathodoluminescent Properties at Nanometer Resolution Through Z-contract Scanning Transmission Electron Microscopy," *Appl. Phys. Lett.* **77**, 594 (2000)

*Gao, H. J., D. Kumar, K. G. Cho, P. H. Holloway, R. K. Singh, X. D. Fan, Y. Yan, and S. J. Pennycook, "Epitaxial Growth of Y2O3: Eu Thin Films On $LaAlO_3$," *Appl. Phys. Lett.* **75**, 2223 (1999).

*Gao, H. J., K. Sohlberg, Z. Q. Xue, H. Y. Chen, S. M. Hou, L. P. Ma, S. W. Fang, S. J. Pang, and S. J. Pennycook, "Reversible, Nanometer-Scale Conductance Transitions in an Organic Complex," *Phys. Rev. Lett.* **84**, 1780 (2000).

* Maiti, A., S. T. Pantelides, M. F. Chisholm, and S. J. Pennycook, "Damage Nucleation and Vacancy-Induced Structural Transformation in Si Grain Boundaries," *Appl. Phys. Lett.***75**, 2380 (1999).

†*Nellist, P. D. and S. J. Pennycook, "The Principles and Interpretation of Annular Dark-Field Z-Contrast Imaging," *Adv. Imag. Elect. Phys.* **113**, 147 (2000).

*Pennycook, S. J., M. F. Chisholm, Y. Yan, G. Duscher, and S. T. Pantelides, "A Combined Experimental and Theoretical Approach to Grain Boundary Structure and Segregation," *Physica B* **274**, 453 (1999).

†*Pennycook, S. J., B. Rafferty, and P. D. Nellist, "*Z*-Contrast Imaging in an Aberration-Corrected Scanning Transmission Electron Microscope," *Microsc. Microanal.* **6**, 343 (2000).

*Puretzky, A. A., D. B. Geohegan, X. Fan, and S. J. Pennycook, "Dynamics of Single-Wall Carbon Nanotube Synthesis by Laser Vaporization," *Appl. Phys. A - Mater.* **70**, 153 (2000).

*Puretzky, A. A., D. B. Geohegan, X. Fan, and S. J. Pennycook, "In Situ Imaging and Spectroscopy of Single-Wall Carbon Nanotube Synthesis by Laser Vaporization," *Appl. Phys. Lett.* **76**, 182 (2000).

†*Rafferty, B., S. J. Pennycook, and L. M. Brown, "Zero Loss Peak Deconvolution for Bandgap EEL Spectra," *J. Electron Microscopy* **49**, 517 (2000).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

*Sinnott, S. B., R. F. Wood, S. J. Pennycook, "Ab Initio Calculations of Rigid-Body Displacements at the (Sigma)5 (210) Tilt Grain Boundary in TiO_2," *Phys. Rev. B* **61**, 15645 (2000).

*Sohlberg, K., S. J. Pennycook, and S. T. Pantelides, "Explanation of the Observed Dearth of Three-Coordinated Al on _-Alumina Surfaces," *J. Am. Chem. Soc.* 121, 10999 (1999).

†*Sohlberg, K., S. J. Pennycook, and S. T. Pantelides, "The Bulk and Surface Structure of _-Alumina," *Chem. Eng. Commun.* **181**, 107 (2000).

†*Xin, Y., E. M. James, I. Arslan, N. D. Browning, S. J. Pennycook, and S. Sivananthan, "Atomic Resolution Z-Contrast Imaging of Semiconductors," *J. Electron Microsc.* **49**, 231 (2000).

*Xin, Y., E. M. James, I. Arslan, S. Sivananthan, N. D. Browning, S. J. Pennycook, F. Omnes, B. Beaumont, J.-P. Faurie, and P. Gibart, "Direct Experimental Observation of the Local Electronic Structure at Threading Dislocations in Metalorganic Vapor Phase Epitaxy Grown Wurtzite GaN Thin Films," *Appl. Phys. Lett.* **76**, 466 (2000).

†*Xu, Z. K., S. M. Gupta, D. Viehland, Y. F. Yan, and S. J. Pennycook, "Direct Imaging of Atomic Ordering in Undoped and La-doped $Pb(Mg_{1/3}Nb_{2/3})O_{-3}$," *J. Am. Ceram. Soc.* **83**, 181 (2000).

*Yan, Y., M. M. Al-Jassim, K. Matsuda, H. Tatsuoka, H. Kuwabara, and S. J. Pennycook, "The Mechanism for the High-Quality Single-Phase Growth of MnSi Films on Si (111) in the Presence of SB Flux," *Appl. Phys. Lett.* **75**, 2894 (1999).

*Yan, Y., M. F. Chisholm, S. J. Pennycook, and S. T. Pantelides, "Structures of Pure and Ca-segregated MgO (001) Surfaces," *Surf. Sci.* **442**, 251 (1999).

*Yan, Y. and S. J. Pennycook, "Structural Model for the $A1_{72}Ni_{20}Co_8$ Decagonal Quasicrystals," *Phys. Rev. B* **61**, 14291 (2000).

†*Yan, Y. and S. J. Pennycook, "Alloys—Atomic Structure of the Quasicrysta $A1_{72}Ni_{20}Co_8$," *Nature* **403**, 266 (2000).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

Excitonic Effects in Core-Excitation Spectra of Semiconductors

R. Buczko,* G. Duscher, S. J. Pennycook, and S. T. Pantelides
Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235
and Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831
(Received 21 March 2000)

Core-electron excitation spectra are used widely for structural and chemical analysis of materials, but interpretation of the near-edge structure remains unsettled, especially for semiconductors. For the important Si $L_{2,3}$ edge, there are two mutually inconsistent interpretations, in terms of effective-mass excitons and in terms of Bloch conduction-band final states. We report *ab initio* calculations and show that neither interpretation is valid and that the near-edge structure is in fact dominated by short-range electron-hole interactions even though the only bound excitons are effective-mass-like.

PACS numbers: 71.35.Gg, 71.20.Mq

Excitations of electrons from core levels to available empty states are the key ingredient of several powerful spectroscopies that probe the electronic and atomic structure of bulk solids, surfaces, and interfaces. The energy thresholds for such core excitations are unambiguous signatures of the chemical identity of the elements that are present [1], while small shifts in the thresholds are indicative of different bonding coordination or charge state [2]. Because core levels are dispersionless, the spectra reflect the available excited states. Interpretation of these spectra, however, has been a very difficult task because electron-hole correlations may produce bound excitons below the band edge and also distort the continuum spectrum. The task has been complicated by several factors: broadened excitons merge with continuum excitations so that they cannot be unambiguously distinguished, band-threshold calculations are not sufficiently accurate, and the inclusion of excitonic effects in theoretical spectra was until recently [3] pursued only in limiting cases [4–7].

In the early 1970s, calculated conduction-band densities of states of ionic solids were compared with measured x-ray absorption spectra by treating the band threshold as a free parameter. It was concluded that the spectra could be interpreted in terms of band final states without excitonic effects [8]. Subsequently, it was shown that band thresholds could be established by a combination of x-ray photoemission and optical absorption data, leading to conclusions that excitonic effects in fact dominate the spectra [9]. Since then, it appears that the freedom to adjust the band threshold has been reintroduced, leading to mixed results about the role of excitonic effects in both insulators and semiconductors [6,7,10–12]. Calculations that use the $Z + 1$ approximation for excitons (modeling a core hole as an additional proton on the nucleus, i.e., increasing by one the atomic number of the atom being excited) have also produced mixed results [6,7]. More recently, Shirley [3] reported calculations of x-ray spectra of several materials by treating electron-hole interactions explicitly using the Bethe-Salpeter equation and demonstrated that large excitonic effects are present.

For semiconductors, the role of excitons in core spectra has been particularly controversial. Though calculations including the effect of core-holes demonstrated substantial excitonic effects in several systems [13], the issue is far from settled. The best illustration of the problem is provided by the Si $L_{2,3}$ edge, namely, transitions from the $2p$ core level (which is split into $2p_{1/2}$ and $2p_{3/2}$ levels) that the spectrum is a superposition of two virtually identical spectra separated by a small energy). When Brown and Rustgi [14] first reported the x-ray absorption (XAS) spectrum in 1972, they compared it with the calculated total conduction-band density of states (DOS) and noted that the latter could not account for the observations. The DOS exhibits a slow rise to a broad peak ~2 eV above threshold, whereas the observed spectrum exhibits a very steep rise and then a relative flattening. The same year, Altarelli and Dexter (AD) [4] incorporated excitonic effects in the effective-mass approximation on the basis that a core hole in Si resembles a shallow P impurity and obtained a steep initial rise in agreement with the data. In 1989, Weng *et al.* [11] revisited the issue and pointed out that one should compute a projected DOS (PDOS) because, by symmetry, $2p$ electrons sample only the s-like and d-like components of the final states. They computed the projected DOS for Si and compared it with electron-energy-loss spectra (EELS) [12]. They concluded that they could account for the observed spectra without any excitonic effects. So far no reconciliation of these contradictory points of view has been proposed. In subsequent work, Batson [15,16] claimed that a model band structure combined with AD-like excitonic effects can be used to fit the experimental data with high accuracy.

In this paper we present a series of calculations in terms of which we show that both of the above conflicting interpretations of the Si $L_{2,3}$ spectra fail to capture the correct physics. We then report *ab initio* calculations of the actual core-excitation spectra with and without the effect of the core hole and establish that the near-edge structure of the observed spectra is dominated by short-range electron-hole effects produced by the "central cell" part of the

 0031-9007/00/85(10)/2168(4)$15.00 © 2000 The American Physical Society

electron-hole potential. Effective-mass excitonic effects due to the long-range screened Coulombic potential play a secondary role. Thus, the dominance of short-range excitonic effects is likely to occur in other semiconductors and, even more so, in insulators where screening is far less effective.

First, we address the two conflicting interpretations of the Si $L_{2,3}$ spectrum. We note that the PDOS used by Weng *et al.* [11,12] is in fact not very different from the total DOS used by Brown and Rustgi. In Ref. [11], Weng *et al.* arrived at their conclusion by comparing the PDOS only with EELS spectra which have a much larger intrinsic broadening than XAS spectra (part of it is instrumental and part is caused by smaller electron and hole lifetimes in an EELS experiment because of the much higher particle densities). The slow rise of the PDOS definitely cannot account for the steep rise in the XAS spectrum. Furthermore, in order to match the PDOS with the EELS spectrum, one must assume a band threshold at 99.0 eV, significantly lower than the XAS threshold of 99.8 eV, which agrees with the photoemission data of Himpsel *et al.* [2]. In Ref. [12], Weng *et al.* shifted the theoretical spectrum by a smaller amount and compared with higher-resolution data, but agreement between experiment and theory is not satisfactory for either the initial rise or the position of the first peak. Overall we conclude that the PDOS cannot provide a consistent interpretation of all the relevant data. As further confirmation of this conclusion we computed the actual x-ray absorption spectrum without electron-hole interactions, which is what PDOS is supposed to approximate. The calculations were performed using an all-electron energy-band code (full-potential linearized augmented plane wave method [17]). The resulting spectrum is shown in Fig. 1 using a threshold of 99.8 eV. The absolute magnitude of the absorption coefficient has not been adjusted. It is clear that spectra calculated without electron-hole interactions are in poor agreement with the experimental spectrum.

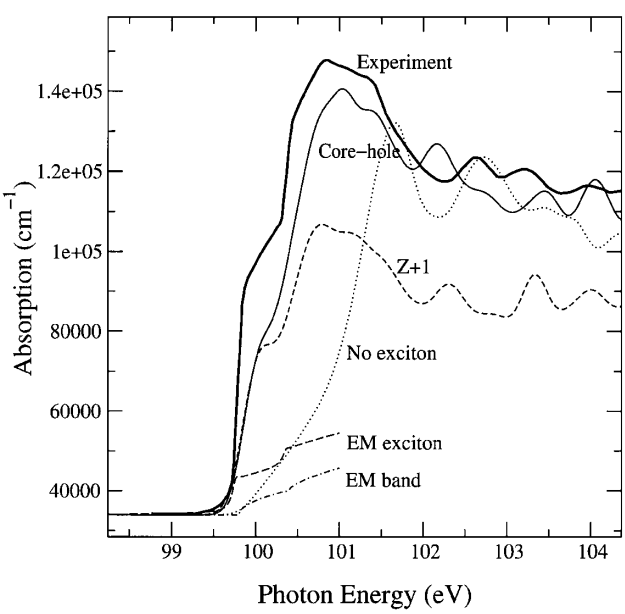


FIG. 1. Experimental and theoretical x-ray absorption spectra of the Si $L_{2,3}$ edge. Thick curve: experiment (Ref. [14]); solid curve: full calculation including electron-hole interactions with a self-consistent hole orbital; dashed curve: calculation using the $Z + 1$ model; dotted curve: calculation of excitation spectrum without electron-hole interactions; dash-dotted curve: effective-mass (parabolic band) spectrum without electron-hole interactions; long-dashed curve: effective-mass spectrum with hydrogenic excitonic effects as first done by Altarelli and Dexter (Ref. [4]).

Altarelli and Dexter [4] accounted for the Brown-Rustgi spectrum in terms of effective-mass excitons. The authors confined their attention to only the first 1 eV or so where the effective-mass approximation for the conduction-band density of states is valid. They adjusted the matrix element for transitions from the $2p$ core level to Bloch functions at the bottom of the conduction band and demonstrated a good fit to the near-edge experimental data. They argued that the resulting value of the matrix element is plausible. In Fig. 1 we show AD-like spectra using a matrix element extracted from the same calculation that produced the dotted curve. The dash-dotted line represents the case without effective-mass excitonic effects (i.e., the density of states multiplied by the matrix element) whereas the long-dashed line includes these effects. It is clear that, despite the initial steep rise, the AD spectrum cannot account for the observed spectrum. Note that one cannot invoke rescaling of the matrix element to bring the AD spectrum in agreement with the observed spectrum because the effective-mass excitonic correction is in principle tied to the initial rise of the calculated band spectrum (dotted curve in Fig. 1). Rescaling of this initial portion of the spectrum would require severe energy-dependent rescaling in the opposite sense at higher energies.

In principle, inclusion of electron-hole interactions in the excitation spectrum requires solution of the Bethe-Salpeter equation, as recently carried out both for valence and core-excitation spectra in a pseudopotential framework [3,18]. For core spectra, however, the hole is completely localized on a single atom so that the Bethe-Salpeter equation can be reduced to a single-particle calculation: For the Si $L_{2,3}$ edge, we impose a $1s^2 2s^2 2p^5$ configuration on one of the atoms of a large supercell and maintain that configuration throughout the self-consistency cycle of an all-electron calculation. We then calculate the corresponding one-electron excitation spectrum between the core state [19] and the available empty states that now reflect the effect of electron-hole interactions. The well-known $Z + 1$ model invokes an additional approximation in that it replaces the core hole by an extra proton on one of the nuclei in the supercell. Otherwise, the two types of calculations are identical in all respects. In both cases, there is an electron in the otherwise empty states above the valence bands that now reflect the presence of the hole or extra proton.

We have performed calculations using (a) a self-consistent hole in the core and (b) the $Z + 1$ model as described above. We employed supercells of 16, 32, and 64 atoms to test the convergence of the results. The differences between 32-atom and 64-atom results are very small. In addition to the spectrum, the calculation with the explicit core hole allows us to determine band threshold by subtracting the total energies of the initial (perfect crystal) and final (excited crystal) states. This threshold in principle would allow a positioning of the theoretical spectrum on the same energy axis as the experimental spectrum. The theoretical threshold is 100.4 eV, very close to the threshold extracted from both the XAS spectra [14] and photoemission spectra [2] experimental. Because of the theoretical error bar that may easily be of order 1 eV, we display the theoretical spectra using the experimental value for the threshold.

In Fig. 1 we show the results from a full-core-hole calculation and a $Z + 1$ calculation. We compare with the XAS spectra of Brown and Rustgi [14]. No adjustment has been made in the absolute value of the spectrum. The agreement between the full-core-hole calculation and experiment is indeed remarkable and is in sharp contrast to the other two approximations, namely, no electron-hole interactions and only effective-mass electron-hole interactions. The agreement is particularly good in the first two eV or so. The small-scale structure in the dropoff after the first peak is not reproduced in detail, which can be traced to numerical problems (convergence is very slow with number of **k** points in the Brillouin zone) and the well-known limitations of the local-density approximation in getting all the details of the conduction bands right. We further see that the $Z + 1$ approximation produces an excellent spectrum except for a scaling factor. Finally, the extended Coulomb tail, which is left out in a supercell calculation, in principle produces effective-mass excitons with very small binding energy and further steepening of the initial rise of the spectrum as it does in the AD effective-mass calculation. As Fig. 1 illustrates, any such additional effect will be only a small correction.

The net conclusion is unambiguous that the central-cell potential produced by the core hole has a substantial effect on the continuum final states, even though the long-range Coulombic tail produces only very shallow hydrogenic bound excitons. The origin of this effect can be traced to the fact that a short-range potential perturbs the conduction-band continuum by introducing broad resonances [20] (a similar effect is predicted for small quantum dots embedded in a bulk semiconductor [21]). Because the central-cell potential is relatively weak, these resonances do not descend into the band gap to produce deep states (a sulfur impurity, for which the central-cell potential is roughly double that of phosphorus, in fact causes an s-like deep state [20]). Thus, if it were not for the screened Coulombic tail of the potential, the extra electron of a P impurity would simply occupy the lowest conduction-band Bloch state that is largely unaffected by the central-cell potential. The screened Coulombic tail, however, binds it into hydrogenic effective-mass shallow levels. These results also suggest that core spectra from P impurities would result in strongly bound excitons below the continuum spectrum [22].

The all-electron calculations of core-excitation spectra are demanding even with high-performance parallel computers. Similar calculations would be even more demanding for SiO_2 and much worse for more complex systems such as the Si-SiO_2 and SiC-SiO_2 interfaces [23]. The preferred experimental method is EELS because of its spatial resolution [24]. Theoretical EELS spectra need not be as accurate as XAS spectra because of the increased broadening. For these reasons, we investigated the suitability of simply using projected densities of states of the relevant $Z + 1$ system instead of doing the full x-ray absorption calculation. One advantage of such an approach is the ability to use pseudopotential codes that are generally more efficient. The one ambiguity in such a calculation is the relative contributions of the s-projected and d-projected densities of states. Our full calculations for the Si $L_{2,3}$ spectrum in crystalline Si yield a s/d ratio of 1.1. We used this ratio with projected densities of states computed with a first-principles ultrasoft pseudopotential code [25]. The resulting spectrum, suitably broadened for an EELS spectrum is shown in Fig. 2. Agreement between experiment and theory is satisfactory.

In summary, we have demonstrated that short-range electron-hole interactions dominate the core-excitation spectrum of the $L_{2,3}$ edge in Si. These results have

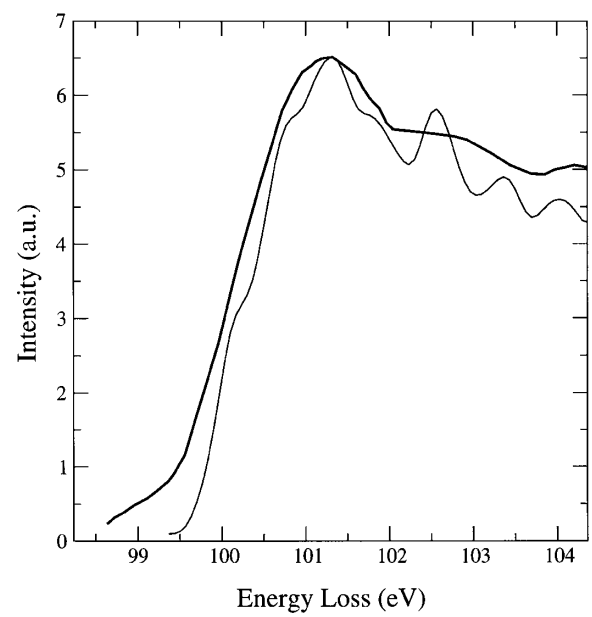


FIG. 2. Experimental and theoretical EELS spectra. Thick curve: experimental curve obtained with a VG Microscope HB501 UX using 100 kV acceleration voltage and collection semiangle of 13 mrad; solid curve: theoretical spectrum using the PDOS as described in the text.

profound implications for other semiconductors and especially for insulators for which dielectric screening is weaker, resulting in even stronger core-hole potentials. It is, therefore, imperative to avoid a free adjustment of theoretical band spectra on the energy axis, claiming agreement with data and hence no excitonic effects. Clear conclusions about the role of electron-hole interactions require at least $Z + 1$ calculations and/or determination of the band threshold using photoemission data.

This work was supported in part by AFOSR MURI Grant No. F-49620-99-1-0289, DARPA Grant No. MDA972-98-1-0007, EPRI Grant No. WO806905, NSF Grant No. DMR-9803768, the Division of Materials Sciences, U.S. Department of Energy under Contract No. DE-AC05-960R22464, and the William A. and Nancy F. McMinn Endowment at Vanderbilt University. Some of the calculations were done at the National Energy Research Scientific Computing Center, supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, and at the National Computational Science Alliance supported by the National Science Foundation.

*On leave from the Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland.

[1] R. F. Egerton, *Electron Energy-Loss Spectroscopy in the Electron Microscope* (Plenum Press, New York, 1996), 2nd ed.
[2] F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, and J. M. Woodall, Phys. Rev. B **38**, 6084 (1988).
[3] E. L. Shirley, Phys. Rev. Lett. **80**, 794 (1998).
[4] M. Altarelli and D. L. Dexter, Phys. Rev. Lett. **29**, 1100 (1972).
[5] H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B **24**, 6010 (1981).
[6] Th. Lindner, H. Sauer, W. Engel, and K. Kambe, Phys. Rev. B **33**, 22 (1986).
[7] P. Rez, J. Bruley, P. Brohan, M. Payne, and L. A. J. Garvie, Ultramicroscopy **59**, 159 (1995); L. A. J. Garvie, P. Rez, J. R. Alvarez, P. R. Buseck, A. J. Craven, and R. Brydson, Am. Mineral. **85**, 732 (2000).
[8] F. C. Brown, C. Gahwiller, A. B. Kunz, and N. O. Lipari, Phys. Rev. Lett. **25**, 927 (1970); A. B. Kunz, Phys. Rev. B **2**, 5015 (1970).
[9] S. T. Pantelides, Phys. Rev. B **11**, 2391 (1975).
[10] R. Brydson, H. Sauer, W. Engel, J. M. Thomas, E. Zeitler, N. Kosugi, and H. Kuroda, J. Phys. Condens. Matter **1**, 797 (1989).
[11] X. Weng, P. Rez, and O. F. Sankey, Phys. Rev. B **40**, 5694 (1989).
[12] X. Weng, P. Rez, and P. E. Batson, Solid State Commun. **74**, 1013 (1990).
[13] A. Zunger, Phys. Rev. Lett. **50**, 1215 (1983); S. B. Zhang, S.-H. Wei, and A. Zunger, Phys. Rev. B **52**, 13 975 (1995).
[14] F. C. Brown and O. P. Rustgi, Phys. Rev. Lett. **28**, 497 (1972).
[15] P. E. Batson, Phys. Rev. B **47**, 6898 (1993).
[16] P. E. Batson, J. Electron. Microsc. **49**, 267 (2000).
[17] P. Blaha, K. Schwarz, and J. Luitz, WIEN97, a full-potential linearized augmented plane wave package for calculating crystal properties (Technical University, Wien, Austria, 1999), ISBN 3-9501031-0-4, updated version of P. Blaha, K. Schwarz, P. Serantin, and S. B. Trickey, Comput. Phys. Commun. **59**, 399 (1990); the effective energy cutoff used in the calculations was 160 eV.
[18] L. X. Benedict, E. L. Shirley, and R. B. Bohn, Phys. Rev. Lett. **80**, 4514 (1998); M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998).
[19] The many-body, "initial state" is the perfect crystal. The "excited state" contains a core hole and an electron in the "conduction" states in the presence of the hole. The one-electron transition matrix element contains the perfect-crystal core state wave function and the conduction state wave function of the excited crystal.
[20] J. Bernholc, N. O. Lipari, S. T. Pantelides, and M. Scheffler, Phys. Rev. B **26**, 706 (1982).
[21] R. Buczko and F. Bassani, Phys. Rev. B **54**, 2667 (1996).
[22] S. T. Pantelides, Solid State Commun. **16**, 217 (1975).
[23] R. Buczko, S. J. Pennycook, and S. T. Pantelides, Phys. Rev. Lett. **84**, 943 (2000).
[24] G. Duscher, N. D. Browning, and S. J. Pennycook, Phys. Status Solidi a **166**, 327 (1998).
[25] These calculations were performed using VASP, developed at the Institute fuer Theretische Physik of the Technische Universitaet Wien; see G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993); **49**, 14 251 (1994); G. Kresse and J. Furthmuller, Comput. Mater. Sci. **6**, 15 (1996); Phys. Rev. B **54**, 11 169 (1996); the kinetic energy cutoff was set at 150 eV. 216 k-points were used in the irreducible wedge for k-space integration.

Bonding Arrangements at the Si-SiO_2 and SiC-SiO_2 Interfaces and a Possible Origin of their Contrasting Properties

Ryszard Buczko,* Stephen J. Pennycook, and Sokrates T. Pantelides
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831
and Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235
(Received 15 June 1999)

We report *ab initio* calculations designed to explore the relative energetics of different interface bonding structures. We find that, for Si (001), abrupt (no suboxide layer) interfaces generally have lower energy because of the surface geometry and the softness of the Si-O-Si angle. However, two energetically degenerate phases are possible at the nominal interface layer, so that a mix of the two is the likely source of the observed suboxide and dangling bonds. In principle, these effects may be avoidable by low-temperature deposition. In contrast, the topology and geometry of SiC surfaces is not suitable for abrupt interfaces.

PACS numbers: 68.35.Ct, 68.35.Bs, 81.65.Mq, 82.65.Dp

A key component of metal-oxide-semiconductor field-effect transistors (MOSFETs) is the semiconductor-oxide interface. The predominance of Si in microelectronics is due primarily to the properties of the Si-SiO_2 interface, which can easily be made very abrupt (minimal suboxide layer) and smooth (minimal steps) with a minimal density of point defects [1]. As scaling laws are pushing the technology to ultrathin SiO_2 layers, understanding and control of the interface on the atomic scale remain open challenges. For power MOSFETs, a semiconductor with a wider band gap is preferred. SiC, whose native oxide is also SiO_2, is one of the options, but efforts to develop SiC-based MOSFETs have been thwarted by poor-quality SiC-SiO_2 interfaces.

The abruptness and smoothness of the Si-SiO_2 interface are puzzling because SiO_2 is amorphous. In addition, O has a high solubility and diffusion constant in Si. In contrast, O is not known to be an abundant impurity in SiC, and a recent calculation finds that O has very low solubility in SiC [2]. Yet, SiC interfaces are generally found to be rough for MOSFETs. Photoemission [3] has provided information on bonding arrangements at the Si-SiO_2 interface while microscopy and theory led to several interface models [4–10]. Total-energy calculations have accounted for the observed low density of interface point defects [11] and have found that, during thermal oxidation, lateral growth is preferred [10]. Yet, the mechanisms that control abruptness and smoothness remain an open issue that is critical for the ultimate control of ultrathin gate oxides.

In this Letter, we address the mechanisms that control abruptness at the Si-SiO_2 and SiC-SiO_2 interfaces. In the case of the Si (001) surface, a set of judiciously chosen *ab initio* calculations suggest that abrupt interfaces generally have lower energy. The origin of this result can be traced to the softness of the Si-O-SiO_2 angle and the particular geometry of this surface, which imposes order in the nominal interface layer (NIL). These effects make suboxide bonds energetically costly, thus setting the stage for potentially perfect interfaces, except for entropic considerations that normally introduce defects (in this case, suboxide bonds) to lower the free energy at finite temperatures, as is true of "perfect" crystals. Unfortunately, however, two distinct but energetically degenerate ordered structures are possible at the NIL. Inevitably, the NIL of real Si-SiO_2 structures is made of two types of domains. It is these domain boundaries, plus entropy effects within domains, that are the likely cause of the observed suboxide [3] and dangling bonds [12]. Thus, theory suggests that the high quality of Si-SiO_2 interfaces can be traced to processing choices, made empirically, that enable large single domains to form. It may, therefore, be possible to further suppress suboxide bonds, steps, and dangling bonds with low-temperature nonequilibrium deposition techniques that avoid entropic effects. In contrast, in the case of SiC, neither the (0001) surface of the hexagonal phase nor the (001) surface of the cubic phase offer a suitable geometry and bond lengths for abrupt models. Geometry alone suggests that suboxide bonding is inevitable.

The calculations were carried out using state-of-the-art codes based on density functional theory, the local-density approximation for exchange correlation, ultrasoft pseudopotentials, and plane waves [13]. The energy cutoff for the plane waves was set at 380 eV, as determined by extensive convergence studies. Integrations over the Brillouin zone were done using the Pack-Monkhorst scheme with two k-points in the relevant irreducible wedge. The supercells differed in size as described below. All of them had 7–9 Si layers separating the SiO_2 layers. In all cases, all the atoms and the z dimension of the supercell were allowed to relax to equilibrium values.

Though several bonding arrangements have been proposed for the Si-SiO_2 interface, comparisons of their

 © 2000 The American Physical Society

relative total energies have not been possible so far. Most of the earlier work sought to match a Si substrate with a crystalline form of SiO_2. We approached the problem from a different perspective, namely, layer-by-layer deposition, leaving open the possibility that, after several layers, the film may turn out to be one of the known crystalline polymorphs. We emphasize at this point that we are describing a model construction process to obtain models whose relative total energies can be computed and compared. We will discuss the feasibility of achieving such a process in practice later in the paper.

It is well known that the (001) Si surface undergoes reconstruction by forming rows of dimers, eliminating half of the dangling bonds. Insertion of an O atom in a dimer leads to a stable configuration [6]. We investigated ordering of Si-O-Si bridges and found that rows [Fig. 1(a)] and a checkerboard pattern [Fig. 1(b)] have the same energy (the difference, 0.02 eV per Si-O-Si is smaller than the numerical error bar). The basic reason for this equality is that, in both patterns, every surface Si atom takes part in one Si-O-Si bridge and has one dangling bond. Since a Si atom cannot take part in bridges on both sides because of the bond length values, a random pattern of Si-O-Si bridges would lead to a highly defective interface. Domains separated by NIL domain boundaries or steps are the only alternative to a single pattern. We will first examine deposition of SiO_2 on each of the two distinct patterns and then examine domain boundaries.

The next step in our gedanken layer-by-layer deposition is to connect the remaining dangling bonds (one per atom) with longer bridges such as the following: (a) bridges containing one extra Si atom, namely, Si-O-Si-Si or Si-O-Si-O-Si; the first of these corresponds to suboxide bonding; (b) bridges containing two extra Si atoms, namely, Si-O-Si-O-Si-Si or Si-O-Si-O-Si-O-Si; again, the first of these corresponds to suboxide bonding.

Clearly, one can proceed with a random combination of such bridges to build an amorphous oxide [14]. In order to gain insights into energetically preferred bonding arrangements, we explored ordered structures containing only one type of long bridge at a time. We were able to form several superstructures. Each type of long bridge yields patterns that are reminiscent of bonding in different crystalline polymorphs. In addition, we inserted an additional one or two "bulk" layers within the SiO_2 film with bonding patterns reminiscent of β quartz, α cristobalite, and β tridymite. All three of them can be connected to the checkerboard pattern, but only tridymite can be connected to rows. Several of the resulting superstructures are shown in Fig. 2. We note that the ideal interface has a NIL in which all Si atoms are in the Si^{2+} oxidation state without any Si^{1+} or Si^{3+} states. In contrast, every suboxide bond,

FIG. 1. Ordered structures on a Si (100) surface with half-monolayer O coverage. (a) Rows and (b) checkerboard pattern of Si-O-Si bridges. All atoms have a dangling bond.

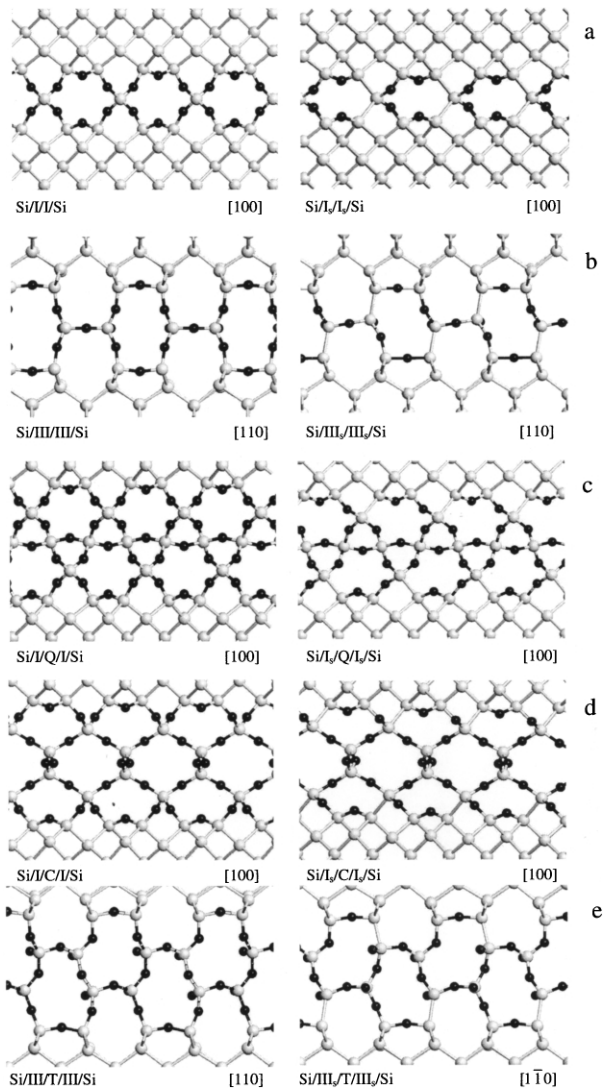


FIG. 2. Examples of Si-SiO_2-Si superstructures with one [(a), (b)] and two [(c), (d), (e)] oxide layers. The left panels are abrupt interfaces; the right panels have suboxide bonding.

TABLE I. Interface energies for studied structures. The numbers shown correspond to the sum of the two interface energies for each superstructure, normalized per 1×1 cell of the Si surface. For the structures with crystalline interlayer, the strain energy of the interlayer was subtracted as described in the text. Q stands for quartzlike bonding (the c axis of quartz lies along the [100] axis of Si). C stands for cristobalitelike bonding, and T stands for tridymitelike bonding. Interfaces I and II correspond to the checkerboard pattern of short bridges; III and IV to rows.

Superstructure		Interface energy (eV) 1 layer	2 layers	3 layers
Interfaces with suboxide layer	$Si/I_S/Q/I_S/Si$	2.1	2.2	2.0
	$Si/I_S/C/I_S/Si$		2.8	2.9
	$Si/III_S/T/III_S/Si$	2.0	2.0	1.8
Abrupt interfaces	Si/I/Q/I/S	1.3	1.4	1.4
	Si/I/C/I/Si		2.1	1.8
	Si/II/T/II/Si	1.7	1.7	1.8
	Si/III/T/III/Si	1.7	1.5	1.5
	Si/IV/T/IV/Si	1.4	1.5	1.4

as defined here, corresponds to a pair of Si^{1+} and Si^{3+} states, so that our model interfaces with 50% suboxide bonds correspond to a 1:1:1 ratio of oxidation states. This happens to be the observed ratio. We would need significantly larger supercells, beyond current capabilities, to study models with other ratios of oxidation states. Nevertheless, our calculations will be informative about the energetics and origins of suboxide bonding.

Calculation of the total energies of the superstructures described above enables us to extract and compare "interface energies." We first compute the total energy of the relaxed superstructure and then subtract $N\varepsilon + N'\varepsilon'$ where N and N' are the number of Si-Si and Si-O-Si bonds, respectively, in the superstructure, and ε and ε' are the corresponding energies per bond in perfect Si and SiO_2 crystals (β quartz, the lowest-energy form). In the cases that contain the extra SiO_2 interlayer(s), we subtracted the strain energy of the oxide, which was estimated by using suitably stressed perfect crystals (it ranges from 0.0 to 0.5 eV for one to three layers). This construction enables for the first time the extraction of the excess interface energy in different interface bonding structures.

The calculated interface energies are shown in Table I. We note that they are of the same order as those of grain boundaries [15], suggesting that these ordered interfaces are not unrealistic. Furthermore, the interface energy is always higher in all interfaces with suboxide bonding. Since the effect is true independent of whether the oxide is in compression (tridymitelike) or tension (cristobalitelike or quartzlike), we conclude that the result is quite general, though a truly amorphous structure may allow additional relaxations. There is a simple explanation for this result: Si-Si bonds cannot be bent very easily, whereas Si-O-Si bridges have a soft angle at the O atom and pivot quite freely about the O (Fig. 3). Thus, they can be both stretched and bent with minimal energy cost. For the model interfaces studied so far, the results of Table I show that insertion of O atoms in the suboxide Si-Si bonds always lowers the energy. We have, therefore, identified a mechanism that tends to make the Si-SiO_2 interface abrupt. We note further from Table I that suboxide bonding is most favored when the local SiO_2 structure is tridymitelike.

Before we consider interface domains and steps, we note from Table I that quartzlike and tridymitelike bonding are energetically preferred. Furthermore, quartzlike bonding entails tension along the c axis [parallel to the Si (100) direction], whereas tridymitelike bonding entails biaxial compression. The best choice, therefore, would be to have a checkerboard interface with a coexistence of small quartzlike and tridymitelike domains to counterbalance the overall strain. Because of the large strain in the bulk, identifiable phase domains are not likely, in favor of "intimate phase mixing," in effect amorphous bonding. Such cells can probably be constructed, but calculations would have to wait for even higher-performance computers.

We now consider the effect of interface domain boundaries and steps. The NIL atomic arrangements of domain boundaries that run parallel to the interface Si-O-Si bridges do not contribute any energy cost because every

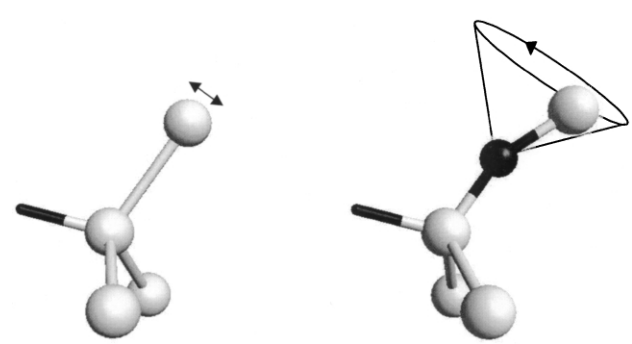


FIG. 3. Schematic illustration of the pivoting relaxation allowed by the Si-O-Si bridges. In addition, the Si-O-Si bridge can be easily stretched because the angle at the O atom is soft.

Si atom is still taking part in only one such bridge (see Fig. 1). Domain boundaries in any other direction, however, inevitably result in a $\ldots SLSXLSLS \ldots$ sequence of short (S) and long (L) bridges, where X denotes a frustrated bridge (must be both S and L to satisfy both sides). If an ideal NIL is retained, such domain boundaries are lines of dangling bonds. Using stick-and-ball models, we found that reconstruction is very easy if suboxide bonds are allowed. Calculations are not feasible, but it is safe to conclude that such "suboxide intrusions" would be energetically preferred over dangling bonds. Thus, whereas we found that suboxide bonds are energetically costly within domains, they form naturally at domain boundaries. Each such suboxide bond, however, does not necessarily correspond to a pair of Si^{1+} and Si^{3+} oxidation states any more. In agreement with the theory of Ref. [11], it seems that dangling bonds are accidental defects and can potentially be suppressed by nonequilibrium deposition.

Steps require even larger calculations to assess quantitatively, but they are fundamentally suboxide intrusions. The pivoting relaxation mechanism of Fig. 3 suggests that steps are not favored, but calculations are not feasible so that no definitive conclusion can be drawn. The results of Ref. [10] regarding lateral growth are consistent with this notion.

Putting all the above considerations together, we conclude the following: The Si (001) surface imposes order at the interface layer requiring either rows or a checkerboard pattern of Si-O-Si bridges. Furthermore, the unique contrast between Si-Si and Si-O-Si bonds exerts a powerful force to keep the interface abrupt and smooth. Suboxide intrusions corresponding to the observed Si^{1+} and Si^{3+} oxidation states are most likely within domains for entropic reasons and at domain boundaries between two energetically degenerate interface structures (rows and checkerboad pattern). It appears, therefore, that a suboxide layer can, in principle, be avoided if conditions are found to grow large interface domains at low temperatures, e.g., by atomic layer epitaxy.

We close the discussion of Si by commenting briefly on the feasibility of fabricating Si-SiO_2-Si superstructures that could lead to revolutionary Si-based devices, such as single-electron transistors. Our calculations suggest that metastable ordered structures can be formed layer by layer if one starts with a checkerboard pattern of Si-O-Si bridges. Thus, a Si-SiO_2-Si superstructure may be realizable. It is a major challenge to practitioners of atomic-layer epitaxy because the structures are metastable (O would like to penetrate the Si substrate as in thermal oxidation), requiring a nonequilibrium pathway, but the rewards would be enormous.

In contrast to Si (001), in the Si-terminated cubic SiC (001) surface, the distance between dangling bonds is $\sim$3 Å, so that one can easily insert an O atom between all pairs [i.e., have an O in both the "black" and "white" sites of the checkerboard pattern of Fig. 1(b)]. All dangling bonds are then saturated and growth is terminated. The same total-saturation effect is favored by geometry in the (0001) surface of hexagonal SiC and has been observed experimentally [16]. Careful examination of the hexagonal SiC (0001) surface revealed several other modes of total saturation, but no bonding arrangements that yield abrupt and smooth interfaces. Suboxide bonds seem inevitable. Such structures have been constructed, but the cells are too large for practical calculations. Visual inspection of the models suggests that elimination of the suboxide bonds is unlikely because of severe geometric constraints. We conclude that geometric consideration alone imply that suboxide bonds, either isolated or in the form of steps, are inevitable at SiC-SiO_2 interfaces.

We acknowledge useful discussions with L. C. Feldman and his group and with M. Di Ventra. This work was supported in part by the Division of Materials Sciences, U.S. Department of Energy under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp., the Office of Naval Research under Grant No. N00014-96-1-1286, a joint grant from the Advanced Research Projects Agency (MDA972-98-1-0007) and the Electrical Power Research Institute (W0806905), an Air Force Office of Scientific Research grant (F-49620-99-1-0289), and the William A. and Nancy F. McMinn endowment at Vanderbilt University.

*On leave from the Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland.

[1] S. T. Pantelides and M. Long, *The Physics of SiO_2 and its Interfaces* (Pergamon, New York, 1978).
[2] M. DiVentra and S. T. Pantelides, Phys. Rev. Lett. **83**, 1624 (1999).
[3] F. J. Himpsel *et al.*, Phys. Rev. B **38**, 6084 (1988); A. Pasquarello *et al.*, Phys. Rev. Lett. **74**, 1024 (1995).
[4] A. Ourmazd *et al.*, Phys. Rev. Lett. **59**, 213 (1987).
[5] H. Akatsu and I. Ohdomari, Appl. Surf. Sci. **41/42**, 357 (1989).
[6] A. B. Gurevich *et al.*, Phys. Rev. B **58**, R13 434 (1998).
[7] P. V. Smith and A. Wander, Surf. Sci. **219**, 77 (1089); Y. Miyamoto and A. Oshiyama, Phys. Rev. B **41**, 12 680 (1990); T. Hoshino *et al.*, Phys. Rev. B **50**, 14 999 (1994); N. A. Modine, G. Zumbach, and E. Kaxiras (unpublished).
[8] F. Herman and R. V. Kasowski, J. Vac. Sci. Technol. **19**, 395 (1985).
[9] A. Pasquarello, M. S. Hybertsen, and R. Car, Appl. Phys. Lett. **68**, 625 (1996); Phys. Rev. B **53**, 10 942 (1996); A. Pasquarello, M. S. Hybertsen, and R. Car, Nature (London) **396**, 58 (1998).
[10] H. Kageshima and K. Shiraishi, Phys. Rev. Lett. **81**, 5936 (1998).
[11] M. Ramamoorthy and S. T. Pantelides, Appl. Phys. Lett. **75**, 115 (1999).
[12] E. H. Poindexter *et al.*, J. Appl. Phys. **52**, 879 (1981).
[13] M. C. Payne *et al.*, Rev. Mod. Phys. **64**, 1045 (1992).
[14] Ohdomari, Mihara, and Kai [J. Appl. Phys. **60**, 3900 (1986)] used Si-O-Si-Si bridges to build a model interface.
[15] Y. Yan, M. F. Chisholm, G. Duscher, A. Maiti, S. J. Pennycook, and S. T. Pantelides, Phys. Rev. Lett. **81**, 3675 (1998).
[16] J. Bernhardt *et al.*, Appl. Phys. **74**, 1084 (1999).

Hierarchically Imprinted Sorbents for the Separation of Metal Ions

Sheng Dai,*,† M. C. Burleigh,†,§ Y. H. Ju,†,§ H. J. Gao,‡ J. S. Lin,‡ S. J. Pennycook,‡ C. E. Barnes,§ and Z. L. Xue§

Chemical Technology and Solid State Divisions
Oak Ridge National Laboratory
P.O. Box 2008, Oak Ridge, TN 37831-6181
Department of Chemistry, University of Tennessee
Knoxville, TN 37996-1600

Received September 2, 1999

Molecular imprinting involves arranging monomers of polymerization synthesis around a template molecule so that complexes between the monomer and template molecules are formed.[1−7] Subsequent polymerization of the monomer molecules results in trapping template molecules in a highly cross-linked amorphous polymer matrix. Extraction of the imprint molecules leaves a predetermined arrangement of ligands and a tailored binding pocket. Such imprinted polymers have been used to mimic antibody functions, to resolve racemates, and to separate mixtures of metal cations.[1] Thus far, the organization of precursor monomers has been achieved mainly in *inhomogeneous* organic polymer matrices, with little control over structural parameters, such as pore sizes and surface areas. Here, we report the first synthesis of imprinted hybrid sorbent materials with precise control of not only adsorption sites but also pore structures. The concepts behind our multilevel imprinting are as follows. Surfactant micelles and metal ions both act as templates in these *hierarchically imprinted* sorbents. The metal ion and the surfactant are removed from the silica matrix via acid leaching and ethanol extraction, respectively. This results in the formation of different-sized imprints within the silica matrix, each with a specific function. On the microporous level, the removal of the metal ion from the complex leaves cavities (1−3 Å) that exhibit ionic recognition. These pores give the sorbent enhanced selectivity for the given ion. On the mesoporous level, the removal of the surfactant micelles results in the formation of relatively large, cylindrical pores (diameter 25−40 Å) that give the gel an overall porosity which includes large surface areas and excellent metal ion transport kinetics.[8−16] This combination of high capacity and selectivity, coupled with fast kinetics,[17] makes these materials ideal candidates for many applications. Because the whole process utilizes *template or imprinting synthesis* twice and on different-length scales, it can be viewed as a hierarchical double-imprinting process.

Recently, Yang *et al.*,[18] and Schacht *et al.*[19] have demonstrated the hierarchical organization of three-dimensional structures patterned over multiple-length scales (10 nm to several micrometers) through the combination of micromolding and cooperative assembly of inorganic sol−gel species with amphililic triblock copolymers. Holland *et al.*[20] reported the synthesis of porous materials with simultaneous microscale and macroscale orders. To our knowledge, structures on the mesoscale and microscale with a hierarchical organization have yet to be synthesized. Our work reported here bridges this gap by combining molecular-scale imprinting and mesophase template synthesis with surfactants. These materials should find extensive applications in sensors and separations.

To illustrate the potential of this double-imprinting methodology, Cu^{2+}-selective, ordered mesoporous sorbents were chosen to demonstrate the basic principles. An extensive literature study[21,22] already exists concerning *singly* imprinted polymers for copper separations. The synthetic procedure to produce doubly imprinted sorbents is similar to that of co-assembling synthesis.[11−15] The bifunctional ligand used to complex the Cu^{2+} template is 3-(2-aminoethylamino) propyltrimethoxysilane, $H_2NCH_2CH_2$-$NHCH_2CH_2CH_2Si(OMe)_3$ (aapts). The ethylenediamine group in this ligand is known to form strong bidentate dative bonding with many metal ions. The imprinting complex precursor used in this study is $Cu(aapts)_2^{2+}$, which was synthesized by standard literature procedures.[23] The typical procedure involves mixing cetryltrimethylammonium bromide (CTAB), tetraethylothosilicate (TEOS), $Cu(aapts)_2^{2+}$, water, and base (NaOH) in relative molar ratios of 0.12:1.0:0.15:130:0.7. The mixture is heated and stirred at 100 °C for 24 h. The solid blue product is then recovered by filtration. This is then refluxed in ethanol/HCl to extract the surfactant templates with ethanol and to strip the copper-ion templates by protonation of the diamine functional groups. The final material is washed with copious amounts of 1 N HNO_3 to ensure the complete removal of the copper-ion templates. The control blank samples are prepared using an identical procedure without the addition of the Cu^{2+} template. Doubly imprinted sorbents containing the anionic sodium dodecylsulfate (SDS)[24] and the neutral dodecylamine (DDA)[25] surfactants were prepared using a similar protocol. The use of surfactants in the sorbent synthesis is crucial. The sorbents *singly* imprinted with $Cu(aapts)_2^{2+}$ all have very small surface areas (<20 m^2/g), the Cu^{2+}-uptake capacity is one-fifth of those of the doubly imprinted sorbents.

Doubly imprinted and control blank functionalized mesoporous silica samples prepared by means of this procedure have surface areas in range of 200−600 m^2/g.[26] Although the pore-size

* To whom all correspondence should be addressed.
† Chemical Technology Division, Oak Ridge National Laboratory.
‡ Solid State Division, Oak Ridge National Laboratory.
§ University of Tennessee.

(1) Wulff, G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1812−1832.
(2) Shea, K. J. *Trends Polym. Sci.* **1994**, *2*, 166−173.
(3) Mosbach, K. *Trends Biochem. Sci.* **1994**, *19*, 9.
(4) Steinke, J.; Sherrington, I. R.; Dunkin, D. C. *Adv. Polym. Sci.* **1995**, *123*, 81−125.
(5) *Molecular and ionic recognition with imprinted polymers*; Bartsch, R. A., Maeda, M., Eds; American Chemical Society: Washington, DC, 1998.
(6) Shi, H.; Tsai, W. B.; Garrison, M. D.; Ferrari, S.; Ratner, B. D. *Nature* **1999**, *398*, 593−597.
(7) D'Souza, S. M.; Alexander, C.; Carr, S. W.; Waller, A. M.; Whitcombe, M. J.; Vulfson, E. N. *Nature* **1999**, *398*, 312−316.
(8) Kresge, C. T.; Leonowicz, M. E.; Roth, W. J.; Vartuli, J. C.; Beck, J. S. *Nature* **1992**, *359*, 710−712.
(9) Yang, P.; Zhao, D.; Margolese, D. I.; Chmelka, B. F.; Stucky, G. D. *Nature* **1998**, *396*, 152.
(10) Tanev, P. T.; Pinnavaia, T. J. *Science* **1995**, *267*, 865−867.
(11) Hall, S. R.; Fowler, C. E.; Lebeau, B.; Mann, S. *Chem. Commun.* **1999**, 201−202.
(12) Brown, J.; Mercier, L.; Pinnavaia, T. J. *Chem. Commun.* **1999**, 69−70.
(13) Burkett, S. L.; Sims, S. D.; Mann, S. *Chem. Commun.* **1997**, 1769−1770.
(14) Macquarrie, D. J. *Chem. Commun.* **1996**, 1961−1962.
(15) Lim, M. L.; Blanford, C. F.; Stein, A. *J. Am. Chem. Soc.* **1997**, *119*, 4090−4091.
(16) Yang, H.; Kuperman, A.; Coombs, N.; Mamiche-Afara, S.; Ozin, G. A. *Nature* **1996**, *379*, 703−705.
(17) See Supporting Information for metal ion adsorption kinetics.
(18) Yang, P. D.; Deng, T.; Zhao, D. Y.; Feng, P.; Pine, D. S.; Chmelka, B. F.; Whitesides, G. M.; Stucky, G. D. *Science* **1998**, *282*, 2244−2246.
(19) Schacht, S.; Huo, Q.; Voigt-Martin, I. G.; Stucky, G. D.; Schuth, F. *Science* **1996**, *273*, 768−771.
(20) Holland, B.T.; Abrams, L.; Stein, A. *J Am Chem. Soc.* **1999**, *121*, 4308−4309.
(21) Kuchen, W.; Schram, J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1695−1679.
(22) Uezu, K.; Yoshida, M.; Goto, M.; Furusaki, S. *CHEMTECH* April, 12−18, 1999.
(23) De, G. T.; Epifani, M.; Licciulli, *J. Non-Cryst. Solids* **1996**, *201*, 250−255.
(24) Yada, M.; Ohya, M.; Machida, M.; Kijima, T. *Chem. Commun.* **1998**, 1941−1942.
(25) Karakassides, M. A.; Fournaris, K. G.; Travlos, A.; Petridis, D. *Adv. Mater.* **1998**, *10*, 483.
(26) See Supporting Information.

10.1021/ja993168x CCC: $19.00 © 2000 American Chemical Society
Published on Web 01/19/2000

Table 1. Competitive Loading of M_1 (Cu^{2+}) and M_2 (Zn^{2+}) by Copper-imprinted and Control Blank Mesoporous Sorbents Made with the Precursor Silica and Functional Ligand Molar Ratio 1.0:0.15 at pH 5.0 (Acetic Acid/Sodium Acetate Buffer)

type	solution[a] Cu (M)	solution[a] Zn(M)	% Cu Abs[d]	% Zn Abs[d]	Cu K_d (mL/g)	Zn K_d (mL/g)	k	k'
nonimp-aapts (CTAB)[b]	0.001	0.001	99.66	99.77	29000	44000	0.66	
imp-aapts (CTAB)[c]	**0.001**	**0.001**	**99.17**	**10.13**	**11000**	**67**	**160**	**240**
nonimp-aapts (CTAB)	0.001	0.01	99.67	55.81	30000	130	240	
imp-aapts (CTAB)	**0.001**	**0.01**	**99.17**	**10.13**	**13000**	**11**	**1100**	**4.6**
nonimp-aapts (SDS)	0.001	0.001	96.07	44.59	2400	71	34	
imp-aapts (SDS)	**0.001**	**0.001**	**99.56**	**13.31**	**23000**	**15**	**1500**	**44**
nonimp-aapts (SDS)	0.001	0.01	93.37	10.64	1400	10	140	
imp-aapts (SDS)	**0.001**	**0.01**	**99.98**	**2.40**	**82000**	**2.5**	**33000**	**240**
nonimp-aapts (DDA)	0.001	0.001	99.70	90.40	32000	940	34	
imp-aapts (DDA)	**0.001**	**0.001**	**99.60**	**40.00**	**25000**	**66**	**370**	**11**
nonimp-aapts (DDA)	0.001	0.01	96.00	18.00	2400	22	109	
imp-aapts (DDA)	**0.001**	**0.01**	**99.50**	**12.20**	**19000**	**14**	**1300**	**12**

[a] Initial solutions. [b] Nonimp-aapts (CTAB) = nonimprinted sorbent synthesized with CTAB surfactant. [c] Imp-aapts (CTAB) = imprinted sorbent synthesized with CTAB surfactant. [d] Percentage metal ion absorbed (± 0.01).

distributions of the imprinted and control blank samples are very similar, the surface areas of the imprinted samples are less than those of the control blanks by 10−20%. Small-angle X-ray scatterings show a peak around $Q = \sim 1.2-2.4\ nm^{-1}$, which agrees with the corresponding pore size determined by nitrogen adsorption experiments.[26] The close match between the UV−vis spectrum of $Cu(aapts)_2^{2+}$ in methanol and that of $Cu(b\text{-}aapts)_2^{2+}$ (b-aapts = aapts ligand covalently bonded to silica) indicates that the stereochemical environments of the copper ion in the two systems are similar and that the complex is doped into the bulk silica matrix.[26] The absorption peak position is consistent with the fact that Cu^{2+} is coordinated with 4-amine ligands since both 3- and 2-amine-coordinated copper complexes have absorption peaks in much longer wavelength.[27]

To test the selectivity of our new sorbents, we conducted competitive ion-binding experiments using aqueous Cu^{2+}/Zn^{2+} mixtures and a typical batch procedure.[28] This system constitutes one of the most stringent tests for ion-binding selectivity because both ions have identical charge, are of similar size, and exhibit high affinities for amine ligands. If a sorbent is simultaneously exposed to mixtures of different metal ions, preferential binding of those ions with the highest binding affinity will be observed. The selectivity coefficient, k, for the binding of a specific metal ion in the presence of competitor species can be obtained from equilibrium binding data according to methods given previously.[21]

Table 1 summarizes values for the distribution constant (K_d), selectivity coefficient of the sorbent toward Cu^{2+} (k), and the relative selectivity coefficient (k') obtained in these competitive ion binding experiments between zinc and copper ions. A comparison of the k values for the Cu^{2+}-imprinted sorbents with the corresponding control samples shows a significant increase in k for Cu^{2+} through imprinting, with the largest k values over 10 000 and the largest k' over 200. We have recently reported a synthesis of imprint-coated sorbents with the largest k value around 90 and the largest k' value around 40, which are, to our knowledge, the highest k and k' values currently achieved for molecular imprinting of metal ions.[29] The new double-imprinting approach offers a significant improvement in selectivity.

In conclusion, a novel design for template-selective recognition sites on mesoporous sorbents through hierarchical double imprinting is described. Unlike the mesoporous sorbents synthesized by surface functionalization,[30,31] precise control of the stereochemical arrangement of ligands on the adsorption sites can be achieved with this new methodology. Surface areas, controllable pore sizes, structural rigidities, and thermal stabilities achieved with the current silica-based imprinted sorbents are superior to those of imprinted organic polymers. This methodology has resulted in a new class of ordered mesoporous sorbents with molecular recognition capabilities. In addition to selective sorbents, the design principles illustrated by these results may lead us to new applications in areas such as chemical sensors and catalysis for these *hierarchically imprinted* materials.

Acknowledgment. This work was conducted at the Oak Ridge National Laboratory and supported by the Division of Chemical Sciences, Office of Basic Energy Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. and the Environmental Management Science Program (EMSP), U.S. Department of Energy.

Supporting Information Available: Schematic representation of the double imprinting process. UV−vis spectra of $Cu(aapts)_2^{2+}$ in methanol and that of $Cu(b\text{-}aapts)_2^{2+}$ (b-aapts = aapts ligand covalently bonded to silica). Adsorption kinetics curve of Cu^{2+} on Cu^{2+} imprinted dodecylamine/aapts sorbent. Small-angle X-ray scatterings of hierarchially imprinted sorbents. Surface areas and porosity data of imprinted and nonimprinted sorbents (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA993168X

(27) Baker, A. *J. Chem. Edu.* **1998**, *75*, 98.
(28) In all batch experiments, 0.1 g of each sample was equilibrated with 10 mL of the buffer solution in stoppered plastic vials, and these mixtures were stirred for 1 h at room temperature. The uptakes of Cu^{2+} and Zn^{2+} by sorbents were measured with a Perkin-Elmer Plasma 400 ICP/AE spectrophotometer.
(29) Dai, S.; Burleigh, M. C.; Shin, Y. S.; Morrow, C. C.; Barnes, C. E.; Xue, Z. L. *Angew. Chem., Int. Ed.* **1999**, *38*, 1235−1239.
(30) Feng, X.; Fryxell, G. E.; Wang, L. Q.; Kim, A. Y.; Liu, J.; Kemner, K. M. *Science* **1997**, *276*, 923−926.
(31) Mercier, L.; Pinnavaia, T. J. *Adv. Mater.* **1997**, *9*, 500−504.

PERGAMON

PII: S1359-6454(99)00266-9

Acta mater. Vol. 47, Nos 15, pp. 4061–4068, 1999
Published by Elsevier Science Ltd
On behalf of Acta Metallurgica Inc.
Printed in Great Britain
1359-6454/99 $20.00 + 0.00

DIRECT ATOMIC-SCALE IMAGING OF CERAMIC INTERFACES

E. C. DICKEY[1], **X. FAN**[1, 2] **and S. J. PENNYCOOK**[2]†

[1]Department of Chemical and Materials Engineering, University of Kentucky, Lexington, KY 40506-0046, U.S.A. and [2]Solid State Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, U.S.A.

Abstract—Understanding the atomic structure and chemistry of internal interfaces is often critical to developing interface structure–property relationships. Results are presented from several studies in which Z-contrast scanning transmission electron microscopy (STEM) and electron energy loss spectroscopy (EELS) have been employed to solve the atomic structures of oxide interfaces. The Z-contrast imaging technique directly reveals the projected cation sublattices constituting the interface, while EELS provides chemical and local electronic structure information. Because Z-contrast imaging and EELS can be performed simultaneously, direct correlations between structure and chemistry can be made at the atomic scale. The utility of Z-contrast imaging and EELS is demonstrated in three examples: a ZrO_2 24° [100] symmetric tilt grain boundary, a NiO–cubic ZrO_2 eutectic interface and a Ni–cubic ZrO_2 metal–ceramic interface. The power and versatility of Z-contrast and EELS for solving interface structures in oxide systems is clearly demonstrated in these three material systems. *Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.*

Keywords: Ceramic; Interface; Scanning transmission electron microscopy (STEM); Electron energy loss spectroscopy (EELS)

1. INTRODUCTION

The structure and chemistry of interfaces are known to affect the mechanical and electrical behaviors of oxide ceramics [1, 2]. It is therefore useful to obtain experimental data regarding the atomic structure and chemistry of these interfaces that can be used as the basis of atomistic simulations [3, 4]. Z-contrast STEM imaging has become increasingly popular among materials scientists for structure determination because the images are more directly interpretable than conventional high-resolution TEM images. Z-contrast imaging is an incoherent imaging process; thus the phase ambiguity inherent to HREM images is removed. Moreover, because the high-angle, or Rutherford scattered, electrons are used for image formation, there is chemical sensitivity in the images. For oxide materials, Z-contrast imaging has been a useful technique for understanding interface structures and defects. In particular, Z-contrast images directly reveal the projected cation sublattices of the material. While almost no information about the oxygen positions is contained in the Z-contrast images, positions can often be inferred or EELS can provide some information regarding the local coordination around oxygen atoms. One caveat for obtaining atomic-scale structural data by Z-contrast imaging, or in fact by any electron imaging technique, is that the two phases constituting the interface must have crystallographic orientation relationships that lead to the parallel alignment of low-index zone axes. Z-contrast imaging is particularly sensitive to crystal tilt since it necessitates channeling conditions. If, however, the two crystals are well oriented, then atomic-scale images and spectra can be obtained along the common zone axes of the two crystals.

Below we present case studies from three classes of oxide interfaces: (1) a grain boundary, (2) an oxide–oxide heterophase interface and (3) a metal–oxide heterophase interface. The three examples have a common attribute in that they all contain yttria-stabilized cubic ZrO_2 as one of the constituent phases (all references to ZrO_2 below are to cubic ZrO_2). The grain boundary is a ZrO_2 24° symmetric tilt [100] grain boundary. The NiO–ZrO_2 interfaces are formed by directional solidification of a NiO–ZrO_2 eutectic, and the Ni–ZrO_2 interfaces are formed by the reduction of the NiO–ZrO_2 eutectic. In all three cases, atomic-scale models of the interfaces are generated from Z-contrast images.

2. EXPERIMENTAL TECHNIQUES

All Z-contrast imaging studies were carried out on the 300 kV VG Microscopes HB603U dedicated STEM ($C_s = 1$ mm) at Oak Ridge National Laboratory. In the Z-contrast imaging technique, transmitted electrons are detected by a high-angle annular dark field detector. Although the electron scattering process in STEM has both coherent and

†To whom all correspondence should be addressed.

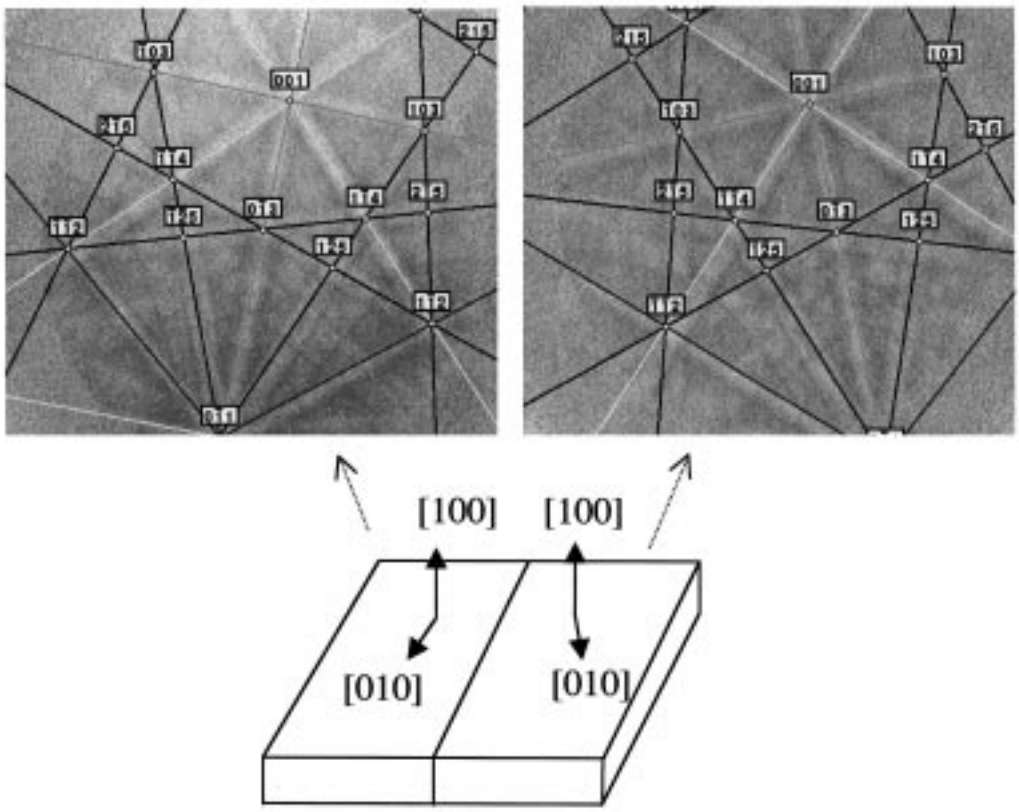


Fig. 1. Electron backscattered diffraction patterns from a 24° symmetric tilt [100] ZrO_2 bicrystal.

incoherent (thermal diffuse scattering) components, a transversely incoherent image can be produced by the large-angle annular detector which effectively breaks the transverse coherence of the signal by averaging over the interference fringes in the transverse plane [5, 6]. Because the signal is incoherent, the intensity in a Z-contrast image, $I(\mathbf{R})$, is given by the convolution of the incident probe intensity, $P^2(\mathbf{R})$, with the crystal object function, $O(\mathbf{R})$:

$$I(\mathbf{R}) = O(\mathbf{R}) * P^2(\mathbf{R}).$$

Even under dynamical diffraction conditions, it has been shown that s-type Bloch states, which are tightly bound to individual atomic columns, are predominantly responsible for the image intensity [7–9]. Consequently, the object function is localized on atomic columns. It should also be noted that the cross sections for scattering to high angles are much greater for higher atomic number elements so that the Z-contrast image intensity scales with atomic number, producing peaks of intensity at the cation sites while yielding almost no intensity at low-Z column positions [6].

Maximum entropy reconstructions were performed to quantify atomic positions from the Z-contrast images [10, 11]. Using a Lorentzian profile for the probe function, $P^2(\mathbf{R})$, the object function, $O(\mathbf{R})$, corresponding to cation positions was reconstructed by maximum entropy image analysis [12, 13]. The resulting object function was used for quantitative analysis. Note that all object functions presented below are convoluted with a Gaussian for better visibility.

To assess the positional accuracy of the technique, images of the perfect crystal, far from any defect, were analyzed. The object function positions were compared to the known atomic column positions based on the crystal structure and a standard deviation (σ_d) between the two positions was calculated. Positional accuracies do depend on signal-to-noise ratios in the image and therefore vary with imaging conditions and the scattering power of the material [13]. Reference [13] gives a thorough review of the factors that affect the accuracy of reconstructed Z-contrast images. In $SrTiO_3$ the $3\sigma_d$ for Sr columns is 0.019 nm and for Ti columns (having less scattering power) is 0.024 nm [13]. In the [112] NiO lattice projection, discussed below, we find that object functions derived from images recorded at a magnification of 10^7 times yield positional accuracies ($3\sigma_d$) of 0.021 nm far away from the interface. Although the ultimate accuracy of reconstructed Z-contrast images is consistently found to be on the order of 0.02 nm for cation col-

umns, STEM images, because they are serially acquired, are susceptible to mechanical and electronic instabilities. Any such perturbations will, of course, distort the lattice image and degrade the positional accuracy.

To characterize the chemistry of the interfaces, EEL spectra were collected on a 100 kV HB501UX STEM. Because an annular dark field image may be acquired simultaneously, the probe could be positioned accurately to provide EEL spectra from well-defined areas of the specimen. The energy resolution of the parallel energy loss spectrometer, as determined by the full-width half-maximum of the zero-loss peak, was 1.1 eV and the dispersion was 0.303 eV/channel over 385 channels. Each spectrum was corrected for the gain variation across the detector array and the background was fitted to a power law over a 50 eV window preceding the edge-onset and subtracted from each spectrum.

The 24° symmetric tilt [100] ZrO_2 bicrystal used in this study was purchased in bulk from Shinkosha Co. Ltd, Japan. Backscattered electron diffraction patterns shown in Fig. 1 illustrate and confirm the orientation relationship between the two crystals. Specimens were prepared normal to the tilt axis, along [100].

The NiO–ZrO_2 eutectic interfaces were produced by directional solidification as described elsewhere [14, 15]. Each grain within the bulk eutectic sample contains lamellae of the two phases which have well-defined crystallographic orientation relationships:

$$[1\bar{1}0]_{NiO}//[100]_{ZrO_2}//\text{growth direction}$$

$$[111]_{NiO}//[010]_{ZrO_2}.$$

Cross-sectional samples were prepared from two orthogonal directions so that the three-dimensional structure of the interface could be examined.

The Ni–ZrO_2 interfaces were formed by reducing the NiO–ZrO_2 eutectics [16]. The reductive phase transformation is electrochemical in nature and results in oxygen removal from NiO to the outside gas phase through the fast-ion conducting ZrO_2. The Ni remains well aligned crystallographically with respect to the ZrO_2 although a twin of the cation sublattice results from the phase transformation (see Ref. [17] for details). The final orientation relationship between the Ni and ZrO_2 is

$$[\bar{1}10]_{Ni}//[100]_{ZrO_2}$$

$$[111]_{Ni}//[010]_{ZrO_2}.$$

Cross-sectional samples were prepared along the $[\bar{1}10]_{Ni}//[100]_{ZrO_2}$ orientation.

All of the TEM samples were prepared by mechanically thinning interface cross sections to approximately 10 μm. For the heterophase interfaces, it was crucial that the samples be <10 μm thick before ion milling because of the differential thinning rates of the two phases. The cross sections were subsequently ion milled to electron transparency with 5 keV Ar^+ ions at liquid nitrogen temperature. To remove any amorphous surface layers, the foils were finally milled with 3 keV Ar^+ ions.

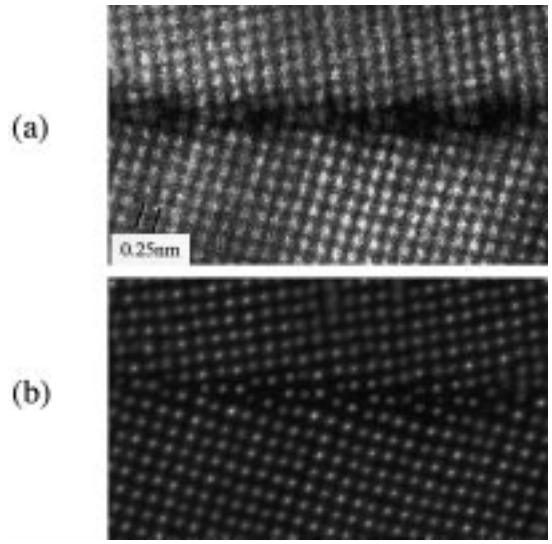


Fig. 2. Atomic-scale image of 24° symmetrical [100] tilt cubic-ZrO_2 boundary: (a) Z-contrast STEM image; (b) object function derived from maximum entropy reconstruction.

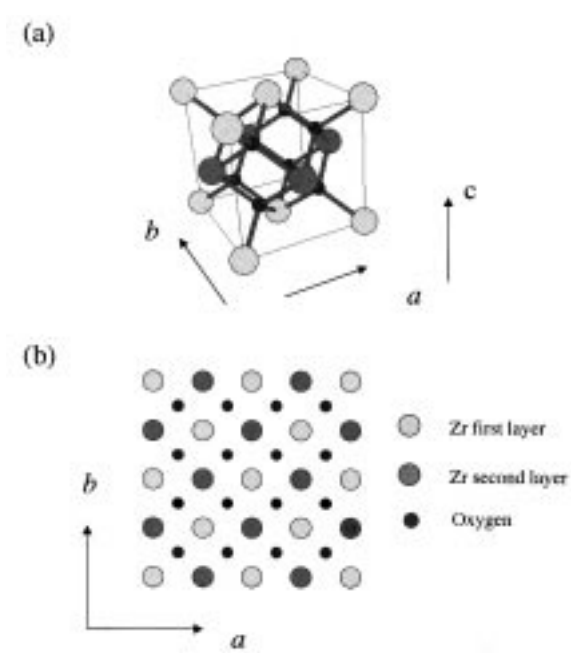


Fig. 3. Unit cell of ZrO_2: (a) perspective view; (b) [001] projection. Notice that [001] has two different Zr layers a half unit cell apart. We distinguish them by two shades of gray.

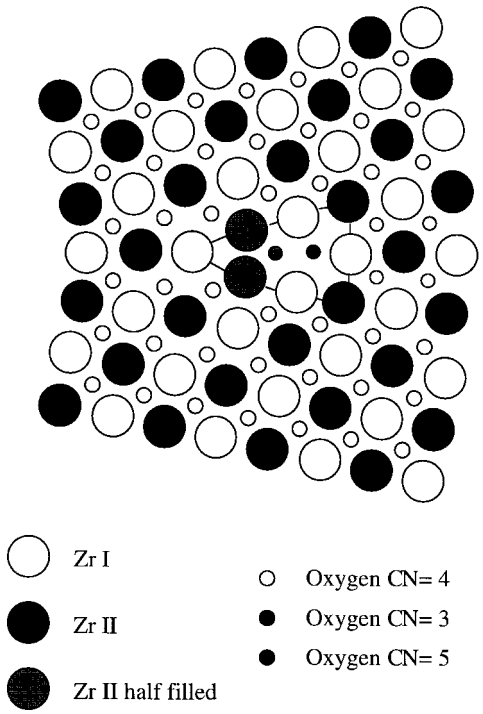


Fig. 4. Boundary structure model determined from Z-contrast image. Cation sites are determined directly by maximum entropy, oxygen sites are assumed. The solid lines outline the defect structural unit which is stoichiometric and therefore uncharged.

3. RESULTS AND DISCUSSION

3.1. 24° ZrO_2 tilt boundary

The Z-contrast image of the 24° symmetric tilt [100] ZrO_2 grain boundary is shown in Fig. 2 along with the corresponding object function obtained by maximum entropy reconstruction. The white dots in the image correspond to Zr columns, or the cation sublattice. Within the field of view, the 24° tilt boundary is composed of a periodic array of a basic structural unit (these are indicated by the open circles on Fig. 2). Since the image is only a representation of the projected structure, it is important to appreciate the cubic-ZrO_2 crystal structure (fluorite structure) in three dimensions. Figure 3 shows a ZrO_2 unit cell in which the Zr atoms are distinguished by their position along the [100] direction, the beam direction. The Zr atoms denoted in dark gray are a half unit cell below the Zr atoms in light gray.

Based on the Z-contrast images, a structural model of the grain boundary was developed and is shown in Fig. 4. Between each structural unit, there is a continuity of the (200) planes across the boundary indicating no in-plane rigid body translation. Within the defect structural unit, two of the Zr columns (both having Zr atoms at the same depth) are very close to each other in the boundary core. Electrostatically, this situation would be energetically unfavorable because of the close proximity of the Zr cations. However, also note that the intensities of these two columns in Fig. 2 are lower than those of the other Zr columns. The lower intensities can be interpreted as partially occupied Zr columns. Partial occupancy would allow the Zr atoms in the two adjacent columns to stagger along the beam direction and thus avoid close cation–cation positions. Similar observations have been made in other grain boundary studies including those of $SrTiO_3$ [18, 19] and $YBa_2Cu_3O_{7-x}$ [20]. Partial oc-

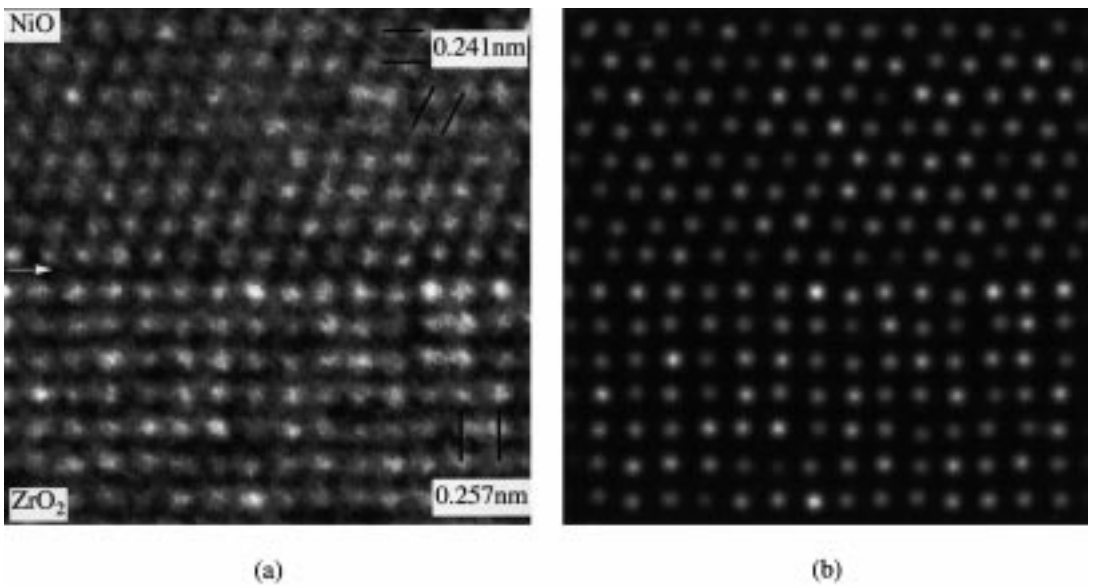


Fig. 5. Atomic-scale image of NiO–ZrO_2 interface along $[1\bar{1}0]_{NiO}//[100]_{ZrO_2}$: (a) Z-contrast STEM image; (b) object function derived from maximum entropy reconstruction.

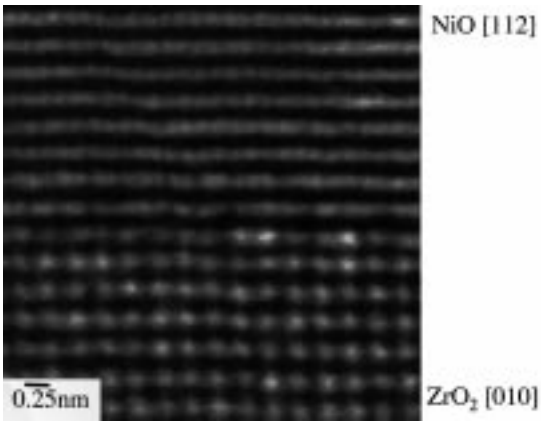


Fig. 6. Z-contrast image of NiO–ZrO_2 interface along $[112]_{NiO}//[010]_{ZrO_2}$ reveals displacement of the Ni atoms from their bulk lattice sites near the interface.

cupancy to avoid cation crowding is emerging as a common relaxation mechanism for oxide grain boundaries.

The oxygen positions in the grain boundary model were chosen such that they were most similar to the positions in bulk ZrO_2. In the fluorite structure, the oxygen ions are tetrahedrally coordinated by Zr ions. Most of the oxygen sites near the boundary have very similar environments to bulk oxygen sites in that they maintain a distorted tetrahedral coordination (these are indicated in white in Fig. 4). In the core of the structural unit, if we assume full oxygen columns, then the number of nearest neighbors would vary from that observed in the bulk. The oxygen sites indicated in gray in Fig. 4 would be deficient in Zr nearest neighbors while those shown in black would have an excess of nearest neighbors. Since the Z-contrast image does not give us information regarding the oxygen ions, it will be necessary to further explore the coordination of the oxygen ions at the boundary core with a combination of EELS and atomistic simulations. Having accurate coordinates for the cation positions from the Z-contrast image, however, provides a good starting model of the interface from which further refinements can be made.

3.2. NiO–ZrO_2 heterophase interfaces

As a result of the crystallographic orientation relationship between NiO and ZrO_2, it was possible to image the edge-on interface along two orthogonal directions, allowing a three-dimensional model of the interface to be produced [15]. Figure 5 shows the Z-contrast image of the NiO–ZrO_2 interface along the $[1\bar{1}0]_{NiO}//[100]_{ZrO_2}$ projection. The bright spots, again, correspond to the projected cation sublattices. Note that even though the atomic number of Zr is 40 and Ni is 28, the intensities of the Zr and Ni columns are very similar. This is because the atomic density of the Zr columns is much less than that of the Ni columns. The image shows that the interface is atomically flat and abrupt with a separation of 0.271 ± 0.004 nm between the interface Ni and Zr planes. This spacing is larger than that of either Zr (200) or NiO (111), so some volume expansion at the interface plane is observed. The boundary appears completely coherent with no misfit dislocations—the lattice misfit along this direction is only 0.4%.

Although the interface appears extremely ordered and coherent along the projection discussed above, the same is not true in three dimensions. When viewed along an orthogonal projection,

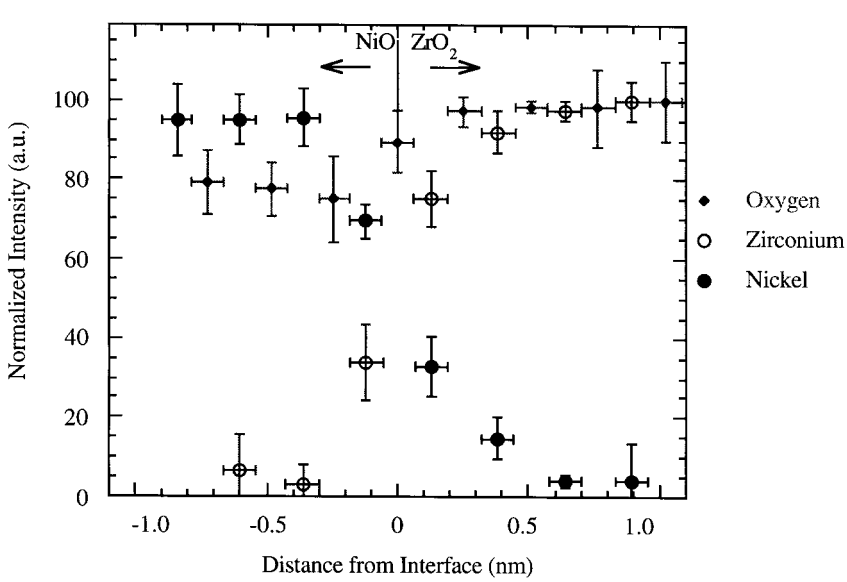


Fig. 7. EELS elemental profiles taken in steps of interplanar spacing across a NiO–ZrO_2 interface.

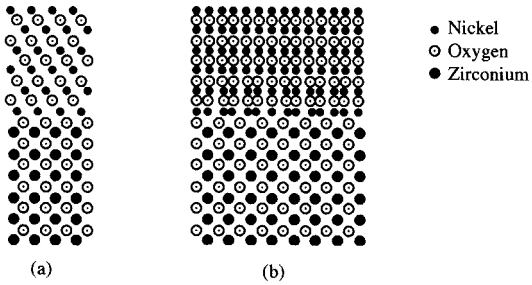


Fig. 8. Three-dimensional model of the NiO–ZrO_2 interface derived from Z-contrast images: (a) $[1\bar{1}0]_{NiO}//[100]_{ZrO_2}$ projection; (b) $[112]_{NiO}//[010]_{ZrO_2}$ projection.

$[112]_{NiO}//[010]_{ZrO_2}$, where there is a 14% lattice mismatch between the two lattices, structural disorder is observed near the boundary predominantly in the NiO (see Fig. 6). The NiO $(2\bar{2}0)$ lattice fringes running perpendicular to the interface have a 0.148 nm spacing that was successfully imaged by the VG HB603U machine. When the atomic positions are determined from maximum entropy reconstructions, a 0.02 nm root mean square (r.m.s.) relaxation of the $(2\bar{2}0)_{NiO}$ in the plane adjacent to the boundary is quantified. Because the interface is incommensurate along this direction, the relaxation is aperiodic.

EELS profiles acquired in steps of interplanar spacing across the interface indicate that the boundary is chemically abrupt as well. Figure 7 shows the chemical profiles, determined from integrated core-loss EELS edges for the various elements, in steps of interplanar spacing across the interface. Note that there is a complete transition from NiO to ZrO_2 in two interplanar steps, indicating no appreciable intermixing between the two phases across the boundary plane. The oxygen intensity at the interface is intermediate to that of either phase.

A three-dimensional model of the boundary based on the experimental data is shown in Fig. 8. Note several salient features of the model. First, the boundary consists of a single oxygen plane that is shared between the two crystals and continues the anion–cation sequence of planes seamlessly across the interface. Secondly, this common oxygen plane, as projected along $[1\bar{1}0]_{NiO}//[100]_{ZrO_2}$, is continuous with the oxygen sublattices of both phases, suggesting that the anion sublattice, and not the cation sublattice, is controlling the rigid body translation. Finally, the r.m.s. strain in the NiO along $[1\bar{1}0]_{NiO}$ has been incorporated into the model. Although the projected strains are on the order of 15%, the interatomic distances change by a maximum of $\pm 4.6\%$. The model not only provides us with an accurate, experimentally based interpretation of the interface structure, but also demonstrates the necessity of obtaining three-dimensional information from heterophase interfaces. Although Z-contrast images only provide two-dimensional data, as has been illustrated, if more than one edge-on cross section can be made, three-dimensional information from the interface can be obtained.

3.3. *Ni–ZrO_2 interface*

The interface between Ni and ZrO_2 [17] provides another challenge not yet encountered in the oxide–oxide interfaces. Because the ZrO_2 (100) surface is polar, it can terminate at the Ni–ZrO_2 interface with either an O or Zr plane. The former would result in Ni–O bonds across the boundary and the latter in Ni–Zr bonds. Figure 9 shows a Z-contrast

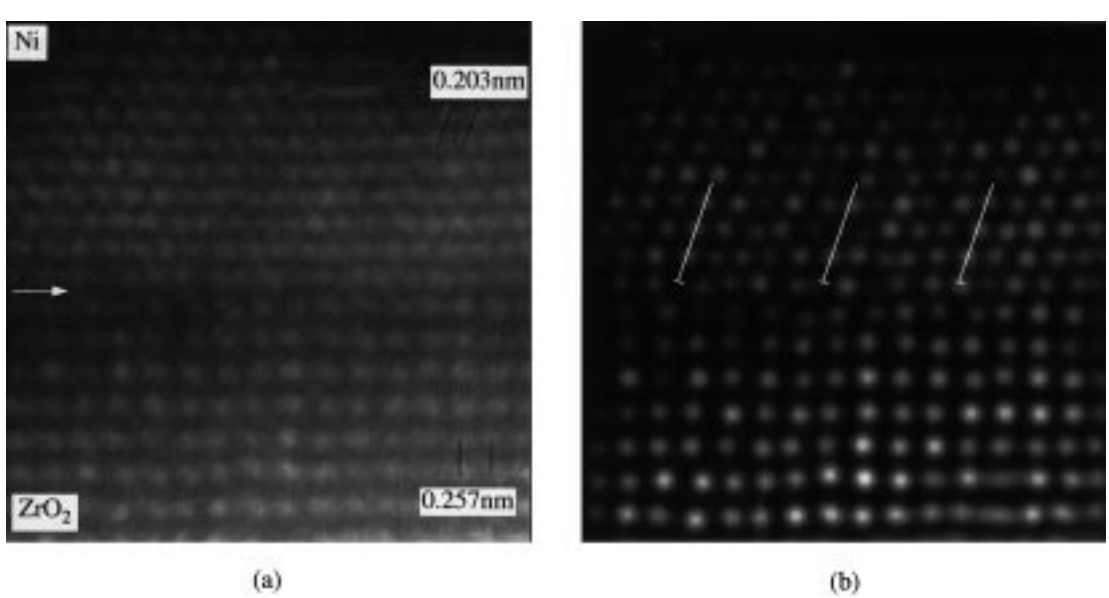


Fig. 9. Z-contrast STEM image of interface between Ni and ZrO_2 along $[110]_{Ni}//[100]_{ZrO_2}$.

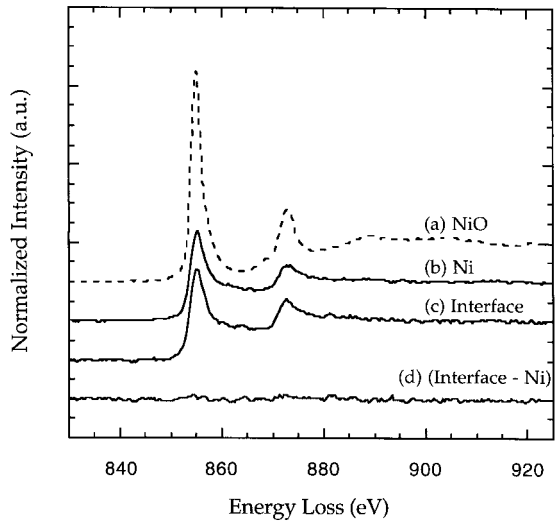


Fig. 10. Ni $L_{2,3}$ EEL spectra from: (a) bulk NiO; (b) bulk Ni; (c) the Ni–ZrO_2 interface. All spectra have been normalized to the contribution from the continuum. Spectrum d, the difference spectrum, results from subtracting the bulk Ni spectrum (b) from the interface spectrum (c).

image of the Ni–ZrO_2 interface along with the maximum entropy reconstruction along the $[1\bar{1}0]_{Ni}$ zone axis. Upon reduction, there is now a substantial 17% lattice mismatch between the two phases that leads to the misfit dislocations identified in Fig. 9.

Although the image does directly reveal the presence or absence of oxygen between the terminating Ni and Zr planes at the interface, the interplanar spacing between the two cation planes does provide some indirect evidence. The spacing between the terminating Ni and Zr planes was quantified from the maximum entropy reconstruction and found to be 0.244 ± 0.004 nm. This compares to a spacing of 0.271 ± 0.004 nm in the oxide–oxide boundary before reduction. The substantial contraction at the interface suggests that oxygen has been lost from the boundary plane. To substantiate this contention, we also collected EELS spectra from the boundary plane to investigate the type of bonding present at the interface. A metal–oxygen bond should show some ionicity while a metal–metal bond would not. We used the Ni $L_{2,3}$ EELS edge as a signature since its fine structure is very sensitive to the valence state of the Ni (see Fig. 10). The L_3/L_2 ratio of Ni^{2+} is 20–30% greater than that of metallic Ni [21–23]. The L_3 peak is also much sharper and more symmetrical in NiO than Ni due to the presence of a core exciton in the oxide [23]. As shown in Fig. 10, the EELS spectrum from the interface plane is nearly identical to the metallic Ni reference spectrum and therefore shows no evidence for an ionic bond at the interface. This is consistent with the observed contraction in the boundary plane observed in the Z-contrast images. We can

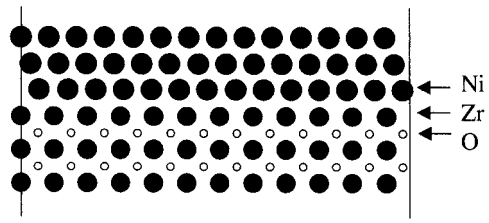


Fig. 11. Unrelaxed structure of Ni–ZrO_2 interface illustrating metal–metal termination at the boundary.

therefore conclude that metal–metal bonding predominates at this interface. Unfortunately, for this interface it was not possible to obtain lattice images from two edge-on boundary orientations. Although the $[010]_{ZrO_2}$ and $[112]_{Ni}$ zone axes are parallel, the $(220)_{Ni}$ lattice spacing is below the spatial resolution of the instrument. Based on the available data, however, an unrelaxed boundary structure was developed that incorporates the metal–metal termination at the interface; this is shown in Fig. 11.

4. CONCLUSIONS

As illustrated in the previous three examples, Z-contrast STEM imaging is an extremely powerful technique for elucidating the atomic structures of oxide interfaces. When coupled with maximum entropy analysis, quantitative structural information can be extracted. Structural units comprising grain boundaries can be directly determined from object function coordinates. Because the scattered cross sections scale with atomic number, partially occupied columns can be identified from lower intensities. When multiple edge-on projections of an interface can be obtained, Z-contrast images can provide quantitative three-dimensional structural information including rigid body translations and local atomic relaxations. In the case of metal–oxide interfaces, accurate determination of interplanar spacings can be obtained by maximum entropy analysis and can distinguish between metal and oxygen termination. Atomic resolution EELS is a powerful complementary means for quantifying atomic bonding at such interfaces. The experimentally based atomic models will serve as excellent starting models for atomistic calculations that can give further insight into the physical properties of the interfaces.

REFERENCES

1. Norton, M. G. and Carter, C. B., in *Materials Interfaces*, ed. D. Wolf and S. Yip. Chapman & Hall, London, 1992, pp. 151–189.
2. Dimos, D., *Phys. Rev. B*, 1990, **41**, 4038.
3. Yan, Y., Chisholm, M. F., Duscher, G. and Pennycook, S. J., *Phys. Rev. Lett.*, 1998, **81**, 3675.

4. Pennycook, S. J., Dickey, E. C., Nellist, P. D., Chisholm, M. F., Yan, Y. and Pantelides, S. T., *J. Eur. Ceram. Soc.*, 1999, **19**, 2211
5. Jesson, D. E. and Pennycook, S. J., *Proc. R. Soc. Lond. A*, 1993, **441**, 261.
6. Jesson, D. E. and Pennycook, S. J., *Proc. R. Soc. Lond. A*, 1995, **449**, 273.
7. Pennycook, S. J. and Jesson, D. E., *Acta metall. mater.*, 1992, **40**, S149.
8. Pennycook, S. J. and Jesson, D. E., *Phys. Rev. Lett.*, 1990, **64**, 938.
9. Pennycook, S. J. and Jesson, D. E., *Ultramicroscopy*, 1991, **37**, 14.
10. Gull, S. F. and Daniell, G. J., *Nature*, 1978, **272**, 686.
11. Gull, S. F. and Skilling, J., *IEE Proc.*, 1984, **131F**, 646.
12. Nellist, P. D. and Pennycook, S. J., *J. Microsc.*, 1998, **190**, 159.
13. McGibbon, A. J., Pennycook, S. J. and Jesson, D. E., *J. Micros.*, 1999, **195**, 44
14. Dhalenne, G. and Revcolevschi, A., *J. Cryst. Growth*, 1984, **69**, 616.
15. Dickey, E. C., Dravid, V. P., Nellist, P. D., Wallis, D. J. and Pennycook, S. J., *Acta mater.*, 1998, **46**, 1801.
16. Revcolevschi, A. and Dhalenne, G., *Nature*, 1985, **316**, 335.
17. Dickey, E. C., Dravid, V. P., Nellist, P. D., Wallis, D. J., Pennycook, S. J. and Revcolevschi, A., *Microsc. Microanal.*, 1997, **3**, 443.
18. McGibbon, M. M., Browning, N. D., Chisholm, M. F., McGibbon, A. J. and Pennycook, S. J., *Science*, 1994, **261**, 102.
19. McGibbon, M. M., Browning, N. D., McGibbon, A. J. and Pennycook, S. J., *Phil. Mag. A*, 1996, **73**, 625.
20. Browning, N. D., Nellist, P. D., Norton, D. P., Chisholm, M. F. and Pennycook, S. J., *Physica C*, 1998, **294**, 183.
21. Waddington, W. G., Rez, P., Grant, I. P. and Humphreys, C. J., *Phys. Rev. B*, 1986, **34**, 1467.
22. Sparrow, T. G., Williams, B. G., Ral, C. N. R. and Thomas, J. M., *Chem. Phys. Lett.*, 1984, **108**, 547.
23. Leapman, R. D., Grunes, L. A. and Fejes, P. L., *Phys. Rev. B*, 1982, **26**, 614.

Atomic Arrangement of Iodine Atoms inside Single-Walled Carbon Nanotubes

X. Fan,[1,2] E. C. Dickey,[1,*] P. C. Eklund,[3,†] K. A. Williams,[3] L. Grigorian,[3,‡] R. Buczko,[2,4,§] S. T. Pantelides,[4,2] and S. J. Pennycook[2]

[1]*Department of Chemical and Materials Engineering, University of Kentucky, Lexington, Kentucky 40506*
[2]*Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*
[3]*Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506*
[4]*Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235*
(Received 9 November 1999)

We report atomic resolution Z-contrast scanning transmission electron microscopy images that reveal the incorporation of I atoms in the form of helical chains inside single-walled carbon nanotubes. Density functional calculations and topological considerations provide a consistent interpretation of the experimental data. Charge transfer between the nanotube walls and the I chains is associated with the intercalation.

PACS numbers: 61.72.Ss, 61.16.Bg, 61.48.+c

Single-walled carbon nanotubes (SWNTs) have been shown to be amphoteric in character; i.e., they can exchange electrons with a dopant atom (or molecule) to form the corresponding positively or negatively charged counterion [1,2]. Little is known, however, about the actual dopant sites and distributions. In this Letter, we present Z-contrast scanning transmission electron microscopy (STEM) images and complementary theoretical calculations that reveal the distribution of I atoms within I-doped SWNT bundles.

SWNT ropes were made by an arc-discharge technique [3] and doped by immersion into molten iodine as described previously [4]. The intercalation was found to be reversible and Raman scattering showed charged polyiodide ions to be distributed throughout the bundles [4]. It has been generally assumed that the intercalant enters the interstitial channels between tubes. This assumption was based on the belief that the ends of the as-prepared tubes are closed, unless opened by a postsynthesis treatment, e.g., refluxing in concentrated acid [5,6]. Here we report Z-contrast images of individual SWNTs protruding from the ends of SWNT bundles that reveal iodine chains to be incorporated inside the tubes. Images were taken with a VG microscopes HB603U STEM, operating at 300 kV, equipped with a high-resolution objective lens capable of forming a probe 0.13 nm in diameter. This microscope has demonstrated information transfer in Z-contrast images at the sub-angstrom level [7] and direct imaging of individual heavy atoms on a light support [8]. Of particular relevance to this work is the fact that the Rutherford scattering used to form Z-contrast images scales with the square of the atomic number [9] resulting in large contrast between I ($Z = 53$) and C ($Z = 6$). To reduce the noise in some of the images, maximum entropy refinement was utilized [10–12].

In order to guide interpretation of the experimental data, we performed a number of theoretical calculations using density functional theory with the local density approximation of exchange correlation, norm-conserving pseudopotential plane waves, and supercells [13]. The use of complete nanotube supercells was impractical because of the large number of atoms required. Calculations were therefore performed for free I chains and for I chains on flat graphene sheets or on graphene sheets that were bent to simulate the curvature of the interior of a nanotube.

Figure 1 compares bright field and Z-contrast images from regions in the sample where several individual tubes can be distinguished. Whereas the bright field image shows very little contrast and is ambiguous as to the position or even presence of I, the dark field image directly reveals the projected I distribution. Comparing the two images, it is seen that the iodine is incorporated within one particular nanotube (labeled A) but is not present in others in the bundle (e.g., nanotube B). The diameter of the A tube is 1.36 nm and assuming that the π electrons extend 0.17 nm from the wall, as in graphite, the inner tube diameter would be $\approx$1.03 nm. From the width of the bright stripe, 0.65 nm, it is clear that the iodine is incorporated inside the single-walled tube.

The internal iodine is confirmed in the higher resolution Z-contrast images of Figs. 2 and 3, where two chains of iodine atoms are observed inside the nanotubes. The projected spacing between the two chains is observed to modulate along the tube axis, suggesting that the chains have a helical structure in three dimensions. The period of the modulation is measured to be about 5 nm in Fig. 2, but other tubes show different periodicities. For example, Fig. 3(a) shows a tube containing a double helix with a periodicity of about 12.5 nm, and Fig. 3(b) shows a double helix with an even longer periodicity. The observation of different periodicities suggests that the iodine chains adopt a preferred orientation with respect to the nanotube wall, and the different periods correspond to tubes with different chirality. Direct attempts to confirm the chirality of individual tubes by electron nanodiffraction were unsuccessful, however.

The observed discontinuities and variations in intensities along the length of the iodine chains hamper direct

 © 2000 The American Physical Society

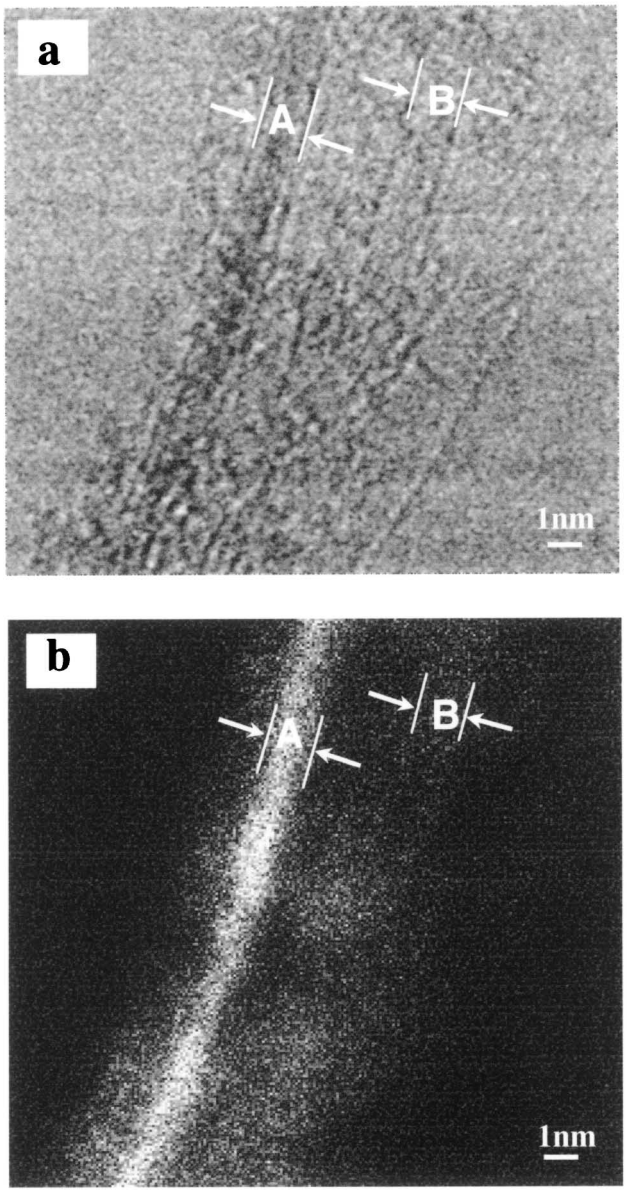


FIG. 1. (a) Bright field phase contrast image and (b) dark field Z-contrast image of a SWNT rope, taken simultaneously on the STEM. The conventional phase contrast image is useful for locating the nanotube walls; the precise registry of the two images therefore allows the iodine distribution to be determined to high precision. Iodine is clearly inside some tubes, such as A but not others, for example, B.

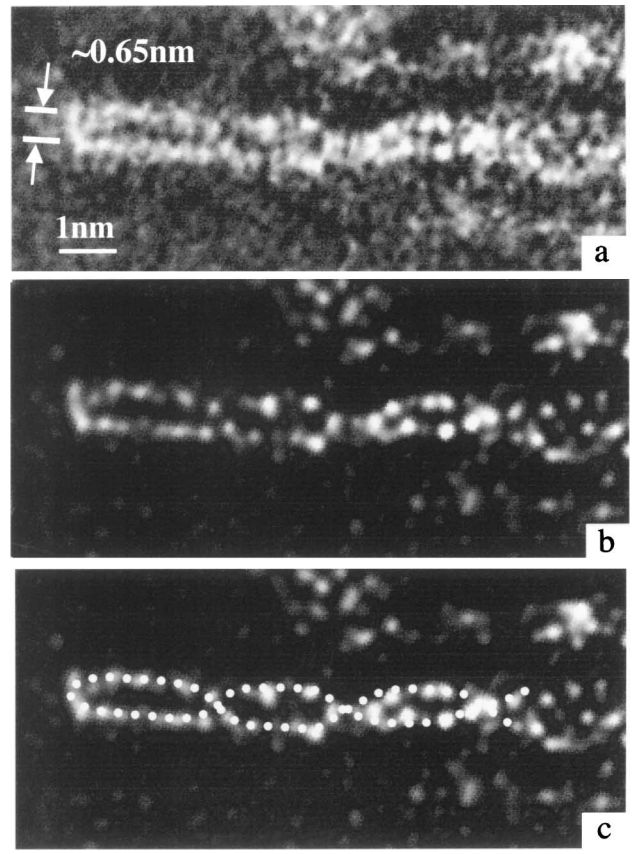


FIG. 2. (a) Higher resolution Z-contrast image of a SWNT containing two strands of iodine. (b) Maximum entropy processed image to reduce the noise. Two constrictions can be seen, consistent with the projection of a double helix configuration with the two chains spiraling along the inside walls of the nanotube as shown in the overlay in (c). The period of the helix is about 5 nm, and the maximum separation of the two strands is ~0.65 nm, much less than the 1.3–1.4 nm diameter of the SWNT.

structure determination. These intensity variations are believed to result from electron displacement damage, which occurs during the imaging process. Since we are imaging single atoms, and not columns of atoms as in crystalline materials, displacements of individual atoms have more severe consequences for the image resolution. The maximum energy transferred from the incident electrons to the iodine nucleus, T_m, assuming a relativistic elastic collision between the electron and a stationary nucleus is 3.3 eV for 300 keV electrons incident on an I atom [14,15]. The energy barrier necessary to remove an iodine atom from an I_3^- chain was calculated to be ~1.5 eV by first-principles total energy calculations. Taking into account the 0.3 eV/atom binding energy of the chain to the tube wall (to be discussed in detail below), we place an upper estimate on the displacement energy (E_d), or the energy necessary to displace an I atom from the chain, of 1.8 eV. Since $E_d/T_m \sim 0.5$ significant beam damage will occur during the imaging process. These atomic displacements lead to increased noise and thus variations in projected positions and intensities, which are apparent in Figs. 2 and 3.

To guide interpretation of the experimental data, we performed several first-principles calculations that shed light on interactions between iodine and the nanotube wall. Similar continuous, one-dimensional iodine structures to those observed in this study have been reported in the polymer literature where iodine has been used to increase the electrical conductivity of polymers by orders of magnitude [16–18]. Diffraction studies of iodine-doped polymers indicate that the iodine intercalates the one-dimensional

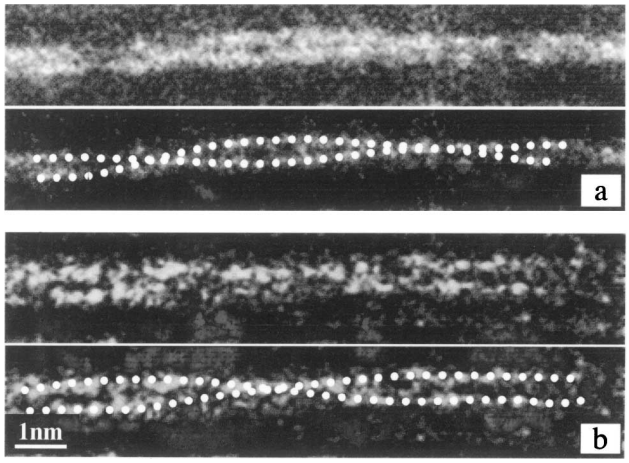


FIG. 3. Images of iodine doped SWNT ropes showing different periodicities, shown schematically above the tubes; (a) an iodine helix showing 12–13 nm periodicity; (b) a helix of longer periodicity.

interstitial channels of crystalline polymers in the form of continuous linear arrays. Furthermore, Raman spectroscopy shows the charge transfer to the iodine to be typically 0.33 eV/iodine or 0.2 eV/iodine, corresponding to I_3^- and I_5^- species, respectively [16–18]. Since we observe similar linear arrays of iodine and, as reported earlier [4], the Raman modes of the I-doped SWNTs agree with those reported for the I_3^- and $(I_5)^-$ arrays [16–18], we conclude that the iodine is in a similar state to that found in doped crystalline polymers. We therefore performed calculations for chains of three and five iodine atoms above both a flat graphene sheet and a curved sheet, as illustrated by Fig. 4. This configuration would correspond to one unit cell of the continuous iodine chain. In all cases, we found an equilibrium distance of ~0.35 nm between the I chain and graphene sheet with a net binding of ~0.3 eV per I atom, provided the initial position of the I chain was close enough to the graphene sheet to allow charge transfer to take place. The chain prefers a straight configuration. We contrast this finding with the fact that neutral chains of three I atoms in free space are bent whereas charged chains are straight. We conclude that charge transfer from the tube wall to the I chain stabilizes the I chains. Our calculated equilibrium distance (0.35 nm) agrees precisely with the distance obtained from the observed 0.65 nm spacing of the chains in Fig. 2, if we assume the double helix is located inside a (10, 10) tube of 1.36 nm diameter.

Another indication of the influence of the curved nanotube walls on the formation of the helical chains is provided by the topological considerations. I_2 does not intercalate graphite because of its large van der Waals diameter (0.396 nm) [17] and its 0.27 nm interatomic spacing does not correlate with any spacing in the basal plane of graphite [19–21]. However, the space available for intercalation is much greater in nanotubes and the radially projected interatomic spacings of *curved* graphene sheets better match the preferred I-I spacing. (10, 10) tubes are thought to be a common type of nanotube in the arc-derived material, and in Fig. 4 we show three possible correlations between the nanotube lattice and the iodine spacings, which are nearly commensurate. Continuous chains bound to these sites result in helical arrangements with periods of 12.5, 5, and 1.5 nm. First-principles calculations show the I chains to be stable in these commensurate orientations. The longest two periods for the (10, 10) tube are in excellent agreement with the helices imaged in Figs. 3(b) and 2. The shorter period helix would project as a rather uniform arrangement of iodine, which may explain images of the type shown in Fig. 1(b). A complete structural model for a (10, 10) tube with a 5 nm double helix of iodine inside is given in Fig. 5. Although we have focused our discussion on (10, 10) tubes because of their relatively simple, symmetric structure, typical nanotube samples are believed to have a distribution of diameters and chiralities. Tubes with different chirality would, of course, be expected to show different chain periodicities because of the importance of the chain-(tube wall) interaction.

Two continuous strands of iodine with the 5 nm period give a composition of IC_{20}, but it would appear feasible to incorporate additional spirals of iodine, up to five before the distance between neighboring strands become too close, giving a composition IC_8. This provides an alternative explanation for the lack of an observed helical structure in Fig. 1(b). For the 1.5 nm helix there would be sufficient room for only two chains, giving a saturation doping of IC_{15}. The overall composition determined experimentally by weight uptake measurements and thermogravimetric analysis is IC_{12} [4], but since we observe many tubes to be unfilled, this suggests substantial iodine is also incorporated in the interstitial positions between tubes in a bundle. We would anticipate that polyiodide chains in the

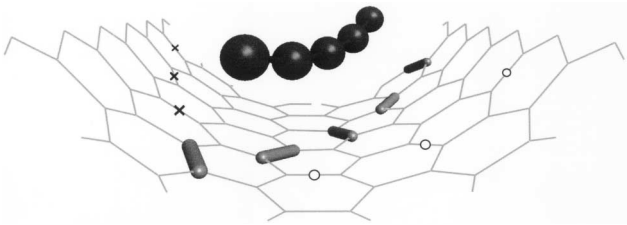


FIG. 4. Bent graphene sheet used in the calculations with periodic boundary conditions. The curvature corresponds to the correct curvature of an actual (10, 10) nanotube. The distances between the C-C bonds that are highlighted with rods 3.2 Å) project radially to 0.28 nm, very close to the preferred I-I spacing of 0.29 nm, at a distance of 0.35 nm where the I chain prefers to rest. This results in a helical arrangement of iodine with a 5 nm period, as seen in Fig. 2. Two other near-commensurate directions are possible. The direction marked by crosses has an I-I spacing of 0.31 nm and a period of 12.5 nm, as seen in Fig. 3(a). The direction shown by circles has an I-I spacing of 0.30 nm and a much shorter period, 1.5 nm.

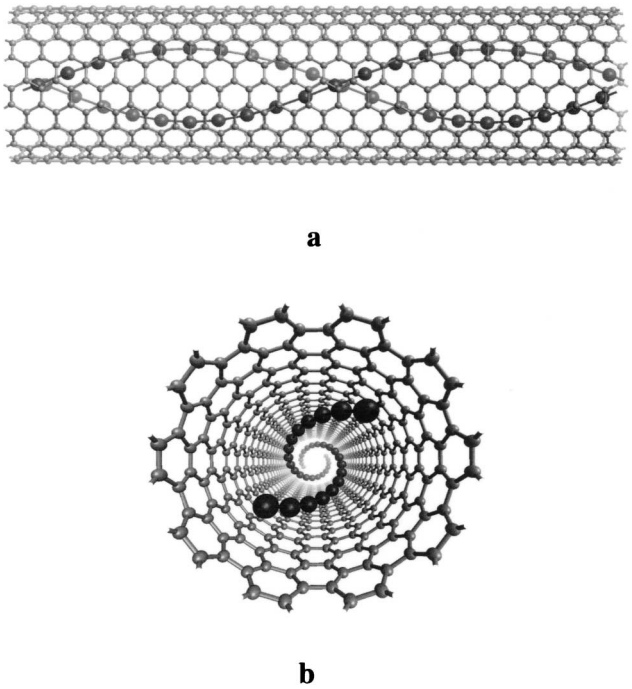


FIG. 5. Structure model obtained by extrapolating the 5 nm helix of Fig. 4 to a full scale (10, 10) nanotube, (a) side view and (b) top view.

interstitial locations between tubes could bond by a similar mechanism although for (10, 10) armchair tubes there are no good correlations with the projected nanotube lattice.

It is interesting to note that previous studies did not observe any charge transfer between SWNTs and iodine when the iodine was present in the vapor phase [2]. Molten iodine is quite different from iodine vapor. In particular, iodine molecules in molten iodine tend to dissociate into highly reactive iodinium (I^+) and polyiodide $(I_3)^-$ species [22]. On the other hand, the iodine vapor contains only nondissociated iodine molecules, which are less chemically active as compared to ionic forms. The (I^+) ions in the molten I are very strong oxidizing agents and could attack the remaining catalyst particles at the tube ends or the defect sites in the tube walls. These sites could subsequently serve as entry points for the polyiodide species. Although we have no conclusive evidence for this intercalation mechanism, x-ray diffraction spectra from the samples do show all of the metal catalyst to be oxidized to their metal-iodide phases. We do not yet understand why some of the tubes are intercalated and others are not. Such understanding awaits more thorough understanding of the tube end structures.

This research was sponsored by the NSF-MRSEC Advanced Carbon Materials Center under Grant No. DMR-9809686 and by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

*To whom correspondence should be addressed. Electronic address: ecdickey@engr.uky.edu
†Current address: 104 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802-6300.
‡Current address: Fundamental Research Lab, Honda R&D Americas, Inc., 100 Town Center, Suite 2050, Southfield, MI 48075.
§Permanent address: Institute of Physics, Polish Academy of Sciences, 02-668, Warsaw, Poland.

[1] R. S. Lee, H. J. Kim, J. E. Fischer, A. Thess, and R. E. Smalley, Nature (London) **388**, 255 (1997).
[2] A. M. Rao, P. C. Eklund, S. Bandow, A. Thess, and R. E. Smalley, Nature (London) **388**, 257 (1997).
[3] C. Journet *et al.*, Nature (London) **388**, 756 (1997).
[4] L. Grigorian, K. A. Williams, S. Fang, G. U. Sumanasekera, A. L. Loper, E. C. Dickey, S. J. Pennycook, and P. C. Eklund, Phys. Rev. Lett. **80**, 5560 (1998).
[5] J. Liu *et al.*, Science **280**, 1253 (1998).
[6] J. Sloan *et al.*, Chem. Commun. **3**, 347 (1998).
[7] P. D. Nellist and S. J. Pennycook, Phys. Rev. Lett. **81**, 4156 (1998).
[8] P. D. Nellist and S. J. Pennycook, Science **274**, 413 (1996).
[9] D. E. Jesson and S. J. Pennycook, Proc. R. Soc. London A **449**, 273 (1995).
[10] S. F. Gull and G. J. Daniell, Nature (London) **272**, 686 (1978).
[11] S. F. Gull and J. Skilling, IEE Proc. F, Radar Signal Process. **131**, 646 (1984).
[12] A. J. McGibbon, S. J. Pennycook, and D. E. Jesson, J. Microsc. **195**, 44 (1999).
[13] P. Hohenberg and W. Kohn, Phys. Rev. **140**, B1133 (1964).
[14] L. Reimer, *Transmission Electron Microscopy* (Springer-Verlag, New York, 1993), Chap. 5.
[15] A. Maiti, M. F. Chisholm, S. J. Pennycook, and S. T. Pantelides, Appl. Phys. Lett. **75**, 2380 (1999).
[16] R. C. Teitelbaum, S. L. Ruby, and T. J. Marks, J. Am. Chem. Soc. **101**, 7568 (1979).
[17] B. H. Baughman *et al.*, in *Molecular Metals*, edited by W. E. Hatfield (Plenum Press, New York, 1979), p. 187.
[18] M. J. Nilges, E. K. Barefield, R. L. Belford, and P. H. Davis, J. Am. Chem. Soc. **99**, 749 (1977).
[19] S. Nakashima *et al.*, Chem. Phys. Lett. **268**, 359 (1997).
[20] H. Werner, M. Wesemann, and R. Schlogl, Europhys. Lett. **20**, 107 (1992).
[21] S. Flandrois, C. Hauw, and B. François, J. Phys. (Paris) **44**, C3-523 (1983).
[22] S. A. Jenekhe, S. T. Wellinghoff, and J. F. Reed, Mol. Cryst. Liq. Cryst. **105**, 175 (1984).

Z-contrast imaging and electron energy-loss spectroscopy analysis of chromium-doped diamond-like carbon films

X. Fan[a] and E. C. Dickey[b]
Department of Chemical and Materials Engineering, University of Kentucky, Lexington, Kentucky, 40506

S. J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

M. K. Sunkara
Department of Chemical Engineering, University of Louisville, Louisville, Kentucky 40292

(Received 30 April 1999; accepted for publication 31 August 1999)

Metal-doped diamond-like carbon films were produced for the purpose of an electrochemical nanoelectrode. In this study we use *Z*-contrast scanning transmission electron microscopy to directly observe metal cluster formation and distributions within the chromium-doped carbon films. At low doping (~6 at. % Cr), Cr is uniformly distributed within the *C* matrix; at high doping (~12 at. % Cr), Cr-rich clusters are formed. Analyzing electron energy loss spectroscopy Cr $L_{2,3}$ white line ratios, we find that the Cr tends to be metallic-like when it is uniformly distributed in the *C* matrix and carbide-like in the Cr-rich clusters. The carbon is more diamond-like at low doping and more graphite/carbide-like at high doping according to the sp^2/sp^3 electron percentage measurements.
© *1999 American Institute of Physics.* [S0003-6951(99)00844-X]

Diamond-like carbon (DLC) is a very promising material with a wide variety of potential applications due to its unique diamond-like properties.[1] Doping with metals could create a two dimensional array of metal clusters within the diamond-like carbon matrix. Such composite structures are of interest for use as nanoelectrodes in electrochemistry.[2] The formation and distribution of the metal clusters within the carbon film are fundamentally important for the physical properties of the material. Here we study the structure and chemistry of chromium-doped DLCs using *Z*-contrast electron imaging and electron energy loss spectroscopy (EELS). We assess the structure of the diamond matrix by quantifying the percentage of sp^2 bonding from the EELS *C K* edge, and we investigate the distribution and bonding of Cr within the DLC matrix.

In this work we employ the *Z*-contrast electron imaging method because it is more directly interpretable than conventional phase contrast imaging.[3–5] Using an annular dark-field detector in the scanning transmission electron microscope (STEM), only the high angle scattering of electrons are collected, where the total scattering intensity is proportional to the square of the atomic number, *Z*. The *Z*-contrast image can directly reveal atomic positions as the brighter spots correspond to the positions of heavier atoms. The EELS data can be collected simultaneously through the hole of the dark-field detector allowing direct correlations between the images and spectra.

Diamond-like carbon films of 0.1–5.0 μm thickness were grown on alumina substrates using a plasma chemical deposition technique described elsewhere.[6] The doping of chromium was performed by coevaporation of chromium into the reactor volume while growing the carbon film.[7,8] The carbon film was prepared for transmission electron microscopy by mechanical dimpling from the substrate side and ion beam milling to electron transparency. Two samples doped at low (~6 at. % Cr) and high levels (~12 at. % Cr) were examined by *Z*-contrast imaging in a VG HB603 STEM and a VG HB501. The EELS data were collected on VG HB501 STEM and Philips CM200 transmission electron microscopes.

Figure 1 is a phase contrast image of highly doped sample which is similar to a typical amorphous structure. The Cr is not readily distinguishable. Figure 1(b) is a dark-field *Z*-contrast image of the same area. Since Cr atoms are heavier than C, the brighter areas directly correspond to Cr-rich regions. The average Cr-rich cluster size is estimated to be 2 nm with a separation of 3.5 nm between clusters. The low-doped sample does not show any Cr-rich regions, indicating uniform Cr distribution throughout the DLC matrix (picture not shown).

The EELS spectra of carbon the *K* edge and Cr *L* edge were collected from both samples, as well as from reference samples of chromium carbide Cr_3C_2, metallic Cr and synthetic graphite. Spectra, shown in Figs. 2(a)–2(e), were taken in image mode from a scanned area of approximately 100 nm^2. Since the highly doped film has uneven Cr distribution, spatially resolved EELS were also collected on the VG HB501 with a probe size less than 3 Å from both Cr-rich regions and Cr-deficient regions, as shown in Fig. 3. The probe position was determined from the *Z*-contrast image taken simultaneously with the EELS spectra. Using the intensity ratio of I_C and I_{Cr} and calibrating with the reference sample Cr_3C_2, the concentrations of Cr are found to be an average of 6% for the low doped sample. For the highly doped sample, bright regions show 7–17 at. % Cr whereas dark regions show 1–3 at. % Cr.

[a]Also at Solid State Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831.
[b]Corresponding author; electronic mail: ecdickey@engr.uky.edu

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

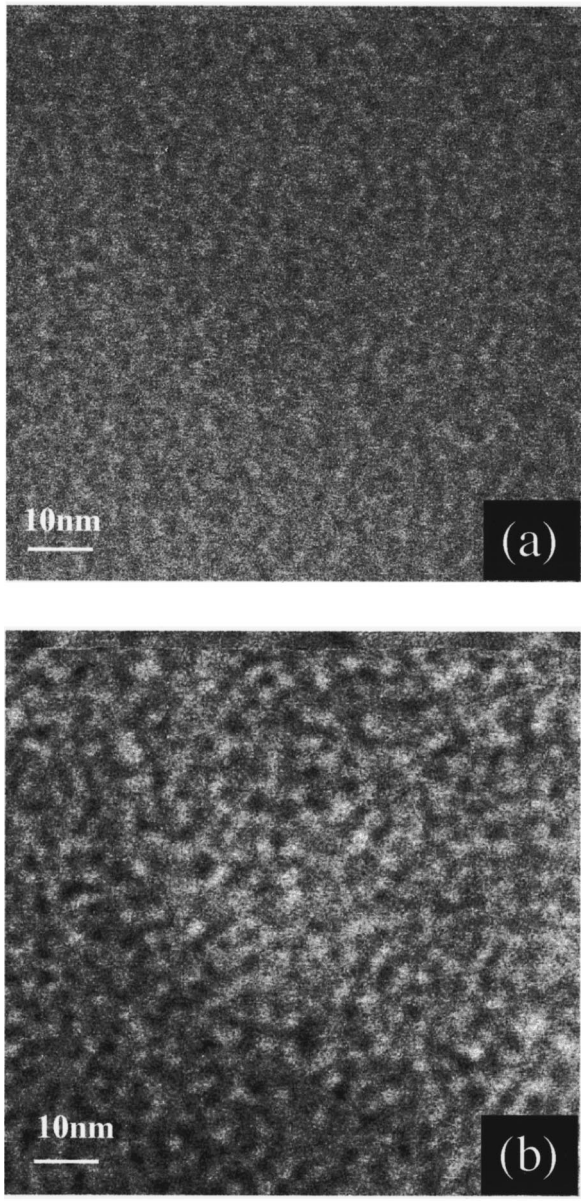


FIG. 1. Chromium doped diamond-like carbon. (a) Phase contrast image showing amorphous carbon matrix. (b) *Z*-contrast image showing Cr clusters.

All chromium L edges have similar features comprising two sharp peaks L_2 and L_3, known as ''white lines,'' and a continuum background following the edge. The L_3/L_2 white line ratio, R_w, is correlated to the electron occupancy and spin pairing in $3d$ band.[9–12] The variations of the white line ratio are due to the interaction between chromium and surrounding atoms. For example, a large increase of R_w in the chromium alloy $Cr_{20}Au_{80}$ reflects the dramatic increase of the magnetic moment.[9] However, it is also found that the white line ratio is sensitive to Cr cluster size when deposited on a graphite substrate.[10]

In our case it is interesting to compare the Cr-doped DLCs with metallic Cr and chromium carbide. The white line ratios, R_w, calculated using the standard method,[12] are listed in Table I. The white line ratios in low-doped samples are close to metallic Cr. In the highly doped sample, they are close to carbide in the Cr rich region and metallic in the Cr deficient region. So, it appears that after a critical doping

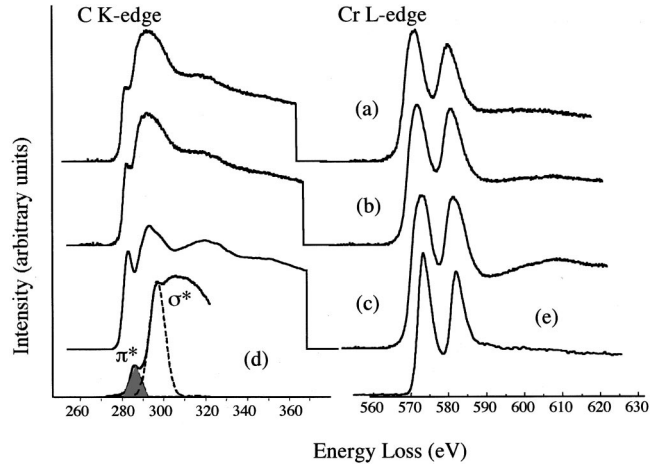


FIG. 2. EELS of carbon K edge(left) and Cr L edge(right). (a) Low doped DLC, (b) highly doped DLC, (c) chromium carbide, and (d) graphite. The σ^* peak is partially fitted by a Gaussian function as shown by the dashed curve. The shaded area is the π^* peaks obtained by subtracting the σ^* fitting. (e) Metallic Cr.

level, Cr-carbide clusters precipitate from the DLC matrix.

Since the introduction of chromium may change the structure of the carbon, it is also important to evaluate the electronic properties of the carbon. The major difference between a graphite-like carbon and a diamond-like carbon is the absence of π electrons in diamond-like material. This feature is shown in the EELS spectra in Figs. 2(a)–2(d) where the π^* peak is at about 284 eV. By measuring the intensity ratio of this peak against the following broad σ^* peak, and then calibrating with the known graphite structure, the percentage of sp^2 bonding in the many forms of pure carbon can be determined.

Several methods have been used to measure the sp^2 percentage from the EELS C K edge which involve selection of a fixed window for the π^* peak integrated intensity.[13,14] We find these methods to be insufficient for our study because they do not take into account potential contributions to π^* integrated intensity from the overlapping σ^* peak. As can be seen, some of our spectra have higher relative intensity in the π^* peak region than that of graphite. This leads to the unrealistic result of higher sp^2/sp^3 ratio than graphite. It is believed that the broadening of the σ^* peak causes the increased intensity at π^* peak region. It was therefore

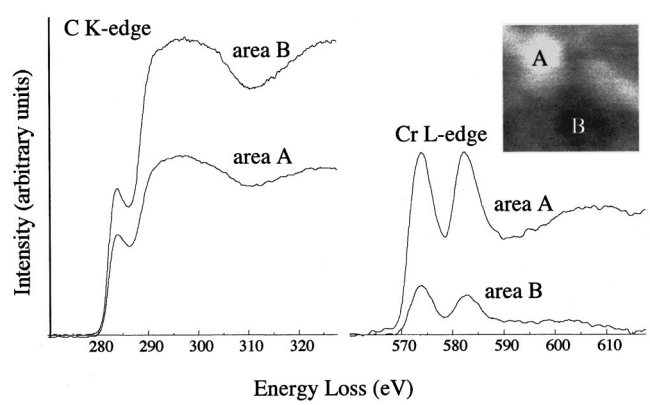


FIG. 3. The inset is a portion of the *Z*-contrast image. Spatially resolved EELS spectra of the C K edge and Cr L edge were collected at bright area A and dark area B, respectively.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

TABLE I. The Cr concentration % Cr, peak intensity ratio $I_\pi/I_{\pi+\sigma}$, sp^2 percentage and Cr white line ratio R_w, obtained from low doped (DLC-1), highly doped (DLC-h) at Cr rich area (A), and Cr deficient area (B), the reference samples chromium carbide (Cr_3C_2), metallic Cr and graphite.

	at. % Cr	R_w	$I_\pi/I_{\pi+\sigma}$	% sp^2
DLC-1	6%	1.7	0.047	56%
DLC-h(A)	7%–17%	1.4	0.13	⋯
DLC-h(B)	1%–3%	1.7	0.11	100%
Cr_3C_2	60%	1.35	0.15	⋯
metallic Cr	100%	1.63	⋯	⋯
graphite	0%	⋯	0.11	100%

necessary to use a modified version that takes into account this broadening effect.[15] We demonstrate this method for the case of graphite in Fig. 2(d). The front edge of the σ^* peak is partially fitted with a Gaussian function as indicated by the dashed curve. The π^* function is obtained by subtracting the σ^* function in the π^* region, resulting as the shaded small peak. Then the relative peak ratio $I_\pi/I_{\pi+\sigma}$ is defined as the total intensity of the π^* peak divided by the total intensity of a fixed energy window in the original spectrum. We choose this range from 280 to 300 eV consistently for all the spectra. For pure carbon materials, the actual sp^2/sp^3 ratio can be determined by calibrating with pure graphite for which the percentage of sp^2 is 100%.

Our measurements on DLC samples are summarized in Table I. The low-doped sample shows about 56% sp^2, i.e., 44% sp^3 bonding. In the highly doped material we have the additional complication of carbide formation. The Cr_3C_2 standard shows a $C-K$ edge structure with a strong peak at about 284 eV, the same energy as the π^* electron peak in graphite [see Fig. 2(c)]. Therefore, the sp^2/sp^3 ratio cannot be determined in Cr-rich regions where the Cr edge indicated carbide formation. The Cr-deficient region shows no evidence of carbide formation, the Cr edge indicating metallic bonding. Here the sp^2/sp^3 ratio could be analyzed, and the percentage of sp^2 bonded carbon is found to be almost 100%. We do not believe beam broadening to be a significant factor limiting our spatial resolution; we estimate beam broadening in a 20-nm-thick film to be only ~2 nm, which is comparable to the size of the Cr-rich and Cr-deficient regions. However, if the Cr signal from the Cr-deficient region was actually due to beam broadening into the adjacent Cr-rich area, the white line ratios would be identical in the two areas. In fact, the white line ratio indicates a change in bonding type, implying we do have the necessary spatial resolution.

In summary, we have directly imaged the Cr distribution in Cr-doped diamond-like carbon films by *Z*-contrast imaging. There is a critical change in Cr distribution from low to high level doping. The carbon in a low doped sample maintains its diamond-like features; the Cr is uniformly distributed throughout the matrix and maintains metallic bonding. This implies no strong interaction between carbon and Cr. High Cr doping leads to the formation of chromium carbide clusters. The residual Cr in between the clusters remains metallic-like, however the C loses its sp^3-type bonding. Therefore the diamond-like carbon network breaks down at high Cr doping.

This research was sponsored by the NSF-MRSEC Advanced Carbon Materials Center under Grant No. DMR-9809686, and by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

[1] J. Robertson, Pure Appl. Chem. **66**, 1789 (1994).
[2] R. M. Penner, M. J. Heben, T. L. Longin, and N. S. Lewis, Science **250**, 1118 (1990).
[3] S. J. Pennycook and D. E. Jesson, Phys. Rev. Lett. **64**, 938 (1990).
[4] S. J. Pennycook and D. E. Jesson, Ultramicroscopy **37**, 14 (1991).
[5] S. J. Pennycook, ''STEM: *Z*-contrast,'' in *Handbook of Microscopy*, edited by S. Amelinckx, D. van Dyck, J. van Landuyt, and G. van Tendeloo (VCH, Weinheim, 1997), pp. 595–620.
[6] V. F. Dorfman and B. N. Pypkin, Surf. Coat. Technol. **48**, 193 (1991).
[7] M. Sunkara, E. Yeap, S. V. Babu, V. G. Ralchenko, B. N. Pypkin, and B. L. Shubekin, ''Electrochemical behavior of metal-doped diamondlike carbon (Me-DLC) films,'' in *Diamond Materials V*, edited by J. L. Davidson, W. D. Brown, A. Gicquel, B. V. Spitsyn, and J. C. Angus (Electrochemical Society, Pennington, NJ, 1997), pp. 297–308.
[8] P. Koduri, M. Sunkara, E. C. Dickey, C. Frasier, and S. V. Babu, ''Structural and electrochemical characterization of metal-doped diamondlike (Me-DLC) films,'' in *Proceedings of Twelfth International Symposium On Surface Modification Technologies*, edited by T. S. Sudarshan, K. A. Khor, and M. Jeandin (Institute of Materials and American Society of Metals London, U.K., 1998), pp. 357–364.
[9] D. M. Pease, S. D. Bader, M. B. Brodsky, J. I. Budnick, T. I. Morrison, and N. J. Zaluzec, Phys. Lett. A **114**, 491 (1986).
[10] L. Lozzi, M. Passacantando, P. Picozzi, S. Santucci, and M. De Crescenzi, Z. Phys. D **20**, 387 (1991).
[11] T. I. Morrison, M. B. Brodsky, and N. J. Zaluzec, Phys. Rev. B **32**, 3107 (1985).
[12] D. H. Pearson, C. C. Ahn, and B. Fultz, Phys. Rev. B **47**, 8471 (1993).
[13] S. D. Berger, D. R. Mckenzie, and P. J. Martin, Philos. Mag. Lett. **57**, 285 (1988).
[14] J. Bruley, D. B. Williams, J. J. Cuomo, and D. P. Pappas, J. Microsc. **180**, 22 (1995).
[15] B. Rafferty, Ph.D. thesis, Cambridge University, 1997.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

Atomic structure of $Ba_{0.5}Sr_{0.5}TiO_3$ thin films on $LaAlO_3$

H.-J. Gao[a)]
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030

C. L. Chen
Department of Physics and the Texas Center for Superconductivity, University of Houston, Houston, Texas 77204

B. Rafferty and S. J. Pennycook[b)]
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030

G. P. Luo and C. W. Chu
Department of Physics and the Texas Center for Superconductivity, University of Houston, Houston, Texas 77204

(Received 26 April 1999; accepted for publication 16 August 1999)

Perovskite barium strontium titanate $Ba_{0.5}Sr_{0.5}TiO_3$(BST) thin films were grown on (001) $LaAlO_3$ (LAO) using pulsed-laser ablation. The microstructures of the *as*-grown BST films were studied with selected electron diffraction, transmission electron microscopy, and scanning transmission electron microscopy. The BST thin films are oriented with their [001] directions parallel to the ⟨102⟩ directions of the LAO. Both cross-sectional and plan-view studies show the BST films to be single crystals with smooth surfaces. The interfaces were seen to be atomically sharp by cross-sectional, high-resolution electron microscopy. The density of misfit dislocations was consistent with the 4.3% lattice mismatch, and they were found to be dissociated into partials. © *1999 American Institute of Physics.* [S0003-6951(99)00443-X]

Thin films of ferroelectric barium strontium titanates $(Ba_{1-x}Sr_x)TiO_3$ (BST) have recently become very attractive to the microelectronic industry as good candidates for device applications,[1] for example, as high density dynamic random access memories (DRAM),[2] smart card memories,[3] and tunable microwave devices.[4–9] They exhibit not only high values of the relative dielectric constant (ϵ_r), but also large tunability of ϵ_r through the application of an electric field. The tunability of BST thin films offers unique opportunities for the development of various microwave devices, such as microstrip line phase shifters, tunable filters, and high-Q resonators.

BST films with various stoichiometries have been grown on $LaAlO_3$ (LAO) by techniques such as sputtering, metal-organic chemical vapor deposition, laser ablation, and others.[10–13] LAO is also a widely used substrate for $PbZr_xTi_yO_3$ (PZT) films and for high temperature superconductors.[14–16] Application of BST to tunable microwave devices, however, requires high quality epitaxial BST films with atomically smooth surfaces and atomically sharp interfaces to minimize the loss tangent. Thus, investigating the microstructures between the as-grown films and the substrates and the epitaxial behavior is particularly important for BST thin film synthesis and device engineering. We have studied the microstructures and interface behavior of epitaxial BST films on (001) LAO by using selected area electron diffraction, transmission electron microscopy (TEM), and high resolution scanning transmission electron microscopy (STEM). In this letter, we present results from both cross-sectional and plan-view TEM studies of epitaxial BST films including the direct observation of dislocation cores at the interfaces between the films and substrates.

The BST samples used in this research were grown on (001) LAO by using pulsed laser deposition. Details on the growth procedure and the physical properties of the as-grown BST films can be found in the literature.[9] Briefly, the BST films are grown in oxygen pressures from 200 to 350 mTorr and at 750 to 830 °C to a thickness of typically 200 nm. X-ray diffraction indicates the films to be oriented with the [001] directions parallel to the ⟨102⟩ directions of the rhombohedral LAO. As the cell of LAO is close to cubic, this orientation is equivalent to a cube-on-cube orientation for a cubic substrate. Rocking curve measurements indicate a full width at half maximum (FWHM) of 0.15°. Samples for cross-sectional TEM studies were prepared using a standard procedure consisting of gluing, cutting, mechanical polishing, dimpling, and ion milling. Plan-view specimens were simply prepared by mechanical polishing and dimpling to thin out the substrate, followed by Ar-ion milling. Both plan-view and cross-sectional samples were ion milled by using an E.A. Fichione Ion Polishing System with an initial acceleration voltage of 3.5 kV, reducing to 1.0 kV at the end. Electron microscopy was carried out using a Philips EM-400 electron microscope operated at 100 kV for bright field images and selected area electron diffraction (SAED) patterns and a VG HB603 STEM at 300 kV for Z-contrast high-resolution images.

Figure 1 shows a bright field TEM image of a plan-view BST sample [Fig. 1(a)] and the selected area electron diffraction pattern taken from the same sample [Fig. 1(b)]. As seen in Fig. 1(a), the film consists of aligned rectangular-shaped

[a)]On leave from Beijing Laboratory of Vacuum Physics, Institute of Physics & Center for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100080, P.R. China.
[b)]Electronic mail: pky@ornl.gov

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

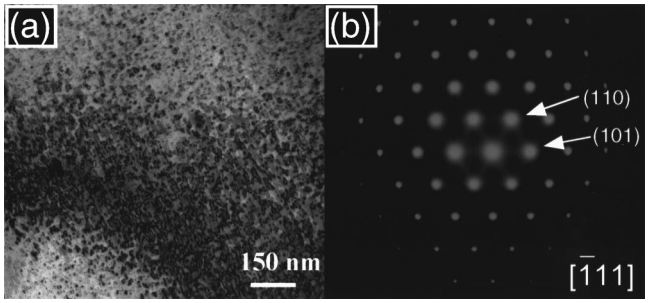


FIG. 1. TEM micrograph (a) and corresponding electron diffraction pattern (b) of a laser ablation deposited $Ba_{0.5}Sr_{0.5}TiO_3$ thin film, showing the formation of a single crystalline $Ba_{0.5}Sr_{0.5}TiO_3$ film. The electron projection is along the ⟨111⟩ zone axe of the $Ba_{0.5}Sr_{0.5}TiO_3$ thin film.

domains formed during the film growth. The SAED pattern was taken along the BST [111] direction. The sharp diffraction spots show that the *as*-grown BST films are good single crystals.

In order to understand better the film quality and the epitaxial behavior of the BST films on (001) LAO substrates, investigations of cross-sectional specimens were conducted using electron diffraction, bright field imaging, and *Z*-contrast imaging. Figure 2(a) is a low magnification bright field TEM image of the cross-sectional sample. The as-grown film has a flat surface with about 1 nm roughness, a sharp interface with the substrate, and a very uniform thickness of about 200 nm over the entire specimen. The columnar texture is due to small rotations between the grains, and is consistent with the FWHM of the x-ray rocking curve results. No large angle grain boundaries were found in the entire cross-sectional sample, indicating a near perfect single crystal BST film. Figure 2(b) is a SAED pattern taken in the area covering both the film and the substrate and along a

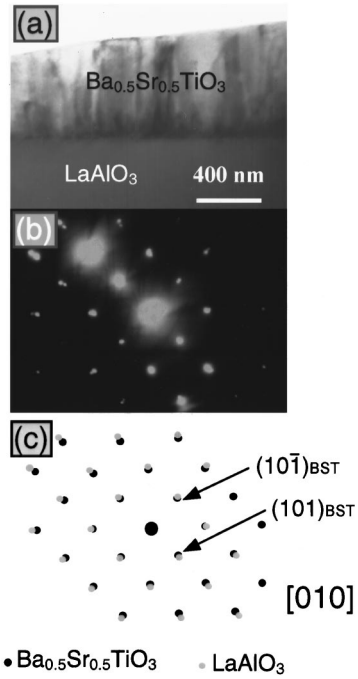


FIG. 2. TEM micrograph (a) and corresponding electron diffraction pattern (b) of both the substrate and the grown $Ba_{0.5}Sr_{0.5}TiO_3$ film. The crystalline orientation relation of the substrate and the film is deduced in (c).

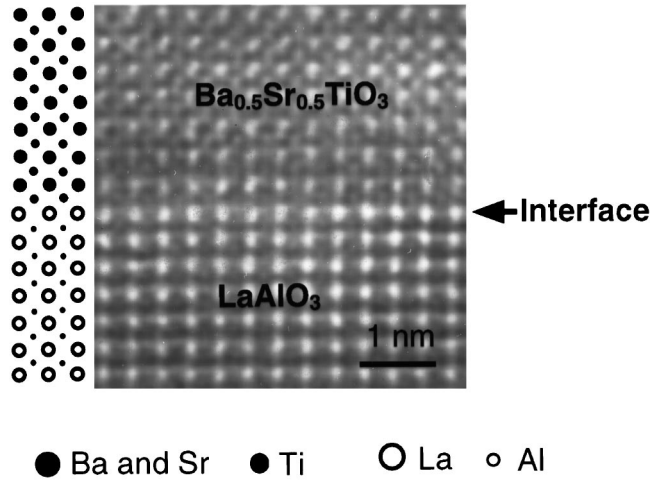


FIG. 3. *Z*-contrast STEM dark field image of the interface structure between the substrate $LaAlO_3$ and the grown $Ba_{0.5}Sr_{0.5}TiO_3$ film. Atom column positions can be clearly seen in both the substrate and the film, as shown in the schematic.

[010] direction of the LAO (using pseudocubic notation). The superposition of patterns from the BST film and the LAO substrate is indexed in Fig. 2(c). The ''cube-on-cube'' orientation is clearly observed with the epitaxial relationship $[010]_{BST} \| [010]_{LAO}$ and $(101)_{BST} \| (101)_{LAO}$. The lattice mismatch was estimated to be about 4.5% from the high order spots in the diffraction patterns, which is in good agreement with the BST and LAO lattice parameters. This result implies that the film is fully relaxed, with edge dislocations between the film and the substrate every 22 unit cells along the BST [100] direction.

High-resolution, cross-sectional electron microscopy studies of the as-grown films provide detailed information on the atomic structure of the interface. Figure 3(a) shows a *Z*-contrast STEM[17,18] image taken along the $[010]_{LAO}$ zone axis, that shows the interface between the film and the substrate is atomically sharp and flat. The stacking sequence across the interface is clearly visible. The last plane of the substrate is seen to be the La plane, while growth of the BST is seen to initiate with the Ti plane. No precipitates or other phases are present in the film or at the interface.

Edge dislocations were seen periodically along the entire interface, with a period of the edge dislocations of about 22 unit cells in BST lattice or 23 unit cells in the LAO lattice, which is in good agreement with the 4.3% lattice mismatch. This confirms that the BST film is essentially fully relaxed. Figure 4(a) shows a *Z*-contrast STEM image of the interface between LAO and the BST, with the approximate locations of misfit dislocation cores arrowed. Their cores lie a few unit cells inside the BST film. Between the cores the $Ba_{0.5}Sr_{0.5}$ plane of the BST is matched to the La plane of the LAO, and the Ti with the Al, as expected. Even in thin regions the misfit dislocation cores did not show the clear reconstructed form seen in $SrTiO_3$ grain boundaries[19] and $SrTiO_3/SrZrO_3$ interfaces.[20] As seen in the higher magnification images of Fig. 4(b), there is a strong tendency for the dislocations in this material to decompose into partials, $a[100] \rightarrow a/2[101] + a/2[10\text{-}1]$.

It should be pointed out that no twins were observed

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

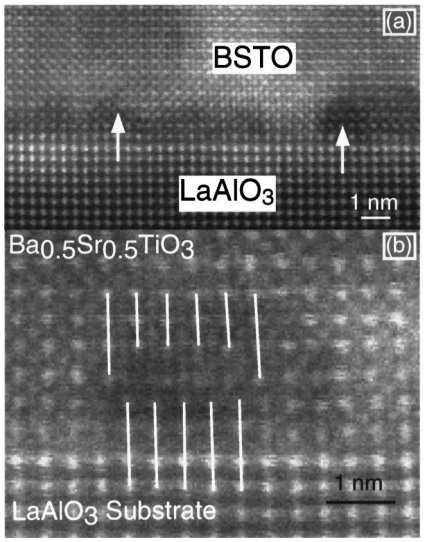


FIG. 4. (a) *Z*-contrast STEM dark field image of the interface between the substrate $LaAlO_3$ and the $Ba_{0.5}Sr_{0.5}TiO_3$ film showing misfit dislocation spacing, (b) high-resolution *Z*-contrast image showing a misfit dislocation decomposed into two partial dislocation cores separated by a small segment of stacking fault. The cores are located ~3 unit cells inside the film.

inside the BST films grown under these conditions. Threading dislocations must be present at the small angle grain boundaries between the columnar grains. Using the 0.15° FWHM and the ~50 nm grain size observed from TEM, a dislocation density of approximately $2\times10^9\,\mathrm{cm}^{-2}$ is estimated. Therefore the BST films have excellent single crystal quality. Although it is not well understood how the physical properties of high quality BST films relate to their microstructures, we would anticipate that single crystal films of this quality should display similar behavior and physical parameters with bulk materials. This is why the as-grown BST films exhibit a very high dielectric constant ϵ_r of about 1500 that is competitive with that of BST bulk single crystals, a large tunability of 33% under an electric field of 2.33 V/μm at room temperature, and very low loss tangent δ of only about 0.007.[9]

In summary, single crystalline $Ba_{0.5}Sr_{0.5}TiO_3$ thin films epitaxially grown on (001) LAO by pulsed laser deposition were characterized by SAED, TEM, and STEM. The as-grown films had a columnar structure, with an orientation relationship of $[010]_{BST}\|[010]_{LAO}$ and $(100)_{BST}\|(100)_{LAO}$. *Z*-contrast STEM images reveal the misfit dislocations to be dissociated, but the interfaces were seen to be atomically smooth and atomically abrupt.

This research was partially sponsored by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation, and by appointment to the ORNL. Postdoctoral Research Program administered jointly by ORISE and ORNL. The research at University of Houston was also partially supported by the NSF MRSEC research program of No. DMR-9632667, and the state of Texas through the Texas Center for Superconductivity at the University of Houston.

[1] J. S. Scott, Ferroelectr. Rev. **1**, 1 (1998).
[2] For example, *Epitaxial Oxide Thin Films III* edited by D. G. Schlom, C. B. Eom, M. E. Hawley, C. M. Foster, and J. S. Speck (Materials Research Society, Pittsburg, 1997), Vol. 474; *Ferroelectric Thin Films II*, edited by A. I. Kingon, E. R. Myers, and B. A. Tuttle (Materials Research Society, Pittsburgh, 1991), Vol. 243.
[3] O. Auciello, J. F. Scott, and R. Ramesh, Phys. Today **51**, 22 (1998).
[4] A. T. Findikoglu, Q. X. Jia, I. H. Campbell, X. D. Wu, D. Reagor, C. B. Mombourquette, and D. McMurry, Appl. Phys. Lett. **66**, 3674 (1995).
[5] F. W. Van Keuls, R. R. Romanofsky, D. Y. Bohman, M. D. Winters, F. A. Miranda, C. H. Mueller, R. E. Treece, T. V. Rivkin, and D. Galt, Appl. Phys. Lett. **71**, 3075 (1997).
[6] F. A. Miranda, R. R. Romanofsky, F. W. Van Keuls, C. H. Mueller, R. E. Treece, and T. V. Rivkin, Integr. Ferroelectr. **17**, 231 (1997).
[7] J. S. Horwitz, J. M. Pond, B. Tadayan, R. C. Y. Auyeung, P. C. Dorsey, D. B. Chrisey, S. C. Qadri, and C. Muller, Mater. Res. Soc. Symp. Proc. **361**, 515 (1995).
[8] K. R. Carroll, J. M. Pond, D. B. Chrisey, J. S. Horwitz, R. E. Leuchtner, and K. S. Grabowski, Appl. Phys. Lett. **62**, 1845 (1993).
[9] C. L. Chen, H. H. Feng, Z. Zhang, A. Brazdeikis, Z. J. Huang, W. K. Chu, C. W. Chu, F. A. Miranda, F. W. Van Keuls, R. R. Romanofsky, and Y. Liou, Appl. Phys. Lett. **75**, 412 (1999).
[10] J. S. Horwitz, W. Chang, A. C. Carter, J. M. Pond, S. W. Kirchoefer, D. B. Chrisey, J. Levy, and C. Hubert, Integr. Ferroelectr. **22**, 279 (1998).
[11] K. R. Carrol, J. M. Pond, D. B. Chrisey, J. S. Horwitz, R. E. Leuchtner, and K. S. Grabowski, Appl. Phys. Lett. **62**, 1845 (1993).
[12] S. G. Yoon, J. C. Lee, and A. Safari, J. Appl. Phys. **76**, 2999 (1994).
[13] K. Abe and S. Komatsu, Jpn. J. Appl. Phys., Part 1 **33**, 5297 (1994).
[14] X. D. Wu, S. R. Foltyn, R. C. Dye, Y. Coulter, and R. E. Muenchausen, Appl. Phys. Lett. **62**, 2434 (1993).
[15] F. Wang, A. Uusimaki, and S. Leppavuori, Appl. Phys. Lett. **67**, 1692 (1995).
[16] C. Bjormander, A. M. Grishin, B. M. Moon, J. Lee, and K. V. Rao, Appl. Phys. Lett. **64**, 3646 (1994).
[17] S. J. Pennycook, ''STEM: Z-contrast,'' in *Handbook of Microscopy*, edited by S. Amelinckx, D. van Dyck, J. van Landuyt, and G. van Tendeloo (VCH, Weinheim, Germany, 1997), p. 595.
[18] A. J. MaGibbon, S. J. Pennycook, and J. E. Angelo, Science **269**, 519 (1995).
[19] Miyoung Kim, N. D. Browning, S. J. Pennycook, K. Sohlberg, and S. T. Pantelides, Microsc. Microanal. **5** (suppl.2), 110 (1999).
[20] S. J. Pennycook, P. D. Nellist, N. D. Browning, P. A. Langjahr, and M. Rühle, in Proceedings of the 1996 Ann. MSA Meeting, San Francisco, California, 1996 (unpublished), p. 104.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

Cathodoluminescent properties at nanometer resolution through *Z*-contrast scanning transmission electron microscopy

H.-J. Gao,[a] G. Duscher, M. Kim, and S. J. Pennycook[b]
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030

D. Kumar, K. G. Cho, and R. K. Singh
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611-6400

(Received 31 January 2000; accepted for publication 31 May 2000)

We report the observation of porous structures in laser-ablation-deposited Y_2O_3:Eu thin films and their correlation with luminescent properties by a combination of transmission electron microscopy and *Z*-contrast scanning transmission electron microscopy (*Z*-STEM). Depending on growth conditions, a large density of voids is incorporated into the films, which leads to a much increased surface area. Cathodoluminescence imaging in the STEM directly reveals a 5 nm ''dead layer'' around each void, which is responsible for the observed reduction in luminescence efficiency. © *2000 American Institute of Physics.* [S0003-6951(00)05229-3]

Rare earth-activated oxide-based phosphor thin films are promising candidates for high-resolution display devices such as cathode ray tubes, graphics, field emission displays (FEDs), and thermomechanical devices.[1–5] However, thin-film phosphors typically have a significantly reduced brightness compared to equivalent powder phosphor materials.[6,7] Several possible explanations have been suggested for the lower brightness including internal reflection or channeling of the emitted light along transverse axes parallel to the surface instead of vertical emission from the surface, and the small interaction volume between the incident beam and the solid. In this letter, we show another factor to be crucial to external radiative efficiency, the porosity of the films. Porosity creates internal surface area, and the associated dead layer decreases the emission efficiency. The dead layer is directly observed by simultaneous cathodoluminescence (CL) imaging and *Z*-contrast imaging in the scanning transmission electron microscope (STEM), and quantitatively accounts for the reduction of luminescent efficiency.

Eu-activated Y_2O_3 thin films with thickness of about 200 nm were deposited on (001) $LaAlO_3$ substrates by laser ablation at different temperatures (735, and 775 °C) and different laser pulse frequencies (5 and 10 Hz).[8,9] A Lambda Physik 305I laser ($\lambda = 248$ nm, $\tau = 25$ ns) is used to irradiate the Eu-doped Y_2O_3 phosphor target (YOE), creating an expanding plume of atomic species. Detailed information about the deposition can be found elsewhere.[10] Cross-sectional slices were obtained by cutting the sample along the [100] or [010] directions of YOE, and then gluing face to face. Both plan-view and cross-section specimens for transmission electron microscopy (TEM) and/or STEM observations were prepared by mechanical grinding, polishing, and dimpling, followed by Ar-ion milling using a E. A. Fichione Ion Polishing System first at a voltage of 3.5 kV and an angle of 12°, and finally at 1.0 kV and 10°. TEM bright field images and the electron diffraction patterns were recorded in a Philips EM-400 electron microscope operated at 100 kV. *Z*-contrast imaging was conducted in VG HB603 STEM at 300 kV,[11,12] while the CL imaging was conducted in a VG HB501 STEM at 100 kV. The CL emission was collected by a lens system and detected by a photomultiplier.[13]

The porous structure of YOE thin films was first demonstrated by TEM images of plan view samples. For comparison, three TEM plan view samples of YOE thin films were prepared at different substrate temperatures and laser pulse rates: (a): 735 °C, 10 Hz, (b) 735 °C, 5 Hz (c) 775 °C, 10 Hz. As shown in Fig. 1(a), the sample prepared at 735 °C, 10 Hz shows a high density of pores at the edge of the sample, with an average size of about 25 nm. When the pulse rate was lowered to 5 Hz, but the substrate temperature was kept at 735 °C, the pore size decreased to 17 nm, as shown in Fig. 1(b). When the substrate temperature was increased to 775 °C with 10 Hz pulse rate, the voids formed were much smaller, ~10 nm, as shown in Fig. 1(c). These observations show that the porosity of the film decreases with higher substrate temperature and with lower laser pulse rate.

In order to check the distribution of the pores through the film thickness, cross-section samples were examined. A *Z*-contrast STEM image of the sample grown at 775 °C, 10 Hz is shown in Fig. 2. Figure 2(a) shows a large area dark field *Z*-contrast STEM image in which both the substrate and the film can be seen. Above the substrate and through the film, some columnar structures and prolonged darker areas are clearly seen extending from the surface to the interface. The brightness in the *Z*-contrast STEM image depends on both the atomic number and the sample thickness. At a specific thickness condition, the brightness will only be determined by Z^2 (where Z is the atomic number), but for a constant Z, the brightness will reflect the film thickness. So, in our case, the darker regions imply that the pores are columnar in nature. This is confirmed by the following two higher magnification images. In Fig. 2(b), two columnar pores are seen. In the thinnest region (near the top of the image), they both form holes in the specimen. As the sample becomes thicker, a columnar track is seen extending from the left hand

[a]On leave from Beijing Laboratory of Vacuum Physics, Institute of Physics & Center for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100080, China.
[b]Electronic mail: pyk@ornl.gov

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

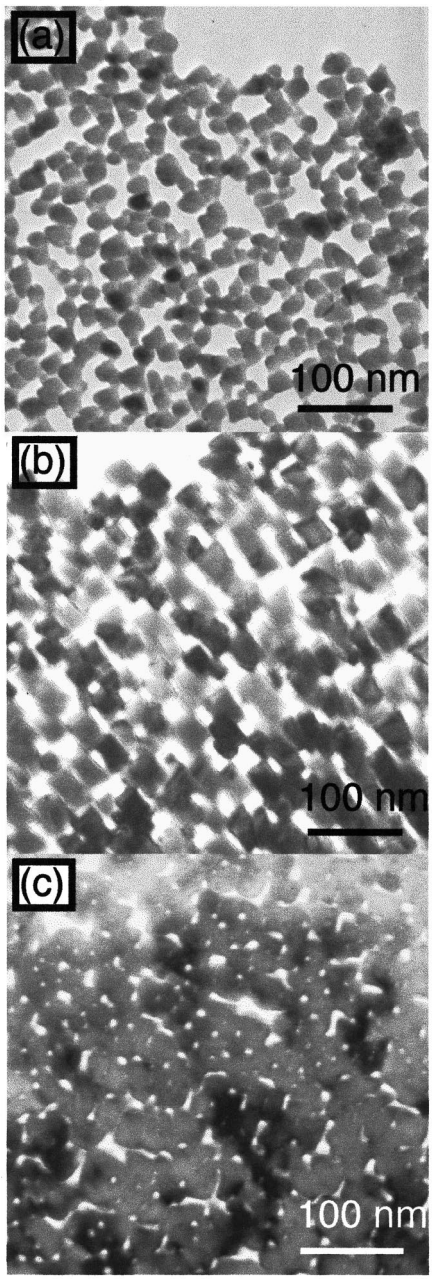


FIG. 1. Three plan-view TEM micrographs of Y_2O_3:Eu thin films prepared at different deposition conditions, showing clearly the variation of void size with substrate temperature and deposition rate, (a), 735 °C, 10 Hz, (b) 735 °C, 5 Hz, (c) 775 °C, 10 Hz.

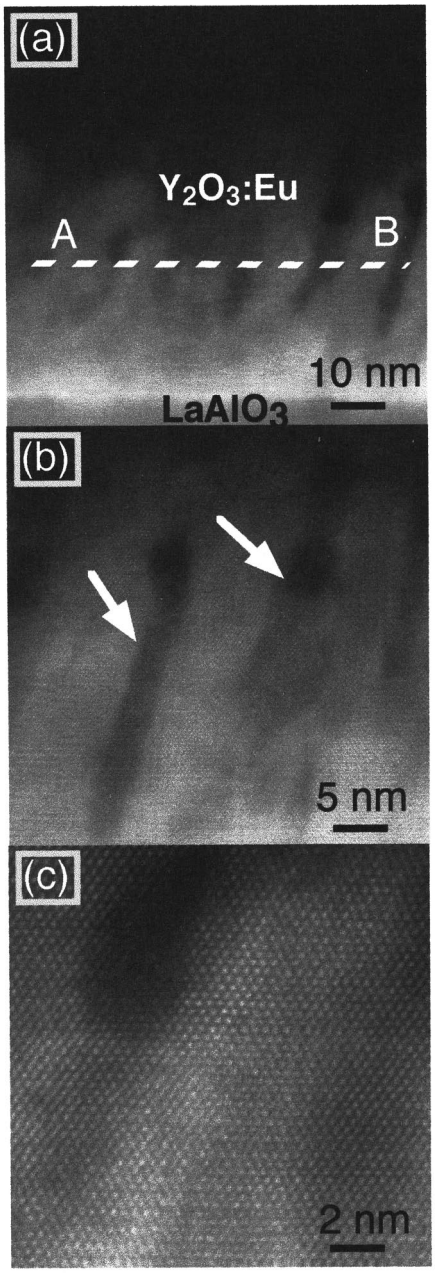


FIG. 2. (a) Dark field *Z*-contrast STEM image of a cross-section Y_2O_3:Eu sample prepared by laser ablation at a substrate temperature of 775 °C and in fast deposition mode (10 Hz), showing the columnar nature of the pores. (b) and (c) are higher magnification images of the region in (a).

pore (arrowed), while the track from the other intersects the sample surface. This contrast shows the pores to be columnar in nature, which is consistent with our previous TEM investigation showing the columnar structure to be formed during the laser ablation film growth.[10] In Fig. 2(c), we can see clearly the well ordered atomic structure around a pore. In agreement with our previous observation,[10] we see that the film is an almost perfect single crystal.

The porous structure has a dramatic effect on the external radiative efficiency of the film. This is demonstrated by measuring the efficiency of different samples prepared under different deposition condition, and, therefore, containing different pore sizes.[5,7,14] The photoluminescence efficiency correlates directly with average pore size, as shown in Fig. 3, and we would expect the CL efficiency to be similar. The reason for this correlation is expected to be the existence of a ''dead layer'' near the specimen surface, resulting from nonradiative recombination of electron–hole pairs via surface states.[15–17] Increasing the pore size would create more internal surfaces per unit volume, resulting in lower overall efficiency.

Direct measurement of the dead layer at the surface of a pore has been achieved by comparing the *Z*-contrast image with the CL image obtained simultaneously. Figure 4(a) shows a *Z*-contrast image of a region of film containing some relatively widely spaced pores. The corresponding CL image in Fig. 4(b) shows the variation in CL intensity. It is immediately clear that the edges of the holes in the CL image are much less sharp. Intensity profiles across the hole marked are shown in Figs. 4(c) and 4(d). The width of the pore in the

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

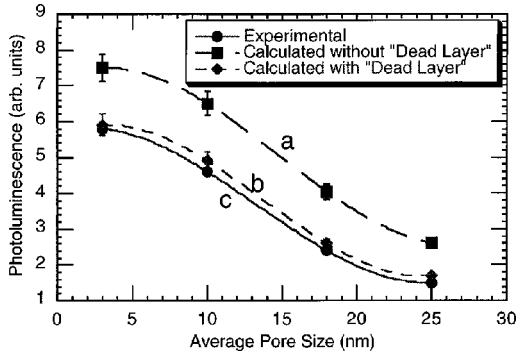


FIG. 3. Relationship between the photoluminescence efficiency and the pore size, showing the decrease in efficiency with increasing pore size. Samples are those shown in Fig. 1, with the addition of a film grown at 775 °C, 5 Hz, which has the highest efficiency. Also shown are theoretical curves with and without the effect of the dead layer.

Z-contrast image is 10 nm, whereas in the CL image is more than doubled to 20 nm. Clearly, if the grain size of a film is comparable to the extent of the dead layer, then the efficiency will be substantially reduced. This explains the order of magnitude reduction in efficiency observed experimentally in Fig. 3. It is clear that because of this dead layer porous structures can lead to greatly reduced emission efficiencies.

The relationship between the decrease of emission efficiency and the pore size can be estimated directly from our plan view images. Assuming the grains to be columnar, extending throughout the film thickness, the relative efficiency can be estimated from the total area of grains observed in the image per unit area of film, giving curve (a) in Fig. 3. Now, if the area of each grain is reduced to take into account the 5 nm dead layer, the relative efficiency is significantly reduced (curve b), resulting in values that agree well with our experimental observations (curve c).

Singh and co-workers[5,7,14,18] have shown that the surface roughness of a film has a close relationship with the luminescent properties, i.e., increased roughness of the film will increase the efficiency. Now, based on our TEM and *Z*-contrast STEM observations, we see that another critical issue is the porosity of the films.

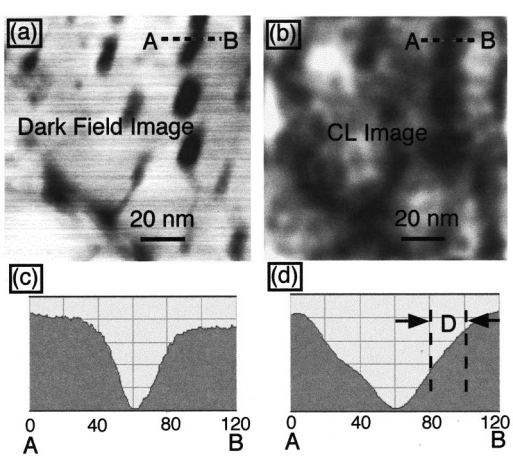


FIG. 4. *Z*-contrast image (a), CL image, (b) and corresponding intensity profiles (c) and (d) across the same hole, showing the ''dead layer'' to be about 5 nm in width (labeled D).

In summary, we have directly correlated the pore structure in laser ablation deposited YOE thin films with luminescent efficiency. A high density of voids creates more internal surface in the films per unit volume. The resulting increase in the surface dead layer reduces external luminescence efficiency. By simultaneous *Z*-contrast and CL imaging, we have shown the extent of the dead layer to be ~5 nm, and obtained quantitative agreement between the measured pore size and luminescent efficiency. We suggest that in the preparation of YOE and related thin films, attention should be paid to avoiding the formation of internal voids in order to increase the radiative efficiency.

The authors are grateful to A. Kadavanich for assistance with the CL detection system. The work at ORNL was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation, and by appointment to the ORNL Postdoctoral Research Program administrated jointly by ORISE and ORNL. The work at University of Florida was supported by the Phosphor Technology Center of Excellence by DARPA Grant No. MDA972-93-1-0030 and Department of Energy, under Contract No. DE-FG05-95ER45533.

[1] K. A. Wickersheim and R. A. Lefever, J. Electrochem. Soc. **111**, 47 (1964).
[2] M. Sayer and K. Sreenivas, Science **247**, 1056 (1990).
[3] G. Blasse and B. C. Grabmaier, *Lumni. Mater.* (Springer, Berlin, 1994).
[4] S. J. Duclos, C. D. Greskovich, and C. R. O'Clair, Mater. Res. Soc. Symp. Proc. **348**, 503 (1994).
[5] S. L. Jones, D. Kumar, R. K. Singh, and P. H. Holloway, Appl. Phys. Lett. **71**, 404 (1997).
[6] R. P. Rao, Solid State Commun. **99**, 439 (1996).
[7] K. G. Cho, D. Kumar, S. L. Jones, D. J. Lee, P. H. Holloway, and R. K. Singh, J. Electrochem. Soc. **145**, 3456 (1998).
[8] J. Fitz-Gerald, T. Trottier, R. K. Singh, and P. H. Holloway, Appl. Phys. Lett. **72**, 1838 (1998).
[9] D. Kumar, J. Fiz-Gerald, and R. K. Singh, Appl. Phys. Lett. **72**, 1451 (1998).
[10] H-J. Gao, D. Kumar, K. G. Cho, P. H. Holloway, R. K. Singh, X. D. Fan, Y. Yan, and S. J. Pennycook, Appl. Phys. Lett. **75**, 2223 (1999).
[11] S. J. Pennycook, in *Handbook of Microscopy*, edited by S. Amelinckx (VCH, Weinheim, 1997), p. 595.
[12] N. D. Browning, M. F. Chisholm, and S. J. Pennycook, Nature (London) **366**, 143 (1993).
[13] H.-J. Gao, G. Duscher, X. D. Fan, S. J. Pennycook, D. Kumar, K. G. Cho, P. H. Holloway, and R. Singh, Mater. Res. Soc. Symp. Proc. (in press).
[14] K. G. Cho, D. Kumar, P. H. Holloway, and R. K. Singh, Appl. Phys. Lett. **73**, 3058 (1997).
[15] J. Yuan, S. D. Berger, and L. M. Brown, J. Phys.: Condens. Matter **1**, 3253 (1989).
[16] B. G. Yacobi and D. B. Holt, *Cathodoluminescence Microscopy of Inorganic Solids* (Plenum, New York, 1990).
[17] D. B. M. Klaassen and D. M. de Leeuw, J. Lumin. **37**, 21 (1987).
[18] K. G. Cho, D. Kumar, D. J. Lee, S. L. Jones, P. H. Holloway, and R. K. Singh, Appl. Phys. Lett. **71**, 3335 (1997).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

Epitaxial growth of Y_2O_3:Eu thin films on $LaAlO_3$

H-J. Gao[a)]
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030 and Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611-6400

D. Kumar, K. G. Cho, P. H. Holloway, and R. K. Singh
Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611-6400

X. D. Fan, Y. Yan,[b)] and S. J. Pennycook[c)]
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030

(Received 25 May 1999; accepted for publication 20 August 1999)

We report the epitaxial growth of europium-activated yttrium oxide (Y_2O_3:Eu) (001) thin films on $LaAlO_3$ (001) using laser ablation deposition at a substrate temperature of 775 °C and 10 Hz pulse repetition rate. The orientation relationship between the films and the substrates is $[110]Y_2O_3 \| [100]LaAlO_3$ and $[-110]Y_2O_3 \| [010]LaAlO_3$ which results in a lattice mismatch of only 0.8%. Transmission electron microscopy (TEM) of the films reveals the single crystalline Y_2O_3:Eu thin film to contain small pores. Scanning transmission electron microscopy (STEM) imaging of the films shows the substrate always terminates with the Al sublattice. Moreover, the STEM reveals that no precipitates of Eu had formed in the films. © *1999 American Institute of Physics.* [S0003-6951(99)05441-8]

Numerous efforts have been made to prepare yttrium oxide (Y_2O_3) thin films for applications in ultralarge scale integration (ULSI) gate insulators, ULSI capacitors, and, by addition of a suitable dopant species, for electroluminescence devices.[1–4] Single crystalline thin films are of great importance for devices because of their reduced defect content which can result in improved properties. Thus, growth of single crystal films has attracted much attention, and in the past decade many different deposition techniques and substrates have been tried.[5–13] Fukumoto, Choi, and co-workers[5,6] reported the heteroepitaxial growth of Y_2O_3 (YO) films on silicon, but the atomic structure of the interface was not determined. In this letter, we report the epitaxial growth of YO thin films doped with ~4 wt % Eu on a $LaAlO_3$ substrate. *Z*-contrast scanning transmission electron microscopy (STEM)[14] is employed to directly image the atomic structure of the film/substrate interface.

It is well known that the Y_2O_3 has a *C*-type rare-earth sesquioxide structure, closely related to the fluorite structure with a cell parameter $a = 1.060$ nm and space group $T_h^7(Ia3)$.[15–18] In the fluorite lattice, each cation is surrounded by eight anions located at the corner of a cube. The *C*-type structure is derived by removing one quarter of oxygen atoms and slightly rearranging the remaining ones.[8,19] For 75% of the cations the vacancies lie at the ends of a face diagonal, while for the other 25% they lie at the ends of a body diagonal. Therefore, each yttrium atom is surrounded by only six oxygen neighbors forming two different types of distorted octahedral structure in the unit cell, called S_6 and C_2.[20] Eight yttrium atoms have the S_6 symmetry and the other 24 atoms have the C_2 symmetry. From the crystallographic structure one can deduce that the distance of two neighboring Y atoms along the ⟨100⟩ direction of YO is 0.5302 nm, and along the ⟨110⟩ direction it is 0.375 nm. $LaAlO_3$ (LAO) is a rhombohedral structure with lattice parameters $\mathbf{a} = 0.378$ nm, $\theta \leqslant 90.5°$, very close to a cubic structure. The lattice mismatch with the ⟨110⟩ direction of the YO is therefore less than 0.8%, and so we would anticipate epitaxial growth of single crystalline YO thin films on the LAO (001) substrate to be feasible.

Eu activated YO thin films were deposited by laser ablation[21–23] on (001) LAO substrates at a temperature of 775 °C and a deposition rate of 10 Hz. Pulses from a Lambda Physik 305i laser (wavelength 248 nm, pulse length 25 ns) are passed through a quartz window to irradiate the phosphor target material, YO:Eu, creating an expanding plume. The typical distance between the target and the substrate was ~5 cm, and energy densities were approximately 2.0–3.5 J/cm^2. The films were grown on LAO to a total thickness of 0.3 μm, in an oxygen ambient pressure that ranged from 50 to 600 mTorr. Rocking curve measurements indicate a full width at half maximum (FWHM) of 0.1°. Cross sectional slices were obtained by cutting the LAO along the [100] or [010] directions (using pseudocubic indexing) and then gluing face to face in the usual way. Both plan view and cross section specimens were prepared for transmission electron microscope (TEM) and/or STEM observations by mechanical grinding, polishing, and dimpling, followed by Ar-ion milling using an E. A. Fichione ion polishing system, a 13° incident angle, and a beam voltage of 3.5 kV initially, reducing to 1.0 kV for final milling. TEM bright field images and electron diffraction patterns were recorded in a Philips EM-400 electron microscope operated at 100 kV. *Z*-contrast imaging was conducted in a VG HB603 STEM at 300 kV.[14,24]

Figure 1 is a low magnification TEM micrograph and

[a)]On leave from Beijing Laboratory of Vacuum Physics, Institute of Physics & Center for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100080, P.R. China.
[b)]Now at National Renewable Energy Laboratory, 1617 Cole Blvd. Golden, CO 80401-3303.
[b)]Electronic mail: pky@ornl.gov

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

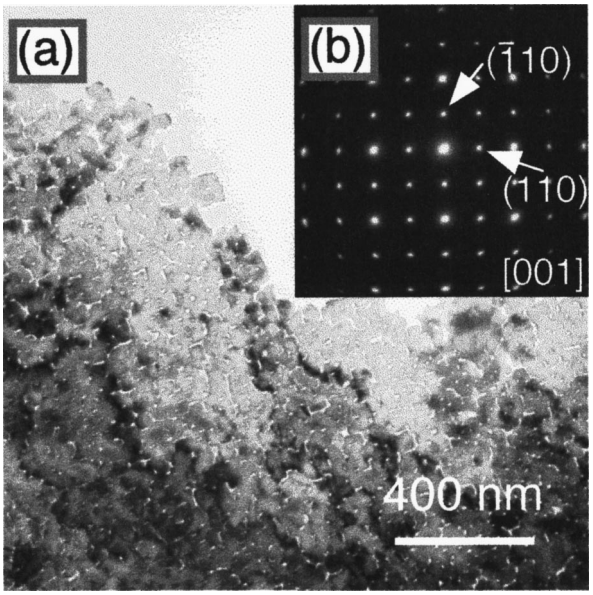


FIG. 1. Plan view TEM image (a) and corresponding electron diffraction pattern (b) of a laser ablation deposited YO:Eu thin film, showing the formation of a good single crystalline film containing numerous voids. The electron projection is along the [001] zone axes of the YO:Eu crystalline thin film.

corresponding selected area electron diffraction pattern (EDP) of a plan view sample of the YO:Eu thin film. The diffraction pattern indicates an almost perfect single crystal film, but the image shows numerous small voids, suggesting an island growth mechanism with incomplete coalescence of the islands.

A cross section image of the sample is presented in Fig. 2(a), showing the smooth surface, sharp interface, and a uniform thickness of 300 nm maintained over the entire region. Figure 2 also shows selected area EDPs of the film (b) and the LAO substrate (c), showing the orientation relationship to be [110]YO‖[100]LAO and [−110]YO‖[010]LAO. The columnar structure of the film is also apparent from the cross section image, with small rotations between neighboring grains giving the strong diffraction contrast. Each individual column, however, appears to be a good single crystal, which implies that the presence of the voids may avoid the need for a high density of dislocations between the grains to accommodate the rotations, and/or a high level of stress within the grains. The dominant direction of the voids is not crystallographic, suggesting that it is related to the deposition direction not being normal to the substrate.[25] The sample was not rotated during film deposition.

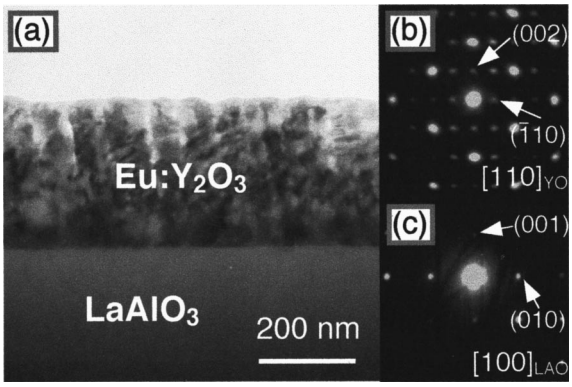


FIG. 2. Cross section TEM image (a) and corresponding electron diffraction patterns of the as-grown YO:Eu film (b) and the LAO substrate (c) showing the orientation relationship to be [110]YO‖[100]LAO and [−110]YO‖[010]LAO.

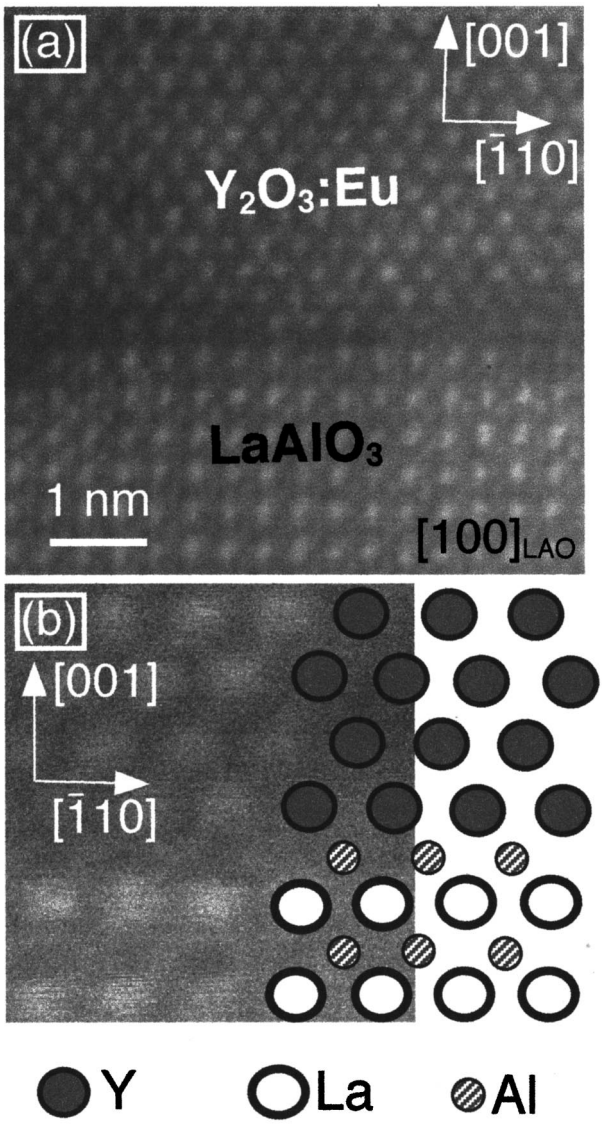


FIG. 3. (a) *Z*-contrast STEM dark field image showing the atomically abrupt interface, (b) higher magnification image showing clearly the Al terminated substrate, as shown in the schematic.

In order to determine the detailed interface atomic structure, high-resolution *Z*-contrast STEM imaging of the samples was carried out. The *Z*-contrast image is a direct image with intensity highly localized about the atomic column positions and approximately proportional to the mean square atomic number (Z). Thus the La and Al columns in LAO, and the Y columns in YO, are directly distinguishable in a *Z*-contrast image taken along the [010] zone axis of the LAO substrate. Figure 3(a) is an atomic resolution *Z*-contrast STEM image of the film/substrate interface. The bright spots in the film are Y columns, the bright spots in the substrate La columns, and the less bright spots Al columns. The O columns are not visible. Also shown in Fig. 3(b) is a higher magnification *Z*-contrast image that shows clearly the atomic structure of the interface. The substrate is seen to terminate with the Al plane, which matches directly onto the Y layer of the film as shown in the schematic.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

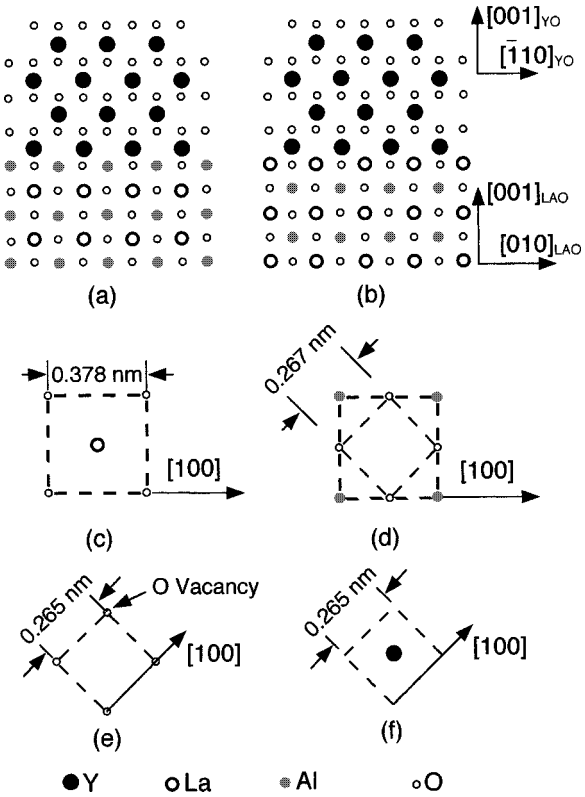


FIG. 4. Schematic interface structures for YO on the Al terminated surface (a), and on the La terminated surface of LAO (b). Atomic arrangements on the substrate (001) surface are shown for La termination (c), and Al termination (d). The atomic arrangements on the alternating oxygen and Y {004} planes in YO are shown in (e) and (f).

Figure 4 is a schematic of two possible interface structures corresponding to the two possible terminations of the (001) substrate, either the $(AlO_2)^-$ or $(LaO)^+$ planes. Figures 4(c) and 4(d) show these two planes, while Figs. 4(e) and 4(f) show the two (001) planes of YO, comprising pure Y and O. The full unit cell of the YO structure is four times the dimensions shown due to ordering of the O vacancies. For $(AlO_2)^-$ termination of the substrate, the four oxygen positions match almost exactly the oxygen positions in the YO. The Y atoms can sit over the center of the four O positions in Fig. 4(d), directly over the La site in the plane below, as seen in the image of Fig. 3(b). The interfacial Y is then coordinated by seven oxygen atoms instead of six, which may be compensated by some additional oxygen vacancies. In contrast, if the substrate is terminated by LaO, each Y sitting directly over one oxygen in the LaO plane, then each Y is substantially undercoordinated. This explains the observed termination and the fact that no single layer height steps were observed.

Another important aspect of the work concerns the possibility of Eu precipitation in the YO films. *Z*-contrast STEM is very sensitive to Eu with $Z=63$. In our observations, however, no sign of any precipitation was found, or of any preferential segregation to the surfaces of the columnar grains. This implies that the Eu was distributed uniformly within the films.

In conclusion, we have successfully grown epitaxial single crystalline YO films doped with Eu on LAO substrates by laser ablation. *Z*-contrast STEM reveals the substrates are always Al terminated, and no evidence of nonuniform Eu distribution was found.

The work at ORNL was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation, and by appointment to the ORNL Postdoctoral Research Program administrated jointly by ORISE and ORNL. The work at the University of Florida was supported by the Phosphor Technology Center of Excellence by DARPA Grant No. MDA972-93-1-0030 and the Department of Energy, under Contract No. DE-FG05-95ER45533.

[1] L. Manchanda and M. Gurvitch, IEEE Electron Device Lett. **9**, 180 (1988).
[2] T. S. Kalkur, Y. R. Kwor, and C. A. Paz de Araujo, Thin Solid Films **170**, 185 (1989).
[3] S. J. Duclos, C. D. Greskovich, and C. R. O'Clair, Mater. Res. Soc. Symp. Proc. **348**, 503 (1994).
[4] G. Blasse and B. C. Grabmaier, *Lumni. Mater.* (Springer, Berlin, 1994).
[5] H. Fukumoto, T. Imura, and Y. Osaka, Appl. Phys. Lett. **55**, 360 (1989).
[6] S. C. Choi, M. H. Cho, S. W. Whangbo, C. N. Whang, S. B. Kang, S. I. Lee, and M. Y. Lee, Appl. Phys. Lett. **71**, 903 (1997).
[7] R. P. Rao, Solid State Commun. **99**, 439 (1996).
[8] K.-I. Onisawa, M. Fuyama, K. Tamura, K. Taguchi, T. Nakayama, and Y. A. Ono, J. Appl. Phys. **68**, 719 (1990).
[9] A. F. Jankowski, L. R. Schrawyer, and J. P. Hayes, J. Vac. Sci. Technol. A **11**, 1548 (1993).
[10] W. M. Cranton, D. M. Spink, R. Stevens, and C. B. Thomas, Thin Solid Films **226**, 156 (1993).
[11] S. L. Jones, D. Kumar, R. K. Singh, and P. H. Holloway, Appl. Phys. Lett. **71**, 404 (1997).
[12] K. G. Cho, D. Kumar, D. J. Lee, S. L. Jones, P. H. Holloway, and R. K. Singh, Appl. Phys. Lett. **71**, 3335 (1997).
[13] K. G. Cho, D. Kumar, S. L. Jones, D. J. Lee, P. H. Holloway, and R. K. Singh, J. Electrochem. Soc. **145**, 3456 (1998).
[14] S. J. Pennycook, ''STEM: *Z*-contrast,'' in *Handbook of Microscopy*, edited by S. Amelinckx, D. van Dyck, J. van Landuyt, and G. van Tendeloo (VCH, Weinheim, Germany 1997), p. 595.
[15] J. L. Daams, P. Villars, and J. H. N. Vanvucht, *Atlas of Crystal Structure Types for Intermetallic Phases* (ASM International, Materials Park, OH, 1994), p. 6706.
[16] M. G. Paten and E. N. Maslen, Acta Crystallogr. **19**, 307 (1965).
[17] B. H. O'Conner and T. M. Valentine, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **25**, 2140 (1969).
[18] M. Faucher and J. Pannetier, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. **36**, 3209 (1980).
[19] W. v. Schaik and G. Blasse, Chem. Mater. **4**, 410 (1992).
[20] F. Jollet, C. Noguera, N. Thromat, M. Gautier, and J. P. Duraud, Phys. Rev. B **42**, 7587 (1990).
[21] J. Fitz-Gerald, S. J. Pennycook, H. Gao, V. Krishnamoorthy, J. Marcinka, W. Glenn, and R. Singh, Mater. Res. Soc. Symp. Proc. **502** (1998).
[22] J. Fitz-Gerald, T. Trottier, R. K. Singh, and P. H. Holloway, Appl. Phys. Lett. **72**, 1838 (1998).
[23] D. Kumar, J. Fitz-Gerald, and R. K. Singh, Appl. Phys. Lett. **72**, 1451 (1998).
[24] N. D. Browning, M. F. Chisholm, and S. J. Pennycook, Nature (London) **366**, 143 (1993).
[25] E. S. Machlin, *Materials Science in Microelectronics* (Giro, Croton-on-Hudson, 1995).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

Reversible, Nanometer-Scale Conductance Transitions in an Organic Complex

H. J. Gao,[1,2] K. Sohlberg,[1] Z. Q. Xue,[2] H. Y. Chen,[2] S. M. Hou,[2] L. P. Ma,[2] X. W. Fang,[3] S. J. Pang,[2] and S. J. Pennycook[1]

[1]*Solid State Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6031*

[2]*Beijing Laboratory of Vacuum Physics, Institute of Physics and Center for Condensed Matter Physics, Chinese Academy of Sciences, Beijing 100080, People's Republic of China*

[3]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637*

(Received 23 June 1999)

Reversible conductance transitions are demonstrated on the molecular scale in a complex of 3-nitrobenzal malononitrile and 1,4-phenylenediamine, by application of local electric field pulses. Both macroscopic and local current-voltage (I/V) measurements show similar electrical bistability behavior. The mechanism of the electrical bistability is discussed.

PACS numbers: 73.61.Ph

Conjugated organic materials have aroused much attention recently [1]. Of particular interest is the utilization of individual molecules or molecular complexes as electronic device elements. Aviram *et al.* [2] proposed the donor-bridge-acceptor (D-Σ-A) model for molecular rectification using conjugated molecules. A crucial issue in this regard is whether or not a conductance transition can be induced in single molecules or nanoscale regions of a thin film, and the mechanism of the transition. The pioneering work of Aviram *et al.* [3] reported the first evidence of switching and rectification in an organic thin film, but the results were not conclusive [4]. More recently, Potember *et al.* [5] have shown a field-induced conductance transition on a 500 nm scale, but did not demonstrate *local* reversibility. The reverse transition was induced only by application of a broad laser pulse or heat. Here we report a *reversible* conductance transition in a crystalline fully organic complex, on a scale close to the dimensions of the unit cell.

The two conjugated organic compounds shown in Fig. 1, 3-nitrobenzal malonitrile (NBMN) and 1,4-phenylenediamine (pDA), were used to form the complex. NBMN was prepared as in the literature [6], and the two materials were mixed together in a 1:1 molar ratio and vacuum evaporated (10^{-6}–10^{-5} Torr). High resistivity polycrystalline NBMN-pDA films were prepared by evaporation onto room temperature substrates at a rate $\leq$5 nm/min. Films for conductance transition experiments were deposited to a thickness of 20 nm on freshly cleaved, highly oriented pyrolytic graphite (HOPG). Samples for macroscopic four-probe I/V measurements were deposited to a thickness of 200 nm on substrates of glass coated with conductive indium-tin-oxide films. For transmission electron microscopy (TEM), films were simultaneously deposited onto carbon films and subsequently examined in a JEOL 200CX. Coverage was found to be very uniform. X-ray diffraction showed the polycrystalline films to have a triclinic structure with unit cell parameters: $a = 0.7823(2)$ nm, $b = 0.8730(2)$ nm, $c = 0.7286(1)$ nm, $\alpha = 105.90(3)°$, $\beta = 101.49(3)°$, $\gamma = 78.75(3)°$. In contrast, for deposition rates $>$6 nm/min, the resulting films were conductive, and structural characterization showed these more rapidly deposited films to be amorphous. As discussed below, this result is a clue to the mechanism of the conductance transition.

We first performed standard four-point probe measurements of *I-V* relationships and transient conductance on the 200 nm thick films. The *I-V* measurements, shown in Fig. 2(a), demonstrate electrical bistability. Increasing the voltage stepwise from zero produced the open points in Fig. 2(a), but when the voltage was increased above 3.2 V, the film abruptly switched to a conductive state. In the conductive state, stepwise sampling of the current as a function of applied voltage produced the solid points in Fig. 2(a). The high-impedance state has a resistivity of $\approx 10^{-8}$ $\Omega \cdot$ cm, reducing to $\approx 10^{-4}$ $\Omega \cdot$ cm in the conductive state. We note that our films are significantly more conductive than those reported by Alvarado *et al.* [7]. The transition time of the film was measured to be about 80 ns, as shown in Fig. 2(b).

Local conductance transitions were induced in the high-resistivity films by a scanning tunneling microscope (STM). A typical image of the film is shown in Fig. 3(a). The dimensions of the surface unit cell are $b = 0.93$ nm, $c = 0.78$ nm, and the angle between $\hat{b}$ and $\hat{c}$ is about 100°, closely matching the unit cell dimensions determined by x-ray diffraction. This image was taken at a bias (V_b) of 0.1 V and tunneling current (I_t) of 0.4 nA. As shown in Figs. 3(b) and 3(c), a 3 × 3 array and an "A"

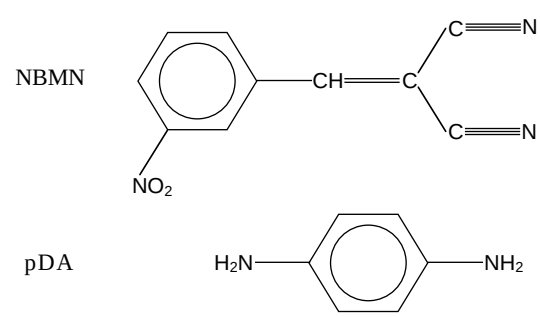


FIG. 1 Structures of the conjugated molecules NBMN and pDA.

 0031-9007/00/84(8)/1780(4)$15.00 © 2000 The American Physical Society

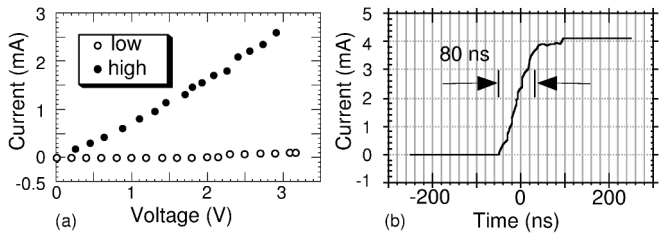


FIG. 2 (a) *I-V* relation of a 200 nm film, showing the low and high conductance states. The voltage threshold is 3.2 V. (b) Transient conductance measurements on the 200 nm film showing a transition time of 80 ns.

pattern could be formed by applying voltage pulses to the STM tip in constant height mode. Each bright spot in the image represents a high-conductance region, i.e., regions that have been switched from insulator to conductor. The shadow effect on each mark is due to the fact that the feedback circuits were not completely nulled during the scan. All marks were erased on heating the sample *in situ* above 423 K, when the conductance of the local regions recovered their original insulating states. It was possible to induce the reverse transition in *individual* marks by applying a reverse-polarity voltage pulse of −4.5 V for 50 μs, as shown in Figs. 3(d) and 3(e). If the same pulse was applied without reversing the polarity, the mark was incompletely erased. Therefore the process is due to the combined effect of applied field and a local heating induced by the current passing through the conducting region. In principle, the ultimate resolution should be the size of the molecular complex, ≈1 nm. Figure 3(f) demonstrates resolution on this length scale. The marks are separated by approximately 1.7 nm.

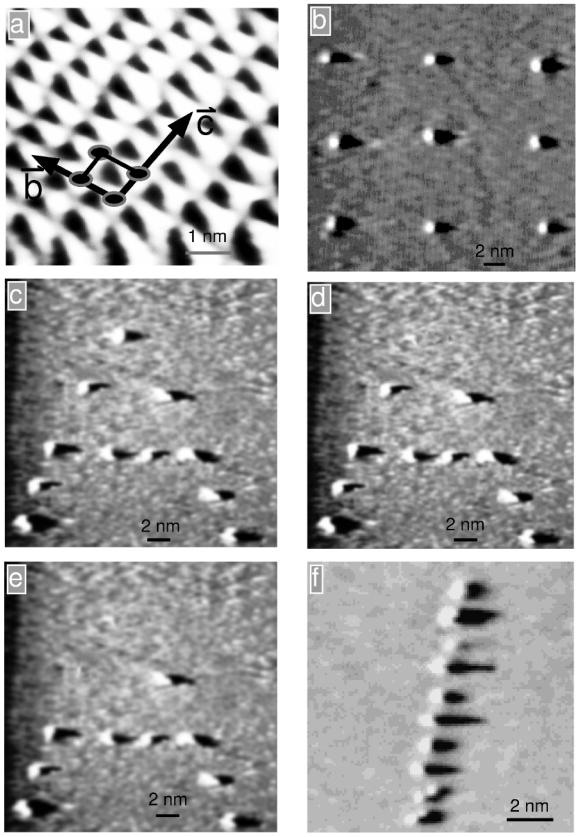


FIG. 3 STM images of the NBMN-pDA film on HOPG. (a) An image of the surface of the film showing crystalline order, image size: 6 nm × 6 nm; (b) a 3 nm × 3 nm array formed by voltage pulses of 4 V and 1 μs; (c) an "A" pattern formed by voltage pulses of 3.5 V and 2 μs; (d) and (e) STM images after erasing marks one at a time with reverse-polarity voltage pulses of −4.5 V and 50 μs; (f) resolution test using voltage pulses of 4.2 V and 10 μs. The distance between neighboring marks is 1.7 nm. Scan conditions are $V_b = 0.1$ V, $I_t = 0.4$ nA for (a); and $V_b = 0.19$ V, $I_t = 0.19$ nA for (b), (c), (d), (e), and (f); constant height mode.

Finally, local *I-V* characteristics of the film measured before and after mark formation confirm the conductance transition. As shown in curve *a* of Fig. 4, with the NBMN-pDA in its initial high-impedance state, the applied voltage can be varied from 0 to 2.1 V but the tunneling current remains small (about 0.2 nA). Beyond a threshold of 2.1 V, the film changes to a low-impedance state with the *I-V* relation of curve *b*. Curve *b* clearly demonstrates conductivity and eliminates deposited charge as a possible mechanism of the mark formation. For comparison, the *I-V* curve from the graphite substrate is shown in curve *c* of Fig. 4. The *I-V* relation for the substrate is inconsistent with that of the film in its conducting state, eliminating the possibility that the conducting state is reached by simply burning a hole through the film.

Several possible mechanisms for the conductance transition were considered resulting in the identification of a candidate mechanism that is in agreement with all available experimental data for this system. The first candidate considered was the charge transfer (ct) mechanism proposed by Ma *et al.* [8] for conductance transitions in a thin organic film of N-(3-nitrobenzylidene)-*p*-phenylenediamine (NBPDA). In the ct mechanism, hypothetically, a voltage pulse above some threshold brings about a shift in

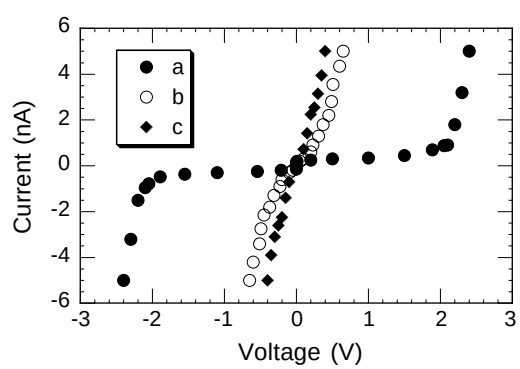


FIG. 4 Typical STM current-voltage relations. Curve *a*: Before the voltage pulse, with the film in a highly insulating state, which becomes conductive above a voltage threshold of 2.1 V; Curve *b*: *I-V* relation after the voltage pulse, indicating a transition to the conductive state; Curve *c*: linear *I-V* relation from the HOPG substrate.

electron density into a charge-separated state. While such an excited electronic state could presumably impart conductance properties to the molecule different from those of the ground state, in the present case, the ct mechanism is excluded by the transition time measurements. As shown in Fig. 2(b), the transition time from the initial high-impedance state to the conducting state is about 80 ns, many orders of magnitude longer than even the most sluggish of molecular electronic transitions. Furthermore, it is highly unlikely that an appreciable population could be sustained in the excited (ct) state for the hours to days that the film remains in the conducting state [8].

As a second possible mechanism it was considered that the individual molecular complexes of NBMN-pDA might undergo an internal rotation into an excited molecular conformation whose geometry completes the conjugated π structure *across the entire complex*. This is essentially the "conformational switching" mechanism of Joachim and Launay [9]. In a molecule containing two or more benzene rings in sequence, the electrical resistance increases by about $\times 10^4$ when the angle between the planes of adjacent rings is changed from 0 to 90° [10]. In the present case, the individual components contain extended π-conjugated structures. If these components were brought into alignment in an excited conformation, completing the π structure across the entire complex, it would result in greatly enhanced conduction in the excited conformation.

The present system proved too large to test the conformational isomerization hypothesis computationally, but we were able to carry out electronic structure calculations on a closely related system that exhibits the same electrical bistability behavior. Local conductance transitions have been demonstrated on the nanometer scale in NBPDA in the same laboratory as the present work [8]. The structure of NBPDA is shown in Fig. 5 (inset). We searched for stable conformations of this molecule at the Hartree-Fock/SCF level of theory [11]. Earlier calculations at this level resulted in excellent accuracy for both structure and intramolecular conformational energy differences in two conjugated organic systems similar to NBPDA [12]. As shown in Fig. 5, NBPDA exhibits two stable conformers. The predicted harmonic vibrational frequencies for the more stable of these two conformers reproduce the dominant peaks and features of the FTIR spectrum of crystalline NBPDA very well [13]. In both conformers, however, the two benzene rings are coplanar. Since the π-π conjugation extends over the entire NBPDA system in both conformers, we conclude that local conductance transformations observed in NBPDA do not arise from conformational isomerization, and by analogy it seems highly unlikely that such a mechanism would be at the root of the same phenomena in the very closely related NBMN-pDA system.

The most likely source of the local conductance transition in NBMN-pDA thin films is a mechanism whereby the applied electric pulse leads to a reorientation of one or more *entire molecules*, in effect introducing local disorder into the crystalline thin film. The experiments demonstrate that the crystalline films are insulating and the amorphous films are relatively conducting. This result is less surprising when one notes that this is an organic system, and carbon, while an excellent insulator when in the form of diamond, is much more conducting in various less-ordered forms. It follows naturally that if a local area of the film is structurally disordered, that region will become conducting. Since the NBMN-pDA system has a permanent dipole, an electric field pulse could lead to molecular reorientation. Such a reorientation would effectively introduce local disorder into the film. This "induced local disorder" hypothesis is also consistent with the mechanism for erasure. Only reversed polarity pulses completely reverse the conductance transition, and they must be applied for a longer duration than the "write" pulse. Obviously, if the "write" pulse grabs the permanent dipole and twists the molecule in one direction, reversing the twist will require the application of force of the opposite sign. Furthermore, the duration must be sufficient to allow time for the order to redevelop. Finally, the 80 ns characteristic transition time for the conductance transition, although much too slow for an electronic or conformational transition, is perfectly reasonable for *reorientation* of a large molecule. To test the hypothesis that the conductive region of the film is disordered, TEM and electron-diffraction studies were carried out on a film before and after switching. The results are shown in Fig. 6. As anticipated, the nonconductive film is crystalline, but after switching to the conductive state, the TEM and diffraction studies indicate an amorphous film.

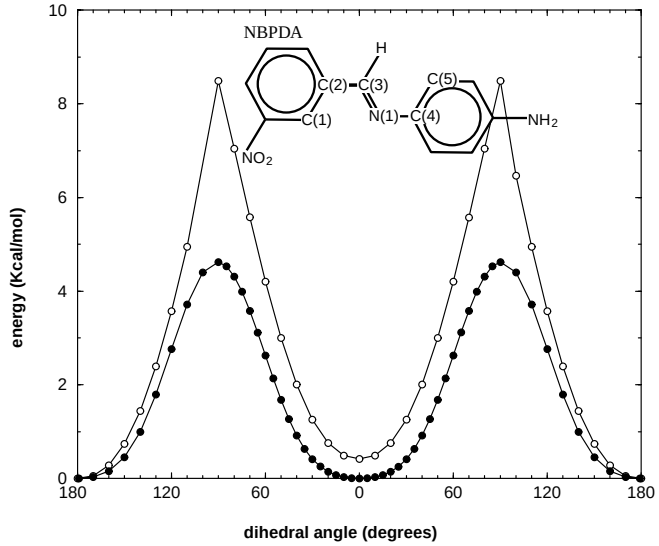


FIG. 5 The molecular structure of NBPDA (inset) and potential energy curves for internal rotations of NBPDA. Solid circles: rotation about the ($C_{(3)}$, $N_{(1)}$, $C_{(4)}$, $C_{(5)}$) dihedral angle. Open circles: rotation about the ($C_{(1)}$, $C_{(2)}$, $C_{(3)}$, $C_{(5)}$) dihedral angle.

In summary, we have demonstrated that conductance transitions can be induced by a local electric field in the

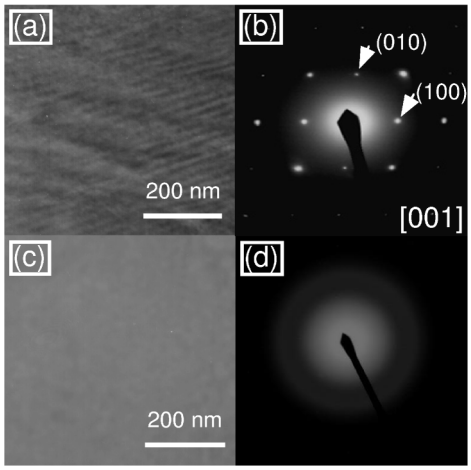


FIG. 6 Structural characterization of NBMN-pDA films. Before switching (nonconducting): (a) TEM, (b) diffraction. After switching (conducting): (c) TEM, (d) diffraction.

NBMN-pDA crystalline organic complex on the nanometer scale, and the transition can be reversed by the application of reverse pulses for a longer duration. An analysis of several possible mechanisms suggests that the conductance transition arises from the electric field pulse inducing local disorder into the crystalline thin film.

The authors thank Dr. T. Thundat for helpful discussions and W. J. Yang for experimental assistance. K. S. was supported by ORNL and ORAU, and thanks Dr. A. Kadavanich for several fruitful brainstorming sessions. This work was supported by Oak Ridge National Laboratory, managed by Lockheed Martin Energy Research Corporation under DOE Contract No. DE-AC05-96OR22464, and the National Science Foundation of People's Republic of China.

[1] T. Nakamura, T. Akutagawa, K. Honda, A. E. Underhill, A. T. Coomber, and R. H. Friend, Nature (London) **394**, 159 (1998); F. Zwick, D. Jerome, G. Margaritondo, M. Onellion, J. Voit, and M. Grioni, Phys. Rev. Lett. **81**, 2974 (1998); T. Kawamura, Y. Miyazaki, and M. Sorai, Chem. Phys. Lett. **273**, 435 (1997); H. Sirringhaus, N. Tessler, and R. H. Friend, Science **280**, 1741 (1998); A. Köhler, D. A. dos Santos, D. Beljonne, Z. Shuai, J.-L. Brédas, A. B. Holmes, A. Kraus, K. Müllen, and R. H. Friend, Nature (London) **392**, 903 (1998); J. M. Williams, A. J. Schultz, U. Geiser, K. D. Carlson, A. M. Kini, H. H. Wang, W.-K. Kwok, M.-H. Whangbo, and J. E. Schirber, Science **252**, 1501 (1991); D. Jéerome, Science **252**, 1509 (1991); G. Yu, J. Gao, J. C. Hummelen, F. Wudl, and A. J. Heeger, Science **270**, 1789 (1995); Y. Yang and A. J. Heeger, Nature (London) **372**, 344 (1994); G. M. Whitesides, J. P. Mathias, and C. T. Seto, Science **254**, 1312 (1991); M. C. Lonergan, Science **278**, 2103 (1997); M. Strukelj, F. Papadimitrakopoulos, T. M. Miller, and L. J. Rothberg, Science **267**, 1969 (1995).

[2] A. Aviram and M. A. Ratner, Chem. Phys. Lett. **29**, 277 (1974); A. Aviram and P. E. Seiden, U.S. Patent No. 3 953 874 (1976).

[3] A. Aviram, C. Joachim, and M. Pomerantz, Chem. Phys. Lett. **146**, 490 (1988).

[4] A. Aviram, C. Joachim, and M. Pomerantz, Chem. Phys. Lett. **162**, 416 (1989).

[5] R. S. Potember, T. O. Poehler, and D. O. Cowan, Appl. Phys. Lett. **34**, 405 (1979); R. S. Potember, T. O. Poehler, and R. C. Benson, Appl. Phys. Lett. **41**, 548 (1982); R. S. Potember, T. O. Poehler, A. Rappa, D. O. Cowan, and A. N. Bloch, Synth. Met. **4**, 371 (1982); R. S. Potember, T. O. Poehler, A. Rappa, D. O. Cowan, and A. N. Bloch, J. Am. Chem. Soc. **102**, 3659 (1980); R. S. Potember, T. O. Poehler, D. O. Cowan, F. L. Carter, and P. Brant, *Molecular Electronic Devices*, edited by F. L. Carter (Marcel Dekker, New York, 1983), p. 73; S. Yamaguchi and R. S. Potember, Mol. Cryst. Liq. Cryst. **267**, 241 (1995).

[6] B. B. Corson and R. W. Stoughton, J. Am. Chem. Soc. **50**, 2825 (1928).

[7] S. F. Alvarado, P. F. Seidler, D. G. Lidzey, and D. D. C. Bradley, Phys. Rev. Lett. **81**, 1082 (1998).

[8] L. P. Ma, W. J. Yang, Z. Q. Xue, and S. J. Pang, App. Phys. Lett. **73**, 850 (1998).

[9] C. Joachim and J. P. Launay, J. Mol. Electron. **6**, 37 (1990).

[10] M. P. Samanta, W. Tian, S. Datta, J. I. Henderson, and C. P. Kubiak, Phys. Rev. B. **53**, R7626 (1996).

[11] Hartree-Fock self-consistent field calculations, geometry optimizations, and harmonic vibrational analysis carried out with the code GAMESS: M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993). Original program assembled by M. Dupuis, D. Spangler, and J. J. Wendoloski. Polarized double-zeta (3-21G*) basis used for all atoms: J. S. Binkley, J. A. Pople, and W. J. Hehre, J. Am. Chem. Soc. **102**, 939 (1980); W. J. Pietro, M. M. Francl, W. J. Hehre, D. J. DeFrees, J. A. Pople, and J. S. Binkley, J. Am. Chem. Soc. **104**, 5039 (1982).

[12] K. Sohlberg, B. L. Baker, S. P. Leary, N. L. Owen, J. C. Facelli, and B. A. Trofimov, J. Mole. Struct. **354**, 55 (1995); K. Sohlberg, S. Leary, N. L. Owen, and B. Trofimov, Vib. Spectrosc. **13**, 227 (1997).

[13] As is conventional, HF-SCF frequencies were scaled $\times 0.89$: M. Diem, *Introduction to Modern Vibrational Spectroscopy* (Wiley, New York, 1993).

Damage nucleation and vacancy-induced structural transformation in Si grain boundaries

A. Maiti[a] and S. T. Pantelides
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831
and Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235

M. F. Chisholm and S. J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

(Received 13 February 1999; accepted for publication 16 August 1999)

Atomic resolution *Z*-contrast scanning transmission electron microscopy reveals preferential nucleation of electron-beam-induced damage in select atomic columns of a Si tilt grain boundary. Atomic scale simulations find that the region of initial damage nucleation corresponds to columns where the formation energies of vacancies and vacancy complexes are very low. The calculations further predict that vacancy accumulation in certain pairs of columns can induce a structural transformation to low-density dislocation ''pipes'' with all atoms fourfold coordinated. © *1999 American Institute of Physics.* [S0003-6951(99)00742-1]

Beam-induced processing is a widely studied phenomenon,[1] and it has long been known that beam-induced amorphization nucleates preferentially at extended defects such as grain boundaries.[2] The electron microscope has also been widely used to induce and study damage.[3–7] In this letter we report experimental data showing that the energetic electron beam of a transmission electron microscope induces nucleation of damage at specific dislocation cores in a Si grain boundary. We also report atomistic simulations that provide an explanation for this effect and further predict a structural transformation that may occur under suitable experimental conditions.

The experimental study was performed on a silicon bicrystal wafer containing an isolated 16° symmetric tilt boundary with the tilt axis parallel to the $\langle 001\rangle$ direction ($\Sigma=25\{710\}\langle 001\rangle$). Plan view samples with a $\langle 001\rangle$ surface normal were prepared by mechanical polishing and ion milling with 1–3 keV Ar ions, which was followed by characterization with high resolution *Z*-contrast imaging using a HB603U scanning transmission electron microscope operating at 300 kV.

A symmetric $\Sigma=25$ boundary with a common $\{710\}$ plane can be constructed by rotating two crystalline regions about the $\langle 001\rangle$ axis through equal and opposite angles of 8.13°. The boundary structure is periodically repeated, with a period of one conventional lattice parameter ($a=5.43$ Å) parallel to the tilt axis, and a period of 38.4 Å ($a\sqrt{50}$) in a direction normal to the tilt axis. In each repeat period normal to the tilt axis the grain boundary consists of two extended cores, one being a mirror image of the other, separated by perfect crystalline regions. Figures 1(a) and 1(b) display, in the projection normal to the tilt axis, electron micrographs of one of these extended cores at two stages of exposure to electron irradiation: (a) a nearly unaffected core with appreciable intensity from all the columns; and (b) a partially affected core with a few of the columns having reduced brightness. Figure 2(a) displays in the same projection a three dimensional computer model of the boundary structure.

A closer look at the structure of Fig. 2(a) reveals that in the undamaged grain-boundary core there are four dislocation cores,[8] each dislocation core appearing as a ''pentagon-triangle'' combination in the projection normal to the tilt axis. All these cores have Burgers vectors in a plane normal to the tilt axis and are therefore designated perfect edge dislocations. Two of the cores, $D1$ and $D2$ are adjacent to each other and have opposite Burgers vectors, thereby constituting a dislocation dipole. The other two edge dislocation cores $E1$ and $E2$ have Burgers vectors $a/2\langle 110\rangle$ and $a/2\langle 1\bar{1}0\rangle$, respectively, where the x direction is taken normal to the boundary plane and the y direction taken in the grain boundary plane normal to the tilt axis. The Burgers vectors of $E1$ and $E2$ add up to a, which is the total dislocation content of the boundary core. We note that all the columns with reduced brightness in the damaged core are located on and around the edge dislocations $E1$ and $E2$. In the following, we discuss results of an extensive computational analysis to understand: (1) why a small region localized around $E1$ and

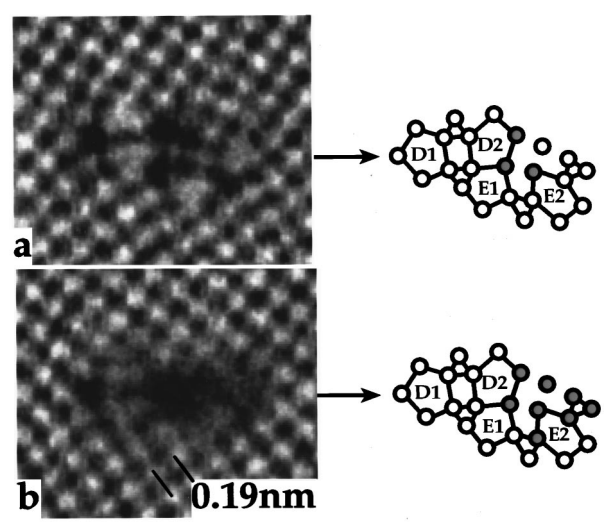


FIG. 1. *Z*-contrast images and derived structures of the $\Sigma=25$ $\{710\}$ $\langle 001\rangle$ symmetric tilt boundary at two stages of exposure to electron irradiation: (a) a nearly unaffected core with all columns visible but those shaded showing reduced intensity; (b) a partially affected core with several columns appearing darker.

[a]Present address: Molecular Simulations, Inc., Burlington, MA 01803.

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

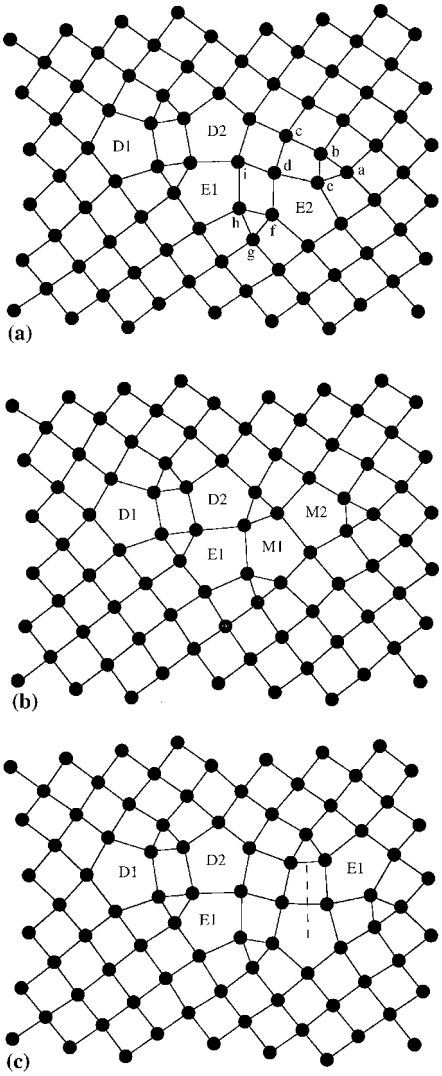


FIG. 2. (a) Three-dimensional atomic model of the boundary structure identifying the individual dislocations cores and labeling various sites. (b) The atomic structure of the transformed grain boundary resulting from the segregation of the [*b*,*c*] divacancy chain. The edge dislocation core $E2$ of the original grain boundary dissociates into two mixed dislocation cores $M1$ and $M2$. (c) The transformed structure resulting from the segregation of the [*a*,*b*] divacancy chain creating an extended edge dislocation core.

$E2$ is preferentially affected by electron irradiation, leaving the rest of the extended core unchanged; (2) the atomic scale processes that underlie the observed phenomenon; (3) any possible low-energy defect configuration that might lead to a structural transformation of the boundary.

In a relativistic elastic collision between an electron and an atomic nucleus initially at rest, the energy required to knock an atom out of its position and create structural damage, T_m must be greater than the ''displacement energy'' E_d. The displacement energy is somewhat larger than the formation energy of a Frenkel pair because of kinetic barriers and has been determined experimentally to be 15 ± 5 eV for a bulk Si crystal.[9,10] It is expected to be smaller at a grain boundary where the formation energy of a Frenkel pair is generally smaller (by 2–4 eV; see below). The fraction of atoms in a column initially knocked out by the electron beam (i.e., the fraction of Frenkel pair defects) is given by

$$f_F = n_{e1}\sigma_d. \tag{1}$$

Here, n_{e1} is the number of electrons passing through per unit area, estimated to be $\sim 10^5$ Å^2 from the scan time and the electron current in the microscope probe; σ_d is the total Mott scattering cross-section of the electron[11] which decreases rapidly with increasing E_d. Since Si is a light element, the Mott cross-section can be accurately computed using the McKinley–Feshbach formula.[12] Using appropriate values of E_d, we find that the electron beam creates 2%–5% vacancies in bulk atomic columns and as many as 20% in grain boundary columns. Point defects in Si are, however, known to be very mobile under electron irradiation conditions.[13] Therefore, these initial defects would either recombine, escape to the surface, or segregate at low-energy sites in the grain boundary. This analysis suggests that the low-intensity columns in Fig. 1(b) correspond to sites with large segregation energies of certain point defects.

To address the nature of these defects, we have computed the formation energies of both vacancies and interstitials in the bulk crystal and at various sites within the grain boundary core. The segregation energy is the difference between formation energies in the grain boundary and the bulk crystal. To describe the interatomic interaction for Si we used the many-body potential due to Tersoff,[14] which reproduces well the local density approximation (LDA) results for vacancy and interstitial formation energies in the bulk crystal (see Table III of Ref. 14 for more details). We tested the potential further against LDA results for vacancies at various sites in the $\Sigma=5\{310\}\langle 001\rangle$ tilt boundary, which has a smaller repeat cell than the $\Sigma=25$ boundary. Agreement was satisfactory for both formation energies and atomic relaxations.

For the $\Sigma=25$ boundary, we computed the formation energy of isolated vacancies, isolated divacancies, chains of monovacancies (i.e., removal of an entire column of atoms) and pairs, triplets, etc., of such chains. In each case, the relevant atom(s) were removed and the surrounding lattice fully relaxed using a combined scheme of simulated annealing and conjugate gradients. Similar calculations were done for self-interstitials.

Table I lists the formation energies (normalized per vacancy) of the most stable vacancy structures in various configurations. Figure 2(a) displays the site designations. The vacancy formation energies at other sites are at least 2.5 eV or higher. It is clear from the table that all the low-energy structures occur in a localized region of the core in and around the dark region in the micrograph of Fig. 1(b). From a similar computational study on the interstitials, we find that the lowest formation energies are around 2.5 eV, and the low-energy sites are scattered all over the dislocation core, showing no particular affinity to the low-intensity region of Fig. 1(b). These results lead to the conclusion that interstitials are not trapped in specific regions and either recombine with vacancies or move to the surface. We conclude that vacancies segregated in various low-energy configurations are primarily responsible for the low-intensity columns observed in Fig. 1(b).

A closer look at Table I shows that the chain configurations have lower formation energy than the isolated vacancies, which indicates a significant attraction between neighboring vacancies. However, of the many different chain configurations explored, two divacancy chain structures, i.e.,

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

TABLE I. The formation energies of the most stable vacancies in various configurations. Configurations are isolated unless indicated explicitly as a chain configuration. The letters indicate sites at which the vacancies are initially placed, as designated in Fig. 2(a).

Vacancy configurations	Formation energy (eV/vacancy)
bulk	3.7
[*b*]	1.7
[*d*]	1.7
[*i*]	1.8
[*a*,*b*]	1.0
[*b*,*c*]	1.6
[*b*] chain	0.8
[*d*] chain	1.4
[*d*,*i*] chain	1.3
[*a*,*b*] chain	0.3[a]
[*b*,*c*] chain	0.2[a]
[*a*,*b*,*e*] chain	1.2
[*f*,*g*,*h*] chain	1.4
[*b*,*c*,*d*,*e*] chain	1.3
[*d*,*f*,*h*,*i*] chain	1.3

[a]Structural transformation.

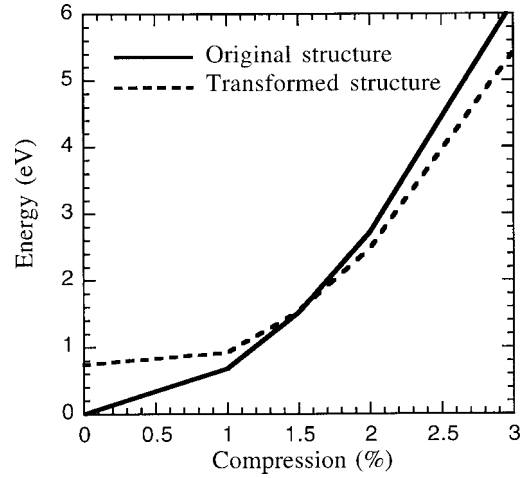


FIG. 3. A plot of the change in energy as a function of uniaxial compression perpendicular to the boundary plane of the original boundary of Fig. 2(a) and the transformed grain boundary of Fig. 2(b). Values are normalized per periodic segment of the grain boundary and are relative to the energy of the original boundary structure. A stability crossover occurs at a compression of ~1.5%.

the [*a*,*b*] chain and the [*b*,*c*] chain have extremely low formation energies, an order of magnitude smaller than the vacancy formation energy in the bulk.[15] Figures 2(b) and 2(c) display the structures following the relaxation of the [*b*,*c*] and the [*a*,*b*] divacancy chains, respectively. Both these structures have a noteworthy feature: all atoms are fourfold coordinated, i.e., the segregation of a chain of divacancies has led to a new structure without any dangling bonds!

A comparison of the dislocation cores of Figs. 2(a) and 2(b) indicates that the pure edge dislocation core $E2$ with Burger's vector $\mathbf{b}=a/2\langle 1\bar{1}0\rangle$ has dissociated into two mixed dislocations $M1(\mathbf{b_1}=a/2\langle 10\bar{1}\rangle)$ and $M2(\mathbf{b_2}=a/2\langle 0\bar{1}1\rangle)$, while all the other cores remain unchanged. The atomic arrangement shown in Fig. 2(c) is more complicated—here $E2$ splits up into an edge dislocation with Burger's vector $\mathbf{b}=a/2\langle \bar{1}10\rangle$ and an ''extended'' edge dislocation with $\mathbf{b}=a\langle 010\rangle$ indicated by the dotted line. These reconstructed dislocation cores have lower atomic density. In addition, we found that the formation energy of interstitials in these cores is reduced by almost an electron volt compared to that in the untransformed grain boundary, making them effective diffusion pipes. Another interesting feature is the presence of a chain of ''long'' bonds ~2.6 Å, which provide ideal sites for the cooperative segregation of certain dopants in the form of substitutional dimers.[16]

All the foregoing calculations were carried out under zero external stress. Figure 3 plots the change in energy as a function of uniaxial compression normal to the grain boundary plane for the original grain boundary [Fig. 2(a)] and the transformed grain boundary of Fig. 2(b). There is a stability crossover at a compression of ~1.5%. A large local compressive stress can thus drive the formation energy of the [*b*,*c*] divacancy chains negative and bring about a spontaneous structural transformation of the boundary. Such compression is commonly achieved in strained layer epitaxy.

As is evident from the micrographs of Fig. 1, electron irradiation can induce vacancy segregation in the select regions of the grain boundary core, but does not lead to the predicted reconstruction. The key difference is that the simulations involved annealing at high temperatures whereas the experiment was done at room temperature. Further experiments with different conditions would be needed to explore the relevant phase space, e.g., irradiation at elevated temperatures and/or pressures, which is not currently feasible on our instrument. Finally, external stress might be used to induce a spontaneous transformation.

We would like to thank Mark Robinson for useful discussions. This research was supported in part by Lockheed Martin Energy Research Corp. under DOE Contract No. DE-AC05-96OR22464, and ONR Grant No. N00014-96-1-1286, and by an appointment to the ORNL Postdoctoral Research Associates Program administered jointly by ORNL and ORISE.

[1] See, e.g., *Ion Implantation and Beam Processing*, edited by J. S. Williams and J. M. Poate (Academic, New York, 1984).
[2] G. J. Clark, A. D. Marwick, F. LeGoues, R. B. Laibowitz, R. Koch, and P. Madakson, Nucl. Instrum. Methods Phys. Res. B **32**, 405 (1988).
[3] L. W. Hobbs, Ultramicroscopy **3**, 381 (1978); A. Howie, Rev. Phys. Appl. **15**, 291 (1980).
[4] D. A. Muller and J. Silcox, Philos. Mag. A **71**, 1375 (1995).
[5] L. M. Wang, Nucl. Instrum. Methods Phys. Res. B **141**, 312 (1998).
[6] M. Kiritani, Ultramicroscopy **39**, 135 (1991).
[7] L. Reimer, Ultramicroscopy **14**, 291 (1984).
[8] We refer to these ''structural units'' as dislocations because they retain the characteristic ''edge,'' ''screw,'' or ''mixed'' character that isolated dislocations in bulk crystals exhibit. They are also characterized by Burger's vectors. This practice goes back to the classic paper by Hornstra [J. Hornstra, Physica **26**, 198 (1959)].
[9] L. Reimer, *Transmission Electron Microscopy* (Springer, New York, 1993), Chap. 5.
[10] J. W. Corbett and J. C. Bourgoin, in *Point Defects in Solids*, edited by J. H. Crawford and L. M. Slifkin (Plenum, New York, 1975), Vol. 2, Chap. 1, Sec. 7.
[11] N. F. Mott, Proc. R. Soc. London, Ser. A **135**, 429 (1932).
[12] W. A. McKinley and H. Feshbach, Phys. Rev. **74**, 1759 (1948).
[13] G. D. Watkins, in *Radiation Damage in Semiconductors* (Dunod, Paris, 1964), p. 97.
[14] J. Tersoff, Phys. Rev. B **38**, 9902 (1988).
[15] Chains of interstitials were also found to be more stable than isolated interstitials. However, the formation energy of the most stable chain was found to be around 0.9 eV per interstitial, significantly higher than the formation energy of the most stable divacancy chains.
[16] A. Maiti, M. F. Chisholm, S. J. Pennycook, and S. T. Pantelides, Phys. Rev. Lett. **77**, 1306 (1996).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

ELSEVIER

Physica B 273–274 (1999) 453–457

PHYSICA B

www.elsevier.com/locate/physb

A combined experimental and theoretical approach to grain boundary structure and segregation

S.J. Pennycook[a,b,*], M.F. Chisholm[a], Y. Yan[c], G. Duscher[a,b], S.T. Pantelides[a,b]

[a]*Solid State Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831-6030, USA*
[b]*Department of Physics and Astronomy, Vanderbilt University, Nashville TN 37235, USA*
[c]*National Renewable Energy Laboratory, Golden, CO 80401, USA*

Abstract

Recently, the scanning transmission electron microscope has become capable of forming electron probes of atomic dimensions. Through the technique of Z-contrast imaging, it is now possible to form atomic resolution images with high compositional sensitivity from which atomic column positions can be directly determined. An incoherent image of this nature also allows atomic resolution chemical analysis to be performed, by locating the probe over particular columns or planes seen in the image while electron energy loss spectra are collected. Such data represent either an ideal starting point for first-principles theoretical calculations or a test of theoretical predictions. We present several examples where theory and experiment together give a very complete and often surprising atomic scale view of complex materials. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Grain boundaries; Atomic structure; Impurities; Segregation; Transformation; Electron microscopy; Total energy calculations

1. Direct imaging with electrons

The images we see with our eyes generally appear to be a direct representation of the world around us. This is because they are formed *incoherently*; objects are illuminated by light over a large range of directions and the intensity scattered depends primarily on the nature of the object and not on the direction of illumination. If objects are illuminated by coherent laser light they show a speckle pattern, which is due to interference effects. These are directly related to the object, but not in a manner that can be directly interpreted. Except for special applications, optical instruments such as a projector or a microscope invariably use incoherent illumination to maintain this capability for direct interpretation.

It is over one hundred years since Lord Rayleigh first explained the difference between coherent and incoherent imaging in the context of the light microscope. He clearly pointed out the advantages of incoherent imaging: absence of interference artifacts, and in addition a factor of two improved resolution [1]. Ideally, a self-luminous object is required for perfect incoherent imaging, so that each point will emit light independently. But Lord Rayleigh showed how the condenser lens can be used to give a close approximation to incoherent imaging. If the condenser lens subtends a sufficiently large range of angles, then we approach the incoherent imaging conditions of normal vision. Specifically, for an illumination semiangle θ, the transverse coherence length in the sample, ℓ_T, is $0.61\lambda/\theta$, which can be made comparable to the resolution limit.

In the conventional transmission electron microscope (CTEM), historically, a small condenser aperture has been used to give an approximation to perfect coherent imaging. This is necessary for low-resolution diffraction contrast imaging of defects such as dislocations, but this

* Correspondence address. Solid State Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6030, USA. Tel.: + 1-423-574-5504; fax: + 1-423-574-4143.
E-mail address: pky@ornl.gov (S.J. Pennycook)

0921-4526/99/$ - see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0921-4526(99)00521-9

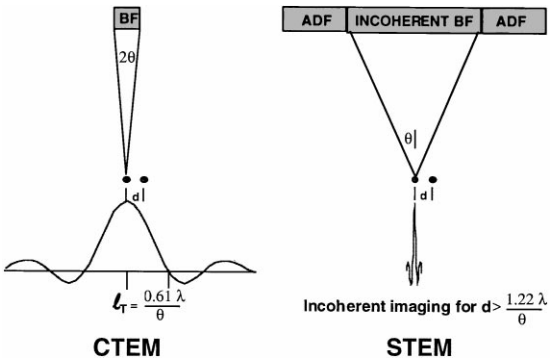


Fig. 1. Schematic comparing the transverse coherence lengths for coherent and incoherent imaging. In bright field (BF) conventional TEM the coherence length is large compared to interatomic spacings, whereas in the STEM, a large-angle BF or annular dark field (ADF) detector effectively breaks the coherence between neighboring atoms.

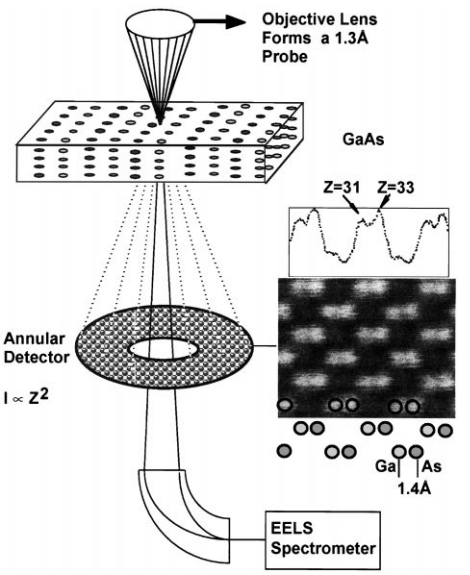


Fig. 2. Schematic showing a fine probe formed by the objective lens of an STEM. A Z-contrast image is formed by mapping high-angle scattered electrons as the probe is scanned across the sample, while atomic resolution electron energy loss spectroscopy is possible with the probe stationary on specific columns selected from the image. The Z-contrast image of GaAs shows As columns with higher intensity than Ga.

leads to a transverse coherence length much greater than typical interatomic spacings as shown in Fig. 1. As electron microscope resolution gradually improved it became possible to resolve atomic-scale features, but the coherent imaging conditions were maintained. In a coherent image (referred to as a phase contrast image) atomic columns can be bright or dark depending on specimen thickness and objective lens focus (which alter the relative phases of the scattered beams). It is clear that simply increasing the illumination aperture will reduce the transverse coherence length. Eventually, it will reduce below the typical interatomic spacings of materials, and the coherence between columns will be broken. We would then achieve incoherent imaging, and we would expect atomic images to become directly interpretable. This has indeed proved to be the case. In practice, however, it is more efficient and results in higher image contrast to reverse the direction of the electrons, and to use the complementary high-angle annular *detector* shown in Fig. 1. Then the image is formed by scanning the illumination across the sample, a scanning transmission electron microscope (STEM).

The principle of Z-contrast imaging in the STEM is shown schematically in Fig. 2. A small electron probe is scanned across a thin specimen, and the Z-contrast image results from mapping the intensity of electrons reaching the annular detector [2–7]. The detector performs the function of Lord Rayleigh's condenser lens. But it not only breaks the coherence in the *transverse* plane [8], it enforces high scattering angles, so that Rutherford scattering dominates and atoms contribute to the image with a brightness determined by their mean square atomic number Z. For this reason the image is referred to as a Z-contrast image. In a crystal, Rutherford scattering becomes phonon scattering, and effectively breaks the coherence through the *thickness* of the sample [9]. The thickness dependence of the image becomes relatively intuitive, and atomic images from thicker materials can be interpreted equally well on the basis of Z-contrast.

In the image of GaAs shown in Fig. 2, bright features correspond directly to columns of As, and the less bright features to columns of Ga. Unlike the coherent imaging of conventional high-resolution electron microscopy (HREM), the positions of atomic columns can be determined directly and uniquely from the image to a high accuracy, without the need for extensive image simulations of model structures. Incoherent imaging effectively bypasses the phase problem of HREM, a particular advantage for complex materials. The VG Microscopes HB603U STEM at ORNL has a 300 kV accelerating voltage, and a directly interpretable resolution of 1.26 Å, although recently information transfer at 0.78 Å has been demonstrated in Si⟨1 1 0⟩ [10]. This also demonstrates the factor of two improved resolution available with incoherent imaging; the comparable phase contrast image resolution on this microscope is 1.93 Å. There are now many examples where the high resolution and the direct interpretability of Z-contrast imaging have proved very successful. Examples include the direct determination of dislocation core structures in GaN [11] and at CdTe/GaAs interfaces [12], and in imaging structure and

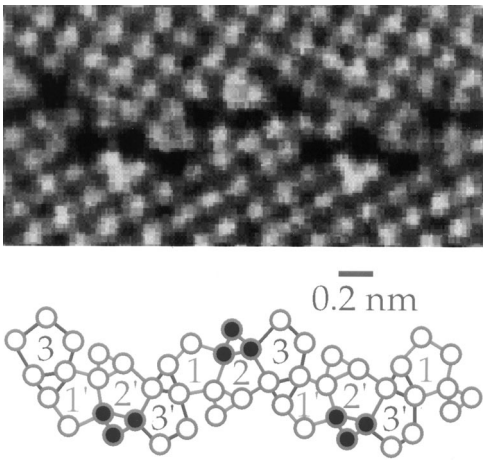


Fig. 3. Z-contrast image of a 23° ⟨0 0 1⟩ tilt grain boundary in Si showing its unexpectedly complex structure. The five-fold rings (with black centers in the image) are dislocation cores arranged in a repeating sequence along the boundary. Columns shown black in the schematic are those seen brighter in the image due to segregated As.

impurity sites at grain boundaries, as shown in the examples below.

The ability to retrieve atomic structures directly from experiment is a great advantage for first principles simulations, as it avoids the need to calculate large numbers of trial structures. This is especially true for complex materials where there are a great many possibilities. A further advantage of the STEM is that it allows electron energy loss spectroscopy (EELS) to be performed simultaneously with the Z-contrast image, allowing compositional analysis and local band structure to be determined at atomic resolution [13,14].

2. Arsenic segregation sites at a silicon grain boundary

Z-contrast imaging enables low concentrations of high-Z impurities to be directly observed. A recent example of this capability is shown in Fig. 3, a Z-contrast image from a 23° grain boundary in Si, after doping with As [15].

The atomic structure of the boundary is directly determined from the positions of the bright features in the image, and is different from all structures proposed previously. It comprises a continuous sequence of dislocation cores, a perfect edge dislocation (1) and two perfect mixed dislocations (2,3) arranged as a dipole, followed by the same sequence (1′,2′,3′) mirrored across the boundary plane. In the ⟨0 0 1⟩ projection, these dislocations appear as a connected array of pentagonal and triangular arrangements of atomic columns. The presence of the dipoles is surprising, as being of equal and opposite Burgers vector they could equally well be replaced by perfect crystal. However, precisely the same atomic arrangement is seen in the undoped boundary given the same annealing treatment.

Looking closely at the relative intensities of the columns in the doped sample, it is seen that one of the dislocation cores contains columns that are 20% brighter on average than other similar columns. This must be due to the presence of the As dopant. Taking into account the scattering cross section, the increased intensity corresponds to an average of only 5% As concentration, approximately two As atoms in each atomic column.

In a previous theoretical study of the shorter-period 36° grain boundary, it was found that isolated As atoms have only a small segregation energy of ~ 0.1 eV, too small to account for the concentrations observed experimentally [16]. However, as arsenic prefers to be three-fold coordinated, calculations were performed for arsenic dimers. It was found that the two As atoms repel and become three-fold coordinated without having to create an Si dangling bond. Thus binding of the dimer occurs through repulsion. After the image of Fig. 3 was obtained, calculations were repeated for the 23° boundary, and preference was found for those sites seen bright in the image. The segregation energy was again increased, becoming consistent with the As solubility limit in the bulk at the annealing temperature of 700°C. This combined use of experimental and theoretical techniques produced a remarkably detailed and consistent atomic-scale picture of impurity segregation at this grain boundary.

Very recently, an extensive ab initio study has confirmed the observed grain boundary structure, with its redundant dislocations, to be energetically preferred in the undoped grain boundary [17]. Further theoretical work could build on these results to determine grain boundary diffusion coefficients, as well as to extend studies to other boundaries and polycrystalline materials.

3. Impurity-induced grain boundary transformation in MgO

In Fig. 4, STEM imaging of an MgO grain boundary [18] reveals a structure that is inconsistent with the widely accepted structure of the boundary proposed by Harris et al. [19], based on theoretical modeling using classical potentials. The observed structure is similar to that proposed much earlier by Kingery [20]. The large empty core of the Harris structure is obviously very different from the more dense core of the Kingery model. On careful examination of the intensity in the experimental image, it can be seen that certain specific atomic columns at the grain boundary are again significantly brighter than neighboring columns, as arrowed in the figure. This suggests that impurities, with $Z > 12$, may be segregated at these sites. EELS measurements indeed

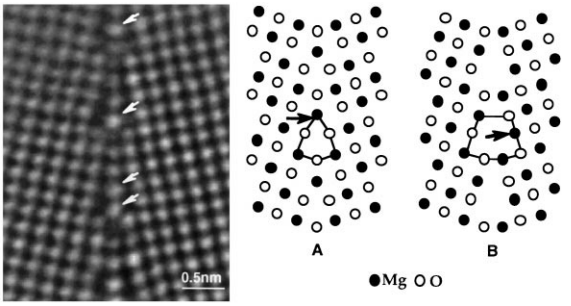


Fig. 4. Z-contrast image from a 24° ⟨0 0 1⟩ tilt grain boundary in MgO showing occasional bright atomic columns at the grain boundary (arrowed), compared to two structures for the 36° ⟨0 0 1⟩ tilt grain boundary proposed by (a) Kingery [20] and (b) Harris et al. [19]. Sites of Ca segregation are arrowed.

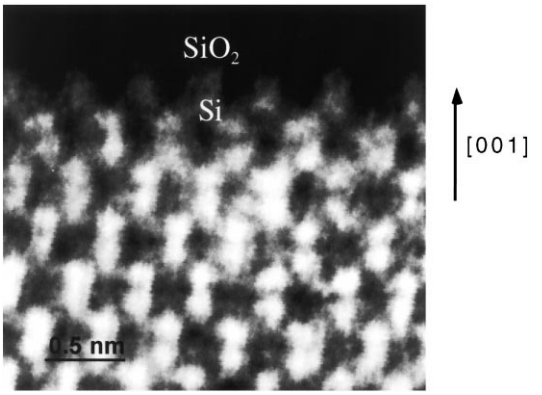


Fig. 5. Z-contrast image of a sharp Si/SiO_2 interface showing the last monolayer of Si in atomic positions close to those of the bulk.

established that significant concentrations of Ca were present in the grain boundary, approximately 0.3 monolayers, consistent with the bright intensity in the image.

To reconcile these observations with the apparently conflicting prior experimental and theoretical work, first-principles theoretical calculations were performed. These calculations in fact reproduced the results of the classical potential calculations for the clean grain boundary, indicating the open structure to be 0.5 eV lower in energy per periodic repeat unit. Theory further determined that Ca has a large segregation energy in both boundary structures, but *significantly higher* in the dense structure, sufficient to make the dense structure the lower energy boundary. These calculations therefore established that the dense structure is in fact *stabilized* by the Ca segregation, an example of a segregation-induced structural transformation [18]. Examination of the electronic charge distribution revealed just a small perturbation to the oxygen ions next to the Ca atom, indicating that the transformation is structural and not electronic in origin, i.e. it is driven by the size difference between Ca and Mg ions.

4. Structure and composition of the Si–SiO_2 interface

In the case of an amorphous material, no channeling of the probe can occur, and a Z-contrast image reflects only changes in projected atomic density. This is in marked contrast to conventional HREM images where amorphous materials always show a speckle pattern due to random interferences, and interfaces show Fresnel fringes and other coherent interference phenomena that can obscure the structure. In the Z-contrast image of Fig. 5, the last monolayer of the crystalline Si is clearly visible, and the Si columns are in almost exactly the expected positions for bulk Si [21]. Here again theory can be very useful in establishing the relative energies of various possible interface structures. Recent calculations have established that atomically abrupt interfaces are in fact energetically preferred over structures involving suboxides [22]. This is due to the softness of the Si–O–Si bond which allows strain relaxation to occur. Experimentally, however, suboxides are always found, presumably due to the non-equilibrium nature of the oxidation process and indicating the potential for improved characteristics. EELS provides a sensitive means to detect suboxides at the Si/SiO_2 interface, with high sensitivity and spatial resolution [23]. An example of an extended zone of suboxide is seen in Fig. 6, where Si-L edge spectra are plotted for a series of beam positions moving from the crystal into the oxide. The edge onset is 99 eV in the crystal, moving to 104 eV for stoichiometric SiO_2, but for approximately 2 nm both features are seen indicative of suboxide. The width of the suboxide zone is an order of magnitude greater than the beam size; this is confirmed also by profiles of the O-K edge [21], where there can be no contribution from the crystalline Si.

5. Conclusions

It is now possible to determine, without prior knowledge, the structure, impurity content and local electronic states, at grain boundaries, dislocations and interfaces by the combination of Z-contrast imaging and electron energy loss spectroscopy in the STEM. Atomic-scale total energy calculations are a natural complement and extension to these STEM techniques, giving critical insight into the underlying physics through the ability to study segregation energies, electronic states, and atomic-scale dynamics.

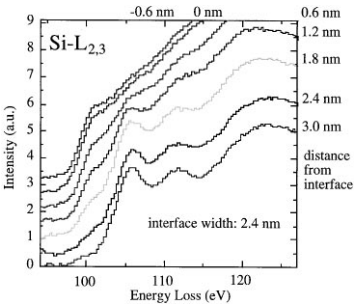


Fig. 6. Si-$L_{2,3}$ EELS profiles across a Si/SiO_2 interface showing evolution of the SiO_2 band gap. The full gap is not established until 2.4 nm into the oxide.

Acknowledgements

This research was supported by Lockheed Martin Energy Research Corp. under DOE Contract No. DE-AC05-96OR22464, ONR Grant No. N00014-95-1-0906, AFOSR Grant No. F-49620-99-1-0289 and by appointment to the ORNL Postdoctoral Research Associates Program administered jointly by ORNL and ORISE.

References

[1] L. Rayleigh, Philos. Mag. 42 (5) (1896) 167.
[2] S.J. Pennycook, in: S. Amelinckx, G. Van Tendeloo, D. Van Dyck, J. Van Landuyt (Eds.), Handbook of Microscopy, VCH Publishers, Weinheim, Germany, 1997, pp. 595–620.
[3] S.J. Pennycook, D.E. Jesson, Phys. Rev. Lett. 64 (1990) 938.
[4] S.J. Pennycook, D.E. Jesson, Ultramicroscopy 37 (1991) 14.
[5] S.J. Pennycook, D.E. Jesson, Acta. Metall. Mater. 40 (1992) S149.
[6] R.F. Loane, P. Xu, J. Silcox, Ultramicroscopy 40 (1992) 121.
[7] P.D. Nellist, S.J. Pennycook, Ultramicroscopy 78 (1999) 111.
[8] D.E. Jesson, S.J. Pennycook, Proc. Roy. Soc. A 441 (1993) 261.
[9] D.E. Jesson, S.J. Pennycook, Proc. Roy. Soc. A 449 (1995) 273.
[10] P.D. Nellist, S.J. Pennycook, Phys. Rev. Lett. 81 (1998) 4156.
[11] Y. Xin, S.J. Pennycook, N.D. Browning, P.D. Nellist, S. Sivananthan, F. Omnès, B. Beaumont, J.-P. Faurie, P. Gibart, Appl. Phys. Lett. 72 (1998) 2680.
[12] A.J. McGibbon, S.J. Pennycook, J.E. Angelo, Science 269 (1995) 519.
[13] N.D. Browning, M.F. Chisholm, S.J. Pennycook, Nature 366 (1993) 143.
[14] G. Duscher, N.D. Browning, S.J. Pennycook, Phys. Stat. Sol. A 166 (1998) 327.
[15] M.F. Chisholm, A. Maiti, S.J. Pennycook, S.T. Pantelides, Phys. Rev. Lett. 81 (1998) 132.
[16] A. Maiti, M.F. Chisholm, S.J. Pennycook, S.T. Pantelides, Phys. Rev. Lett. 77 (1996) 1306.
[17] J.R. Morris, Z.-Y. Lu, D.M. Ring, J.B. Xiang, K.M. Ho, C.Z. Wang, C.L. Fu, Phys. Rev. B 77 (1998) 11241.
[18] Y. Yan, M.F. Chisholm, G. Duscher, A. Maiti, S.J. Pennycook, S.T. Pantelides, Phys. Rev. Lett. 81 (1998) 3675.
[19] D.M. Harris, G.W. Watson, S.C. Parker, Philos. Mag. A 74 (1996) 407.
[20] W.D. Kingery, J. Am. Cer. Soc. 57 (1974) 1.
[21] G. Duscher, S.J. Pennycook, N.D. Browning, R. Rupangudi, C. Takoudis, H.-J. Gao, R. Singh, in: D.G. Seiler, A.C. Diebold, W.M. Bullis, T.J. Shaffner, R. McDonald, E.J. Walters (Eds.), Characterization and Metrology for ULSI Technology: 1998 International Conference, The American Institute of Physics, Woodbury, New York, 1998, pp. 191–195.
[22] R. Buczko, S. J. Pennycook, S. T. Pantelides, submitted for publication.
[23] D.A. Muller, T. Sorsch, S. Moccio, F.H. Baumann, K. Evans-Lutterodt, G. Timp, Nature 399 (1999) 758.

Appl. Phys. A 70, 153–160 (2000) / Digital Object Identifier (DOI) 10.1007/s003390000257

Applied Physics A
Materials
Science & Processing

Dynamics of single-wall carbon nanotube synthesis by laser vaporization

A.A. Puretzky, D.B. Geohegan, X. Fan, S.J. Pennycook

Solid State Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6056, USA
(Fax: +1-423/576-3676)

Received: 20 September 1999/Accepted: 20 September 1999/Published online: 21 January 2000 – © Springer-Verlag 2000

Abstract. The key spatial and temporal scales for single-wall carbon nanotube (SWNT) synthesis by laser vaporization at high temperatures are investigated with laser-induced luminescence imaging and spectroscopy. Graphite/(Ni, Co) targets are ablated under typical synthesis conditions with a Nd:YAG laser at 1000 °C in a 2-in. quartz tube reactor in flowing 500-Torr Ar. The plume of ejected material is followed for several seconds after ablation using combined imaging and spectroscopy of Co atoms, C_2 and C_3 molecules, and clusters. The ablation plume expands in stages during the first 200 μs after ablation and displays a self-focusing behavior. Interaction of the plume with the background gas forms a vortex ring which segregates and confines the vaporized material within a $\sim$ 1-cm^3 volume for several seconds. Using time-resolved spectroscopy and spectroscopic imaging, the time for conversion of atomic and molecular species to clusters was measured for both carbon (200 μs) and cobalt (2 ms) at 1000 °C. This rapid conversion of carbon to nanoparticles, combined with transmission electron microscopy analysis of the collected deposits, indicate that nanotube growth occurs over several seconds in a plume of mixed nanoparticles. By adjusting the time spent by the plume within the high-temperature zone using these in situ diagnostics, single-walled nanotubes of controlled ($\sim$ 100 nm) length were grown and the first estimate of a growth rate on single laser shots (0.2 μm/s) was obtained.

PACS: 42.62.Fi; 52.70.Kz; 81.05.Tp

Single-wall carbon nanotubes (SWNTs) exhibit remarkable electronic and structural properties which promise to revolutionize application areas from nanoscale electronics to ultralightweight structural materials [1]. In the six years since SWNTs were discovered [2, 3], research on their electronic properties and chemical functionalization has provided strong evidence that carbon nanotubes (in striking contrast to earlier fullerenes) will enable major advances in nanoscale electronic devices [4], micro electro mechanical systems (MEMS) [5, 6], biological probes [7], and field emission devices [8, 9]. Possibly of greatest significance, nanotubes may serve as molecular wires for input/output paths and interconnects between molecular electronic components or devices. Recent experiments also show that SWNTs are highly thermally conductive [1], resistant to high temperatures and harsh chemicals [10], lightweight [11], incredibly strong in the axial direction (1-TPa Young's modulus) and resilient in the transverse [6], and their hollow structure makes them ideal for hydrogen storage [11] and drug delivery. These properties broaden the scope of nanotube applications into aerospace, transportation, and biotechnology.

However nanotube growth is not controlled or understood. In situ diagnostics have not been applied during growth, and therefore no mechanisms have been determined to control the atomic-scale structure (chirality) of nanotubes during synthesis. Since the electronic properties of chemically unmodified SWNTs are entirely determined by their chirality, electronic transport in as-grown nanotubes is currently uncontrolled. Furthermore, approaches for large-scale production of SWNTs for structural applications await measurements to determine their growth rate using current synthesis techniques, and methods to increase it toward the theoretical maximum.

Up until now, three principal methods have been used to synthesize carbon nanotubes. These include laser vaporization (LV) [12], dc-arc vaporization (AV) [2, 3, 13], and chemical vapor deposition (CVD) [14, 15]. Unlike CVD, both LV and AV now produce nearly exclusively SWNT. However, very little is known about where and when the SWNTs are formed in LV or AV growth chambers, the identity of the precursor species, and the growth rates.

Laser vaporization is arguably the best method to grow high-quality, high-purity SWNTs. As typically employed, a small amount of material ($\sim 10^{16}$ carbon atoms and $\sim 10^{14}$ metal catalyst atoms) is laser-vaporized inside a hot oven with gently flowing inert gas. On a single laser shot [16], this material self-assembles to form an astounding fraction of SWNTs (70–90 vol. % yield) under the correct conditions [17, 18]. Despite the wide range of metal catalyst nanoparticle diameters produced during pulsed-LV, only SWNTs are formed. Although SWNTs grow easily by nanosecond time-scale laser vaporization followed by annealing in a hot oven, they can also be grown by LV at room temperature where it appears that sufficient heating of the target or ejecta is accomplished by high-repetition-rate [19] or long-pulse ($\sim$ 10 ms) [20] lasers.

Pulsed-LV with ns lasers is especially amenable to diagnostic investigations. The vaporizing pulse lasts only ∼ 10 ns and SWNT growth then can occur undisturbed from further excitation, even for single laser ablation events [16]. Recent spectroscopic measurements of the luminous laser plasma have been made for pulsed-LV at early times after Nd:YAG laser ablation (< 200 μs) [21] and after long-pulse CO_2 laser ablation at room temperature [20]. However, these measurements were limited to times while the ablated material is still quite hot.

Recently laser-induced luminescence (LIL), gated intensified charge-coupled device (ICCD) imaging, and optical emission spectroscopy were applied to understand how nanoparticles form and grow in pulsed laser ablation plumes at room temperature [22–24]. In this study these techniques are applied to understand how SWNTs grow from laser-vaporized carbon and catalyst species under typical growth conditions at high temperatures.

1 Experimental

The SWNT growth setup in Fig. 1 was fashioned after that used by Guo et al. [12]. It consists of a quartz tube (2 in. diameter, 24 in. length) mounted inside a hinged tube furnace (12 in. length) operating at 1000 °C. The quartz tube was O-ring sealed to standard 4.5-in. Conflat vacuum components. The ablation and probe laser beams entered through the same Suprasil window which was mounted in a vacuum flange. Argon gas was introduced around this window, controlled at 100 sccm to maintain a 500 Torr pressure, and pumped out through a needle valve downstream of a brass, water-cooled collector which was inserted into the quartz tube and positioned just outside the oven.

A 1-in. diameter graphite target containing 1 at. % each of Ni (Alfa, 2.2–3.0 μm, 99.9%) and Co (Alfa, 1–6 μm, 99.8%) powders was prepared with the carbon cement (Dylon GC) procedure described in [12]. The target was screwed onto a 0.25-in. diameter graphite rod and was rotated during operation. This rod was mounted along the tube axis through a hole in the collector. The ablation laser (Nd:YAG, 8-ns FWHM pulse) beam was focused to a 1.6 mm-diameter, donut-shaped spot on the target. Both the 1.06-μm fundamental and the 532-nm 2nd-harmonic were used (with no time delay) to provide a total energy of 140 mJ at the target. An unfocused, time-delayed XeCl-laser pulse (308 nm, 30-ns FWHM, 20 mJ/cm^2) was used to induce luminescence in the ablation plume.

To permit imaging of the plume with a gated ICCD-camera system (Princeton Instruments, 5-ns minimum gate, 200–820 nm spectral range) the furnace was opened for ∼ 3 seconds/per image. Alternatively, at each plume location lenses collected the plasma emission and/or laser-induced luminescence for optical emission spectroscopy using a 0.3-m spectrometer (Acton VM-503, resolution either 10 nm or 1.3 nm) with an intensified, gated diode array (Princeton Instrument IRY-700RB).

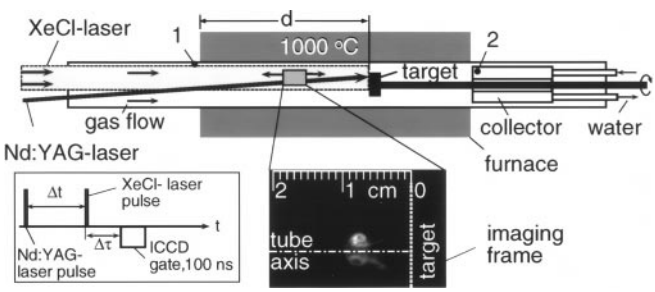


Fig. 1. Schematic of the 2″-diameter quartz tube and hot furnace used for laser vaporization growth of SWNT with in situ LIL-imaging and spectroscopy diagnostics. Beam geometries and imageable area are indicated. The *black dots* and the *numbers* show the collection points of the ablated material: 1-upstream; 2-collector. The C/Ni/Co target was positioned at two distances, d, from the front of the furnace. The *inset on the left* shows the relative timing between ablation (Nd:YAG) and and LIL-probe (XeCl) laser pulses (Δt), and the ICCD gate delay after the XeCl laser, ($\Delta\tau$)

2 Single-wall carbon nanotube deposits

Carbon nanotubes were synthesized under a variety of laser repetition rates, flow conditions, target positions, and numbers of total shots on the target. Following each run, deposits were collected at various points in the reactor and analyzed by bright-field TEM for correlation with the transport dynamics observed during the run with time-resolved imaging and/or spectroscopy.

When the target was positioned $d = 21$ cm from the front of the furnace, ∼ 10 μm-long SWNT bundles comprised very high fractions of the deposit found downstream on the collec-

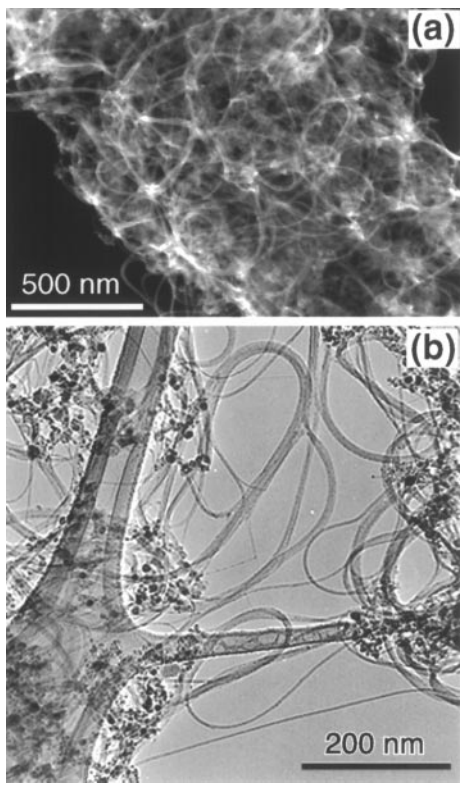


Fig. 2. **a** Field-emission SEM image of SWNT bundles, metal catalyst particles, and unconverted carbon soot produced by laser ablation of a C/Ni/Co target at 1000 °C. **b** TEM images of the raw soot collected: downstream on the collector (point 2, for $d = 21$ cm in Fig. 1), showing a very high fraction of SWNT bundles along with metal nanoparticles (*black dots*)

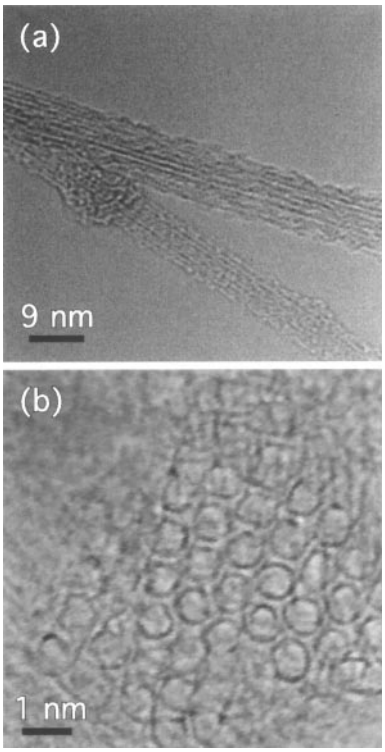


Fig. 3. a HRTEM images of bundles of SWNTs grown by laser ablation at 1000 °C. **b** A cross section of a nanotube bundle, revealing the shape and diameters of the hollow SWNT

tor. Figure 2a shows a field-emission SEM image of a deposit produced at a laser repetition rate of only 0.016 Hz. Over 90% of the deposit appears to consist of SWNT bundles as estimated from an areal analysis. Bright-field TEM images as in Fig. 2b show that metal catalyst nanoparticles and small amounts of amorphous carbon account for the rest of the deposit (as described in the literature) [12, 17, 18].

High-resolution scanning TEM images such as those in Fig. 3a confirm that the bundles consist of SWNTs. Cross-sectional images of the bundles permit an estimate of the nanotube diameters, as shown in Fig. 3b.

The ability to correlate these TEM and FESEM images of deposits with the plume transport dynamics (observed with in situ diagnostics, described below) permits conclusive assessment of the growth conditions. For example, the laser repetition rate for the deposit of Fig. 3 was chosen to ensure that the plume observed by imaging had completely cleared the growth chamber, so that the long nanotubes in Fig. 3 reflected SWNTs created on single laser shots.

3 Plume dynamics

To understand where and when nanotubes grow, ICCD imaging and spectroscopy of the C/Ni/Co plume were performed at different times after laser vaporization, Δt, spanning $20\,\text{ns} \leq \Delta t \leq 3\,\text{s}$. At early times, the ablated material is primarily in excited states and can be imaged directly from the plasma emission alone (without LIL excitation). Figure 4 compares the plume propagation at room temperature and at 1000 °C for $\Delta t \leq 200\,\mu\text{s}$. In each case, the plume exhibits oscillations in both axial and radial directions. Just after ablation, the plume expands both radially and axially and compresses the background gas. However, the initial kinetic energy of the plume contributes to its overexpansion into the background gas, and it expands past the position where the plume pressure equals that of the surrounding ambient. Recovering from this overexpansion, a backward motion is induced in the plume. In addition, a backward motion in the radial direction results in an axial focusing of the plume, as shown in Fig. 4.

The position of the leading edge of the plasma plumes is plotted versus time in Fig. 5. At 1000 °C, four axial oscillations, decaying in strength, are evident (at 40 ns, 200 ns, 2 μs, and 200 μs). At room temperature, three oscillations can be seen at 600 ns, 10 μs, and 300 μs. These oscillations result from the propagation of internal shock waves within the plume material, caused when the strong external shock wave is generated and detaches to propagate through the background gas. The plume material reflected from the contact

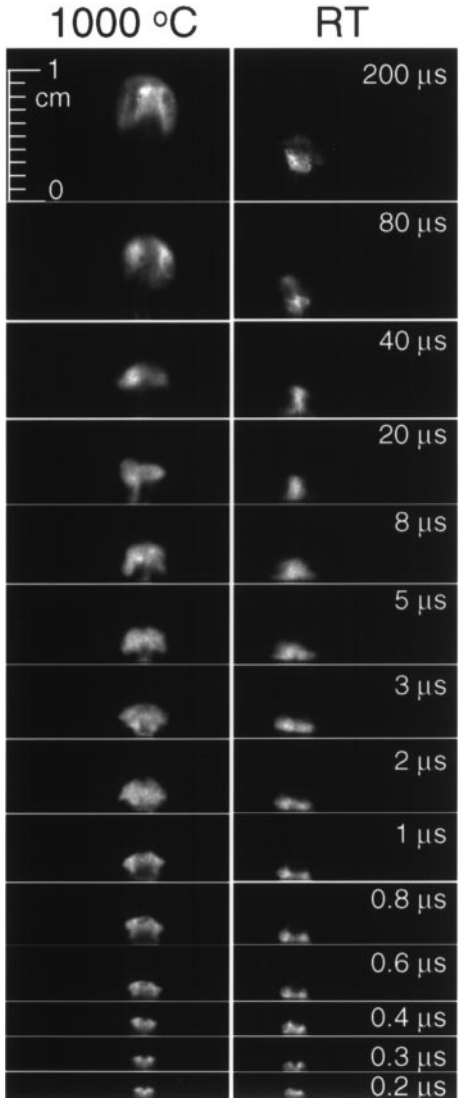


Fig. 4. ICCD images of the nascent plasma emission from the plume of vaporized C/Ni/Co target material at 1000 °C and at room temperature. The images show oscillations and self-focusing effects during the early dynamics of both plumes

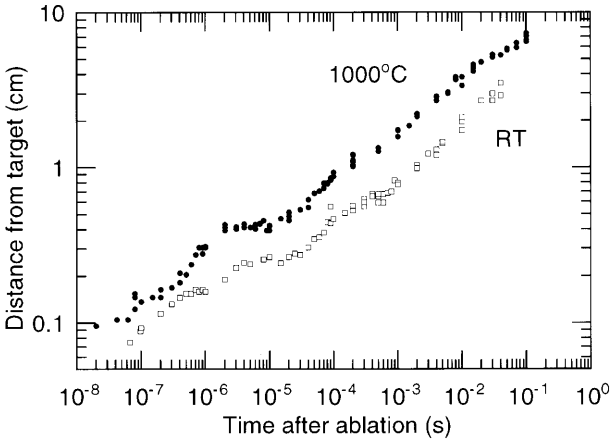


Fig. 5. Propagation of the leading edge of the ablation plumes shown in Fig. 4

surface between the plume and the background gas forms the internal shocks within the plume. These internal shocks serve to mix and additionally heat the atomic and molecular carbon and the catalyst species within the plume. This oscillatory behavior of laser plasmas propagating into background gases has been observed before [25, 26] and has been described theoretically [27].

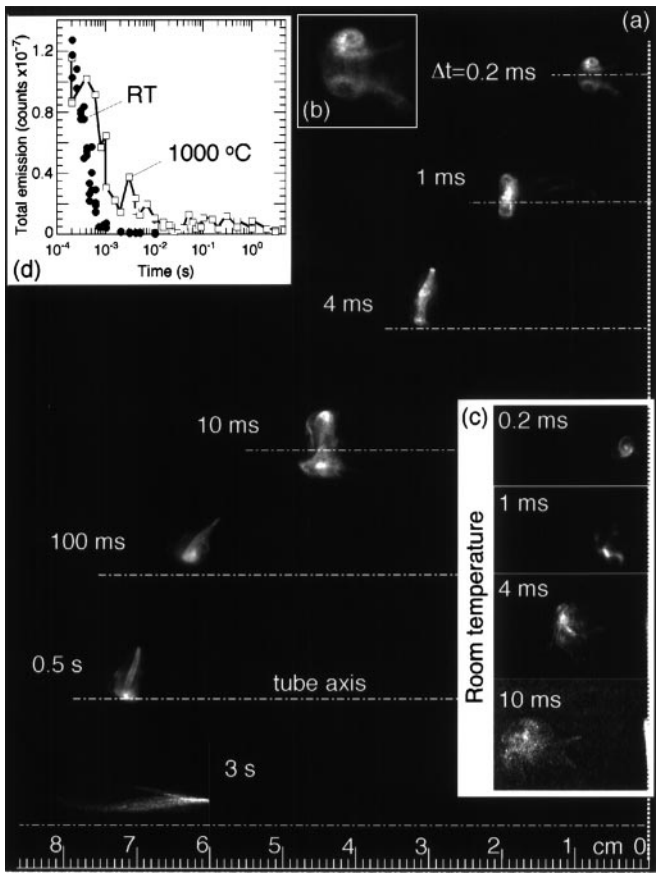


Fig. 6. a Laser-induced luminescence (LIL) images record the dynamics of the C/Ni/Co plume during SWNT synthesis. A Nd:YAG-laser vaporizes a C/Ni/Co target (right edge of figure) inside a 2-in. quartz tube at 1000 °C in 500-Torr Ar (flowing to the right at 100 sccm). Each image represents a different ablation event (100 ns gate width, opened $\Delta\tau = 100$ ns after the XeCl-laser pulse). **b** Enlarged view (2X) of the plume at 0.2 ms showing vortex ring. **c** LIL images at room temperature in 500-Torr Ar. **d** Integrated total emission from LIL-images acquired at the indicated times at 1000 °C and room temperature

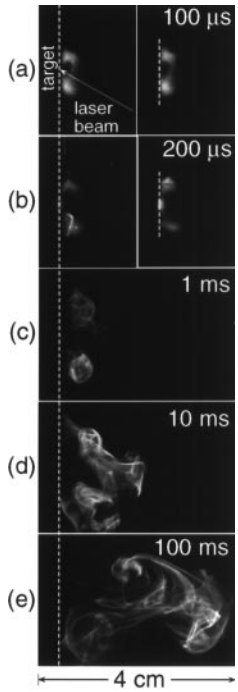


Fig. 7a–d. Images of Rayleigh-scattered (RS) light (308 nm XeCl-laser, 26 ns, 20 mJ/cm^2) and plasma emission from a carbon vapor plume generated by KrF-laser (248 nm, 28 ns FWHM, 7 J/cm^2) ablation of a pyrolitic graphite target into 300-Torr Ar at room temperature. The onset of clustering can be gauged by comparing **a–b** RS (*left*) vs. plasma emission images (*right*). **c–d** RS images at later times show the highly turbulent motion of plume material and its confinement within thin sheets

Following these plume oscillations [25, 26] during $\Delta t \leq 0.2$ ms, the ablated material is segregated into a vortex ring [27, 28] (or "smoke ring") [22] shown at $\Delta t = 0.2$ ms in Fig. 6a,b. This vortex ring is generated because of the viscous interaction between the plume and the background gas. Vortex formation efficiently mixes the ablated species with the background gas, promoting clustering of the plume species through three-body collisions. Then the vortex motion efficiently traps aggregated nanoparticles in a confined volume for long times (~ 3 s within ~ 1 cm^3 in this study) as shown in Fig. 6a.

The leading edge of the plume in Fig. 6a propagates with velocities of: 10^3 cm/s between $200\,\mu\text{s} < \Delta t < 2$ ms; 50 cm/s for $10\text{ ms} < \Delta t < 50$ ms; and 6 cm/s during $30\text{ ms} < \Delta t < 200$ ms. After $\Delta t = 2$ s the plume stops moving upstream, and the plane of the vortex ring tilts toward the tube axis, possibly due to flow currents or thermophoretic forces. The plume is then dragged by the gas flow back to the collector with an estimated flow velocity of 0.6 cm/s. Finally, nanotubes and unconverted soot deposit on the cool collector surface by thermophoresis [24].

At room temperature the plume dynamics at later times are completely different. As shown in Fig. 6c, the plume propagates slower in the axial direction and the motion of the material within the plume is highly turbulent. This turbulent behavior can be seen more explicitly using Rayleigh scatter-

ing from the nanoparticles which form during the cooling of the laser plasma in the background gas.

Rayleigh-scattering images of the plume evolution from a pure graphite target after KrF-laser ablation into 300-Torr Ar, are shown in Fig. 7. By comparing images of the plasma with and without secondary laser-irradiation, the emergence of Rayleigh scattering can be used to estimate the onset of plume condensation into nanoparticles [24]. Under these conditions, this time is estimated at 150 μs after ablation from these images. The highly turbulent behavior evident in the images likely results from the higher background gas density and smaller gas viscosity compared to those at 1000 °C. Despite the plume expansion, the ablated material remains confined to a relatively small volume within the thin sheets of multiple vortices.

4 Plume spectroscopy

Once the plume was located with ICCD imaging, several spectroscopic techniques were applied to probe its composition, including: optical emission spectroscopy (OES) from excited states in the plasma; optical absorption spectroscopy (OAS) and laser-induced fluorescence (LIF) from ground-state atoms and molecules; laser-induced blackbody emission and Rayleigh scattering from clusters, nanoparticles, and nanotubes.

To understand the nature of the carbon species ejected from the target, OAS was performed in vacuum [29]. Figure 8a shows a low-resolution (2.6 nm) absorption spectrum of the plume from a pure graphite target. The spectrum is dominated by C_3 absorption via $A\,^1\Pi_u \leftarrow X\,^1\Sigma_g$ and C_2 absorption: in the $\Delta v = -2$, -1, 0, and $+1$ sequences of the Swan $d\,^3\Pi_g \leftarrow a\,^3\Pi_u$ system; and the $\Delta v = 0$ sequence of the Mulliken $D\,^1\Sigma_u{}^+ \leftarrow X\,^1\Sigma_g{}^+$ system. The dominance of C_3 as the principal component of the plume is consistent with mass spectrometric measurements under the same conditions which show that the composition of the plume is primarily C_3, C_2, and C with very few higher order clusters (intensity ratio, $C{:}C_2{:}C_3C_4{:}C_5 = 56{:}35{:}100{:}2.5{:}1.6$) [30].

Ground-state C_3 in the plume was probed by laser-induced fluorescence with a 308-nm XeCl laser. Figure 8b shows the LIF spectrum of vaporized graphite in vacuum, displaying predominantly C_3 $A\,^1\Pi_u \rightarrow X\,^1\Sigma_g$ emission (resolution 10 nm). In addition, weak $\Delta v = 0$ C_2 Swan $d\,^3\Pi_g \rightarrow a\,^3\Pi_u$ emission can be observed at 516 nm. The XeCl laser can also excite vibronically induced absorption of C_3 via $^1\Pi_g$ and/or $^1\Delta_u \leftarrow X\,^1\Sigma_g$ [31].

In addition, the XeCl laser can induce emission from atomic Co (via $a\,^4F_J \leftarrow y\,^4G_J{}^o$) and from atomic Ni (via $a\,^3D_1 \leftarrow y\,^3D_2{}^o$). It therefore serves as an excellent probe of the ground state atoms and molecules in the plume, both for laser-induced luminescence (LIL) spectroscopy, and spectroscopic imaging (see Sect. 5) of the separate components.

From sets of images as in Fig. 6, both OES and LIL-spectra were acquired at 1000 °C and at room temperature (RT). Figures 9 and 10 present a short summary of these spectra, with nascent plasma emission plotted along with the LIL-emission whenever both could be detected. At early times in the plume expansion, close to the target while the plasma is very hot, the plume species are primarily electronically excited and this emission from

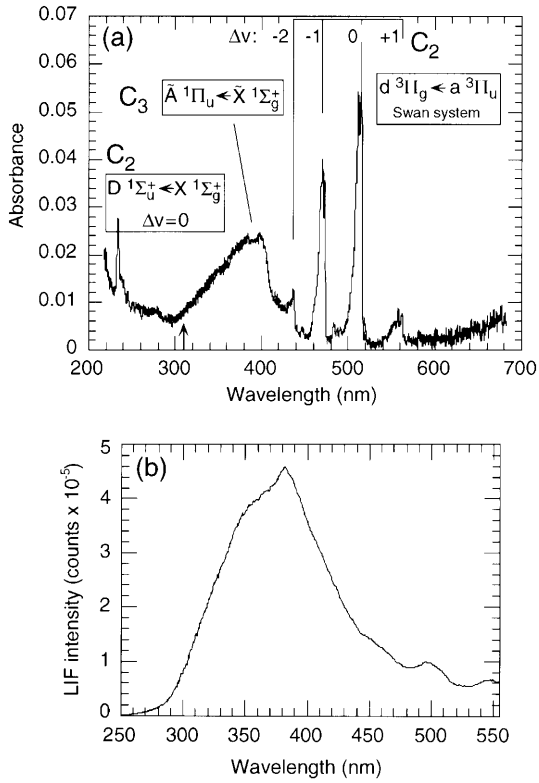


Fig. 8. **a** Gas-phase absorption spectrum of the carbon vapor plume generated by KrF laser (1.6 J/cm^2) ablation of a pyrolitic graphite target into vacuum. The spectrum was acquired $d = 0.5$ cm from the target surface, with a 400-ns gate width beginning 1 μs after the KrF-laser pulse. **b** LIF (XeCl-laser, 20 mJ/cm^2)-spectrum of C_3 in the carbon vapor plume generated by KrF-laser in vacuum

excited states dominates any laser-induced luminescence (from the ground states) both at 1000 °C (Fig. 9) and at RT (Fig. 10). Bright nascent emission from C_2 ($d\,^3\Pi_g \rightarrow a\,^3\Pi_u$, Swan system) and C_3 ($A\,^1\Pi_u \rightarrow X\,^1\Sigma_g$) dominate at these times. As the plasma expands, cools, and recombines, the ground states become populated, and LIL-emission emerges to compete with the nascent plasma emission. At 1000 °C, LIL from atomic Co in the 320–380 nm range (see high-resolution spectrum in Fig. 9) is clearly visible. Finally, the nascent plasma emission completely disappears, and only LIL from ground states remains. From previous measurements of clustering in laser plasmas expanding into background gases, this disappearance of the nascent plasma emission usually signals the onset of nanoparticle formation [22–24].

Laser-induced blackbody radiation can be used to signal the presence of carbon clusters [32], nanoparticles [33], and nanotubes. The intensity of this blackbody emission, I, is defined by

$$I = Ar^3(T_0 + \Delta T)^5 \,, \tag{1}$$

where T_0 is the initial temperature of the cluster, ΔT is the temperature increase due to laser heating, A is a constant, and r is the cluster radius.

Induced blackbody radiation becomes observable for $\Delta t > 200$ μs at 1000 °C (Fig. 9) and for $\Delta t > 50$ μs at room temperature (Fig. 10), coincident with the disappearance

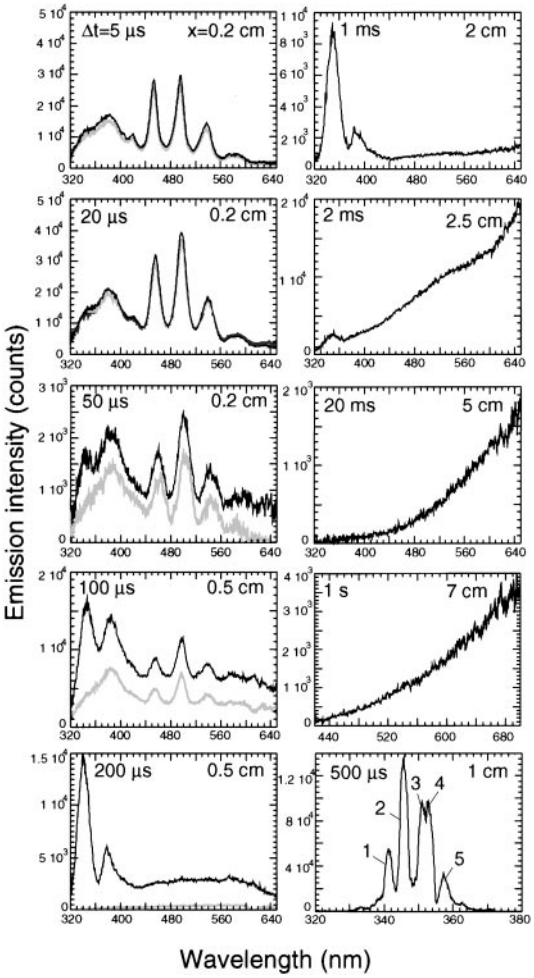


Fig. 9. Plasma emission (*lighter curve*) and laser-induced luminescence (*black curve*) spectra measured at different time delays after the ablation laser pulse, Δt, and distances, x, from the target at 1000 °C during SWNT synthesis: Acquisition times of 100 ns (for $\Delta t \leq 1$ ms) and 3.5 μs (for $\Delta t > 1$ ms) began 50 ns after the XeCl-laser pulse. Higher resolution (1.3-nm) LIL-spectrum in the 320–380 nm region (*bottom*, *right*) show the following transitions of atomic Co: (1) $b\,^4F_{7/2}$–$y\,^2G_{9/2}{}^0$ at 341.23 nm (2) $b\,^4F_{9/2}$-$y\,^4G_{11/2}{}^0$ at 345.35 nm (3) $b\,^4F_{5/2}$–$y\,^4G_{7/2}{}^0$ at 350.98 nm (4) $b\,^4F_{7/2}$–$y\,^4G_{9/2}{}^0$ at 352.98 nm, and (5) $b\,^4F_{5/2}$–$y\,^4D_{5/2}{}^0$ at 357.50 nm

of the C_2 and C_3 bands in both the plasma-emission and the laser-induced spectra for $\Delta t > 200$ μs at 1000 °C (and $\Delta t > 50$ μs at room temperature). From these two coincident measurements, we conclude that the carbon in the plume has all converted into clusters or larger aggregates by these times.

However, the atomic Co in the plume remains in atomic form long after the atomic and molecular carbon has disappeared. Integrating the emission in the 320–380 nm range from the LIL-spectra, the ground-state Co population is observed to peak at $\Delta t = 0.8$ ms and drop by an order of magnitude by $\Delta t = 2$ ms, permitting estimates of the Co clustering time of ~ 2 ms at 1000 °C (~ 1 ms at room temperature). A similar estimation for the Co clustering time can be performed using spectroscopic imaging of the Co vapor plume (see Fig. 11c in Sect. 5).

Blackbody radiation remains the only feature of the spectra taken for $\Delta t > 2$ ms at 1000 °C (and images, see Fig. 6d). At room temperature the laser-induced blackbody radiation

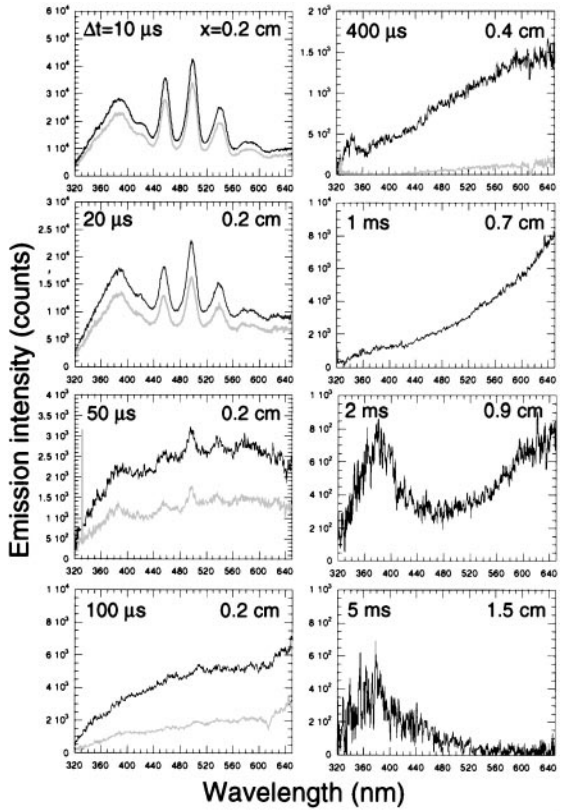


Fig. 10. Plasma emission (*lighter curve*) and laser-induced luminescence (*black curve*) spectra measured at room temperature at different time delays after the ablation laser pulse, Δt, and distances, x, from the target at (100 ns acquisition times began 35–75 ns after the XeCl-laser pulse)

disappears rapidly ($\Delta t \sim 5$ ms) for the low XeCl-laser intensity employed here (chosen for use at $T_0 = 1000$ °C in (1)). The last recognizable feature of the spectrum (see 5 ms image of Fig. 10) is a broad band at 380 nm, which is currently unidentified.

Hot particulates from the target would be easily imaged via their blackbody emission [34] or by Rayleigh scattering. It is important to note that we observed very few particulate ejecta for the properly outgassed, hard, Dylon-fabricated targets used in this study.

5 Spectroscopic imaging

Once sharp spectral features have been identified throughout a set of spectroscopic data (such as those described in Sect. 4), optical filters can be used in conjunction with ICCD imaging to selectively image different constituents of the plume. Here, the 320–380 nm spectral region was imaged at 1000 °C to locate the ground-state atomic Co in the plume.

Figure 11a,b shows these images along with corresponding images from a pure carbon plume (to assess the extent of blackbody emission from carbon species in the same spectral region). The total integrated luminescence from both sets of images is plotted in Fig. 11c. Again, the ground-state atomic Co population is observed to rise (due to population of the ground states via cooling of the hot plasma) and then decay as these atoms become incorporated in clusters. As in Fig. 6, the

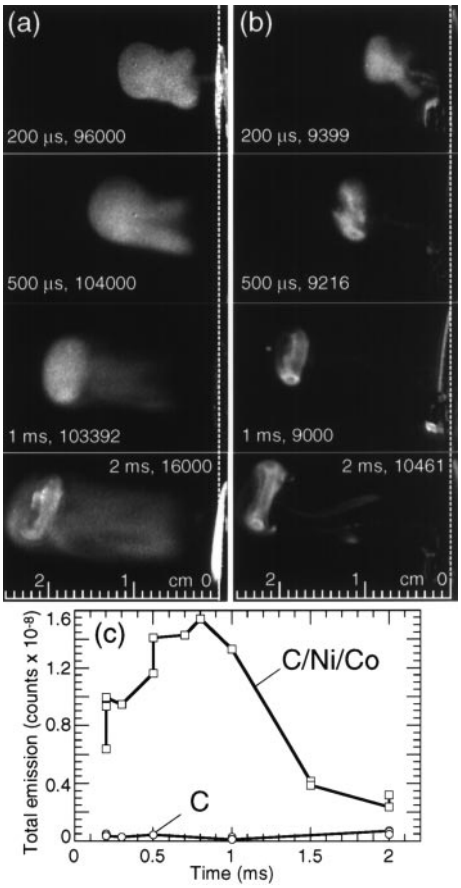


Fig. 11a–c. Selective imaging in the 320–380 nm spectral range at high temperature in 500-Torr Ar (100 ns gate width, $\Delta\tau = 0$, peak image intensities listed) **a** to locate ground-state atomic Co in the plume under conditions for SWNT synthesis using a graphite/Ni/Co target, and **b** to locate carbon species in the same region using a pure pyrolitic graphite target. **c** Total emission intensities from sets of images as shown in **a**, **b** compare the ground-state atomic Co temporal history (from C/Ni/Co target) and the background blackbody emission from carbon species (from C target) in the same spectral region

carbon in the plume clusters much more quickly and forms a vortex ring.

Further evidence of the sequential condensation of carbon and cobalt into clusters is the relatively uniform spatial distribution of atomic Co in the plume for $\Delta t < 2$ ms compared to the vorticity of the clustered carbon material (compare Fig. 11a,b). We believe that the higher diffusivity of the atomic Co effectively competes with the hydrodynamic trapping within the vortex during this time. Only the leading edge of the atomic-Co plume overlaps the carbon clusters during the condensation of the Co atoms. Assuming that Co clustering initiates nanotube formation, it appears that nanotubes grow from a feedstock of aggregated nanoparticles during seconds of time, confined first within the vortex ring and then by thermophoresis and fluid flow in the quartz tube.

6 Controlled SWNT growth with in situ diagnostics

In order to check this conclusion and estimate the growth rate, the target was positioned closer to the front furnace edge ($d = 12.5$ cm in Fig. 1) such that the plume spent only ~ 0.5 s within the hot zone before exiting the furnace in the upstream direction (as shown in Fig. 12a). From estimates of heat transfer coefficients and the thermal diffusivity of the gas (convection currents were not considered) the thermal gradient near the upstream edge of the furnace was estimated to extend ~ 5 cm into the tube ($d > 7$ cm in Fig. 12a). Thermophoresis is a powerful driving force for gas-suspended nanoparticles, driving them toward cooler regions of the growth chamber (and permitting capture of the nanomaterials on the collector) [24]. The intention in this experiment was to utilize the natural plume velocity in the upstream direction (determined from Fig. 6) to drive the plume into the region of the thermal gradient near the upstream edge of the furnace. LIL-ICCD imaging recorded the plume dynamics from ablation to deposition.

The plume motion in Fig. 12a is quite similar to that of Fig. 6 for $\Delta t < 100$ ms. However, for $\Delta t > 100$ ms the plume propagation changes dramatically, i.e. the plane of the ring vortex tilts relative to the tube axis and the ring elongates along this axis. Between 0.5–0.7 s the plume exits the furnace in this tilted orientation to deposit onto the upper surface of the quartz tube (at point 1 in Fig. 1).

A transmission electron microscopy (TEM) image of this deposit is shown in Fig. 12b. The collected material consists of aggregated carbon and metal-catalyst nanoparticles, and thin SWNT bundles of only ~ 100 nm length. The relative yield of the carbon particles is larger than the yield of the carbon nanotubes which clearly shows that the time spent by the plume in the hot zone (~ 0.5 s) was not sufficient to convert all of the carbon material into nanotubes. This combined

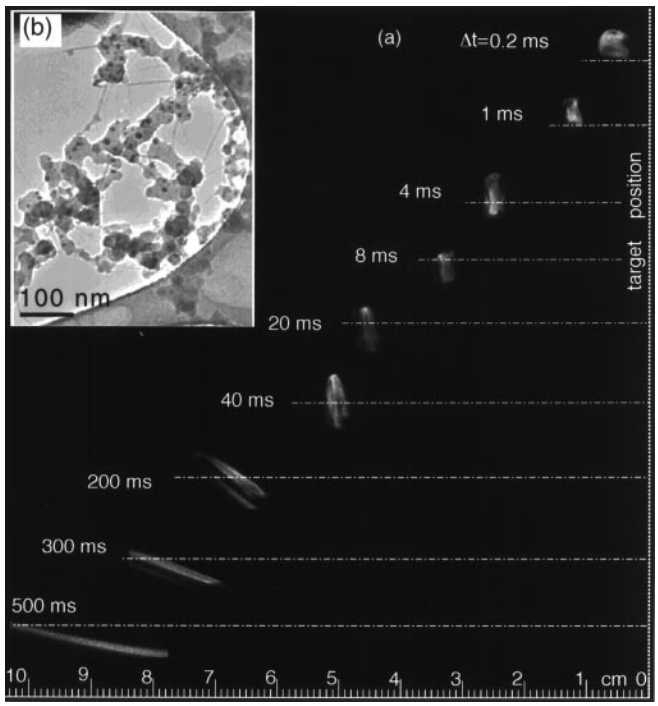


Fig. 12. a Laser-induced luminescence (LIL) images of the C/Ni/Co plume during synthesis of SWNT with controlled growth times of ~ 0.5 s. The target is positioned at $d = 12.5$ cm in Fig. 1. **b** Corresponding deposit collected at point 1 in Fig. 1, showing short (~ 100 nm) SWNT in the early stages of growth

imaging and TEM analysis permits the average growth rate at 1000 °C to be estimated at $\sim 0.2\,\mu m/s$.

7 Conclusions

In summary, we have used laser-induced imaging and spectroscopy diagnostic techniques, along with ex situ TEM, to determine that single-walled carbon nanotubes form over several seconds inside the hot furnace after laser vaporization. Size-controlled SWNTs were formed in short 100-nm lengths by controlling the growth time with these diagnostics to ~ 0.5 s, yielding $0.2\,\mu m/s$ for the first estimate of the SWNT growth rate by LV. The spectroscopy at early times after laser ablation indicates that the plume initially consists of atomic and molecular species, with no evidence of hot molten particulates which were recently suggested as the primary ejecta [35]. Condensation of carbon occurs within 0.2 ms after ablation, whereas atomic Co condenses much later (between 1.5 and 2 ms) at 1000 °C. The nanotubes grow within a vortex ring which trap clusters and aggregates within a $\sim$ 1-cm^3 volume during very long periods of time (~ 3 s). Assuming that metal catalyst clusters are required before carbon nanotube growth begins, we conclude that feedstock for nanotube growth is this mixture of carbon and metal catalyst nanoparticles. It is quite possible that the carbon clusters serve as the condensation centers for the metal clusters, and mixed carbon–catalyst alloy clusters are produced. Finally, we produced high-purity SWNT deposits with laser repetition rates as low as 0.016 Hz. Through these images of the growth process, we conclude that long ($\sim 10\,\mu m$) SWNTs can form from the small amount of material vaporized in a single-laser shot, a remarkable feat of self-assembly.

Acknowledgements. The authors gratefully acknowledge research assistance by M.L. Simpson, M. Guillorn, V.I. Merkulov, and D.J. Rader. This work was sponsored by the Division of Material Science, U.S. Department of Energy under Contract No. DE-AC05-96-R22464 with Lockheed Martin Energy Research Corp.

References

1. B.I. Yakobson, R.E. Smalley: Am. Sci. **85**, 324 (1997)
2. S. Iijima, T. Ichihashi: Nature **363**, 603 (1993)
3. D.S. Bethune, C.H. Kiang, M.S. de Vries, G. Gorman, R. Savoy, J. Vazquez, R. Beyers: Nature **363**, 605 (1993)
4. Ph. Avouris, T. Hertel, R. Martel, T. Schmidt, H.R. Shea, R.E. Walkup: Appl. Surf. Sci. **141**, 201 (1999)
5. R.H. Baughman, C. Cui, A.A. Zakhidov, Z. Iqbal, J.N. Barisci, G.M. Spinks, G.G. Wallace, A. Mazzoldi, D. De Rossi, A.G. Rinzler, O. Jaschinski, S. Roth, M. Kertesz: Science **284**, 1340 (1999)
6. P. Poncharal, Z.L. Wang, D. Ugarte, W.A. de Heer: Science **283**, 1513 (1999)
7. S.S. Wong, E. Joselevich, A.T. Woolley, C.L. Cheung, C.M. Lieber: Nature **394**, 52 (1998)
8. H. Schmid, H.-W. Fink: Appl. Phys. Lett. **70**, 2679 (1997)
9. J. Li, C. Papadopoulos, J.M. Xu, M. Moskovits: Appl. Phys. Lett. **75**, 367 (1999)
10. J. Liu, A.G. Rinzler, H. Dai, J.H. Hafner, R.K. Bradley, P.J. Boul, A. Lu, T. Iverson, K. Shelimov, C.B. Huffman, F. Rodriguez-Macias, Y.-S. Shon, T.R. Lee, D.T. Colbert, R.E. Smalley: Science **280**, 1253 (1998)
11. Y. Ye, C.C. Ahn, C. Witham, B. Fultz, J. Liu, A.G. Rinzler, D. Colbert, K.A. Smith, R.E. Smalley: Appl. Phys. Lett. **74**, 2307 (1999)
12. T. Guo, P. Nikolaev, A. Thess, D.T. Colbert, R.E. Smalley: Chem. Phys. Lett. **236**, 419 (1995)
13. C. Journet, W.K. Maser, P. Bernier, A. Loiseau, M. Lamy de la Chapelle, S. Lefrant, P. Deniard, R. Lee, J.E. Fisher: Nature **388**, 756 (1997)
14. H.M. Cheng, F. Li, G. Su, H.Y. Pan, L.L. He, X. Sun, M.S. Dresselhaus: Appl. Phys. Lett. **72**, 3282 (1998)
15. B.C. Satiskumar, A. Govindaraj, R. Sen, C.N.R. Rao: Chem. Phys. Lett. **293**, 47 (1998)
16. M. Yudasaka, T. Ichihashi, T. Komatsu, S. Iijima: Chem. Phys. Lett. **299**, 91 (1999)
17. A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y.H. Lee, S.G. Kim, A.G. Rinzler, D.T. Colbert, G.E. Scuseria, D. Tomanek, J.E. Fisher, R.E. Smalley: Science **273**, 483 (1996)
18. A.G. Rinzler, J. Liu, H. Dai, P. Nikolaev, C.B. Huffman, F.J. Rodriguez-Macias, P.J. Boul, A.H. Lu, D. Heymann, D.T. Colbert, R.S. Lee, J.E. Fisher, A.M. Rao, P.C. Eklund, R.E. Smalley: Appl. Phys. A **67**, 29 (1998)
19. A.C. Dillon, P.A. Parilla, K.M. Jones, G. Riker, M.J. Heben: In *Advances in Laser Ablation of Materials*, ed. by R.K. Singh, D.H. Lowndes, D.B. Chrisey, E. Fogarassy, J. Narayan (Warrendale, PA, USA: Mater. Res. Soc. 1998) pp. 403–408
20. F. Kokai, K. Takahashi, M. Yudasaka, R. Yamada, T. Ichihashi, S. Iijima: J. Phys. Chem. B **103**, 4346 (1999)
21. S. Arepalli, C.D. Scott: Chem. Phys. Lett. **302**, 139 (1999)
22. D.B. Geohegan, A.A. Puretzky, G. Duscher, S.J. Pennycook: Appl. Phys. Lett. **72**, 2987 (1998)
23. D.B. Geohegan, A.A. Puretzky, G. Duscher, S.J. Pennycook: Appl. Phys. Lett. **73**, 438 (1998)
24. D.B. Geohegan, A.A. Puretzky, D.J. Rader: Appl. Phys. Lett. **74**, 3788 (1999)
25. D.B. Geohegan, A.A. Puretzky, R.L. Hettich, X.-Y. Zheng, R.E. Haufler, R.N. Compton: In Advanced Materials '93, IV/Laser and Ion Beam Modification of Materials, ed. by I. Yamada, et al., Trans. Mater. Res. Soc. Jpn. **17**, 349 (1994)
26. Images also available online at www.ornl.gov/∼odg
27. A.V. Bulgakov, N.M. Bulgakova: J. Phys. D: Appl. Phys. **31**, 693 (1998)
28. F. Garrelie, C. Champeaux, A. Catherinot: Appl. Phys. A **69**, 45 (1999)
29. D.B. Geohegan, A.A. Puretzky: Appl. Phys. Lett. **67**, 197 (1995)
30. D. Krajnovich: J. Chem. Phys **102**, 726 (1995)
31. E.A. Rohlfing: J. Chem. Phys. **91**, 4531 (1989)
32. D.B. Geohegan, A.A. Puretzky: Mater. Res. Symp. Proc. **397**, 55 (1996)
33. E.A. Rohlfing: J. Chem. Phys. **89**, 6103 (1988)
34. D.B. Geohegan: Appl. Phys. Lett. **62**, 1463 (1993)
35. M. Yudasaka, R. Yamada, N. Sensui, T. Wilkins, T. Ichihashi, S. Iijima: J. Phys. Chem. B **103**, 6224 (1999)

In situ imaging and spectroscopy of single-wall carbon nanotube synthesis by laser vaporization

A. A. Puretzky, D. B. Geohegan,[a] X. Fan, and S. J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

(Received 9 September 1999; accepted for publication 10 November 1999)

The synthesis of single-wall carbon nanotubes by Nd:YAG laser vaporization of a graphite/(Ni, Co) target is investigated by laser-induced luminescence imaging and spectroscopy of Co atoms, C_2 and C_3 molecules, and clusters at 1000 °C in flowing 500 Torr Ar. These laser-induced emission images under typical synthesis conditions show that the plume of vaporized material is segregated and confined within a vortex ring which maintains a ~1 cm^3 volume for several seconds. Using time-resolved spectroscopy and spectroscopic imaging, the time for conversion of atomic and molecular species to clusters was measured for both carbon (200 μs) and cobalt (2 ms). This rapid conversion of carbon to nanoparticles, combined with transmission electron microscopy analysis of the collected deposits, indicate that nanotube growth occurs over several seconds in a plume of mixed nanoparticles. By adjusting the time spent by the plume within the high-temperature zone using these *in situ* diagnostics, single-walled nanotubes of controlled length were grown at an estimated rate of 0.2 μm/s. © *2000 American Institute of Physics.* [S0003-6951(00)02102-1]

The remarkable electronic and structural properties of carbon nanotubes promise to revolutionize application areas from nanoscale electronics to ultralightweight structural materials.[1] Since electronic transport in single-wall carbon nanotubes (SWNT) is directly related to their atomic structure, it is essential to understand what controls nanotube size and structure during synthesis. Furthermore, methods for large-scale production of SWNT have not been developed because so little is understood about the synthesis process.

Up until now, three principal methods have been used to synthesize carbon nanotubes. These include laser vaporization (LV),[2] dc-arc vaporization (AV),[3,4] and chemical vapor deposition (CVD).[5,6] Unlike CVD, both LV and AV now produce nearly exclusively SWNT. However, very little is known about where and when the SWNT are formed in LV or AV growth chambers, the identity of the precursor species, and the growth rates.

Pulsed LV is especially amenable to diagnostic investigations. The vaporizing pulse lasts only ~10 ns and SWNT growth then can occur undisturbed from further excitation, even for single laser ablation events.[7] Recent spectroscopic measurements of the luminous laser plasma have been made for pulsed LV at early times after Nd:YAG laser ablation (<200 μs)[8] and after long-pulse CO_2 laser ablation at room temperature.[9] However, these measurements were limited to times while the ablated material is still quite hot.

Recently, laser induced luminescence (LIL), gated intensified charge-coupled device (ICCD) imaging, and optical emission spectroscopy were applied to understand how nanoparticles form and grow in pulsed laser ablation plumes at room temperature.[10–12] In this study these techniques are applied to investigate SWNT growth under typical ablation conditions at high temperatures.

The SWNT growth setup in Fig. 1 was fashioned after that used by Guo *et al.*[2] It consists of a quartz tube (2 in. diameter, 24 in. length) mounted inside a hinged tube furnace (12 in. length) operating at 1000 °C. The quartz tube was O-ring sealed to standard 4.5-in.-Conflat vacuum components. The ablation and probe laser beams entered through the same Suprasil window which was mounted in a vacuum flange. Argon gas was introduced around this window, controlled at 100 sccm to maintain a 500 Torr pressure, and pumped out through a needle valve downstream of a brass, water-cooled collector which was inserted as shown.

A 1-in.-diameter graphite target containing 1 at. % each of Ni (Alfa, 2.2–3.0 μm, 99.9%) and Co (Alfa, 1–6 μm, 99.8%) powders was prepared with the carbon cement (Dylon GC) procedure described in Ref. 2. The target was screwed onto a 0.25-in.-diameter graphite rod and was rotated during operation. The ablation laser beam (single 8 ns pulses, 140 mJ of combined 1.06 μm and 532 nm outputs) was focused to a 1.6-mm-diameter donut-shaped spot on the target. An unfocused, time-delayed XeCl-laser (308 nm, 30 ns pulse, 20 mJ/cm^2) was used to induce luminescence in the ablation plume.

The furnace was opened for ~3 s to permit imaging of the plume with a gated ICCD-camera system (Princeton Instruments, 5-ns-minimum gate, 200–820 nm spectral range).

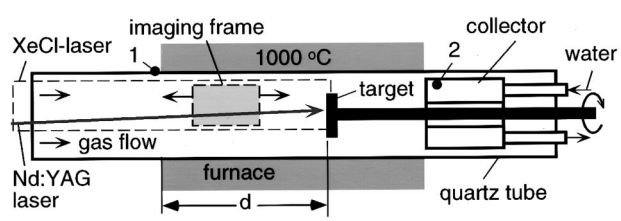


FIG. 1. Schematic of the 2-in.-diameter quartz tube and hot furnace used for laser vaporization growth of SWNT with *in situ* LIL imaging and spectroscopy diagnostics. Beam geometries and imageable area are indicated. The black dots and the numbers show the collection points of the ablated material: (1) upstream; (2) collector. The C/Ni/Co target was positioned at different distances, *d*, from the front of the furnace.

[a]Electronic mail: odg@ornl.gov

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

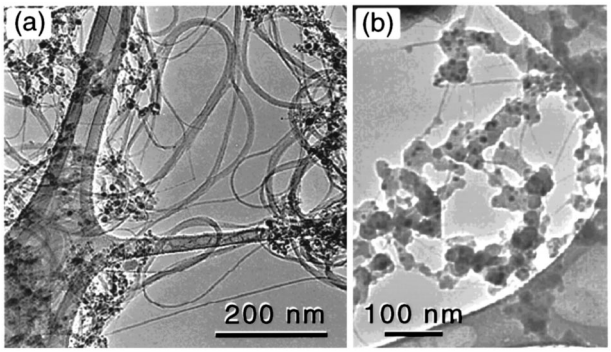


FIG. 2. TEM images of the raw soot collected: (a) downstream on the collector (point 2, for $d=21$ cm in Fig. 1), showing a very high fraction of SWNT bundles along with metal nanoparticles (black dots); (b) upstream (point 1, for $d=12.5$ cm in Fig. 1) showing short (~100 nm) SWNT in the early stages of growth, controlled by limiting the growth time to 0.5 s.

At each plume location, the plasma emission and/or laser-induced luminescence was collected for optical emission spectroscopy using a 0.3 m spectrometer (Acton VM-503, resolution either 10 or 1.3 nm) with an intensified, gated diode array (Princeton Instrument IRY-700RB).

As shown in Fig. 2(a), ~10-μm-long SWNT bundles comprised very high fractions of the deposit found downstream on the collector (>70%–90% as estimated from field-emission SEM analysis) when the target was positioned $d=21$ cm from the front of the furnace. Metal catalyst (<20 nm diameter) and small amounts of amorphous carbon nanoparticles accounted for the rest of the deposit (as described in the literature).[2] EELS investigation of individual metal-catalyst nanoparticles (using a 0.3-nm-diameter electron beam) revealed that their compositions were reproducibly $Ni_xCo_y(x\approx y\approx 0.5)$, with both Ni and Co uniformly distributed inside each nanoparticle to within sub-nm dimensions.

To determine where and when SWNT growth occurred, LIL-ICCD imaging and spectroscopy of the C/Ni/Co plume were performed at different times after laser vaporization, Δt, spanning $20\,\text{ns}\leqslant\Delta t\leqslant 3\,\text{s}$. For $\Delta t\leqslant 0.2$ ms, a series of shocks within the plume[13,14] results in segregation of the ablated material into the vortex ring[15,16] (or ''smoke ring'')[10] shown at $\Delta t=0.2$ ms in Figs. 3(a) and 3(b). As shown in Fig. 3(a), the vortex motion efficiently traps aggregated nanoparticles in a confined volume for long times (~3 s within ~1 cm^3 in this study).

The leading edge of the plume propagates with velocities of: 10^3 cm/s between $200\,\mu\text{s}<\Delta t<2$ ms; 50 cm/s for $10\,\text{ms}<\Delta t<50$ ms; and 6 cm/s at $30\,\text{ms}<\Delta t<200$ ms. After $\Delta t=2$ s the plume stops moving upstream, and the plane of the vortex ring tilts toward the tube axis, possibly due to flow currents or thermophoretic forces. The plume is then dragged by the gas flow back to the collector with an estimated flow velocity of 0.6 cm/s. Finally, nanotubes, metal catalyst particles and unconverted carbon soot deposit on the cool collector surface by thermophoresis.[12] LIL-ICCD imaging ensured that all deposits and images were collected without interference from previous laser shots using laser repetition rates less than 0.016 Hz.

Optical emission spectra were obtained for each image. Figure 4 presents a short summary. At early times in the plume expansion, emission from excited states in the plasma

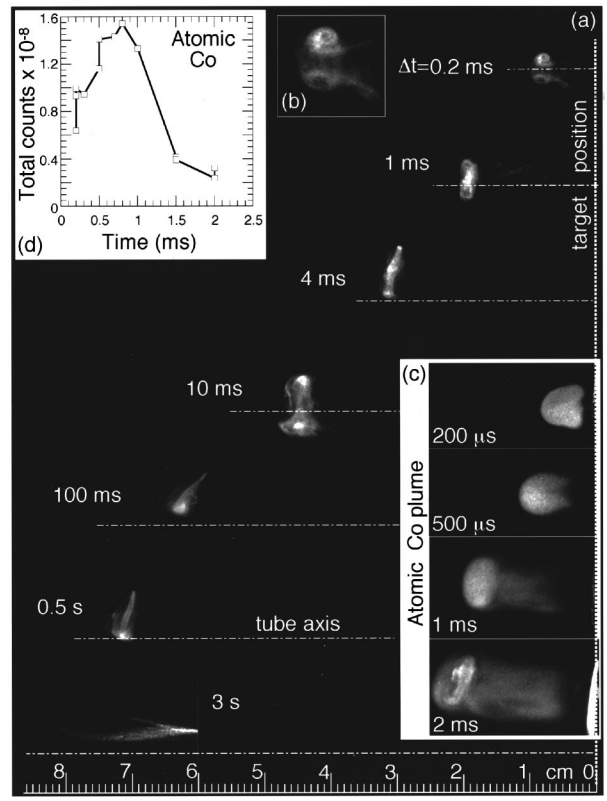


FIG. 3. Laser-induced luminescence (LIL) images of the C/Ni/Co plume during SWNT synthesis. A Nd:YAG laser vaporizes a C/Co/Ni target (right edge of figure) inside a 2-in.-quartz tube at 1000 °C in 500 Torr Ar (flowing to the right at 100 sccm, ~0.6 cm/s at the center of the tube). Each image represents a different ablation event. (a) Unfiltered images show the location of all C/Co/Ni species during nanotube growth (100 ns gate width, opened $\Delta\tau=100$ ns after the XeCl-laser pulse); (b) selective imaging in the 320–380 nm spectral region to locate ground-state atomic Co in the plume (100 ns gate width, $\Delta\tau=0$); (c) temporal history of the total number density of ground-state atomic Co, determined from integrating the LIL from a complete set of atomic-Co images shown in (b).

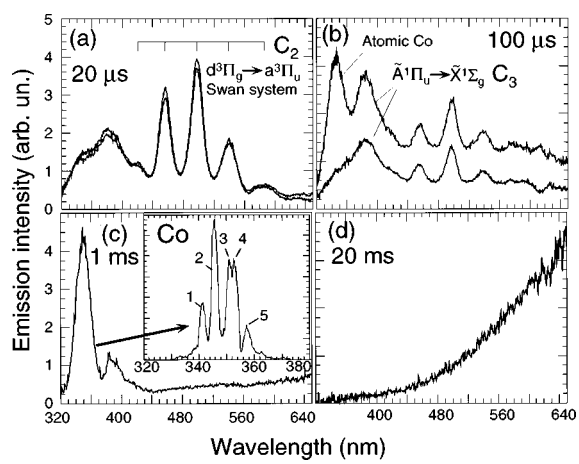


FIG. 4. Plasma emission (dashed) and laser-induced luminescence (solid) spectra measured at different time delays after the ablation laser pulse, Δt, and distances, x, from the target: (a) $\Delta t=20\,\mu$s, $x=0.2$ cm; (b) $\Delta t=100\,\mu$s, $x=0.5$ cm; (c) $\Delta t=1$ ms, $x=2$ cm; and (d) $\Delta t=20$ ms, $x=5$ cm. Acquisition times of 100 ns [(a), (b)] and 3.5 μs [(c), (d)] began 50 ns after the XeCl-laser pulse. The inset in (c) shows a 1.3-nm-resolution spectrum of induced fluorescence from the following transitions: (1) $b^4F_{7/2}-y^2G^0_{9/2}$ at 341.23 nm; (2) $b^4F_{9/2}-y^4G^0_{11/2}$ at 345.35 nm; (3) $b^4F_{5/2}-y^4G^0_{7/2}$ at 350.98 nm; (4) $b^4F_{7/2}-y^4G^0_{9/2}$ at 352.98 nm; and (5) $b^4F_{5/2}-y^4D^0_{5/2}$ at 357.50 nm.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

dominates any laser-induced luminescence. As in Fig. 4(a), the $d^3\Pi_g \rightarrow a^3\Pi_u$ Swan system of C_2 and the $A^1\Pi_u \rightarrow X^1\Sigma_g$ band C_3 are prominent at these times. Later when the plasma expands, cools, and recombines, the 308-nm-XeCl laser can induce emission by pumping transitions from the ground states: of C_3 (via $A^1\Pi_u \leftarrow X^1\Sigma_g$); of atomic Co (via $a^4F_J \leftarrow y^4G_J^0$); and of atomic Ni (via $a^3D_1 \leftarrow y^3D_2^0$).

In addition, blackbody radiation can be induced from carbon clusters,[17] nanoparticles,[18] and nanotubes. The intensity of this blackbody emission, I, is defined by $I = Ar^3(T_0 + \Delta T)^5$ where T_0 is the initial temperature of the cluster, ΔT is the temperature increase due to laser-heating, A is a constant, and r is the cluster radius. This radiation becomes observable coincident with the disappearance of the C_2 and C_3 bands in both the plasma emission and the laser-induced spectra for $\Delta t > 200\ \mu$s at 1000 °C (and $\Delta t > 100\ \mu$s at room temperature). We conclude that nearly all of the carbon in the plume has converted into clusters or larger aggregates by these times.

However, the Co in the plume remains in atomic form until much later. As shown in Figs. 3(c) and 4(c), the ground-state Co population peaks at $\Delta t = 0.8$ ms and drops by an order of magnitude by $\Delta t = 2$ ms, permitting estimates of the Co clustering time of ~2 ms at 1000 °C (~1 ms at room temperature). Blackbody radiation remains the only feature of the spectra taken at all later times [Fig. 4(d)].

Further evidence of the sequential condensation of carbon and cobalt into clusters is the relatively uniform spatial distribution of atomic Co in the plume for $\Delta t < 2$ ms compared to the vorticity of the clustered carbon material [compare Figs. 3(a) and 3(b)]. We believe that the higher diffusivity of the atomic Co effectively competes with the hydrodynamic trapping during this time. Only the leading edge of the atomic-Co plume overlaps the carbon-cluster vortex ring during the condensation of the Co atoms. The Ni_xCo_y alloy particles indicate similar condensation times for Ni and Co. It is quite possible that the carbon clusters serve as condensation centers for metal cluster growth.

Both imaging and spectroscopy indicate that within the first few milliseconds after laser ablation, atoms and molecules of both carbon and metal catalyst disappear due to condensation into nanoparticles. Unless SWNTs grow very rapidly from atoms and molecules within these first few milliseconds, the majority of growth appears to occur from a feedstock of mixed nanoparticles over seconds of annealing time.

In order to check this conclusion and estimate the SWNT growth rate, the target was positioned at several distances from the front furnace edge to adjust the time spent within the hot zone before exiting the furnace in the upstream direction. LIL-ICCD imaging recorded the plume dynamics from ablation to deposition and the collected deposits were examined by TEM.

Figure 2(b) shows a TEM image of a deposit collected for a target position of $d = 12.5$ cm (see Fig. 1), where LIL-ICCD imaging measured a time of 0.5–0.7 s spent by the plume inside the furnace. The collected material consists of aggregated carbon and metal-catalyst nanoparticles, and thin SWNT bundles of only ~100 nm length, yielding an average growth rate of 0.2 μm/s at 1000 °C. The relative yield of the carbon particles is much larger than the yield of the carbon nanotubes which clearly shows that the time spent by the plume in the hot zone (~0.5 s) was not sufficient to convert all of the carbon material into nanotubes.

In summary, these results confirm the conclusions of the time-resolved imaging and spectroscopy measurements (along with *ex situ* TEM, EELS, and FESEM). Although SWNT may initiate growth during the first few milliseconds after laser ablation, the majority of growth occurs inside the oven from a mixture of gas-suspended carbon and metal-catalyst nanoparticles which are hydrodynamically trapped in a vortex ring within a ~1 cm^3 volume for several seconds. The spectroscopy at early times after laser ablation indicates that the plume initially consists of atomic and molecular species, with no evidence of hot molten particulates which were recently suggested as the primary ejecta.[19] Condensation of carbon occurs within 0.2 ms after ablation, while atomic Co condenses much later (between 1.5 and 2 ms). Through these images of the growth process, we conclude that long (~10 μm) SWNT can form from the small amount of material vaporized in a single-laser shot, a remarkable feat of self-assembly.

The authors gratefully acknowledge research assistance by M. L. Simpson, M. Guillorn, and helpful conversations with D. J. Rader. This work was sponsored by the Division of Material Science, U.S. Department of Energy under Contract No. DE-AC05-96-R22464 with Lockheed Martin Energy Research Corp.

[1] B. Yakobson and R. E. Smalley, Am. Sci. **85**, 324 (1997).
[2] T. Guo, P. Nikolaev, A. Thess, D. T. Colbert, and R. E. Smalley, Chem. Phys. Lett. **236**, 419 (1995).
[3] T. S. Iijima and T. Ichihashi, Nature (London) **363**, 603 (1993).
[4] C. Journet, W. K. Maser, P. Bernier, A. Loiseau, M. Lamy de la Chapelle, S. Lefrant, P. Deniard, R. Lee, and J. E. Fisher, Nature (London) **388**, 756 (1997).
[5] H. M. Cheng, F. Li, G. Su, H. Y. Pan, L. L. He, X. Sun, and M. S. Dresselhaus, Appl. Phys. Lett. **72**, 3282 (1998).
[6] B. C. Satiskumar, A. Govindaraj, R. Sen, and C. N. R. Rao, Chem. Phys. Lett. **293**, 47 (1998).
[7] M. Yudasaka, T. Ichihashi, T. Komatsu, and S. Iijima, Chem. Phys. Lett. **299**, 91 (1999).
[8] S. Arepalli and C. D. Scott, Chem. Phys. Lett. **302**, 139 (1999).
[9] F. Kokai, K. Takahashi, M. Yudasaka, R. Yamada, T. Ichihashi, and S. Iijima, J. Phys. Chem. **B103**, 4346 (1999).
[10] D. B. Geohegan, A. A. Puretzky, G. Duscher, and S. J. Pennycook, Appl. Phys. Lett. **72**, 2987 (1998).
[11] D. B. Geohagan, A. A. Puretzky, G. Duscher, and S. J. Pennycook, Appl. Phys. Lett. **73**, 438 (1998).
[12] D. B. Geohegan, A. A. Puretzky, and D. J. Rader, Appl. Phys. Lett. **74**, 3788 (1999).
[13] D. B. Geohegan, A. A. Puretzky, R. L. Hettich, X.-Y. Zheng, R. E. Haufler, and R. N. Compton, Trans. Mater. Res. Soc. Jpn. **17**, 349 (1994).
[14] Images also available online at www.ornl.gov/~odg.
[15] A. V. Bulgakov and N. M. Bulgakova, J. Phys. D: Appl. Phys. **31**, 693 (1998).
[16] F. Garrelie, C. Champeaux, and A. Catherinot, Appl. Phys. A: Mater. Sci. Process. **69**, 45 (1999).
[17] D. B. Geohegan and A. A. Puretzky, Mater. Res. Soc. Symp. Proc. **397**, 55 (1996).
[18] E. A. Rohlfing, J. Chem. Phys. **89**, 6103 (1988).
[19] M. Yudasaka, R. Yamada, N. Sensui, T. Wilkins, T. Ichihashi, and S. Iijima, J. Phys. Chem. B **103**, 6224 (1999).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

Ab initio calculations of rigid-body displacements at the Σ5 (210) tilt grain boundary in TiO_2

S. B. Sinnott*
The University of Kentucky, Department of Chemical and Materials Engineering, Lexington, Kentucky 40506-0046

R. F. Wood and S. J. Pennycook
Oak Ridge National Laboratory, Solid State Division, Oak Ridge, Tennessee 37831-6032
(Received 22 November 1999)

Through extensive *ab initio* total energy calculations of the Σ5 (210) grain boundary in TiO_2 (rutile), the remarkably large rigid-body contractions found by high resolution electron microscopy and *Z*-contrast imaging are explored. Unlike previous calculations, this work emphasizes the importance of the displacements of the two halves of the bicrystal towards one another. Previous experimental work suggests that the stochiometry and ionic charges at the boundary are preserved at their bulk values and therefore vacancies and impurities within the grain boundary core were not considered. The calculations predict no significant rigid-body contraction or expansion. The implications of this result for both computational and experimental studies of rigid-body displacements at grain boundaries are discussed.

It is well known that the grain boundaries in polycrystalline materials have a profound effect on their electronic and vibrational properties, but the exact nature of these defects is not well understood even in Si, the most studied electronic material. Transition metal oxides have many physical properties that are important for such diverse applications as catalysis,[1,2] electronic devices,[3] and electrochemical cells,[4] to name a few. One of the most widely used transition metal oxides is TiO_2, a semiconducting ceramic used in catalytic processing of organic wastes,[5] photovoltaic cells,[6] protective coatings and paints,[7] and gas sensors.[8] Many other technologically important materials have the rutile and similar structures.[9] In these materials, as in Si, it is essential to understand grain boundaries on the atomic level, where the electronic and vibrational states are determined and to relate this structure to the various physical properties.

From a fundamental perspective, TiO_2 has become of great importance because well characterized bicrystals of it can be readily grown and studied by a variety of powerful experimental techniques such as *Z*-contrast scanning transmission electron microscopy (STEM), electron energy loss spectroscopy (EELS), and high resolution electron microscopy (HREM). Unfortunately, on the theoretical side, TiO_2, like all oxides, is difficult to work with because of the O ions whose electronic structure is quite sensitive to the crystalline environment (the O^{2-} ion is unstable in free space). This problem becomes potentially severe near surfaces and interfaces where the crystalline environment may differ greatly from that of the bulk material. It was found by Dahman *et al.*[10] using HREM that the lattice around the Σ5 (210) TiO_2 grain boundary is *contracted* through a rigid-body shift by a remarkably large amount (~1 Å). This result has defied theoretical explanation in terms of either *ab initio* or semiempirical calculations, leading Dahman *et al.* to attribute it to defects (a loss of oxygen) and/or nonstochiometry (excess Ti with a +3 valence). But the subsequent EELS measurements of Wallis *et al.*[11] gave details about the oxidation states at cation sites and the atomic structure around anion sites from multiple scattering theory. It was found that to a large degree the stochiometry and ionic charges at the boundary are preserved at their bulk values. Further, the *Z*-contrast image of the grain boundary in Ref. 11 agrees with the HREM results[10] in that a large rigid-body contraction is indicated. In addition, it is clear from the *Z*-contrast image that some of the projected Ti-Ti distances across the grain boundary are significantly contracted (11–18 %) compared to the bulk. In the work described here, we explore the structure at the grain boundary through very extensive *ab initio* calculations, with particular emphasis on the relative stabilities of grain boundary structures that include rigid-body shifts of various magnitudes.

Structural calculations have generally proceeded by two approaches. One uses semiempirical methods such as the shell model developed by Dick and Overhauser[12] for the calculation of phonon dispersion curves, while the other uses first principles calculations, perhaps augmented by semiempirical considerations.[13] In this work only first principles calculations are used. It should be noted that a paper by Dawson *et al.*[14] employed techniques quite similar to those used here to study the same grain boundary in TiO_2 but with a few significant differences as discussed below.

Our calculations, like many others, are based on density functional theory using the local density approximation combined with nonlocal, norm-conserving pseudopotentials and plane wave expansions.[15] The calculations used the band-by-band conjugate gradient technique to minimize the total energy with respect to plane-wave coefficients. Pseudopotentials of the Kleinman-Bylander[16] representation were generated in real space using a well established optimization scheme. The O potential was generated using the reference atomic configuration $2s^2 2p^4$ for the *s* and the *p* angular momentum components. For the *d* component, $2s^1 2p^{1.75} 3d^{0.25}$ was used. For all three components, the core radii were 1.8 a.u. The Ti potential used $3d^2 4s^2$ for the *s* and *d* components and $3d^2 4s^{0.75} 4p^{0.25}$ for the *p* component. Core radii of 2.5 a.u. were used for all three components. Brillouin-zone sampling used the lowest-order set of *k* points.[17] After extensive experimentation, a plane-wave cutoff of 1000 eV was used.

 ©2000 The American Physical Society

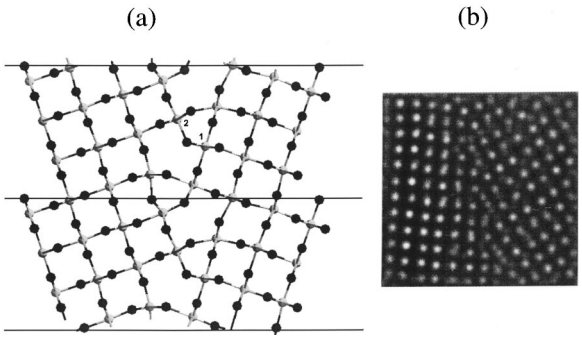


FIG. 1. The Σ5 (210) grain boundary in TiO_2: (a) sixty-six-atom unit cell used in the preliminary calculations where the lines indicate the unit cell boundaries and (b) *Z*-contrast image.

The energy per unit cell was converged to 0.001 eV/atom (0.066 eV/unit cell) and residual forces on the ions were less than 0.1 eV/Å.

Before the grain boundary calculations, a bulk TiO_2 unit cell was allowed to relax and the results are as follows:

Parameter	Theory (this work)	Theory (Ref. 14)	Experiment (Ref. 18)
a (Å)	4.603	4.603	4.594
c (Å)	2.976	2.976	2.956
Ti-O (short) (Å)	1.953	1.961	1.949
Ti-O (long) (Å)	1.989	1.978	1.980

It should be noted that the calculations in Ref. 14 used a 500 eV cutoff rather than the 1000 eV cutoff used in this work.

Figure 1(a) shows the grain boundary structural model used in the initial calculations. The large unit cell of these calculations is enclosed by the solid lines. The grain boundary structure was generated by matching the two halves of the bicrystal at one of the O sites in the bulk unit cell, carrying out the necessary rotations to get the Σ5 coincident site lattice and removing all ions that were within 0.5 Å of one another. The large unit cell contains sixty-six ions and is electrically neutral for formal valences of +4 and −2 for Ti and O, respectively. We also constructed structural models by taking the match points at the Ti ions, but this generated only symmetric boundaries that did not agree with the microscopy results unless an additional translation along the boundary was introduced, as others[10] have noted.

It is seen that our unit cell does not contain two grain boundaries of opposite orientation as in Ref. 14. Instead we include several ''layers of vacuum'' before repeating the structure to form a periodic slab configuration. While this approach introduces surface states, it allows more atoms per grain boundary unit cell to be treated and eliminates the interaction of close-lying, oppositely directed grain boundaries. Numerous calculations were made to study the effects of varying the vacuum layer thickness. As the bicrystal was expanded and contracted, the number of vacuum layers also increased and decreased such that the length of the bicrystal minus the length of the vacuum layers is constant at 5.5 Å. This was done to satisfy long-range interactions of the crystal atoms and at the same time prevent orbital overlap at the surface edges. This approach will be discussed in detail in a later paper.

Initially, the connectivity in the *c* direction was the same as in the bulk but subsequently the two sides of the grain boundary were displaced relative to one another along *c*, as described below. The ions on the outer extremities of the slab were fixed and all others were allowed to relax. The very lengthy calculations (carried out on an Intel Paragon supercomputer) with this structure gave results so similar to those of Dawson *et al.* that they will not be discussed further here. While the grain boundary structure is quite similar to the experimental images, reproduced in Fig. 1(b), the large inward rigid-body shift could not be obtained. Although these calculations, like those of Ref. 14, were unsuccessful in this respect, they provided some important insights. It became clear that only those ions in the immediate neighborhood of the grain boundary plane undergo significant displacements during relaxation. In contrast, other ions near the grain boundary plane moved only slightly and still more distant ones hardly at all. This implied that some other approach would have to be taken if the large observed contraction was to be obtained.

The results of Ref. 10 suggest that one side of the grain boundary is simply rigidly displaced toward the other while maintaining essentially the bulk distances within each half of the bicrystal. Thus, all atoms on one side of the bicrystal must undergo large displacements relative to those on the other side. Obviously, then one cannot fix one or more layers of ions on both sides of the grain boundary, as is sometimes done, and expect to find the observed result. To implement this rigid displacement constraint in our calculations, the number of atoms in the unit cell was reduced to forty-eight. This simplification, which appears well justified in view of the results for the sixty-six ion calculation, made it possible, though computationally still extremely intensive, to carry out the series of calculations discussed next.

It should be emphasized that, with the exception of the rigid-body displacements of the two sides of the grain boundary relative to one another, the atoms at the slab/vacuum surfaces are held rigid throughout the calculations. This is because the slab is cut out of, and is meant to mimic, a bulk TiO_2 material. The surface states are the same for each structure considered and it is assumed that their effect on the *relative stabilities* should be negligible. Some evidence for this is that the relaxation of grain boundary core atoms in the sixty-six atom unit cell (with more rigid ''bulklike'' atoms between the grain boundary core and the surface of the slab) was approximately the same as the relaxation of grain boundary core atoms in the comparable forty-eight atom unit cell (with fewer ''bulklike'' atoms between the grain boundary core and the surface of the slab).

Figure 2 summarizes the results of the calculations. Curve 1 shows the total energy for the forty-eight ion configuration discussed above and shown in Fig. 3 as a function of the rigid projected displacement, x_p, of Ti atoms *at the edges of the slab* towards and away from one another. The atoms at the edge are shifted in such a way that their relative position in the *c* direction is maintained at the level given in Fig. 1(a). This means that the projected distance observed by TEM is shortened while keeping the +4 Ti atoms fairly far apart since they are on different levels in the *c* direction. The atoms that were allowed to relax are the twenty-seven inside and at the perimeter of the grain boundary, indicated by the

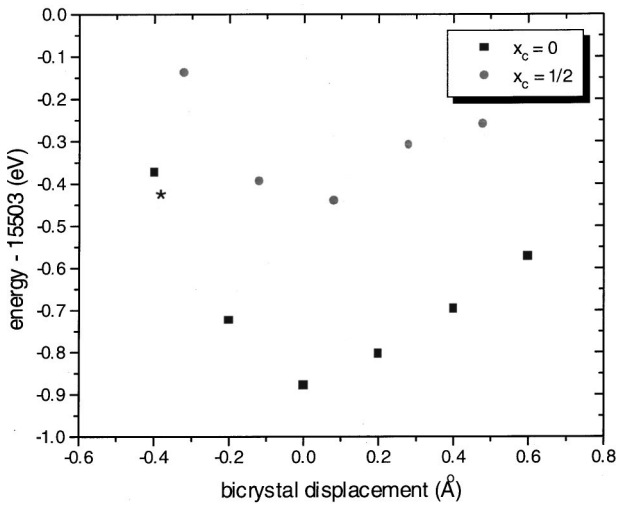


FIG. 2. Relative stability of grain boundary structures as the two sides of the boundary are shifted relative to one another. The two curves represent structures that are shifted relative to one another in the *c* direction as discussed in the text. The asterisk indicates a contraction comparable to that found experimentally.

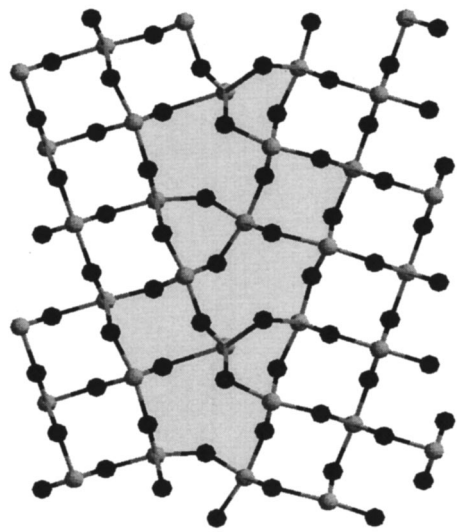


FIG. 3. The minimum-energy structure for the forty-eight-atom unit cell where $x_c = 0$. The shaded areas mark the grain boundary core where atoms were allowed to relax. Two unit cells are shown for clarity.

shading in Fig. 3. In addition to the rigid-body displacement, we calculated the relaxed projected Ti-Ti distance across the grain boundary. The results show no rigid-body contraction of the bicrystal but there is a large contraction of the Ti-Ti projected distance across the grain boundary of 12% (0.4 Å) that does agree with the *Z*-contrast image obtained by Wallis *et al.* However, the calculations also give a projected Ti-Ti distance that is 10% (0.33 Å) greater than the bulk on one side of the grain boundary region not seen in the *Z*-contrast image. These expansions and contractions effectively cancel each other out.

In the second series of calculations, the two sides of the grain boundary were displaced relative to one another by $c/2$. This is equivalent to interchanging the up-down positions of all ions in the bulk unit cells as the grain boundary is crossed. It also puts the 1 and 2 Ti atoms shown in Fig. 1(a) on the same level in the *c*-direction. Thus, in this case the actual and projected Ti-Ti distances are the same. The results are shown in curve 2 on Fig. 2. The total energy is ~0.44 eV (0.009 eV/atom) higher than that of the preceding calculations and the minimum again occurs at a position where there is little or no rigid-body contraction of the bicrystal. A smaller contraction of the Ti-Ti projected distances across the grain boundary relative to the bulk of 4.3% (0.14 Å) is predicted for this structure.

The results of these calculations argue strongly that the experimentally observed rigid-body contractions must be due to some phenomena not considered in these calculations, such as half-filled columns, vacancies, or impurities that are present at this grain boundary and not detected in the EELS work of Wallis *et al.*[10] Whatever the origin of the contraction, the implication for the computational side is disturbing in that long series of calculations for even more complex structures will have to be considered, requiring greatly increased computational power. However, for the experimental side, the implications may be even more significant. Clearly, the whole question of the growth of bicrystals, which is still in its infancy, becomes crucial. How can it be established that an ''intrinsic'' grain boundary, devoid of impurities, vacancies, half columns, etc. has been obtained under given growth conditions? The crucial need for microscopes with even higher resolution that are able to detect all the ionic species in a given sample is evident. Recent progress in correcting the large spherical aberration of present objective lenses promises to make such microscopes a reality.[19,20]

The authors thank G. Mahan, D. Wallis, Y. Yan, S. Pantelides, and E. C. Dickey for many helpful discussions. This work was supported in part by the Basic Energy Sciences Division, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464. S.B.S. also acknowledges partial support from the National Science Foundation (DMR-9976851). This work was started while S.B.S. was a U.S. Department of Energy Faculty Research Participant at Oak Ridge National Laboratory through the Laboratory Co-op Faculty Research Participation program administered by the Oak Ridge Institute of Science and Education.

*Author to whom all correspondence should be addressed.

[1] S. J. Tauster, Acc. Chem. Res. **20**, 389 (1987).
[2] R. L. Kurtz and V. E. Henrich, Phys. Rev. B **28**, 6699 (1983).
[3] S. Nakabayashi, A. Fujishima, and K. Honda, J. Am. Chem. Soc. **107**, 250 (1985).
[4] C. Bechinger, S. Ferrere, A. Zaban, J. Sprague, and B. A. Gregg, Nature (London) **383**, 608 (1996).
[5] A. Harstrick, O. M. Kut, and E. Heinzle, Environ. Sci. Technol. **30**, 817 (1996).
[6] B. O'Regan and M. Gratzel, Nature (London) **353**, 737 (1991).
[7] K. Akiyama, N. Toyama, and K. Muraoka, J. Am. Ceram. Soc. **81**, 1071 (1998).
[8] *Current Topics in Materials Science*, edited by E. Kaldis (North-Holland, Amsterdam, 1981), Chap. 3, pp. 148–482.

[9] See, for example, *The Surface Science of Metal Oxides*, edited by V. E. Henrich and P. A. Cox (Cambridge University Press, Cambridge, 1996).
[10] U. Dahmen, S. Paciornik, I. G. Solorzano, and J. B. Van der Sande, Interface Sci. **2**, 125 (1994).
[11] D. J. Wallis, N. D. Browning, P. D. Nellist, S. J. Pennycook, I. Majid, Y. Liu, and J. B. Vander Sande, J. Am. Ceram. Soc. **80**, 499 (1997).
[12] B. G. Dick and A. W. Overhauser, Phys. Rev. **112**, 603 (1958).
[13] J. Purton, D. W. Bullett, P. M. Oliver, and S. C. Parker, Surf. Sci. **336**, 166 (1995).
[14] I. Dawson, P. D. Bistrowe, M.-H. Lee, M. C. Payne, M. D. Segall, and J. A. White, Phys. Rev. B **54**, 13 727 (1996).
[15] M. C. Payne, M. P. Teter, D. C. Allen, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[16] M.-H. Lee, J. S. Lin, M. C. Payne, V. Heine, V. Milman, and S. Crampin (unpublished).
[17] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[18] K. M. Glassford and J. R. Chelikowsky, Phys. Rev. B **46**, 1284 (1992).
[19] M. Haider *et al.*, Nature (London) **392**, 768 (1998).
[20] O. L. Krivanek, N. Delby, and A. Lupini, Ultramicroscopy **78**, 1 (1999).

Explanation of the Observed Dearth of Three-Coordinated Al on γ-Alumina Surfaces

Karl Sohlberg,*,† Stephen J. Pennycook,†,‡ and Sokrates T. Pantelides†,‡

Contribution from the Solid State Division, P.O. Box 2008, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6031, and Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37235

Received July 26, 1999

Abstract: NMR data have revealed that γ-alumina surfaces do not contain any three-coordinated Al atoms, despite the fact that their presence is expected on the basis of the bulk structure of the material. We report density-functional calculations of γ-alumina (110) surfaces that reveal a massive spontaneous reconstruction of the ideally terminated bulk. The three-coordinated Al atoms initially in the surface layer drop into vacant octahedral interstices in the second layer where they become six-coordinated. The predicted reconstruction is also fully consistent with IR studies of hydroxylated γ-alumina surfaces and reveals the physical origin of results obtained by earlier computational simulations.

1. Introduction

The aluminas are among the technologically most important catalytic materials, serving both as catalytic supports[1] and as catalysts in their own right.[2] They are employed in applications from petroleum refining[2] to automotive emissions control.[3] Accordingly, there has been prolonged widespread interest in the structure and function of alumina surfaces.[4] Of particular importance is the distribution of surface Al atoms and their coordination states, since coordinatively unsaturated surface Al atoms are the source of the surface Lewis acidity, which is central to catalysis.[5] One of the most powerful techniques for elucidating the distribution of Al atoms of various coordination on the surface is ^{27}Al cross-polarization NMR spectroscopy.[6] Meticulous studies of γ-alumina with this technique have produced a perplexing result.[5] No three-coordinated Al is observed, despite the fact that the presence of three-coordinated Al on the surface would be expected based on the bulk structure. This result is particularly important because the presence of three-coordinated Al has often been assumed for the purposes of assigning IR bands of the hydroxylated surface.[2,4,7,8] We have resolved this conundrum with density-functional-theory (DFT) calculations. We find that when the (110C) layer of γ-alumina is exposed, it undergoes a massive spontaneous reconstruction. The three-coordinated Al atoms, initially in the surface layer, drop from the surface layer into vacant octahedral interstices in the next lower layer.

2. Computational Method

The density functional theory[9] calculations employed the generalized gradient approximation (GGA) to the exchange-correlation energy,[10] as described in the review by Payne et al.[11] The electron−ion interactions were described with nonlocal reciprocal-space pseudopotentials[12] in the Kleinman-Bylander form.[13] A plane wave basis was used to describe the electronic density with a cutoff energy of 1000 eV as determined by convergence studies. Integrations over the Brillouin zone employed a grid of k points with a spacing of 0.1 Å^{-1} chosen according to the Monkhorst-Pack scheme.[14] The structural relaxation studies employed a 56-atom unit cell consisting of a slab of γ -alumina 4 layers thick. The atoms in the bottom layer were frozen, as were the dimensions (a, b, c, α, β, γ) of the unit cell. The system was infinitely periodic with a vacuum spacing between slabs of 10 Å. The starting structure from which the slab was generated was that of fully relaxed γ-alumina in its lowest energy form.[15]

3. Results

γ-Alumina is derived from the thermal dehydration of boehmite.[16] Its structure is very closely related to that of Mg-spinel,[17] which has the primitive unit cell $Mg_2\ Al_4\ O_8$. In Mg-spinel, the oxygen atoms are cubic close-packed, the Mg atoms are tetrahedrally (T_d) coordinated by oxygen, and the Al atoms

† Oak Ridge National Laboratory.
‡Vanderbilt University.

(1) Satterfield, C. N. *Heterogeneous Catalysis in Practice*; McGraw-Hill: New York, 1980; Section 4.5.
(2) Knözinger, H.; Ratnasamy, P. *Catal. Rev.—Sci. Eng.* **1978**, *17*, 31.
(3) Taylor, K. C. *Catal. Rev.—Sci. Eng.* **1993**, *35*, 457.
(4) Tsyganenko, A. A.; Mardilovich, P. P. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 4843.
(5) Coster, D.; Blumenfeld, A. L.; Fripiat, J. J. *J. Phys. Chem.* **1994**, *98*, 6201.
(6) Morris, H. D.; Ellis, P. D. *J. Am. Chem. Soc.* **1989**, *111*, 6045.
(7) Liu, X.; Truitt, R. E. *J. Am. Chem. Soc.* **1997**, *119*, 9856.
(8) Tsyganenko, A. A.; Filimonov, V. N. *J. Mol. Struct.* **1973**, *19*, 579.
(9) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140A*, 1133.
(10) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(11) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045.
(12) Al pseudopotential: Winkler, B.; Milman, V.; Hennion, B.; Payne, M. C.; Lee, M. H.; Lin, J. S. *Phys. Chem. Miner.* **1995**, *22*, 461. H pseudopotential: CASTEP default reciprocal-space. O pseudopotential: Winkler, B.; Milman, V.; Hennion, B.; Payne, M. C.; Lee, M. H.; Lin, J. S. *Phys. Chem. Miner.* **1995**, *22*, 461.
(13) Kleinman, L.; Bylander, D. M. *Phys Rev. Lett.* **1982**, *48*, 1425.
(14) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B: Solid State* **1976**, *13*, 5188.
(15) Sohlberg, K.; Pennycook, S. J.; Pantelides, S. T. *J. Am. Chem. Soc.* **1999**, *121*, 7493.
(16) Wefers, K.; Misra, C. *Oxides and Hydroxides of Aluminum*; Alcoa, 1987.
(17) Wyckoff, R. W. G. *Crystal Structures*; Interscience Publishers: New York, NY, 1963.

10.1021/ja9926358 CCC: $18.00 © 1999 American Chemical Society
Published on Web 11/04/1999

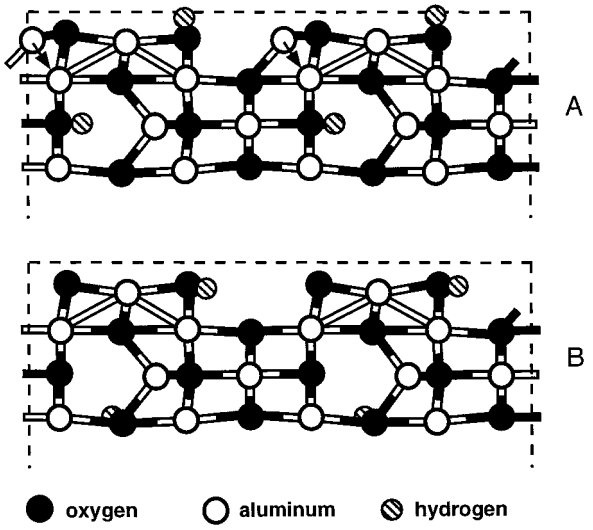


Figure 1. Edge-on view of the top four layers of γ-alumina before (a) and after (b) relaxation. Note that the three-coordinated Al are absorbed into the first subsurface layer. The bottom layer of atoms was frozen during the calculation.

are octahedrally (O_h) coordinated by oxygen. Recently, we have shown that γ-alumina has a range of valid stoichiometries $H_{3m}Al_{2-m}O_3$, ($0 \leq m \leq 1/3$) but the lowest energy form has the stoichiometry of a hydrogen−aluminum spinel.[15] The primitive unit cell of the lowest energy form is HAl_5O_8, where the H atom and one Al atom occupy (nominally) the Mg sites in the spinel structure, and the remaining 4 Al atoms occupy the Al sites in the spinel structure. In HAl_5O_8, 20% of the Al atoms (1 in 5) are tetrahedrally coordinated by oxygen and 80% (4 in 5) of the Al atoms are octahedrally coordinated by oxygen. It is important to note, however, that throughout the range of valid stoichiometries ($0 \leq m \leq 1/3$), including the fully dehydrated limit ($m = 0$), the structure contains Al in both tetrahedral and octahedral coordination. In the present work, calculations were performed for the HAl_5O_8 stoichiometry. We will show, however, that our main results have general validity.

Although early surface studies focused on the (100) face,[18] it is now known that the (110) face of γ-alumina is preferentially exposed.[2] For example, Z-contrast scanning transmission electron microscopy (Z-STEM) images of γ-alumina, in application as an automotive catalytic support, showed the (110) surfaces to be exposed.[19] There are, however, two (110) layers in γ-alumina, denoted (110C) and (110D).[2] We carried out structural relaxations of these two surfaces and found that it is energetically preferred to expose the (110C) layer. Structural relaxation of the (110C) surface of γ-alumina also showed that it undergoes an interesting and important reconstruction. The unrelaxed (110C) layer exposes both three-coordinated Al atoms (Al in T_d coordination with one oxygen removed) and four-coordinated Al atoms (Al in O_h coordination with two oxygen atoms removed). When this structure is relaxed, the three-coordinated Al atoms drop from the surface layer into vacant octahedral interstices in the next lower layer. An edge-on view of this reconstruction is shown in Figure 1. In addition to the large physical displacement exhibited in Figure 1, the reconstruction entails a large energy gain, about 10 eV for the unit cell employed here, which translates into 2 $J \cdot m^{-2}$.

(18) Peri, J. B. *J. Phys. Chem.* **1965**, *69*, 220. Peri, J. B. *J. Phys. Chem.* **1965**, *69*, 231.

(19) Nellist, P. D.; Pennycook, S. J. *Science* (Washington, D.C.) **1996**, *274*, 413.

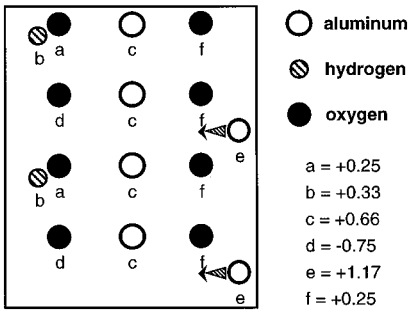


Figure 2. Top-down view of the unrelaxed (110C) surface of γ-alumina. Note that the charge on the three-coordinated Al is +1.17 prior to reconstruction. After reconstruction the three-coordinated Al atoms drop from the surface layer into the first subsurface layer, after which their charge is zero. The observed reconstruction thereby eliminates the electrostatically most unstable species from the surface.

The reconstruction can be rationalized by observing the electrostatic distribution. We use the following definition of effective charge introduced by Knözinger and Ratnasamy,[2] based on the Pauling electrostatic valence rules:[20] (*atom charge*) = (*preferred valence of atom*) + (*sum of strengths of bonds to atom*). Here, (*bond strength*) = (*preferred charge of adjacent atom*)/ (*coordination of adjacent atom*). Figure 2 shows the unrelaxed (110C) surface of γ-alumina along with charges assigned as above. The arrows indicate the motion of the three-coordinated Al atoms upon structural relaxation. Note that, prior to relaxation, the three-coordinated Al atoms are assigned a charge of 1.17, the largest (in magnitude) of any surface atom. After relaxation, however, these Al atoms are six-coordinated by oxygen. Since every oxygen atom in the bulk is four-coordinated, the new charge on the Al atom after reconstruction is, by the above scheme, $(+3) + 6(-2/4) = 0$. The reconstruction therefore completely eliminates the electrostatically most unstable surface species, three-coordinated Al.

The predicted reconstruction is consistent with results obtained by earlier theoretical studies of γ-alumina surfaces. Alvarez et al.[21] carried out molecular dynamics (MD) studies of γ-alumina surfaces using empirical potentials. In MD simulations starting from truncated bulk surfaces, they reported the loss of Al from tetrahedral cation positions on the γ-alumina (110C) surface. More recently, similar behavior was reported by Gunji et al.[22] While no mechanism for the observed depletion was proposed, it is now clear that their observed result is a manifestation of the spontaneous reconstruction reported here.

4. Discussion

The surface reconstruction reported here provides a natural explanation for the perplexing experimental result[7] mentioned in the Introduction. As already noted, the (110C) layer contains both four-coordinated Al and six-coordinated Al. When exposed, one would therefore expect both three-coordinated and four-coordinated Al atoms on the surface as shown in Figure 3. Cross-polarizing ^{27}Al NMR studies,[6] however, found no observable three-coordinated Al.[5] This mystery is explained by our surface-reconstruction studies. The three-coordinated Al atoms drop from the surface layer into vacant octahedral interstices in the next lower layer.

(20) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960.

(21) Alvarez, L. J.; Sanz, J. F.; Capitán, M. J.; Centeno, M. A.; Odriozola, J. A. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 3623.

(22) Gunji, I.; Teraishi, K.; Endou, A.; Miura, R.; Yin, X. L.; Yamauchi, R.; Kubo, M.; Chatterjee, A.; Miyamoto, A. *Appl. Surf. Sci.* **1998**, *132*, 549.

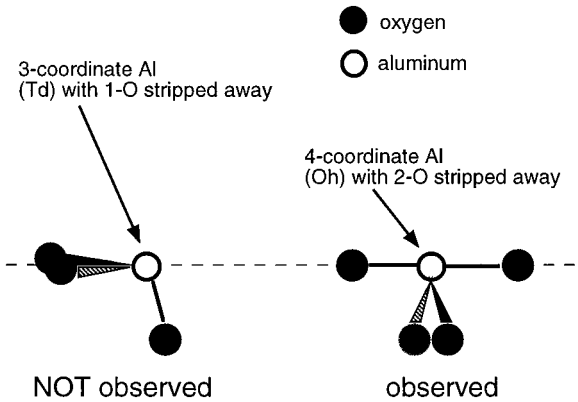


Figure 3. Two types of Al expected on the surface of γ-alumina when the (110C) layer is exposed. ^{27}NMR studies find four-coordinated Al, but not three-coordinated Al,[5] a quandary resolved by the reconstruction reported here.

Further confirmation of the surface reconstruction reported here comes from IR studies of hydroxylated γ-alumina surfaces. Dry γ-alumina is well-known to chemisorb water.[18] This chemisorption leads to the formation of surface hydroxyl groups that have been the subject of numerous studies by IR spectroscopy.[2,4,7,8,18,23] These studies clearly document that there are three prominent bands in the IR spectrum of surface OH groups of γ-alumina. Tsyganenko and Filimonov[8] appear to be the first to have proposed that these three bands arise from three basic types of chemisorbed OH. These three types, denoted I, II, and III, correspond to the O atom of the OH group in question being bound to one, two, or three surface Al atom(s), respectively. Later, Knözinger and Ratnasamy[2] subdivided types I and II into I_a, I_b, II_a, and II_b, where the subscript *a* denotes that (one of) the bonding Al atom(s) is itself in tetrahedral coordination, and the subscript *b* indicates that all of the bonding Al atoms are in octahedral coordination. More recently, Tsyganenko and Mardilovich[4] have replaced *a* and *b* with integer subscripts to indicate the exact coordination of each of the Al atoms that are bonding to the chemisorbed OH. Each digit of the integer subscript denotes the coordination of one of the Al atoms bonding to O of the chemisorbed OH. (O from the chemisorbed OH is included in this count.) Equating the designations of Knözinger and Ratnasamy with this most recent system we see that $I_a = I_4$, $I_b = I_6$, $II_a = II_{64}$, $II_b = II_{66}$, etc. Liu and Truitt[7] have added a few minor refinements to this model for the type III case and give a complete review of the incremental improvements of the model.

In the most modern assignment, the three major bands are assigned to I_6, II_{66}, and III, with two other minor bands being assigned to hydrogen-bonded OH and/or various type III groups.[7] Conspicuous by their absence are OH groups of type I_4. If the three-coordinated Al shown in Figure 3 were present on the γ-alumina surface, upon hydroxylation one would expect type I_4 groups to form. This is particularly true in light of recent ab initio results of Tachikawa and Tsuchida.[24] The ab initio studies showed that surface Al in tetrahedral sites is a stronger Lewis acid than surface Al in octahedral sites. Type I_4 OH groups should therefore form even more easily than those of type I_6. The fact that the latter give rise to one of the three prominent bands in the IR spectrum of surface OH, and the former are not observed, is completely consistent with our surface-reconstruction studies, which find that three-coordinated Al atoms at the (110C) surface of truncated bulk γ-alumina are depleted by a spontaneous reconstruction.

Finally, we address stoichiometries other than HAl_5O_8, namely $H_{3m}Al_{2-m}O_3$, with $m \neq 1/8$.[15] These materials also have a spinel structure. For $m < 1/8$, a fraction of the cation sites are vacant, thus a (110C) surface occasionally has a cation vacancy. The existence of such cation vacancies has no bearing on the mechanism we report. If a three-coordinated Al (see Figure 3) is absent, obviously there is no Al to reconstruct. If a four-coordinated Al (see Figure 3) is absent, it does not impact the reported reconstruction since three-coordinated Al and four-coordinated Al are not adjacent on γ-alumina (110C) surfaces. For $m > 1/8$, there is excess interstitial hydrogen in the bulk. Once more, this has no bearing on our results regarding the relaxation of surface Al atoms. These observations are consistent with those of Tsyganenko and Mardilovich,[4] who noted that different hydrogen contents do *not* lead to any new structures on the hydroxylated surface.

5. Conclusions

We have carried out DFT studies of the energetically preferred (110C) surface of γ-alumina. It is found that that when the (110C) layer is exposed, it undergoes a massive spontaneous reconstruction. The three-coordinated Al atoms initially in the surface layer drop into vacant octahedral interstices in the first subsurface layer. In addition to the large physical displacement of the initially three-coordinated Al, the reconstruction liberates an energy of 2 $J \cdot m^{-2}$. This spontaneous reconstruction of the truncated bulk explains the perplexing experimental observation that no three-coordinated Al is observed on γ-alumina surfaces[5] despite the fact that their presence should be expected on the basis of simple truncation of the bulk at the (110C) surface, the preferentially observed surface. The predicted reconstruction is also fully consistent with IR studies of hydroxylated surfaces and gives new insight into earlier computational simulations.

Acknowledgment. This work was supported by the US DOE, Office of Basic Energy Sciences under contract no. DE-AC05-96OR22464 with the Lockheed-Martin Energy Research Corp., by the NSF Grant DMR-9803768, by the William A. and Nancy F. McMinn Endowment at Vanderbilt University, and by financial support from Dupont through an ATE grant. Computations were partially supported by the National Center for Supercomputing Applications under Grant CHE980043N to K.S. and for them the SGI Origin2000 at the University of Illinois was utilized. K.S. was supported by an appointment to the ORNL Postdoctoral Research Program administered jointly by ORNL and Oak Ridge Associated Universities.

JA9926358

(23) Okamoto, Y.; Imanaka, T. *J. Phys. Chem.* **1988**, *92*, 7102.
(24) Tachikawa, H.; Tsuchida, T. *J. Mol. Struct.* **1995**, *96*, 277.

Direct experimental observation of the local electronic structure at threading dislocations in metalorganic vapor phase epitaxy grown wurtzite GaN thin films

Y. Xin, E. M. James, I. Arslan, S. Sivananthan, and N. D. Browning[a)]
Department of Physics, University of Illinois at Chicago, Chicago, Illinois 60607-7059

S. J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6031

F. Omnès, B. Beaumont, J-P. Faurie, and P. Gibart
CRHEA-CRNS, rue Bernard Gregory, 06560 Valbonne, France

(Received 14 October 1999; accepted for publication 24 November 1999)

The electronic structure of pure edge threading dislocations in metalorganic vapor phase epitaxy grown wurtzite GaN thin films has been studied directly by atomic resolution *Z*-contrast imaging and electron energy loss spectroscopy in a scanning transmission electron microscope. Dislocation cores in *n*-type samples grown in N-rich conditions show no evidence for the high concentration of Ga vacancies predicted by previous theoretical calculations. Nitrogen *K*-edge spectra collected from edge dislocation cores show a sudden and significant increase in the intensity of the first fine-structure peak immediately above the edge onset compared to the bulk spectra. The origin of this increase is discussed. © *2000 American Institute of Physics.* [S0003-6951(00)03404-5]

The benign effect of high density of threading dislocations[1] on GaN high efficiency light emitting diodes has led to considerable interest in determining their fundamental properties. Although early work[1,2] suggested that threading dislocations are not nonradiative recombination centers, the increase in optical properties that is achieved[3] by reducing the threading dislocation density in the epitaxial lateral overgrowth method,[4] provides circumstantial evidence that dislocations do have a deleterious effect. Furthermore, recent more direct evidence indicates that threading dislocations *can* be optically and electrically active. In particular, atomic force microscopy (AFM) combined with cathodoluminescence (CL)[5] and plan-view transmission electron microscopy (TEM) combined with CL[6] clearly show threading dislocations to be related to dark spots in band edge emission CL images. This effect is consistent with theoretical calculations suggesting that dislocations should be charged depending on doping and growth conditions.[7,8] The presence of charged dislocations is further supported by low transverse mobility measurements,[9] near surface electrical properties observed by AFM and scanning capacitance microscopy,[10] and selective photoelectrochemical etching.[11] These observations have been used in calculations of dislocation scattering.[12] However, our detailed experimental results show that the density of acceptor states must be less than is assumed in these calculations.

Our experiments used both a 200 kV JEOI 2010F and a 300 kV VG HB603U field emission microscope, which have point resolutions of ~0.14 nm operating in scanning model.[13,14] This experimental setup allows atomic resolution imaging and spectroscopy to be obtained. The incoherent ''*Z*-contrast'' image[15] of a threading dislocation along the *c* axis is used to determine the core structure and position the electron probe for electron energy loss spectroscopy (EELS).[16] Core loss EELS probes the unoccupied density of states near the conduction band minimum and is exactly analogous to near edge x-ray absorption spectroscopy (XAS), but with a much higher spatial resolution afforded by the microscope. Here, EELS spectra were acquired directly from the dislocation core and compared with bulk spectra taken from no more than 2 nm away. The sample under investigation in this work was grown on a sapphire substrate by low-pressure metalorganic vapor phase epitaxy under a nitrogen rich (high V/III ratio) growth condition and doped with Si at a level of $\sim 2\times 10^{18}\,\mathrm{cm}^{-3}$. The free carrier mobility is about 200 $\mathrm{cm}^2/\mathrm{V\,s}$. The threading dislocation density was determined by TEM to be $\sim 3\times 10^{11}\,\mathrm{cm}^{-2}$, with the majority being pure edge dislocations.

Pure edge dislocations can be distinguished from those with a screw component by differences in the strain contrast in TEM images.[17] Figure 1(a) shows a two beam bright field image of the sample tilted close to the zone axis $[1\bar{1}02]$ and imaged with $\mathbf{g}=(1\bar{1}0\bar{1})$. Both pure edge and mixed type dislocations are visible in this image, but they have markedly different diffraction contrast. The mixed type exhibits strong contrast at both ends, as is indicated by the black arrows in Fig. 1(a), while the pure edge dislocations show no such contrast (as is indicated by a white arrow). This same lack of strain contrast is also observed in the *Z*-contrast image which means that by choosing only the dislocations that exhibit the smallest contrast in the end-on *Z*-contrast images [Fig. 1(b)], we are able to select pure edge dislocations.

The core structure of such a dislocation is shown in Fig. 2(a), and clearly shows the eight-fold ring structure, as previously reported.[13] On the basis of the prior theoretical calculations, this *n*-type sample grown under N-rich growth conditions is expected to show the Ga-vacancy core structure. Since atoms contribute to the image according to their

[a)]Electronic mail: browning@uic.edu

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

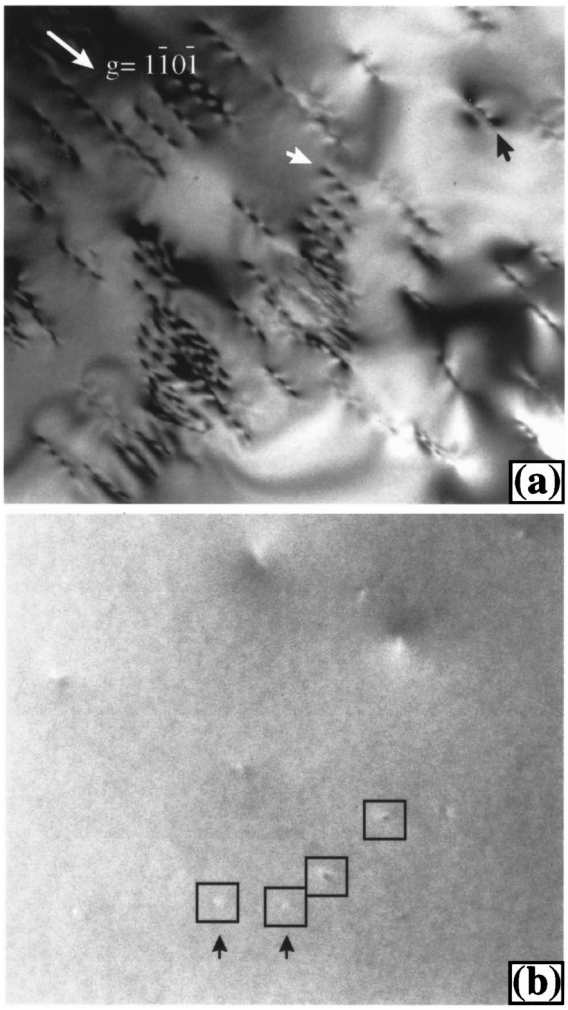


FIG. 1. (a) Two beam bright field image of a plan view sample tilted close to the [1–102] zone axis showing the different diffraction contrast from a pure edge dislocation (indicated by a white arrow) and a mixed type dislocation (indicated by a black arrow). (b) *Z*-contrast image obtained at the exact [0001] zone axis. The end-on dislocations are indicated for clarity.

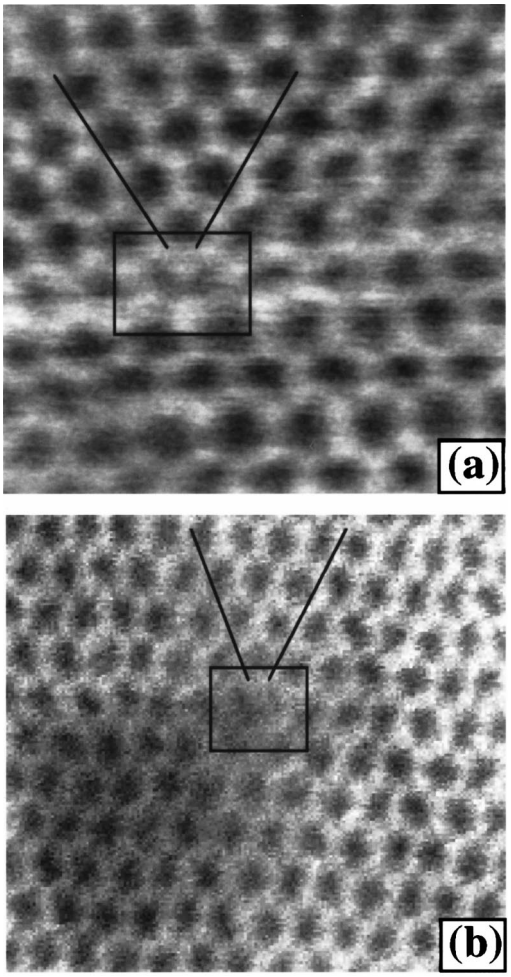


FIG. 2. *Z*-contrast images of pure edge dislocations obtained with (a) the 0.14 nm probe in the VG Microscopes HB603 showing the presence of a core column and (b) the 0.2 nm beam in the JEOL 2010F used for EELS.

mean square atomic number (Z), the light N atoms contribute only 5% of the image contrast in a full GaN atomic column. Thus, it is immediately apparent from the image that the Ga vacancy concentration must be much less than 100%. At the typical thicknesses used for *Z*-contrast imaging, the image intensity is roughly proportional to the number of atoms in a column.[15] Therefore, our sensitivity to Ga vacancies is limited just by the statistical varyiations in column intensities close to the dislocation core. Integrated columnar intensities were obtained from a maximum entropy analysis[13] and showed a mean deviation of 15%. From this data we therefore estimate Ga vacancy concentration in the range 0%–15% per *c*-axis repeat, implying a significantly lower line charge than previously assumed for threading dislocations.

In order to investigate the local electronic structure of the dislocations, we chose to use the nitrogen *K*-edge spectrum. This absorption edge gives information on transitions from the $1s$ core level to the unoccupied density of states in the conduction band with $2p$ symmetry. Figure 2(b) shows the dislocation image obtained from the 2010F with the larger beam diameter used for EELS, and Fig. 3(a) shows experimental nitrogen *K*-edge spectra obtained from several pure edge dislocation cores as well as nearby perfect regions of the film for comparison. Each spectrum is the summation of 13 spectra from 8 dislocation cores with an acquisition time of 5 s for each individual spectrum. The energy resolution is 1.2 eV (as measured from the full width at half maximum of the zero loss peak). Other than a power law background subtraction,[18] the spectra are as acquired.

It can be clearly seen that there are dramatic changes in the fine structure of the nitrogen *K*-edge spectra on and off the dislocation core. In particular, in the core spectrum, the first peak 2.2 eV above edge onset (398 eV) rises significantly relative to the second peak. It should be noted that the fine structure changes observed here are significantly different from the effects of beam damage, which can therefore be excluded.[19] In addition, the change in fine structure is only observed if we position the probe exactly on the core (a misplacement of the probe by as little as 0.6 nm means that the spectrum does not show the increase in the first peak). Simulations of the electronic structure using the multiple scattering methodology of the FEFF eight codes[20] are shown in Fig. 3(b). These simulations take account of the collection conditions on the relative weighting of the in-plane and *c*-axis component of the experimental spectrum[21] and show a similar trend to the experimental results.

The increase in absorption just beyond the band gap is most likely due to the broken symmetry in the region of the

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

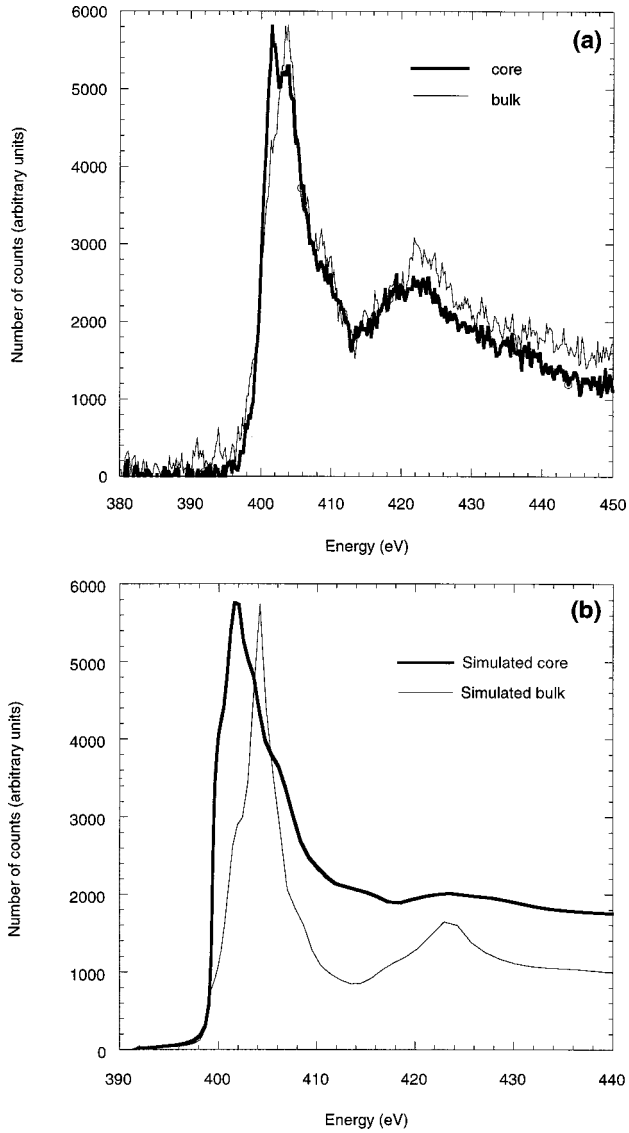


FIG. 3. (a) EELS spectra obtained from the pure edge dislocation (arrowed) and (b) nearby bulk compared to multiple scattering simulations.

core. Studies of the polarization dependence of XAS have shown the first peak of the spectrum to correspond to transitions to final states in the *c* direction, i.e., to states that originate from the mixing of *s* and p_z molecular orbitals, while the second peak is related to final states in the $a-b$ plane, i.e., from the mixing of p_x and p_y orbitals.[22] By comparison of Fig. 3(a) with XAS nitrogen *K*-edge spectra[23] it is clear that the general shape of the acquired EELS data is very similar to the x-ray absorption spectrum obtained when the photoelectric vector, **E**, is perpendicular to the *c* axis. This is not surprising as the effective collection angle used in this experiment ($\sim$40 mrad) is much larger than the characteristic angle, $\theta_E = 1$ mrad, and therefore momentum transfer perpendicular to the beam (i.e., *c* axis) should swamp the component parallel to the *c*-axis.[21] This strong increase of states along the *c* axis, i.e., more $s-p_z$ hybridization, is also seen in the cubic form of GaN, and it therefore most likely just reflects the broken C_{6v} symmetry of the core. It is noteworthy that no absorption is seen in the band gap. This is consistent with any Ga vacancy induced acceptor states being full, as expected if the dislocation line is charged, or may be simply due to a low matrix element between the core states and the defect states. A full understanding of this subtlety requires extensive *ab initio* simulations.

In conclusion, we show that pure edge threading dislocations in *n*-type GaN grown under N-rich conditions have a Ga vacancy concentration of 0%–15% in the core, significantly less than previously assumed. EELS studies show a significant increase in the first peak above the nitrogen *K*-edge onset at the dislocation core compared to the perfect region, reflecting the reduced C_{6v} symmetry of the core. No absorption was observed below the band gap, consistent with the dislocation line being negatively charged.

The authors are grateful to A. F. Wright of Sandia National Laboratory for helpful discussions during the course of this work, which was funded by NSF Grant No. DMR-9733895 and by DOE Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

[1] S. D. Lester, F. A. Ponce, M. G. Craford, and D. A. Steigerwald, Appl. Phys. Lett. **66**, 1249 (1995).
[2] J. Elsner, R. Jones, P. K. Sitch, V. D. Porezag, M. Elstner, T. Frauenheim, M. I. Heggie, S. Oberg, and P. R. Briddon, Phys. Rev. Lett. **79**, 3672 (1997).
[3] S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, Y. Sugimoto, T. Kozaki, H. Umemoto, M. Sano, and K. Chocho, Appl. Phys. Lett. **72**, 211 (1998).
[4] O. Nam, M. D. Bremser, and R. F. Davis, Appl. Phys. Lett. **71**, 2638 (1997).
[5] S. J. Rosner, E. C. Carr, M. J. Ludowise, G. Girolami, and H. I. Erikson, Appl. Phys. Lett. **70**, 420 (1997).
[6] T. Sugahara, H. Sato, M. S. Hao, Y. Naoi, S. Kurai, S. Tottori, K. Yamashita, K. Nishino, L. T. Romano, and S. Sakai, Jpn. J. Appl. Phys., Part 2 **37**, L398 (1998).
[7] A. F. Wright and U. Grossner, Appl. Phys. Lett. **73**, 2751 (1998).
[8] J. Elsner, R. Jones, M. I. Heggie, P. K. Sitch, M. Haugk, V. D. Porezag, Th. Frauenheim, S. Oberg, and P. R. Briddon, Phys. Rev. B **58**, 12571 (1998).
[9] N. G. Weimann, L. F. Eastman, D. Doppalapudi, H. M. Ng, and T. D. Moustakas, J. Appl. Phys. **83**, 3656 (1998).
[10] P. J. Hansen, Y. E. Strausser, A. N. Erickson, E. J. Tarsa, P. Kozodoy, E. G. Brazel, J. P. Ibbetson, U. Mishra, V. Narayanamurti, S. P. DenBaars, and J. S. Speck, Appl. Phys. Lett. **72**, 2247 (1998).
[11] C. Youtsey, L. T. Romano, and I. Adesida, Appl. Phys. Lett. **73**, 797 (1998).
[12] D. C. Look and J. R. Sizelove, Phys. Rev. Lett. **82**, 1237 (1999).
[13] Y. Xin, S. J. Pennycook, N. D. Browning, P. D. Nellist, S. Sivananthan, J.-P. Faurie, and P. Gibart, Appl. Phys. Lett. **72**, 2680 (1998).
[14] E. M. James and N. D. Browning, Ultramicroscopy **78**, 125 (1999).
[15] S. J. Pennycook and D. E. Jesson, Ultramicroscopy **37**, 14 (1991).
[16] N. D. Browning, M. F. Chisholm, and S. J. Pennycook, Nature (London) **366**, 143 (1993).
[17] M. Hao, T. Sugahara, H. Sato, Y. Morishima, Y. Naoi, L. T. Romano, and S. Sakai, Jpn. J. Appl. Phys., Part 2 **37**, L291 (1998).
[18] R. F. Egerton, *Electron Energy Loss Spectroscopy in the Electron Microscope*, 2nd ed. (Plenum, New York, 1996).
[19] T. J. Eustis and J. Silcox, Microsc. Microanal. **5**, 642 (1999).
[20] A. L. Ankudinov, B. Ravel, J. J. Rehr, and S. D. Conradson, Phys. Rev. B **58**, 7565 (1998).
[21] N. D. Browning, J. Yuan, and L. M. Brown, Ultramicroscopy **38**, 291 (1991).
[22] M. Katsikini, E. C. Paloura, and T. D. Moustakas, J. Appl. Phys. **83**, 1437 (1998).
[23] K. Lawniczak-Jablonska, T. Suski, Z. Liliental-Weber, E. M. Gullikson, J. H. Underwood, R. C. C. Perera, and T. J. Drummond, Appl. Phys. Lett. **70**, 2711 (1997).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

The mechanism for the high-quality single-phase growth of MnSi films on Si (111) in the presence of Sb flux

Y. Yan[a] and M. M. Al-Jassim
National Renewable Energy Laboratory, Golden, Colorado 80401

K. Matsuda, H. Tatsuoka, and H. Kuwabara
Faculty of Engineering, Shizuoka University, Hamamatsu 432-8561, Japan

S. J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

(Received 2 August 1999; accepted for publication 7 September 1999)

The microstructures of high-quality single-phase MnSi layers grown on Si (111) by Mn deposition and reaction with Si in the presence of Sb flux are characterized by Z-contrast imaging. It is found that there is a transition layer consisting of two Sb monolayers sandwiching a Mn layer in between the Si substrate and the single-phase MnSi film. This Sb–Mn–Sb sandwich layer effectively prevents deposited Mn atoms from direct reaction with Si atoms in the substrate to form Mn silicides. This explains why high-quality single-phase MnSi layers can be grown with remarkably smooth interface on Si (111) substrates. © *1999 American Institute of Physics.*
[S0003-6951(99)01645-9]

Metal–silicide/silicon interfaces have drawn extensive attention because of their microelectronics applications.[1,2] While many studies have been carried out on W, Ti, Co, and Ni silicides,[3–5] much less attention has been paid to silicides of the group VIIA elements, such as Mn. It is reported that MnSi, a cubic FeS structure, shows a magnetic transition at $T_c = 29.1$ K from a paramagnetic state to a helicoidal order. Its interesting magnetic properties make it a promising material for use in electronic and optoelectronic devices.[6–8] However, there are few reports to date on the growth of Mn-based layers on Si substrates. The main reason is due to the difficulties of growing a high-quality single-phase layer with a smooth interface. In the most common growth procedures, Mn atoms easily react with Si atoms in the substrate to form many different polycrystalline Mn–silicides phases, resulting in a very rough interface.[9–12] Very recently, the growth of high-quality single-phase MnSi films on Si (111) has been reported with remarkably smooth interfaces.[13,14] The critical factor for the growth is the presence of an Sb flux. Without the presence of Sb, polycrystalline Mn–silicides form leading to rough interfaces. Thus, understanding the effects of the Sb flux during the growth will enable the growth of higher-quality films. However, so far no conclusive mechanism has been given.

In this letter, we report our Z-contrast imaging studies on the interfaces of high-quality single-phase MnSi films grown on Si (111), which reveal the growth mechanism for high-quality films.

The MnSi films were grown on Si (111) substrates by hot-wall epitaxy. The substrates were solvent degreased, etched in HF solution, and rinsed with de-ionized water. The oxide layer was etched away by immersing the wafers in a dilute hydrofluoric acid solution ($HF:H_2O = 1:10$), and dried prior to loading into the vacuum chamber. Growth was performed by simultaneous exposure of the Si substrates, held at a temperature of 350 °C, to Mn and Sb atoms. The Mn flux used was 2×10^{14} atom/cm^2 s, while the Sb/Mn flux ratio supplied to the Si surface was 10.

Specimens were prepared for electron microscopy by first mechanical polishing to ~100 μm, then dimpling the central portion of the specimens to ~5 μm. Samples were thinned to electron transparency using a 4 kV Ar ion beam at 14° inclination and then cleaned at lower voltage (1.5 kV). A liquid N_2 cooling stage was used in order to minimize the damage during the ion milling. The Z-contrast images were formed by scanning a 1.26 Å probe across a specimen and recording the transmitted high-angle scattering with an annular detector (inner angle ~45 mrad). The image intensity can be described accurately as a convolution between the electron probe and an object function. Thus, the Z-contrast image gives a directly interpretable, atomic resolution map of the columnar scattering cross section in which the resolution is limited by the size of the electron probe.[15,16]

The quality and orientation of the Mn–Si film and the smoothness of the interface have been characterized by conventional transmission electron microscopy previously.[13,14] The Mn–Si layer was confirmed to be the single MnSi phase, with the epitaxial relationship between the MnSi layer and the Si substrate given by $(111),[\bar{1}\bar{1}2]\text{MnSi} \parallel (111),[\bar{1}10]\text{Si}$. The MnSi/Si interface is remarkably smooth, as seen in the low magnification Z-contrast image of Fig. 1. The electron beam is parallel to the $[\bar{1}10]$ zone axis of the Si substrate. The MnSi layer gives higher intensity than the Si substrate because Mn atoms have a higher atomic number ($Z=25$) than Si ($Z=14$). It is seen that the interface is very smooth, consistent with previous results.

Figure 2(a) shows a higher-magnification Z-contrast image of the MnSi/Si interface viewed along the $[\bar{1}10]$ zone axis of Si. The lattice image of Si is clearly seen, in which the bright spots correspond to atomic columns. It is seen that

[a]Electronic mail: yyan@nrel.gov

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

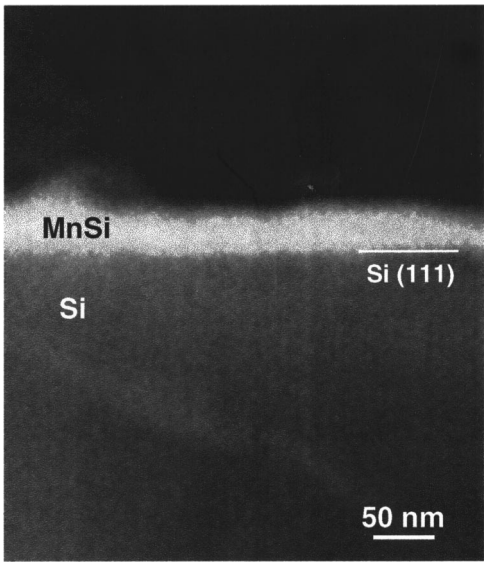


FIG. 1. Low-magnification Z-contrast image of an epitaxial MnSi film grown on Si(111) showing a very smooth interface.

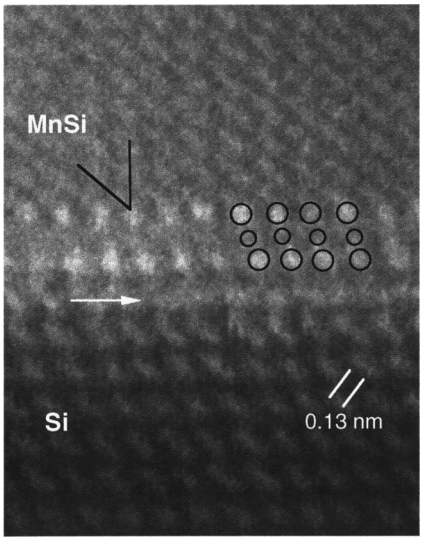


FIG. 3. High-resolution Z-contrast image of the MnSi/Si interface showing the Sb–Mn–Sb sandwich structure at the interface.

the MnSi/Si interface is atomically smooth. It is surprising that the first two planes of the grown film are much brighter than either the Si or the MnSi layer. Figure 2(b) shows the intensity profile measured from the dotted box region across the interface. The position of the interface is indicated by an arrow. The two peaks at the interface have much higher intensities than bulk MnSi or Si. This strongly suggests that Sb atoms have segregated to the interface. This is consistent with SIMS measurements showing Sb segregation on the interface.[17]

Figure 3 is an enlargement of the image showing details of the structure of the interface. The Si dumbbells (pairs of atomic columns with a separation of 0.136 nm) are clearly visible. Also seen is a monolayer of Sb atoms bound to Si atoms on the (111) surface of the substrate (indicated by larger circles). The separation of two adjacent Sb columns is the same as that of two adjacent dumbbells. The first monolayer of Sb modifies the surface of the substrate: the Si columns at the surface no longer have the dumbbell shape as indicated by the white arrow. The small increase of intensities for some Si columns at the surface suggests a small occupation of Sb in those columns. Above the first monolayer of Sb is another monolayer as indicated by smaller circles. The intensities of columns in this layer are higher than Si columns in the substrate but lower than Sb columns, indicating that they are Mn columns. This layer of Mn is sandwiched by two Sb monolayers. The MnSi film then grows epitaxially on top of the second Sb monolayer. The two dark lines show how two planes of MnSi match the Sb–Mn–Sb structure.

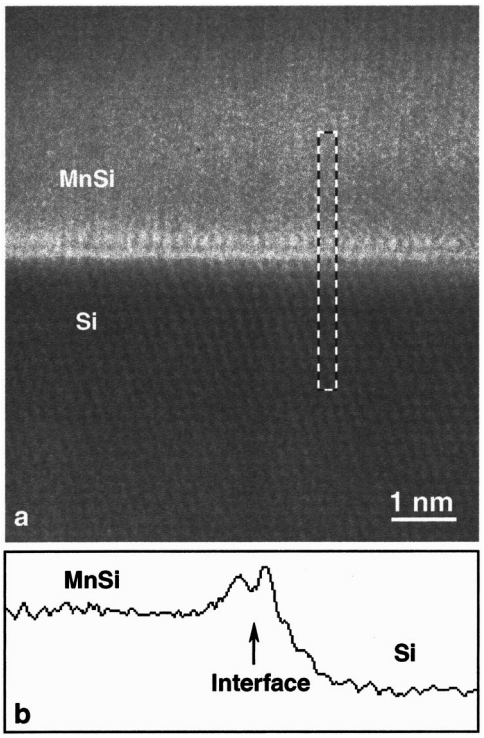


FIG. 2. Higher-resolution Z-contrast image of the MnSi/Si interface showing a narrow band of Sb segregation at the interface.

It has been reported that Sb–Si adsorption can easily occur on Si surfaces even at low temperature.[18] The desorption of Sb on Si (111), i.e., the breaking of Sb–Si bonds, takes place only above a peak temperature of $T_p = 950\,°\mathrm{C}$. This indicates that the Sb–Si bonds on the Si (111) surface are energetically very stable. Since the Sb/Mn flux ratio supplied to the Si surface was 10, it would greatly enhance the formation of Sb–Si bonding. The substrate temperature (350 °C) is too low to break the Sb–Si bonds. Thus, a monolayer of Sb can easily form and firmly cover the Si (111) surface. However, it is known that the free energy of formation of MnSi (7.25 kcal/g atom) is significantly higher than that of MnSb (3.25 kcal/g atom).[19] Thus, the formation of MnSi is thermodynamically favored as compared with MnSb. A thick layer of MnSb is therefore not expected at the interface. This is consistent with the fact that only a Sb–Mn–Sb sandwich layer is observed at the interface.

It is believed that reaction between deposited Mn and Si in the substrate is responsible for the formation of many different polycrystalline Mn–silicide phases that lead to rough, wavy interfaces. The growth mechanism for the high-quality MnSi films on Si (111) in the presence of Sb flux has not so far been understood. However, it is now seen from the

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

above observations, that in the presence of the Sb flux, a Sb–Mn–Sb sandwich structure forms and covers the surface prior to growth of the MnSi film. This Sb–Mn–Sb structure prevents the diffusion of Mn into the substrate to react with Si and form Mn silicides. This forces the epitaxial growth of the single MnSi phase with an atomically smooth interface. Thus, it is clear that the mechanism for the growth of high-quality single-phase MnSi film with a smooth interface is the formation of the Sb–Mn–Sb sandwich structure as a result of the presence of Sb flux during the growth. This is very similar to the recent report on the growth of Ge on Si:[20,21] With the presence of Sb as a surfactant, the inner diffusion of epitaxial Ge atoms into the Si substrate is greatly suppressed, leading to the formation of a very abrupt Ge/Si interface. In both cases, excess surfactant floats on the surface of the growing film. In our case, however, some is used to form the Sb–Mn–Sb sandwich structure at the film/substrate interface.

In conclusion, we have studied the structure of high-quality smooth MnSi/Si interface using Z-contrast imaging. We found that the formation of a Sb–Mn–Sb sandwich layer on the Si (111) surface is the reason that high-quality single-phase MnSi films can be grown with atomically smooth interfaces.

The work at NREL was supported by the U.S. Department of Energy under Contract No. DE-AC36-98GO 10337; at ORNL the work was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

[1] A. Hiraki, Appl. Surf. Sci. **56/58**, 370 (1992).
[2] H. Lange, Mater. Res. Soc. Symp. Proc. **402**, 307 (1995).
[3] F. M. D'Heurle and C. S. Petersson, Thin Solid Films **128**, 283 (1985).
[4] K. N. Tu, *Advances in Electronic Materials*, edited by B. W. Wessels and G. Y. Chin (American Society for Metals, Meta Park, OH, 1986), Chap. 7.
[5] C. I. Gregory, D. B. Lambrick, and N. R. Bernhoeft, J. Magn. Magn. Mater. **104/107**, 689 (1992).
[6] P. Lerch and T. Jarlborg, J. Magn. Magn. Mater. **131**, 321 (1994).
[7] P. Harris, B. Lebech, H. S. Shin, K. Mortensen, and J. S. Pedersen, Physica B **213/214**, 357 (1995).
[8] C. Thessieu, J. Flouquet, and G. Lapertot, Solid State Commun. **95**, 707 (1995).
[9] K. E. Sundstrom, S. Petersson, and P. A. Tove, Phys. Status Solidi A **20**, 563 (1973).
[10] M. Eizenberg and K. N. Tu, J. Appl. Phys. **53**, 6885 (1982).
[11] L. Zhang and D. G. Ivey, J. Mater. Res. **6**, 1518 (1991).
[12] J. Wang, M. Hirai, M. Kusaka, and M. Iwami, Appl. Surf. Sci. **113/114**, 53 (1997).
[13] H. Tatsuoka, K. Isaji, K. Sugiura, H. Kuwabara, P. B. Brown, Y. Xin, and C. J. Humphreys, J. Appl. Phys. **83**, 5504 (1998).
[14] K. Matsuda, H. Katsuoka, K. Matsunaga, K. Isaji, H. Kuwabara, P. B. Brown, Y. Xin, R. Dunin-Borkowski, and C. J. Humphreys, Jpn. J. Appl. Phys., Part 1 **37**, 6556 (1998).
[15] S. J. Pennycook and D. E. Jesson, Phys. Rev. Lett. **64**, 938 (1990).
[16] S. J. Pennycook and D. E. Jesson, Ultramicroscopy **37**, 14 (1991).
[17] K. Matsuda, Y. Takano, K. Kuwahara, H. Tatsuoka, Y. Yan, S. J. Pennycook, and H. Kuwabara (unpublished).
[18] R. A. Metzger and F. G. Allen, Surf. Sci. **137**, 397 (1984).
[19] O. Kubaschewski and C. B. Alcock, in *International Series on Materials Science and Technology: Metallurgical Thermo-chemistry*, 5th ed. (Pergamon, Oxford, 1979), Vol. 24, p. 294.
[20] K. Fujita, S. Fukatsu, H. Yaguchi, T. Igarashi, Y. Shiraki, and R. Ito, Jpn. J. Appl. Phys., Part 2 **29**, L39 (1990).
[21] Z. Jiang, A. Xu, D. Hu, H. Zhu, X. Liu, X. Wang, M. Mao, X. Zhang, J. Hu, D. Hung, and X. Wang, Thin Solid Films **321**, 116 (1998).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

ELSEVIER

Surface Science 442 (1999) 251–255

surface science

www.elsevier.nl/locate/susc

Structures of pure and Ca-segregated MgO (001) surfaces

Yanfa Yan [a,*], M.F. Chisholm [a], S.J. Pennycook [a, b], S.T. Pantelides [a, b]

[a] *Solid State Division, Oak Ridge National Laboratoty, Oak Ridge, TN 37831, USA*
[b] *Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235, USA*

Received 8 April 1999; accepted for publication 5 August 1999

Abstract

The structures of pure and Ca-segregated MgO (001) surfaces have been studied using first-principles density-functional theory. The relaxation and rumpling for the pure surface are found to be 0.48% and 1.62%, respectively. Ca segregation significantly modifies the surface structure. The surface-segregated Ca atoms protrude outwards owing to the size mismatch between Ca and Mg. Consequently, their nearest neighbor oxygen atoms are pulled up. The value of the protrusion of Ca atoms is strongly dependent on the Ca coverage of the surface. © 1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: Ab initio calculations; Magnesium oxide; Segregation; Surface

1. Introduction

The MgO (001) surface is widely used as a substrate for the epitaxial growth of superconductors and metals [1–3]. The structure and morphology of the surface play a critical role in the growth of thin films. It is therefore important to understand the details of the surface structure for the growth of high quality thin films. Extensive studies have been carried out on the pure MgO (001) surface both experimentally and theoretically [4–9]. However, a consistent picture for describing the surface structure has not yet been produced. It is known that low concentration of impurities that segregate to the surface can change the structure of the surface, and thus affect the growth of thin films. MgO is a material that is extremely difficult to purify. It always contains a small amount of impurities (in the parts per million range) such as Al, Ca, C and Fe. Among these impurities, Ca often segregates to grain boundaries and surfaces [10,11]. Very recently, the effects of Ca segregation on the MgO (001) surface have been studied using time-of-flight impact-collision ion scattering spectroscopy (IOF-ICISS), Rutherford backscattering spectroscopy [12], and grazing-incidence X-ray scattering (GIXS) [13]. These studies showed that Ca atoms substitute for Mg atoms at the topmost layer, and protrude outwards owing to the size mismatch between Ca and Mg. However, the value measured by GIXS, 0.63 Å, is much larger than that measured by IOF-ICISS, 0.4 Å. It is reported that the coverage of Ca on the surface of the sample measured by IOF-ICISS is about 20%, while it is about 50% for the sample measured by GIXS. Thus, it is reasonable to speculate that the structure of the Ca-segregated surface is dependent on the coverage of Ca. However, so far, there are neither systematic exper-

* Corresponding author. Present address: National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO 80401, USA. Fax: +1-303-384-6604.
E-mail address: yanfa_yan@nrel.gov (Y. Yan)

0039-6028/99/$ - see front matter © 1999 Published by Elsevier Science B.V. All rights reserved.
PII: S0039-6028(99)00923-1

imental studies nor first-principles calculations on the Ca-segregated surface. Only one empirical potential calculation has been carried out by Masri et al. [14]. The purpose of this paper is to investigate the structure of the pure MgO (001) surface, the effect of Ca segregation on the surface and the dependence of the surface structure on the coverage of Ca using first-principles density-functional theory.

2. Method

In our calculations, we use density-functional theory with the exchange-correlation potential treated in the generalized gradient approximation (GGA) [15]. The atomic cores are represented by nonlocal, norm-conserving pseudopotentials in a separable form [16]. The Ca pseudopotential includes the 3p states as valence electrons. We relax the surface structures by a periodic slab geometry. Each slab contains 11 layers and a (001) surface on each side of the slab. The vacuum region between such slabs has a thickness of about 17 Å. To optimize the structure, atoms are relaxed along the calculated forces until the remaining forces are all within 0.1 eV $Å^{-1}$. The wave functions were expanded in a plane wave basis set with an energy cutoff of 600 eV and the integration over the Brillouin zone was performed using three special k points chosen according to the Monkhorst–Pack scheme [17]. We calculated the lattice constant for a perfect MgO crystal and the structure of a pure MgO (001) surface using both GGA and local density approximation (LDA) [18]. Consistent results were obtained. The calculated value of the lattice parameter, $a_0 = 4.20$ Å, is satisfactorily close to the experimental value of 4.21 Å.

3. Results

We first calculated the structure of the pure MgO (001) surface. The terms rumpling and relaxation are commonly used to describe the structure of metal oxide surfaces. For a pure surface, the perpendicular displacement of anions Z_a and that of the cations Z_c can be significantly different. The difference between the displacements $Z_a - Z_c$ is called the surface rumpling, whereas the mean movement of the surface layer $(Z_a + Z_c)/2$ is called the surface relaxation. Normally both are described in percentages of the bulk nearest-neighbor spacing. Our fully relaxed slab shows a very slight rumpling of the surface, with oxygen ions displacing outwards by 0.022 Å and magnesium ions displacing inward by 0.012 Å with respect to the unrelaxed surface. Such displacements occur only at the topmost layer. The values of rumpling and relaxation were calculated to be 1.62 and 0.48%, respectively.

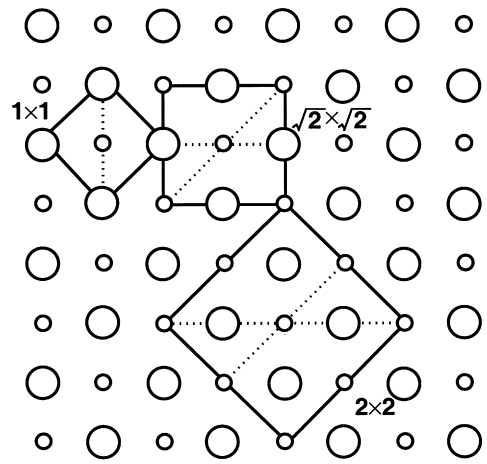


Fig. 1. Structure of the perfect MgO (001) surface showing the 1×1, $\sqrt{2} \times \sqrt{2}$ and 2×2 cells.

Calculations were then performed on the Ca-segregated MgO (001) surfaces with three coverages of Ca. Ca atoms are considered to substitute for Mg atoms in the top layer. We use three different cell sizes on the surface to represent the coverages of Ca, as shown in Fig. 1. The large circles indicate oxygen atoms and the small ones magnesium atoms. If we substitute Mg with Ca only at the center of these cells, the three cells corresponding to 1×1, $\sqrt{2} \times \sqrt{2}$ and 2×2 have 100, 50 and 25% coverages of Ca. Because, the calculations are very time consuming, we did not calculate beyond the 25% coverage. Our calculations showed that the segregated Ca not only modifies the topmost layer, but also affects several layers underneath. For each unit cell, sections

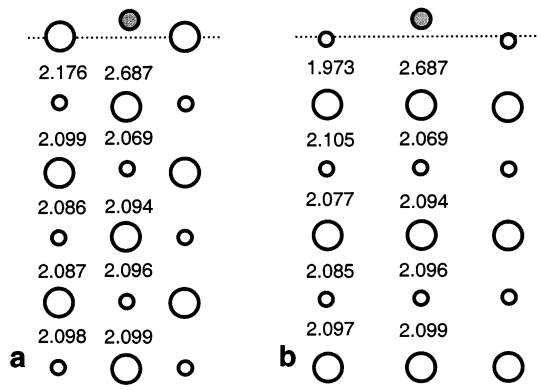


Fig. 2. Atomic positions on (a) (100) section and (b) (110) section of the relaxed $\sqrt{2}\times\sqrt{2}$ structure with 50% coverage of Ca. The dotted lines indicate the position of the unrelaxed surface. The vertical inter-distances are given in Å.

indicated by dotted lines cover all inequivalent sites. Thus, the relaxation below the surface can be described by giving the atomic position in these sections.

Fig. 2 shows the atomic positions on both the (100) and (010) sections for the relaxed $\sqrt{2}\times\sqrt{2}$ structure with 50% coverage of Ca. Ca atoms are indicated by broken circles. It is seen that Ca protrudes outwards by 0.547 Å and its neighboring O on the surface is consequently pulled out by 0.046 Å with respect to the unrelaxed surface, as indicated by dotted lines in Fig. 2. The neighboring Mg on the top layer is pushed inwards by 0.163 Å. The effect of Ca-induced relaxation extends further into the crystal; rearrangements are seen within four layers.

The vertical displacements of atoms of the first two layers of the $\sqrt{2}\times\sqrt{2}$ surface structure with 50% coverage of Ca were measured by Robach et al. using GIXS [13]. Comparison of their measured and our calculated results is shown in Table 1. It is seen that the calculated values for the protrusion of Ca, 0.547 Å, and the displacement of Mg, −0.163 Å, are slightly smaller than the measured values, 0.63 ± 0.03 and -0.066 ± 0.14 Å, respectively. This difference is probably caused by procedure used to fit the data: the O atoms in the surface and all atoms below the second layer were fixed, which is inconsistent with our calculation. Our relaxed structure (Fig. 2) clearly shows that O atoms in the top layer are pulled outwards by 0.046 Å, and the relaxation extends into the bulk for four layers. If we renormalize the calculated displacements by adding the displacement of O in the top layer, 0.046 Å, the calculated displacements (the forth row in Table 1) for atoms in the top layer agree very well with the experimental data.

Fig. 3 shows the (100) and (110) sections of the relaxed 2×2 surface structure with 25% coverage of Ca. It is seen that the relaxed structure differs from that of the surface with 50% coverage of Ca, as follows. The Ca-induced relaxation extends into the crystal for only three layers. The protrusion of Ca, 0.477 Å, is much smaller than for the $\sqrt{2}\times\sqrt{2}$ structure with 50% coverage of Ca. Oxygen atoms are pulled outwards by 0.073 Å with respect to the unrelaxed surface, slightly smaller than for the $\sqrt{2}\times\sqrt{2}$ structure. The first nearest neighbor Mg in the top layer is pushed inwards by 0.058 Å, while the second nearest neighbor Mg is pushed inwards by 0.009 Å. Souda et al. [12] have measured the structure of the MgO (001) surface with 20% coverage of Ca. The measured value for the Ca protrusion, 0.4 ± 0.01 Å, is only slightly smaller than our calculated value of 0.477 Å, consistent with the slightly reduced coverage of Ca in the experiment.

Table 1
Comparison of measured and calculated vertical displacements of atoms at the first two layers of the $\sqrt{2}\times\sqrt{2}$ structure with 50% coverage of Ca. The numbers labeling the atoms are the same as those used in Ref. [13]

	1:Ca	2:Mg	3:O	4:O	5:Mg	6:Mg	7:O	8:O
exp	0.63 ± 0.03	$-0..066\pm0.14$	fixed	fixed	0.20 ± 0.03	-0.42 ± 0.18	-0.42 ± 0.32	0.42 ± 0.12
cal.	0.547	−0.163	0.046	0.046	−0.030	−0.030	−0.036	−0.042
renorm.	0.593	−0.117	0.0	0.0	0.016	0.016	0.010	0.004

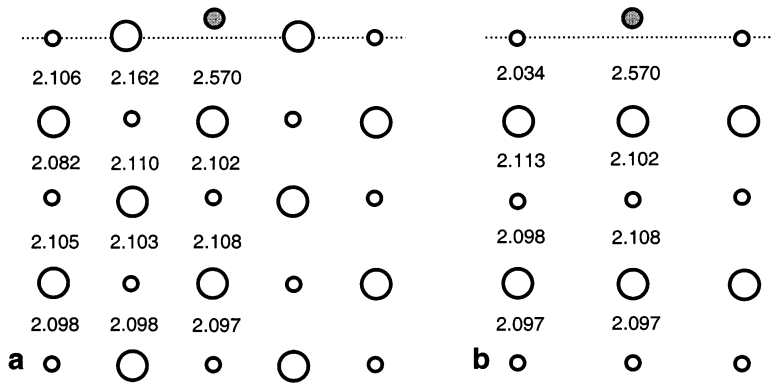


Fig. 3. Atomic positions on (a) (100) section and (b) (110) section of the relaxed 2×2 structure with 25% coverage of Ca. The dotted lines indicate the position of the unrelaxed surface.

To strengthen our conclusion that the Ca-segregated surface structure is dependent on the coverage of Ca, we calculated the surface structure with another coverage of Ca. Fig. 4 shows the atomic positions on the (010) section for the relaxed 1×1 structure with 100% coverage of Ca. As we expected, owing to the higher Ca coverage, the Ca-induced relaxation extends into the crystal for more than five layers. Ca protrudes outwards by only 0.244 Å, but O in the top layer is pulled outwards by 0.071 Å with respect to the unrelaxed surface as indicated by dotted lines in Fig. 4. This again clearly shows that the surface structure is dependent on the coverage of Ca. It should be noted that, below the second layer, both anions and cations are forced to move inwards. This distortion would require a large amount of energy. We therefore expect that the MgO (001) surface with 100% coverage of Ca would never occur experimentally.

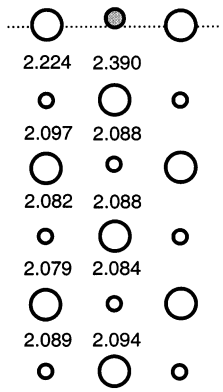


Fig. 4. Atomic positions on (010) section the relaxed 1×1 structure with 100% coverage of Ca. The dotted lines indicate the position of the unrelaxed surface.

4. Conclusions

The density-functional calculations show that the pure MgO (001) surface has a rumpling of 1.62% and a relaxation of 0.48%. Ca segregation significantly roughens the surface. The relaxation induced by Ca segregation propagates further into the bulk crystal. The surface structure is strongly dependent on the coverage of Ca.

Acknowledgements

The work was supported in part by the Division of Materials Science, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation, by an appointment to the Postdoctoral Research Program administered jointly by ORNL and ORISE, and by NSF grant DMR-9803768.

References

[1] Y.C. Lee, P. Tong, P.A. Montano, Surf. Sci. 181 (1987) 559.
[2] P. Guenard, G. Renaud, B. Villette, Physica B 221 (1996) 205.
[3] C. Li, R. Wu, A.J. Freeman, C.L. Fu, Phys. Rev. B 48 (1993) 8317.
[4] P.A. Maksym, Surf. Sci. 149 (1985) 157.
[5] D.L. Blanchard, D.L. Lessor, J.P. LaFemina, D.R. Baer, W.K. Ford, T. Guo, J. Vac. Sci. Technol. A 9 (1991) 1814.
[6] J.P. LaFemina, C.B. Duke, J. Vac. Sci. Technol. A 9 (1991) 1847.
[7] S. Pugh, M.J. Gillan, Surf. Sci. 320 (1994) 331.
[8] J. Goniakowski, C. Noguera, Surf. Sci. 323 (1995) 129.
[9] Y. Li, D.C. Langreth, M.R. Pederson, Phys. Rev. B 55 (1997) 16 456.
[10] R.C. McCune, P. Wynblatt, J. Am. Ceram. Soc. 66 (1983) 111.
[11] Y. Yan, M.F. Chisholm, G. Duscher, A. Maiti, S.J. Pennycook, S.T. Pantelides, Phys. Rev. Lett. 81 (1998) 3675.
[12] R. Souda, Y. Hwang, T. Aizawa, W. Hayami, K. Oyoshi, S. Hishita, Surf. Sci. 387 (1997) 136.
[13] O. Robach, G. Renaud, A. Barbier, Surf. Sci. 401 (1998) 227.
[14] P. Masri, P.W. Tasker, J.P. Hoare, J.H. Harding, Surf. Sci. 173 (1986) 439.
[15] J.P. Perdew, Y. Wang, Phys. Rev. B 45 (1992) 13 244.
[16] L. Kleinmann, D.M. Bylander, Phys. Rev. Lett. 48 (1982) 1425.
[17] H.J. Monkhorst, J.D. Pack, Phys. Rev. B 13 (1976) 5188.
[18] J.P. Perdew, A. Zunger, Phys. Rev. B 23 (1981) 5048.

Structural model for the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystals

Yanfa Yan
National Renewable Energy Laboratory, Golden, Colorado 80401

Stephen J. Pennycook
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

(Received 15 February 2000)

We propose a structure model for the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystals based on its τ^2-inflated $Al_{13}Co_4$ approximant phase: Applying a 10_5 screw operation on a monolayer obtained from the approximant reproduces almost all features of 2-nm clusters seen in atomic-resolution *Z*-contrast images of $Al_{72}Ni_{20}Co_8$ decagonal quasicrystals. The exception is the central ring, where the symmetry is broken due to chemical ordering of Al and transition metals. By restricting possible overlaps, this enforces the perfect quasiperiodic tiling.

Decagonal quasicrystals can be described by quasiperiodic packing of atomic clusters.[1] For example, all Al-based decagonal quasicrystals (e.g., Al-Ni-Fe, Al-Ni-Co, Al-Cu-Co, ...) comprise distinct 2-nm clusters.[2–5] The detailed determination of the atomic arrangements within the 2-nm clusters is critical for understanding why quasicrystals form. So far, many structural models have been proposed for decagonal quasicrystals based upon x-ray and neutron diffraction, high-resolution electron microscopy, and theoretical calculations.[6–11] However, controversies still remain among these models. It is well known that quasicrystals and their approximant phases have very similar chemical compositions, densities, and electron diffraction patterns, and reversible transformations between the two have been reported.[12] These observations indicate that quasicrystals and their approximant phases have many similar *local* structures. Thus an alternative approach to build the structure of quasicrystals is to understand the differences and relationships between quasicrystals and their approximant phases.

In the present paper, we use this approach to propose a structure model for the 2-nm clusters of the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal. We take pentagonal structural features present in its closely related τ^2-inflated $Al_{13}Co_4$ approximant phase to propose a cluster monolayer. Applying a 10_5 screw operation to this layer then gives a perfect fit to all features seen in atomic-resolution *Z*-contrast images obtained from the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal, *except* within the central ring. Here we observe broken symmetry induced by chemical ordering. This ordering restricts possible cluster overlaps and so enforces the perfect quasiperiodic tiling. At the same time, imperfect ordering throughout the structure creates the entropic stabilization necessary to explain why this phase is only stable at high temperatures.

We focus on the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal because of its high perfection[5] and because the structure of its closely related τ^2-inflated $Al_{13}Co_4$ approximant phase (a phase with *a* and *c* lattice parameters τ^2 times greater than those of monoclinic $Al_{13}Co_4$, where τ is the golden mean) has been well determined.[13] So far, $Al_{72}Ni_{20}Co_8$ is the only high-perfection decagonal quasicrystal ever found, showing no diffuse scattering in electron diffraction. This indicates that it has a periodicity of about 0.4 nm along its periodic axis. Our atomic-resolution *Z*-contrast images[14] [Figs. 1(d)–1(f)] show an important difference from the most popular models for the 2-nm clusters of Al-Ni-Co and Al-Cu-Co de-

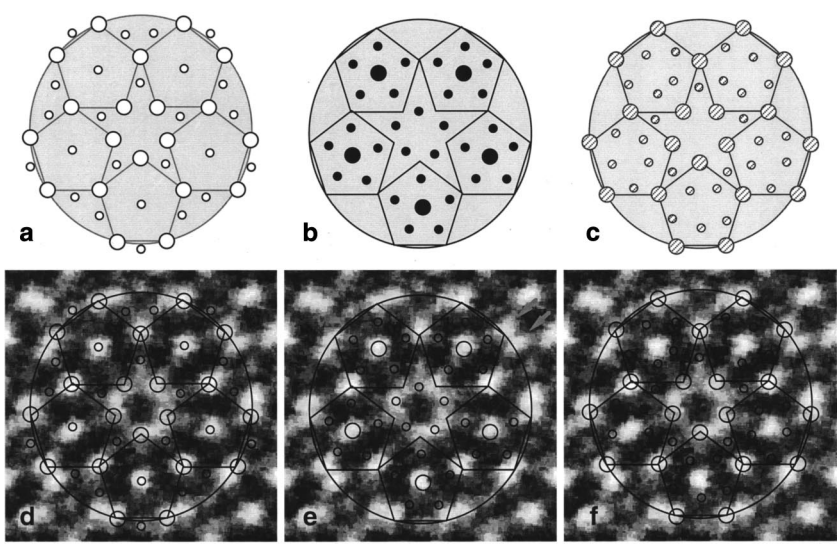


FIG. 1. (a)–(c) Three layers of the 2-nm clusters seen in τ^2-inflated $Al_{13}Co_4$ approximant (from Ref. 13) with (d)–(f) superimposition onto a high-resolution *Z*-contrast image of a 2-nm cluster of the high-perfection decagonal quasicrystal $Al_{72}Ni_{20}Co_8$, viewed along the tenfold axis.

 ©2000 The American Physical Society

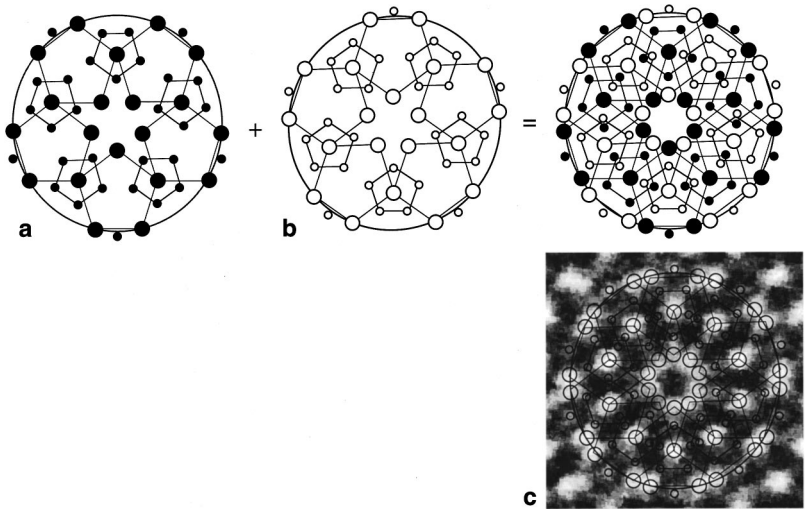


FIG. 2. (a) A layer structure obtained from the $z=0$ layer of the τ^2-inflated $Al_{13}Co_4$ approximant. Only the Al atoms in the third ring are rearranged to form small pentagons. (b) The layer after applying a 10_5 screw operation on (a). (c) Both layers (a) and (b) are superimposed on the *Z*-contrast image. The TM columns are the bright features, while the less intense intensity maxima (gray) are Al columns.

cagonal quasicrystals: These proposed models have one more ring containing transition metals (see, for example, the second outermost ring in the model proposed by Burkov[7]). We now analyze the relationships between the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal and its τ^2-inflated $Al_{13}Co_4$ approximant phase. It has been reported that the τ^2-inflated $Al_{13}Co_4$ approximant phase consists of fivefold and pseudofivefold 2-nm clusters.[13] The fivefold 2-nm clusters are more similar to the 2-nm clusters seen in the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal, and their atomic structure has been determined recently.[13] They contain four layers, since the periodicity along the columnar axis is 0.8 nm. The various layers are shown in Fig. 1, (a) corresponding to $z=0$, (b) to $z=1/4,3/4$, and (c) to $z=1/2$, which are stacked in the sequence $a,b,c,b,\ldots$. It is seen that the (a) and (c) layers differ only in the Al atom positions, which are more disordered in (c).

All Al-Ni-Co decagonal quasicrystals have been reported to contain a 10_5 screw axis,[2–5] and many of the previously proposed models for the 2-nm clusters have a 10_5 screw axis through their centers.[6–10] Thus the structure of the 2-nm cluster is solved once the structure for a single layer is known. We therefore need to find which layer of the fivefold 2-nm cluster of the τ^2-inflated $Al_{13}Co_4$ approximant phase is similar to the 2-nm clusters of the Al-Ni-Co decagonal quasicrystals. Figures 1(d)–1(f) show superimpositions of the three approximant layers on the *Z*-contrast image of a 2-nm cluster from the $Al_{72}Ni_{20}Co_8$ quasicrystal, taken along the periodic axis. It is seen that all sites in the pentagonal (a) layers match sites in the quasicrystal. All transition-metal (TM) atom sites show brightly in the image corresponding to high intensity, while all Al sites are seen as gray, but are still local maxima or saddle points in the image intensity, consistent with their lower *Z*. Because the (b) layer structure has no atoms in the outermost ring and the (c) layer structure has no Al atoms in the outermost ring, we do not consider these two structures further.

We now construct a cluster monolayer based on the (a) approximant layer. This monolayer must keep the key structural features of the approximant layer, e.g., all the TM atoms located at pentagonal sites. Such pentagonal shapes giving local fivefold symmetry are common in decagonal quasicrystals. We next assume that Al atoms also locate in pentagonal arrangements. The Al atoms in the third ring counting from the center are located at intensity maxima. We thus split these Al sites into two to form small Al pentagons as shown in Fig. 2(a). Now we see that this overlapping array of large TM and small Al pentagonal clusters fits *precisely half* the brightest features in the *Z*-contrast image. If we now apply the 10_5 screw operation to this monolayer, we obtain the layer shown in Fig. 2(b). Superimposing these two rotational variants on the *Z*-contrast image matches *all* image features [Fig. 2(c)].

This 10_5 screw axis can be understood as rotational ordering which occurs under certain conditions, presumably a specific range of chemical composition. The structure explains the origin of the ten columns in the central ring and the ten closely spaced TM column pairs around the 2-nm ring. Our model is similar to the model proposed by Burkov:[7] both have ten TM columns at the central ring and ten closely spaced Al and TM column pairs at the third and outermost rings, respectively. The ten closely spaced TM columns pairs at the outermost ring have also been observed by x-ray diffraction[6] and are predicted in the model of Cockayn and Widom.[10] It is noted that Burkov's model predicts a stoichiometric composition of $Al_{60}M_{40}$, which is low in Al com-

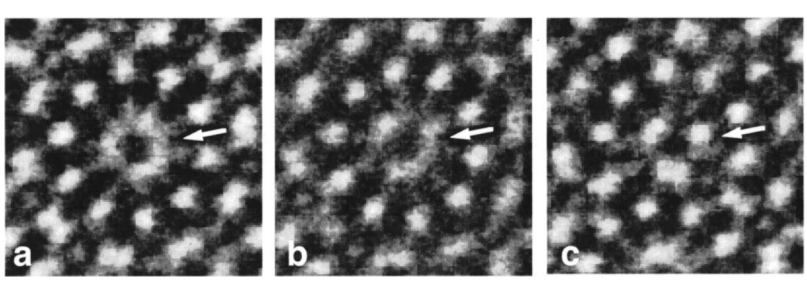


FIG. 3. *Z*-contrast images showing three 2-nm clusters with different degrees of chemical ordering at the central rings: (a) low, (b) medium, and (c) high. Chemical ordering breaks the symmetry of the central rings.

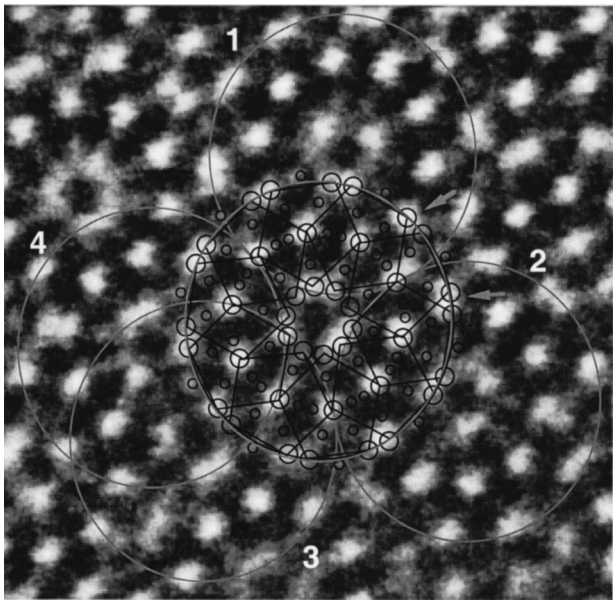


FIG. 4. *Z*-contrast image showing the four intercluster separations. The central circle shows a cluster with only four neighbors nucleated due to the broken symmetry of the central ring. These four clusters now show all the intercluster separations necessary for quasiperiodic tiling: clusters 1, 3 and 2, 4 show the 3.44-nm separation. Clusters 1 and 2 are 2.06 nm apart, as are 2 and 3 and also 1 and 4. Clusters 1, 2, 3, and 4 all show a 1.23-nm separation from the central cluster, while a 0.74-nm separation is seen between clusters 3 and 4.

pared to the $Al_{72}Ni_{20}Co_8$ decagonal quasicrystal. A major difference between our model and Burkov's model is the absence of the ten TM atoms in the fourth ring counting from the center in our model (for which there is no evidence in our *Z*-contrast images). Thus our model significantly improves the stoichiometry.

It is important to note that none of the central rings show the uniform strong intensity characteristic of ten TM columns. In Fig. 3(a), the central ring shows a relatively uniform and high intensity, but in comparison to the ring of ten TM columns, the intensity is clearly not high enough to correspond to ten full TM columns. Partial substitution by Al has taken place, which we referred to previously as chemical disorder and proposed as the source of entropy to stabilize the phase at high temperatures.[16] Chemical ordering has also been reported in other icosahedral quasicrystals.[17] The central rings of some clusters show much stronger broken symmetry, as shown in Fig. 3(c). Similar broken symmetry has also been reported previously.[3,4,15] In our images we can see that *five* columns show an intensity sufficiently high (bright) to correspond to full TM columns, one being a single column, the others forming closely spaced column pairs. The other five columns have lower intensity (gray), indicating high Al occupancy. The average composition of the ordered ring is therefore 50% TM and 50% Al. In Fig. 3(b), the central ring shows an intermediate case, also with some broken symmetry. In the disordered case, the intensities are consistent with the *same* average composition for the ring, 50% TM and 50% Al. Assuming this to be the case, the stoichiometry of our model becomes $Al_{73}M_{27}$, which is very close to the nominal composition $Al_{72}M_{28}$.

The intermediately ordered cases might also be interpreted as resulting from phason flips within individual ordered cluster columns.[18] However, we find no correlation between these different degrees of order and the configuration of overlapping clusters, as would be expected in this case. We propose therefore that these more uniform rings represent true chemical disorder. Further evidence of chemical disorder is seen in Fig. 4, where large intensity variations exist between different TM sites. These intensity variations are many times too large to be attributed to statistical noise in the image. They must represent real variations in composition in particular atomic columns, another indication of a high chemical entropy contribution.

Atomistic explanations for the overlapping cluster network in quasicrystals have been proposed previously.[19] In our case, the closely spaced pairs of TM columns that are present in the ordered central rings and in the outer rings of the clusters provide an ideal atomistic mechanism to interlink clusters and explain the matching rules.[20] Such a growth process naturally produces overlaps with intercluster separations in the ratio τ, exactly the 0.74-, 1.23-, 2.06-, and 3.44-nm spacings seen in the Gummelt coverage model of the ideal quasicrystal.[21–23] If there is no chemical ordering occurring, each 2-nm cluster will have many different ways to overlap with the next. Thus the clusters will form a random tiling. However, chemical ordering breaks the symmetry in the central rings and limits the freedom in forming overlaps. As seen from Fig. 4, the ordered central cluster has only four possible overlaps, which are in precisely the configuration of the Gummelt coverage model for the ideal quasicrystal. The overall structure would therefore show ideal quasicrystalline coverage only when complete chemical ordering occurs at the center of *all* the 2-nm clusters.

In the case of the approximant phase, if we try to overlap its clusters without the rotational variants, the density of common sites is very low. The approximant phase therefore grows only by the addition of *individual* small pentagonal units around its basic 2-nm cluster, as found by Saitoh *et al.*[3] in the case of the τ^2-inflated $Al_{13}Co_4$ approximant. Thus we believe that rotational ordering of 2-nm clusters, with chemical ordering in the central rings, is fundamental to stabilizing the quasicrystalline form. At the same time, substantial variations in site occupation occur throughout the cluster, which explains the entropic stabilization. Order and disorder coexist.

The work at NREL was supported by the U.S. Department of Energy under Contract No. DE-AC36-99GO10337. The work at ORNL was supported by the U.S. Department of Energy under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. We thank Dr. A. P. Tsai of the National Research Institute for Metals, Japan for the provision of high-perfection decagonal quasicrystal $Al_{72}Ni_{20}Co_8$ studied in this paper.

[1] P. Guyot, and M. Audier, Philos. Mag. B **52**, L15 (1985); V. Elser and C. L. Henley, Phys. Rev. Lett. **55**, 2883 (1985); W. Steurer, Acta Crystallogr., Sect. B: Struct. Sci. **B45**, 534 (1989); H.-C. Jeong and S. J. Steinhardt, Phys. Rev. Lett. **73**, 1943 (1994).
[2] K. Hiraga, F. J. Lincoln, and W. Sun, Mater. Trans., JIM **32**, 308 (1991).
[3] K. Saitoh, K. Tsuda, M. Tanaka, A.-P. Tsai, A. Inoue, and T. Masumoto, Philos. Mag. A **73**, 387 (1996); K. Saitoh, K. Tsuda, M. Tanaka, K. Kaneko, and A.-P. Tsai, Jpn. J. Appl. Phys., Part 2 **36**, L1400 (1998).
[4] K. Tsuda, Y. Nishida, K. Saitoh, M. Tanaka, A. P. Tsai, A. Inoue, and T. Masumoto, Philos. Mag. A **74**, 697 (1996).
[5] S. Ritsch, C. Beeli, H.-U. Nissen, T. Godecke, M. Scheffer, and R. Luck, Philos. Mag. Lett. **74**, 99 (1996).
[6] W. Steurer and K. H. Kuo, Acta Crystallogr., Sect. B: Struct. Sci. **B46**, 703 (1990).
[7] S. E. Burkov, Phys. Rev. Lett. **67**, 614 (1991).
[8] A. Yamamoto, Sci. Rep. Res. Inst. Tohoku Univ. A **42**, 207 (1996).
[9] K. Saitoh, K. Tsuda, and M. Tanaka, Philos. Mag. A **76**, 135 (1997).
[10] E. Cockayne and M. Widom, Phys. Rev. Lett. **81**, 598 (1998).
[11] P. J. Steinhardt, H.-C. Jeong, K. Saitoh, M. Tanaka, E. Abe, and A. P. Tsai, Nature (London) **396**, 55 (1998).
[12] P. A. Bancel, Phys. Rev. Lett. **63**, 2741 (1989); E. R. Abe and A. P. Tsai, *ibid.* **83**, 753 (1999).
[13] K. Saitoh, T. Yokosawa, M. Tanaka, and A.-P. Tsai, J. Electron Microsc. **48**, 105 (1999).
[14] S. J. Pennycook and D. E. Jesson, Phys. Rev. Lett. **64**, 938 (1990); D. E. Jesson and S. J. Pennycook, Proc. R. Soc. London, Ser. A **449**, 273 (1995).
[15] C. Beeli and S. Horiuchi, Philos. Mag. B **70**, 215 (1994).
[16] Y. Yan, S. J. Pennycook, and A.-P. Tsai, Phys. Rev. Lett. **81**, 5145 (1998).
[17] M. Boudard, M. de Boissieut, C. Janot, G. Heger, C. Beeli, H.-U. Nissen, H. Vincent, R. Ibberson, M. Audier, and J. M. Dubois, J. Phys.: Condens. Matter **4**, 10149 (1992).
[18] S. Ritsch, H.-U. Nissen, and C. Beeli, Phys. Rev. Lett. **76**, 2507 (1996).
[19] C. Janot and M. de Boissieu, Phys. Rev. Lett. **72**, 1674 (1994); C. Janot, J. Phys.: Condens. Matter **9**, 1493 (1997).
[20] The Burkov and Cockayne-Widom models use Cu/Co ordering as a source of matching rules.
[21] P. Gummelt, Geom. Dedicata **62**, 1 (1996).
[22] M. Duneau, in *Proceedings of the Fifth International Conference on Quasicrystals*, edited by C. Janot and R. Mosseri (World Scientific, Singapore, 1995), p. 116.
[23] P. J. Steinhardt and H.-C. Jeong, Nature (London) **382**, 433 (1996).

KC 02 01 02 0
ERKCM04 (Radiation Effects)

Refereed Journal Publications

Alves, A., C. J. McHargue, R. C. Silva, C. Jesus, O. Conde, M. F. da Silva, and J. C. Soares, "Structural and Magnetic Studies of Fe-Implanted Alpha-Al_2O_3," *Surface and Coatings Technology* **128-129**, 434-39 (2000).

Farrell, K., R. E. Stoller, P. Jung, and H. Ullmaier, "Hardening of Ferritic Alloys at 288°C by Electron Irradiation," *J. Nucl. Mater.* **279**, 77-83 (2000).

Kim, I.-S., J. D. Hunn, N. Hashimoto, D. L. Larson, P. J. Maziasz, K. Miyahara, and E. H. Lee, "Defect and Void Evolution in Oxide Dispersion Strengthened Ferritic Steels Under 3.2 MeV Fe^+ Ion Irradiation with Simultaneous Helium Injection," *J. Nucl. Mater.* **280**, 264-74 (2000).

Lee, E. H., T. S. Byun, J. D. Hunn, N. Hashimoto, and K. Farrell, "A Method to Study Deformation Mechanisms for Irradiated Steels Using a Disk-Bend Test," *J. Nucl. Mater.* **281**, 65-70 (2000).

Lee, E. H., J. D. Hunn, T. S. Byun, and L. K. Mansur, "Effects of Helium on Radiation Induced Defect Microstructure in Austenitic Stainless Steel," *J. Nucl. Mater.* **280**, 18-24 (2000).

Lee, E. H., J. D. Hunn, N. Hashimoto, L. K. Mansur, "Hardness and Defect Structures in EC316LN Austenitic Alloy Irradiated under a Simulated Spallation Neutron Source Environment using Triple Ion-Beams," *J. Nucl. Mater.* **278**, 266-72 (2000).

Lee, E. H. and L. K. Mansur, "Fe-15Ni-13Cr Austenitic Stainless Steels for Fission and Fusion Reactor Applications—Part I: Effects of Minor Alloying Elements On Precipitate Phases in Melt Products and Implication in Alloy Fabrication," *J. Nucl. Mater.* **278**, 1-10 (2000).

Lee, E. H. and L. K. Mansur, "Fe-15Ni-13Cr Austenitic Stainless Steels for Fusion and Fission Reactor Applications—Part II: Effects of Minor Elements On Precipitate Phase Stability During Thermal Aging," *J. Nucl. Mater.* **278**, 11-19 (2000).

Lee, E. H. and L. K. Mansur, "Fe-15Ni-13Cr Austenitic Stainless Steels for Fusion and Fission Reactor Applications—Part III: Phase Stability During Heavy Ion Irradiation," *J. Nucl. Mater.* **278**, 20-29 (2000).

McHargue, C. J., J. D. Hunn, E. Alves, M. F. daSilva, and J. C. Soares, "Influence of Oxygen Ion Implantation on the Damage and Annealing Kinetics of Iron-implanted Sapphire," *Nucl. Instr. and Meth. B* **166-167**, 188-92 (2000).

(*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

McHargue, C. J. and L. J. Romana, "Influence of Tin Implantation on the Damage and Annealing Kinetics of Sapphire," *Nucl. Instr. Methods Phys. Res. B* **166/167**, 193-97 (2000).

(A) Stoller, R. E., "The Role of Cascade Energy and Temperature in Primary Defect Formation in Iron," *J. Nucl. Mater.* **276**, 22-32 (2000).

(A) Stoller, R. E., "Evaluation of Neutron Energy Spectrum Effects and RPV Thru-Wall Attenuation Based on Molecular Dynamics Displacement Cascade Simulations," *Nucl. Eng. and Des.* **195**, 129-36 (2000).

(*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

ELSEVIER

Surface and Coatings Technology 128–129 (2000) 434–439

SURFACE & COATINGS TECHNOLOGY

www.elsevier.nl/locate/surfcoat

Structural and magnetic studies of Fe-implanted α-Al_2O_3

E. Alves[a,b,*], C. MacHargue[c], R.C. Silva[a,b], C. Jesus[a,b], O. Conde[d], M.F. da Silva[a,b], J.C. Soares[a,b]

[a] *Instituto Tecnológico e Nuclear, Department of Physics, EN 10, 2686-953 Sacavém, Portugal*
[b] *CFNUL, Av. Prof. Gama Pinto 2, 1699 Lisbon, Portugal*
[c] *Center for Materials Processing, University of Tennessee, Knoxville, TN 37996-2350, USA*
[d] *Department of Physics-FCUL, Ed. C1, Campo Grande, 1700 Lisbon, Portugal*

Abstract

High fluence implantation of Fe in α-Al_2O_3 and post-implantation annealing leads to the formation of various phases in the near-surface region whose nature depend upon the annealing atmosphere and crystal orientation. The structural evolution of the near-surface region and the magnetic properties of samples oriented with either the *c*-axis or *a*-axis normal to the surface and annealed in oxidizing and reducing atmospheres were studied with glancing incidence X-ray diffraction (GIXRD), transmission electron microscopy (TEM), Rutherford backscattering (RBS)/channelling and vibrating sample magnetometry (VSM). The as-implanted samples contained metallic iron clusters in a damaged matrix. This matrix became amorphous for fluences of 2×10^{17} Fe^+/cm^2 and above. The annealing in oxidizing atmosphere of the samples with the *a*-axis normal to the surface leads to the formation of a mixed layer containing α-Al_2O_3 and Fe_2O_3. The magnetization measurements indicate the presence of a large superparamagnetic contribution and a small ferromagnetic component. Annealing in a reducing atmosphere promotes the in-diffusion of iron and the magnetization decreases. A similar annealing of the sample with *c*-axis normal to the surface induces coalescence of precipitates in a discontinuous buried film of metallic iron parallel to the *c*-plane with a strong ferromagnetic signal and a small superparamagnetic contribution. The magnetization increases with the annealing temperature. © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Fe precipitates; Magnetization; Sapphire; Orientation dependence

1. Introduction

The formation of buried magnetic layers in oxides allows the development of very attractive structures for magnetic recording technologies [1]. Due to its high chemical and mechanical stability, sapphire (α-Al_2O_3) is one of the most important oxides for such applications. The production of magnetic buried layers can be achieved through the implantation of high fluences of magnetic ions. By choosing beam energy and ion fluence it is possible to control the depth and thickness of the magnetic layers. Indeed, MacHargue et al. observed the formation of elongated platelets in high fluence iron-implanted sapphire [2], the dimensions and distribution of which depend on the implantation conditions (e.g. energy and ion fluence). They noticed also that the final state of the system was influenced by the annealing conditions [3,4]. In all this work no evidence was found for any effect of the substrate orientation on the iron behaviour. Furthermore, although some studies on the magnetic properties of iron implanted sapphire have been done [5,6], the influence of substrate orientation, annealing conditions and ion fluence on these properties was not investigated to date.

In this study we report on the magnetic and struc-

* Corresponding author. Tel: +351-1-9946086; fax: +351-1-9941525.
E-mail address: ealves@itn1.itn.pt (E. Alves).

0257-8972/00/$ - see front matter © 2000 Elsevier Science S.A. All rights reserved.
PII: S0257-8972(00)00578-8

tural behaviour of iron implanted *a*- and *c*-cut α-Al_2O_3. Samples implanted with different fluences were annealed in oxidizing and reducing atmospheres. The structural changes were followed by glancing incidence X-ray diffraction (GIXRD), transmission electron microscopy (TEM) and Rutherford backscattering spectrometry (RBS), while a vibrating sample magnetometer (VSM) was used to measure the magnetic behaviour. In this way, we could follow in detail the relationship between the microstructural evolution and magnetic properties.

2. Experimental

High purity α-Al_2O_3 single crystals with optically polished surfaces cut perpendicularly to the $\langle 0001 \rangle$ axis (*c*-samples) and $\langle 11\bar{2}0 \rangle$ axis (*a*-samples) were implanted with 160 keV Fe^+ ions to fluences in the range of 1×10^{17} to 4×10^{17} at./cm^2. The samples were tilted 7° with respect to the ion beam to avoid channelling and held at room temperature. The range of the implanted ions was 72 nm. Thermal annealings were carried out at 700°C, 800°C and 1100°C for 1 h in reducing (4% hydrogen–argon mixture) and oxidizing atmospheres.

RBS/channeling studies were performed with a 1.6 and 2 MeV He^+ beam after implantation and after each stage of thermal annealing. The backscattered particles were detected at 160° and close to 180° using silicon surface barrier detectors with resolutions of 13 and 18 keV, respectively. In order to minimize the effects of charge accumulation on the Al_2O_3 surface during analysis, the crystal was masked with a thin Ni foil with a 4-mm hole. A Siemens D-5000 spectrometer, allowing both conventional (Bragg–Brentano) and grazing incidence geometries, was used to perform the X-ray diffraction studies. The diffraction patterns were recorded using the Cu Kα radiation at 0.5° angle of incidence to the specimen surface. The identification of crystalline phases was done using the JCPDS database cards [7]. Transmission electron microscopy observations were done with a Hitachi HF-2000 field emission electron microscope operating at 200 kV. The magnetization curves were obtained with a Digital Measurement System magnetometer, model DMS 880.

3. Results

3.1. (0001) α-Al_2O_3

The RBS spectra of Fig. 1 shows the annealing behaviour under reducing conditions of two samples implanted with 1×10^{17} Fe^+/cm^2 and 2×10^{17} Fe^+/cm^2. We can see that no significant change occurs

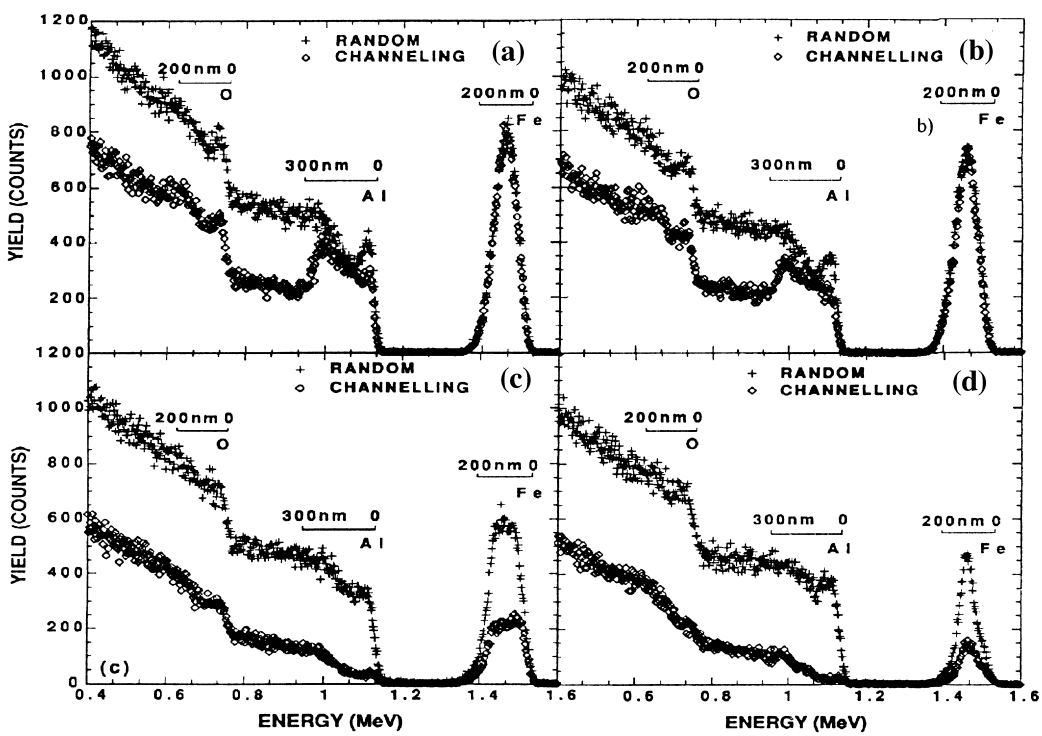


Fig. 1. Aligned and random RBS spectra of *c*-axis-oriented sapphire implanted with 2×10^{17} Fe^+/cm^2 before (a) and after the annealing at 700°C (b) and 1100°C (c) in a reducing atmosphere. The results of a sample implanted with 1×10^{17} Fe^+/cm^2 were also shown in (d) for comparison.

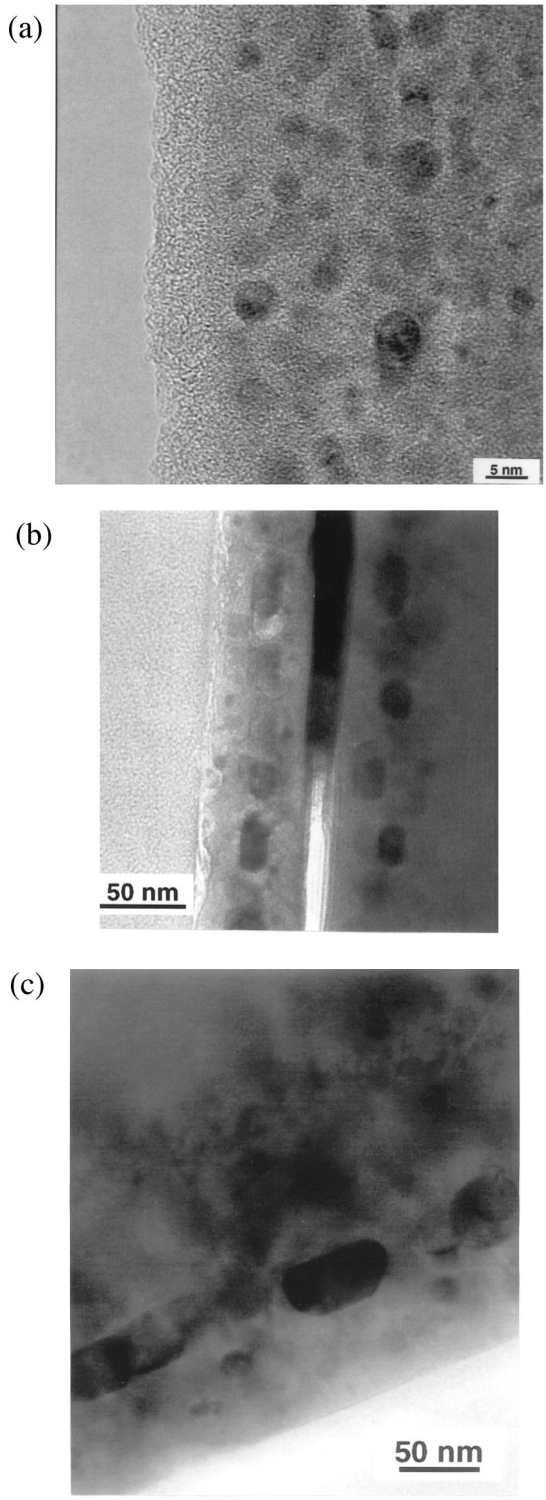


Fig. 2. Micrographs of the as-implanted (a) and annealed at 1100°C (b) in reducing atmosphere of *c*-axis-oriented samples implanted with 2×10^{17} Fe^+/cm^2. The micrograph of a sample implanted with 1×10^{17} Fe^+/cm^2 and submitted to the same annealing is shown in (c).

up to 700°C. The annealing at 1100°C produces a great reduction of the implantation damage, visible through the decrease of the yield on the sapphire aligned spectrum. However, the as-implanted Gaussian profile of iron became rectangular and the maximum peak concentration decreased but no loss of iron was detected. The reduction observed in the iron yield in the aligned spectrum reveals that approximately 50% of the ions were incorporated at substitutional lattice sites or interstitial sites shadowed by the matrix atoms along this direction. These changes are compatible with the formation of iron precipitates coherently aligned with the $\langle 0001\rangle$ axis of α-Al_2O_3.

The presence of the precipitates is revealed in the micrographs of the cross-sectioned samples of Fig. 2. Precipitates with diameters in the range of 5–20 nm were found in samples implanted with 2×10^{17} Fe^+/cm^2 even after implantation (Fig. 2a). This is in agreement with previous studies by Ren et al. who showed that the precipitates consist mainly of crystalline body-centred cubic α-Fe [8]. The annealing induces the coalescence of the precipitates leading to the formation of two distinct regions. At the front and rear edges of the iron profile we see a disperse distribution of small precipitates. Large precipitates with a rectangular shape were found at a depth close to the range where the maximum Fe concentration was located forming a discontinuous band parallel to the surface. The length of the precipitates increases with the fluence of the implanted ions. The formation of such precipitates suggests that iron diffuse faster parallel to the *c*-plane.

Fig. 3 shows the influence of the annealing temperature on the magnetic behaviour of the sapphire sample implanted with 2×10^{17} Fe^+/cm^2. For comparison we show the result obtained for a sample implanted with 1×10^{17} Fe^+/cm^2 after annealing at 1100°C, since at lower temperatures only the diamagnetic contribution of sapphire is seen. After implantation the curve of the sample implanted with 2×10^{17} Fe^+/cm^2 shows the presence of a diamagnetic contribution of the sapphire and the presence of a small ferromagnetic signal responsible for the hysteresis of the curve. The magnetization increases with the annealing temperature, with the ferromagnetic contribution being dominant with a small superparamagnetic signal, which causes the 'S' shape of the curve. The increase of the magnetization with the annealing temperature and the dose is clearly correlated with the growing of the precipitates due to the coalescence mechanism.

3.2. $(11\bar{2}0)$ α-Al_2O_3

Fig. 4 compares the annealing behaviour of sapphire implanted with 1×10^{17} Fe^+/cm^2 and 4×10^{17} Fe^+/cm^2 annealed in reducing and oxidizing atmospheres. For the samples implanted with the lower dose iron starts to diffuse to the surface during the annealing at 800°C in both annealing conditions (not shown).

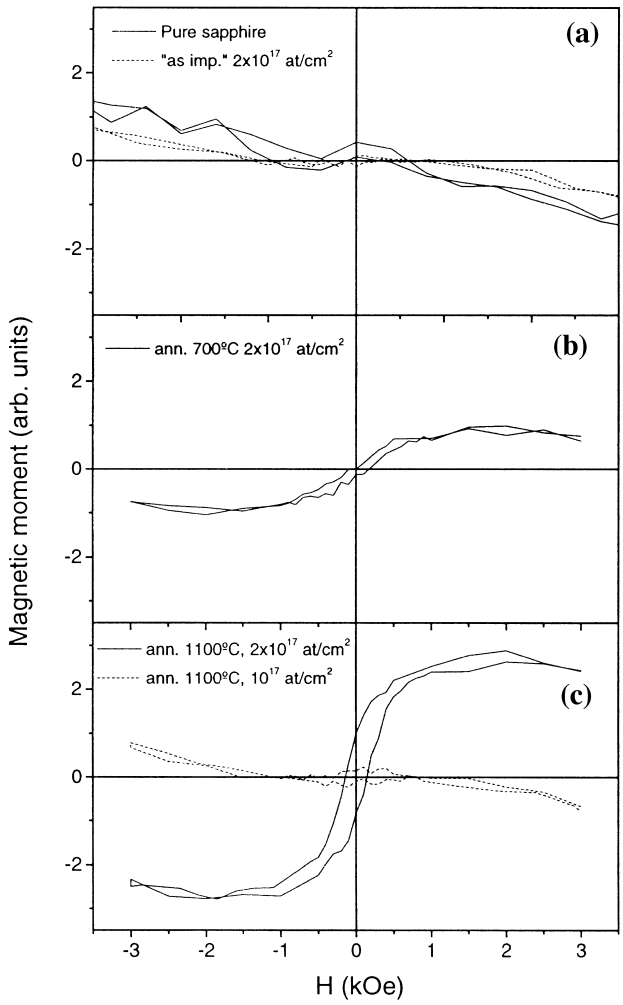


Fig. 3. Magnetization curves for the virgin and as-implanted *c*-axis-oriented sapphire with 2×10^{17} Fe^+/cm^2 (a) and after annealing at 700°C (b) and 1100°C (c) in reducing atmosphere. The dashed curve in (c) was obtained for a sample implanted with 1×10^{17} Fe^+/cm^2.

The amount of iron at the surface for the sample annealed in an oxidizing atmosphere increased with the annealing temperature (Fig. 4b); while for the samples annealed in reducing atmosphere a large redistribution of iron occurred and approximately 60% of the Fe was lost (Fig. 4c). Under oxidizing conditions the presence of free oxygen allows the formation of iron oxides at the surface and may explain this different behaviour. The changes in the Fe profile agree with the assumption made above that the diffusion along the *c*-plane is much easier (notice that in these samples the *c*-plane is perpendicular to the surface). The damage recovery is also different as is evident from the large difference in the minimum yields along the $\langle 11\bar{2}0\rangle$ axis of the samples annealed in the oxidizing (35%) and reducing atmospheres (3%). Combining these results with those found for the *c*-samples we conclude that the microstructure developed in the implanted region inhibits the damage recovery as was also observed in other works [9]. The presence of Fe_2O_3 in the samples annealed in oxidizing atmosphere was confirmed by GIXRD. Fig. 5 shows the X-ray spectra of the samples implanted with 4×10^{17} Fe^+/cm^2 after implantation and after annealing at 1100°C in an oxidizing atmosphere, respectively. The presence of metallic iron precipitates in the samples was observed immediately after implantation in agreement with TEM observations. The annealing in oxidizing atmosphere induces the formation of hematite (Fe_2O_3), as shown by the X-ray diffraction spectrum. This transformation was also noticed by the change in coloration from metallic to

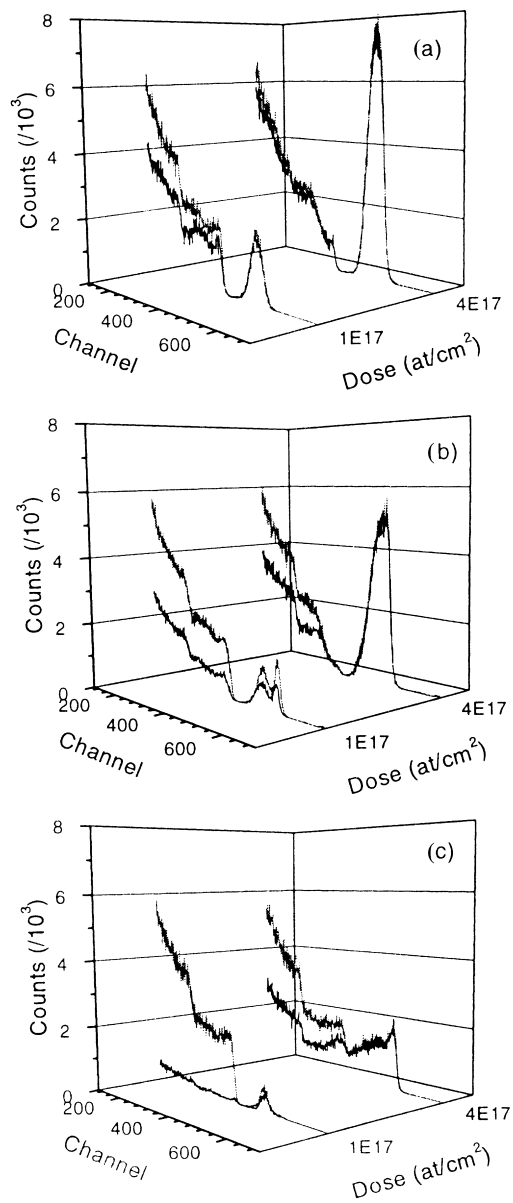


Fig. 4. Aligned and random RBS spectra of *a*-axis-oriented sapphire implanted with 1×10^{17} Fe^+/cm^2 and 4×10^{17} Fe^+/cm^2 before (a) and after the annealing in reducing (b) and oxidizing atmosphere (c).

reddish brown. The X-ray diffraction spectrum of the sample annealed in a reducing atmosphere does not reveal the presence of any iron compound.

Fig. 6 shows the magnetization curves measured for the samples implanted with 4×10^{17} Fe^+/cm^2 before and after different annealings. The curve obtained after implantation shows the formation of a ferromagnetic system. This is due to the high dose of iron present at the surface allowing the creation of an almost continuous nearly metallic layer. The magnetization decreases drastically after the annealing at 1100°C in reducing atmosphere. One explanation for this decrease could be the redistribution of the iron ions in the implanted layer as observed by RBS (Fig. 4c), which was accompanied by the loss of the implanted Fe. The remaining iron is more diluted, dispersed in small precipitates or isolated ions, leading to a decrease of the magnetic component of the magnetization curve. This redistribution and loss of iron, not found in the *c*-cut samples, explains the totally different magnetic properties observed in the two kinds of samples after annealing. The anisotropy in the iron diffusion could be the reason for that. In the samples annealed in oxidizing atmosphere the ferromagnetic component disappears after the first annealing and only the diamagnetic contribution of sapphire was observed.

4. Conclusions

Implantation of high fluences of iron in sapphire leads to the formation of metallic precipitates during the implantation. Above 2×10^{17} Fe^+/cm^2 a layer with metallic characteristics develops at the surface, leading to a strong ferromagnetic behaviour. The microstructures developed during annealing are strongly dependent on the substrate orientation and annealing atmosphere. For (0001) α-Al_2O_3 samples annealed in reducing atmosphere we observe the formation of elongated metallic iron precipitates parallel to the surface whose dimension increases with ion fluence. The samples display ferromagnetic behaviour after annealing at 1100°C. The ferromagnetism in the $(11\bar{2}0)$ α-Al_2O_3 samples decreases after similar annealing due to diffusion and loss of iron. Annealing in an oxidizing atmosphere leads to the formation of non-magnetic mixed oxide compounds in the implanted region.

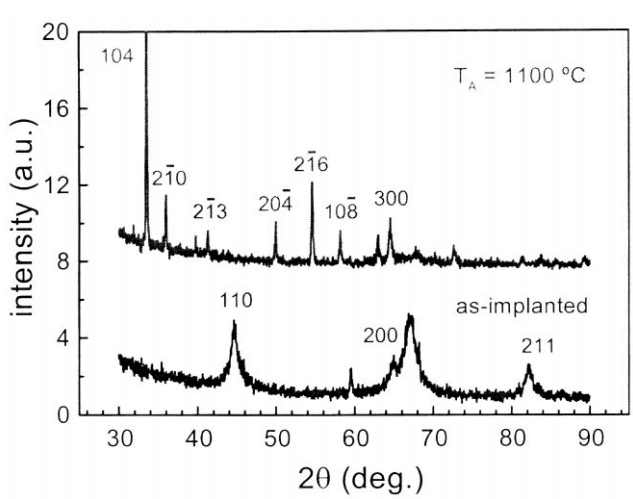


Fig. 5. XRD spectra for *a*-axis-oriented sapphire implanted with 4×10^{17} Fe^+/cm^2 before and after annealing at 1100°C in an oxidizing atmosphere.

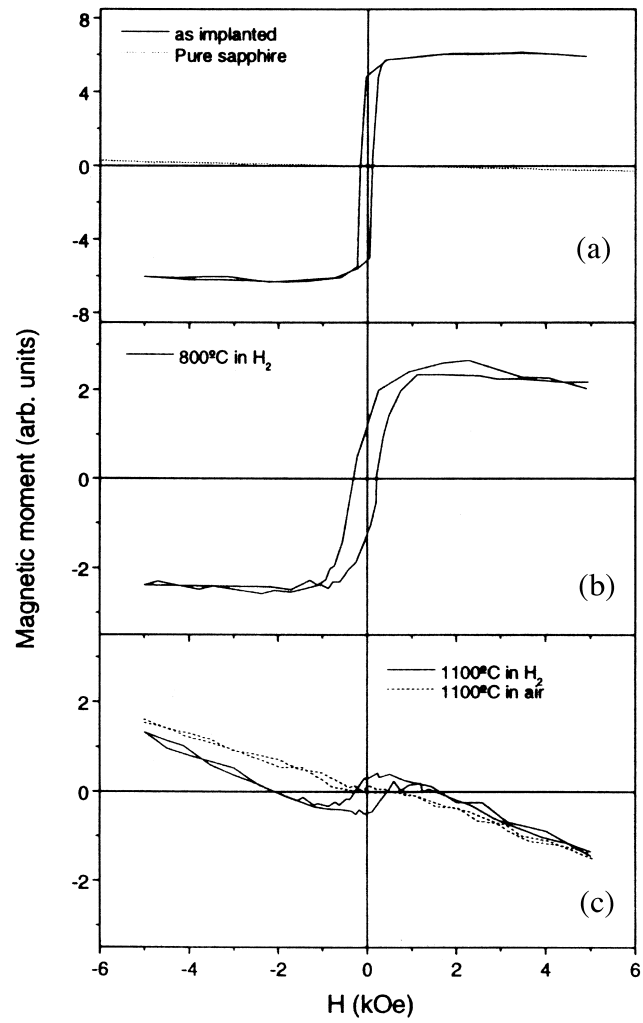


Fig. 6. Magnetization curves for *a*-axis-oriented sapphire after implantation with 4×10^{17} Fe^+/cm^2 (a) and after annealing at 800°C (b) and 1100°C (c) in a reducing atmosphere. The dashed curve in (c) was obtained for a sample annealed in an oxidizing atmosphere.

Acknowledgements

This work was partially supported by PRAXIS XXI under contract C/CTM/12067/1998 and the US Department of Energy, Division of Materials Sciences

under contract DE-AC05-96OR22464 with the Lockheed Martin Energy Research Corporation.

References

[1] T. Miyazaki, N. Tezuka, J. Magn. Magn. Mater. 139 (1995) L231.
[2] C.J. MacHargue, S.X. Ren, P.S. Sklad, L.F. Allard, J. Hunn, Nucl. Instrum. Methods B 116 (1996) 173.
[3] C.J. MacHargue, G.C. Farlow, P.S. Sklad et al., Nucl. Instrum. Methods B 19/20 (1987) 813.
[4] C.J. MacHargue, P.S. Sklad, C.W. White, J.C. McCallum, A. Perez, G. Marest, J. Mater. Res. 6 (1991) 2160.
[5] M. Ohkubo, T. Hioki, J.-i. Kawamoto, J. Appl. Phys. 62 (1987) 3069.
[6] I. Sakamoto, S. Honda, H. Tanoue, N. Hayashi, H. Yamane, Nucl. Instrum. Methods B 148 (1999) 1039.
[7] Joint Committee on Powder Diffraction Standards, Powder Diffraction File, ASTM, Philadelphia, PA, 1992.
[8] S.X. Ren, C.J. MacHargue, L.F. Allard et al., Mater. Res. Soc. Symp. Proc. 373 (1995) 305.
[9] C.J. MacHargue, E. Alves, M.F. da Silva, J.C. Soares, Nucl. Instrum. Methods B 148 (1999) 730.

ELSEVIER

Journal of Nuclear Materials 279 (2000) 77–83

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Hardening of ferritic alloys at 288°C by electron irradiation ☆

K. Farrell [a,*], R.E. Stoller [a], P. Jung [b], H. Ullmaier [b]

[a] *Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831, USA*
[b] *Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany*

Received 22 February 1999; accepted 4 November 1999

Abstract

Tensile specimens of a commercially pure iron, an iron–copper alloy and two ferritic pressure vessel steels, were irradiated at 288°C with 2.5 MeV electrons to doses of 2.82×10^{23} and 9.35×10^{23} e^-/m^2, corresponding to calculated atomic displacement doses of 9.53×10^{-4} and 3.16×10^{-3} dpa, respectively. Tensile tests at room temperature showed dose-dependent increases in yield stress and ultimate tensile stress and reductions in uniform elongation, compatible with literature data for A533B steel neutron-irradiated at 288°C to similar displacement levels. No systematic effect of copper content was discerned in these electron irradiations, contrary to expectations based on neutron irradiations. For the limited dose range over which direct comparison can be made, it is concluded that the hardening efficiency of electron irradiations per unit dpa at 288°C is similar to that for neutron irradiations. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Radiation hardening in metals is caused by clusters of point defects introduced in the metal lattice by atomic displacements; such clusters may be complexed with foreign atoms. Although most experimental investigations have involved neutron irradiation, it has long been known that irradiations with energetic electrons can also cause hardening [1,2]. Electron displacements are involved in nuclear reactor irradiations, where electrons are generated via the processes of Compton scattering and electron–positron pair production by high energy gamma rays emanating from materials in the core region. The relative contributions of various sources of atomic displacements such as fast neutrons, thermal neutron capture recoils, energetic particles created by transmutation reactions and gamma-induced electrons, have been quantified as a function of irradiation environment [3]. The contribution of gamma (electron)-induced displacements to the total atomic displacements responsible for hardening and embrittlement in most reactor pressure vessels (RPVs) is usually negligible compared to the displacements from fast neutrons [3,4]. The atomic displacement cross section for 1 MeV gamma rays in iron alloys is only about 1 barn versus about 1500 barns for 1 MeV neutrons. Since, in most reactors, gammas and neutrons with energies greater than about 1 MeV are produced in roughly equal amounts and arrive at the vessel in almost equal quantities, the damage from gamma displacements to the RPV is relatively small. However, in special situations where there is a long path of water between the core and the RPV, the neutrons will be attenuated more strongly than the gamma rays and the fraction of atomic displacements in the RPV contributed from gamma (electron) displacements will increase and in exceptional circumstances may exceed that from the neutrons [5–7]. Such was the case [5,6] for the RPV of the High Flux Isotope Reactor (HFIR).

In neutron irradiations the point defect clusters responsible for hardening are considered to develop in two ways. One way is directly within the large displacement cascades generated by the energetic neutrons and knock-on atoms. These clusters are left as almost instantaneous

☆ The submitted manuscript has been authored by a contractor of the US Government under contract No. DE-AC05-96OR22464. Accordingly, the US Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

* Corresponding author.

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(99)00270-6

debris from the creation and collapse of the cascades. The other way clusters are formed is by a more gradual process of migration and coalescence of freely migrating point defects (fmpds). The fmpds comprise those surviving from the cascades and those resulting from relatively 'soft' displacement events, i.e. low pka energy events such as the gamma displacements and the recoils from capture of thermal neutrons. Since both of these cluster formation routes occur inseparably in fast neutron irradiations, it is difficult to discriminate the relative effects of the cascade debris and the fmpd clusters on irradiation hardening.

A discrimination of the hardening efficiencies of these two sources of clusters is fundamentally interesting because there are grounds for believing that a greater fraction of the point defects from soft displacements may avoid recombination than point defects from large cascades. This belief stems from the recognition that most of the point defects generated in energetic displacement cascades are mutually annihilated by virtue of the close spatial correlation of vacancies and self-interstitial atoms within the cascade. In soft displacements the fraction of point defects per unit initial displaced atom (dpa) that survives immediate local annihilation is much greater. These 'extra' fmpds from the soft irradiations are acknowledged to be the reason why bombardments with 1 MeV electrons and protons cause much greater degrees of void swelling and radiation-assisted solute segregation per dpa than do irradiations with 1 MeV neutrons or heavy ions. By corollary, it has been suggested [8] that the higher survival fractions of point defects in soft irradiations might result in more fmpd clusters and greater hardening per unit dpa. However, analysis of the HFIR RPV embrittlement [6] indicated that the irradiation strengthening efficiency of the gamma (electron) displacements per dpa, seemed to be the same as that of fast neutrons.

One way to test these hypotheses and deductions is to compare the hardening responses of neutron irradiated materials with those irradiated to equal atomic displacement levels in an electron accelerator. In electron irradiations made at energies below about 5 MeV, the energy of the atomic recoils are too low to create true cascade-like behavior and most point defects are created as fmpds in the form of isolated Frenkel pairs. For example, if the electron displacment cross sections of Oen [9] are applied for iron with a 40 eV displacement threshold, the cross sections at 5 MeV are about 43 and 69 barn for primary and total displacments, respectively. At 2 MeV, the cross section for primary displacements is about 31 b and 35 b for total displacements. Since primary Frenkel pair production dominates at these electron energies, there are few clusters formed directly as a result of the displacement events and any hardening must ensue solely from clusters formed by diffusion and agglomeration of fmpds.

The first comparison of electron- and neutron-irradiation hardening was made on copper by Makin and Blewitt [2] and suggested a low hardening efficiency for electron displacements. More recently, tensile specimens of the archive steel of the HFIR vessel and other ferritic alloys were subjected to electron bombardments at about 60°C in a particle accelerator and their properties were compared with those for the same materials irradiated to the same displacement levels ($1.4 \times 10^{-3} - 5.3 \times 10^{-3}$ dpa) with neutrons [10]. The electron hardening per dpa was found to be equal for the electron and neutron irradiations. These latter hardening experiments were performed at a temperature of about 60°C, in keeping with the low operating temperature of the HFIR RPV. For the RPVs of commercial power reactors a temperature of 288°C is more relevant. Hence it is desired to know the hardening efficiency of electron displacements (fmpds) at 288°C. This paper describes the results of new electron irradiations made on ferritic alloys at 288°C.

2. Experimental

Chemical compositions of the four test materials are given in Table 1. Materials 2A2 and 1B2 are, respectively, pure iron and a Fe–0.28%Cu alloy kindly provided by J.R. Hawthorne of Materials Engineering Associates. A212B is the archive steel of the HFIR RPV. A533B is the reference HSST-Plate 02 correlation monitor steel used in numerous investigations of commercial RPV steels. Two flat SS-3 type tensile specimens of each material, with gauge sections of 1.52×0.76 mm^2 and gauge length of 7.6 mm, were mounted in an upright array in the target chamber of the van-de-Graaff accelerator at the Institut für Festkorperforschung, Forschungszentrum, Julich. The specimens were irradiated with an incident beam of 2.5 MeV electrons through a thin foil window for periods of 57.5 and 222 h to electron doses of 2.82×10^{23} and 9.35×10^{23} e$^-$/m^2. The average current density in the few mm diameter beam spot was 0.2 A/m^2 and the beam spot was rastered horizontally and vertically over the specimen gauge areas (height $\times$ width $= 12.5 \times 12$ mm^2) at a rate of 500 scans per second, avoiding Lissajou effects. To homogenize the small gradient in damage through the thickness of the specimens due to their attenuation of the beam, the specimen assembly was rotated 180° halfway through each run. The electron energy and atomic displacement rates were calculated as a function of depth through the specimens by accounting for both ionization and radiative (Bremsstrahlung) losses. A simple finite difference program was written which implemented well-known expressions for these loss terms [11]. After accounting for beam energy loss in

Table 1
Chemical compositions (wt%) and heat treatments of the test materials[a]

Element	2A2(Fe)[b]	1B2(Fe–0.28Cu)[b]	A212B[c]	A533B[d]
C	0.013	0.013	0.26	0.23
Al	<0.005	0.007	0.07	
Co			0.015	
Cr			0.075	0.04
Cu	<0.005	0.28	0.15	0.14
Mn	0.018	0.013	0.85	1.55
Mo			0.02	0.53
Nb			<0.001	
Ni	0.018	0.012	0.20[e]	0.67
Si			0.29	0.20
Sn			0.02	
Ti			0.01	
V			0.0005	0.003
W			<0.005	
Zr			<0.001	
P	0.003	0.004	0.006	0.009
S			0.04	0.014
As			0.007	
B			<0.0005	
N			0.0060	
O			0.0024	

[a] AC = air-cooled; FC = furnace-cooled; WQ = water-quenched.
[b] Anneal 1/2 h at 914–926°C, AC.
[c] Anneal 1093°C; reheat 3 h at 899°C, WQ; reheat 663°C, AC.
[d] Anneal 4h, at 857–885°C, AC; 4 h at 649–677°C, AC; 20 h at 607–635°C, FC to 315°C, AC.
[e] Believed to be too high; independent analysis at another laboratory showed 0.09.

the 25 μm-thick Hastelloy beam window, the thickness-weighted average electron energy in the specimens was 1.98 MeV. Using the calculations of Oen [9] and the ASTM-recommended displacement threshold of 40 eV for iron [12], a displacement cross section of 33.8 b is obtained. Based on total electron fluences of 2.82×10^{23} and 9.35×10^{23} e^-/m^2, the calculated atomic displacement doses were 9.53×10^{-4} and 3.16×10^{-3} dpa, respectively, at an average displacement rate of 4.1×10^{-9} dpa/s.

Cooling was provided by water-chilled helium gas pumped across the faces of the specimens via five nozzles located on each vertical edge of the specimen assembly. Temperatures were measured with a scanning infrared pyrometer normalized to the readings from a thermocouple welded to the gauge section of one of the specimens and were held in the range 278–298°C.

Unirradiated control specimens were sealed in evacuated glass ampoules and were furnace-heated at 288°C for periods matching the irradiation periods. Tensile tests were made in a screw-driven Tinius–Olsen machine under computer control at room temperature and at a strain rate of 1.1×10^{-3} s^{-1}. Tensile yield strengths were read at the lower yield point inflection or at 0.2% strain offset when there was no yield inflection.

3. Results

The full set of data is given in Table 2. The superscripts s and l for the thermal controls indicate the short and long aging periods corresponding to the two irradiation periods. During irradiation slight discoloration of the specimens occurred. The discoloration was removed from one of each pair of specimens by dry sanding with 600 grit paper. No effects of the discoloration on tensile properties were discerned. Thermal aging at 288°C did not alter the properties of the control specimens, as is evident by comparing the control data with previous data we have measured on unaged materials labeled various in the first row for each material.

The electron irradiations caused well-defined and reproducible hardening and loss in ductility. Examples of typical stress–elongation curves are shown in Fig. 1 for the pure iron and the A533B steel. A peculiarity of the curves for all four materials and especially for the two steels, is that the increases in UTS are roughly comparable with those for the yield strengths. Such behavior is in contrast to curves for neutron irradiated materials which normally exhibit significantly less increase in UTS than in yield strength.

The increases in yield stress due to electron irradiation at 288°C are dose dependent and show some response to chemical composition and/or metallurgical

Table 2
Tensile properties

Mater.	Spec. ID	dpa	Surface condition	YS (MPa)	UTS (MPa)	Unif. elong. (%)	Total elong. (%)
Fe	Various	0	Unaged	166–171	272–283	22.4–24.6	36.7–41.2
	K36	0	As-aged[s]	158	276	25.3	42.0
	K37	0	Sanded[s]	160	274	25.0	39.3
	K30	9.5E−4	As-irrad.	200	287	21.5	34.9
	K31	9.5E−4	Sanded	207	303	20.7	36.5
	K35	0	Sanded[l]	151	269	23.8	38.9
	K34	0	As-aged[l]	201	297	18.1	37.8
	K32	3.2E−3	As-irrad.	255	337	14.7	29.5
	K33	3.2E−3	Sanded	249	338	15.3	29.5
Fe–0.25Cu	Various	0	Unaged	221–227	313–322	14.0–16.0	27.3–29.3
	H41	0	As-aged[s]	214	321	15.5	24.9
	H42	0	Sanded[s]	214	321	17.9	30.5
	H32	9.5E−4	As-irrad.	273	367	19.7	30.4
	H36	9.5E−4	Sanded	274	374	20.7	34.4
	H39	0	As-aged[l]	229	328	19.7	31.1
	H40	0	Sanded[l]	214	319	17.4	31.3
	H37	3.2E−3	As-irrad.	307	397	14.6	25.5
	H38	3.2E−3	Sanded	304	402	15.9	28.7
A212B	Various	0	Unaged	318–345	542–569	14.1–17.7	24.7–30.1
	A24	0	As-aged[s]	325	544	17.0	22.9
	A105	0	Sanded[s]	358	577	14.3	25.3
	A39	9.5E−4	As-irrad.	389	584	13.4	23.1
	A89	9.5E−4	Sanded	394	593	13.6	23.7
	A104	0	As-aged[l]	356	566	14.8	25.1
	A23	0	Sanded[l]	333	558	16.7	27.9
	A102	3.2E−3	As-irrad.	441	632	13.1	23.3
	A103	3.2E−3	Sanded	445	633	13.5	23.9
A533B	Various	0	Unaged	433–457	594–618	10.2–12.3	16.2–19.7
	X11	0	As-aged[s]	459	614	10.9	20.4
	X12	0	Sanded[s]	462	616	10.7	20.3
	X1	9.5E−4	As-irrad.	489	634	10.9	18.5
	X2	9.5E−4	Sanded	492	637	10.7	18.2
	X9	0	As-aged[l]	463	616	11.4	20.9
	X10	0	Sanded[l]	471	626	10.9	18.3
	X3	3.2E−3	As-irrad.	505	649	9.3	17.5
	X4	3.2E−3	Sanded	517	654	10.0	16.9

condition. In Fig. 2(a) the increases are plotted versus the yield strengths of the unirradiated material. The increases are consistently larger for the higher electron dose. They are also seemingly insensitive to the original strength for the Fe, Fe–0.28Cu and A212 materials, but are notably lower for the highest strength material, the A533B steel. Expressed as percentage increases in yield strength they are progressively smaller with increasing original yield strength, Fig. 2(b). These observations imply that the degree of electron radiation damage is not strongly sensitive to chemical composition and microstructure for the first three materials but is reduced in the A533B material, presumably because of its higher content of alloying elements and/or its metallurgical condition. The A533B steel has a much finer grain size and a more uniformly distributed carbide phase than the other materials.

4. Discussion

The observation that the UTS of each material is increased to roughly the same extent as the increase in yield strength concurs with a recent report of similar behavior in ferritic alloys electron irradiated at 35–60°C [13]. There it was suggested that the post-yield strain hardening mechanism in electron irradiated alloys is similar to that for unirradiated materials and is different than that for neutron irradiations. The yield strengths for electron- and neutron-irradiated materials were similar.

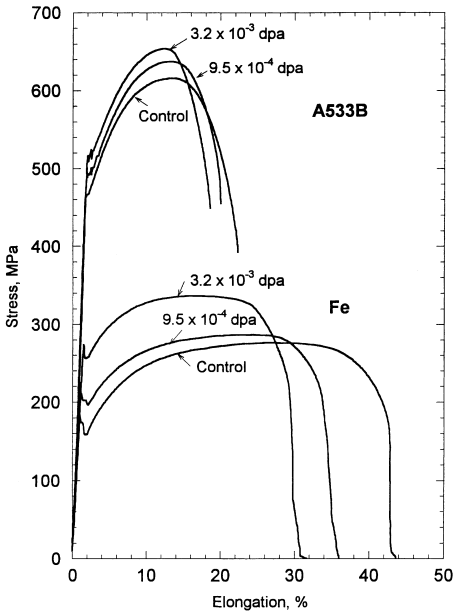


Fig. 1. Examples of tensile test curves for iron and A533B steel electron irradiated at about 288°C and tested at room temperature.

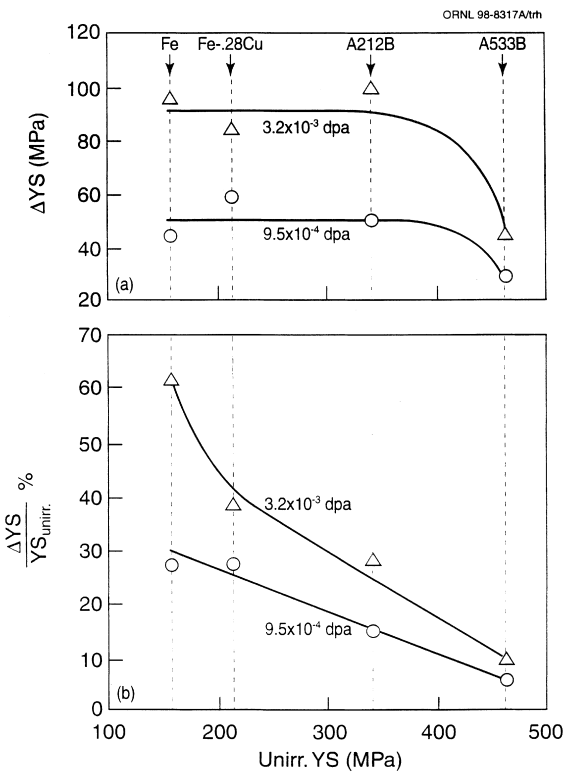


Fig. 2. (a) Increases in yield strengths versus unirradiated yield strengths; (b) ratios of increases in yield strengths to unirradiated yield strengths versus unirradiated yield strengths.

Of the four alloys studied here, only the A533B steel has a sufficient data base of neutron irradiations at 288°C to permit an assessment of the relative efficiencies of radiation hardening by electrons and neutrons at 288°C. In Fig. 3 the electron-induced increases in yield strengths for the A533B steel from the present work are compared with the yield strength changes in neutron irradiated tensile specimens made from the same HSST-02 plate material and from other A533B plates, all irradiated with neutrons at 288°C [14,15]. It is evident that the electron hardening for the A533B steel is the same as the neutron hardening at the same dpa levels. This agrees with the observation made for irradiations of steels at temperatures <100°C [10]. The neutron-irradiation data in Fig. 3 include results for the HSST-02 plate irradiated in test reactors and other A533B plate data with similar copper content from commercial reactor surveillance programs (PR-EDB). The ORR and BSR designations refer to the Oak Ridge Research and Bulk Shielding reactors, respectively. The surveillance data cover a range of fast fluxes ($E > 1$ MeV) from about 6.7×10^{12} to 1.6×10^{15} n/m^2/s, while the test reactor data fluxes are in the range of 2×10^{15}–1.3×10^{16} n/m^2/s. The agreement within the neutron data for neutron fluxes in this range is consistent with both theoretical and other analyses of radiation-induced hardening in these materials [16–19]. These same data support the comparison of the HSST-02 plate and other materials with similar copper content since fluence and copper are the variables that contribute most significantly to the hardening. The 2.5 MeV electron-induced displacement rate corresponds to a higher effective neutron flux of about 4×10^{16} n/m^2/s ($E > 1$ MeV), which is slightly above the flux-independent regime identified in

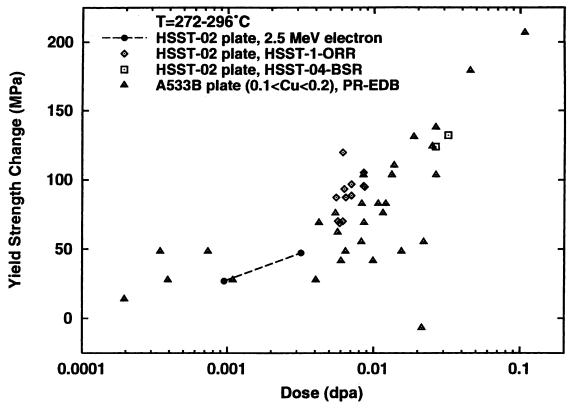


Fig. 3. Comparison of increases in yield strength for the electron-irradiated A533B steel (HSST-02 plate) with A533B steels neutron irradiated at 272–296°C. The neutron data are taken from [14,15]. The neutron fluences in [14,15] were originally in units of n ($E > 1$ MeV) /cm^2; they were converted to dpa by multiplying them by 1.5×10^{-21} dpa/unit fluence.

Refs. [17,18]. Since the predicted impact of the higher displacement rate would be only a modest increase in the hardening [17], the comparison of the electron- and neutron-irradiated data is sensible.

It is also obvious that the dose range for which agreement between electron and neutron hardening is demonstrated in Fig. 3 is quite narrow and is at low dpa levels. This is a regrettable shortcoming of the electron irradiations, which would require unreasonably lengthy irradiation times to attain displacement levels of 0.01–0.1 dpa. This severely limits the generality of our apparent conclusion that there is apparently a one-to-one relationship between the hardening efficiencies of electron and neutron induced displacements in these materials. Unfortunately, there is a lack of definitive data from other sources to confirm or refute this equivalence. We were able to find only two publications containing comparisons of radiation hardening data for iron-based materials irradiated with electrons and neutrons at about 288°C [20,21].

In contrast with our results, the data from both Refs. [20,21] show significantly greater hardening per dpa for the electron irradiations. However, a direct comparison of these data with ours is confounded by several factors. First, the authors of Ref. [20] indicate that it is only the hardening rate that increases under electron irradiation and that saturation of hardening is reached earlier. They state that the saturation level of hardening is higher under neutron irradiation. We have not observed saturation in our experiments. Second, only hardness data are available in Refs. [20,21], while the data shown in Fig. 3 are all tensile data. There are well established correlations that can be applied to convert Vickers hardness changes to equivalent yield strength changes, typically ΔYS (Mpa) $\approx$ 3.3–3.6 ΔH_v. However, when such a correlation is applied to the hardness data from Ref. [21], the resulting yield strengths are much higher than those measured in the current experiment at the same doses. This result may partially arise from the lower irradiation temperatures employed in the experiment reported in Ref. [21], although this is contrary to the weak irradiation temperature dependence observed in that experiment.

A more important difference between our data and that reported in Refs. [20,21] is that their materials were all Fe–Cu binary alloys that were solution annealed and rapidly cooled to maintain a high copper supersaturation in the solid solution. Copper impurities have long been recognized as a cause of accelerated hardening and embrittlement in ferritic steels neutron irradiated at 288°C and the extra hardening and embrittlement are attributed to the formation of a fine dispersion of Cu-rich clusters or precipitates. In the materials used in Refs. [20,21] copper-rich precipitates dominate the hardening, as evidenced by alloys of very low copper levels in the experiments which showed relatively minor hardening. During electron irradiations of supersaturated alloys the high concentrations of fmpds introduced by the electrons can be expected to increase the migration of copper atoms and thereby encourage the formation of copper-rich particles and associated hardening. In neutron irradiations the lesser quantities of fmpds and the destructive action of displacement cascades on precipitate nuclei will no doubt give fewer precipitate particles per dpa and less hardening. It is now generally agreed that a steel's sensitivity to copper impurity depends more on the amount of copper in solid solution than on the total copper content. The alloys in the present experiments were not heat treated to maximize the dissolved copper concentrations and we found no sensitivity to copper content. The increment of radiation hardening in our Fe–0.28Cu alloy is the same as in the iron which contains almost no copper and had the same heat treatment. This equality of hardening illustrates unambiguously that copper-rich precipitates and clusters of other solutes are not likely to be the sole sources of radiation hardening in commercial RPV steels; the matrix defect clusters should not be downplayed. As a whole, these electron irradiation experiments on Fe–Cu alloys support the view that formation of copper-rich phases in steels during reactor irradiation is controlled by diffusion processes, not by in-cascade events.

The present and earlier [10] observations that the hardening efficiency of electron displacements in ferritic steels is equal to the hardening efficiency of fast neutrons per dpa is interesting on four counts. First, Ref. [10] strongly supports the conclusion that the embrittlement of the HFIR RPV is dominated by gamma (electron) displacements. Second, both observations demonstrate that clusters generated directly in large displacement cascades are not necessarily the primary source of radiation hardening and associated embrittlement in ferritic alloys. This conclusion is particularly important for the HFIR RPV investigations because there it was not clear whether the gamma-induced atomic displacements were generating their own hardening clusters or whether they were simply stabilizing or enhancing clusters of cascade debris created by the neutrons. Third, it challenges the postulation that soft irradiations might have greater hardening efficiencies than irradiations with fast neutrons. Fourth, it conflicts with the benchmark finding by Makin and Blewitt [2] of relatively weak electron irradiation hardening efficiency in copper which had set a tone of expectation of low electron irradiation hardening efficiencies for other metals.

5. Conclusions

The hardening efficiencies of electron irradiations per dpa in ferritic materials irradiated at 288°C are the same

as those of neutron irradiations, at least at low doses. This agrees with earlier conclusions for irradiations made at 60°C. These observations indicate that point defect clusters created in displacement cascades are not the only source of radiation hardening centers in RPVs. The insensitivity of radiation hardening to copper impurity content indicates the importance of dissolved copper and suggests that matrix defect clusters might play a larger role than hitherto acknowledged. The results do not support the contention that point defects produced in soft displacements might be more efficient than those generated in large displacement cascades.

Acknowledgements

Research sponsored in part by the Division of Materials Sciences, US Department of Energy under contract DE-ACO5-96OR22464 with Lockheed Martin Energy Research Corporation and in part by Association EURATOM-KFA.

References

[1] C.E. Dixon, in: Progress in Nuclear Energy, Series V, vol. 2 Pergamon, New York, 1959, p. 475.
[2] M.J. Makin, T.H. Blewitt, Acta Metall. 10 (1962) 241.
[3] L.K. Mansur, K. Farrell, J. Nucl. Mater. 244 (1997) 212.
[4] R. Gold, J.H. Roberts, D.G. Doran, in: Reactor Dosimetry: Methods, Applications and Standardization, ASTM STP 1001 (1989) 603.
[5] K. Farrell, S.T. Mahmood, R.E. Stoller, L.K. Mansur, J. Nucl. Mater. 210 (1994) 268.
[6] I. Remec, J.A. Wang, F.B.K. Kam, K. Farrell, J. Nucl. Mater. 217 (1994) 258.
[7] D.E. Alexander, L.E. Rehn, J. Nucl. Mater. 217 (1994) 213.
[8] L.K. Mansur, K. Farrell, J. Nucl. Mater. 170 (1990) 236.
[9] O.S. Oen, Cross Sections for Atomic Displacements in Solids by Fast Electrons, ORNL-4897, Oak Ridge National Laboratory, Oak Ridge, TN, 1973.
[10] D.E. Alexander, L.E. Rehn, K. Farrell, R.E. Stoller, J. Nucl. Mater. 228 (1996) 68.
[11] R.D. Evans, The Atomic Nucleus, McGraw-Hill, New York, 1972.
[12] ASTM E521, Standard Practice for Neutron Radiation Damage Simulation by Charged Particle Irradiation, Annual Book of ASTM Standards, vol. 12.02.
[13] R.J. DiMelfi, D.E. Alexander, L.E. Rehn, J. Nucl. Mater. 252 (1998) 171.
[14] F.W. Stallman, J.-A. Wang, F.B.K. Kam, B.J. Taylor, PR-EDB: Power Reactor Embrittlement Data Base, NUREG/CR-4816 Revision 2, ORNL/TM-10328/R2, Oak Ridge National Laboratory, Oak Ridge, TN, 1994.
[15] F.W. Stallman, J.-A. Wang, F.B.K. Kam, TR-EDB: Test Reactor Embrittlement Data Base, NUREG/CR-6076, ORNL/TM-12415, Oak Ridge National Laboratory, Oak Ridge, TN, 1994.
[16] E.D. Eason, J.E. Wright, G.R. Odette, Improved Embrittlement Correlations for Reactor Pressure Vessel Steels, NUREG/CR-6551, MCS 970501, US Nuclear Regulatory Commission, November 1998.
[17] R.E. Stoller, in: Effects of Radiation on Materials, ASTM STP 1270, American Society for Testing and Materials, Philadelphia, 1996, p. 25.
[18] G.R. Odette, E.V. Mader, G.E. Lucas, W.J. Phythian, C.A. English, in: Effects of Radiation on Materials, ASTM STP 1270, American Society for Testing and Materials, Philadelphia, 1996, p. 373.
[19] T.J. Williams, P.R. Burch, C.A. English, P.H.N. de la Cour Ray, in: Environmental Degradation of Materials in Nuclear Power Systems: Water Reactors, Third International Symposium, The Metallurgical Society, Warrandale, PA, 1998, p. 121.
[20] P. Auger, P. Pareige, M. Akamatsu, J.-C. Van Duysen, J. Nucl. Mater. 211 (1994) 194.
[21] T. Tobita, M. Suzuki, Y. Idei, A. Iwase, K. Aizawa, Study on γ-ray Induced Embrittlement of Light Water Reactor Pressure Vessel Steel, in: Proceedings of the 15th International Conference on Structural Mechanics in Reactor Technology, 15–20 August, 1999, Seoul, Korea (in press).

ELSEVIER

Journal of Nuclear Materials 280 (2000) 264–274

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Defect and void evolution in oxide dispersion strengthened ferritic steels under 3.2 MeV Fe^+ ion irradiation with simultaneous helium injection

I.-S. Kim [a], J.D. Hunn [b], N. Hashimoto [b], D.L. Larson [b,1], P.J. Maziasz [b], K. Miyahara [a], E.H. Lee [b,*]

[a] *Department of Molecular Design and Engineering, Nagoya University, Nagoya, Japan*

[b] *Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, P.O. Box 2008, TN 37831-6376, USA*

Received 8 December 1999; accepted 27 April 2000

Abstract

In an attempt to explore the potential of oxide dispersion strengthened (ODS) ferritic steels for fission and fusion structural materials applications, a set of ODS steels with varying oxide particle dispersion were irradiated at 650°C, using 3.2 MeV Fe^+ and 330 keV He^+ ions simultaneously. The void formation mechanisms in these ODS steels were studied by juxtaposing the response of a 9Cr–2WVTa ferritic/martensitic steel and solution annealed AISI 316LN austenitic stainless steel under the same irradiation conditions. The results showed that void formation was suppressed progressively by introducing and retaining a higher dislocation density and finer precipitate particles. Theoretical analyses suggest that the delayed onset of void formation in ODS steels stems from the enhanced point defect recombination in the high density dislocation microstructure, lower dislocation bias due to oxide particle pinning, and a very fine dispersion of helium bubbles caused by trapping helium atoms at the particle–matrix interfaces. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 81.20.Ev; 81.30.Mh; 61.80.Jh; 61.70.Yq

1. Introduction

In recent years, oxide dispersion strengthened (ODS) ferritic steels have been considered for structural materials for advanced fast breeder reactors and future fusion first wall applications [1–3]. These efforts are motivated mainly by the fact that ODS steels have superior creep resistance which permits operation at higher temperatures [1–4], and they have outstanding stability under irradiation [5], and meet low activation criteria for fusion power system [3]. In a 4 MeV Ni ion irradiated 12Cr–8Mo–0.1Y_2O_3 ODS ferritic steel, no voids were found after irradiation up to 300 dpa with 4 MeV Ni^{3+} ions at 525°C, and it was alluded that the void suppression might be prolonged indefinitely in such an ODS steel without considering the effect of helium generation under neutron irradiation [5]. On the other hand, electron irradiation of a 13Cr–0.23Y_2O_3 steel at 400°C produced a high number density of dislocation loops and faceted voids at oxide particle–matrix interfaces, but void formation was suppressed with increasing irradiation temperature, and ceased at 550°C [4].

In general, ferritic/martensitic steels have shown excellent swelling resistance under both neutron [6–10] and ion irradiation [5,11,12]. It is, however, well known that voids do form in ferritic steels under gas generating environments such as in neutron irradiation [6–10] or ion irradiation with helium injection [11,12]. Although

* Corresponding author. Tel.: +1-865 574 5058; fax: +1-865 574 0641.

E-mail address: leeeh@ornl.gov (E.H. Lee).

[1] Present address: Seagate Recording Heads, Mailstop NRWTRI, 7801 Computer Avenue South, Minneapolis, MN 55435-5489, USA.

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(00)00066-0

bias-driven void growth [13] is generally delayed in ferritic steel, evidences suggest that the void growth rate is not much slower than that in austenitic stainless steels once a bias-driven void growth regime is established [14,15].

It is the aim of this paper to elucidate the nature of swelling resistance of ferritic steels and the potential benefit for the inclusion of Y_2O_3 particles in ODS steel. For this study, a set of ODS steels were irradiated together with a tempered martensitic steel, 9Cr–2WVTa, and a solution annealed AISI 316LN austenitic stainless steel under the same irradiation condition by employing 3.2 MeV Fe^+ and 330 keV He^+ ion beams simultaneously. After irradiation, the damage and cavity microstructures were studied with transmission electron microscopy (TEM). The results are explained first by establishing a strong correlation between microstructure and void formation behavior and then analyzing the information under a powerful theoretical framework based on a rate theory model [13].

2. Experimental

The three ODS steels (17Y3, 12Y1, 12YWT) investigated in this work were provided by Kobe Specialty Tube Company in Japan, and the 9Cr–2WVTa ferritic/martensitic steel (Heat #3791) and AISI 316LN austenitic stainless steel (Heat #18474) were procured by Oak Ridge National Laboratory. The alloy compositions are tabulated in Table 1.

The ODS steels were prepared by a mechanical alloying (MA) process, which involved a preparation of pre-alloyed powder of ≈ 70 μm diameter particle size via an Ar gas atomization process, mixing the metal powder with ≈ 20 nm Y_2O_3 powder in Ar gas atmosphere, milling with high energy attrition type balls for 48 h, degassing for 2 h at 400°C in $< 2 \times 10^{-2}$ Pa vacuum, hot isostatic press (HIP) canning, hot extrusion to a bar at 1150°C, hot rolling into plate at 1150°C, thickness reduction with cold rolling for 17Y3 and 12Y1 and warm rolling at 600°C for 12YWT, final annealing at 1050°C for 1 h, and air cooling. The detailed alloy preparation procedures can be found in [16]. The 9Cr–2WVTa alloy was prepared by austenitizing at 1050°C for 30 min under flowing He, fast inert gas cooling, tempering at 750°C for 1 h, and fast inert gas cooling. The AISI 316LN steel was solution annealed at 1050°C for 30 min in vacuum.

TEM disks were prepared from each steel by mechanically polishing followed by electrochemical polishing prior to irradiation. Irradiation was carried out with 3.2 MeV Fe^+ and 330 keV He^+ ion beams simultaneously using the Triple Ion Facility (TIF) at ORNL. Details of the triple ion beam facility can be found in [17]. The ion energies were chosen to render the maximum damage and gas atom deposition near a depth of 750–800 nm determined by the computer code, stopping and range of ions in matter (SRIM, 1998 version) [18]. An Fe^+ ion beam was applied to give a peak damage rate of $\approx 10^{-3}$ dpa/s (displacements per atom/s) together with ≈ 5 appm He/dpa injection rate at 650°C. The final dose attained was 50 dpa with 260 appm He at the peak damage region. The procedure for the dpa calculation is described in [19].

TEM specimens were prepared by electrochemically removing ≈ 700 nm from the ion bombarded side and then thinning from the unirradiated side until perforation occurred. This procedure produced TEM foils with thicknesses of ≈ 100 nm on average and allowed examination of the microstructure between 700 and 800 nm original depth. Damage microstructure was examined with JEOL FX2000 and Philips CM30 electron microscopes operated at 200 keV.

The nature of the oxide particles after MA were characterized by an energy-compensated 3-dimensional atom probe field ion microscopy (APFIM). The detailed APFIM study results are published in a separate paper [20]. Here, only brief experimental findings are included in Section 4 to aid the understanding of the state of oxide particle dispersion in the mechanically alloyed ODS steels and its impact on void formation during irradiation.

Table 1
Alloy compositions in wt%[a]

Alloy	Fe	Cr	Ni	Mn	Mo	W	V	Ta	Ti	Si	C	P	Y_2O_3
316 LN	Bal	16.31	10.20	1.75	2.07	–[b]	–	–	–	0.39	0.009	0.029	–
9Cr–2WVTa	Bal	8.72	–	0.43	–	2.09	0.23	0.075	–	0.23	0.10	0.015	–
17Y3	Bal	17.07	–	< 0.1	–	–	–	–	–	< 0.1	< 0.05	< 0.01	0.239
12Y1	Bal	12.35	–	< 0.1	–	–	–	–	–	< 0.1	< 0.05	< 0.01	0.248
12YWT	Bal	12.29	–	< 0.1	–	–	–	–	0.39	< 0.1	< 0.05	< 0.01	0.249

[a] Trace inclusions such as B, Co, Ni, Cu, N, S (< 0.01) etc., exist also.
[b] Not added.

3. Results

The initial microstructure of solution annealed AISI 316LN austenitic stainless steel was characterized by well defined, clean grains with a very low density of dislocations ($\ll 10^{13}$ m/m^3). On the other hand, the 9Cr–2WVTa ferritic/martensitic steel was characterized by complex lath subgrains with a higher dislocation density ($\lesssim 10^{15}$ m/m^3). The dislocation density was higher ($\approx 10^{15}$ m/m^3) still in the 17Y3 and 12Y1 ODS steels. Finally, among the three ODS steels, the 12YWT had the highest as-processed dislocation density, which varied slightly from grain to grain ($10^{15} - 10^{16}$ m/m^3). These differences in the initial microstructures have a profound impact on the evolution of defects and cavities during irradiation.

Figs. 1 and 2 compare the oxide particle dispersions in 12Y1 and 12YWT, respectively. Electron diffraction indicated that most oxide particles in 12Y1 were in crystalline form, the particle sizes ranging from 10 to 40 nm diameter. Particle distribution was non-uniform, with the number density in the range of 10^{20}–10^{21} m^{-3}. Besides the large particles, numerous tiny particles a few nm in diameter were also present. Although not shown here, the particles in 17Y3 steel were coarser and fewer in number compared with those in 12Y1. By comparison, much finer (median 3–5 nm diameter) and finely dispersed particle-like objects were found in the 12YWT steel. Surprisingly, however, except for diffraction spots from occasional large particles, neither diffraction spots nor amorphous rings were derived from these small objects. A dark-field image from a matrix diffraction spot revealed profuse white specks, presumably arising from the strain contrast around these objects (Fig. 2), suggesting that these objects may be clusters of atoms which reside as solid solution with lattice mismatch around them. The cluster number density in 12YWT steel was estimated to be in the range 1–2×10^{23} m^{-3} by counting the white strain-contrast specks in the weak-beam dark-field image.

During irradiation, dislocation microstructures were modified and cavities formed. Figs. 3 and 4 are low magnification TEM micrographs for the five steels irradiated to 50 dpa and 260 appm He at 650°C. The microstructure of the irradiated solution annealed AISI 316LN showed bubbles ($r < 0.5$ nm), voids (0.5 nm $\lesssim r < 100$ nm), faulted interstitial loops, and segments of line dislocations. Here, the term 'bubble' is used to describe cavities smaller than a certain critical size where a bimodal cavity size distribution is observed, and 'void' describes cavities larger than the critical size, which grow via a bias-driven mechanism (by absorption of vacancies without the aid of gas atoms). The definition of the critical cavity size and a theoretical basis for the

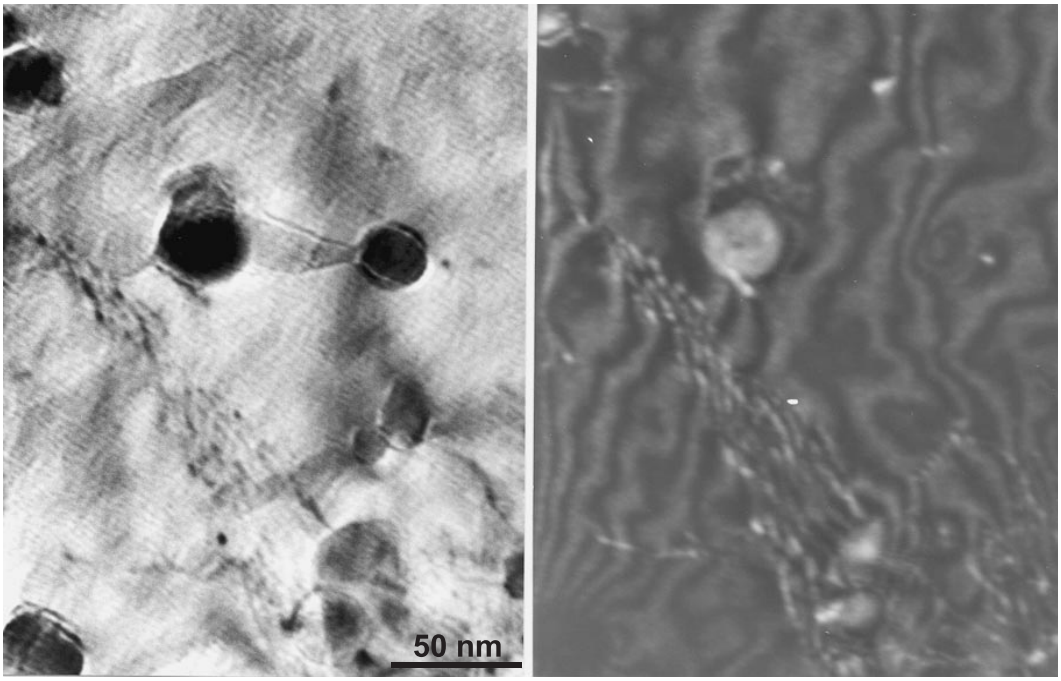


Fig. 1. Microstructure of unirradiated 12Y1 ODS steel: bright- and dark-field TEM micrographs taken near beam direction $B \approx \langle 1\,2\,2 \rangle$. The weak-beam dark-field image is taken at $g = 110$, $(g, 4g)$ diffracting condition. Y_2O_3 precipitate particle sizes are in the range of a few tens of nanometers in diameter.

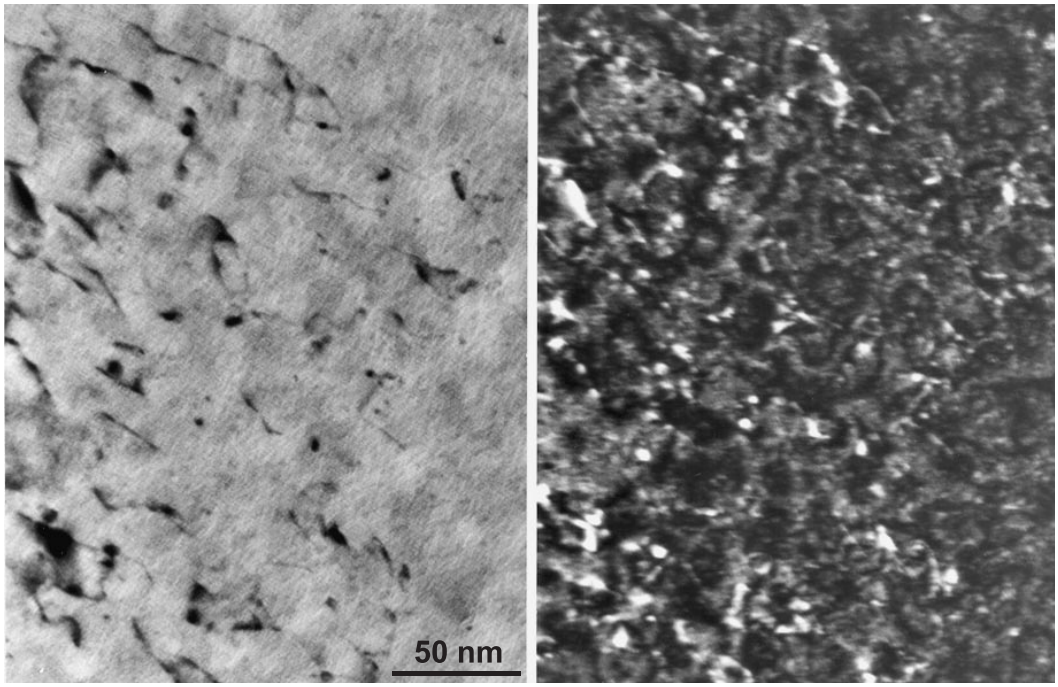


Fig. 2. Microstructure of unirradiated 12YWT ODS steel: bright- and dark-field micrographs taken near beam direction $B \approx \langle 1\,0\,0 \rangle$. The weak-beam dark-field image is taken at $g = 110$, $(g, 4g)$ diffracting condition. The average cluster size is in the range of 3–5 nm. No diffraction spots were seen from these particles, but the presence of high number density clusters is apparent from the strain contrast of clusters in dark-field image.

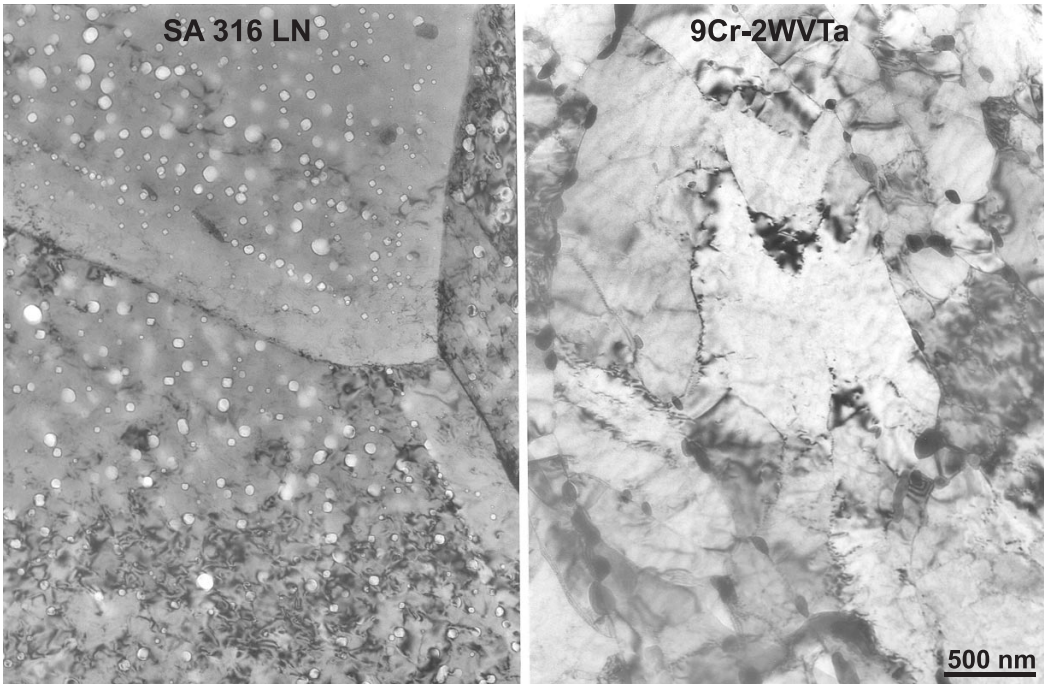


Fig. 3. Low magnification microstructure of solution annealed AISI 316LN austenitic stainless and 9Cr–2WVTa ferritic/martensitic steels after irradiation to 50 dpa (260 appm He) at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions.

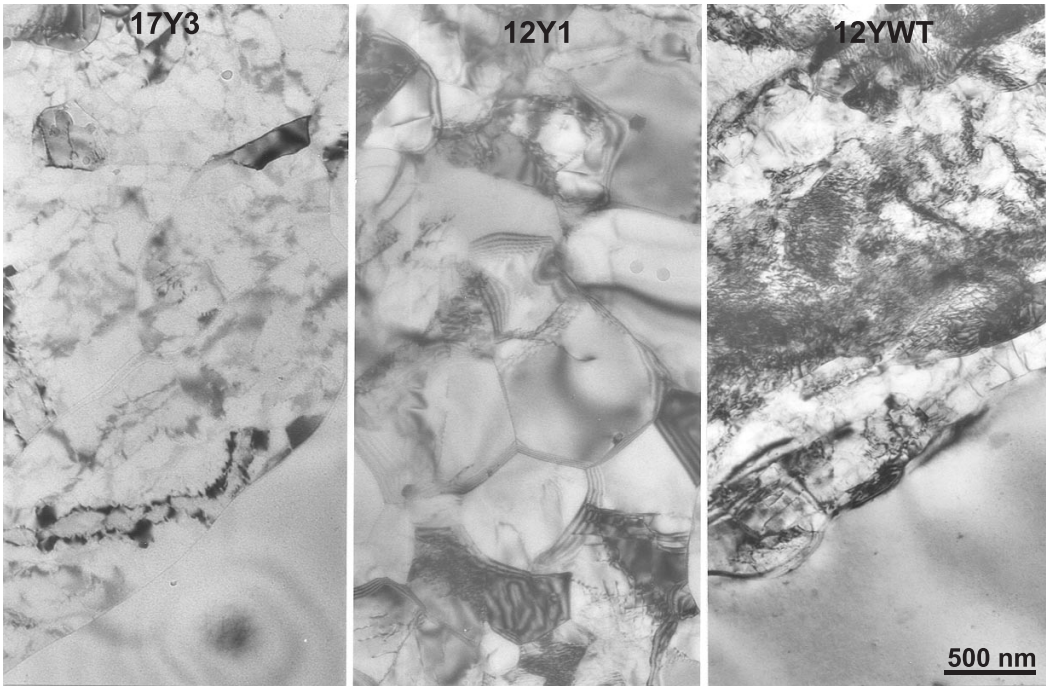


Fig. 4. Low magnification microstructure of 17Y3, 12Y1, and 12YWT ODS ferritic steels after irradiation to 50 dpa (260 appm He) at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions.

appearance of a bimodal cavity size distribution will be further clarified in Section 4. In the ferritic steel subgrains, varying degrees of dislocation recovery were observed after irradiation, but only a few subgrains showed complete recovery. The least amount of radiation-induced recovery occurred in the 12YWT ODS steel. Although bubbles and voids were not clearly visible at low magnification, high magnification micrographs showed that bubbles and voids were also present in the 9Cr–2WVTa and ODS steels, mainly in the recovered areas.

Figs. 5 and 6 show high magnification microstructures of the irradiated steels corresponding to Figs. 3 and 4. Bimodal cavity size distributions can be seen clearly in the solution annealed AISI 316LN and 9Cr–2WVTa steels in the magnified insets in Fig. 5. In the 9Cr–2WVTa steel, dislocation loops were nucleated in the recovered area, in which bubbles ($r < 0.5$ nm) and voids (0.5 nm $\lesssim r < 10$ nm) were also found, mainly associated with the irradiation induced dislocation loops and remnant dislocations. In both the 17Y3 and 12Y1 steels, voids were very small ($r < 5$ nm) but a careful examination clearly indicated bimodal cavity size distributions. The 12YWT steel contained the highest concentration and the finest bubbles ($r < 1$ nm), but there was no indication of bimodal cavity development. Overall, the microstructural evidence indicated that bimodal cavity evolution occurred earliest in the solution annealed AISI 316LN austenitic stainless steel, next in the 9Cr–2WVTa ferritic/martensitic steel, and last in the 17Y3 and 12Y1 ODS steels.

However, a bimodal cavity distribution did not develop in the 12YWT ODS steel. Although bubbles in the 12YWT were too small to make a precise counting, very high number density of bubbles seems to present in association with dislocations and clusters. It is well known that dislocations and misfit particle interfaces are good gas atom trap sites. Therefore, in consideration of the initial high density of dislocations and the high number of misfit clusters confirmed by APFIM and TEM, the bubble number density in 12YWT is believed to be in the range of 10^{23}–10^{24} m^{-3}.

4. Discussion

Before discussing the radiation-induced microstructures, the findings on the nature of oxide particles by APFIM and X-ray are reviewed here first to better understand the nature of mechanically alloyed ODS steels.

In an independent APFIM study [20] of the 12YWT steel, (Ti, O, Y, C, B, Cr)-rich nebulous clusters were found to be present, embedded in the matrix of the host Fe, with atomic Y and O scattered around them, rather than as well-defined isolated particles, Fig. 7, confirming the TEM diffraction results. The cluster number density

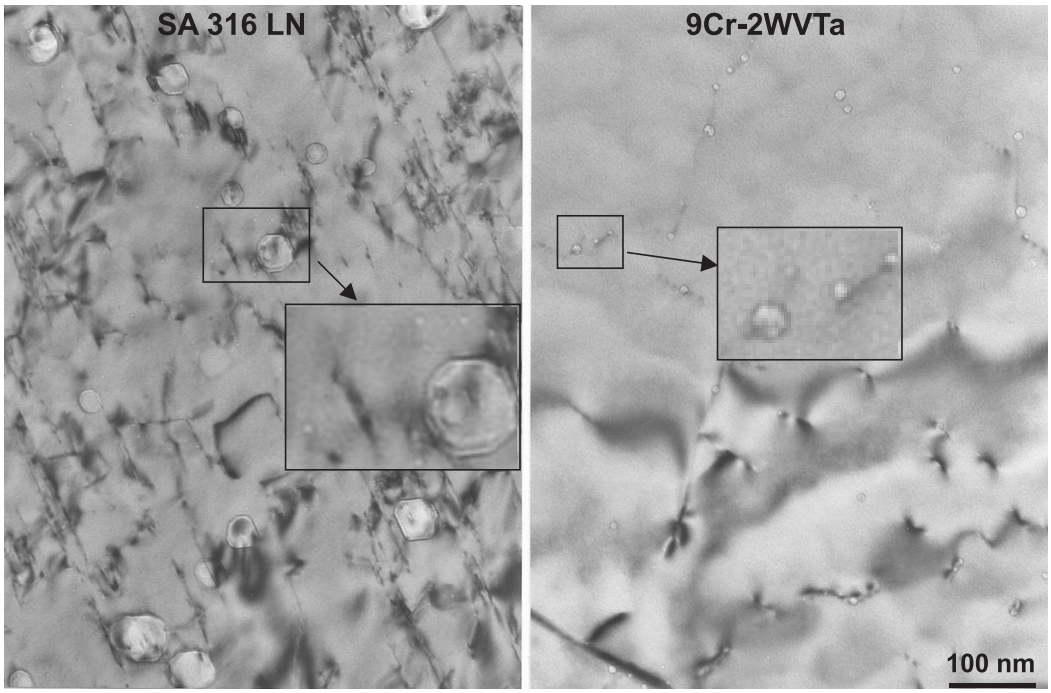


Fig. 5. High magnification microstructure of solution annealed AISI 316LN austenitic stainless and 9Cr–2WVTa ferritic/martensitic steels after irradiation to 50 dpa (260 appm He) at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions.

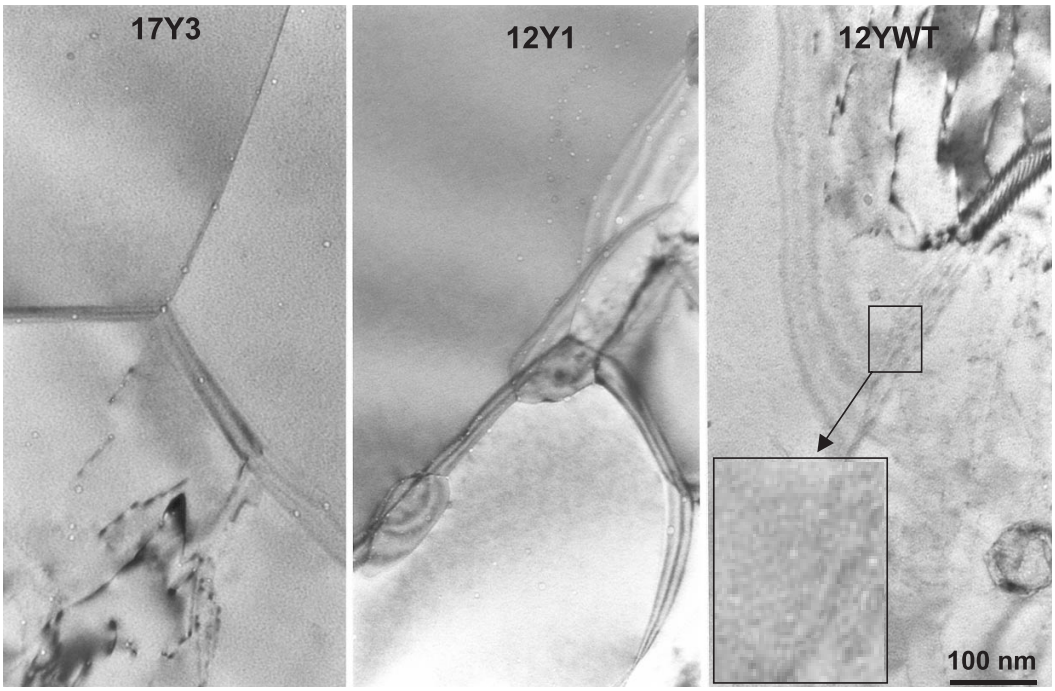


Fig. 6. High magnification microstructure of 17Y3, 12Y1, and 12YWT ODS ferritic steels after irradiation to 50 dpa (260 appm He) at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions.

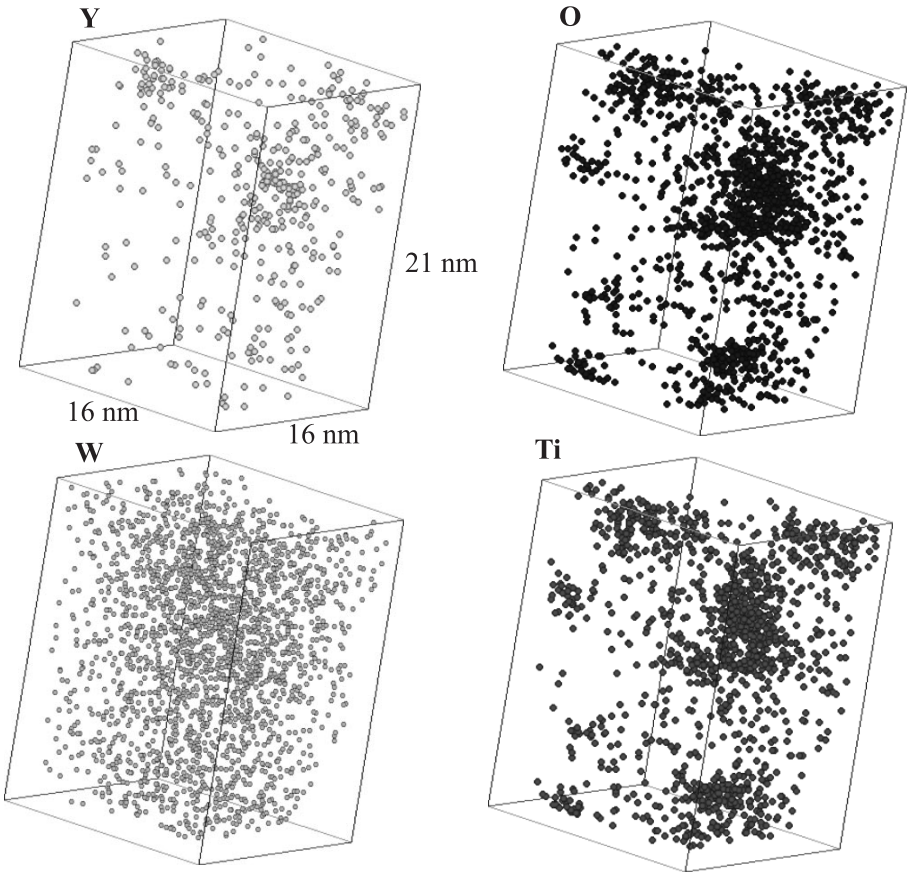


Fig. 7. APFIM image of the oxide clusters in 12YWT ODS steel. The fine (1–7 nm) oxide cluster regions are depleted in Fe and enriched in Y, Ti, O, Cr, C, and B. The composition of the arrow-marked cluster is in at.% $Fe_{bal}Cr_{15.3}W_{0.70}Y_{1.21}Ti_{4.32}O_{3.78}C_{0.57}Si_{0.19}B_{0.13}$. The scatter of Y and O atoms around the clusters and the enrichment of Ti, C and B in the clusters suggest that the original Y_2O_3 particles were smashed into atomic scale during mechanical alloying and reconstructed as new clusters during the final annealing process at 1050°C. The refinement of particle size occurred only in the presence of Ti.

was estimated to be about 1×10^{24} m^{-3}, an order of magnitude higher than those observed by TEM. In particular, O was found to be in strong association with Ti, seemingly as a single TiO_x complex or a cluster of such. Interestingly, tungsten was distributed uniformly in the matrix showing no particular association with any specific element, but refinement of cluster sizes were observed in ODS steel with W but no Ti added [20]. The sizes of the enriched core were in the range of a few to several nm in diameter, a similar size scale found in TEM.

In 12Y1 steel, no fine clusters were found, but rather large ($\approx$ 30 nm) and well-defined (Y, O, C)-rich clusters were present in association with scattered Y, O, and C atoms around them, Fig. 8. These results suggest that a substantial fraction of the crushed and smeared particles during MA recrystallized into a larger cluster by accommodating other alloying elements in the matrix during the final annealing at 1050°C. This is significantly different from the clustering behavior observed in Ti-added 12YWT steel, indicating that Ti might have a significant role in the refinement of oxide particles during annealing. Finally, it should be noted that thermally produced precipitates such as $M_{23}C_6$ and MC appear as well-defined isolated particles in APFIM.

In X-ray diffraction analyses, it was found that Y_2O_3 particles broke down into smaller sizes with increasing MA time, eventually below the detection limit after 48 h of milling in various ODS steels [21]. X-ray diffraction analyses of the mechanically alloyed Ti-containing ODS steel showed the characteristic b.c.c. iron matrix peaks but no Y_2O_3 peaks, suggesting that Y_2O_3 dissociated into atomic yttrium and oxygen during the MA process, eventually producing a single phase solid solution. Very weak new peaks corresponding to Y_2O_3, Y_2TiO_5 ($Y_2O_3 + TiO_2$), and $Y_2Ti_2O_7$ ($Y_2O_3 + 2TiO_2$) began to

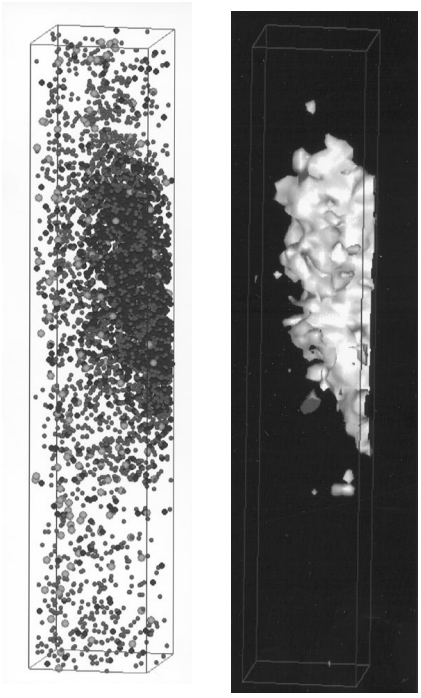


Fig. 8. APFIM image of the oxide particles in 12Y1 ODS steel. Shown here is an oxide particle in a ($10 \times 10 \times 50$ nm^3) box: light dots – Yttrium ($\sim$ 1%), medium dots – Carbon ($\sim$ 5%) and dark dots – Oxygen ($\sim$ 10%). The white particle on the right is the contour of 5% Y + O + C surface. Dispersion of constituent elements of Y_2O_3 is evident around the particle suggesting that atomization of Y_2O_3 particle also occurred during mechanical alloying. The particles are enriched in Y, O and C, much different from that of original Y_2O_3, again suggesting that the clustering (recrystallization in this case) occurred during the secondary annealing process. These particles are found to be in crystalline form under TEM. Particle sizes in ODS steels without Ti were larger than those in ODS steel with added Ti.

emerge after annealing at 1000°C for 1 h, but no new peaks were observed for 900°C annealing. The Y_2O_3 peak after 1000°C annealing was very small and broad, indicating either a smaller fraction or a finer dispersion of the dissociated Y_2O_3 recrystallized at this temperature. The intensity of the new peaks increased with increasing annealing temperature, becoming prominent above 1200°C. Most particle sizes were initially smaller than 1 nm, smaller even than the unit cell dimension of cubic Y_2O_3 (1.06 nm), indicating that Y_2O_3 particles were almost atomized during MA. Electron diffraction and X-ray analyses of the extracted particles from the1100°C annealed ODS steels also indicated that Y_2O_3 particles were in the form of fine ($\approx$ 3 nm) Y_2TiO_5 and $Y_2Ti_2O_7$ complex clusters in the presence of Ti, whereas crystalline coarse ($\approx$ 13nm) Y_2O_3 and Y_2SiO_5 particles were present in 12Cr–Y_2O_3 steel with no added Ti [22], consistent with the new APFIM findings.

The microstructures of irradiated austenitic stainless, ferritic/martensitic, and ODS ferritic steels were examined after a displacement damage of 50 dpa and 260 appm helium implantation at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions. The results showed that void formation in the irradiated steels diminished with increasing complexity of the initial microstructure as summarized in Table 2.

The most conspicuous void formation occurred in solution annealed AISI 316LN austenitic stainless steel with the lowest dislocation density, then 9Cr–2WVTa ferritic/martensitic steel with more dislocations, then the 17Y3 ODS steel with the coarser oxide dispersions and higher dislocation density, then 12Y1 ODS steel, and finally only bubbles were present in 12YWT ODS steel with the finest oxide cluster dispersions and the highest dislocation density. The development of a bimodal cavity size distribution occurred in all but the 12YWT. Void formation in the AISI 316LN was accompanied by the formation of interstitial faulted loops followed by the development of prismatic loops and line dislocations, and void formation in ferritic steels occurred mainly in the recovered areas in the subgrains. The most refined bubble distribution and the least dislocation recovery were found in the 12YWT steel, which contained the finest oxide dispersion.

Table 2
Summary of microstructure before and after irradiation

Steels	Before irradiation	After irradiation
SA AISI 316LN	Well-defined clean grains low dislocation density, $\rho_l \ll 10^{13}$ m/m^3	Evolution of black dots, faulted loops, prismatic loops, line dislocations, and bimodal cavities (< 100 nm radius)
9Cr–2WVTa	Complex lath subgrains, $\rho_l \lesssim 10^{15}$ m/m^3	Partial recovery of dislocations bimodal cavities (< 10 nm radius)
17Y3	Finer complex subgrains with coarse and fine particles, $\rho_l \approx 10^{15}$ m/m^3	Partial recovery of dislocations bimodal cavities (<5 nm radius)
12Y1	Finer complex subgrains with coarse and fine particles, $\rho_l \approx 10^{15}$ m/m^3	Partial recovery of dislocations bimodal cavities (< 5 nm radius)
12YWT	Finest complex subgrains with finest particles or clusters, $\rho_l = 10^{15} - 10^{16}$ m/m^3	Partial recovery of dislocations unimodal bubbles (< 1 nm radius)

Generally, void formation in ferritic/martensitic steel is delayed compared to that in solution annealed 300 series austenitic stainless steel under the same irradiation condition. However, this difference in void formation between the two different types of steels is not considered to stem from fundamental differences in material parameters (i.e., crystal structure, point defect diffusivity, migration energy, formation energy, etc.). Rather, the apparent void formation resistance of the ferritic/martensitic steels originates from the difference in microstructure, namely high density dislocations in the subgrains in the ferritic/martensitic structure. It has been well documented that void formation was also suppressed in austenitic stainless steel when a high density of dislocations were introduced by cold working, and/or the microstructure contained a high number density of precipitates and/or He-bubbles [14,23].

During irradiation, bubbles nucleate and grow by forming vacancy–gas-atom complexes as they accumulate gas atoms. When bubbles exceed a certain critical size or accumulate a certain number of gas atoms, they can grow inexorably by absorbing vacancies without the need for gas atoms. This process is often referred to as a 'bias-driven mechanism.' When bubbles are smaller than a certain size, thermal emission prevails over the absorption of vacancies. This demarcates bubble sizes below and above the critical size, which causes development of a bimodal cavity size distribution. Numerous examples of bimodal cavity distributions are documented in [24]. Here, cavities below the critical size are simply called 'bubbles,' and the larger ones 'voids.' A more rigorous definition of the critical cavity and the related theoretical concepts can be found in [25–27].

The critical cavity size is not a fixed number, but varies dynamically with material parameters, irradiation conditions, and microstructure. Specifically, vacancy supersaturation, and correspondingly void formation, are affected by material parameters, temperature, defect and gas generation rates, and sink number density (dislocations, precipitates, cavities). In the presence of a high number of sinks, most of the interstitials and vacancies produced by radiation recombine at the sinks and fewer vacancies are available for cavity growth. This causes the critical cavity size to increase, and void formation becomes more difficult. In the absence of sinks, most point defects recombine in the matrix and the same effect on void formation results. Dislocation bias (preferential absorption for interstitials over vacancies or vice versa) is not a fixed value either, and it also varies with the state of dislocation evolution. Tangled or pinned (by particles or bubbles) dislocations are difficult to climb up or down, and this makes the bias in a high dislocation density microstructure initially very low, particularly for dislocations immobilized by pinning. Pinned dislocations not only have a low bias, but also have difficulty recovering, as observed in the 12YWT ODS steel. Precipitate–matrix interfaces are particularly good trap sites for gas atoms, and thus very fine dense precipitates lead to a microstructure with a high concentration of fine bubbles. In such a system, the available helium atoms are diluted over many bubbles as the bubbles themselves also act as recombination sites. In effect, such feedback makes the critical cavity size large, which requires more helium for conversion from bubbles to voids.

Critical quantities associated with cavity formation are calculated by employing the rate theory formulation described in [14]. The results of the calculations are shown in Fig. 9. The calculations require detailed knowledge of material parameters, irradiation variables, and microstructural information. The material parameters used in the calculation are summarized in the figure caption. Although precise material parameters are still unknown, the material parameters used here have been justified previously. Moreover, the outcomes of the calculated values are determined most sensitively by the dominant sink strength (dislocation density and cavity number density), and the use of a wide range of material parameter values does not greatly alter the calculations. In the irradiated ODS steels, the cavity sizes were too small to derive microstructural statistics from the TEM micrographs. So, the microstructural data used for calculations are approximate and the results are qualitative. However, they do clearly show that there exists a fundamental physical basis to explain the development of a bimodal cavity distribution, its delay, or perhaps near indefinite delay in systems with very high number density of fine precipitates, such as the oxide clusters found in the 12YWT steel.

In Fig. 9, the variations of critical cavity radius are plotted as a function of dislocation density for two different cavity number densities and dislocation biases with the corresponding helium concentration levels on the right. The helium concentration that is required for all cavities to reach a critical size ($N_g^* \times N_c$, number of helium atoms per critical cavity multiplied by the cavity number density) is expressed in appm.

The implications of the plots are the following: (1) in a solution annealed or low dislocation system, the critical cavity radius (r_c^*) is initially large because most point defects recombine in the matrix. r_c^* decreases as dislocation microstructure evolves (i.e., black dots faulted interstitial loops → unfaulted prismatic loops → line dislocations); (2) in a cold worked (CW) system, r_c^* is initially large, this time because most point defects recombine at dislocations and r_c^* decreases as dislocations recover by either climbing up or down by absorbing vacancies or interstitials during irradiation; (3) when the cavity number density is high, cavities themselves become the point defect recombination sites and r_c^* becomes large, particularly in the low dislocation density regime; (4) when dislocation bias decreases (i.e., by

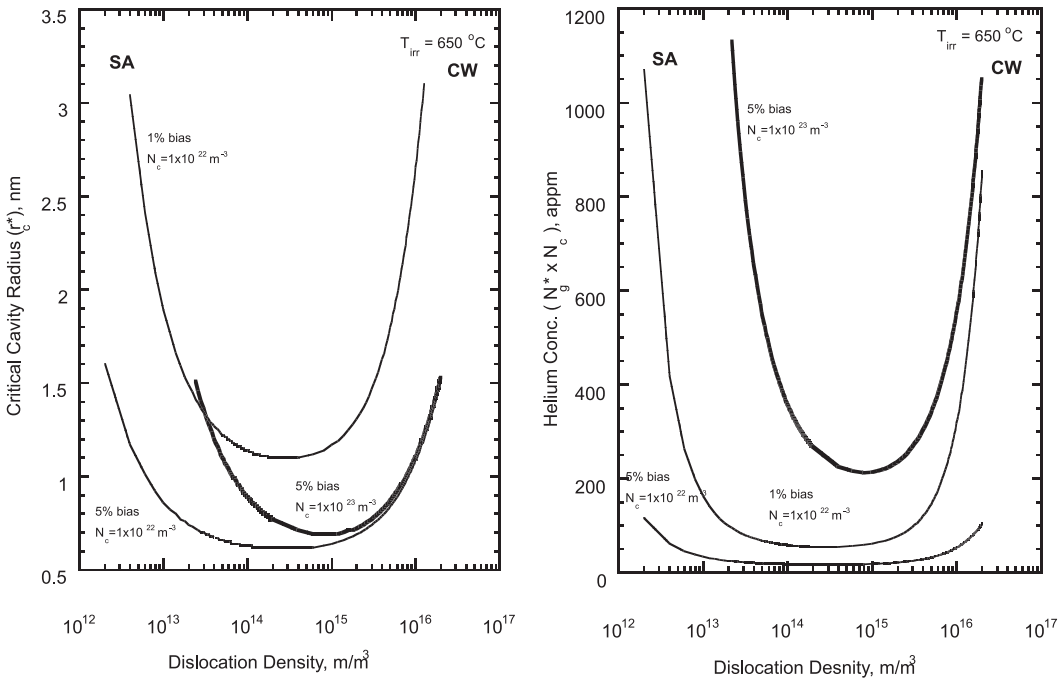


Fig. 9. Variation of critical cavity radius as a function of dislocation density and the corresponding helium concentration ($N_g^* \times N_c$, number of helium atoms per critical cavity multiplied by the cavity number density). The plots are made for two different cavity number densities and dislocation biases. The parameters used in this computation are displacement damage rate $G = 1 \times 10^{-3}$ dpa/s, vacancy diffusivity pre-exponential $D_v^0 = 1 \times 10^{-6}$ m^2/s , interstitial diffusivity pre-exponential $D_i^0 = 1 \times 10^{-6}$ m^2/s, vacancy migration energy $E_v^m = 1.2$ eV, interstitial migration energy $E_i^m = 0.15$ eV, vacancy formation energy $E_v^f = 1.6$ eV, vacancy formation entropy $S_v^f = 1.29 \times 10^{-4}$ eV/K, and surface energy $\gamma = 1.5$ J/m^2. Dislocation bias changes with microstructural evolution and thus precise bias values are not known. For simulation, 1% and 5% bias values were used.

pinning dislocations), defect recombination is enhanced and r_c^* becomes large; (5) finally, when the cavity number density increases (i.e., by introducing a high number of gas atom trap sites such as fine particles in the ODS steel), r_c^* increases and the helium required for the onset of bias-driven cavity growth increases for the reasons mentioned already.

From the above considerations it can be said that: (1) solution annealed AISI 316LN austenitic stainless steel, defect evolution was rapid and bias-driven void growth started earlier; (2) the 9Cr–2WVTa ferritic/martinsitic steel, bias-driven void growth was delayed because of the initial high density of dislocations and slow recovery; (3) the 17Y3 and 12Y1 ODS steels, bias-driven void growth was further delayed because of slow recovery and a low dislocation bias due to oxide particle pinning; (4) in the 12YWT ODS steel, the latter effects were further amplified by the presence of a very high number density of fine oxide clusters. While microstructural evolution can be studied in more detail as a function of dose or the bias change in association with precipitate pinning, overall response of the materials to irradiation is generally in accord with theoretical expectations.

5. Conclusions

Void formation mechanisms in solution annealed AISI 316LN austenitic stainless steel, 9Cr–2WVTa ferritic/martensitic steel, and three (17Y3, 12Y1 12YWT) ODS ferritic steels were examined after irradiation to 50 dpa with 260 appm simultaneous helium injection at 650°C with 3.2 MeV Fe^+ and 330 keV He^+ ions. These results showed that void growth was delayed with increasing microstructural complexity, namely by increasing the dislocation density and immobilizing dislocations via nanoscale oxide particles and clusters. The most significant void suppression occurred in 12YWT ODS steel, which possessed the highest dislocation density and the finest and highest number of oxide clusters. Analyses based on a rate theory model suggested that the sink density in the system dictates point defect recombination and void formation, and that change in bias could also be a significant factor in controlling void formation and growth. The suppression of void formation in the ODS steels is attributable to the initial high density of dislocations, slow dislocation recovery by oxide particle pinning, low bias of immobilized or pinned dislocations, and to the nucleation of a

high number density of cavities by trapping helium atoms at particle–matrix interfaces. Overall the data suggest that, within a practical limit, void formation can be suppressed if the dislocation microstructure is stabilized and helium bubbles are dispersed on a fine scale via fine stable oxide particles or atom clusters. However, for steel with such a microstructure, radiation hardening-induced embrittlement could become a concern. Further research is clearly needed prior to the application of ODS steels for fission and fusion structural materials.

Acknowledgements

This research was sponsored in part by the Division of Materials Sciences, US Department of Energy under contract No. DE-AC05-00OR22725 and in part by the Office of Fusion Energy, US Department of Energy, under contract No. W-7405-eng-26 with UT-Battelle, LLC. One of the authors, Kim, thanks ORNL for the support during his graduate research. The authors thank Drs Kleuh and Rowcliffe for technical review of the manuscript.

References

[1] S. Ukai, T. Nishida, H. Okuda, M. Inoue, M. Fujiwara, T. Okuda, K. Asabe, in: Proceedings of the International Symposium on Materials Chemistry in Nuclear Environment, Materials Chemistry'96, Tsukuba, Ibaraki, Japan, 14–15 March 1996, p. 891.
[2] S. Ukai, T. Nishida, H. Okada, T. Okuda, M. Fujiwara, K. Asabe, J. Nucl. Sci. Technol. 34 (1997) 256.
[3] D.K. Mukhopadhayay, F.H. Froes, D.S. Gelles, J. Nucl. Mater. 258–263 (1998) 1209.
[4] T. Okuda, S. Nomura, M. Nakanishi, M. Fujiwara, H. Takahashi, in: Proceedings of the First Japan International SAMPE Symposium, 28 November–1 December 1989, p. 1616.
[5] J. Ohta, T. Ohmura, K. Kako, M. Tokiwai, T. Suzuki, J. Nucl. Mater. 225 (1995) 187.
[6] D.S. Gelles, in: F.A. Garner, N.H. Packan, A.S. Kumar (Eds.), Proceedings of the 13th International Symposium (Part I), ASTM STP 995, ASTM, Philadelphia, PA, 1987, p. 560.
[7] E.A. Little, D.A. Stow, Mater. Sci. (March) (1980) 89.
[8] J.M. Vitek, R.L. Klueh, J. Nucl. Mater. 122&123 (1984) 245.
[9] Ulrich Stamm, H. Schroeder, J. Nucl. Mater. 155–157 (1988) 1059.
[10] J.J. Kai, R.L. Klueh, J. Nucl. Mater. 230 (1996) 116.
[11] L.L. Horton, L.K. Mansur, in: F.A. Garner, J.S. Perrin (Eds.), Proceedings of the 12th International Symposium, ASTM STP 870, ASTM, Philadelphia, PA, 1985, p. 344.
[12] K. Farrell, E.H. Lee, in: F.A. Garner, J.S. Perrin (Eds.), Proceedings of the 12th International Symposium, ASTM STP 870, ASTM, Philadelphia, PA, 1985, p. 383.
[13] L.K. Mansur, J. Nucl. Mater. 206 (1993) 306.
[14] E.H. Lee, L.K. Mansur, Metall. Trans. A 21 (1990) 1021.
[15] L.K. Mansur, E.H. Lee, J. Nucl. Mater. 179–181 (1991) 105.
[16] T. Okuda, S. Nomura, S. Shikakura, K. Asabe, S. Tanoue, M. Fujiwara, in: A.H. Clauer, J.J. deBarbadillo (Eds.), Minerals, Metals and Materials Society, TMS, USA, 1990, p. 195.
[17] M.B. Lewis, W.R. Allen, R.A. Buhl, N.H. Packan, S.W. Cook, L.K. Mansur, Nucl. Instrum. and Meth. B 43 (1989) 243.
[18] J.F. Ziegler, J.P. Biersack, U. Littmark, The Stopping and Range of Ions in Solids, Pergamon, Oxford, 1985.
[19] E.H. Lee, Nucl. Instrum. and Meth. B 151 (1999) 29.
[20] D.J. Larson, P.J. Maziasz, I.-S. Kim, K. Miyahara, Three dimensional atom probe observation of nanoscale titanium–oxygen clustering in an oxide-dispersion-strengthened Fe–12Cr–3W–0.4Ti + Y_2O_3 ferritic alloy, Scripta Mater. (submitted).
[21] T. Okuda, M. Fujiwara, J. Mater. Sci. Lett. 14 (1995) 1600.
[22] S. Nomura, T. Okuda, S. Shikakura, M. Fujiwara, K. Asabe, in: A.H. Clauer, J.J. deBarbadillo (Eds.), Minerals, Metals and Materials Society, TMS, USA, 1990, p. 203.
[23] E.H. Lee, L.K. Mansur, Metall. Trans. A 23 (1992) 1977.
[24] L.K. Mansur, E.H. Lee, P.J. Maziasz, A.F. Rowcliffe, J. Nucl. Mater. 141–143 (1986) 633.
[25] L.K. Mansur, Nucl. Technol. 40 (1978) 5.
[26] L.K. Mansur, in: G.R. Freeman (Ed.), Kinetics of Nonhomogeneous Process, Wiley, 1989, p. 377 (Chapter 8).
[27] L.K. Mansur, W.A. Coghlan, J. Nucl. Mater. 119 (1983) 1.

ELSEVIER

Journal of Nuclear Materials 281 (2000) 65–70

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

A method to study deformation mechanisms for irradiated steels using a disk-bend test

E.H. Lee *, T.S. Byun, J.D. Hunn, N. Hashimoto, K. Farrell

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 16 March 2000; accepted 2 May 2000

Abstract

There is considerable interest in understanding how dislocations interact with defect clusters produced by radiation to evaluate and remedy the radiation induced degradation of mechanical properties. In this paper, it is demonstrated that deformation mechanisms for irradiated materials can be studied with small TEM disks by employing a disk-bend method. Results showed that dominant deformation modes in ion-irradiated austenitic steel were planar glide and microtwinning. It was concluded that the disk-bend method will be a powerful tool in studying deformation microstructures of irradiated materials. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Ductility loss and embrittlement of steels associated with radiation induced defects and helium accumulation have been an important subject of investigation in the fission and fusion reactor materials research communities [1]. Recently, the development of accelerator based spallation neutron sources are underway in the US, Europe, and Japan [2–4]. In such a system, as a result of nuclear reactions, helium will be generated at a rate of ≈200 appm/dpa, which is about 100 times higher than in fission reactor materials and about 10 times higher than in fusion-irradiated materials. Hence, the effects of helium build-up in the target vessel steel have come under intense scrutiny. In recent studies on proton and heavy ion irradiated AISI 316 austenitic stainless steel, a significant finding was that radiation induced hardening was dramatically augmented at high helium concentrations [5,6].

In the past, various in-reactor irradiation experiments have been carried out to study or monitor the radiation effects on structural components subjected to a radiation environment. At the same time, many simulation experiments have been conducted for various Fe–Cr–Ni base austenitic steels using accelerators to understand the radiation induced phenomena [7]. However, most simulation studies were confined to an examination of the radiation induced microstructure of helium implanted 316L austenitic stainless steel [8] and the measurement of surface hardness with a sensitive hardness tester, such as the Nanoindenter® for 316LN austenitic stainless steel irradiated by H, He, and Fe ions in single, dual, and triple ion-beam irradiation modes [9]. Thus far, only a few experiments have been conducted to study the deformation mechanisms for ion irradiated materials (i.e., 304L austenitic stainless steel) [10], mainly because the ion damage is confined to the near-surface region, not to the bulk. Although the deformation microstructure has been examined for various in-reactor irradiated and tensile tested specimens including Fe–Cr–Ni base steels, and Al, Cu, Nb, Mo, Re, and V alloys [11–13], such a process is very time consuming, expensive, and strenuous because of long irradiation time and difficulty in handling radioactive materials. Moreover, a well-designed, systematic study is often difficult in reactor irradiation because of system constraints, such as controlling temperature and dose.

In this work, a method to study deformation mechanisms using a small ion irradiated transmission electron microscopy (TEM) disk is developed for the first time, although many investigators have employed the

* Corresponding author. Tel.: +1-865 574 5058; fax: +1-865 574 0641.

E-mail address: LeeEH@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(00)00130-6

indentation technique to study mechanical properties in the past [14]. The method involves irradiation of a small steel disk under controlled conditions, deformation of the disk to a desired level by the disk-bend method, electro-chemical thinning of the strained disk to reveal the irradiated and deformed region, and examination of the microstructure. The results show that deformation mechanisms can be studied economically and systematically by employing ion irradiation, with easy variation of irradiation conditions (e.g., temperature, dose, gas atom injection etc.). Thus far, several series of experiments have been conducted to investigate the effects of radiation induced defects on mechanical deformation mechanisms. In this report, a procedure for specimen preparation and test method are described to demonstrate the feasibility of this new approach. An example of the test results is presented to demonstrate the power and convenience of this new approach in studying the effects of radiation on deformation mechanisms.

2. Experimental

One alloy investigated in this experiment was a solution annealed AISI 316LN austenitic stainless steel (Jessop Steel Company Heat # 18474), which is a candidate material for the future spallation neutron source (SNS) target container vessel. The nominal composition of the alloy was, in weight percent, 16.3 Cr, 10.2 Ni, 2.01 Mo, 1.75 Mn, 0.39 Si, 0.009 C, 0.11 N, 0.029 P with the balance Fe. Disks of 3 mm diameter and 0.25 mm thickness were prepared by mechanical and electrochemical polishing prior to irradiation. Irradiation was carried out with 360 keV He^+ ions at 200°C using a 2.5 MV Van de Graaff accelerator at ORNL. The 360 keV ion energy resulted in a maximum gas atom deposition near a depth of 750–850 nm, according to a depth calculation by the computer code, Stopping and Range of Ions in Matter (SRIM, 1998 version) [15]. The 360 keV He^+ ion also produced atomic displacements at a rate of $\simeq 7.5 \times 10^{-5}$ dpa/appm at this depth. The specimen received about 2000 appm or 0.2 at.% helium and 0.15 dpa at the peak damage region. The procedure to calculate dpa is described in Ref. [16].

The irradiated disk was clamped into a recessed circular holder with a central opening area of 1.6 mm diameter, as illustrated in the schematic drawing in Fig. 1. The irradiated side of the disk was placed face-down, and a tungsten carbide (WC) ball of 1 mm diameter was placed on top of the unirradiated side of the TEM disk. A load up to ≈30 N was applied on the WC ball at a strain rate of $\approx 10^{-2}$ per second at room temperature using a Tinius Olsen testing machine until the average plastic strain of the center region of the disk attained about 7%.

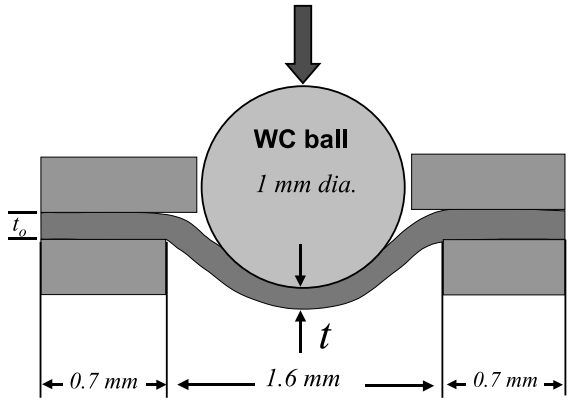


Fig. 1. Schematic drawing of a disk bend setup. The thickness change after deformation was measured by the displacements of the machine cross-head in contact with the WC ball and a LVDT placed underneath the disk.

The average plastic strain (ε_p) was calculated by employing a volume constant criterion and the relation, $\varepsilon_p = \ln(t_0/t)$, where t_0 and t are the thickness of the center region of the disk in contact with the WC ball before and after deformation, respectively. The thickness change after relaxation (zeroing the load) was determined by reading the displacement of the tensile machine cross-head bar in contact with the top of the WC ball and the displacement of a linear variable differential transducer (LVDT) placed underneath the TEM disk.

In this experimental set-up, the irradiated side of the TEM disk surface was strained in multiaxial tension. The irradiated surface strain was more than that at the surface in contact with the WC ball. A comparison of relative displacements of landmarks on the specimen surface before and after disk bend straining indicated that the local strains were not uniform due to varied grain orientations and slip directions. Such uneven strain is also true in a conventional tensile tested specimen, but it is not of great concern because the tensile properties are derived from an average strain of the entire test gauge. However, this inherent variation can be a significant source of uncertainty when a very limited area is examined, as in TEM. In consideration of these uncertainties, the strain at the irradiated surface was estimated to be in the range of 7–13%.

Specimen thinning for TEM was carried out by two methods. First, a thin TEM foil was prepared by electrochemically removing ≃700 nm from the ion bombarded side of the disk and then thinning from the unirradiated side until a perforation occurred. This procedure produced TEM foils with thicknesses of ≃100 nm on average and allowed examination of the microstructure at the peak damage region between 700 and 800 nm of the original depth. Secondly, the

irradiated and deformed TEM disk was Ni-plated electrochemically, sliced across the disk thickness in the ion beam direction, and electrochemically thinned until a perforation occurred across the irradiated region. This cross-sectioned and thinned specimen allowed us to examine the full damage profile and the transition of deformation modes from irradiated to unirradiated areas. Microstructures were examined with a Philips CM12 electron microscope operated at 120 keV and a JEOL FX2000 electron microscope operated at 200 keV. Optical microscopy was also carried out to examine the surface morphology after disk-bend deformation.

3. Results

Microstructures of unimplanted and 0.2 at.% He-implanted TEM disks were examined before and after deformation. Optical microscopy examination revealed a noticeable difference in slip characteristics between the unirradiated and He-implanted specimens as shown in Fig. 2; the bubble-like specks on the surface are stains caused by moisture condensation, which were removed during polishing for TEM foil preparation. The surface morphology of the pristine disk was characterized by numerous very fine slip steps with frequent cross-slip lines, whereas the implanted specimen showed fewer and coarser slip bands. The number of optically visible slip traces, including very faint traces, was about 5 per μm in the pristine specimen. About the same number of slip traces was observed by TEM. One notable observation was that the implanted specimen showed highly localized, large surface steps with very shallow slip traces between them, suggesting that the strain was confined mostly to larger slips and that such localized strain could be a degrading factor on mechanical properties.

Fig. 3 compares the TEM microstructures corresponding to Fig. 2 for the pristine and implanted specimens after deformation. A major deformation feature in the pristine specimen was random glide dislocations and formation of dislocation network cells. Despite the presence of random cross-slip, there was a tendency of dislocations to align along the {1 1 1} slip planes, a characteristic of fcc steels with low stacking fault energy.

At a helium level of 0.2 at.%, bubbles could not be resolved clearly by TEM, although a grainy bubble-like feature at high magnification suggested the presence of small bubbles. However, a high number density of extremely fine black-dots (radiation induced interstitial defect clusters) were present, Figs. 3(b) and 4(d). The deformation mode in the He-implanted specimen was characterized by extensive dislocation pileups on glide planes and formation of microtwins and stacking faults resulting from planar glide (movement of widely separated partial dislocations). TEM diffraction analyses revealed that twin formation occurred exclusively as a result of dislocation glide on {1 1 1} planes. Cross-slip was severely restricted in the presence of radiation induced defects.

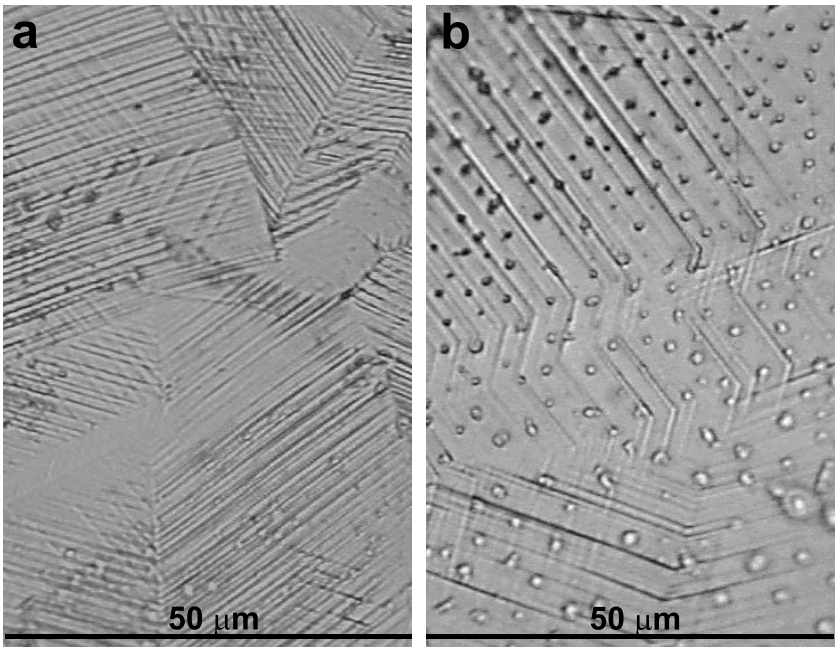


Fig. 2. Optical micrographs of AISI 316 LN austenitic stainless steel after deformation to an average strain of about 7%: (a) unirradiated and (b) implanted with 0.2 at.% helium by 360 keV He-ions at 200°C.

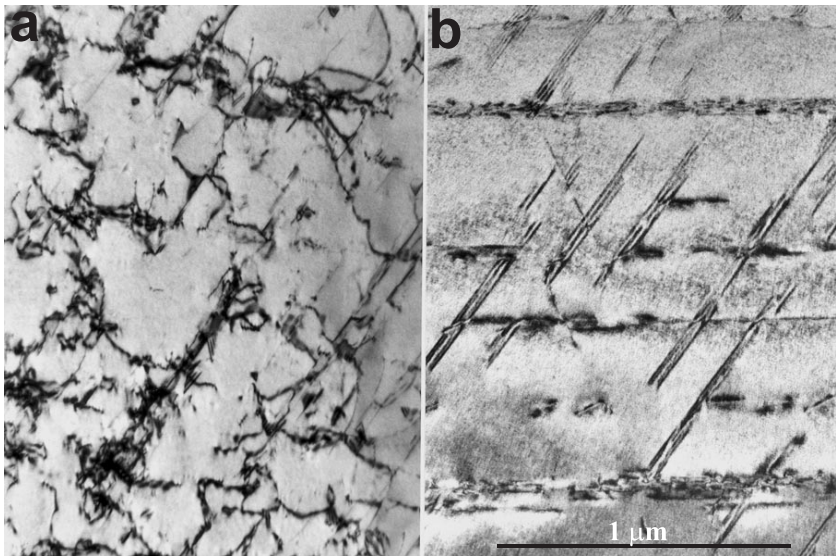


Fig. 3. TEM micrographs of deformation microstructure corresponding to Fig. 2: (a) unirradiated and (b) 0.2 at.% helium implanted with 360 keV He ions at 200°C.

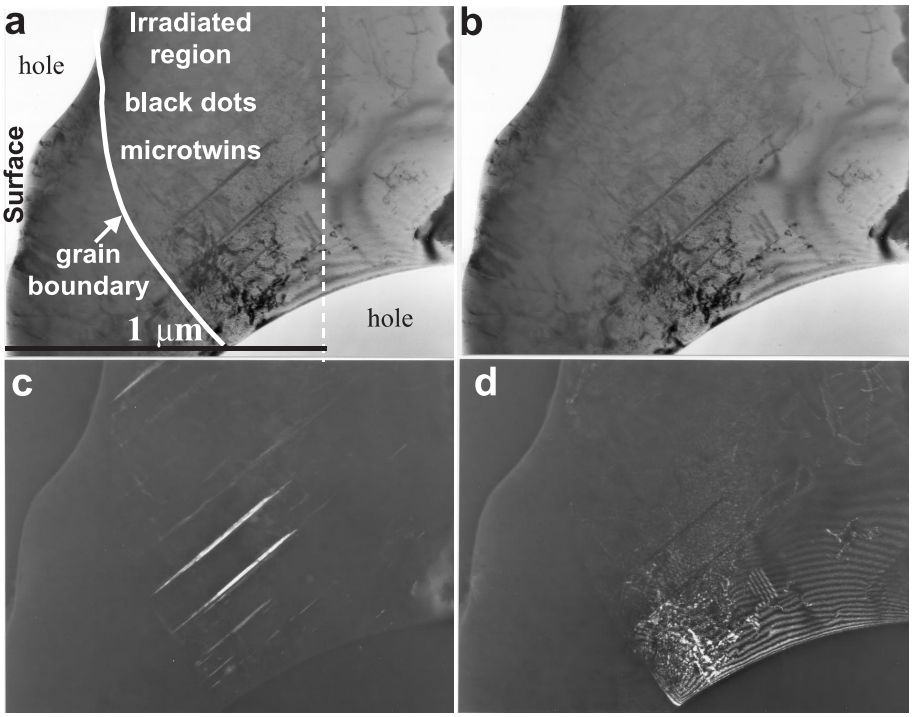


Fig. 4. TEM cross-sectional view of deformation microstructure for the specimen with 0.2 at.% helium implanted with 360 keV He ions at 200°C: (a) description of TEM viewing area; (b) bright field image taken near beam direction $B = [0\,1\,1]$ at $g = 200$ diffracting condition; (c) image taken from twin streaks in diffraction pattern, and (d) weak beam dark field (WBDF) image taken at $g = 200$, $g/5g$ condition.

Fig. 4(a) shows a thin cross-sectioned area with two holes around it because of the uneven polishing among the Ni-plated, He-implanted, and unirradiated bulk regions. The bright field micrograph showed the presence of black-dot defects and deformation microtwins in the He-implanted region (Fig. 4(b)). The dark-field micro-

graph taken from the twin streaks from the diffraction pattern further confirmed the presence of twin bands in the irradiated region (Fig. 4(c)). A weak beam dark field image showed the white strain-contrast spots arising from the black-dots, a few dislocation lines, and a grain boundary in the implanted region (Fig. 4(d)). In the left side of the irradiated grain, twins were also present but not visible because of the difference in grain orientation. A particularly interesting feature was that the deformation twins were confined only to the region where the radiation induced black-dot defects were present, and the twins were not present in the unirradiated region, even in the same grain.

4. Discussions

The deformation mechanism in the pristine specimen was characterized by dislocation multiplication and dislocation network cell formation. Although dislocations tended to align on {1 1 1} glide planes, the presence of randomly distributed dislocations and the slip offsets seen in the optical micrograph indicated that a significant fraction of dislocations cross-slipped during deformation. On the other hand, deformation in the He-implanted specimen was characterized by extensive linear dislocation pileup and microtwin formation. Cross-sectional examination of the irradiated and deformed areas showed that the deformation twins were observed only in the region with radiation induced defects, namely black dot interstitial clusters (perhaps also small helium bubbles), attesting that the twin formation was associated with radiation induced obstacles.

The shear displacement associated with plastic deformation occurred primarily by the motion of dislocations. It is well established that plastic deformation in fcc crystals occurs mainly by dislocation glide on energetically favorable close packed {1 1 1} planes in ⟨1 1 0⟩ directions [17]. Furthermore, in steels with low stacking fault energy such as austenitic stainless steel ($E_{sf} < 10$ J/m^2) [18], ⟨1 1 0⟩ **b** type dislocations tend to dissociate into two ⟨1 1 2⟩ **b** type Schockley partials having a stacking faulted (also twinned) region between them. That is, atoms move by zig-zagging along the energetically favorable valleys (⟨1 1 2⟩ **b**) on close packed {1 1 1} planes instead of directly over the hills (⟨1 1 0⟩ **b**). The partial dislocations repel one another to reduce the stress while the driving force to minimize the stacking faulted surface area between the two partials makes them attract one another. Thus, the spacing between the two leading (zig) and trailing (zag) partials is determined by the forces of repulsion and attraction balance at an equilibrium spacing that depends on the specific stacking-fault energy. The movement of either partial on a different plane would involve energetically unfavorable atomic movements, so the movement of the two widely separated Shockley partial dislocations (planar glide) occurs predominantly on the close packed planes. For these reasons, most slip steps (multiples of the unit dislocation vector **b**) appear straight in stainless steels. In fcc crystals, the close-packed {1 1 1} glide planes are also the coherent twin planes. Therefore, when a layer of atoms is removed from the normal order sequence by dislocation movement, stacking faults and twins are formed as a result of the violation of close packing on the fault plane [17].

Consistent with the mechanisms stated above, most slip traces in the pristine austenitic stainless steel were straight (Fig. 2). However, the presence of slip offsets indicated that cross-slip also occurred. TEM microstructure confirmed that although most dislocation lines aggregated on {1 1 1} planes, various segments of dislocation lines cross-slipped to different planes becoming screw or mixed-type dislocations. As a result, very few stacking faults and twins were observed.

In the He-implanted specimen, however, the alignment of slip traces was enforced, and micro-twins and stacking faults became the dominant deformation microstructures. It is known that cross-slip of an extended screw dislocation around obstacles is difficult, and shear occurs mostly by the successive motion of twinning dislocations, each on a glide plane one interplanar spacing removed from its predecessor [17,18]. Less frequently, however, application of suitably large stress can squeeze the partial dislocations together against a barrier to form a whole dislocation and facilitate cross-slip. Such an incidence often leads to a bifurcation of a channel band and becomes one of the mechanisms of widening a channel band. Clearly, our experimental observations indicated that the restricted dislocation motion on easy glide planes and the limited cross-slip observed in the He-implanted specimen were attributable to radiation induced obstacles (black dots). Although helium bubbles were not visible, probably small bubbles also played a role in restricting dislocation motion. Moreover, stacking faults (twins) themselves are important barriers to dislocation motion, and glide-twinning and twin–twin interaction lead to an opening of a gap in the twin if untwinning occurs successively on twin planes [19].

5. Conclusions

Radiation effects on deformation mechanisms were studied by employing a disk bend method for AISI 316LN stainless steel irradiated with 360 keV He ions at 200°C. The dominant deformation modes in an unirradiated specimen were dislocation multiplication and formation of dislocation network cells, and those in a He-implanted specimen were planar dislocation pileups and the formation of microtwins and stacking faults. Cross-slip occurred readily in unirradiated materials, but the shear deformation was dictated by planar glide

in the He-implanted specimen, suggesting that dislocation motions were confined mostly on glide planes due to the presence of radiation induced obstacles. The combined effects of radiation induced defects and deformation induced microstructure are considered to be a contributing factor to the degradation of mechanical properties of irradiated material.

An important outcome of this experiment is the demonstration that radiation effects on deformation mechanisms can be studied with small TEM disks by employing ion irradiation and a disk-bend test. This method can also be applied to in-reactor irradiated disks for speedy screening tests and for studying the radiation effects on mechanical deformation. At this time, an environment chamber capable of controlling the test temperature from −150°C to 400°C is being installed to study temperature effects. This new approach can be a powerful tool for studying deformation mechanisms of irradiated materials effectively, economically, and systematically under various simulation conditions.

Acknowledgements

This research was sponsored in part by the Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, US Department of Energy, under contract No. DE-AC05-900OR22725, and in part by the NERI Project IWO MSF 99-0072 with UT-Battelle, LLC. The authors thank Drs R.L. Klueh and A.F. Rowcliffe for technical review of the manuscript.

References

[1] Proceedings of the Eight International Conference on Fusion Reactor Materials (ICFRM-8), Sendai, Japan, 26–31 October 1997, J. Nucl. Mater. 258–263, 271&272 (1998).
[2] AccApp'98, American Nuclear Society, La Grange, IL, 1998.
[3] AccApp'99, American Nuclear Society, La Grange, IL, 1999.
[4] Second International Topical Meeting on Nuclear Applications of Accelerator Technology, Gatlinburg, Tennessee, 1998, American Nuclear Society, La Grange Park, Illinois, 20–23 September.
[5] H. Ulmaier, E. Camus, J. Nucl. Mater. 251 (1997) 262.
[6] E.H. Lee, J.D. Hunn, G.R. Rao, R.L. Klueh, L.K. Mansur, J. Nucl. Mater. 271&272 (1999) 385.
[7] E.H. Lee, L.K. Mansur, Metall. Trans. A 23 (1992) 1997.
[8] C.H. Zhang, K.Q. Chen, Y.S. Wang, J.G. Sun, J. Nucl. Mater. 258–263 (1998) 1623.
[9] E.H. Lee, G.R. Rao, J.D. Hunn, P.M. Rice, M.B. Lewis, S.W. Cook, K. Farrell, L.K. Mansur, in: M.S. Wechsler, L.K. Mansur, C.L. Snead, W.F. Sommer (Eds.), Materials for Spallation Neutron Sources, TMS, Warrendale, PA, 1998, p. 57.
[10] S.M. Bruemer, J.I. Cole, R.D. Carter, G.S. Was, Mater. Res. Soc. Symp. Proc. 439 (1997) 437.
[11] M.S. Wechsler, The Inhomogeneity of Plastic Deformation, American Society for Metals, Metals Park, OH, 1971, p. 19 (Chapter 2).
[12] P.D.K. Nathanson, P.J. Jackson, D.R. Spalding, Acta. Metall. 28 (1980) 823.
[13] N. Yamamoto, J. Nagakawa, Y. Murase, H. Shiraishi, J. Nucl. Mater. 258–263 (1998) 1628.
[14] P. Jung, H. Ullmaier (Eds.), Miniaturized specimen for testing of irradiated materials, in: Proceedings of the IEA International Symposium, Forschungszentrum Jülich GmbH, Germany, 1995.
[15] J.F. Ziegler, J.P. Biersack, U. Littmark, The Stopping and Range of Ions in Solids, Pergamon, Oxford, 1985.
[16] E.H. Lee, Nucl. Instrum. and Meth. B 151 (1999) 29.
[17] J.P. Hirth, J. Lothe, Theory of Dislocations, McGraw-Hill, New York, 1968 (Chapters 10 & 23).
[18] R.W. Hertzberg, Deformation and Fracture Mechanics of Engineering Materials, Wiley, New York, 1976 (Chapter 3).
[19] P. Müller, C. Solenthaler, Mater. Sci. Eng. A 230 (1997) 107.

ELSEVIER

Journal of Nuclear Materials 280 (2000) 18–24

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Effects of helium on radiation-induced defect microstructure in austenitic stainless steel

E.H. Lee *, J.D. Hunn, T.S. Byun, L.K. Mansur

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 26 January 2000; accepted 7 March 2000

Abstract

In the construction materials surrounding the spallation neutron source (SNS) mercury target, considerable quantities of transmutation products, particularly hydrogen and helium, will be generated due to the exposure to a high flux of 1 GeV protons and associated neutrons. In an effort to investigate the effects of high helium, therefore, bubble formation and defect clustering processes in AISI 316 LN austenitic steel were studied as a function of helium concentration and displacement damage dose with 360 keV He^+ and 3500 keV Fe^+ ion beams at 200°C. Helium irradiation was less effective in producing defects such as black dots and dislocation loops than Fe^+ ion irradiation at equivalent displacement dose. On the other hand, the formation of helium bubbles produced a strong depressive effect on the growth of loops and the evolution of line dislocations. The results indicated that the effect of helium bubbles was augmented as the bubble number density and size increased with increasing helium beyond 1 atomic percent (at.%). In such a case, the effect of helium bubbles can be more important than that of radiation-induced defects on the evolution of microstructure and the change in mechanical properties. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

For future neutron scattering research, the construction of an accelerator based spallation neutron source (SNS) facility, operating at approximately 2 MW, is presently under design and construction [1]. The materials R&D program that is underway for the SNS target includes major efforts on spallation radiation effects [2]. Type AISI 316 austenitic stainless steel is the candidate material for the containment vessel of the mercury (neutron source). The containment vessel will be exposed to an intense pulsed 1 GeV proton beam and to neutrons generated by spallation reactions in the mercury. It is anticipated that atomic displacement rates will be high, up to $\approx 10^{-2}$ dpa/s (displacements per atom/second) during the microsecond beam pulse period, with a 60 Hz pulse frequency. The time average displacement rate would be in the range of $\approx 10^{-6}$ dpa/s. The operating temperature is expected to be below 200°C. Besides the radiation-induced displacement damage, various transmutation products will be generated as a result of nuclear reactions. In particular, hydrogen and helium will be produced at rates of ≈1000 and ≈100 appm/dpa, respectively.

One of the major concerns is the ductility loss of the vessel material due to radiation-induced hardening and helium effects, which could limit the serviceable lifetime of the vessel. In recent work [3], a dramatic increase in hardness was observed in association with bubble formation when helium was implanted to a level of 10 at.%. In our on-going work in parallel to the present study, to be published in a separate paper, it was found that the hardness increase by helium implantation alone began to exceed the saturation hardness value achieved by displacement damage with 3500 keV Fe^+ ions at 1 at.% He level and reached a saturation hardness value which was twice as high as that by Fe ions at 10 at.% He level. However, it is important to note that for the presently anticipated end-of-life dose for the target vessel of SNS (<10 dpa), the helium level would be below 2000 appm. Nonetheless, there is little systematic information to

* Corresponding author. Tel.: +1-865 574 5058; fax: +1-865 574 0641.

E-mail address: leeeh@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(00)00038-6

evaluate the effects of the high helium to dpa rates expected under the SNS condition. The aim of this work is to study the agglomeration process of helium as a function of helium concentration and its impact on microstructural change in order to provide information to guide design and lifetime estimates for the SNS target vessel.

In this work, AISI 316 LN austenitic stainless steel was irradiated with 360 keV He^+ or 3500 keV Fe^+ ion beams at 200°C with systematic increment in helium concentration and displacement dose. Microstructural evolution during helium irradiation was studied by comparing with the displacement induced defect microstructure by Fe^+ ions.

2. Experimental

A solution annealed AISI 316LN stainless steel (Jessop Steel Company Heat # 18474) was used for this investigation. The nominal composition of the alloy was, in weight percent, 16.3 Cr, 10.2 Ni, 2.01 Mo, 1.75 Mn, 0.39 Si, 0.009 C, 0.11 N, 0.029 P with the balance Fe. Disks of 3 mm diameter and 0.25 mm thickness were prepared by mechanical and electrochemical polishing prior to irradiation. Irradiation was carried out at 200°C with 360 keV He^+ or 3500 keV Fe^+ ion beams using the triple ion facility (TIF) at ORNL. Details of the TIF can be found in [4]. The ion energies were chosen for the maximum gas atom deposition and damage to occur near a depth of 750–850 nm. For the depth calculations, the computer code, Stopping and Range of Ions in Matter (SRIM, 1998 version) was used [5]. The procedure for the dpa calculation is described in [6]. The He^+ ion beam was applied with a deposition rate of $\simeq 0.2$ appm/s at the peak, which also produced a displacement damage at a rate of $\simeq 7.5 \times 10^{-5}$ dpa/appm or $\simeq 1.5 \times 10^{-5}$ dpa/s. This gives the helium appm/dpa ratio almost 100 times higher than that expected in the actual SNS target. The implanted helium concentration levels were from 2 appm (2×10^{-4} at.%) to 500 000 appm (50 at.%) and the corresponding doses were from $\simeq 1.5 \times 10^{-4}$ to $\simeq 37$ dpa. To appraise the damage microstructure produced by helium, a set of reference irradiation was carried out separately with 3500 keV Fe^+ ions at a rate of 2–3 $\times 10^{16}$ ions/m^2/s ($\simeq 1 \times 10^{-3}$ dpa/s at the peak).

Transmission electron microscopy (TEM) specimens were prepared by electrochemically removing $\simeq 700$ nm from the ion bombarded side of the disks and then thinning from the unirradiated side until perforation occurred. This procedure produced TEM foils with thicknesses of $\simeq 100$ nm on average and allowed examination of the microstructure at the peak damage region between 700 and 800 nm original depth. Damage microstructure was examined with a Philips CM12 electron microscope operated at 120 keV.

3. Results

The bubbles and defect clusters were examined after helium irradiation to various concentration levels. At low helium concentration levels ($\lesssim 2$ at.%), it was very difficult to image bubbles clearly by TEM because bubbles were too small in size ($r < 0.2$ nm) and too high in number density. However, the grainy microstructure revealed increasingly bubble-like features with increasing helium concentration in high magnification micrographs. Discernible bubbles could be imaged at a helium level of 5 at.% ($\sim 1.5 \times 10^{23}$ He/m^2). Patches of a well-defined lattice of bubbles were visible at 5 at.% helium level as indicated in the magnified inset of Fig. 1, although complete bubble alignment was not observed as found in 30 keV He^+ ion irradiated copper elsewhere [7]. The bubble lattice spacing was in the range of 1.5 ± 0.3 nm, the bubble radii were smaller than 0.3 nm, and the bubble number density was estimated to be in the range 2–6 $\times 10^{26}$ m^{-3}. With increasing helium concentration, bubbles assumed a more distinctive spherical shape, but bubble coalescence was insignificant up to 10 at.% helium. A significant fraction of bubbles coalesced at 20 at.% helium as shown in Fig. 2 and a severe blistering occurred at the irradiated surface at 50 at.%. Although the blistering might have caused some variation in dose from area to area, the larger and coarser bubble microstructure in the 50 at.% He-implanted specimen indicated that the examined area appeared to be an area not affected by the blistering.

As already mentioned, helium implantation also produced atomic displacement damage at a rate of $\simeq 7.5 \times 10^{-5}$ dpa/appm. Black dots (vacancy and interstitial clusters) and Frank type faulted interstitial loops evolved on {1 1 1} planes with increasing helium concentration. It is known that most surviving defect clusters are interstitial clusters [8,9]. Fig. 3 illustrates defect microstructures for 0.02, 0.2, 2, and 10 at.% helium, which also produced 0.015, 0.15, 1.5 and 7.5 dpa damage, respectively. Very few defects were visible at 0.02 at.% (0.015 dpa). At 0.2 at.% (0.15 dpa), a high number density of black dots evolved but were too small to produce good images within the TEM resolution limit. At 2 at.% (1.5 dpa), black dots were small ($\lesssim 1$ nm) but the number density approached a saturation level and occasional small loops began to develop. The number density of black dots and loops was estimated to be in the range 2–4 $\times 10^{23}$ m^{-3} from high magnification micrographs. We use the term loops when defect clusters become large (>5 nm) and take on a planar disk morphology. With increasing helium fluence, some black dots were converted to Frank type interstitial faulted loops as they grew but there was very little indication of unfaulting to prismatic loops even at 50 at.%. Loop growth was noticeably suppressed during He^+

irradiation compared to that of Fe^+ irradiation, as described below.

Although a similar defect evolution occurred during 3500 keV Fe^+ ion irradiation, black dot formation and loop growth occurred at somewhat lower dose in Fe^+ irradiation than in He^+ irradiation. Neither bubbles nor cavities were observed after Fe^+ irradiation. Fig. 4 illustrates the defect microstructure after

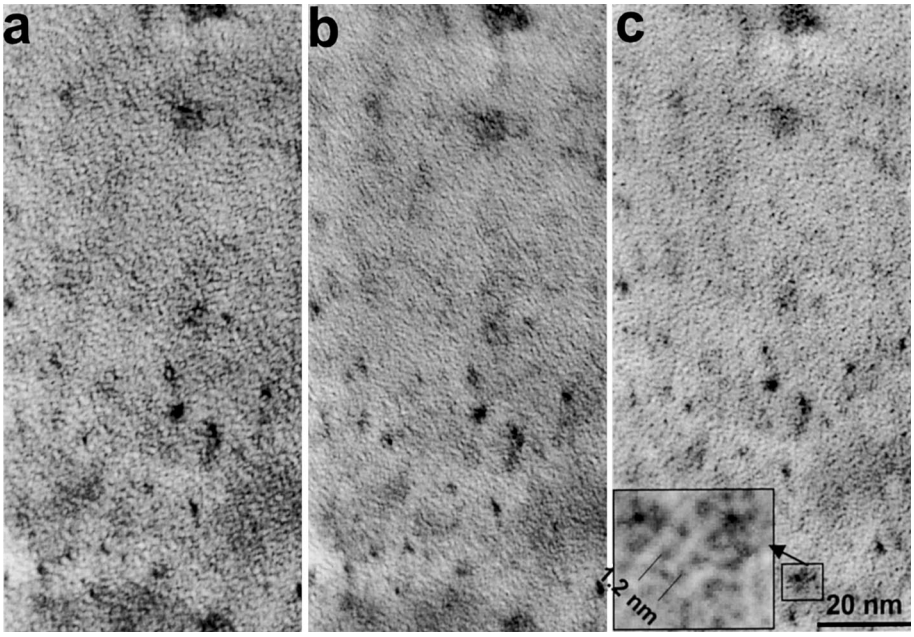


Fig. 1. Bubble microstructure after 5 at.% helium implantation with 360 keV He^+ ions at 200°C. The micrographs are taken near $Z = [110]$: (a) under-focus, (b) near-focus, and (c) over-focus images. The magnified inset shows an area of aligned bubbles in which the spacing between the bubble lattices is found to be in the range of 1.5 ± 0.3 nm and the bubble number density is estimated to be in the range of $2–6 \times 10^{26}$ m^{-3}.

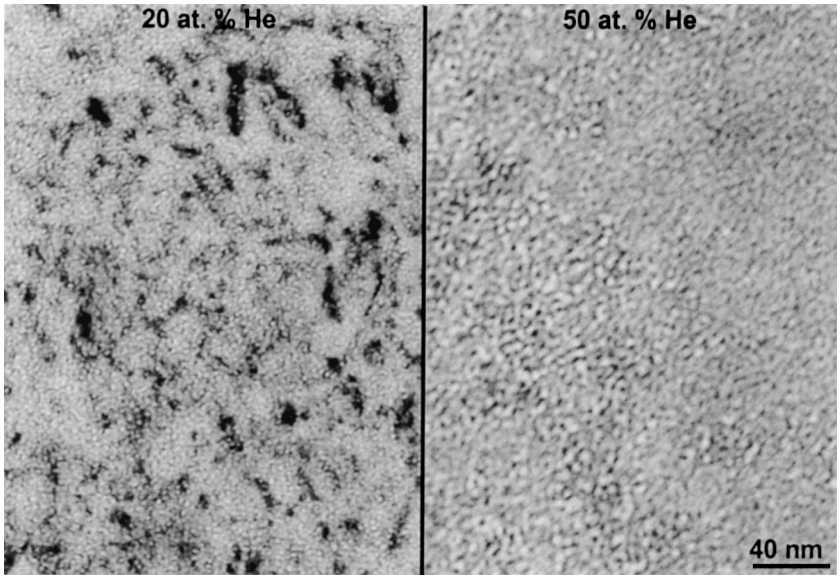


Fig. 2. Bubble microstructure after (a) 20 and (b) 50 at.% helium implantation with 360 keV He^+ ions at 200°C. A significant fraction of bubbles coalesced at 20 at.% and a severe blistering occurred on the irradiated surface at 50 at.%.

0.01, 0.1, 1, and 10 dpa Fe^+ ion irradiation. At 0.01 dpa, defects were not visible. At 0.1 dpa, black dots were well established and some small loops began to evolve. At 1 dpa, small but copious well-defined loops were visible. At 10 dpa, some large faulted loops began to unfault to prismatic loops. The black dot and loop number density appeared to saturate between 0.1 and 1 dpa. Here, it should be pointed out that the image of a small loop or black dot comes from the strain-field around the cluster and varies strongly with Bragg diffracting condition, so it is often difficult to make a precise size determination from a TEM image; the defect cluster sizes at 0.1 dpa appeared to be larger than at a dose of 1 dpa because of the difference in diffracting conditions and partly due to streaking.

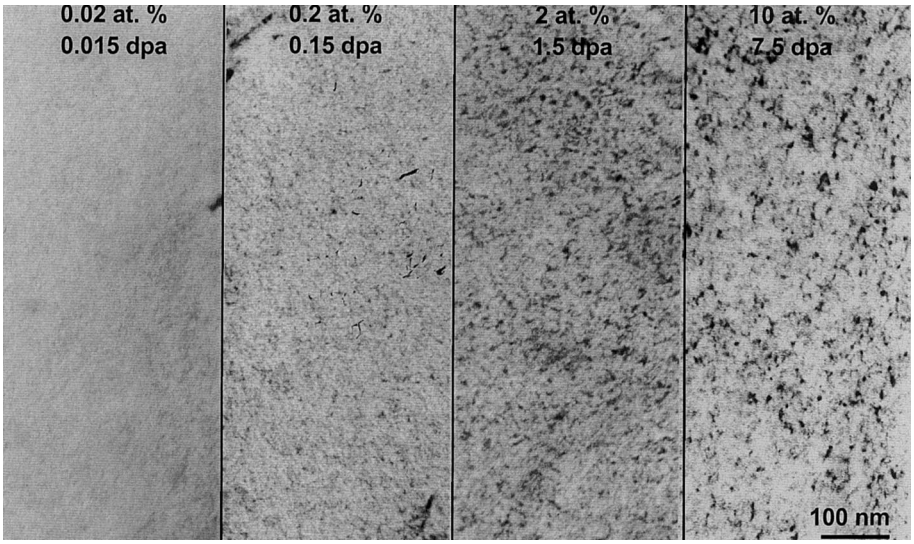


Fig. 3. Defect microstructure after 0.02, 0.2, 2, and 10 at.% (or 0.015, 0.15, 1.5, and 7.5 dpa) irradiation with 360 keV He^+ ions at 200°C.

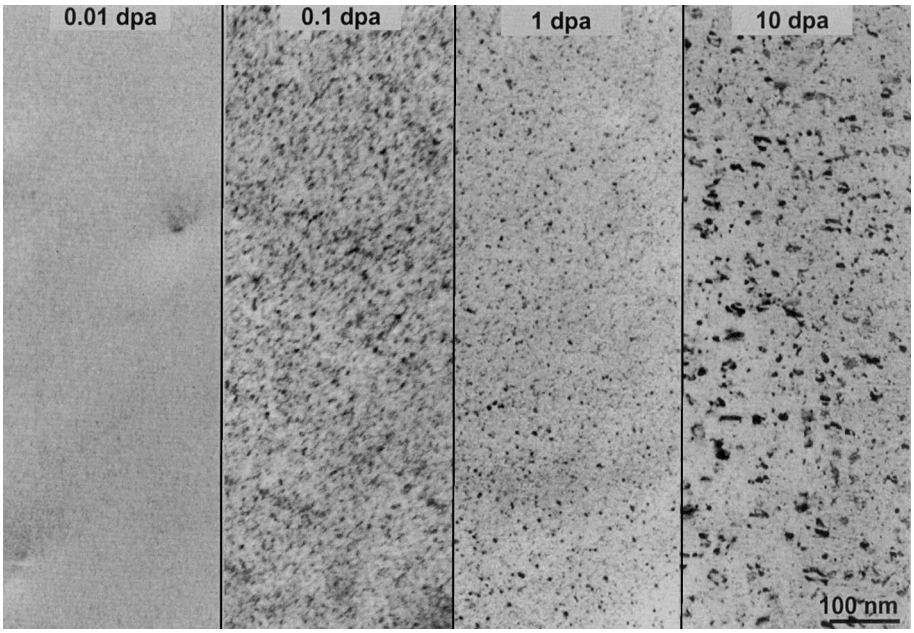


Fig. 4. Defect microstructure after 0.01, 0.1, 1, and 10 dpa irradiation with 3500 keV Fe^+ ions at 200°C.

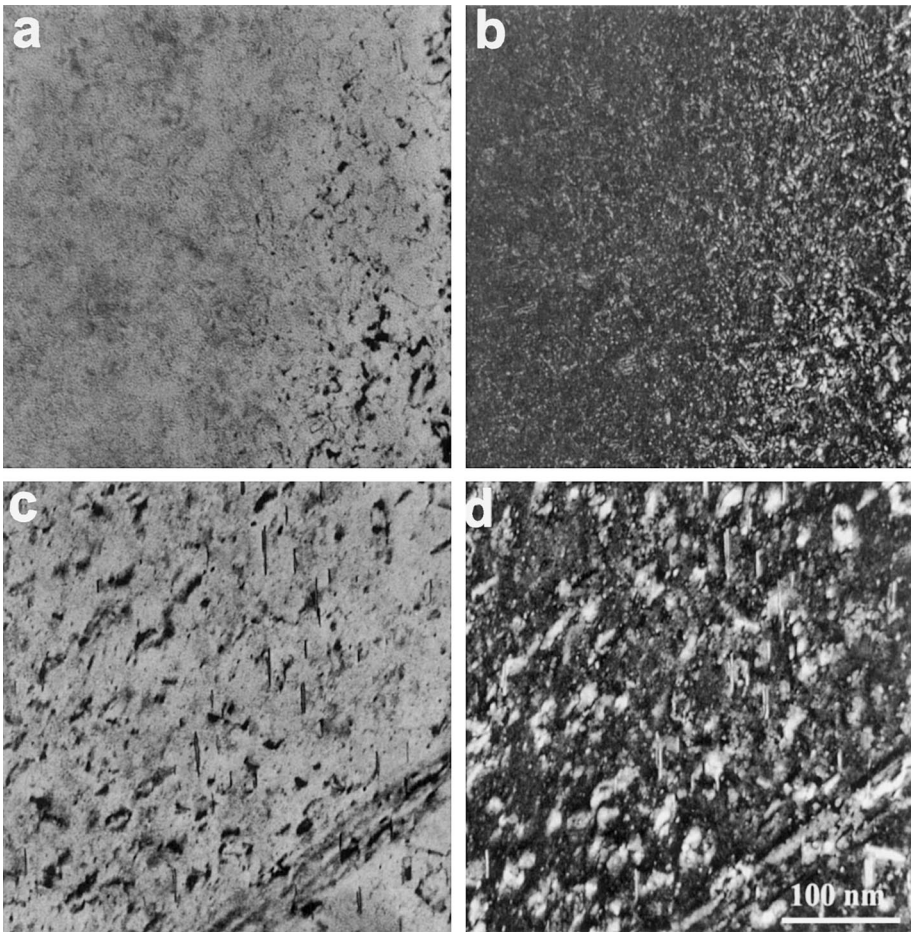


Fig. 5. Bright and dark field defect microstructure (a) and (b) after 50 at.% (37 dpa) He implantation, and (c) and (d) after10 dpa Fe^+ irradiation. Black dot and loop sizes in (a) and (b) are smaller despite the higher displacement damage dose.

Under He^+ irradiation, not only black dots evolved slowly but also loop growth was suppressed noticeably. Fig. 5 compares the defect microstructures for 50 at.% (37 dpa) He versus 10 dpa Fe. The defect clusters in the He-implanted specimen were noticeably smaller than those in the Fe irradiated one, even though the displacement dose for He was almost four times higher.

4. Discussion

As described in Section 2, the microstructures in the depth between 700 and 800 nm depth were examined after removing about 700 nm surface region. A calculation based on a rate theory model showed that point defects at this depth would not see the free surface effect, particularly at low temperature and high sink density conditions where most point defects were recombined or captured by sinks before any significant diffusion could occur [10]. Thus, the microstructures in the depth examined here were considered to represent the bulk behavior. The most notable observation was that the black dot evolution and loop growth was slower for He irradiation compared with Fe irradiation at equivalent dpa dose. The suppressed defect evolution is attributable to the lower displacement damage rate for He irradiation, smaller cascade energies produced by a lower mass and lower energy He^+ ion, enhanced point defect recombination at helium bubbles, and dislocation pinning by bubbles. These effects are described further below.

During irradiation, vacancies and interstitials are generated as a result of displacement damage. Some of these point defects are absorbed by pre-existing or generated defects such as dislocations, grain boundaries, precipitates, and cavities, and the rest are either annihilated via direct recombination or form clusters. At a given temperature, the point defect supersaturation decreases with decreasing damage rate, but the fraction of surviving point defects is higher at a lower damage rate, since the probability of point defect recombination is

lowered in proportion to the product of vacancy and interstitial concentration, which are lower at lower displacement rate [11]. However, the nucleation rate of defect clusters is lower at lower damage rate because the absolute point defect flux to sinks is lower. The employed He^+ ion energy (360 keV) was about 10 times lower than that of the Fe^+ ion (3500 keV) and the damage rate from He^+ ($\simeq 1.5 \times 10^{-5}$ dpa/s) was about two orders of magnitude lower than that from Fe^+ ($\simeq 1 \times 10^{-3}$ dpa/s). A slower defect-clustering rate is thus expected for He^+ ion irradiation because of the lower damage rate.

The second reason for the slower defect nucleation rate in He^+ irradiation is due to the smaller cascade energy of He^+ ions because of the lower ion mass and energy. When an ion penetrates into a material, the incident ion energy is transferred to the target atoms via nuclear and electronic energy loss processes. Atoms displaced from their lattice sites by elastic collisions with the ion are called primary knock-on atoms (PKAs). The PKAs have various energies depending upon the energy received from the ion. PKA energies are similarly transferred to other target atoms in subsequent secondary knock-on processes. The PKAs produced by an ion can be simulated in a full cascade SRIM calculation [5]. A 360 keV He^+ ion produces about 35 PKAs with energy between 40×10^3 and 5.2×10^3 eV, and a 3500 keV Fe^+ ion produces about 350 PKAs with energy between 40×10^6 and 1.6×10^6 eV. Fe^+ ions produce more and higher energy PKAs. About 40% of the PKA energy lies below 100 eV for both ions. Since an average displacement threshold energy for stainless steel is about 40 eV [12], the interstitial clustering for PKA energies below 100 eV is inconsequential within the damage rate range of interest. The value 40 eV was used for all SRIM simulations in this work. In a recent molecular dynamic (MD) simulations of cascade process in bcc iron [9], it was shown that the total number of surviving point defects and the resulting fraction of interstitials in clusters in a cascade increased with increasing cascade energy. Although the austenitic stainless steel studied in this work has a fcc crystal structure, for 3500 keV Fe^+ ion irradiation, a larger number of interstitial clusters per dpa is expected to form because of the larger PKA energy.

In a previous study, we found that defect evolution and void growth were suppressed by a high helium injection rate or in the presence of high bubble number density because of enhanced point defect recombination at bubble sinks [13,14]. Since bubbles were nucleated in high number during the 360 keV He^+ irradiation, point defect recombination at bubble sinks may be an important factor for the slow evolution of black dots and loops. An impeded loop growth has also been observed in the presence of high helium because of pinning by helium bubbles [15]. On the other hand, there is evidence that helium often promotes loop formation by retaining vacancies during the process of helium-vacancy clustering and by punching-out interstitials by over-pressurized helium bubbles [16,17]. In the present study, the relative importance between these two competing effects could not be differentiated. Nevertheless, the microstructural evidence indicated that bubble pinning and enhanced point defect recombination played important roles in loop growth because most loops remained small and faulted in the presence of high helium contents.

A notable finding was that very limited loop growth and unfaulting occurred even at 50 at.% (37 dpa) helium, even though the defect cluster number density saturated near 2 at.% (1.5 dpa). In fact, very little change in microstructure occurred beyond the saturation dose except bubble growth. The primary reason for this static defect microstructure is believed to be due to helium bubbles acting as a barrier to loop growth in addition to enhancing point defect recombination. Helium bubbles are considered to be a weak barrier for dislocation motion [18]. Up to now, there has been no information regarding the synergistic effect of bubble size and number density. However, hardness measurement data reported in other work [3] indicated that, when helium bubbles were present in such high concentration, the barrier strength (hardness) was augmented as bubbles grew, particularly when helium clusters began to become identifiable as bubbles under TEM. That is, at a helium level of above 1 at.%. This trend of microstructural evolution confirms the trend of hardness data measured for the 25 MeV helium implanted 316 L stainless steel by other investigators [3,19] and gives insight into the cause of the high hardness values measured in our more recent work.

5. Conclusions

To investigate the impact of high helium concentrations on a spallation neutron source target material, the microstructural evolution in an AISI 316LN stainless steel was studied after irradiation with 360 keV He^+ and 3500 keV Fe^+ ions. The results indicated that defects (black dots and loops) evolved slower in 360 keV He^+ irradiation than in 3500 keV Fe^+ ion irradiation when the defect microstructure was compared at equivalent damage doses. With increasing helium concentrations above 1 at.%, bubbles grew continuously but loop growth was severely impeded. The slower evolution and growth of defect clusters during 360 keV He^+ irradiation was attributed to the lower damage rate, smaller PKA energies, enhanced point defect recombination at bubble sinks, and dislocation pinning by helium bubbles. An important realization is that barrier strength of helium bubbles increase with increasing bubble size, and that helium bubbles impede not only the motion of disloca-

tions but also the evolution of loops and dislocations. Such effects can augment the radiation-induced hardening and ductility loss for steels subjected to a high helium generation environment. However, within the expected service period of the SNS target, hardening by helium appears to be insignificant.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract no. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The authors thank Drs K. Farrell and Roger E. Stoller for reviewing the paper.

References

[1] Spallation Neutron Source Design Manual, ORNL Report, September 1998.
[2] L.K. Mansur, Trans. ANS, TANSAO 80 (1999) 94.
[3] E. Camus, in: Proceedings of the Second International Workshop on Spallation Materials Technology, Ancona, Italy, 19–22 September 1997, p. 445.
[4] M.B. Lewis, W.R. Allen, R.A. Buhl, N.H. Packan, S.W. Cook, L.K. Mansur, Nucl. Instrum. and Meth. B 43 (1989) 243.
[5] J.F. Ziegler, J.P. Biersack, U. Littmark, The Stopping and Range of Ions in Solids, Pergamon, Oxford, 1985.
[6] E.H. Lee, Nucl. Instrum. and Meth. B 151 (1999) 29.
[7] P.B. Johnson, D.J. Mazey, Radiat. Eff. 53 (1980) 195.
[8] M. Horiki, M. Kiritani, J. Nucl. Mater. 212–215 (1994) 246.
[9] R.E. Stoller, J. Nucl. Mater. 276 (1999) 22.
[10] L.K. Mansur, H.M. Yoo, J. Nucl. Mater. 85&86 (1997) 523.
[11] L.K. Mansur, in: G.R. Freeman (Ed.), Kinetics of Nonhomogeneous Processes, Wiley, New York, 1987, p. 377 (Chapter 8).
[12] ASTM E521, Standard Practice for Neutron Radiation Damage Simulation by Charged-Particle Irradiation, Annual Book of ASTM Standards, vol. 12.02, American Society of Testing and Materials, Philadelphia, PA.
[13] E.H. Lee, L.K. Mansur, Metall. Trans. A 21 (1990) 1021.
[14] E.H. Lee, L.K. Mansur, Metall. Trans. A 23 (1992) 1977.
[15] E.H. Lee, N.H. Packan, M.B. Lewis, L.K. Mansur, Nucl. Mater. Meth. B 16 (1986) 25.
[16] K. Farrell, R.W. Chickering, L.K. Mansur, Philos. Mag. A 53 (1986) 1.
[17] H. Trinkaus, Radiat. Eff. 78 (1983) 189.
[18] S.M. Bruemmer, J.I. Cole, R.D. Carter, G.S. Was, Mater. Res. Soc. Symp. Proc. 439 (1997) 437.
[19] H. Ulmaier, E. Camus, J. Nucl. Mater. 251 (1997) 262.

ELSEVIER

Journal of Nuclear Materials 278 (2000) 266–272

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Hardness and defect structures in EC316LN austenitic alloy irradiated under a simulated spallation neutron source environment using triple ion-beams

E.H. Lee *, J.D. Hunn, N. Hashimoto, L.K. Mansur

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 20 May 1999; accepted 17 September 1999

Abstract

For an assessment of the future US spallation neutron source (SNS) target performance, radiation induced hardening and microstructural evolution were investigated as a function of ion dose for EC316LN stainless steel. Irradiation was carried out using 3.5 MeV Fe^+, 360 keV He^+, and 180 keV H^+ simultaneous ion-beams at 200°C to simulate the damage, He and H production in the SNS target vessel wall. At low dose (< 1 dpa), the predominant defects were black dots whose number density saturated rapidly within a few dpa. This was followed by the evolution of interstitial loops whose number density saturated below 15 dpa. Although He-bubbles were not visible, severely scalloped loops suggested that the implanted He/H atoms existed in the form of small clusters. Comparison with reported neutron irradiation data showed that hardening and ductility loss occurred mostly in the black dot regime (< 1 dpa), but that good ductility (>20% elongation) was maintained up to a dose level of $\simeq$ 10 dpa. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Detailed design and construction of an accelerator-based high power spallation neutron source (SNS) for neutron science applications is underway at Oak Ridge National Laboratory. In this system, a type AISI 316 stainless steel vessel containing a liquid mercury target as the neutron source will be subjected to an intense pulsed proton beam and associated neutron flux, accompanied by heat and stress loads. The requirement for the target vessel is to withstand a pulsed beam of 1 GeV energy protons, which produces 2 MW [1].

Under SNS conditions, the peak damage rate at the vessel nose is expected to be $\simeq 1 \times 10^{-6}$ dpa/s on average and $\simeq 1 \times 10^{-2}$ dpa/s during the microsecond beam pulse period. In addition, gaseous atoms will be produced in the vessel wall by transmutation at a rate of 500–1000 appm H/dpa and 50–200 appm He/dpa. The target vessel will be damaged due to atomic displacements by primary protons, spallation neutrons and recoil atoms, and by the high concentrations of transmutation products such as He and H, and cyclic stresses induced by thermal and pressure pulses. The important radiation effects within the SNS operating temperature range will be hardening and ductility loss. Additional possible sources of damage are fatigue loads, liquid metal embrittlement and temperature gradient mass transfer associated with mercury. These effects are expected to determine the lifetime of the target assembly. A wealth of information on the environment and materials considerations is available in the reports of a series of workshops [2–4]. To appraise the damaging effects on the target, EC316LN SS specimens were irradiated under a simulated SNS irradiation environment by simultaneous irradiation with 3.5 MeV Fe^+, 360 keV He^+, and 180 keV H^+ ions. In two previous reports [5,6], the changes in microstructure and hardness by H, He, and Fe ions were systematically investigated by employing single, dual, and triple ion-beam irradiations. In

* Corresponding author. Tel.: +1-423 574 5058; fax: +1-423 574 0641.

E-mail address: leeeh@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(99)00239-1

this work, radiation induced hardening and microstructural evolution were studied as a function of ion dose for triple ion-beam irradiated EC316LN stainless steel.

2. Experimental

The nominal composition of the subject alloy (designated as EC316LN SS by European Community fusion energy materials research group) is, in weight percent, 17.4 Cr, 12.3 Ni, 2.3 Mo, 1.8 Mn, 0.46 Si, 0.02 C, 0.06 N with the balance Fe. Transmission electron microscopy (TEM) disks were annealed at 1050°C for 30 min and electrochemically polished prior to irradiation. Details of the triple ion beam facility can be found in Ref. [7]. The ion energies were chosen to render the maximum damage and gas atom depositions near 800 nm depth, determined by the computer code, Stopping and Range of Ions in Matter (SRIM, 1998 version) [8]. An approximate Fe^+ current to give a peak damage rate of $\simeq 10^{-3}$ dpa/s was applied continuously with 200 appm He/dpa and 1000 appm H/dpa injection rates. A conversion from ion fluence (ions/m^2) to displacement dose (dpa), calculated using the NRT model [9], is given by

$$\text{dpa} = \frac{0.8}{2E_\text{d}}\left(\frac{\text{d}E}{\text{d}x}\right)_\text{nuclear}\frac{\text{ion fluence}}{\text{target atomic density}}, \quad (1)$$

where E_d is a displacement threshold energy (about 40 eV for steel) and $(\text{d}E/\text{d}x)_\text{nuclear}$ is the nuclear energy loss per unit length per ion, fraction of ion energy expended to atomic nuclei by ion beam. In SRIM calculation, this is derived by adding phonon and displacement energies together. Most of the damage is caused by the Fe^+ ions. He^+ and H^+ ions contributed only about 1.8% and 0.6% of the damage, respectively. The dose ranges covered for this investigation were from 0.1 to 80 dpa. The operating temperature of the SNS target vessel is expected to be between 80°C and 130°C. Ion irradiations were thus carried out at 200°C in consideration of an upward shift in peak damage temperature with increasing damage rate [10]. With increasing damage rate, the point defect concentration does not rise in proportion with the dose rate because of enhanced recombination of vacancies and interstitials. Consequently, the relative fraction of point defects arriving at sinks (i.e., loops) decreases. To compensate, the point defect flux to sinks can be increased by increasing the temperature.

Hardness changes were measured using a nanoindentation hardness tester, which records hardness (load divided by the contact area of the indenter) as a function of contact indentation depth. It is known that the stress field under the indenter extends to about 7 times the indentation depth [11]. Consequently, for a material which has only $\simeq 1000$ nm deep irradiated layer, the viscoelastic properties of the underlying substrate are probed increasingly as the indentation depth increases. For this reason, hardness values at 200 nm contact indentation depth were used as reference values to indicate the relative increase in hardness due to the irradiation. The same reference indentation depth was also used in previous work [5,6]. A detailed procedure of hardness measurement and a basis for choosing the reference depth can be found in Refs. [12,13], respectively.

TEM specimens were prepared by electrochemically removing $\simeq 700$ nm from the ion bombarded side and then thinning from the unirradiated side until perforation occurred. This procedure produced TEM foils with thicknesses of $\simeq 100$ nm on average and allowed examination of the microstructure between 700 and 800 nm original depth. For examination, a JEOL FX2000 electron microscope operated at 200 keV was used.

3. Results

Hardness values at 200 nm contact indentation depth are plotted as a function of displacement dose in Fig. 1. The scatter in hardness data was attributable to experimental limits in determining dose, indenter tip geometry, surface and subsurface effects imposed upon shallow irradiated depth, and to non-uniform microstructure such as dislocations and hard inclusions. The shaded area in the figure shows the trend of hardness change with dose and also the estimated range of error. The data showed that, upon irradiation to 0.1 dpa, hardness increased rapidly to above 4 GPa from a pristine value of 2.7 GPa but showed only a small increase from 0.1 to 1 dpa and beyond. The hardness saturation at higher doses becomes more apparent if the data is plotted in a

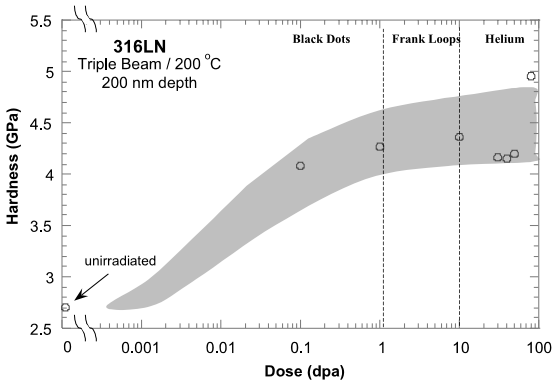


Fig. 1. Hardness variation as a function of ion dose for EC316LN stainless steel irradiated with triple ion-beams at 200°C. The shaded area is the trend of hardness increase with dose and also the range of estimated error.

linear dose scale. More controlled experiments which resolve the scatter are underway. Nonetheless, hardness increase showed a cogent relation to microstructural evolution with dose as described below.

Microstructural examinations were carried out on a different set of TEM disks triple ion-beam irradiated at 200°C to 0.2, 1.7, 15, 30, 36, and 80 dpa. Electron micrographs were taken at zone axis or beam direction **B** near $\langle 1\,1\,0\rangle$ and $\langle 1\,1\,2\rangle$, and the corresponding weak-beam dark-field images were taken at (g, $3g$) or (g, $5g$) diffracting condition, or from the diffraction streaks arising from faulted loops. TEM micrographs taken near $\mathbf{B} \simeq \langle 1\,1\,0\rangle$ are shown for 1.7, 15, and 80 dpa irradiated specimens in Fig. 2.

At 0.2 dpa, the predominant defects were small defect clusters less than a few nm in diameter which are often called black dots, and few loops were visible at this dose. Although not determined in this work, the damage-induced black dots are believed to be vacancy type stacking-fault tetrahedra (SFT) and small interstitial loops, according to Horiki and Kiritani's investigation [14]. At 1.7 dpa, loops became more visible, while black dot number density approached a saturation value of 1×10^{23}–2×10^{23} m^{-3}. Between 1.7 and 15 dpa, interstitial loops evolved continuously but black dot number density remained at about the same level. It should be pointed out that most of the hardness increase occurred in the black dot regime.

At 15 dpa, the loop microstructure appeared to be complete resulting in quite an extensive loop network structure. Overall, the defect microstructures of 15 and 80 dpa were very similar as shown in Fig. 3(a) and (b), respectively, in bright and dark-field images taken near $\mathbf{B} \simeq \langle 1\,1\,2\rangle$. It thus appeared that defect microstructure evolution might have reached a saturation near or below 15 dpa. The fishnet-like loop networks shown in the dark-filed images of Fig. 3 are due to overlapping loop images in projection. The loops were mostly faulted Frank interstitial loops on $\{1\,1\,1\}$ planes with a Burger's vector $\mathbf{b} = a\langle 1\,1\,1\rangle/3$. The complex and distorted Moire fringes shown in the bright field images of Fig. 3 are due to overlapping stacking fault fringes arising from the faulted loops on four different $\{1\,1\,1\}$ planes defined by the Thomson tetrahedron.

The saturation number densities for black dots and Frank loops were 1×10^{23}–2×10^{23} m^{-3} and 22×10^{22}–3×10^{22} m^{-3}, respectively. Black dots were smaller than a few nm in size and Frank loop sizes were in the range 5–60 nm in diameter at 80 dpa. It should be noted that the Frank loops shown in Fig. 2(c) (dark field) are about one-fourth of the total loops, because there are four $\{1\,1\,1\}$ planes in fcc. Almost all loops remained faulted with very little indication of unfaulting to prismatic loops ($\mathbf{b} = a\langle 1\,1\,0\rangle/2$) or growth to line dislocations; perhaps the resistance to unfaulting may have been due to the low irradiation temperature (200°C) and pinning effect by He/H clusters as described below.

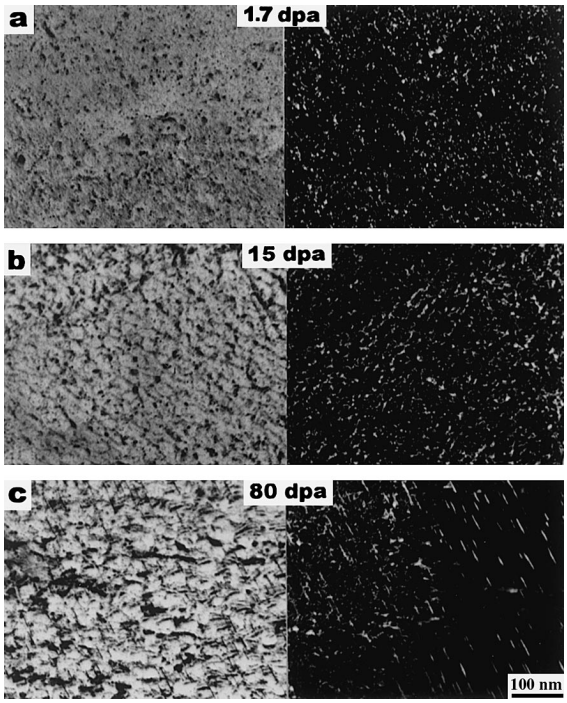


Fig. 2. Bright and dark field TEM micrographs taken near $\mathbf{B} \simeq \langle 1\,1\,0\rangle$ for: (a) 1.7 dpa, (b) 15 dpa, and (c) 80 dpa. Black dot number density approached a saturation near or below 1.7 dpa, and Frank interstitial loops evolved between 1 and 15 dpa.

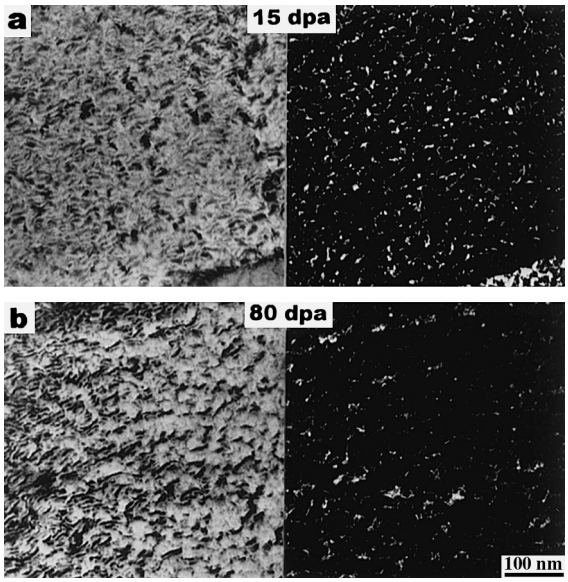


Fig. 3. Bright and dark field TEM micrographs taken near $\mathbf{B} \simeq \langle 1\,1\,2\rangle$ for: (a) 15 dpa, and (b) 80 dpa. The loop microstructures of 15 and 80 dpa were similar.

Despite the high concentration of injected helium (1.6 at.% at 80 dpa), no cavities were observed. However, the finely scalloped appearance of the loops indicated that loops are pinned probably by small invisible gas bubbles, (see the dark field loop images in Fig. 3). Helium bubbles are known to exert a pinning effect on loop growth and produce a scalloped loop structure due to bowing between two pivotal bubbles. Such an example was clearly demonstrated for loops formed in Ti-modified stainless steel irradiated at 675°C with 4 MeV Ni and 0.2–0.4 MeV He dual ion-beams [15] and P7 austenitic alloy neutron irradiated in the Oak Ridge Research Reactor (ORR) [16]. Helium clusters are also believed to contain hydrogen based on nuclear reaction analysis data reported previously, in that hydrogen retention was found to be enhanced in the presence of helium [6].

4. Discussion

In the temperature regime of primary interest for SNS components (<200°C), the triple ion-beam irradiation results showed that the defect cluster (black dot) density increased rapidly at a fraction of a dpa and saturated within a few dpa. Thereafter, a conversion of interstitial defect clusters to Frank loops occurred slowly up to $\lesssim 15$ dpa. Beyond that, no appreciable change in defect microstructure (black dots and Frank loops) occurred. The microstructure at 80 dpa was very similar to that at 15 dpa, suggesting that the defect microstructure had reached a dynamic steady-state near or below 15 dpa. Since black dots and loops are recombination sites for vacancies and interstitials, the negligible change in defect number density between 15 and 80 dpa is attributable to the high sink density attained at saturation and the ensuing enhanced recombination of point-defects.

The primary microstructural features introduced during triple ion-beam irradiation at 200°C were black dots, Frank loops, and, although not visible, small He/H clusters. From a relationship between number density (N) and average distance $1/(N)^{1/3}$ and our experimentally measured defect number densities at saturation, average distances among black dots and loops were found to be $\simeq 22$ and 32 nm, respectively. Similar microstructural evolution was also observed for austenitic steels irradiated by neutrons and ions [17–19], although microstructural saturation occurred at somewhat varying dose levels near 10 dpa, perhaps due to differences in alloy composition, thermomechanical treatment, irradiation temperature, damage rate, and helium generation rate.

During the triple-beam irradiations, hardness increased rapidly within a few tenths of a dpa and approached a saturation near 1 dpa, following the trend of black dot number density. Only a small increase in hardening occurred beyond a few dpa, suggesting that vacancy clusters, interstitial loops and helium made small contributions to hardening. A similar trend was also observed for various AISI 316 austenitic steels neutron irradiated below 100°C [20–23]. This temperature is considered to be comparable to the 200°C ion irradiation temperature because of the shift of peak damage temperature for the higher ion damage rate as mentioned already [10]. Figs. 4 and 5 illustrate the dose dependence of yield strength and elongation, respectively, for AISI 316 austenitic steels neutron irradiated below 100°C. The neutron data were compiled from

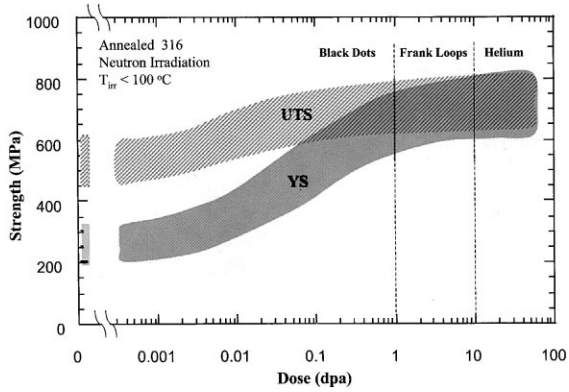


Fig. 4. Yield strength (YS) and ultimate tensile strength (UTS) as a function of dose for AISI 316 stainless steel irradiated by neutrons ($E > 1$ MeV) below 100°C. Both YS and UTS increase rapidly up to 1 dpa (black dot evolution regime) and thereafter show little change. The original data are from Refs. [19–21].

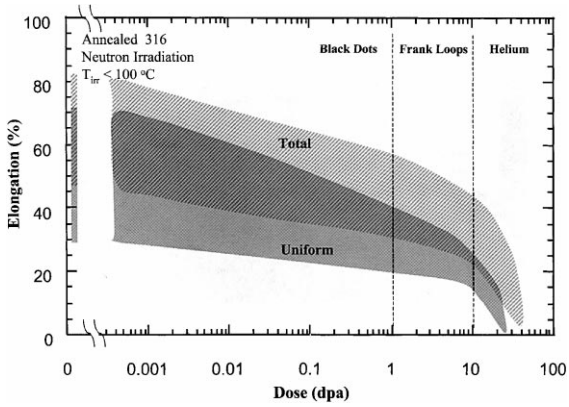


Fig. 5. Uniform and total elongation as a function of dose for AISI 316 stainless steel irradiated by neutrons ($E > 1$ MeV) below 100°C. The most rapid reduction in elongation occurs below 1 dpa (black dot evolution regime) although it declines continuously with increasing dose. The original data are from Refs. [19–21].

various sources, so there is considerable scatter in data. In spite of the scatter, it is clear that most of the yield strength increase occurred very rapidly within a few tenths of a dpa and showed very small increase beyond 1 dpa, similar to the trend of hardness (Fig. 4). Although the elongation decreased continuously with dose until a complete ductility loss occurred above 10 dpa, again the elongation reduction per unit dose was the highest in the black dot evolution regime (< 1 dpa), see Fig. 5.

It is known that radiation induced defects such as voids, bubbles, precipitates, black dots (vacancy clusters and small loops), Frank loops (large loops), and dislocations inhibit or retard dislocation motion, thereby increasing the flow stress of the material. All of these extended defects are responsible for the changes in mechanical properties for alloys subjected to irradiation [24–29]. Of these barriers, Frank loops are considered to be strong barriers for macroscopic deformation, followed by black dots as intermediate, and gas bubbles as weak barriers [29]. However, since the change in mechanical properties (hardness, yield strength, and elongation) was the largest per unit dose in the black dot dominant regime, the black dot number density seems to be the most important factor in affecting mechanical properties at low temperatures; note that the black dot number density was an order of magnitude higher than that of Frank loops.

Deformation in irradiated steel occurs primarily by normal dislocation slip, twinning, and channeling (annihilation of small defect clusters and loops on slip planes by glissile dislocations) [28–31]. In the channeling process, dislocations clear out defects along the path and subsequent dislocations tend to glide along the cleared path resulting in a channel, free of defects [29]. With increasing dose, dislocation motion, cross slip, and jog formation become progressively more difficult because of increasing interaction with defects (black dots and loops). Moreover, in an fcc system, a moving dislocation interacts with loops oriented in four different {1 1 1} planes [32], and the sessile nature of faulted loops, multiplicity of jog formation direction, and pinning by helium clusters all exert augmented barrier effects for moving dislocations. Thus, in a highly defected microstructure, most deformation is likely to occur by channeling with the surrounding defected region remaining undeformed [28]. In irradiated materials, therefore, localized plastic deformation tends to occur on these cleared channels without significant work hardening resulting in a failure with reduced elongation.

Under the SNS condition, the helium generation rate would be 50–200 appm/dpa, over two to three orders of magnitude higher than that of fission neutron irradiation ($\simeq$ 0.4 appm/dpa). Loops pinned by He/H clusters are not only difficult to grow but also difficult to shear and thus exert a stronger barrier effect to moving dislocations. This may accentuate localized deformation induced ductility loss. It is also known that helium accumulation at grain-boundaries causes embrittlement, but probably not at these low temperatures. Thus, a microstructural (black dot and loop) saturation alone may not be a sufficient criterion in evaluating materials; the possible effects of helium/hydrogen must also be considered. In fact, an almost total loss of uniform elongation in tensile tests was observed at a dose level of 3 dpa for austenitic stainless steels irradiated below 100°C by 800 MeV protons at the Los Alamos Neutron Scattering Center (LANSCE) accelerator, where He and H generation rates were comparable to those of SNS [33]. A substantial ductility loss (strain reduction from $\simeq$ 30% to $\simeq$ 8%) was also observed for a AISI 316 L steel implanted with 0.45 at.% helium ($\simeq$ 0.45 dpa) at 300°C [34].

A relation between microstructure and macroscopic deformation characteristics is illustrated in Fig. 6. In our triple-beam experiments, microstructural evidence showed that black dot number density was almost an order of magnitude higher than that of loops at the saturation dose. The initial hardening and yield strength increase can be attributed entirely to black dots (Figs. 1 and 4). Although the elongation reduction rate was the largest in the black dot evolution regime (Fig. 5), it declined continuously with dose throughout the entire range of microstructural evolution up to $\simeq$ 10 dpa. Since black dots evolved at the earliest dose and in the highest number density, they acted as the major barrier to the onset of plastic deformation (< 1 dpa), whereas Frank loops and He/H clusters produced an additional resistance to plastic deformation at higher doses as they evolved. With increasing temperature, however, the ratio of black dots to Frank loops declines as black dots

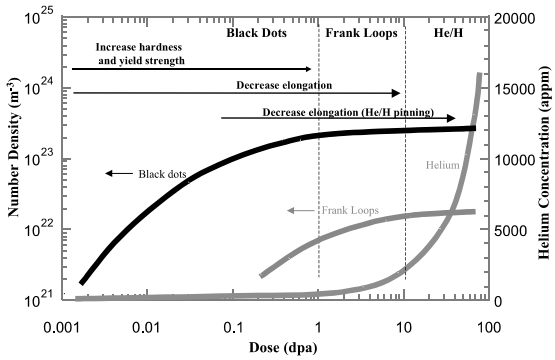


Fig. 6. Schematic representation of microstructural evolution and expected macroscopic mechanical property changes for EC316LN steel irradiated by triple ion-beams at 200°C. This representation may also be relevant for the materials subjected to a fission or fusion reactor irradiation environment, particularly at low temperatures (<400°C) where dynamic recovery is not significant during irradiation.

convert to loops. With further increase of temperature ($\gtrsim 350°C$), loops grow faster as the number density declines due to enhanced kinetics, and cavity formation is also facilitated. Thus a maximum loop and cavity barrier strength appeared around 300–330°C for the neutron case [17,20,29].

Although hardness change showed a similar dose dependence as yield strength, neither hardness nor yield stress alone could provide adequate information to evaluate ductility loss by irradiation. They cannot be used as a criterion to determine the degradation of mechanical performance. Importantly, a comparison between neutron data and triple ion-beam microstructure indicated that a severe reduction in uniform elongation (macroscopic deformation) may occur at a dose level where the defect microstructure (black dots and loops) saturates. However, under SNS conditions, ductility loss may occur at a lower dose level because of the high helium and hydrogen generation. This seems to be consistent with the data from a 800 MeV proton irradiation experiment [33].

5. Conclusions

For an assessment of the performance of future SNS target materials, hardness change and microstructural evolution during irradiation were investigated for EC316LN austenitic steel irradiated by triple ion beams (3.5 MeV Fe^+, 360 keV He^+, and 180 keV H^+) at 200°C. The results showed that black dots evolved very rapidly within a few tenths of a dpa and saturated at 1×10^{23}–2×10^{23} m^{-3} below $\simeq 2$ dpa. Frank interstitial loops started to evolve below $\simeq 1$ dpa and saturated at 2×10^{22}–3×10^{22} m^{-3} near or below 15 dpa. Although cavities were not visible, severely scalloped loop features suggested that loops were decorated by small He/H clusters.

From triple ion-beam, neutron, and 800 MeV proton irradiation data, it was shown that black dots, which appeared at the earliest dose and in the highest number density, dictated the changes in mechanical properties. Thus, most of the increase in hardness and yield strength occurred in the black dot evolution regime (< 1 dpa). Although the highest elongation reduction rate was also in the black dot regime, elongation was reduced continuously as Frank loops and He/H clusters evolved, suggesting that black dots, Frank loops, and He/H clusters, all contributed.

A direct comparison between triple ion-beam and neutron data may not be exact because of the three orders of magnitude difference in damage rate. However, at low temperatures, theoretical calculations based on a kinetic model revealed that the interstitial and vacancy cluster strength would become less sensitive to damage rate because defect supersaturation would dominate over a thermal annealing effect [27]. Guided by the present ion data, and neutron data obtained previously, a significant ductility loss is expected near a dose level of $\simeq 10$ dpa, at which the loop microstructure saturates and He/H accumulation becomes significant. However, the helium and hydrogen effects should be further studied for a better assessment of the SNS target lifetime.

Finally, an important outcome from this work is the suggestion that macroscopic materials properties might be appraised from small TEM specimens. This mapping would require definitive microstructural correlation to macroscopic deformation behavior as well as a theoretical basis to confirm such a relationship. Information gained from this work is not only of immediate interest for the SNS project, but also provides a stepping stone to establish the fundamental basis to map mechanical properties of radiation-induced microstructure for materials subjected to radiation. Such information would also enable better alloy design for fission and fusion reactor materials.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The authors thank Drs K. Farrell and A. F. Rowcliffe for technical review of the manuscript.

References

[1] T.A. Gabriel, in: Proceedings of Third International Workshop on Spallation Materials Technology, Santa Fe, New Mexico, 29 April–4 May 1999.
[2] L.K. Mansur, H. Ullmaier (Eds.), Proceedings of International Workshop on Spallation Materials Technology, Oak Ridge, Tennessee, 23–25 April 1996.
[3] F. Carsugi, L.K. Mansur, W.F. Sommer, H. Ullmaier (Eds.), Proceedings of Second International Workshop on Spallation Materials Technology, Ancona, Italy, 19–22 September 1997.
[4] M.S. Wechsler, L.K. Mansur, C.L. Snead, W.F. Sommer (Eds.), Proceedings of the Symposium on Materials for Spallation Neutron Sources, a publication of TMS, Orlando, Florida, 10–12 February 1997.
[5] E.H. Lee, G.R. Rao, J.D. Hunn, P.M. Rice, M.B. Lewis, S.W. Cook, K. Farrell, L.K. Mansur, in: M.S. Wechsler, L.K. Mansur, C.L. Snead, W.F. Sommer (Eds.), Proceedings of the Symposium on Materials for Spallation Neutron Source, 1997 TMS Annual Meeting, Orlando, Florida, TMS, Warrendale, Pennsylvania, 1998, p. 57.
[6] J.D. Hunn, M.B. Lewis, E.H. Lee, AccApp98 – Second International Topical meeting on Nuclear Applications of Accelerator Technology, American Nuclear Society, La Grange Park, 1998, p. 375.

[7] M.B. Lewis, W.R. Allen, R.A. Buhl, N.H. Packan, S.W. Cook, L.K. Mansur, Nucl. Instrum. and Meth. B 43 (1989) 243.
[8] J.F. Ziegler, J.P. Biersack, U. Littmark, The Stopping and Range of Ions in Solids, Pergamon, NY, 1985.
[9] M.J. Norgett, M.T. Robinson, I.M. Torrens, Nucl. Eng. Design 33 (1975) 50.
[10] L.K. Mansur, J. Nucl. Mater. 206 (1993) 306.
[11] L.E. Samuels, T.O. Mulhearn, J. Mech. Phys. Solids 5 (1957) 125.
[12] W.C. Oliver, G.M. Pharr, J. Mater. Res. 7 (1992) 1564.
[13] E.H. Lee, Y. Lee, W.C. Oliver, L.K. Mansur, J. Mater. Res. 8 (1993) 377.
[14] M. Horiki, M. Kirtani, J. Nucl. Mater. 212–215 (1994) 246.
[15] E.H. Lee, N.H. Packan, M.B. Lewis, L.K. Mansur, Nucl. Instrum. and Meth. B 16 (1986) 251.
[16] R.E. Stoller, G.R. Odette, J. Nucl. Mater. 154 (1988) 286.
[17] S.J. Zinkle, P.J. Maziasz, R.E. Stoller, J. Nucl. Mater. 206 (1993) 266.
[18] N. Hashimoto, E. Wakai, J.P. Robertson, J. Nucl. Mater. 273 (1999) 95.
[19] J. Gan, T. Allen, G.S. Was, Mater. Res. Soc. Symp. Proc. 439 (1997) 445.
[20] J.P. Robertson, I. Ioka, A.F. Rowcliffe, M.L. Grossbeck, S. Jitsukawa, in: R.K. Nanstad, M.L. Hamilton, F.A. Garner, A.S. Kumar (Eds.), Effects of Radiation on Materials: 18th Inter. Symp., ASTM STP 1325, American Society for Testing and Materials, Philadelphia, 1997.
[21] K. Farrell, L.K. Mansur, in: L.K. Mansur, H. Ullmaier (Eds.), International Workshop on Spallation Materials, 23–25 April 1996, Oak Ridge, TN.
[22] A.F. Rowcliffe, S.J. Zinkle, J.F. Stubbins, D.J. Edwards, D.J. Alexander, J. Nucl. Mater. 258–263 (1998) 183.
[23] M.L. Grossbeck, K. Ehrlich, C. Wassilew, J. Nucl. Mater. 174 (1990) 264.
[24] A.L. Bement Jr., in: Proc. on the Strength of Metals and Alloys, ASM, Metals Park, 1970, p. 693.
[25] M.L. Grossbeck, P.J. Maziasz, R.F. Rowcliffe, J. Nucl. Mater. 191–194 (1992) 808.
[26] G.E. Lucas, J. Nucl. Mater. 206 (1993) 287.
[27] R.E. Stoller, in: A.S. Kumar, D.S. Gelles, R.K. Nanstad, E.A. Little (Eds.), 16th International Symposium ASTM STP 1175, American Society for Testing and Materials, Philadelphia, 1993, p. 394.
[28] J.I. Cole, S.M. Bruemmer, J. Nucl. Mater. 225 (1995) 53.
[29] S.M. Bruemmer, J.I. Cole, R.D. Carter, G.S. Was, Mater. Res. Soc. Symp. Proc. 439 (1997) 437.
[30] F.A. Smidt Jr., Dislocation channeling in irradiated metals, NRL Report 7078, 1971.
[31] M.S. Wechsler, The Inhomogeneity of Plastic Deformation, American Society For Metals, Metals Park, OH, 1971, p. 19.
[32] S.G. Song, J.I. Cole, S.M. Bruemmer, Acta Metall. 45 (1997) 501.
[33] S.A. Maloy, LANL Report TPO-RGN-0013 (1/1999).
[34] H. Ullmaier, E. Camus, J. Nucl. Mater. 251 (1997) 262.

ELSEVIER

Journal of Nuclear Materials 278 (2000) 1–10

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Fe–15Ni–13Cr austenitic stainless steels for fission and fusion reactor applications. I. Effects of minor alloying elements on precipitate phases in melt products and implication in alloy fabrication

E.H. Lee *, L.K. Mansur

Metals and Ceramics Division, Oak Ridge National Laboratory , P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 3 May 1999; accepted 13 September 1999

Abstract

In an effort to develop alloys for fission and fusion reactor applications, 28Fe–15Ni–13Cr base alloys were fabricated by adding various combinations of the minor alloying elements, Mo, Ti, C, Si, P, Nb, and B. The results showed that a significant fraction of undesirable residual oxygen was removed as oxides when Ti, C, and Si were added. Accordingly, the concentrations of the latter three essential alloying elements were reduced also. Among these elements, Ti was the strongest oxide former, but the largest oxygen removal (over 80%) was observed when carbon was added alone without Ti, since gaseous CO boiled off during melting. This paper recommends an alloy melting procedure to mitigate solute losses while reducing the undesirable residual oxygen. In this work, 14 different types of precipitate phases were identified. Compositions of precipitate phases and their crystallographic data are documented. Finally, stability of precipitate phases was examined in view of Gibbs free energy of formation. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 64.75.+g; 81.30.Mh

1. Introduction

This work was carried out in the course of developing Fe–Cr–Ni base austenitic stainless steels for fission and possible fusion reactor materials applications. Historically, various alloying elements have been added to steels during fabrication to improve or control ductility, hardenability, formability, and other properties such as resistance to corrosion, oxidation, and creep [1]. For reactor applications, however, additional considerations are required for alloying to mitigate the degradation of alloys by radiation induced cavity swelling, solute segregation, creep, and radiation hardening. Currently available heavy ion and neutron irradiation data suggest that small quantities of impurities or alloying elements can have large effects on the materials' response to radiation [2,3]. In particular, it has been known that while oxygen promotes cavity formation [4,5], fine carbide, phosphide, and silicide particles associated with Ti, C, P, and Si additions suppress cavity growth by trapping and dispersing gas atoms at particle–matrix interfaces (i.e., MC, M_2P, radiation induced γ') [6–8]. Therefore, it is important to minimize residual oxygen and to maintain certain precipitate forming elements to desired levels during alloy fabrication. Moreover, in the presence of excessive oxygen, the crucial alloying elements such as Ti, C, and Si form oxides, and their efficacy as carbide or silicide forming agents diminishes, although they appear to be present in alloys in chemical analysis. Therefore, it is

* Corresponding author. Tel.: +1-423 574 5058; fax: +1-423 574 0641.
E-mail address: LeeEH@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S 0 0 2 2 - 3 1 1 5 (9 9) 0 0 2 3 4 - 2

important to ensure that required alloying elements are maintained to desired levels and kept in an active state or in solution during alloy fabrication. With this objective in mind, a set of 28Fe–15Ni–13Cr base alloys were fabricated. Effects of alloying elements on precipitate phase stability and their impact on the final alloy compositions were systematically investigated. Finally, based on thermodynamic considerations, a proper alloy melting procedure, which facilitates the control of solute concentrations, is described. The same set of alloys was aged at 600°C and 675°C to study thermal effects in Part II (this issue), and also irradiated by heavy ions in Part III (this issue) and neutrons to investigate alloy responses to irradiation. Some of the ion and neutron data were reported previously in Refs. [9,10], respectively.

2. Alloy fabrication and characterization

A ternary Fe–15Ni–13Cr alloy was procured from Combustion Engineering, Chattanooga, Tennessee. The alloy was prepared by melting high purity electrolytic Fe, Cr, and Ni charges using argon blanked induction heating. Table 1 shows the chemical analysis of Heat No. 3756 ingot used in this study. Oxygen was a notable impurity at 0.031 wt%. From this master ternary ingot, 28 compositional variant alloys were fabricated by adding one or several alloying additions. Bars of 500 g each (13 mm × 25 mm × 130 mm) were prepared at ORNL using the non-consumable arc melting and drop casting technique. After removing hot tops, the ingots were subsequently rolled to 0.58 mm sheets having approximately 23% final cold work. A solution annealing temperature of 1200°C was applied twice during the intermediate rolling processes to assure a proper dissolution of precipitates, particularly MC particles.

During the preparation of charges, exact amounts of desired alloying elements were added to produce, in wt%, 15 Ni, 13.5 Cr, 2.2 Mn, 2.0 Mo, 0.22 Ti, 0.044 C, 0.85/0.35 Si, 0.05/0.01 P, 0.002 B, and 0.4/0.2 Nb in the final alloys. The final alloy compositions deviated somewhat from the intended compositions due to the loss of precipitates to the hot tops. The chemical compositions of the sheet alloys (not hot tops) are summarized in Tables 2 (Basic alloy series) and 3 (Engineering alloy series). In particular, oxygen analyses were carried out to confirm the exact residual oxygen contents in the final sheet alloys by using various analytical means such as vacuum fusion, calorimetric, gravimetric, and LECO T-36 (glow discharge gas analysis) methods. Of particular interest, for alloys with Ti, Si, and C additions, the residual oxygen contents in the sheets were substantially less than that in the master ternary stock. This was mainly because large fractions of light oxide and carbide particles floated up during melting and were subsequently removed with hot tops. Details will be discussed in later sections.

3. Precipitate analysis

Precipitate phases in the sheet alloys and hot tops were analyzed to understand the final alloying conditions. Precipitates were extracted by electrolytically dissolving the final sheet materials and slices from selected hot tops. Extraction was done by applying 50 mA current in 10% hydrochloric acid and 90% methyl alcohol electrolyte. Precipitate weight fractions were determined by weighing residues. The precipitate compositions were then analyzed using energy dispersive spectroscopy (EDS) techniques, and crystal structures were determined by X-ray diffraction powder method and transmission electron microscopy (TEM). When particle volume fractions were too small for X-ray analysis or particles were too thick for TEM, crystal structures were derived from the stoichiometric relationship obtained from EDS. A JEM-120CX electron microscope equipped with an energy dispersive X-ray detector was used for EDS and TEM. Details of experimental procedures were described elsewhere [11].

Fourteen precipitate phases were identified in as-fabricated alloys. Table 4 summarizes the types of precipitate phases and their weight fractions found in the sheet alloys. Table 5 summarizes crystallographic data and compositions of the precipitate phases found in all B- and E-series alloys. The precipitate weight fractions in hot tops were several-fold to an order of magnitude higher than those in the corresponding sheet alloys, suggesting that significant fractions of light oxide particles floated up during melting. Specific gravities of TiO, TiC, SiO (SiO_2 and Si mixture), CO are 4.93,

Table 1
Chemical analysis of ingot (wt%)[a]

Fe	Cr	Ni	Mo	Ti	Si	Mn	C	P	Nb	B	N	O
Bal	12.96	15.08	0.01	<0.01	0.04	<0.01	0.008	0.003	<0.01	<0.001	0.012	0.031

[a] Trace elements of S, V, Cu, Al, and Co also exist but are less than 0.01 wt%.

Table 2
Composition of B-series alloys (wt%)[a]

Alloys	Fe	Cr	Ni	Mo	Ti	Si	Mn	C	P	Nb	B	N	O
B1	Bal	13.79	14.95	–	–	–	–	–	–	–	–	–	0.030
B2	Bal	13.69	14.96	–	0.17	–	–	–	–	–	–	–	0.011
B3	Bal	13.66	15.02	–	–	0.83	–	–	–	–	–	–	0.019
B4	Bal	13.74	14.88	–	–	–	–	–	–	–	0.001	–	0.031
B5	Bal	13.68	14.94	–	0.17	–	–	0.040	–	–	–	–	0.004
B6	Bal	13.36	15.04	1.97	0.19	–	–	0.056	–	–	–	–	0.003
B7	Bal	13.55	15.01	–	0.16	0.83	–	0.041	–	–	–	–	0.004
B8	Bal	13.70	14.95	–	0.16	–	–	0.037	–	–	0.001	–	0.006
B9	Bal	13.71	15.23	–	–	–	–	–	0.049	–	–	–	0.031
B10	Bal	13.64	15.15	–	–	0.84	–	–	0.050	–	–	–	0.020
B11	Bal	13.63	15.20	–	0.18	–	–	0.041	0.049	–	–	–	0.013
B12	Bal	13.58	15.15	–	0.17	0.83	–	0.044	0.049	–	–	–	0.004
B13	Bal	13.42	14.88	–	–	–	–	0.025	–	0.40	–	–	0.005
B14	Bal	13.69	14.90	–	–	–	–	–	0.010	–	–	–	0.031

[a] Not added but showed residual impurities: <0.01Mo, <0.01Ti, <0.05Si, <0.01Mn, <0.008C, <0.004P, <0.01Nb, <0.001B, <0.015N.

Table 3
Composition of E-series alloys (wt%)[a]

Alloys	Fe	Cr	Ni	Mo	Ti	Si	Mn	C	P	Nb	B	N	O
E1	Bal	13.14	15.10	1.98	0.19	0.82	2.13	0.059	–	–	–	–	0.004
E2	Bal	13.41	14.97	–	0.19	0.80	2.16	0.055	–	–	–	–	0.003
E3	Bal	13.19	15.07	1.97	0.17	0.82	–	0.053	–	–	–	–	0.004
E4	Bal	13.21	15.11	1.98	–	0.83	2.13	0.040	–	–	–	–	0.006
E5	Bal	13.18	15.15	1.98	0.17	0.83	2.14	–	–	–	–	–	0.008
E6	Bal	13.24	15.07	1.97	0.17	0.37	2.14	0.047	–	–	–	–	0.003
E7	Bal	13.33	15.05	1.97	0.17	–	2.16	0.044	–	–	–	–	0.003
E8	Bal	13.15	15.12	1.98	0.17	0.84	2.14	0.054	0.044	–	–	–	0.004
E9	Bal	13.23	15.11	1.97	0.17	0.35	2.13	0.036	0.043	–	–	–	0.003
E10	Bal	13.30	15.08	1.98	0.18	–	2.13	0.046	0.046	–	–	–	0.004
E11	Bal	13.15	15.23	1.99	–	0.84	2.13	0.035	–	0.45	–	–	0.003
E12	Bal	13.19	15.12	1.95	–	0.35	2.12	0.023	–	0.42	–	–	0.005
E13	Bal	13.20	15.09	1.95	–	–	2.12	0.019	–	0.41	–	–	0.009
E14	Bal	13.20	15.09	1.97	0.09	–	2.15	0.044	–	0.23	–	–	0.003

[a] Not added but showed residual impurities: 0.01Mo, <0.01Ti, <0.01Si, 0.009C, <0.003P, <0.01Nb, <0.002B, <0.011N.

Table 4
Precipitate phases and their weight fractions in Fe–15Ni–13Cr base stainless steels

Alloys	Precipitate phases	Weight fraction (%)	Alloys	Precipitate phases	Weight fraction (%)
B1	Cr_2O_3	0.078	E1	TiO, TiC, MC, $Ti_{10}S_3$	0.187
B2	TiO, Ti_5S_4	0.062	E2	TiO, TiC,Ti_2S	0.225
B3	SiO[a]	0.013	E3	TiO, TiC,Ti_2S	0.167
B4	Cr_2O_3	0.062	E4	Cr_2O_3, G	0.026
B5	TiO, TiC	0.085	E5	TiO	0.107
B6	TiO, TiC	0.066	E6	TiO, TiC, $Ti_{10}S_3$	0.140
B7	TiO, TiC	0.119	E7	TiO, TiC,Ti_2S	0.114
B8	TiO, TiC	0.112	E8	TiO, TiC, MC, Ti_2S	0.032
B9	Cr_2O_3	0.092	E9	TiO, TiC	0.020
B10	SiO[a], Cr_2O_3, CrS	0.030	E10	TiO, TiC	0.082
B11	TiO, TiC, M_2P	0.107	E11	NbO, NbC, SiO[a]	0.166
B12	TiO, TiC, M_2(P, Si)	0.229	E12	NbO, NbC	0.092
B13	NbO, NbC, CrS	0.127	E13	NbO, NbC	0.085
B14	Cr_2O_3, CrS	0.072	E14	NbO, TiO, TiC	0.136

[a] SiO exists in the form of vitreous SiO_2 and Si mixture.

4.93, 2.13, 0.00125, respectively, smaller than that of Fe–Cr–Ni steels, $\simeq$8.9 g/cm^3. Analyses confirmed that precipitate phases in the hot tops were mostly oxide and carbide particles. Consequently, both the solute contents (Ti, C, Si) and the residual oxygen remaining in the sheet alloys were less than original additions.

3.1. Oxygen gettering by alloying elements

The master alloy had $\simeq$0.03 wt% residual oxygen. The residual oxygen content was reproduced in alloys B1, B4, B9, and B14, which had no Ti, C, and/or Si additions. On the other hand, alloys with Ti, C, and/or Si additions showed 30% (B3, B10) to almost 90% (B5, B6, B7, B12) reduction in residual oxygen. During melting 0.22 wt% Ti was added, but analyses showed that only about 0.17 wt% Ti was retained in the final sheet alloys. About 0.05 wt% titanium loss was accompanied by $\simeq$0.02 wt% oxygen loss in B2 alloy, indicating that titanium oxide was in the form of monoxide, TiO. The structure of TiO was confirmed by X-ray diffractometry of the extracted particles. About 0.02 wt% silicon loss was accompanied by 0.01 wt% oxygen reduction in B3 alloy, suggesting that the silicon oxide particles were in the oxygen poor monoxide form, SiO. As will be discussed later, SiO is believed to exist as a condensed mixture of vitreous SiO_2 and Si. In alloy B13, about 0.02 wt% carbon loss was accompanied by 0.023 wt% oxygen loss, suggesting that the oxygen and carbon were lost again as monoxide, CO.

About 30% oxygen reduction was observed in alloy B3, which has only Si added. About 60% oxygen reduction was observed in alloy B2 which has only Ti added. About 80% oxygen reduction was observed in alloy B13, which has C but has neither Ti nor Si. However, when both Ti and C were present together as in alloys B5, B6, B8, and B11, carbon appeared to have little effect on oxygen reduction, since oxygen reduction was accompanied by Ti reduction with little change in C concentration. Although the experimental evidence suggested that Ti was the strongest oxide former followed by C and Si, surprisingly the most efficient oxygen removal was achieved by carbon if Ti was not present. This was mainly because gaseous carbon monoxide, CO, was boiled off easily during melting. As will be discussed further, when Ti was present, TiO formation was preferred to CO, and oxygen removal efficiency was diminished because some fraction of TiO remained in the sheet alloy.

Substantial reduction of residual oxygen was also seen in alloys with Nb and C but no Ti (B13, E12, E13). The oxygen reduction was accompanied by a reduction of carbon but no appreciable niobium reduction was observed, although X-ray analyses revealed the presence of NbO and NbC particles. Specific gravities of NbO and NbC are 7.3 and 7.6 g/cm^3, respectively, close to that of the host ternary. It is thus believed that most oxygen loss occurred by CO emission and the heavy niobium oxide and carbide particles were retained in the sheet alloys.

All E-series alloys showed at least 70% reduction in residual oxygen in response to Ti, C, and/or Si additions. There was no apparent nitrogen effect on alloy compositions although there was about 0.01 wt% trace nitrogen in the master ternary. In most alloys, a few yellow diamond shaped nitride particles could be seen under an optical microscope. In the following, specific characteristics of individual precipitate phases are described.

Table 5
Characterization of thermal precipitate phases in Fe–15Ni–13Cr base austenitic stainless steels

Precipitate phases	Composition (at.%) EDS analysis	Crystal structure	Lattice parameter (nm) – X-ray	Lattice parameter (nm) – powder data	Relevant alloys (found in)
Cr_2O_3	100Cr	Hex., D_{51}, $R\bar{3}c$	$a_o = 0.4954$ $c_o = 1.3584$	$a_o = 0.4954$ $c_o = 1.3584$ card # 6-0604	Alloys with no Ti, C B1, B4, B9, B14, E4
SiO[a]	100Si	No data	Not measured	No lattice data card # 30-1127	Alloys with Si but no Ti, C B3, B10, E11
TiO	100Ti	Cubic, B1, $Fm\bar{3}m$	$a_o = 0.4251 - 0.4293$[b]	$a_o = 0.4293$ card #29-1361	Alloys with Ti
TiC	100Ti	Cubic, B1, Fm3m	$a_o = 0.4251 - 0.4293$[b]	$a_o = 0.4329$ card #6-0614	Alloys with Ti and C
TiN	Not measured	Cubic, B1, $Fm\bar{3}m$	Not measured	$a_o = 0.4246$	Alloys with Ti
MC	89Ti–11Mo 67Ti–5Cr–8Fe–20Mo 91Ti–1Cr–2Fe–6Mo	Cubic, B1, $Fm\bar{3}m$	$a_o \approx 0.4283$[b]	No data	Data from E1 and E8
NbO	100Nb	Cubic, B1, $Fm\bar{3}m$	$a_o \approx 0.4270$[b]	$a_o \approx 0.4210$	Alloys with Nb
NbC	100Nb	Cubic, B1, $Fm\bar{3}m$	$a_o = 0.4429 - 0.4461$	$a_o = 0.4470$ card #10-181	Alloys with Nb and C B13, E11, E12, E13, E14
$M_2(P,Si)$	30P–32Ti–23Cr–12Fe–3Ni 12P–20Si–27Ti–21Cr–9Fe–12Ni	Hex., C_{22}, P321	Not measured	$a_o = 0.5864$ $c_o = 0.3461$ card # 27-1171 (Fe_2P)	B11, B12
Ti_5S_4	45S–50Ti–5Cr	Not measured	Not measured	No lattice data card # 12-402	B2
$Ti_{10}S_3$	24S–62Ti–8Fe–6Mo 24S–72Ti–2Fe–2Mo	Not measured	Not measured	No data	E1, E6
Ti_2S	34Ti–66Ti	Hex.	Not measured	$a_o = 0.3206$ $c_o = 1.1190$ card # 11-664	E2, E3, E7, E8
CrS	45Cr–4Fe–4Nb–47S 47Cr–4Fe–49S	Not measured	Not measured	No data	B13, B14
G-phase	19Si–3Ti–6Cr–72Mn	Cubic, B1, $Fm\bar{3}m$	Not measured	No data	E4

[a] SiO exists in the form of vitreous SiO_2 and Si mixture.
[b] Average values due to an overlap of peaks with similar such lattice parameters from TiO, TiC and NbO.

3.2. Identification of precipitates

Chromium oxide (Cr_2O_3) precipitates were found as large spherical particles, often larger than 3 μm. Composition measured by EDS showed 100% chromium, excluding oxygen which, although present, could not be measured by EDS because of the limitations of the instrument used. X-ray diffractometry confirmed a hexagonal structure consistent with the ASTM Powder Data File of Cr_2O_3. The chromium oxide particles were found mainly in alloys B1, B4, B9, B10, and E4 which lack Ti and/or C, suggesting that Ti and C are stronger oxide former than Cr.

Silicon oxide (SiO, an intimate mixture of vitreous SiO_2 and Si) particles were present in almost perfect spherical shapes, a typical morphology of high temperature melting products. Silicon oxide particles were found in alloys with Si but no Ti or C. Particle sizes ranged from a few tenths of nanometer to a few micrometers. EDS analysis showed 100% silicon, again not counting oxygen due to the detection limitation. As mentioned already, however, the stoichiometry derived from oxygen and silicon losses indicated that the atomic ratio of Si and O in the particles is approximately equal to one ($Si/O \approx 1$). X-ray analysis did not show any discernible peaks from these particles, and thus the crystal structure of these particles could not be determined. Gibbs free energy of formation (Fig. 1, data from [12]) indicates that SiO may form initially as gas phase at high temperatures (>2000 K), but being unstable dissociates into a mixture of vitreous SiO_2 and Si as it condenses [13]. SiO is listed as a solid phase having a melting point greater than 1702°C in the CRC Handbook of Chemistry and Physics. However, the ASTM Powder Data (Card #30-1127) for SiO did not record its crystallographic information, perhaps due to the amorphous nature of the mixture.

Since both titanium monoxide (TiO) and titanium carbide (TiC) have the same $Fm\bar{3}m$ space group cubic crystal structure and very similar lattice parameters, it was difficult to distinguish them using X-rays alone. EDS also showed 100% Ti for both phases, excluding oxygen or carbon. These melt-products of titanium were generally much larger in size compared to MC particles that precipitate at lower temperatures, as described below. X-ray diffraction showed peaks of $Fm\bar{3}m$ space group crystal structure. However, the peak widths were broader than usual due to superposition of peaks from TiO and TiC phases. The average lattice parameters measured were in the range of 0.4251–0.4293 nm, close to or smaller than the lattice parameters recorded in the ASTM Powder data for both TiO (0.4293 nm) and TiC (0.4329 nm). The deviations suggest that perhaps TiO and TiC particles are non-stoichiometric.

Although it will be described further in Part II and III of this issue, here it should be pointed out that TiC, TiO, TiN, NbC, NbO, all share the same crystal structure, have similar lattice parameters, and have a capacity to

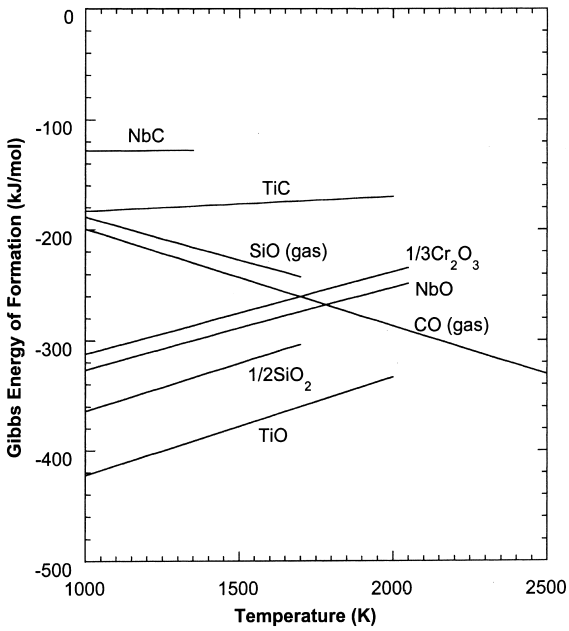


Fig. 1. Gibbs free energy of formation [12] as a function of temperature for various precipitate phases.

accommodate alloying elements such as Cr, Fe, and Mo. Therefore these phases generally occur in a complex form of (Ti, Nb, Mo, Fe, Cr) · (C, N) whenever these constituent elements are present. Such precipitate phases are often designated as MC. X-ray analyses indicated that two distinct peaks appeared when carbide and oxide phases were present together, hinting that oxygen partition in MC phase would be minimal if present, and oxide particles were likely to exist as a separate phase. In the alloys investigated in this work, most Nb containing alloys did not include Ti. In such cases, the precipitate phases are identified as NbC (instead of MC) or NbO although they may include other elements.

In the E-series alloys containing Mo, Ti, and C, EDS confirmed that there were also various Mo-rich titanium carbides (MC) in addition to the pure Ti particles (TiC, TiO) mentioned above. The latter particles were generally smaller in size, suggesting that these particles formed at lower temperatures during the cooling process. X-ray diffraction confirmed $Fm\bar{3}m$ space group crystal structure for MC particles extracted from E1 alloy. The diffraction peaks were split due to peak overlaps, suggesting that MC, TiO, and TiC phases were all present together. The MC, TiO, and TiC phases have the same crystal structure and very similar lattice parameters. The average lattice parameter of these particles was about 0.4283 nm.

All niobium containing alloys showed very irregularly shaped Nb-rich particles. Particle sizes varied from a few tenths of nanometers to several micrometers. EDS analysis showed the composition of these particles was mostly Nb, occasionally with small quantities of Fe and/or Cr. X-ray diffraction showed double peaks due to the overlap of NbC and NbO peaks, but the NbC peak intensities were an order of magnitude larger than those of NbO. Although X-ray powder data did not reveal the structure type of NbO, the parallel nature of the NbC and NbO peaks suggests that NbO has the same $Fm\bar{3}m$ space group cubic crystal structure as NbC. Although NbO phase is known to exist in two types of space group, $Fm\bar{3}m$ and $Pm\bar{3}m$, the lattice parameter measured by X-ray showed a most close match with that of $Fm\bar{3}m$ in the X-ray powder data cards.

Phosphorous-rich phases were found in all phosphorous containing alloys. Compositions of the particles varied considerably but showed approximately an M_2P stoichiometric relation. In establishing the stoichiometric relationship, the elements Ti, Cr, Fe, and Ni were considered to be in the M-sites and Si in the P-sites, after the classical work by Rundqvist [14]. The crystal structure of phosphide could not be determined because precipitate volume fractions were too small for X-ray diffraction. However, as will be shown in Part III (this issue), when the same alloys were irradiated by ions or neutrons, extensive phosphide formation occurred. In such cases, phosphide precipitates could be extracted, and the *d*-spacings matching with the hexagonal Fe_2P powder data were confirmed by TEM diffraction ring analysis.

Sulfide particles were found in various compositional and stoichiometric forms. The major constituents of sulfides were either Ti or Cr with small quantities of Fe, Mo, and Nb. Since sulfur was a trace impurity, sulfide particle number densities were extremely low. Surprisingly, among several hundreds of particles analyzed by EDS, MnS particles could not be spotted, although the formation of MnS is expected in view of thermodynamics. Since most sulfides were found to be in the form of Ti_xS_y, MnS formation appeared to be suppressed in the presence of titanium. Approximate stoichiometric relations derived from EDS analyses are summarized in Table 5.

A few unidentifiable precipitate phases were also present. One of them was a Mn-rich silicide phase having a typical composition of 19Si–3Ti–6Cr–72Mn in atomic percent. This phase was tentatively identified as G-phase, considering that the manganese fraction is similar to that in Th_6Mn_{23} type G-phase with $Fm\bar{3}m$ cubic crystal structure. Yellowish diamond or square shaped particles were observed under optical microscopy examination, particularly for Ti-containing alloys. These particles are believed to be titanium nitride based on their color and morphology. No X-ray confirmation was made because of their small volume fractions.

4. Thermodynamic considerations

Chemical analyses revealed that there was a clear correlation between alloying elements and the residual oxygen contents in the final sheet alloys. The evidence indicated that the oxide forming efficacy decreased in the order of titanium, carbon, silicon, and chromium. However, carbon acted as the most efficient oxygen remover if Ti was not present. This is because the gaseous CO boiled off easily from the system. Most oxide particles were found to be in oxygen poor monoxide forms such as TiO, CO, and SiO (a mixture of amorphous SiO_2 and Si) instead of TiO_2, CO_2, and crystalline SiO_2, perhaps due to the low oxygen levels (<0.03 wt%) in the alloys. The results, however, showed that the oxide forming trends were consistent with Gibbs free energies of formation. Empirical relations of the standard Gibbs free energies of formation (ΔG_f° in J/mol) at a given oxygen level are shown below for TiO, CO (gas), SiO (gas), SiO_2, NbO, Cr_2O_3, TiC, and NbC phases [12].

$Ti + 1/2O_2 \rightarrow TiO$	$\Delta G_f^\circ = 511\,703 + 89.12T$	600–2000 K
$C + 1/2O_2 \rightarrow CO$	$\Delta G_f^\circ = 111\,713 - 87.65T$	298–2500 K
$Si + O_2 \rightarrow SiO_2$	$\Delta G_f^\circ = 902\,070 + 173.64T$	700–1700 K
$SiO_2 + Si \rightarrow 2SiO$	$\Delta G_f^\circ = 709\,606 + 53.97T \log T - 520.07T$	298–1700 K
$Si + 1/2O_2 \rightarrow SiO$	$\Delta G_f^\circ = 96\,232 + 26.99T \log T - 173.22T$	700–1700 K
$2/3Cr + 1/2O_2 \rightarrow 1/3Cr_2O_3$	$\Delta G_f^\circ = 386\,589 + 74.27T$	298–2100 K
$Nb + 1/2O_2 \rightarrow NbO$	$\Delta G_f^\circ = -415\,68024.06T \log T + 161.08T$	298 K
$Ti + C \rightarrow TiC$	$\Delta G_f^\circ = 182\,862 + 10.01T$	298–1155 K
	$= 186\,606 + 13.22T$	1155–2000 K
$Nb + C \rightarrow NbC$	$\Delta G_f^\circ = 130\,122 + 1.67T$	1180–1370 K

Gibbs free energies were plotted as a function of temperature for the above equations in Fig. 1. In agreement with the experimental observations, TiO has the lowest Gibbs free energy of formation within the temperature range of interest, suggesting that TiO would be the most stable phase among them. Although carbon alone acted as the most efficient oxygen removing agent, its oxygen removing capacity was diminished in the presence of titanium because TiO formation became preferred over CO, and TiO particles could not be removed as easily as CO. Formation of chromium oxide was suppressed in alloys with Ti, C , and Si as expected, particularly at high temperatures.

When both carbon and oxide forming elements were present (i.e., NbC vs. NbO or TiC vs. TiO), carbide particles appeared to be more abundant than oxide particles despite the higher ΔG_f° for carbide phase. In general, carbide formation is very fast. In our customary thermal aging process, it was found that carbide formation occurred in less than a minute at 600°C. These fast kinetics might have been responsible for the preponderant carbide. Besides, in the presence of carbon, oxygen concentration was lowered by the loss of CO during alloy melting, which made the oxide formation unfavorable. In fact, the oxygen concentration in B13 alloy was reduced to 0.005 wt% from 0.03 wt% of B1 master ternary. This could have made NbO formation unfavorable because of the depletion of the needed constituent solute.

5. Conclusions

Twenty-eight Fe–15Ni–13Cr base alloys were prepared by adding various alloying elements. Fourteen different types of precipitate phases were identified. Up to four different precipitate phases were found in any one alloy, depending upon the alloy composition. The predominant precipitates were oxides, carbides, and minute amounts of phosphides, and various types of trace sulfides. Chemical analyses revealed that the final alloy compositions were noticeably affected by the formation of oxides and carbides, namely the loss of residual oxygen, Ti, C, and Si. Most losses occurred by emission of gaseous CO and by floatation of the light oxide, carbide, and silicide particles to the hot tops during melting. Although titanium was a stronger oxide former than carbon, carbon alone removed oxygen more efficiently than Ti alone, because gaseous CO molecules could be removed easily from the system. When Ti was present, TiO formation was preferred over CO. In such cases, less residual oxygen was removed because a significant fraction of solid TiO particles was retained in the system. For a similar reason, niobium alone made very little change in residual oxygen because heavy NbO particles were mostly retained in alloys. Silicon also was less efficient in removing oxygen than carbon, since a significant fraction of SiO was retained as a mixture of amorphous SiO_2 and Si.

Gibbs free energies for the oxide formation were in the order of $\Delta G_f^\circ(TiO) < \Delta G_f^\circ(CO) < \Delta G_f^\circ(SiO,\ gas)$, consistent with the experimentally observed stability of these phases. However, the stability of NbC vs. NbO did not follow the trend expected from Gibbs free energy considerations. However, if oxygen levels remaining in alloys and the aspect of formation kinetics are considered, the experimental findings could be justified.

As mentioned already, it is well known that oxygen promotes cavity formation under irradiation [5,6] and that fine carbide and phosphide particles suppress cavity growth by trapping and dispersing gas atoms at their particle–matrix interfaces [7–9,15–18]. It is thus very important to minimize the impurity oxygen level and at the same time to maintain the precipitate forming elements such as Ti, C, Si, and P in solution during alloy fabrication. In this work, however, it was found that these solute elements were reduced or deactivated when oxides formed. It is important to recognize that, if alloying elements are in oxide forms, they cannot be active in carbide or silicide forming processes during irradiation. In such cases, alloys with seemingly adequate bulk compositions may not have effective matrix concentrations of desirable solutes and may not be suitable for applications to critical components of fission and fusion reactors. Fortunately however, carbon was found to act as an extremely efficient oxygen removing agent if Ti was not present. Therefore, in alloy fabrication, it is recommended to remove impurity oxygen first by melting the Fe–Cr–Ni ternary with carbon only, determine the remaining carbon level, supplement the missing portion of carbon, add desired alloying

elements, and finally fabricate alloys with conventional melting practices. Finally it should be pointed out that, as shown in Table 3, all E-series alloys with Ti or C addition have less than 0.006 wt% oxygen. This seemingly low oxygen concentration occurred, however, at the expense of Ti and/or C losses. Better alloys could have been made if oxygen were removed by carbon before adding other alloying elements.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The authors wish to thank Drs P.J. Maziasz and S. Babu for reviewing the manuscript. Also special thanks are due to the editor, Dr H. Kleykamp, for the critical comments on the nature of silicon oxide, which was invaluable in improving the manuscript.

References

[1] F.B. Pickering, Physical Metallurgy and the Design of Steels, Applied Science, Barking, 1978.
[2] J.F. Bates, W.G. Johnston, International Conference on Radiation Effects in Breeder Reactor Structural Materials, Scottsdale, AZ, June 1977, p. 625.
[3] P.J. Maziasz, in: B.L. Bramfit, R.L. Benn, C.R. Brinkman, G.F. Vander Voort (Eds.), ASTM STP 979, American Society for Testing and Materials, Philadelphia, 1988, p. 116.
[4] E.H. Lee, L.K. Mansur, Philos. Mag. A 61 (1990) 751.
[5] S.J. Zinkle, E.H. Lee, Metall. Trans. A 21 (1990) 1037.
[6] E.H. Lee, L.K. Mansur, Metall. Trans. A 23 (1992) 1977.
[7] K. Ehrlich, K. Anderko, J. Nucl. Mater. 124 (1984) 229.
[8] K. Herschbach, W. Schneider, K. Ehrlich, J. Nucl. Mater. (1993) 233.
[9] E.H. Lee, L.K. Mansur, Philos. Mag. A 61 (1990) 733.
[10] E.H. Lee, N.H. Packan, in: N.H. Packan, R.E. Stoller, A.S. Kumar (Eds.), 14th International symposium on Effects of Radiation on Materials, vol. 1, ASTM STP 1046, American Society for Testing and Materials, Philadelphia, 1989, p. 133.
[11] E.H. Lee, P.J. Maziasz, R.F. Rowcliffe, in: J.R. Holland, L.K. Mansur, D.I. Potter (Eds.), Phase Stability during Irradiation, AIME, 1981, p. 191.
[12] O. Kubaschewski, C.B. Alcock, Metallurgical Thermochemistry, 5th Ed., Pergamon, New York, 1979.
[13] JANAF Thermochemical Tables, 2nd Ed., Office of Standard Reference Data, National Bureau of Standards, Washington, DC, 1971.
[14] S. Rundquist, Binary Transition Metal Phosphides, Arkiv För Kemi 20 (1962) 67.
[15] P.J. Maziasz, Scripta Metall. 14 (1980) 1251.
[16] A.F. Rowcliffe, E.H. Lee, J. Nucl. Mater. 108&109 (1982) 306.
[17] W. Kesternich, J. Nucl. Mater. 127 (1983) 153.
[18] E.H. Lee, L.K. Mansur, J. Nucl. Mater. 141–143 (1986) 695.

ELSEVIER

Journal of Nuclear Materials 278 (2000) 11–19

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Fe–15Ni–13Cr austenitic stainless steels for fission and fusion reactor applications. II. Effects of minor elements on precipitate phase stability during thermal aging

E.H. Lee*, L.K. Mansur

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 3 May 1999; accepted 13 September 1999

Abstract

The precipitate phase stability in Fe–15Ni–13Cr base austenitic alloys was investigated as a function of minor alloying additions after thermally aging at 600°C and 675°C for times ranging from 24 h to one year. Seven major precipitate phases were found in aged specimens, including $M_{23}C_6$, Laves, Eta (η), TiO, NbC, MC, and M_2P. The types and amounts of precipitate phases varied with alloying element additions, aging temperature, and aging time. By analyzing the composition of each individual particle, it was possible to determine the essential constituent elements for each phase. From this information, a strategy to promote or suppress certain precipitate phases was developed. Among the seven phases, the most desirable precipitate phases were considered to be MC and M_2P, because these particles form on a fine scale with a high number density and, therefore, can serve as effective gas atom trap sites under irradiation. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 64.75.+g; 81.30.Mh; 81.40.Cd

1. Introduction

Fe–Cr–Ni base austenitic alloys, such as AISI 316 and 304 stainless steels, are used in various industrial applications because they combine outstanding mechanical formability with good higher temperature strength, and oxidation and corrosion resistance. However, it has been found that when these alloys are used under neutron irradiation environment as in nuclear reactor, their serviceable lifetime is greatly compromised due to properties' degradation. The major cause of degradation was attributable to a long term exposure to high energy neutrons and temperature. In the past few decades, considerable progresses have been made in understanding fundamental mechanisms of radiation induced damage and in developing radiation resistant alloys for reactor applications. The most important agenda was to mitigate the adverse effects caused by transmutation induced helium. Theoretical and experimental work has shown that the most damaging effects at elevated temperature come from agglomeration of helium atoms and vacancies leading to cavity swelling and to grain boundary cavitation and embrittlement [1,2]. Fortunately, however, it was found that such detrimental effects could be suppressed by dispersing helium atoms on precipitate–matrix interfaces. This required high number densities of helium traps, which could be provided by number densities of precipitates. MC carbide and M_2P phosphide particles were found to be the most effective precipitate phases, because they form in fine scale and in high number density during thermal aging as well as during irradiation [3–5]. However, inducement of such high number density particles requires a fine tuning of alloy composition, which in turn requires understanding of the synergistic effects of alloying elements, thermal aging, and irradiation on phase stability. In Part I (this issue), we present the results of minor alloying element effects on precipitate phases in the melt products and their implications in

* Corresponding author. Tel.: +1-423 574 5058; fax: +1-423 574 0641.

E-mail address: leeEH@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(99)00235-4

alloy fabrication. In this paper, we present data on the effects of minor alloying elements on phase stability during thermal aging. In Part III (this issue), we present radiation effects on phase stability under heavy ion irradiation. For neutron irradiation effects, some of the Fast Flux Test Facility (FFTF) reactor data were reported previously [6].

In the past, numerous investigators studied precipitate phases in austenitic steels. Weiss and Stickler documented precipitate phases found in AISI 316 stainless steel [7]. Precipitation reactions in iron-base alloys were reviewed by Decker and Floreen [8] and by Edmond and Honeycombe [9]. However, the information in those reviews was gathered from various sources and the investigated alloys covered a wide range of base alloy compositions; thus it was difficult to differentiate the effects of minor alloying elements from the effects of base alloy compositions. Moreover, at that time, it was difficult to study the early stages of phase formation because precipitate volume fractions were often too small to be analyzed by conventional bulk analysis methods. However, recent advancements in analytical electron microscopy have made it possible to analyze the micro-chemistry of individual particles and have allowed us to better understand micro-alloying phenomena. In particular, for specimens aged for very short periods of time, the identification and characterization of precipitate phases owe much to carbon-replica extraction and

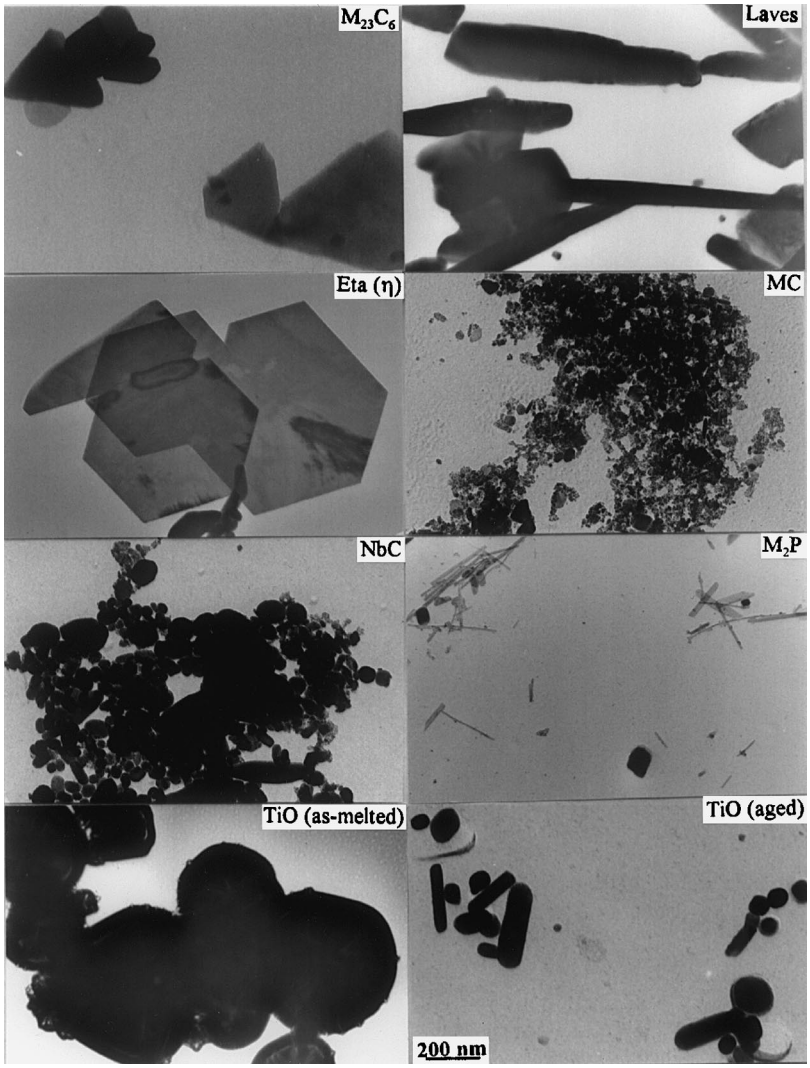


Fig. 1. Typical morphologies of precipitate phases: $M_{23}C_6$ in B12 (500 h/600°C); Laves in E1 (six months/675°C); η in E4 (six months/675°C); MC in B8 (500 h/675°C); NbC in B13 (500 h/675°C); M_2P in E9 (six months/675°C); TiO in B2 (as-melted), and TiO in B2 (500 h/600°C).

energy dispersive X-ray spectroscopy techniques. In this work, alloys with various alloying element additions were thermally aged at 600°C and 675°C for various periods of time, and the precipitate phases in the aged specimens were investigated by using transmission electron microscopy (TEM) and energy dispersive X-ray spectroscopy (EDS) methods.

2. Experimental

The same 28 Fe–15Ni–13Cr based alloys, whose fabrication was described in Part I, were used to investigate thermal aging effects. As previously described, the alloys were prepared by adding one to several alloying elements to the Fe–15Ni–13Cr ternary alloy. Typical concentrations of alloying elements were, in weight percent, 2.2 Mn, 2.0 Mo, 0.22 Ti, 0.044 C, 0.85/0.35 Si, 0.05/0.01 P, 0.002 B, and 0.4/0.2 Nb when they were added. Complete alloy compositions can be found in Part I. Here, all alloy compositions are expressed in weight percent (wt%), and precipitate compositions are in atomic percent (at.%). Carbon and oxygen were not included in precipitate compositions due to the limitation of EDS in detecting low mass elements with the instrument used.

Specimens were prepared in the form of TEM disks from 23% cold-worked sheet alloys. Disks were then encapsulated in evacuated and helium back-filled (25 kPa) quartz capsules, and aged at 600°C and 675°C for the periods of 24 h, 500 h, six months, and one year. Crystal structure and microchemical analyses were

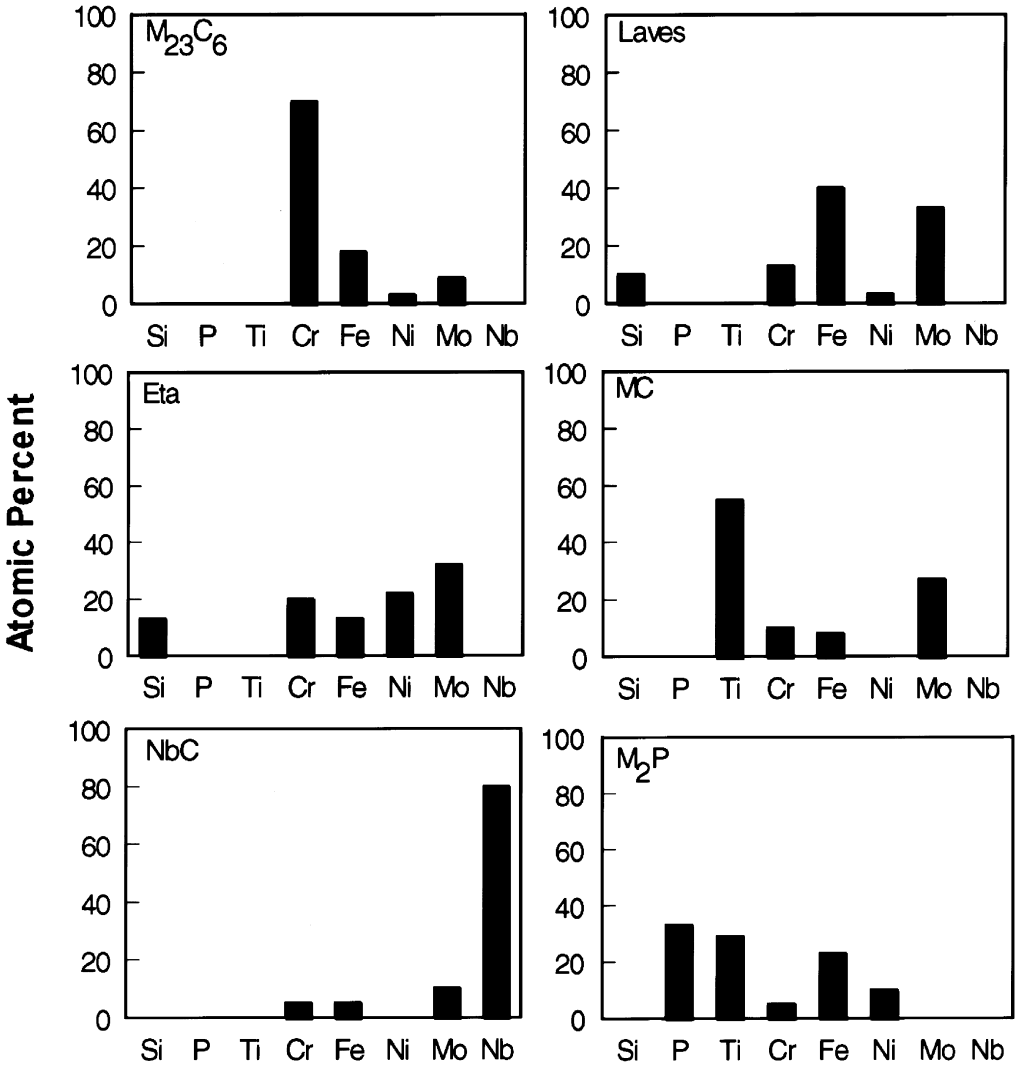


Fig. 2. Typical EDS spectra for the precipitate phases found in Fe–23Cr–15Ni base austenitic steels aged at 600–675°C.

carried out by TEM and EDS for precipitate particles extracted on carbon replicas and also for particles in the thinned TEM foils, following the routine procedures established in previous experiments [10]. Detailed TEM electron diffraction analyses of precipitate phases were also reported in the previous paper and thus are not reiterated here. In general, the morphology and composition of each precipitate phase were so distinct that an experienced microscopist could identify them by their shapes or EDS spectra, even before TEM diffraction analyses. As a guide, typical morphologies of precipitate phases are displayed in Fig. 1 and typical compositional finger prints determined by EDS are illustrated in Fig. 2 . For analyses, a JEM-120CX analytical transmission electron microscope equipped with a beryllium window EDS detector was used. In most cases, it was not difficult to distinguish the melt-product precipitate phases from thermal-aging precipitate phases because their particle sizes and morphologies were noticeably different from each other.

3. Results

In Part I, we reported 14 precipitate phases which occurred during the alloy melting processes. After thermal aging at 600°C and 675°C, seven major precipitate phases were developed, which included $M_{23}C_6$, Laves, Eta (η), TiO, MC, NbC, and phosphides. Among the seven, the first three were not observed in the melt products, whereas the latter four were observed to a lesser degree after melting but developed significantly more in number but smaller in size during thermal aging. The precipitates in the melt-product were clearly recognized by their large and globular morphology. The precipitates produced by thermal aging were somewhat different in composition (accommodating a wide range of impurity elements), morphology (generally smaller in size and having distinct shapes), abundance (higher in number density), and distribution (more uniform in size and space), even though they had the same crystal structure as the corresponding melt product phase. Table 1 summarizes the crystallographic information of these precipitate phases. The types of precipitate phases and amounts found varied from alloy to alloy because of the difference in alloy compositions, as summarized in Table 2. Critical changes in alloying additions are indicated next to the alloy symbols, to better illustrate the relationship between alloy composition and the precipitate phases found. When Laves, $M_{23}C_6$, and η were present, they were large sized particles that comprised most of the precipitate volume fraction. By comparison, MC, M_2P, and NbC precipitates were finer, higher number density phases but comprised only a minor volume fraction of the total precipitation.

All seven precipitate phases could be identified after only 24 h of aging, although their volume fractions were extremely small, particularly for the Laves and phosphide phases. Precipitate volume fractions increased with increasing aging time and temperature. Most precipitate phases became quite copious after 500 h of aging. For a given alloy, the precipitate compositions were rather insensitive to aging time and temperature, but same precipitate phase in two different alloys showed somewhat different compositions. In the following, specific characteristics of each precipitate phase are described.

3.1. $M_{23}C_6$

After thermal aging, profuse $M_{23}C_6$ particles were formed in all carbon containing alloys, mostly at grain boundaries and along deformation bands. A few $M_{23}C_6$ particles were also found even in alloys with no carbon addition because of the presence of trace impurity carbon (<0.01 wt%). The $M_{23}C_6$ phase was characterized by a strong Cr-peak in EDS. Although $M_{23}C_6$ particles accommodated a range of compositions, their constituent elements were strictly limited to Cr, Fe, Ni, and Mo. There were two types of $M_{23}C_6$ depending upon molybdenum addition, namely $(Cr, Fe, Ni)_{23}C_6$ and $(Cr, Fe, Ni, Mo)_{23}C_6$ as summarized in Table 3. When Mo

Table 1
Crystallographic information of precipitate phases

Phases	Crystal structure/space group	Lattice parameter (nm)	Orientation to matrix	Typical morphology
$M_{23}C_6$	Cubic, $D8_4$, $Fm\bar{3}m$	$a_0 = 1.06$	Cube-on-cube or twin	Rhombohedral platelet
Laves	Hex., C14, $P6_3/mmc$	$a_0 = 0.41$	Various orientations	Faulted lath
		$c_0 = 0.77$		
η	Cubic, $E9_3$, $Fd\bar{3}m$	$a_0 = 1.08$	Cube-on-cube	Hexagonal
TiO	Cubic, B1, $Fm\bar{3}m$	$a_0 = 0.43$	Cube-on-cube	Short rod
MC	Cubic, B1, $Fm\bar{3}m$	$a_0 = 0.432$	Cube-on-cube	Small sphere
NbC	Cubic, B1, $Fm\bar{3}m$	$a_0 = 0.447$	Cube-on-cube	Small sphere
M_2P	Hex., C_{22}, P321	$a_0 = 0.604$	$(1\bar{2}10)//(011)$	Thin needle lath
		$c_0 = 0.36$	$(0001)//(001)$	

Table 2
Precipitate phases formed during thermal aging at 600°C and 675°C

Alloys (additions)[a]	Precipitate phases	Alloys (additions)[a]	Precipitate phases
B1 (ternary)	No new phase	E1 (all additions)	Laves, MC, $M_{23}C_6$, η
B2 (Ti)	TiO	E2 (−Mo)	MC, $M_{23}C_6$
B3 (Si)	No new phase	E3 (−Mn)	Laves, MC, $M_{23}C_6$, η
B4 (B)	No new phase	E4 (−Ti)	Laves, $M_{23}C_6$, η
B5 (Ti, C)	MC, $M_{23}C_6$	E5 (−C)	Laves, TiO
B6 (Ti, C, Mo)	MC, $M_{23}C_6$, Laves	E6 (0.3Si)	Laves, MC, $M_{23}C_6$
B7 (Ti, C, Si)	MC, $M_{23}C_6$	E7 (−Si)	Laves, MC, $M_{23}C_6$
B8 (Ti, C, B)	MC, $M_{23}C_6$	E8 (0.8Si, P)	M_2P, Laves, MC, $M_{23}C_6$, η
B9 (P)	No new phase	E9 (0.3Si, P)	M_2P, Laves, MC, $M_{23}C_6$, η
B10 (P, Si)	M_2P	E10 (−Si, P)	M_2P, Laves, MC, $M_{23}C_6$
B11 (P, Ti, C)	M_2P, MC, $M_{23}C_6$	E11 (0.8Si,−Ti, 0.4Nb)	NbC, Laves, $M_{23}C_6$, η
B12 (P, Ti, C, Si)	M_2P, MC, $M_{23}C_6$	E12 (0.3Si,−Ti, 0.4Nb)	NbC, Laves, $M_{23}C_6$
B13 (Nb, C)	NbC, $M_{23}C_6$	E13 (−Si,−Ti, 0.4 Nb)	NbC, Laves, $M_{23}C_6$
B14 (0.01P)	No new phase	E14 (−Si, 0.1Ti, 0.2 Nb)	NbC, Laves, MC, $M_{23}C_6$

[a] Nominal composition of base ternary B1 alloy is Fe–13Cr–15Ni and that of engineering E1 alloy is Fe–13Cr–15Ni–2Mo–2Mn–0.2Ti–0.04C–0.8Si. Added or omitted elements are indicated in the parentheses. Phosphorous additions are about 0.05 wt% unless specified. A small amount of boron (<0.002 wt%) was added in alloys B4 and B8.

Table 3
Composition ranges (at.%) for the precipitate phases formed during thermal aging at 600–675°C[a]

Phases	Si	P	Ti	Cr	Fe	Ni	Mo	Nb	Alloying additions
$M_{23}C_6$	–	–	–	73–82	15–22	2–6	–	–	−Mo
	–	–	–	54–79	12–22	2–11	6–16	–	+Mo
Laves	5–12	–	–	9–19	33–44	3–11	32–37	–	0.8Si
	4–9	–	–	10–14	36–34	6–7	27–39	–	0.3Si
	–	–	–	10–21	37–48	3–8	29–37	–	0.0Si
	0–9	–	–	9–12	36–47	4–9	26–36	2–10	+Nb
η	5–9	–	0–2	19–43	9–16	11–26	23–36	–	+Si + Mo, 600°C
	9–17	–	0–6	12–28	12–17	16–23	32–41	–	+Si + Mo, 675°C
TiO	–	–	100	–	–	–	–	–	−C(B2, E5)
MC	–	–	64–94	0–36	0–2	–	–	–	−Mo
	–	–	29–77	3–20	0–18	–	12–43	–	+Mo
NbC	–	–	–	0–7	0–10	–	9–17	65–83	−Ti
	–	–	7–32	2–16	2–8	–	1–9	62–71	+Ti
	–	–	–	–	–	–	–	100	+C + Nb (B13)
M_2P	–	30–37	17–37	0–14	20–32	0–12	(9–11)[b]	–	+P + Si / Ti, 675°C
M_3P	–	22–26	23–50	5–13	14–29	0–16	–	–	+P + Si / Ti, 675°C
$M_{5-6}P$	–	13–18	0–15	22–34	36–43	10–12	–	–	+P + Si / Ti, 600°C

[a] Carbon and oxygen are not counted in composition because of EDS limitation.
[b] For E-series alloys with Mo.

was present, Cr content in the phase decreased. $M_{23}C_6$ phase accommodated the constituent elements in varying ratios and showed no apparent compositional variation with aging time and temperature. When Nb and/or Ti was present, $M_{23}C_6$ was noticeably reduced because carbon was depleted by NbC and/or MC carbide formation. Phosphorous addition promoted $M_{23}C_6$ formation (B10, B11, B12, E9, E10) or η (E8), because titanium was consumed by Ti-rich phosphide, reducing MC formation, and thus resulting in excess carbon available for $M_{23}C_6$ or η. The enhancement of $M_{23}C_6$ phase was also observed in C + P and C + N + P austenitic steels by Irvine et al. [11], Dulis [12] and Saatinen [13]. They attributed the $M_{23}C_6$ enhancement to the phosphorous entering into the precipitating carbides as $M_{23}(C, P)_6$ [11] or $(Cr, Fe, P)_{23}C_6$ [12,13]. These conclusions were made based on measurements of lattice parameters and *average* compositions of extracted

particles, not on an analysis of individual particles. However, EDS analysis of individual particles clearly revealed that there was no phosphorous in $M_{23}C_6$ phase, but rather M_2P and $M_{23}C_6$ particles coexisted as two distinct phases. Since $M_{23}C_6$ particles generally constituted most of the precipitate volume fraction because of large sizes, averaging particle composition with larger electron probes in previous work might have sampled mostly $M_{23}C_6$, but clearly that is not the case using smaller electron probes with higher spatial resolution in this work.

3.2. Laves

Laves phase particles were found in all alloys with Mo additions. Although Laves formation was sluggish compared to other phases, a few Laves particles were observed after 24 h of aging. The Laves nucleation sites were primarily associated with defects such as dislocations, grain boundaries, twin boundaries, and deformation bands. Laves particles could be identified easily by their blocky or elongated lath morphology with internal faulted defect structure. Analyses revealed that Laves phase had about 33 at.% Mo, approximately in M_2Mo stoichiometric ratio. The two third M sites were occupied mainly by Fe with varying amounts of Cr, Ni, and Si. Although Si was not an essential ingredient for Laves phase, the Si fraction in Laves and the Laves volume fraction increased with increasing Si concentration in alloys. Among 45 spectra recorded, only four showed a small titanium peak, hinting that Ti might not be a Laves ingredient and that the erroneous Ti-peak might have originated from small MC particles attached to Laves during extraction. When Nb was present, Laves formation was enhanced. Analyses of particle compositions revealed $(Mo + Nb) \approx 33$ at.% suggesting that Nb occupied Mo-sites. Laves was suppressed by MC, because MC phase accommodated Mo as high as 43 at.%. Under conditions more favorable for MC (i.e., high Ti supersaturation and high dislocation density), Laves was further suppressed, suggesting kinetic competition between the two phases. MC nucleates in less than a minute at 600°C.

3.3. Eta (η)

η phase particles were found only in alloys having C, Si, and Mo together. Crystal structure of η phase was confirmed to be $Fd\bar{3}m$ diamond cubic by analyzing electron diffraction patterns obtained from the extracted particles. η particles appeared in thin diamond or hexagonal platelet morphology and their detection in TEM was extremely difficult because the thin platelet could not be imaged easily. When either Si or C level was lowered, η was suppressed. Similar to the $M_{23}C_6$ case, phosphorous promoted η because phosphide accommodated Ti, discouraging MC formation and thus leaving more C in solution for η phase. Different from $M_{23}C_6$ and Laves phases, η composition varied with temperature; the (Si, Mo)/Cr ratio increased when aging temperature increased from 600°C to 675°C. Two types of diamond cubic η phase were reported [14], filled type M_6C and unfilled type M_5SiX, with Si being on a metal site and X being a vacant site. Complex carbides derived from the parent lattice Ti_2Ni (η) are also called eta carbides which include $M_3M'_3C$ (η_1), $M_2M'_4C$ (η_2), and $M_6M'_6C$ (η_3) types having lattice parameter in the range of $a_0 = 1.1$–1.2 nm [15]. However, the present work showed that both C and Si were essential elements for η phase. Thus η phase is believed to be a carbosilicide in the form of $(M, Si)_6C_x$, with $x < 1$.

3.4. MC

MC denotes the fcc TiC phase which accommodates other minor metallic elements. Since there exists impurity nitrogen in all alloys, nitrogen is known to be present in the carbide as M(C, N). However, N was not detected in MC phase by electron energy loss spectroscopy (EELS) [4]. As pointed out already, EDS did not include C, O, and N analyses, and thus here we will simply designate the carbide as MC. There were large primary MC particles in all Ti and C containing alloys prior to thermal aging. A very high number density of fine secondary MC particles (20–30 nm) evolved at dislocations during thermal aging. Electron diffraction analyses revealed that the small size MC particles had the characteristic cube-on-cube relation with the austenitic matrix. The major constituents of MC phase were Ti, Mo, and Cr with sporadic inclusions of Fe and Ni. Similar type MC phase was also reported previously [16–18]. Although MC formed without Mo (B5 alloy), MC phase had a strong affinity for Mo and accommodated Mo as high as 43 at.%, such that when MC became dominant, Mo was depleted from solution and Laves was suppressed.

3.5. NbC

When Nb and C were present, NbC phase similar to MC was developed at dislocations during thermal aging. Here again NbC may be in the form of Nb(C, N), but nitrogen will be ignored in the formula. NbC particles were somewhat coarser and lower in number density compared with MC particles. Differing from the NbC phase (100% Nb in EDS excluding C) found in the melt products, thermal-aging-produced NbC particles (E11, E12, E13) accommodated a wide range of metallic elements, namely in decreasing order, Nb, Mo, Cr, Fe, and infrequently Ni and Mn. Similar NbC particles were also observed in niobium modified austenitic stainless steel previously [17]. When Ti was present (E14), (Nb, Ti)C

phase, with Ti as high as 32 at.%, was developed. When Ti was present, Mo in the phase was reduced as Ti substituted on Nb sites. Although both NbC and MC accommodated Mo, Mo solubility in NbC was less than in MC. On the other hand, Nb was soluble in Laves but Ti was not. For these reasons, Laves was promoted when Nb was present. The Ti/Nb ratios in the (Ti, Nb)C carbide particles were in the range 0.1–0.5. The low Ti level in (Nb, Ti)C phase was believed to be due to low Ti (0.1 wt%) in E14 alloy. Interestingly, NbC particles, in B13 alloy which had only Nb and C additions, did not accommodate any other metallic elements. EDS results showed 100% Nb, again not counting C, O, and/or N. Clearly, further studies are need to understand the synergistic effects of Mo and other alloying elements on the phase stability.

3.6. Phosphides

Small globular phosphide particles were found in phosphorous containing alloys prior to thermal aging. After thermal aging, needle-shaped phosphide precipitates were developed at dislocations. The phosphide particles grew with *c*-axis in parallel to the three $\langle 1\,0\,0 \rangle$ orthogonal directions of the matrix, the direction with the least lattice mismatch. The phosphide structure was confirmed to be hexagonal by analyzing the electron diffraction rings obtained from extracted particles. Phosphide formation was extremely sluggish, almost absent in specimens aged for 24 h, but a scant number of phosphide precipitates was found after 500 h of aging at 600°C and 675°C. Phosphide compositions varied with aging temperature. At 675°C, most phosphide particles had an M_2P stoichiometric relation, where M represents a mixture of Ti, Fe, Cr, Ni, and Mo in decreasing order when present. A few M_3P particles were also found. At 600°C, most phosphide particles had initially (500 h) lower phosphorous contents showing $M_{5-6}P$ stoichiometric relations, but the phosphorous fraction in the phase increased with aging time, resulting in the M_2P phase after one year of aging. Although phosphide formation was encouraged with increasing Si, interestingly the thermal-aging-produced phosphides did not include Si; phosphides in neutron irradiated stainless steels included up to 16 at.% Si [10]. Titanium being one of the major constituent elements of phosphide, when phosphide became dominant, MC was suppressed as a result of depleted Ti, whereas $M_{23}C_6$ was encouraged due to conserved C by not forming MC. An excellent review on phosphide phases in binary transition metals was made by Rundquist [19]. M_2P type phosphide was not reported by Rundquist [19], but its presence is indicated in the phase diagram shown in the Binary Alloy Phase Diagrams by Massalski et al. [20]. Profuse needle shaped hexagonal phosphide precipitates were observed in AISI 321 stainless steel held in service at 600°C for 17 yr by Bentley and Leitnaker [21]. The current data, however, suggest that such precipitates might have formed early in the service period. Rowcliffe and Nicholson found Cr_3P phase in Fe–18Cr–10Ni steel and reported the crystal structure being tetragonal [22]. Although the crystal structure of M_3P phase could not be determined in this work, we found that phosphide composition changed to become hexagonal M_2P with increasing aging time.

3.7. TiO

Large TiO particles existed in all Ti-containing alloys prior to thermal aging. Nevertheless, additional smaller size TiO particles were formed during thermal aging, but *only* in alloys B2 and E5, which lacked carbon. This finding suggests that some oxygen might have been in solution instead of in oxide form. When sufficient carbon was present, profuse TiC particles formed instead. Although EDS analyses showed 100% Ti (not counting O, C, and/or N) for both melt product TiO and thermal-aging-produced TiO particles, the latter had various distinctive morphologies such as rod, diamond, sphere etc., and could easily be distinguished from the conspicuously large spherical shape melt product TiO particles as shown at the bottom of Fig. 1. Although TiO, TiN, and TiC all have the same $Fm\bar{3}m$ space group cubic crystal structure and similar lattice parameters, these distinct particles thought to be TiO, since the particles appeared only in B2 and E5 alloys with no carbon addition.

4. Discussion

Seven major precipitate phases were identified in Fe–13Cr–15Ni based alloys after thermal aging at 600°C and 675°C. Among the seven, four were carbides, $M_{23}C_6$, η, MC, and NbC, and the other three were Laves, TiO, and phosphides. Laves phase was the unique intermetallic phase found after thermal aging. Although Sigma (σ) and Chi (χ) phases were often observed in AISI 316 type stainless steels [10], they were not observed in these alloys. Williams investigated precipitate phases in thermally aged austenitic steels [23]. He also found neither σ nor χ below 700°C, but observed σ phase in specimens aged between 700°C and 750°C.

Experimental data showed that virtually all alloying elements were found to participate in the precipitate formation as shown in Table 3. One notable observation was that there were no ideal stoichiometric compounds. The X-ray microchemical analyses revealed that the precipitate phases accommodated a wide variety of atomic species in varying proportions, although each phase had characteristic fingerprints in morphology and

composition as shown in Figs. 1 and 2, respectively. The alloying additions did not produce any new phases that have not been observed in AISI 316 type stainless steels. However, types of precipitate phases and their abundance varied considerably depending upon alloying additions. In this experiment, substantial information was gathered to clarify the roles of each alloying element on phase stability. Table 4 summarizes all experimental findings.

Mo was a promoting or stabilizing agent for $M_{23}C_6$, Laves, and η phase formation, Nb for NbC, Ti for MC, and P for M_2P. Among the seven phases, $M_{23}C_6$, Laves, and η are considered to be undesirable because they occur in large size and in low number density. In such a case, both gas atoms and vacancies are shared by a small number of precipitates, which results in earlier formation of critical cavity size. Moreover, vacancies are collected more efficiently at large particle interfaces (i.e., point-defect collector effect) [24–26], thereby augmenting void swelling. Besides, Laves phase is considered to have little benefit because large particles contribute very little in strengthening mechanical properties.

TiO formation was triggered mainly by the presence of impurity oxygen and would produce no significant impact in suppressing swelling. Since both NbC and MC phases had a similar crystal structure, (Nb, Ti)C type carbides were formed when both Nb and Ti were present. Gas atom trapping efficiency of NbC phase is smaller than that of MC, because NbC particles are generally larger in size but lower in number density compared with those of MC. Therefore, NbC is considered to be a less desirable phase than MC.

Although Si was not an essential element for M_2P phase, it promoted M_2P formation when E8, E9, and E10 alloys were compared (Table 2). Undesirably, Si also promoted Laves and η. Clearly, an optimal level of Si should be determined to minimize formation of such phases.

Among the seven phases, M_2P and MC were considered to be the most desirable phases because they nucleate in fine scale and in high number density. Fortunately, in view of designing radiation resistant alloys, the occurrence of M_2P phase was not affected by other phases because the major constituents of M_2P phase were Fe and P and no other phases required phosphorous as their constituent. Unfortunately, however, MC was somewhat reduced as a result of dominant Laves phase formation in most alloys investigated in this work, particularly in alloys with phosphorous. Nevertheless, as pointed out in Part I, a considerable fraction of the added Ti was depleted from the matrix by TiO formation with impurity oxygen during alloy melting. Therefore, the suppressed MC could be attributable partly to the lower effective Ti and C concentrations in the alloys investigated here. Phosphorous also tended to promote Laves, because Ti-rich M_2P depleted Ti, discouraging Mo-rich MC formation. This, however, raises a question whether a microstructure with suppressed Laves phase could be achieved. Incidently, such example was found in Ti-modified D9 alloy irradiated with neutron at 593°C in the Experimental Breeder Reactor (EBR-II). Laves was suppressed drastically, to the benefit of MC, when Ti supersaturation was increased by dissolving preexisting coarse MC particles back into solution by employing higher solution annealing temperature, 1200°C instead of 1100°C [26].

Mo is known as a solution hardening element and also is important for recrystallization and creep resis-

Table 4
Role of minor alloying elements on precipitate phases during thermal aging

Precipitate phase	Promoted by	Remarks	Suppressed by	Remarks
$M_{23}C_6$	C P (indirect)	M_2P depletes Ti and avails C for $M_{23}C_6$	Ti, Nb	MC depletes C Essential for $M_{23}C_6$
Laves	Mo, Nb, Si	Constituent elements	C, Ti	MC depletes Mo Essential for Laves
η (M_6C)	Mo, Si, C P (indirect)	M_2P depletes Ti and avails C for M_6C	Ti, Nb	MC depletes C Essential for M_6C
TiO	Ti	Residual oxygen No significance	C	C promotes MC
MC[a]	Ti, C, (Mo)	Constituent elements	Nb	Nb promotes Laves and depletes Mo needed for MC
NbC	Nb, C	Constituent elements	Ti	Ti competes with Nb
M_2P[b]	P, Fe	Constituent elements	No apparent suppressor	But require enough P to produce high number density M_2P

[a] MC particles nucleate in high number density very rapidly (<1 min) at dislocations and exhaust nucleation sites, and thus the formation of sluggish large precipitate phases such as Laves, $M_{23}C_6$, and η is suppressed when MC becomes dominant.
[b] M_2P phase accommodates Ti but not C, so if there is not enough Ti, MC phase may be suppressed whereas $M_{23}C_6$ is promoted. An optimum balance of Ti, C, and P is required to achieve a high number density of both MC and M_2P particles.

tance [27]. Although Mo is a common ingredient for Laves, $M_{23}C_6$, and η, careful examination of experimental data suggests that Mo reduction may not be necessary if MC and M_2P were maximized by optimizing P, Ti, and C concentrations. It was reported that MC and M_2P improved hardness, yield strength, tensile strength, and creep properties due to dislocation pinning by high number density small particles [12,23,28]. Enhancement of MC and M_2P by optimizing alloy composition is desirable.

5. Conclusions

Effects of alloying elements on precipitate phase stability were investigated for Fe–13Cr–15Ni based alloys aged at 600°C and 675°C for periods from 24 h to one year. Seven major precipitate phases, $M_{23}C_6$, Laves, Eta (η), TiO, NbC, MC, and M_2P, were evolved during thermal aging. It was found that the stability of precipitate phases depend strongly upon alloy composition and thermo-mechanical treatment (i.e., high temperature solution annealing and cold working). Therefore, alloys with desired precipitate phases could be tailored by optimizing alloying element additions–enhancement of MC and/or M_2P phases by controlling Ti, P, and C levels for example. This, however, requires a detailed knowledge of how minor alloying elements affect the phase stability and what is the impact of thermal aging on precipitate phases. This work establishes a basis to answer these questions and provides guidance in alloy design. The most important and pertinent experimental findings are summarized in Table 4.

Finally, MC and M_2P strengthened alloys are known to have also excellent thermo-mechanical properties [28], and the efforts to develop radiation resistant alloys benefit not only reactor materials applications but also other critical materials applications, such as high temperature service.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The authors wish to thank Drs P.J. Maziasz and S. Babu for reviewing the manuscript.

References

[1] L.K. Mansur, in: Gordon R. Freeman (Ed.), Kinetics of Nonhomogeneous Processes, ch. 8, Wiley, New York, 1987, p. 377.
[2] E.H. Lee, L.K. Mansur, Metall. Trans. 21A (1990) 1021.
[3] W. Kesternich, J. Nucl. Mater. 127 (1983) 153.
[4] P.J. Maziasz, R.L. Klueh, R.F. Rowcliffe, MRS Bull. 14 (1989) 36.
[5] E.H. Lee, L.K. Mansur, Philos. Mag. A 61 (1990) 733.
[6] E.H. Lee, N.H. Packan, in: N.H. Packan, R.E. Stoller, A.S. Kumar (Eds.), Effects of Radiation on Materials: 14th International symposium, ASTM STP 1046, vol. 1, American Society for Testing and Materials, Philadelphia, 1989, p. 133.
[7] B. Weiss, R. Stickler, Metall. Trans. 3 (1972) 181.
[8] R.F. Decker, S. Floreen, in: K.C. Russell, H.I. Aaronson (Eds.), Precipitation Process in Solids, TMS-AIME, 1976, p. 63.
[9] D.V. Edmonds, R.W.K. Honeycombe, in: K.C. Russell, H.I. Aaronson (Eds.), Precipitation Process in Solids, TMS-AIME, 1976, p. 121.
[10] E.H. Lee, P.J. Maziasz, R.F. Rowcliffe, in: J.R. Holland, L.K. Mansur, D.I. Potter (Eds.), Phase Stability During Irradiation, AIME, 1981, p. 191.
[11] J.K. Irvine, D.T. Llewellyn, F.B. Pickering, J. Iron Steel Inst. 199 (1961) 153.
[12] E.J. Dulis, Iron and Steel Inst. Special Report No. 86 (1964) 162.
[13] A.V.A. Saatinen, Scand. J. Met. 2 (1973) 161.
[14] H.H. Stadelmaier, Metal Rich Metal–Metalloid Phases, Developments in the Structural Chemistry of Alloy Phases, Plenum, New York, 1969.
[15] H. Nowotony, P. Ettmayer, J. Inst. Met. 97 (1969) 180.
[16] E.H. Lee, N.H. Packan, L.K. Mansur, J. Nucl. Mater. 117 (1983) 123.
[17] P.J. Maziasz, in: J.R. Holland, L.K. Mansur, D.I. Potter (Eds.), Phase Stability During Irradiation, AIME, 1981, p. 477.
[18] H.O. Andren, A. Henjered, H. Norden, J. Mater. Sci. 15 (1980) 2365.
[19] Stig Rundquist, Ark. Kemi 20 (1962) 67.
[20] T.B. Massalski et al. (Eds.), Binary Alloy Phase Diagrams, American Society for Metals, Metals Park, OH, 1986.
[21] J. Bentley, J.M. Leitnaker, in: E.W. Collings, H.W. King (Eds.), The Metal Science of Stainless Steels, Met. Soc. AIME, 1978, p. 70.
[22] A.F. Rowcliffe, R.B. Nicholson, Acta Metall. 20 (1972) 143.
[23] T.M. Williams, in: H.R. Brager, J.S. Perrin (Eds.), Effects of Radiation on Materials, ASTM STP 782, 1982, p. 166.
[24] L.K. Mansur, Philos. Mag. 44 (1981) 867.
[25] E.H. Lee, A.F. Rowcliffe, L.K. Mansur, J. Nucl. Mater. 103, 104 (1981) 1475.
[26] E.H. Lee, L.K. Mansur, Metall. Trans. 23A (1992) 1977.
[27] R.W.K. Honeycombe, Steels-Microstructure and Properties, American Society for Metals, Edward Arnold, Great Britain, 1981.
[28] P.J. Maziasz, R.W. Swindeman, The International Conference on Advances in Material Technology for Fossil Power Plants, ASM International, Chicago Il., 1–3 September 1987, 8706-018.

ELSEVIER

Journal of Nuclear Materials 278 (2000) 20–29

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

Fe–15Ni–13Cr austenitic stainless steels for fission and fusion reactor applications. III. Phase stability during heavy ion irradiation

E.H. Lee *, L.K. Mansur

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Received 3 May 1999; accepted 13 September 1999

Abstract

The phase stability in Fe–15Ni–13Cr alloys was investigated as a function of minor alloying additions after 4 MeV Ni ion irradiation at 948 K. The results showed that the stability of precipitate phases was dictated mainly by the defects produced by radiation damage and preferential segregation of Si and Ni at defects. In addition, radiation enhanced diffusion and cascade induced dissolution and mixing allowed kinetically sluggish phases to form rapidly under irradiation. These radiation effects caused an enhancement, retardation, or modification of thermal phases, and formation of new phases. The relative stability of precipitate phases varied sensitively with alloy composition. The roles of each alloying element on phase stability and the impact of radiation on the mechanisms of phase evolution were systematically studied and documented. The knowledge obtained from this work provides guidelines for designing alloys that lead to develop desired precipitate microstructures under irradiation. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 64.75.+g; 81.30.Mh; 61.80.-x

1. Introduction

The understanding of phase stability under irradiation is very important to both fission and fusion reactor materials design, because segregation and precipitation lead directly to changes in swelling and mechanical properties. Moreover, the phase stability under irradiation is considerably different from thermal behavior due to radiation induced processes such as enhanced diffusion [1], cascade dissolution [2–4], cascade mixing [5,6], and segregation of various alloying elements [7–10]. Although quite extensive data were documented for the precipitate phases developed under thermal [11] and irradiation [12,13] environments, the data were gathered from various alloy sources. A systematic study with a common base alloy to unveil the effects of minor alloying elements on phase stability is still lacking.

During irradiation, the ability of alloys to sustain the necessary dimensional and mechanical properties to meet design lifetimes and performance is largely dependent upon the manner in which displaced atoms and transmutation helium are accommodated in the microstructure. It was demonstrated that precipitation behavior during irradiation is of major significance because of the relationship between precipitation and swelling behavior [14,15]. It is also realized that precipitation cannot be avoided in multicomponent alloys of technological interest, and that satisfactory alloy performance can be attained only through proper tailoring of precipitate phases. In this work, therefore, an investigation was carried out to understand the phase evolution under irradiation in the Fe–15Ni–13Cr alloy system by systematically varying minor alloying elements.

2. Experimental

The compositions and thermal properties of 28 alloys investigated in this work were described in Parts I and

* Corresponding author. Tel. : +1-423 574 5058; fax: +1-423 574 0641.

E-mail address: leeEH@ornl.gov (E.H. Lee).

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S 0 0 2 2 - 3 1 1 5 (9 9) 0 0 2 3 6 - 6

II. Detailed alloy fabrication procedures were described in Part I, and precipitate phases produced under thermal aging were summarized in Part II. To investigate the precipitate phases evolving during irradiation, specimen disks (3 mm diameter by 0.35 mm thick) were prepared by solution annealing at1373 K. The annealed specimens were bombarded with 4 MeV Ni ions to irradiation doses of 1 and 70 dpa at 948K (675°C) using the triple ion facility (TIF) at ORNL. Some of the specimens were bombarded simultaneously with He and Ni ions. The irradiated specimen disks were then thinned to expose the peak damage region, 0.6–0.8 μm below the surface, for TEM examination. A JEM-120CX analytical electron microscope equipped with an energy dispersive spectrometer was used for analysis. After in-foil analyses of damage microstructure, diffraction patterns, precipitate orientations, and particle compositions for the various precipitate phases were obtained using the carbon replica extraction technique. The compositions and diffraction patterns of extracted particles were recorded, analyzed, and compared with the in-foil data. The general procedures have been described in a previous paper [12].

3. Results

Ten major precipitate phases were found in irradiated specimens. Of these, $M_{23}C_6$, η-phase, TiO, MC, NbC, and M_2P were also found in thermally aged samples, but their relative volume fraction and composition were in most cases altered by radiation. Four new phases, G, γ′, Ni-rich-γ, and CrSi, were also identified in irradiated specimens. Table 1 summarizes the precipitate phases found in each alloy. The qualitative abundances of precipitate phases are indicated by inequality signs. Typical compositions of precipitate phases are listed in Table 2. The details are described in the following.

3.1. $M_{23}C_6$

$M_{23}C_6$ phase was present in all carbon-containing alloys after irradiation, but generally there was less compared with specimens aged for 24 h at 948 K, except for four Ti and C containing alloys, B5(Ti,C), B6(Ti,C,Mo), B7(Ti,C,Si), and B8(Ti,C,B). In these alloys, there was more $M_{23}C_6$ and TiO formation. TiC was reduced, with the largest effect being seen in alloy B6. The apparent source of the oxygen for the formation of TiO was the dissociation of prior Cr_2O_3 oxide particles during irradiation. During thermal aging, both TiO and $M_{23}C_6$ were suppressed in these same alloys by the formation of TiC.

Fig. 1 compares precipitate phases extracted from thermally aged and irradiated B6 alloy specimens: MC dominates the thermally aged microstructure, and lath-shaped $M_{23}C_6$ and cuboidal TiO particles dominate the irradiated microstructure. EDS analyses showed that $M_{23}C_6$ compositions were virtually unchanged by radiation compared with those formed thermally. Mo-free or Mo-containing $M_{23}C_6$ precipitates were found depending upon the presence of molybdenum in the alloy. $M_{23}C_6$ phase was detected only in trace amounts after irradiation in alloys containing niobium, consistent with niobium carbide being much more stable than $M_{23}C_6$, even under irradiation conditions. Although $M_{23}C_6$ formation increased in alloys with phosphorous during thermal aging due to the formation of Ti-rich

Table 1
Precipitate phases in specimens irradiated with Ni ions to 70 dpa at 948 K

Alloys (additions)[a]	Precipitate phases	Alloys (additions)	Precipitate phases
B1(ternary)	No new phase	E1 (all additions)	MC > G, η
B2 (Ti)	TiO	E2 (−Mo)[b]	MC > G
B3 (Si)	Ni-rich $\gamma \gg$ CrSi	E3 (−Mn)	MC > G, η
B4 (B)	No new phase	E4 (−Ti)	No new phase
B5 (Ti, C)	TiO ≫ MC ≫ $M_{23}C_6$	E5 (−C)	TiO, γ′
B6 (Ti, C, Mo)	TiO ≫ $M_{23}C_6$ ≫ MC	E6 (0.3Si)	MC
B7 (Ti, C, Si)	TiO ≫ MC ≫ $M_{23}C_6$	E7 (−Si)	C
B8 (Ti, C, B)	TiO ≫ MC ≫ $M_{23}C_6$	E8 (0.8Si, P)	M_2P ≫ G, MC
B9 (P)	M_2P	E9 (0.3Si, P)	M_2P ≫ MC
B10 (P, Si)	M_2P, Ni-rich γ	E10 (−Si, P)	M_2P ≫ MC
B11 (P, Ti, C)	M_2P ≫ MC ≫ $M_{23}C_6$	E11 (0.8Si, −Ti, 0.4Nb)	G ≫ NbC
B12 (P, Ti, C, Si)	M_2P ≫ MC > G > $M_{23}C_6$	E12 (0.3Si, −Ti, 0.4Nb)	NbC > G
B13 (Nb, C)	NbC	E13 (−Si, −Ti, 0.4 Nb)	NbC
B14 (0.01P)	No new phase	E14 (−Si, 0.1Ti, 0.2 Nb)	NbC, MC

[a] Nominal composition of base ternary B1 alloy is Fe–15Ni–13Cr and that of engineering E1 alloy is Fe–13Cr–15Ni–2Mo–2Mn–0.2Ti–0.04C–0.8Si. Added or omitted elements are indicated in the parentheses. Phosphorous additions are about 0.05 wt% unless specified. Small amounts of boron (< 0.002 wt%) were added in alloys B4 and B8.
[b] "–" signs next to the elements indicate that the alloying elements are subtracted from the master E1 alloy.

Table 2
Typical compositions of the precipitate phases formed during irradiation (at.%)[a]

Phases	Si	P	S	Ti	Cr	Mn	Fe	Ni	Mo	Nb	Reference alloys
$M_{23}C_6$	–	–	–	–	63	–	25	4	8	–	
	–	–	–	–	69	–	26	5	–	–	B12
η	12	–	–	–	33	–	19	28	8	–	E1
	11	–	–	–	33	–	19	14	23	–	E3
TiO	–	–	–	100	–	–	–	–	–	–	B2, B5
	–	–	–	93	2	–	–	–	5	–	B6
	–	–	–	84	6	–	10	–	–	–	B7
MC	–	–	–	55	39	–	6	–	–	–	B5
	–	–	–	27	21	–	15	–	37	–	B6
	[b]	–	–	59	26	–	15	–	–	–	B7
NbC	–	–	–	–	39	–	8	7	–	46	B13
	10	–	–	–	19	–	5	8	8	50	E12
M_2P	–	34	[b]	–	32	–	19	15	–	–	B9
	8	24	5	–	26	–	13	24	–	–	B10
	[b]	32	–	17	22	–	10	19	–	–	B11
	11	24	5	8	14	–	8	28	–	–	B12
	10	22	–	8	14	–	11	27	8	–	E8
	1	29	–	13	20	–	11	19	7	–	E10
G	22	2	1	4	11	–	6	54	–	–	B12
	38	–	–	4	7	6	6	34	5	–	E1
	24	–	–	4	7	7	6	47	5	–	E8
	28	–	–	–	6	7	6	44	4	5	E11
	20	–	–	–	6	7	6	50	3	8	E12
Ni-rich γ	4	–	–	–	9	–	44	43	–	–	B3
	10	–	–	–	8	–	40	42	–	–	B3
	4	–	–	–	9	–	52	35	–	–	B10
	9	–	–	–	13	–	51	27	–	–	B10
CrSi	42	–	–	–	43	–	14	1	–	–	B3

[a] Carbon and oxygen are not included in the carbide and oxide particle compositions.
[b] Trace or occasional inclusion.

FeTiP-type phosphides, which enhanced C in solution, such an increase did not occur during irradiation. Instead, profuse phosphide particles with modified compositions (high Cr but reduced Ti) developed during irradiation.

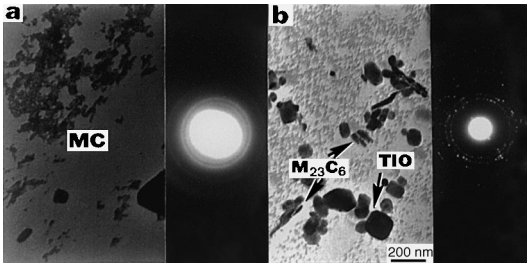


Fig. 1. (a) Extracted MC particles from the B6 alloy aged for six months at 948 K and the corresponding diffraction pattern, and (b) extracted TiO and $M_{23}C_6$ particles from the B6 alloy irradiated to 70 dpa at 948 K and the corresponding diffraction pattern.

3.2. Eta phase

The occurrence of η was suppressed by ion irradiation, and only a few particles were identified in alloys E1 and E3. During thermal aging, η phase readily developed in the alloy lacking titanium (E4) or in phosphorous containing alloys (E8, E9), but such enhancement was not found under ion irradiation. Profuse η precipitates were observed in a neutron irradiated (low damage rate) AISI 316 stainless steel [15,16] and in a pulsed ion beam irradiated Ti-modified austenitic stainless steel, similar to E1 alloy [17]. It may be that the suppression of η phase under ion irradiation is due to the high damage rate and shift of radiation induced segregation (RIS) relative to thermal diffusion effects rather than

insufficient thermal exposure during the irradiation period. In fact, the damage rate dependent stability of η phase was confirmed by the dissolution of the neutron irradiation induced η particles during subsequent Ni ion irradiation [18]. The composition of η phase was not altered by irradiation.

3.3. TiO

During irradiation, copious TiO particles developed in alloys B2 and E5, and to a much greater extent than during thermal aging. Although the former two alloys had no added carbon, surprisingly, TiO formation was enhanced even in alloys with carbon (B5, B6, B7, and B8). In the latter four alloys, small MC (or TiC) particles were also present but their volume fractions were noticeably reduced compared with thermally aged control specimens, indicating that there was a clear reversal of relative stabilities of the TiC to TiO phases under ion-irradiation. This phenomenon was not observed in the complex E-series alloys. When phosphorous was present, however, a Ti-rich phosphide was formed instead and TiO was suppressed.

Fig. 2 shows large cuboidal or rectangular TiO particles extracted from alloys B2 and E5 (no C addition), and B5 and B7 (0.04 wt% C). TiO particle sizes were smaller in alloys with carbon indicating that the growth of TiO particles was somewhat restrained by TiC formation. The electron diffraction patterns from these particles confirmed an fcc crystal structure with a lattice parameter $a = 0.42$ nm. The irradiation-produced TiO particles had somewhat more Cr, Fe, and Mo than thermally-produced TiO. As can be seen in Fig. 2(c), some TiO particles revealed an unusual cavity-like feature at the center of the particles. A similar hallow shape morphology was also observed in NbC particles in a thermally aged Fe–10Cr–6Mo–0.5Nb–0.05C ferritic steel [19]. The nature of these hallow regions has not been determined.

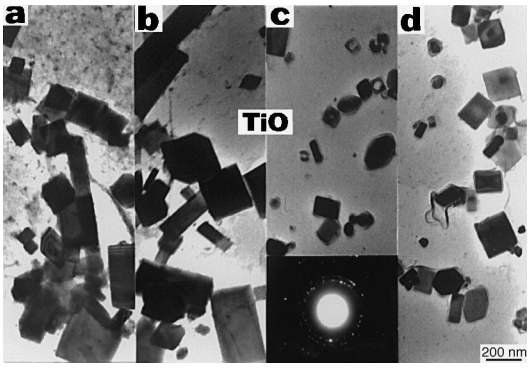


Fig. 2. TiO particles extracted from (a) B2, (b) E5, (c) B5, and (d) B7 alloys 4 MeV Ni-ion irradiated to 70 dpa at 948 K. The former two alloys have no carbon addition and the latter two have 0.04 wt% carbon. The insert is a diffraction ring pattern from the particles in alloy B5.

3.4. MC

Here, MC designates the TiC phase with various transition elements and does not include the Nb-rich carbide, NbC. A very high number density of fine MC particles (<10 nm) were nucleated rapidly during irradiation in all alloys that contained titanium and carbon. The rapid MC formation was due to profuse nucleation sites produced by radiation damage and to radiation-enhanced diffusion. Such a magnitude of MC formation did not occur in solution annealed alloys during thermal aging. The radiation-produced MC particles had a somewhat higher fraction of Cr, Fe, and Mo, but were otherwise virtually the same as the thermally formed MC. As mentioned above, when TiO formation (in B5, B6, B7, B8) became active under irradiation, MC was suppressed; this is an opposite trend to what was observed during thermal aging. In the presence of phosphorous, Ti-rich phosphide became dominant during irradiation and thus MC was reduced.

3.5. NbC

In alloys containing Nb and C, NbC particles precipitated copiously on the radiation-induced dislocations. NbC particles were somewhat coarser (<20 nm) than the MC formed in Ti-containing alloys. The nature of this NbC phase was different from thermally produced NbC in two aspects, namely precipitate size and composition. As the silicon concentration increased, the NbC particle size was refined and the volume fraction was also reduced. This was mainly because Si favored the formation of a Si- and Nb-rich G-phase, and thus niobium was depleted from solution. The radiation-produced NbC particles accommodated more foreign elements such as Cr, Si, and Ni compared to the NbC produced in the same alloy during thermal aging; Si and Ni were not observed in the thermally-produced NbC phase.

3.6. M_2P

In phosphorous-containing alloys, needle-shaped phosphide particles nucleated very rapidly on the loops and dislocations produced during irradiation, although phosphides were the slowest nucleating phase under thermal aging. At 1 dpa, phosphide precipitates were already dominant; such a scale of phosphide formation has not been observed under thermal aging conditions. MC was noticeably suppressed when the phosphide became dominant, partly due to the loss of Ti into the phosphide. Fig. 3 shows the phosphide microstructure at

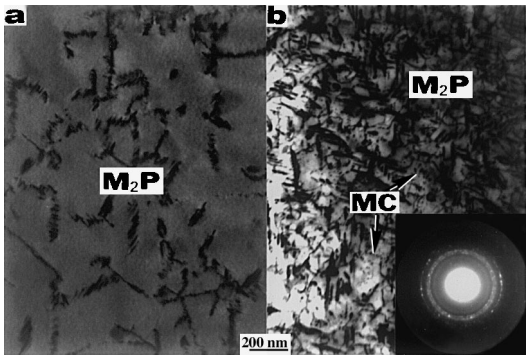


Fig. 3. M_2P precipitates in B12 alloy after (a) 1 and (b) 70 dpa irradiation at 948 K. The insert is a diffraction ring pattern obtained from the extracted particles.

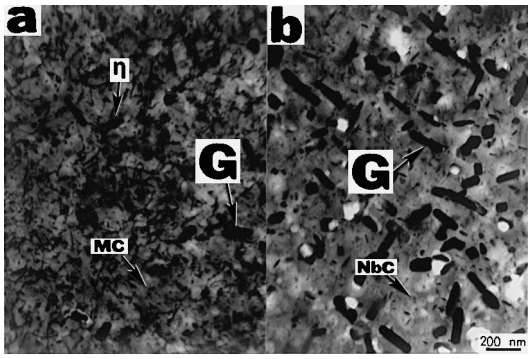


Fig. 4. G-phase particles in (a) E1 (Ti-modified) and (b) E11 (Nb-modified) alloys irradiated to 70 dpa at 948 K.

1 and 70 dpa. Phosphide particle sizes tended to coarsen in the absence of Ti (B9, B10), suggesting that an optimization of the Ti concentration is needed for the refinement of particle sizes and to achieve high precipitate number densities.

EDS analyses revealed that the phosphide compositions had higher Si and Ni concentrations than their thermal counterparts. Electron diffraction patterns from the extracted phosphide particles revealed an Fe_2P-type hexagonal structure, the same as for the thermal phase. Two typical phosphide compositions were found: a Cr-rich phosphide in alloys with no Ti addition (B9, B10) and a Ti-rich phosphide in titanium containing alloys (B11, B12, E8, E9, E10). Phosphide compositions showed an approximate $M_2(P,Si)$ stoichiometric relation where M represents metallic elements, Ti, Cr, Fe, Ni, and Mo. The assignment of Si to a P site was consistent with the report by Rundqvist [20] that Si substitutes on a P site. The phosphide phase accommodated a wide range of elements including S.

3.7. G-phase

G-phase developed in most of the E-series alloys with high Si, although it did not occur during thermal aging. G-phase compositions showed high Ni and Si with varying amounts of Ti, Cr, Mn, Fe, Mo, and Nb. The G-phase particles were generally nucleated near the center of the damage-induced interstitial loops, suggesting that segregation of Si and Ni at defect sinks led to the G-phase formation. EDS analysis confirmed an enrichment of nickel and silicon at such loops [21]. The G-phase was almost completely suppressed when the silicon level was reduced to below 0.3 wt%.

The G-phase was a unique phase which accommodated virtually all alloying elements, and yet not all of these elements were required for G-phase formation. A systematic elimination of alloying elements one at a time from the E1 alloy revealed that there was no particular single element essential for G-phase formation as long as there was enough Si and either Ti or Nb present. Of the latter two, niobium was the stronger G-phase promoter. As shown in Fig. 4, the distribution of G-phase is more extensive in E11 (0.4 wt% Nb, equivalent to 0.2 wt% of Ti in at.%) than in E1 (0.2 wt% Ti), although both alloys have the same base composition except for Ti and Nb. Niobium-rich G-phase was also found in a neutron-irradiated niobium-stabilized austenitic steel by Williams and Titchmarsh [16].

An atom probe field-ion microscopy study by Miller et al. [22] showed that (1 ± 0.7) atomic percent carbon was present in the G-phase particles in a CF 8 stainless steel. On the other hand, Williams and Titchmarsh [23] reported that carbon was unfavorable for G-phase formation because it promoted carbides. In this work, G-phase did not occur in alloy E5, which had all of the alloying elements except carbon, suggesting that the presence of carbon may be necessary for G-phase formation even though it is a silicide. It is, however, not clear whether carbon deficiency or other chemical effects prevented G-phase formation in E5 alloy, since Si and Ti concentrations were also reduced in this alloy due to the formation of γ' (Ni_3Si) and TiO.

3.8. Gamma prime

Gamma prime (γ') particles were observed only in alloy E5, which lacked carbon. Fig. 5 shows bright and dark field micrographs of γ' particles. The electron diffraction pattern showed that γ' particles had a cube on cube relation with the matrix and had an ordered $L1_2$ fcc structure. Unfortunately, the precise composition of γ' particles could not be determined because of difficulties in extraction. However, in-foil analysis of γ' particle compositions indicated an enrichment of Si and Ni. In other work [15], the compositions of γ' have been found to be approximately 65Ni–20Si–15(Cr,Fe,Mn). γ' was found to be a relatively pure Si- and Ni-rich phase

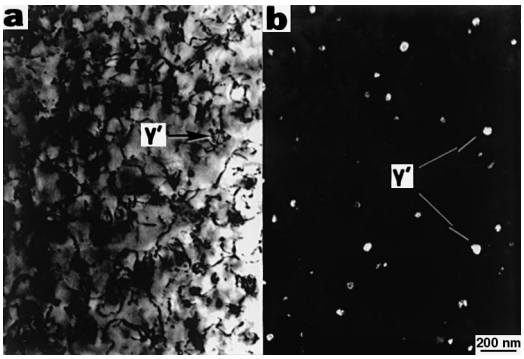


Fig. 5. γ′ particles in E5 (no carbon addition) alloy irradiated to 70 dpa at 948 K.

compared with G-phase, which accommodates a greater fraction of transition elements such as Ti, Mo, and Cr. Of interest was that G-phase formation was inhibited at the expense of γ′ in the E5 alloy, perhaps due to the absence of carbon in this alloy as mentioned already.

Although γ′ phase has been frequently observed in AISI 316 type austenitic stainless steels neutron-irradiated at low temperatures (<550°C) [15,24,25], it has not been observed under heavy ion irradiation where a high damage rate and a high temperature (>500°C) are generally employed. In a separate irradiation for D9C1 alloy, similar to E1 but having higher silicon (1 wt%) and lower molybdenum (1.5 wt%), however, γ′ formation was also observed under the same ion irradiation conditions used for B- and E-series alloys. Therefore, it is believed that γ′ can form even under high damage rate ion irradiation and at higher temperatures, should the Si supersaturation be increased and G-phase formation be suppressed by lowering Mo.

3.9. Ni-rich-gamma

The composition of alloy B3 was chosen to examine the effect of silicon, namely by adding 0.8 wt% silicon to the base ternary. Surprisingly, the expected γ′ phase did not develop in this alloy during Ni ion irradiation, but instead a high-nickel quaternary precipitate phase developed. EDS analysis revealed that a typical composition of the precipitates was 45Ni–42Fe–8Cr–5Si. The Si and Ni concentration in these precipitates were about six- and three-fold higher, respectively, compared with those in the matrix, 71Fe–15Ni–13Cr–0.8Si. The shape of the precipitates appeared generally as spheres or as a coating around cavities when He was coimplanted, often linking-up to form peanut-shaped dumbbells, suggesting that Si and Ni segregation around cavities or defects was the major cause for the formation of this phase. Cavity coatings have been observed also previously [25–28].

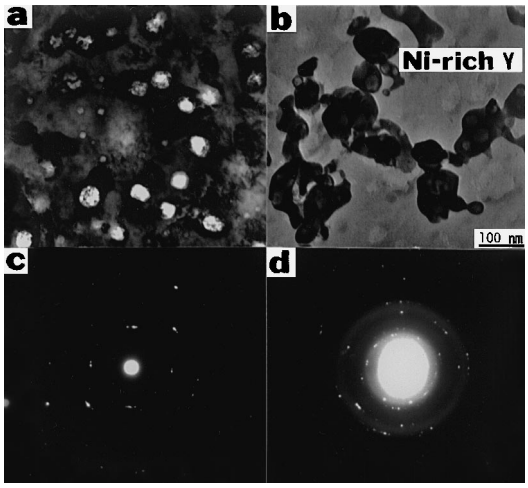


Fig. 6. Alloy B10 irradiated to 60 dpa with Ni-ions and 440 appm He at 948 K: (a) Ni-rich-γ coatings around the cavities, (b) extracted particles on carbon film.

The Ni-rich-γ phase developed also in alloy B10 although it had phosphorous in addition to silicon. The in-foil and extracted Ni-rich-γ particles are shown in Fig. 6(a) and (b), respectively, for alloy B10 irradiated to 60 dpa and 440 appm He at 948 K. Strain contrast around the precipitates is evident due to misfit. Fig. 6(c) and (d) show diffraction patterns for the in-foil and extracted particles, respectively. The diffraction patterns from the foils showed skewed spots superimposed upon the matrix diffraction pattern, indicating that the Ni-rich-γ phase had a skewed cube on cube orientation relative to the matrix due to coherent strain at the phase interface. The degree of misorientation increased as the particles grew. The electron diffraction rings from the extracted precipitates showed an fcc structure, the same as that of the matrix (γ), with the lattice parameter of $a = 0.356$ nm, i.e., about 1% smaller than that of the matrix.

3.10. CrSi

An unusual but quite abundant chromium silicide phase was also formed in alloy B3 (Si) during irradiation. This phase did not develop in alloy B10 (Si, P) because the phosphorous addition promoted phosphide formation. CrSi precipitates were not visible in the TEM foil and were identified from the extraction replica. EDS analysis revealed that the CrSi composition was typically 43Cr–42Si–14Fe–1Ni, having approximately equal Cr and Si stoichiometry. Besides the stoichiometric CrSi phase, there were small fractions of high-Cr phases with various Cr/Si ratios. The electron diffraction from these particles produced very complex powder ring patterns (Fig. 7) from which no specific crystallographic structure

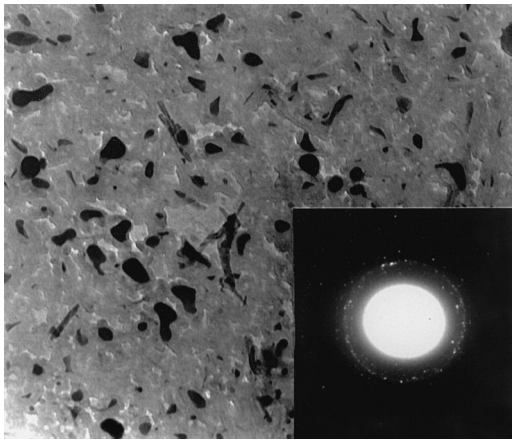


Fig. 7. Extraction replica of chromium silicide particles from B3 irradiated to 70 dpa at 948 K. The insert is a diffraction pattern from the extracted particles.

determination could be made at this time. The structure of CrSi is listed as cubic, FeSi B20 (space group T^4-P213) with lattice parameter $a = 0.462$ nm in Ref. [29].

4. Discussion

Phase transformations under irradiation are very complex phenomena and are affected by various intrinsic and extrinsic factors such as alloy composition, thermomechanical treatment, and irradiation conditions (temperature, damage rate, helium generation, etc.). In this work we have examined the phase stability in 28Fe–15Ni–13Cr alloys as a function of various alloying additions after irradiation with 4 MeV Ni ions at 948 K.

The results showed that there were four types of precipitate phases depending upon the solubility of silicon and nickel in the phases, similar to a pattern laid out in a previous publication [12]. Details are summarized below:

1. Si-free and low-Ni: thermal phases ($M_{23}C_6$, TiC, TiO);
2. high-Si and moderate-Ni: radiation modified phases (η, NbC, M_2P);
3. high-Si and low-Ni: radiation-induced phase (CrSi), and
4. high-Si and high-Ni: radiation-induced phase (G, γ', Ni-rich-γ).

The first type of precipitates are thermal phases whose compositions were not affected by radiation, but their formation was either enhanced or suppressed depending upon alloy composition. The second type of precipitates are radiation modified phases whose compositions were altered by radiation, but again their formation was either enhanced or suppressed depending upon alloy composition and irradiation conditions. The third and fourth types were truly radiation induced phases which were new phases that did not develop during thermal aging of these austenitic steels or similar stainless steels.

Interestingly, among the six thermally stable phases in the first two categories above, the non-carbide (M_2P, TiO) phases were enhanced by radiation, while the remaining four carbide phases ($M_{23}C_6$, η, MC, NbC) showed either enhancement or retardation depending upon the alloy composition. The formation of TiC and NbC was generally enhanced due to profuse nucleation sites produced by radiation damage combined with radiation-enhanced diffusion when there was sufficient carbon. In such a case, $M_{23}C_6$ and η were suppressed as a result of carbon depletion caused by MC formation. The stability of carbides was also affected by the formation of M_2P or TiO, because these phases compete for common solutes, C and Ti. For example, the formation of M_2P and/or TiO phase depleted titanium from solution leaving excess carbon in solution. This resulted in the suppression of MC and the promotion of $M_{23}C_6$.

Phosphide formation was extremely sluggish under thermal aging, but by contrast was remarkably enhanced by radiation. Diffusion coefficients of carbon and phosphorous in γ-iron at 1183 K (910°C) are 1.5×10^{-11} m^2/s and 3.6×10^{-16} m^2/s, respectively [30]. The sluggish formation of phosphide under thermal conditions is thus reasonable. The enhancement of the phosphides was, however, also observed in a high phosphorous (Fe–18Cr–10Ni–0.3P) austenitic stainless steel during quenching and aging experiments [31] This was a critical experiment which demonstrated that vacancy supersaturation and defects induced by quenching are primary reasons for the enhanced phosphide formation. Although the alloys investigated here had several times less phosphorous (0.05 wt%), it is believed that the enhanced diffusion by the increased vacancy supersaturation together with the defects produced by damage attributed to the rapid phosphide formation during irradiation. Furthermore, the high Si-affinity of phosphide encouraged its formation as Si segregated at defects under irradiation. In these ion-irradiation experiments, it was not determined whether the phosphide was a thermal equilibrium phase; i.e., whether the phosphide particles would dissolve again if annealed subsequently. However, an extensive development of phosphide was observed in MC-strengthened austenitic steels subjected to creep test at 700°C [32]. Similar results were observed in an AISI 321 stainless steel (0.03 wt% P) aged for 17 yr at 873 K [33], suggesting that phosphide is a thermal phase.

The NbC phase accommodated an unusually high silicon concentration and showed a refinement of its mean particle size with increasing silicon, but the volume fraction of NbC decreased as a Nb-rich G-phase

developed. This trend is consistent with the nature of competition for niobium between NbC and G-phase. Since the thermal NbC phase did not retain any silicon, it is believed that silicon segregation may have reduced NbC stability in favor of the Si-accommodating G-phase.

The η phase was suppressed during ion irradiation. This result is rather surprising because Si-segregation should have helped the formation of η phase since it has a high solubility for Si. As mentioned already, the damage rate is an important factor for the stability of η phase. It is a thermal carbide phase which requires high Cr, Mo, Si, and Ni. Therefore its stability appeared to be linked to the balance between radiation effects (Si, Ni segregation) and thermal effects (C, Cr, Mo segregation). An increase in damage rate without the corresponding increase in temperature may have reduced the ratio of (Cr, Mo)/(Si, Ni) at sinks, destabilizing the η phase. Moreover, η phase is a carbide and the enhancement of MC (depletion of C) under irradiation may have caused an adverse effect on the stability of η phase.

When thermal processes are sluggish, a true equilibrium state may be unobtainable in a given temperature range. In the vigorous damage environment, however, atoms are constantly displaced, replaced in lattice sites and reconstructed (i.e., each atom experienced on average 70 displacements after 70 dpa) and diffusion is enhanced by the increased point defect supersaturations, expediting the approach to an equilibrium state. Two such examples were observed in this experiment: (1) radiation-enhanced nucleation of phosphide as seen above, and (2) radiation catalyzed transition from a meta-stable TiC to a stable TiO phase.

During irradiation an unusual enhancement of TiO phase was observed. This process was accompanied by a reduction of TiC and a moderate increase in $M_{23}C_6$ phase in comparison to the thermal aging cases. The reversed stability from TiC (thermal) to TiO (radiation) raises the question of which phase is truly a thermally stable phase.

As briefly mentioned already, the source of oxygen for TiO is believed to come from oxide inclusions introduced during alloy melting; this was confirmed in Part I. In Fe–Cr–Ni austenitic steel, most dissolved oxygen exists in the form of chromium oxide because Cr_2O_3 is the most stable phase among Cr_2O_3, Fe_3O_4, and NiO [34]. Under thermal aging the formation of TiC proceeds rapidly because of the fast carbon diffusion, incorporating virtually all of the Ti into this tightly bound compound, while oxygen remains tightly bound in Cr_2O_3. The rates of back reactions releasing Ti and O are evidently negligible at the aging temperature of 948 K. However, under irradiation the atomic bonds are constantly broken in displacement cascades and the ejected atoms from all precipitates are free to diffuse through the matrix and to participate again in all possible reactions. The diffusion processes themselves are also radiation enhanced because of the high radiation-induced supersaturation of point defects. Calculations based on the Gibb's free energy of formation data reveal that TiO is a more stable phase than TiC under thermal conditions. However, it did not form during thermal aging because of low concentration of free oxygen in solution [35]. Under irradiation, however, the reaction between Ti and O is energetically favored because of the availability of oxygen released through cascade dissolution. While it is well known that irradiation can drive a system away from equilibrium, this result showed that irradiation also can catalyze the formation of a more stable phase at the expense of a persistent metastable phase.

There were four phases which were not found in aged specimens, but which are clearly induced by irradiation. These include G, γ′, Ni-rich-γ, CrSi, all rich in Si and/or Ni, showing that Si and Ni segregation during irradiation is a predominant factor in the formation of these particular phases. A verification of the radiation-induced nature of these phases can be made by post-irradiation annealing; G and γ′ phases have been found to dissolve back into solution during annealing at irradiation temperature [13–15]. Although post-irradiation annealing tests have not been carried out for Ni-rich-γ and CrSi phases, they are also believed to be radiation induced phases because they are not found in aged specimens, have unusually high silicon (i.e., 42 at.% in CrSi), and appear as coatings around cavities that are produced only during irradiation (Ni-rich-γ). It is well known that Si and Ni segregate at sinks by preferential solute-defect interactions during irradiation [7–10], thereby resulting in non-equilibrium phase formation. Of interest was that some Ni-rich-γ particles had >40% Ni, but had the same fcc structure and was still in the same γ phase region as the matrix. However, the abrupt phase boundary between matrix and coating indicates that there may exist a miscibility gap between the two distinct phase compositions.

Under neutron irradiation, Laves, η, and σ phases were enhanced and their formation temperature range extended downwards in AISI 316 and 316+Ti steels [12]. Thus these phases can be classified as radiation enhanced thermal phases. Such enhancement was, however, not observed with 4 MeV Ni or Fe ion irradiation. The suppression of these phase under ion irradiation are attributed to sluggish nucleation and growth rates. Extended defects formation and solute segregation under ion irradiation occur in a time scale less than a few hours while equivalent processes take place over several months to a few years for neutron case. Phases with slow kinetics seem to have not enough time to form under ion irradiation, before elemental resources and nucleation sites are exhausted by the formation of fast nucleating precipitate phases such as MC, M_2P, and G phase.

5. Conclusions

Precipitate phases in 28 different Fe–15Ni–13Cr austenitic stainless steels have been investigated as a function minor alloying additions after 4 MeV Ni^+ irradiation at 948 K. The results showed that irradiation caused either enhancement, retardation, modification, or creation of precipitate phases. Principal causes for these changes are attributable to the production of profuse precipitate nucleation sites (loops, dislocations, bubbles, cavities) introduced by radiation damage, the preferential segregation of Si and Ni at defects during irradiation, the radiation-enhanced diffusion by the increased point defect supersaturations during irradiation, and cascade induced mixing and dissolution of metastable phases. Although the phase stability during irradiation showed an extreme sensitivity to the changes in solute supersaturations as various precipitates evolved, Si and Ni segregation is a predominant factor for enhancing Si-accommodating phases and suppressing non-Si phases. Of particular interest was that, although MC formed readily in the complex E-series alloys during irradiation, TiO was favored over TiC in the simple B-series alloys. This is qualitatively different from either a radiation-induced departure from equilibrium or an enhancement of prevalent thermal phases. This is a novel situation in which the stable phase is not produced on thermal aging but only during irradiation, as a result of cascade mixing and dissolution and enhanced kinetics during irradiation.

Although the phase stability under irradiation is complicated, the results provide crucial insights into the energetics of the mechanisms involved, and provide a basis for tailoring alloy compositions to achieve controlled microstructures. For fission and fusion reactor structural materials applications, control of transmutation produced gas atoms is a very important issue in suppressing swelling. Fortunately, M_2P and MC particles can be developed in a fine scale with high number densities during irradiation if the alloy composition is tailored correctly. These are very desirable features because swelling can be suppressed via effective dispersion of gas atoms at the precipitate–matrix interfaces while mechanical strength can be improved by pinning dislocations by these particles [14,32,36–38]. To achieve this end, alloy compositions should be optimized by maximizing the dispersion of M_2P and MC particles.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. The authors wish to thank Dr P.J. Maziasz and Dr A.F. Rowcliffe for reviewing this manuscript.

References

[1] R. Sizmann, J. Nucl. Mater. 69&70 (1978) 386.
[2] R.S. Nelson, J.A. Hudson, D.J. Mazey, J. Nucl. Mater. 44 (1972) 318.
[3] P. Wilkes, J. Nucl. Mater. 83 (1979) 166.
[4] A.F. Rowcliffe, E.H. Lee, J. Nucl. Mater. 108&109 (1982) 306.
[5] P. Chu, N.H. Ghoniem, J. Nucl. Mater. 117 (1983) 55.
[6] A.A. Turkin, C. Abromeit, V. Naundorf, J. Nucl. Mater. 233–237 (1996) 979.
[7] L.E. Rehn, P.R. Okamoto, D.I. Potter, H. Wiedersich, in: Effects of Radiation on Structural Materials, ASTM 683 (1979) 184.
[8] V.K. Sethi, P.R. Okamoto, in: J.R. Holland, L.K. Mansur, D.I. Potter (Eds.), Phase Stability during Irradiation, AIME, 1981, p. 109.
[9] T.R. Allen, G.S. Was, E.A. Kenik, J. Nucl. Mater. 244 (1997) 278.
[10] S. Watanabe, N. Sakaguchi, K. Kurome, M. Nakamura, H. Takahashi, J. Nucl. Mater. 240 (1996) 979.
[11] P.J. Maziasz, C.J. McHargue, Int. Mater. Rev. 32 (1987) 190.
[12] E.H. Lee, P.J. Maziasz, R.F. Rowcliffe, in: J.R. Holland, L.K. Mansur, D.I. Potter (Eds.), Phase Stability during Irradiation, AIME, 1981, p. 191.
[13] P.J. Maziasz, J. Nucl. Mater. 205 (1993) 118.
[14] E.H. Lee, L.K. Mansur, Metall. Trans. 23A (1992) 1977.
[15] P.J. Maziasz, J. Nucl. Mater. 169 (1989) 95.
[16] T.M. Williams, J.M. Titchmarsh, D.R. Arkell, J. Nucl. Mater. 107 (1982) 222.
[17] E.H. Lee, N.H. Packan, L.K. Mansur, J. Nucl. Mater. 117 (1983) 123.
[18] A.F. Rowcliffe, E.H. Lee, P.S. Sklad, in: International Conference on Irradiation Behavior of Metallic Materials for Fast Reactor Core Components, Supplement, Ajacia, Corsica, 1979, p. 1.
[19] K. Farrell, E.H. Lee, Scripta Metall. 17 (1983) 791.
[20] Stig Rundqvist, Ark. Kemi 20 (7) (1962) 67.
[21] E.H. Lee, E.A. Kenik, J. Mater. Res. 3 (1988) 840.
[22] M.K. Miller, J. Bentley, in: G.J. Theus, J.R. Weeks (Eds.), Environmental Degradation of Materials in Nuclear Power System – Water Reactors, The Metallurgical Society, 1988, p. 341.
[23] T.M. Williams, J.M. Titchmarsh, J. Nucl. Mater. 98 (1981) 223.
[24] C. Cawthorne, C. Brown, J. Nucl. Mater. 66 (1977) 201.
[25] H.R. Brager, F.A. Garner, J. Nucl. Mater. 73 (1978) 9.
[26] P.R. Okamoto, H. Widersich, J. Nucl. Mater. 53 (1972) 336.
[27] D.J. Mazey, D.R. Harris, J.A. Hudson, J. Nucl. Mater. 89 (1980) 155.
[28] L.K. Mansur, W.G. Wolfer, J. Nucl. Mater. 69&70 (1978) 825.
[29] A. Taylor, B.J. Kagle, Crystallographic Data on Metals and Alloy Structures, Dover, New York, NY, 1963.

[30] R.W.K. Honeycombe, Steels Microstructure and Properties, American Society for Metals, Metals Park, Ohio, 1981, p. 7.
[31] A.F. Rowcliffe, R.B. Nicholson, Acta Metall. 20 (1972) 143.
[32] R.W. Swindeman, P.J. Maziasz, in: Proceedings of the Fifth International Conference on Creep of Materials, Lake Buena Vista, FL, USA, May 1992, p. 33.
[33] J. Bentley, J.M. Leitnaker, in: E.W. Collings, H.W. Kings (Eds.), The Metal Science of Stainless Steels, Met. Soc. AIME, 1978, p. 70.
[34] O. Kubaschewski, C.B. Alcock, Metallurgical Thermochemistry, 5th Ed., Pergamon, New York, 1979.
[35] E.H. Lee, L.K. Mansur, J. Mater. Res. 2 (1987) 291.
[36] L.K. Mansur, E.H. Lee, P.J. Maziasz, R.F. Rowcliffe, J. Nucl. Mater. 141–143 (1986) 633.
[37] W. Kesternich, J. Nucl. Mater. 127 (1983) 153.
[38] K. Herschbach, W. Schneider, K. Ehrlich, J. Nucl. Mater. 203 (1993) 233.

ELSEVIER

Nuclear Instruments and Methods in Physics Research B 166–167 (2000) 188–192

NIM B
Beam Interactions with Materials & Atoms

www.elsevier.nl/locate/nimb

Influence of oxygen ion implantation on the damage and annealing kinetics of iron-implanted sapphire

Carl J. McHargue [a,*], John D. Hunn [a,1], E. Alves [b], M.F. da Silva [b], J.C. Soares [b]

[a] *Centers for Materials Processing, University of Tennessee, 100 Estabrook Hall, Knoxville, TN 37996-2350, USA*
[b] *Nuclear and Technological Institute, E.N. 10, 2685 Sacavem, Portugal*

Abstract

The effects of implanted oxygen on the damage accumulation in sapphire which was previously implanted with iron was studied for (0 0 0 1) sapphire. The energies were chosen to give similar projected ranges. One series was implanted with a 1:1 ratio (4×10^{16} ions/cm^2 each) and another with a ratio of 2:3 (4×10^{16} Fe^+/cm^2; 6×10^{16} O^+/cm^2). Retained damage, χ, in the Al-sublattice, was compared to that produced by implantation of iron alone. The observed disorder was less for the dual implantation suggesting that implantation of oxygen enhanced dynamic recovery during implantation. Samples were annealed for 1 hour at 800°C and 1200°C in an oxidizing and in a reducing atmosphere. No difference was found in the kinetics of recovery in the Al-sublattice between the two dual implant conditions. However, the rate of recovery was different for each from samples implanted with iron alone. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 61.43.Er; 61.82.Ms
Keywords: Sapphire; Co-implantation iron and oxygen; Annealing

1. Introduction

There is interest in the formation of nanometer-sized phases in insulator matrices due to the possibility of producing quantum confinement effects as the particle size approaches the scale of electron scattering. Nanometer-sized elemental phases and compounds have been produced by ion implantation. Compound formation is achieved by the implantation of two or more chemical species, often followed by annealing.

Because of the very low solubility of impurities, sapphire makes an excellent host for such studies. White and co-workers [1,2] have prepared a number of nanometer-sized semiconducting compounds (e.g., SiGe, GaAs, GaP, GaN, CdS, CdSe, ZnS, PbS) in sapphire by sequential implantation and post-implantation annealing.

The present study explores the amount of the residual damage and the possibility of compound formation for sequential implantation of iron and

* Corresponding author. Tel.: +1-423-974-7680; fax: +1-423-974-4995.
E-mail address: crl@utk.edu (C.J. McHargue).
[1] Present address: Metals and Ceramics Division, Oak Ridge National Laboratory.

0168-583X/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S 0 1 6 8 - 5 8 3 X (9 9) 0 0 6 5 7 - 6

oxygen into sapphire at room temperature (~300 K) and the effect of annealing temperature and atmosphere on damage recovery.

According to the phase diagrams, Al_2O_3 and Fe_2O_3 are soluble in each other to 5–10% at 1100°C [3], thus, a solid solution, $(Al_xFe_{1-x})_2O_3$ may form during the oxygen implantation. Other iron oxides, $Fe_{1-x}O$ and $FeAl_2O_4$, should appear as second phase precipitates [4].

2. Experimental procedures

High purity single crystals of Al_2O_3 with the *c*-axis normal to the surface were obtained from Crystal Systems, with optically polished surfaces. Samples were annealed for 120 h at 1500°C in flowing O_2 to remove any residual polishing damage and surface contamination. One set was implanted at room temperature with 4×10^{16} Fe^+/cm^2 (160 keV) followed by implantation of 4×10^{16} O^+/cm^2 (54 keV), giving an implanted Fe:O ratio of 1:1. Another set was implanted with 4×10^{16} Fe^+/cm^2 followed by implantation of 6×10^{16} O^+/cm^2, giving an implanted Fe:O ratio of 2:3. The energies were chosen such that the ranges were similar. Because of differences in longitudinal straggling, the concentration ratios may vary slightly with depth and differ slightly from these ratios. In order to avoid channeling of the implanted ions, the incident beam direction was 5–7° from the crystal axes.

Samples of each set were annealed for 1 hour at 800°C and 1200°C in an oxidising (oxygen or air) or reducing (Ar–4% H_2) atmosphere. A few samples were annealed at 500°C and 1500°C.

The residual damage was determined from Rutherford backscattering (RBS)-ion channeling using a beam of 1.6 MeV He^+. The incident beam direction was normal to the specimen surface and the scattering angle of detection was 140°. The residual damage, χ, was scaled to that produced by implantation of 4×10^{16} Fe/cm^2 (160 keV) using the value of deposited damage energy (S_D) calculated with the code E-DEP-1 (P-5.0) [5]. Experimental values for samples implanted with only iron were taken from a previous study [6].

3. Results

3.1. Effects of oxygen-implantation on residual damage

There were no significant differences in the RBS-ion channeling spectra from the as-implanted samples having different Fe:O ratios. Values of χ for the Al-sublattice (0.62 and 0.60, 4:4 Fe:O and 4:6 Fe:O, respectively) were lower than for the sample implanted only with 4×10^{16} Fe/cm^2, (0.68). These results are shown in Fig. 1. Also shown in Fig. 1 are the values scaled with total deposited damage energy (S_D). If the disorder scaled directly with damage energy, values of 0.86 and 0.95 would be expected. However, dynamic recovery during implantation makes a linear extrapolation of χ with damage energy unlikely. A previous paper [6] reported χ as a function of iron fluence for similar implantation conditions. Using those data as a basis of extrapolation for the dual implanted samples gives values of $\chi = 0.73$ and 0.75 for the 4:4 and 4:6 samples, respectively.

The observed values are lower than values projected by either method, suggesting that the subsequent implantation of oxygen enhanced recovery in the Al-sublattice. The "error bars" in Fig. 1 indicates values ±5% of the measured or calculated values.

3.2. Damage recovery during post-implantation annealing

Values of χ were determined after annealing for 1 hour at 500°C, 800°C and 1200°C in an oxidising

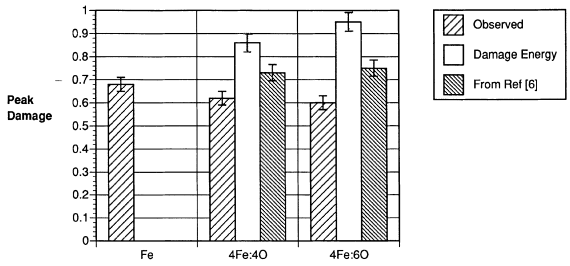


Fig. 1. Damage (χ) in the Al-sublattice measured by RBS-ion channeling and as scaled with calculated damage energy deposition and with damage observed in [6].

Table 1
Fraction damage remaining after annealing

Temperature (°C)	Oxidizing			Reducing		
	Fe only	4:4	4:6	Fe only	4:4	4:6
500	0.9	0.89	0.88	0.9	0.89	0.89
800	0.2	0.19	0.26	0.53	0.19	0.22
1200	0.19	0.10	0.13	0.14	0.08	0.05
1500	0.05			0.05		

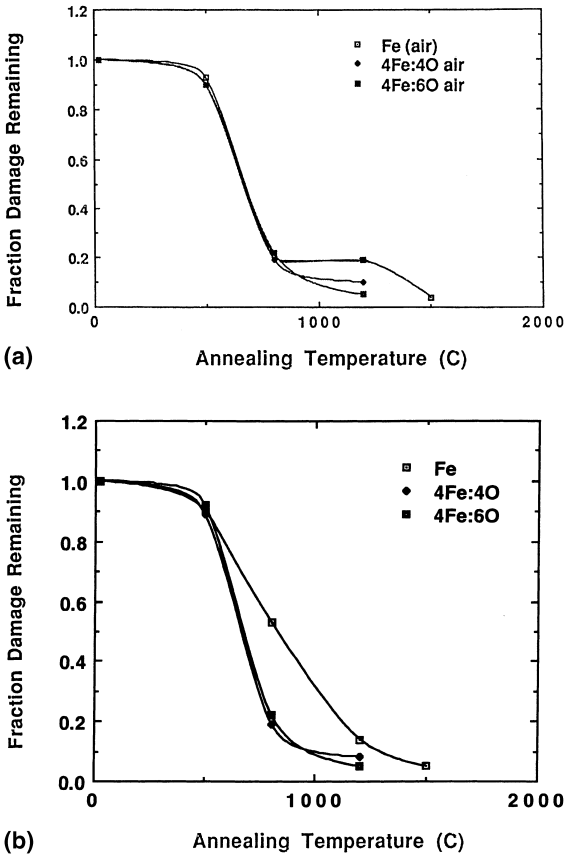


Fig. 2. Recovery of disorder in Al-sublattice during post-implantation annealing: (a) oxidizing atmosphere. (b) reducing atmosphere.

or reducing atmosphere. The data, normalised by dividing $\chi_{annealed}$ by $\chi_{as\text{-}implanted}$, are summarized in Table 1 and plotted in Fig. 2 as the fraction damage remaining versus annealing temperature. Data for samples implanted with iron only, taken from [6], are shown for comparison.

The curves of Fig. 2(a) show that recovery of disorder in the Al-sublattice during annealing in the oxidising atmosphere is essentially the same at temperatures below 800°C for samples implanted with iron only, 4:4 Fe:O and 4:6 Fe:O. As indicated in Table 1, the fraction of retained damage after the 800°C treatment was slightly higher in the 4:6 Fe:O sample than in the others. Little additional recovery of the Fe-implanted sample occurred in the temperature range of 800–1200°C. This observation is consistent with the two-stage recovery process reported earlier [7]. Lattice damage in the dual implanted samples continued to anneal in this temperature range, and similar fractions of the as-implanted disorder remain after the 1200°C anneal.

Fig. 2(b) shows that disorder for both dual implanted conditions recovered at about the same rate in the reducing atmosphere. The samples implanted only with iron recovered at significantly slower rates at 800°C and 1200°C than did the dual implanted samples.

3.3. Iron redistribution during annealing

The RBS-ion channeling spectra for the iron peak are shown in Figs. 3 and 4. There were no significant differences among the samples annealed at 800°C. All exhibited increased channeling in the aligned spectra, but no shift in the position of the peak. No iron was lost from the implanted zone. These observations are in marked contrast to the

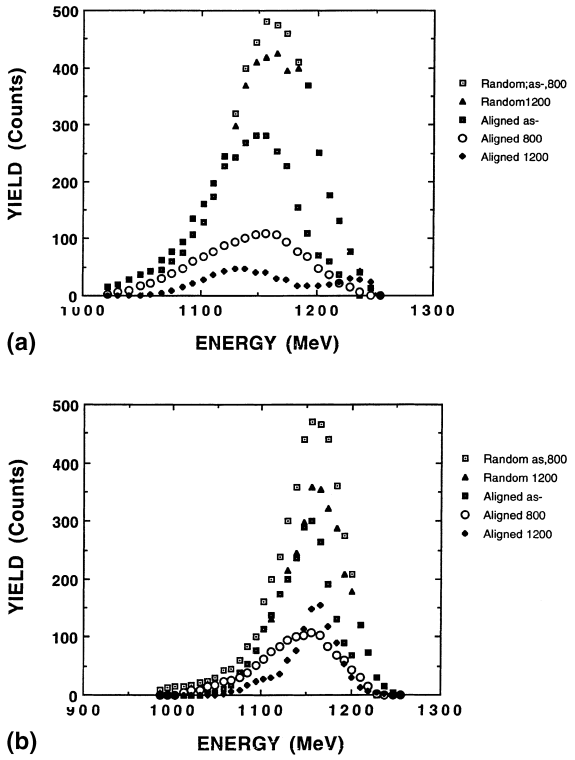


Fig. 3. Iron spectra for 4Fe:4O annealed in (a) air, (b) Ar–4% H_2 (□: random for as-implanted and 800°C; ▲: random for 1200°C; ■: aligned, as-implanted; ○: aligned, 800°C; ◆: aligned, 1200°C).

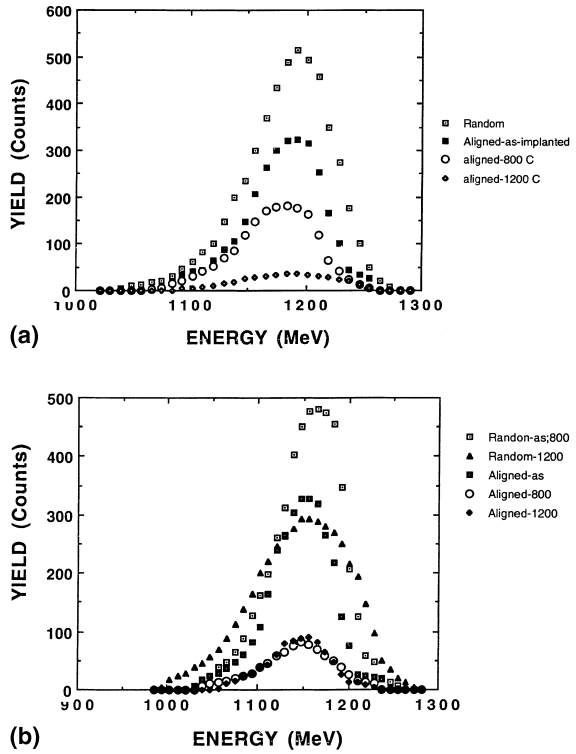


Fig. 4. Iron spectra for 4Fe:6O annealed in (a) air, (b) Ar–4% H_2 (□: random for as-implanted and 800°C; ▲: random for 1200°C; ■: aligned, as-implanted; ○: aligned, 800°C; ◆: aligned, 1200°C).

results for samples implanted only with iron and annealed in oxygen [7]. In that case, a surface peak was present after the 800°C anneal in oxygen, indicating a migration of the implanted iron to the surface.

The 4:4 sample annealed in air at 1200°C developed a surface peak in the iron spectrum but there was no change in the position of the peak for the 4:6 sample annealed under the same conditions. A slight (~5%) amount of iron was lost from the 4:4 sample but there was no detectable loss of iron from the 4:6 sample.

The spectra for the 4:4 sample annealed in Ar–4% H_2 at 1200°C suggest that some iron diffused into and toward the surface of the sapphire during this anneal. On the other hand, the aligned curves were identical for the 4:6 samples annealed in the same atmosphere at 800°C and 1200°C. The peak position did not shift from the as-implanted position, there was only a broadening of the peak. The position of the peak did not change in samples implanted with only iron annealed under similar conditions [7].

Many of the iron atoms, either as isolated-implanted atoms or in iron-containing compounds are shielded by the aluminum atoms as viewed along the *c*-axis. The value of χ_{Fe} for iron was lower for the dual implanted samples than for the iron implanted samples in the as-implanted specimen and after each annealing treatment. Earlier studies found that iron oxides formed during annealing in oxygen have a high degree of coherency in the [0 0 0 1] direction but not along other crystallographic axes [7]. Iron nanocyrstals formed during high fluence implantation and during reducing anneals are also highly oriented along the *c*-axis [8].

4. Discussion

The reduced disorder in the dual as-implanted samples, compared to the single implant of iron, could be due to in situ compound formation or to the implanted oxygen replacing oxygen displaced by knock-ons. At this fluence, most of the iron is present as Fe^{2+} with a small amount present as nanometer size metallic iron precipitates [9]. It was proposed that the Fe^{2+} was present as iron-oxygen vacancy clusters. Oxygen implanted after the iron may annihilate these vacancies, causing the iron to precipitate. The oxygen could also promote the formation of iron oxides. Studies in progress using transmission electron microscopy and grazing angle X-ray diffraction should define the location of the iron.

The lower values of χ_{Fe} for the dual implanted samples is consistent with the formation of coherent second phases.

A previous study [7] of Fe-implanted sapphire found that recovery of the disorder in the Al-sublattice had a plateau between 800°C and 1200°C (see Fig. 2(b)). The TEM studies cited in [7] identified small precipitates of $FeAlO_4$ in regions near the peak iron concentration after an anneal at 700°C in oxygen and precipitates of α-$(Fe,Al)_2O_3$ at and near the surface after a 1200°C anneal. The plateau in the recovery curve was thought to be controlled by the in-diffusion of oxygen to reach a level where the mixed Fe–Al oxide could form. The implantation of oxygen after the iron may provide sufficient oxygen for in situ precipitation. Removal of Fe from the Fe–O vacancy clusters in the presence of implanted oxygen should promote recovery of the Al_2O_3 lattice.

The increased recovery of the dual implanted samples for the 800°C anneals in the reducing atmosphere may also be due to the presence of the implanted oxygen in the region where Fe^{2+} is being reduced to Fe^0, thereby releasing oxygen vacancies from the charge compensating clusters, allowing reformation of the sapphire lattice.

5. Summary

The amount of disorder in the Al-sublattice is less than expected for the implantation of Fe followed by the implantation of oxygen, compared to the results for implantation of Fe only.

The recovery of disorder in the Al-sublattice during post-implantation annealing was enhanced by the dual implantation in both oxidizing and reducing atmospheres. It is suggested that the implanted oxygen reacts with the iron to form iron (or iron-aluminum) oxides which are coherent with the sapphire lattice as viewed along the *c*-axis.

Acknowledgements

This study was supported in part by the Division of Materials Sciences, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. A portion of the work was conducted in the Surface Modification and Characterisation Facility, Oak Ridge National Laboratory.

References

[1] C.W. White, J.D. Budai, S.P. Withrow, J.G. Zhu, E. Sonders, R.A. Zuhr, A. Meldrum, D.M. Hembree, D.O. Henderson, S. Prawer, Nucl. Instr. and Meth. B 141 (1998) 228.
[2] C.W. White, A. Meldrum, J.D. Budai, S.P. Withrow, E. Sonders, R.A. Zuhr, D.M. Hembree, M. Wu, D.O. Henderson, Nucl. Instr. and Meth. B 148 (1999) 991.
[3] A. Muan, C.L. Gee, Am. J. Sci. 256 (1956) 415.
[4] L.M. Atlas, W.K. Sumida, J. Am. Ceram. Soc. 41 (1958) 158.
[5] C.M. Davisson, I. Manning, NRL Report 8859, Naval Research Laboratory, Washington DC, 1986.
[6] C.J. McHargue, E. Alves, M.F. da Silva, J.C. Soares, Nucl. Instr. and Meth. B 148 (1999) 730.
[7] G.C. Farlow, P.S. Sklad, C.W. White, C.J. McHargue, J. Mater. Res. 5 (1990) 102.
[8] C.J. McHargue, S.X. Ren, J.D. Hunn, Mater. Sci. Engrg. A 253 (1998).
[9] C.J. McHargue, P.S. Sklad, C.W. White, G.C. Farlow, A. Perez, G. Marest, J. Mater. Res. 6 (1991) 2145.

ELSEVIER

Nuclear Instruments and Methods in Physics Research B 166–167 (2000) 193–197

NIM B
Beam Interactions with Materials & Atoms

www.elsevier.nl/locate/nimb

Influence of tin ion implantation on the damage and annealing kinetics of sapphire

Carl J. McHargue [a,*], L.J. Romana [b]

[a] *Center for Materials Processing, University of Tennessee, 514 East Stadium Hall, Knoxville, TN 37996-0750, USA*
[b] *Universite des Antilles et de la Guyane, 97159 Pointe-a-Pitre Cedex, Guadeloupe, France*

Abstract

In order to study the amorphization of sapphire during room temperature ion implantation and the recrystallization of ion-implantation-induced amorphization, sapphire single crystals were implanted with tin (180 keV) at room temperature to fluences from 5×10^{15} to 1×10^{17} Sn/cm^2. The calculated damage energy at the amorphous/crystalline interface varied from 85 eV/atom (10^{16} Sn/cm^2) to 1.4 eV/atom (4×10^{16} Sn/cm^2). The concentration of Sn at the interface was less than 2 at%. Samples implanted with 4×10^{16} Sn/cm^2 were annealed for 1 hour in oxidizing or reducing atmospheres at a temperature of 700°C, 900°C, 960°C or 1100°C. Regrowth at temperatures below 960°C followed the sequence: amorphous $\rightarrow \gamma$-Al_2O_3; γ-$Al_2O_3 \rightarrow \alpha$-$Al_2O_3$. Essentially all (97%) implanted tin was lost during the 1100°C anneal in Ar–4% H_2, whereas all Sn was retained in the sample annealed in the oxidizing atmosphere. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 61.43.Er; 61.82.Ms
Keywords: Sapphire; Sn-implantation; Amorphization; Regrowth

1. Introduction

Ion implantation of solids presents a unique method for altering the near-surface structure of a wide range of materials in a manner that is independent of many of the constraint associated with conventional processing methods. Current activities in the area of integrated optics and the tailoring of oxides for specific applications have led to renewed interest in the use of ion implantation to alter the near-surface optical, electrical, magnetic or mechanical properties of amorphous and crystalline oxides.

Sapphire (single crystal α-$A1_2O_3$) can be crystalline with lattice disorder, e.g., damage in the form of dislocations and point defect clusters, or it can be amorphous after implantation [1]. The degree of disorder depends upon the implantation parameters of ion fluence, ion species, target temperature and orientation of the ion beam relative to crystallographic axes. The ion fluence required for amorphization at room temperature is significantly lower for some ion species than for others

* Corresponding author. Tel.: +1-423-974-7680; fax: +1-423-974-4995.
E-mail address: crl@utk.edu (C.J. McHargue).

0168-583X/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0168-583X(99)00654-0

[2]. This observation has led to the speculation of the existence of some "chemical" effect on amorphization.

For many potential applications, it will be necessary to anneal the material after implantation in order to remove the radiation damage, recrystallize the amorphous layer to the crystalline state, and/or redistribute the implanted species. The regrowth of sapphire has been extensively studied only for crystals amorphized by sequential implantation of Al + O in the ratio of 2:3 and the implantation of Fe, both at 77 K [3]. These studies found the amorphous Al_2O_3 first transformed to γ-Al_2O_3 and then the γ-Al_2O_3 transformed to α-Al_2O_3 by epitaxial growth from the α–γ interface in the temperature range of 700–1000°C. There was, however, a difference in the kinetics of the transformations between the (Al + O) and the Fe-implanted samples. The effect of annealing atmosphere (oxidizing or reducing) has received only limited attention.

The purpose of the present study is to study in greater detail amorphization due to implantation of Sn at room temperature and to study the effects of annealing atmosphere on the regrowth of the matrix and the distribution of the tin. An earlier study reported limited results for Conversion electron Mössbauer spectroscopy and transmission electron microscopy (TEM) examinations of a sample implanted with 4×10^{16} Sn/cm^2 [4].

2. Experimental details

High-purity Al_2O_3 single crystals having the *c*-axis normal to the surface were obtained with an optical polish from Crystal Systems, Inc. (Salem, MA). The disks were annealed for 120 h at 1450°C in flowing oxygen to remove any residual polishing damage or surface contamination. Minimum yields, χ_m, were ~2% in the Al sublattice and ~8% in the oxygen sublattice after annealing, indicating a high degree of perfection.

Crystals were implanted with ^{119}Sn (180 keV) at room temperature to fluences of 5×10^{15}, 1×10^{16}, 2×10^{16}, 4×10^{16} and 10^{17} Sn/cm^2 with the beam ~7° off-normal. Values of the projected range, R_p, and straggling, ΔR_p, calculated from the code EDEP-1(P.5) [5] were 53.3 and 7 nm, respectively.

Samples implanted with 4×10^{16} Sn/cm^2 were annealed for 1 h at 700, 900, 960 or 1100°C in an oxidizing (flowing oxygen) or reducing (Ar–4% H_2) atmosphere.

Specimens were analyzed by Rutherford backscattering ion channeling spectroscopy (RBS) using 2.0 MeV He^+ ions at a scattering angle of 160°. Specimens were prepared for TEM by mechanical polishing followed by ion milling.

3. Results and discussion

3.1. As-implanted

The RBS spectra of Fig. 1 shows the aligned spectrum to reach the random value at a fluence between 5×10^{15} and 1×10^{16} Sn/cm^2. The aligned scattering from the Al-sublattice for the 5×10^{15} Sn/cm^2 sample exhibits a strong broad surface peak and a second maximum at a distance approximately equal to $R_p + \Delta R_p$. The amorphous (random) layer extends to a depth of ~70 nm for a fluence of 1×10^{16} Sn/cm^2 and 80 nm at 2×10^{16} Sn/cm^2. This layer was about 95 nm thick at 4×10^{16} Sn/cm^2.

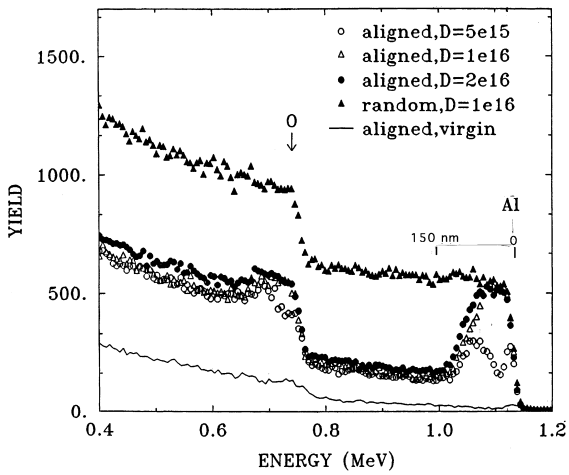


Fig. 1. RBS-ion channeling using 2 MeV He^+ for Al_2O_3 implanted with $^{119}Sn^+$/cm^2 at room temperature.

The value of deposited damage energy, $S(D)$, at the positions of the crystalline–amorphous interfaces, calculated using EDEP-1(P5), has values of 85, 28 and 1.4 eV/atom for fluences of 1, 2 and 4×10^{16} Sn/cm^2, respectively. The position of the interface does not correlate with $S(D)$. If a value of 18 eV is used for the displacement energy (E_d) for an Al ion from its lattice site, a calculation indicates that Al will not be displaced by incident Sn^+ at depths greater than about 60 nm. A similar calculation for oxygen, using $E_d = 72$ eV, suggests that oxygen ions will not be displaced at depths greater than about 40 nm. Ignoring any effect of differences in scattering cross-sections, these calculations suggest that an excess of Al-vacancies over O-vacancies will be created in a zone 40–60 nm from the surface. This corresponds to R_p for the implanted Sn.

The creation of an Al-deficient (or O-rich) zone could explain the TEM observations of the earlier study wherein two "amorphous" were observed after implantation to 4×10^{16} Sn/cm^2 [4]. The contrast in TEM images indicated one phase to be Sn-rich. The relative values of the free energies of formation, ΔG°, of α-Al_2O_3, SnO_2 and Al–Sn–O compounds indicate that Sn cannot reduce α-Al_2O_3. However, if the Al has been removed by energetic collisions, the Sn might react with the excess oxygen. The CEMS results for the same study indicated that amorphous SnO_2 or $SnAl_2O_5$ and $SnAl_2O_4$ were present in this sample.

The Sn/Al ratio at the position of the amorphous–crystalline interface was determined from the RBS-ion channeling measurements to have values of 1–2 at%.

3.2. Post-implantation annealing studies – oxidizing atmosphere

There were no significant changes in the RBS spectra for samples annealed 1 h at 700°C and 900°C in air. The previous CEMS study [4] showed tin to be present as Sn(II) and Sn(IV) in the as-implanted condition and in the same relative amounts after an 700°C anneal. Only the Sn(IV) state was present after anneals at 900°C and higher. In-diffusion of oxygen through the amorphous surface layer oxidized the Sn(II) state to the Sn(IV) state.

There was a slight (~5 nm) movement of the α–amorphous interface toward the surface during the 960°C anneal and movement of the Sn toward the surface. The peak Sn:Al ratio at the surface was 0.09 and 0.10 at R_p. There was no channeling of the Sn.

There were major changes in the RBS after the 1100°C anneal, Fig. 2. The strong peak from the

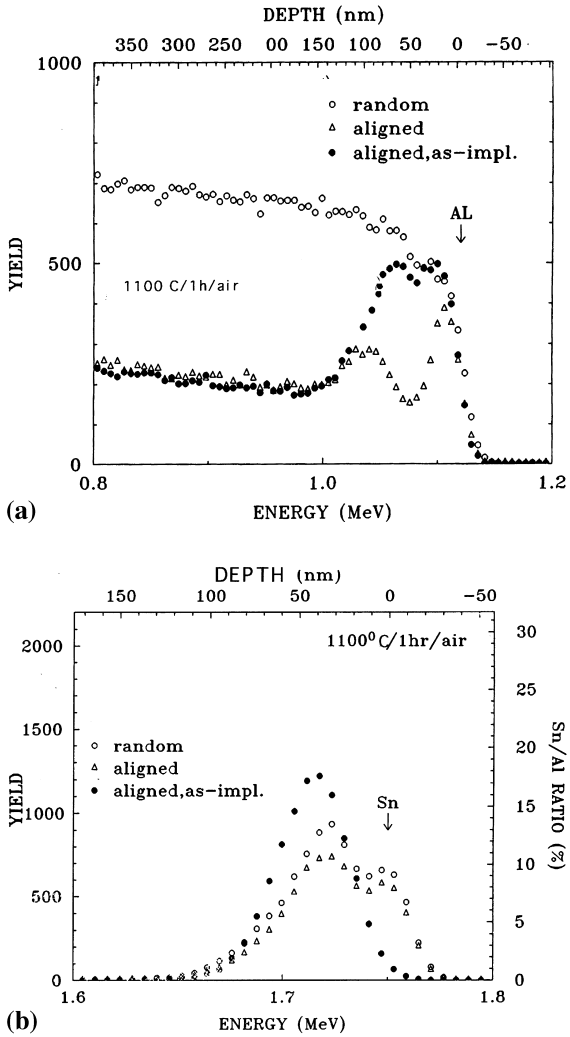


Fig. 2. RBS-ion channeling using 2 MeV He^+ for Al_2O_3 implanted with 4×10^{16} $^{119}Sn^+$/cm^2 at room temperature and annealed 1 h in air at 1100°C. (a) Al-sublattice; (b) Sn.

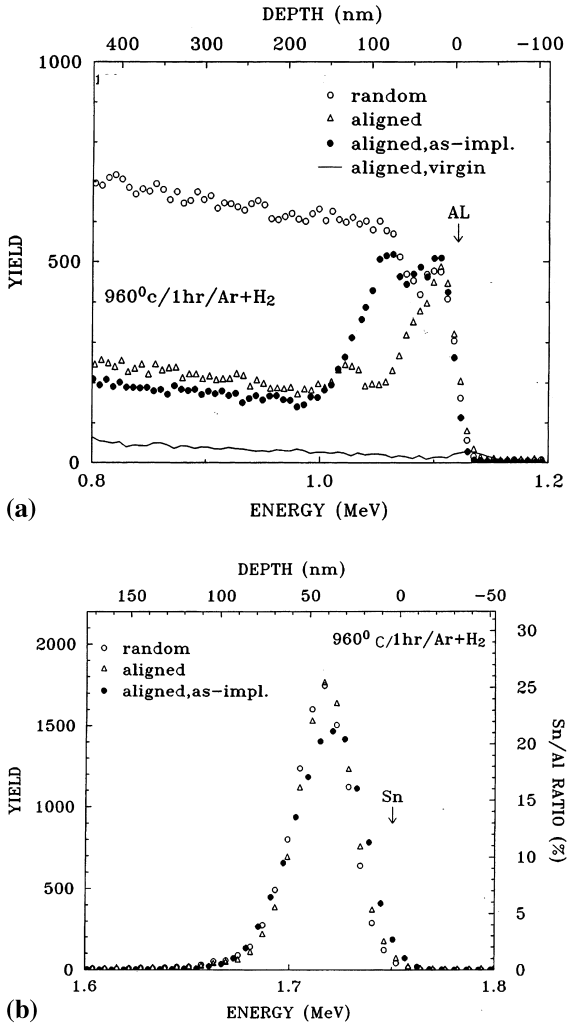


Fig. 3. RBS-ion channeling using 2 MeV He^+ for Al_2O_3 implanted with 4×10^{16} $^{119}Sn^+/cm^2$ at room temperature and annealed 1 h in Ar–4% H_2 at 960°C. (a) Al-sublattice; (b) Sn.

surface to ~30 nm suggests either a microcrystalline structure or, based on earlier studies [3], γ-Al_2O_3 with substructure. An examination by TEM showed that it is γ-Al_2O_3. The region from ~30 nm to the original amorphous/crystalline interface was α-Al_2O_3. A number of voids were detected in the region from ~30 to ~70 nm, corresponding to the region of higher channeling in Fig. 2(a). These voids may be due to condensation of Al-vacancies, as suggested above. Selected area diffraction patterns showed the Sn-rich precipitates to be SnO_2.

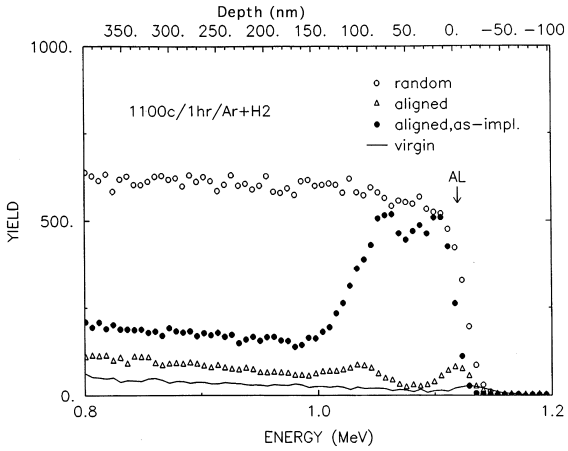


Fig. 4. RBS-ion channeling from Al using 2 MeV He^+ for Al_2O_3 implanted with 4×10^{16} $^{119}Sn^+/cm^2$ at room temperature and annealed 1 h in Ar–4% H_2 at 1100°C.

The precipitates and both Al_2O_3 phases exhibited a high degree of coherency with the following orientation relationships:

$$(2\,0\,0)_{SiO_2} || (0\,0\,0\,1)_{\alpha} || (1\,1\,1)_{\gamma},$$

$$[0\,0\,2]_{SiO_2} || [3\,0\,\bar{3}\,0]_{\alpha} || [4\,4\,0]_{\gamma}.$$

Fig. 2(b) shows a tendency for a surface peak in the Sn spectra as well as movement of the Sn from the position of R_p toward the surface. All of the implanted Sn was retained during this anneal. The melting point of SnO_2 is about 1730°C; thus, this oxide should be stable at this annealing temperature.

3.3. Post-implantation annealing studies – reducing atmosphere

Annealing at 700° did not change the RBS spectra for the Al or Sn. Annealing at 900°C caused the α-amorphous interface to migrate ~15 nm toward the surface. The Al-spectrum did not reach the random level, and the appearance indicates the formation of γ-Al_2O_3. There was no change in the Sn-spectra.

The spectra of Fig. 3(a) indicate that the α–γ interface is ~50 nm from the surface, indicating epitaxial regrowth of the α into the γ during the

960°C anneal. The Al spectrum suggests the regrown α is not as perfect as observed in the study of samples implanted with (Al + O) or Fe [3]. No changes were noted in the Sn spectra (Fig. 3(b)).

The spectra of Fig. 4 indicate complete regrowth to the α-phase occurred during the 1100°C anneal. The regrown phase contained more disorder than the virgin crystal, with values of $\chi \approx 0.26$ at the surface and ≈0.2 at 70 nm. Although 97% of the Sn was lost from the sample during this anneal, the spectra for both Al and Sn show a bimodal shape, suggesting that the disorder remaining in the Al sublattice is connected with the presence of this small amount of tin. Reduction of the Sn from Sn(II) and Sn(IV) to the metallic state is expected for the reducing anneals. Since Sn melts at 232°C, it is not surprising that most of the Sn was lost during this anneal. The retained tin appears to be tied up as $SnAl_2O_4$.

4. Summary

Sapphire crystals having the *c*-axis approximately parallel to the ion beam become amorphous at a fluence between 5×10^{15} and 1×10^{16} ions/cm^2 for implantation of Sn^+ at room temperature. The calculated damage energy and measured Sn concentration at the crystalline–amorphous interface are very low.

Only a slight amount of regrowth from the crystalline substrate into the amorphous zone occurs in 1 h for annealing in a oxidizing atmosphere at temperatures of 960°C and below. Annealing 1 h at 1100°C produces a surface layer of γ-Al_2O_3 from the surface to ~30 nm and regrown α-Al_2O_3 in the region ~30–95 nm. Examination by TEM reveals SnO_2 particles coherent with both phases. There is a tendency for Sn to migrate toward the free surface; however, very little Sn is lost during annealing. The presence of the Sn as SnO_2 precipitates may prevent the loss of Sn by evaporation.

Migration of the α-Al_2O_3 substrate into the amorphous layer is detected after 1 h at 900°C for anneals in a reducing atmosphere. A surface layer of γ-Al_2O_3 (0–50 nm) and epitaxially regrown α-Al_2O_3 (50–95 nm) was present after annealing 1 h at 960°C. Complete regrowth to α-Al_2O_3 was observed after the 1 h anneal at 1100°C; however, this α-Al_2O_3 contained significant disorder. The disorder may be due to the observed presence of $SnAl_2O_4$ precipitates. Approximately 97% of the Sn was lost during the 1100°C anneal.

These observations suggest that the interaction of Sn and/or Sn–Al–O compounds with irradiation-produced defects is very strong, effectively pinning them in place.

Acknowledgements

These studies were supported in part by the Department of Energy, Division of Materials Sciences, under contract DE-AC05-96OR with Lockheed Martin Energy Research Corporation. The implantations were performed in the Surface Modification and Characterization Facility of Oak Ridge National Laboratory. Special thanks go to Dr. P.S. Slad, Oak Ridge National Laboratory for assistance with the electron microscopy.

References

[1] C.J. McHargue, Mater. Sci. Eng. A 253 (1998) 94.
[2] C.J. McHargue, P.S. Sklad, P. Angelini,. C.W. White, J.C. McCallum, The structure of amorphous Al_2O_3 produced by ion implantation, in: P. Borgesen, J.A. Knapp, R.A. Zuhr, (Eds.), Beam–Solid Interactions: Physical Phenomena, Mater. Res. Soc., Pittsburgh, PA, 1990, pp. 505–512.
[3] C.W. White, L.A. Boatner, P.S. Sklad, C.J. McHargue, J. Rankin, G.C. Farlow, M. Aziz, Nucl. Instr. and Meth. B 32 (1988) 11.
[4] C.J. McHargue, P.S. Sklad, J.C. McCallum, C.W. White, A. Perez, E. Abonneay, G. Marest, Nucl. Instr. and Meth. B 46 (1990) 74.
[5] C.M. Davisson, I. Manning, NRL Report 8859, Naval Research Laboratory, Washington, DC, 1986.

ELSEVIER

Journal of Nuclear Materials 276 (2000) 22–32

journal of nuclear materials

www.elsevier.nl/locate/jnucmat

The role of cascade energy and temperature in primary defect formation in iron

Roger E. Stoller*

Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6376, USA

Abstract

Molecular dynamics (MD) simulations have been used to investigate the formation of atomic displacement cascades in iron with energies up to 50 keV (corresponding to a primary knock-on atom (PKA) energy of 79 keV) at 100 K, up to 20 keV at 600 K, and up to 10 keV at 900 K. The cascade damage production has been characterized in terms of several parameters: the number of surviving point defects, the fraction and type of surviving point defects found in clusters, and the size distributions of the in-cascade point defect clusters. A sufficient number of simulations have been completed at each condition to evaluate the statistical variation in these primary damage parameters as a function of irradiation temperature and cascade energy. The energy dependence of stable defect formation can be conveniently separated into three regimes with the number of defects in each regime correlated by a simple power law with a characteristic exponent. The primary effects of cascade energy on defect formation at high energies are explained in terms of subcascade formation. Only a modest effect of temperature is observed on defect survival, while irradiation temperature increases lead to a slight increase in the in-cascade interstitial clustering fraction and a decrease in the vacancy clustering fraction. Cascade energy has little effect on the in-cascade clustering fractions above about 5 keV. However, there is a systematic change in the cluster size distribution, with higher energy cascades producing larger clusters. The loosely coupled nature of the in-cascade vacancy clusters persists at higher energies. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

The method of molecular dynamics (MD) has been broadly applied in recent years to the investigation of primary damage formation in irradiated materials [1–10]. Modern, high speed computers with large amounts of memory have permitted researchers to examine higher energy events with their requirement for larger numbers of atoms. These same computing resources have permitted cascade simulations to be completed in sufficient numbers to obtain statistically meaningful mean values, and for comparisons to be made between different materials. For example, a previously reported comparison of cascade damage formation in iron and copper was based on a database of over 600 cascades [5]. The results presented in that study comprise cascade energies from 60 eV to 10 keV for irradiation temperatures of 100 and 600 K, and up to 5 keV at 900 K. The average number of defects that survived in the iron and copper MD cascade simulations at each energy was also compared with the number of displacements obtained from the standard secondary displacement model by Norgett, Robinson, and Torrens (NRT) [11] with the recommended value of the atomic displacement threshold for both materials [12].

MD investigations of displacement cascades in a number of materials have established several consistent trends in primary damage formation. These trends include: (1) the total number of stable point defects produced follows a power-law dependence on the cascade energy over a broad energy range, (2) the ratio of MD stable displacements divided by the number obtained from the NRT model decreases with energy until subcascade formation becomes prominent, (3) the in-cascade clustering fraction of the surviving defects increases with cascade energy, and (4) the effect of lattice temperature on the MD results is rather weak. Two additional observations have been made regarding in-cascade clustering in iron. First, the interstitial clusters have a complex, three-dimensional morphology, with

* Tel.: +1-423 576 788; fax: +1-423 574 0641.
E-mail address: rkn@ornl.gov (R.E. Stoller)

0022-3115/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0022-3115(99)00204-4

both sessile and glissile configurations. Mobile interstitial clusters appear to glide with a low activation energy similar to that of the mono-interstitial (~0.1–0.2 eV) [13]. Second, the vacancy clustering fraction is much lower than the interstitial clustering fraction. If based on a nearest-neighbor clustering criterion, the fraction of surviving vacancies and interstitials in clusters for high energy cascades is about 3% and 20% of the NRT displacements, respectively.

The results presented here summarize further statistical analysis of the iron cascade database completed at lower energies to provide a more complete characterization of the energy and temperature dependence of cascade damage production. Additional cascade simulations have extended the range of the previous studies to a peak cascade energy of 50 keV at 100 K and 20 keV at 600 K. Completing higher energy simulations provides an opportunity to validate the trends established at lower energies. In particular, they further probe the energy domain of extensive subcascade formation. Both the total number of point defects produced and the fraction of these defects found in clusters at the end of the cascade event appeared to be relatively independent of temperature at the lower energies. The in-cascade cluster size distributions exhibit more temperature dependence at the higher energies, with vacancy clustering decreasing and interstitial clustering increasing as the temperature increases.

2. MD simulation method

The calculations were carried out using the MOLDY [14] MD code and the interatomic potential for iron developed by Finnis and Sinclair for α-iron [15,16] and subsequently modified by Calder and Bacon [3]. This code and potential have been widely used, and are discussed in detail elsewhere [2–10]. The simulations were carried out using periodic boundary conditions and at constant pressure. Since the boundary atoms were not damped to remove heat, the initiation of a cascade event led to some lattice heating. The primary impact of the temperature increase is to extend the lifetime of the thermal spike, it has little or no impact on the ballistic phase of the cascade. The effect of this temperature increase on primary damage formation is not yet clear. There is some evidence that the adiabatic conditions may have an impact on the in-cascade interstitial clustering fraction [17]. In addition, although previous work has not shown a significant impact of temperature on total defect survival [3,5,10], some of the results presented here suggest that relatively large temperature increases may reduce the number of stable point defects produced.

Prior to conducting the cascade simulations, a block of atoms was thermally equilibrated at the temperature of interest for about 10 ps. This equilibrated atom block was used as the starting point for the subsequent cascade simulations; the size of the block is increased for higher energy simulations as shown in Table 1. The simulations were initiated by giving a lattice atom the specified amount of kinetic energy and an initial direction; a high index lattice direction such as ⟨1 3 5⟩ was typically used to minimize channeling. Additional cascades were differentiated by changing the location of the primary knock-on atom (PKA) or by using a different equilibrated atom block. As discussed in Ref. [5], the initial kinetic energy is analogous to the damage energy in the NRT [11] model. The kinetic energy lost to electronic excitation accounts for the difference between the PKA energy and damage energy and this energy loss mechanism is not simulated in the MOLDY code [5]. Both the MD simulation energies and the corresponding PKA energies are listed in Table 1. The neutron energies listed in Table 1 are those for which the PKA energy is the average recoil energy in an elastic collision. Note that these MD simulations account for recoils from neutrons that are well into the high (fast) energy domain, i.e. energies greater than about 0.1 MeV.

Table 1
Typical particle energy parameters, corresponding NRT displacements, and simulation cell size for MD cascade simulations

Neutron energy (MeV)	Average PKA energy (keV)	Corresponding E_{MD} (keV)	NRT displacements	Atoms in simulation
0.0034	0.116	0.1	1	3456
0.0058	0.236	0.2	2	6750
0.014	0.605	0.5	5	6750
0.036	1.24	1	10	54 000
0.074	2.54	2	20	54 000
0.19	6.60	5	50	128 000
0.40	13.7	10	100	250 000
0.83	28.8	20	200	250 000
1.8	61.3	40	200	1 024 000
2.3	78.7	50	500	2 249 728

3. Primary damage formation in MD simulations

The parameters used in the following discussion to describe results of MD cascade simulations are the total number of surviving point defects and the fraction of the surviving defects contained in clusters. The number of surviving defects will be expressed as a fraction of the NRT displacements listed in Table 1, whereas the number of defects in clusters will be expressed as either a fraction of the NRT displacements or a fraction of the total surviving MD defects. The criterion used to define a point defect cluster in this study was that each defect in the cluster must be in a nearest-neighbor position to another defect. As discussed below, defect clusters are not necessarily compact structures under this definition. In order to analyze the statistical variation in the primary damage parameters, the mean value (M), the standard deviation about the mean (σ), and the standard error of the mean (ε) have been calculated for each set of cascades conducted at a given energy and temperature. The standard error of the mean is calculated as $\varepsilon = \sigma/n^{0.5}$, where n is the number of cascade simulations completed [18]. The standard error of the mean provides a measure of how well the sample mean represents the actual mean. For example, a 90% confidence limit on the mean is obtained from $1.86 * \varepsilon$ for a sample size of nine [19]. These statistical quantities are summarized in Table 2 for total defect survival and interstitial clustering. Work is in progress to obtain a more complete temperature comparison, e.g. 10 keV cascades at 600 K and 20 keV cascades at 900 K.

3.1. Total point defect survival

Previous work indicates that the number of stable displacements (vacancy–interstitial pair) remaining at the end of a cascade simulation, N_D, exhibits a power-law dependence on cascade energy [7,8]. For example, after analyzing simulations in iron between 0.5 and 10 keV at 100 K, Bacon et al. found that the total number of surviving point defects could be expressed as [7]

$$N_D = 5.67 E_{MD}^{0.779}, \quad (1)$$

Table 2
Statistics of primary damage parameters derived from MD cascade simulations

Temperature (K)	Energy (keV)	Number of cascades	Surviving displacements (mean/standard deviation/standard error)		Clustered interstitials (mean/standard deviation/standard error)		
			Number	Per NRT defects	Number	Per NRT defects	Per MD surviving defects
100	5	8	17.1	0.343	8.38	0.168	0.488
			2.59	0.0517	1.85	0.0369	0.0739
			0.916	0.0183	0.654	0.0130	0.0261
	10	15	33.6	0.336	17.0	0.170	0.506
			5.29	0.0529	4.02	0.0402	0.101
			1.37	0.0137	1.04	0.0104	0.0261
	20	10	60.2	0.301	36.7	0.184	0.610
			8.73	0.0437	6.50	0.0325	0.0630
			2.76	0.0138	2.06	0.0103	0.0199
	30	9	93.2	0.311	n/c[a]	n/c	n/c
			14.9	0.0498			
			4.98	0.0166			
	40	8	131.0	0.328	74.5	0.186	0.570
			12.6	0.0315	15.0	0.0375	0.102
			4.45	0.0111	5.30	0.0133	0.0361
	50	9	168.3	0.337	93.6	0.187	0.557
			12.1	0.0242	6.95	0.0139	0.0432
			4.04	0.00807	2.32	0.00463	0.0144
600	20	8	55.8	0.279	41.6	0.208	0.747
			5.90	0.0295	5.85	0.0293	0.0782
			2.09	0.0104	2.07	0.0104	0.0276
900	10	7	27.3	0.273	18.6	0.186	0.677
			5.65	0.0565	6.05	0.0605	0.1605
			2.14	0.0214	2.29	0.0229	0.0607

[a] n/c denotes that the analysis of interstitial clusters is not completed at this time.

where the cascade energy, E_{MD} is given in keV. This relationship did not hold below about 0.5 keV because true cascade-like behavior does not occur at these low energies. Subsequent work by Stoller [8,10] indicated that N_D also begins to deviate from this energy dependence above 20 keV when extensive subcascade formation occurs. New results at 30 and 50 keV confirm that the energy dependence of defect survival changes near 20 keV. This is illustrated in Fig. 1(a), where Eq. (1) is plotted along with the average number of surviving interstitials at energies up to 50 keV. At each energy, the data point is an average of between 7 and 26 cascades. Error bars are included for the 10–50 keV data to indicate the standard error of the mean. The range of plus or minus one standard error is barely detectable around the data points, indicating that the change in slope is statistically significant.

In addition to the new results at 30 and 50 keV, two of the values in Fig. 1 are different from those reported previously [3,5,10]. As part of an investigation of the influence of pre-existing damage on defect formation, eight 10 keV simulations at 100 K were carried out in perfect crystal to provide the baseline for comparison. The results of these eight simulations have been averaged together with the seven reported previously. This resulted in a negligible change in the average number of surviving defects from 33.9 to 33.6. This small change in average is consistent with the small standard errors shown in Fig. 1. A more significant change has been made in the value at 0.3 keV. Although 16 cascade simulations were completed at this energy by Calder and Bacon [3], the total point defect survival and interstitial clustering fraction values obtained at 0.3 keV appeared to be inconsistent with the rest of the data. Ten additional simulations have been carried out and the new average of all 26 simulations is used in this paper. With this new data, it appears that three well-defined regions with different energy dependencies exist. A power-law fit to the points in each energy region is shown in Fig. 1(a). The best-fit exponent in the absence of true cascade conditions below 0.5 keV is 0.485. From 0.5 to 20 keV, the exponent is 0.795. This is marginally higher than the value in Eq. (1), possibly because the 20 keV data were used in the current fitting. An exponent of 1.12 was found in the range above 20 keV which is dominated by subcascade formation. In each energy range, the MD results deviate from the linear energy dependence predicted by the NRT model.

The data from Fig. 1(a) are replotted in Fig. 1(b) where the number of surviving displacements is divided by the NRT displacements at each energy. The rapid decrease in this MD defect survival ratio at low energies is well known. The error bars again reflect the standard error and the solid line through the points is a least-squares fit to the energy dependence that was used in an assessment of neutron energy spectrum effects [20]. Note that at very low energies, the number of surviving displacements in the MD simulations exceeds the NRT value. This is a result of using the average displacement threshold energy of 40 eV in the NRT model. The actual displacement threshold varies with crystallographic direction and is as low as ~19 eV in the [1 0 0] direction [21]. This simply implies that the low energy (near threshold) simulations preferentially produce displacements in the 'easy' directions [22].

Just as the 30–50 keV data points fall above the values predicted by the mid-energy power-law curve in Fig. 1(a), the surviving defect fraction shows a slight increase as the cascade energy increases from 20 to 50 keV, and the standard errors shown indicate that the increase is statistically significant. Although it is not possible to quantitatively associate a specific defect survival fraction with a given cascade morphology, it is interesting to compare 10 and 50 keV cascades since the

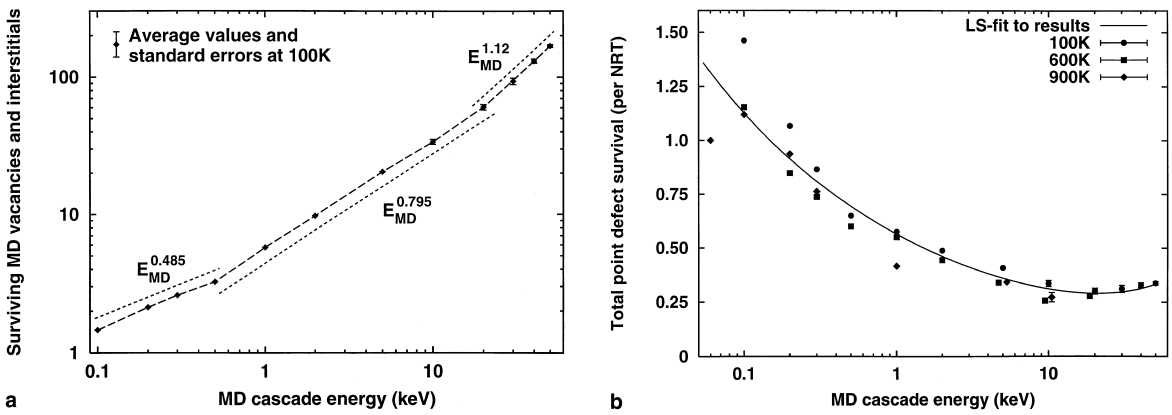


Fig. 1. Cascade energy dependence of stable point defect formation in iron MD cascade simulations: (a) total number of interstitials or vacancies at 100 K; (b) ratio of defects to NRT displacements. Data points indicate mean values at each energy and error bars indicate standard error of the mean.

defect survival fractions are nearly the same at both energies. Such a comparison is shown in Fig. 2, where the peak damage state from typical 10 and 50 keV cascades are displayed. An average defect survival fraction of ~0.3 as shown for the 10–50 keV cascades is consistent with values estimated from resistivity change measurements following low-temperature neutron irradiation [22,23].

Although it is difficult to discern a consistent effect of temperature between the 600 and 900 K data points, the defect survival fraction at 100 K is always somewhat greater than at either of the two higher temperatures. This is in agreement with the effect of temperature reported by Gao et al. [17] in their study of 2 and 5 keV cascades. Calculated mean survival values and standard errors for 10 keV cascades at 900 K, and 20 keV cascades at 600 K are shown in Fig. 1(b). These results also indicate a significant reduction in defect survival with increasing temperature. Further verification of the temperature dependence of defect survival will be provided by work which is currently in progress: a statistical analysis of all the lower energy cascades, and the completion of 10 and 20 keV cascades simulations at 600 and 900 K, respectively.

3.2. In-cascade interstitial clustering

The dependence of in-cascade interstitial clustering on cascade energy is shown in Fig. 3 for simulation temperatures of 100, 600, and 900 K, where the average number of interstitials in clusters has been divided by the corresponding number of NRT displacements. Statistical parameters are listed in Table 2 and the number of cascades is the same as in Fig. 1. The solid line in Fig. 3(a) is again the least-squares fit to the data that was used to help assess neutron energy spectrum effects [20]. Data points at 0.1, 10, and 20 keV have been slightly offset along the energy axis to help visually distinguish between the results. Although no physical significance should be assigned to the structure of the curve above 10 keV in Fig. 3(a), it can be understood by considering Figs. 1(b) and 3(b). The clustering per NRT (Fig. 3(a)) is essentially the product of clustering per MD (Fig. 3(b)) times the surviving MD defects per NRT (Fig. 1(b)).

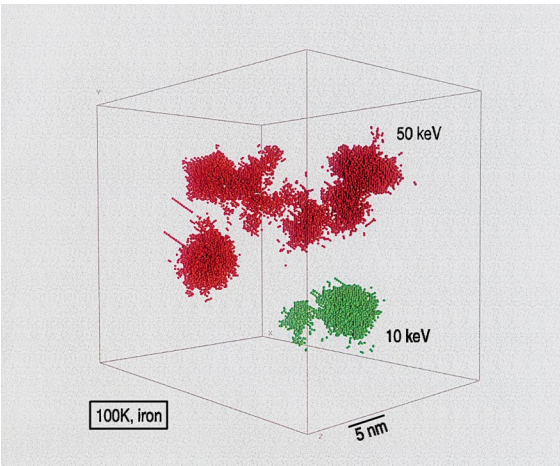


Fig. 2. Illustration of typical cascade morphology and subcascade formation in 10 and 50 keV simulations.

Considerable scatter is observed in the clustering fraction at low energies. This may be partially a result of differences between the tools used to characterize the low and high energy cascades. The higher energy simulations required more sophisticated visualization tools, and this potential difference has prompted the reanalysis of the lower energy cascades which is now in progress. Overall, the interstitial clustering fraction exhibits more variability between cascades at the same energy than does defect survival, and the validity of the indicated fit is therefore less certain. The standard errors in Fig. 3(a) are much larger than the corresponding values for the defect survival fraction. The fraction of interstitials in clusters is zero at the lowest energies and appears to saturate at about 18–20% of the NRT displacements above 5 keV, which corresponds to about 60% of the total surviving interstitials. Although it is not possible to discern a systematic effect of temperature below 10 keV, there is an apparent trend toward greater clustering with increasing temperature at higher energies. This can be more clearly seen in Fig. 3(b), where the ratio of clustered interstitials to surviving interstitials is shown, and in the cluster size distributions that will be discussed below. Such an effect of temperature on interstitial clustering in these adiabatic simulations is consistent with that observed by Gao et al. [17] in their work with a hybrid MD model that extracted heat from the simulation cell, i.e. the found that the clustering fraction increases with temperature.

The interstitial cluster size distributions exhibit a consistent dependence on cascade energy and temperature as shown in Fig. 4 (where a size of 1 denotes the single interstitial). The cascade energy dependence at 100 K is shown in Fig. 4(a), where only the size distributions at 10 and 50 keV are included to simplify the figure. The influence of cascade temperature is shown for 10 keV cascades in Fig. 4(b), and for 20 keV cascades in Fig. 4(c). All interstitial clusters larger than size 10 are combined into a single class in the histograms in Fig. 4. The interstitial cluster size distribution shifts to larger sizes as either the cascade energy or temperature increases. An increase in the clustering fraction at the higher temperatures is most clearly seen as a decrease in the number of mono-interstitials. Comparing Fig. 4(b) and (c) demonstrates that the temperature dependence increases as the cascade energy increases. The largest interstitial cluster observed in these simulations was

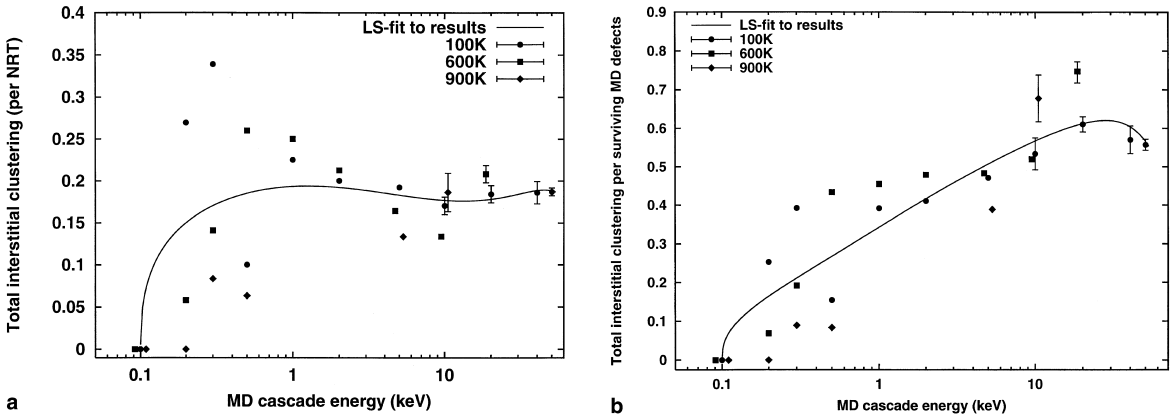


Fig. 3. Cascade energy dependence of interstitial clustering: (a) clustered interstitials divided by NRT displacements; (b) clustered interstitials divided by total surviving interstitials. Data points indicate mean values and error bars indicate standard error of the mean.

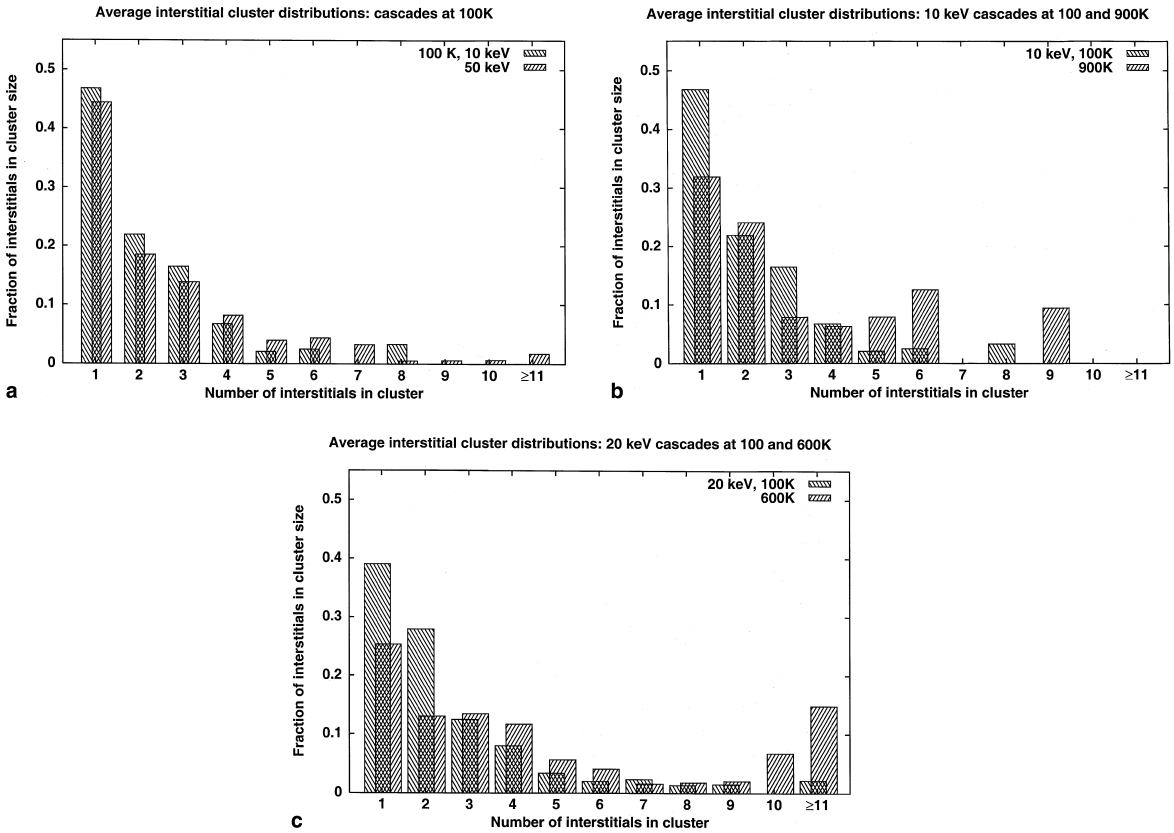


Fig. 4. Fractional size distributions of interstitial clusters formed directly within the cascade, comparison of: (a) 10 and 50 keV cascades at 100 K; (b) 10 keV cascades at 100 and 900 K; (c) 20 keV cascades at 100 and 600 K.

contained in a 20 keV cascade at 600 K as shown in Fig. 5. This large cluster was composed of 33 interstitials (⟨1 1 1⟩ crowdions and dumbbells), and exhibited considerable mobility via what appeared to be one-dimensional glide in a ⟨1 1 1⟩ direction [10].

The increase in the number of large clusters at high energies and temperatures suggested that the in-cascade cluster size distributions may exhibit more sensitivity to neutron energy spectrum than did either total defect survival or the total interstitial clustering fraction. At

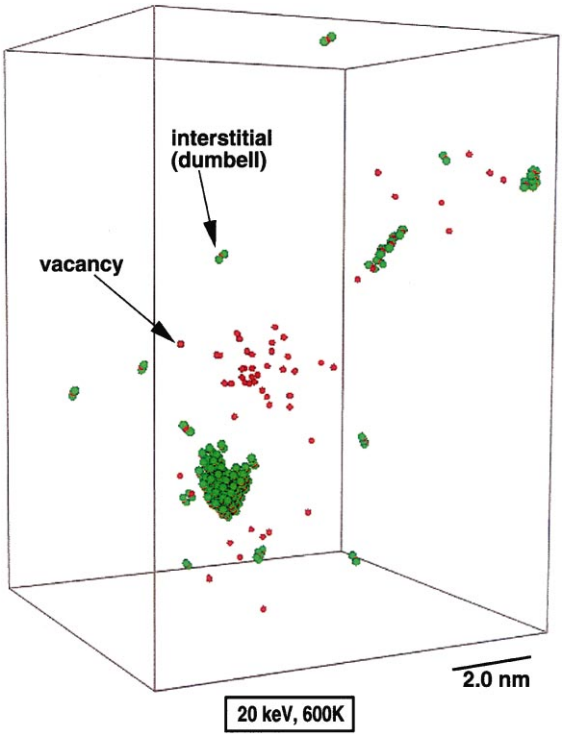


Fig. 5. Residual defects at ~30 ps from a 20 keV cascade at 100 K containing a 33-interstitial cluster.

100 K, there are no interstitial clusters larger than 8 for cascade energies of 10 keV or less. Therefore, the fraction of interstitials in clusters of 10 or more was chosen as an initial parameter for evaluation of the size distributions. This partial interstitial clustering fraction is shown in Fig. 6. Since the large clusters are relatively uncommon, the fraction of interstitials contained in them is correspondingly small. This leads to the relatively large standard errors shown in the figure. The 33-interstitial cluster shown in Fig. 5 plays a significant role in the sharp increase in this clustering fraction observed between 100 and 600 K for the 20 keV cascades. A least-squares fit to the three 100 K results is shown by the solid line in Fig. 6. Although the statistical basis for using the curve in Fig. 6 is more limited than for the curves in Figs. 1(b) and 3, this partial clustering fraction was also included in the evaluation of neutron spectrum effects discussed in Ref. [20]. The average fraction of interstitials in large in-cascade clusters for several neutron irradiation environments was obtained by weighting the function shown in Fig. 6 with their respective PKA spectra. These spectrum-averaged clustering values exhibited a significant variation between the various

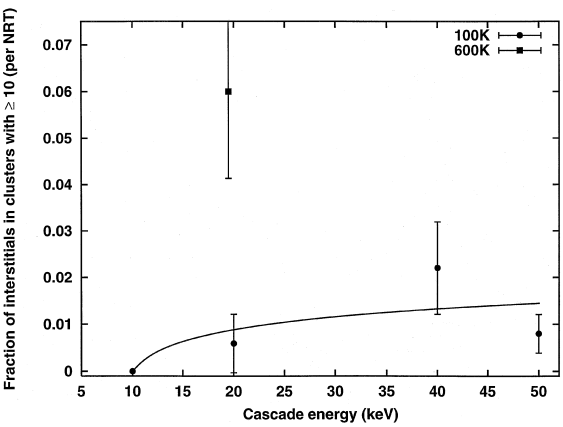


Fig. 6. Cascade energy dependence of interstitials contained in clusters of 10 or more: clustered interstitials divided by NRT displacements. Data points indicate mean values and error bars indicate standard error of the mean.

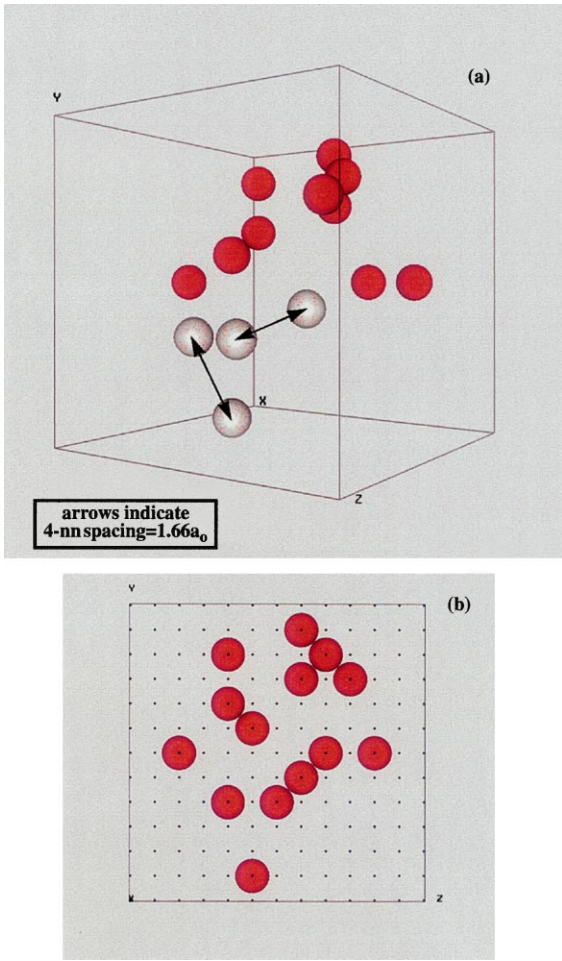


Fig. 7. Typical uncollapsed or nascent vacancy cluster from 50 keV cascade at 100 K; 14 vacancies are contained, each of which is within the 4-nn distance ($1.66a_0$). Perspective view shown in (a) and an orthographic projection in (b).

environments. This is in contrast to the spectrum-averaged total defect survival and total interstitial clustering fractions which show little dependence on neutron energy spectrum.

3.3. In-cascade vacancy clustering

As reported previously [5,10], the in-cascade vacancy clustering fraction in iron is quite low (~3% of NRT) when the nearest-neighbor (nn) criterion for clustering is applied. When an analysis of the spatial arrangement of the surviving vacancies in 10, 20, and 40 keV cascades was conducted, peaks in the distribution were obtained for vacancies in second (2-nn) and fourth (4-nn) nearest neighbor locations [10]. Similar results were obtained from the analysis of the vacancy distributions in 50 keV cascades at 100 K and the 20 keV cascades 600 K. The peak observed for vacancies in 2-nn locations is consistent with the di-vacancy binding energy being greater for 2-nn (0.22 eV) than for 1-nn (0.09 eV) [24]; the reason for the peak at 4-nn is unclear.

An example of these uncollapsed vacancy clusters is shown in Fig. 7 where a collection of 14 vacancies has been extracted from a 50 keV cascade at 100 K. Each of the vacancies has at least one other vacancy within the 4-nn spacing of $1.66a_o$. The 'cluster' is shown in two views; a 3-D perspective view is shown in Fig. 7(a), and an orthographic projection ($-x$) in Fig. 7(b). Such an arrangement of vacancies is similar to some of the vacancy clusters observed by Sato et al. in field ion microscope images of irradiated tungsten [25]. Since the time of the MD simulations is too short to allow vacancies to jump ($\leqslant 100$ ps), the possibility that these closely correlated vacancies might collapse into clusters over somewhat longer times has been investigated using Monte Carlo (MC) simulations. The vacancy coordinates at the end of the MD simulations were extracted and used as the starting configuration in MC cascade annealing simulations. The expectation of vacancy clustering was confirmed in the MC simulations, where many of the isolated vacancies had clustered within 70 μs [24,26].

The energy and temperature dependence of in-cascade vacancy clustering as a fraction of the NRT displacements is shown in Fig. 8 for cascade energies of 10–50 keV. Results are shown for clustering criteria of 1-nn, 2-nn, 3-nn, and 4-nn. A comparison of Figs. 8 and 3 indicates that in-cascade vacancy clustering in iron remains lower than that of interstitials even when the 4-nn criterion is used. This is consistent with the experimentally observed difficulty of forming visible vacancy clusters in iron as discussed in Ref. [5], and the fact that only relatively small vacancy clusters are found in positron annihilation studies of irradiated ferritic alloys [27]. The cascade energy dependence of vacancy clustering is similar to that of interstitials; a slight increase in vacancy clustering is observed as the cascade energy

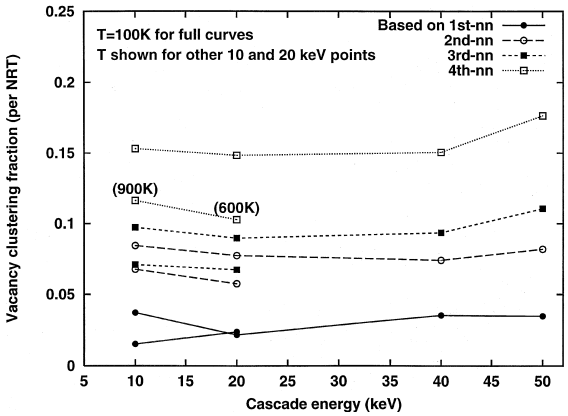


Fig. 8. Cascade energy dependence of vacancy clustering: clustered vacancies divided by NRT displacements. Data points indicate mean values and error bars indicate standard error of the mean.

increases from 10 to 50 keV. However, vacancy clustering decreases as the temperature increases.

Fractional vacancy cluster size distributions are shown in Fig. 9, for which the 4-nn clustering criterion has been used. Fig. 9(a) illustrates that the vacancy cluster size distribution shifts to larger sizes as the cascade energy increases from 10 to 50 keV. This is similar to the change shown for interstitial clusters in Fig. 4(a). There is a corresponding reduction in the fraction of single vacancies. However, as mentioned above, the effect of cascade temperature shown in Fig. 9(b) and (c) is opposite to that observed for interstitials. The magnitude of the temperature effect on the vacancy cluster size distributions also appears to be weaker than in the case of interstitial clusters. The fraction of single vacancies increases and the size distribution shifts to smaller sizes as the temperature increases from 100 to 900 K for the 10 keV cascades, and from 100 to 600 K for 20 keV cascades. As seen for interstitial clusters, the effect of temperature seems to be greater at 20 keV than at 10 keV.

4. Discussion and summary

As part of an ongoing effort to characterize primary damage formation in iron, MD displacement cascade simulations have been conducted at energies up to 50 keV and temperatures up to 900 K. A sufficient number of cascades have been completed to begin a statistical evaluation of the results. For example, the calculated standard errors of the mean values at any one energy are small enough to establish significant trends in the energy dependence. In the case of simulations at 100 K, three well-defined energy regimes are observed. The

total number of surviving defects in each of these regimes can be well fit with a simple power law. The lowest energy dependence, with an exponent of 0.485, is observed at energies below 0.5 keV, where well-developed cascades do not occur. This low energy exponent may arise from the discrete way in which displacements occur in events near the threshold. Once the first displacement is produced, a greater fraction of the damage energy is lost to lattice phonons, as opposed to atomic displacements, until sufficient energy is available to create the second or third displacement. The angular dependence of the displacement threshold may also be a factor. A low energy exponent would be expected when the first displacement is created in a low threshold direction and the second in some higher threshold direction. Such arguments are in qualitative agreement with the displacement damage functions for copper derived by Merkle et al. [28].

Between 0.5 and 20 keV, a power-law exponent of 0.795 is observed. Subcascade formation begins to have a significant influence on the defect formation for cascade energies above 10 keV. A power-law exponent of 1.12 was obtained for simulations between 20 and 50 keV. Although only a relatively small range of energies could be used to obtain the exponent of 1.12, there are two reasons for believing that the slope change at 20 keV is significant. First, the phenomenon of subcascade formation offers a physical basis for the change. The second reason is statistical; standard errors of the mean values obtained at 10, 20, 30, 40, and 50 keV are quite small. One possible explanation for this higher than linear behavior is the relatively low defect density regions that connect the compact subcasades as shown in Fig. 2. Such regions could exhibit higher defect survival efficiencies because their defect density is similar to relatively low energy cascades, and their contribution could slightly raise the average number of stable displacements.

When the number of surviving defects is divided by the predictions of the NRT model, the effect of subcascade formation is seen in the nearly asymptotic value of this ratio obtained at high energies. For simulations at 100 K, a minimum MD/NRT ratio of 0.301 is found at 20 keV. This ratio slightly increases at higher energies, consistent with the high energy power-law exponent of 1.12 just mentioned. Extrapolation of this ratio to higher energies is somewhat uncertain. Observations of how the cascade/subcascade morphology changes between 10 and

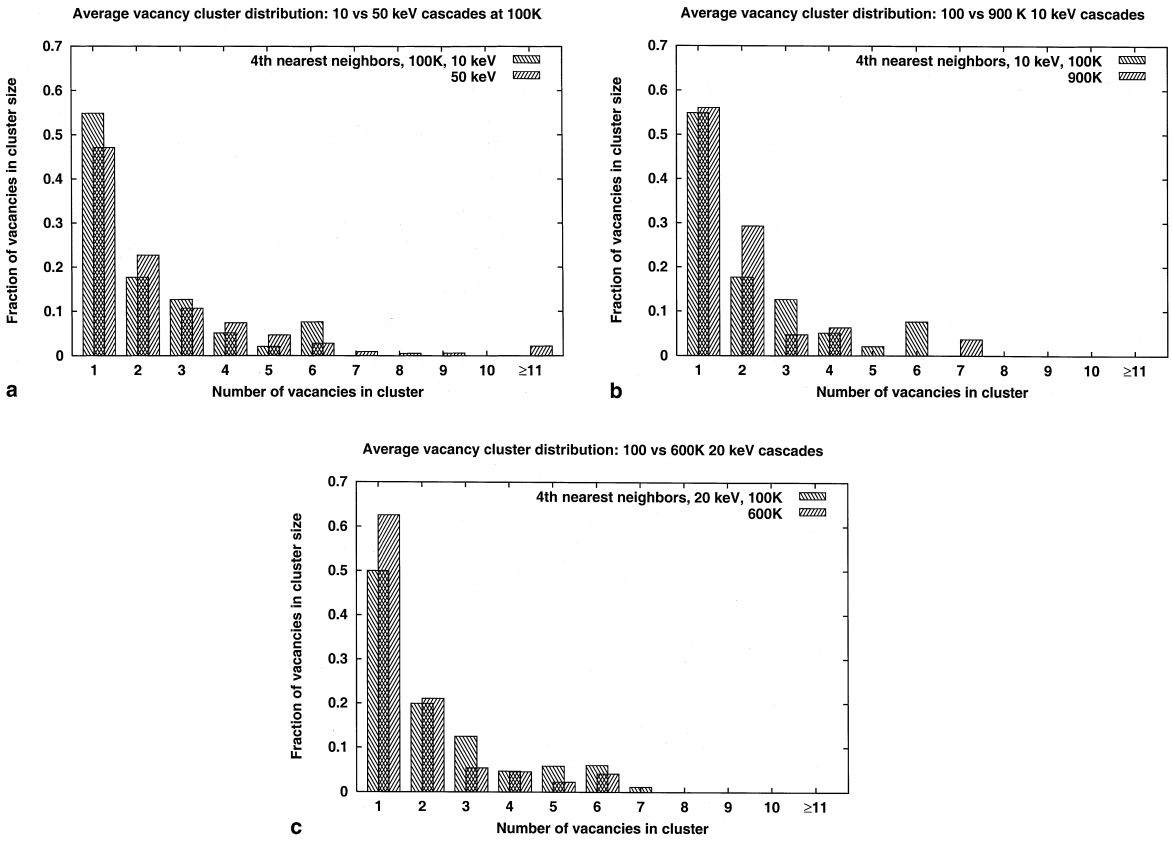


Fig. 9. Fractional size distributions of loosely coupled vacancy clusters (all within 4-nn) formed directly within the cascade, comparison of: (a) 10 and 50 keV cascades at 100 K; (b) 10 keV cascades at 100 and 900 K; (c) 20 keV cascades at 100 and 600 K.

50 keV, and the degree of subcascade formation observed at 50 keV, suggests that higher-energy events should not be qualitatively different. This would imply that the energy dependence should fall to 1.0. However, this conclusion should be verified by at least a limited number of very high energy simulations. In addition, further work is required to confirm that the behavior which is now well documented in 100 K simulations persists at higher irradiation temperatures; some of this work is currently underway.

The effect of subcascade formation is also manifest in the fraction of the defects that form in-cascade clusters. Nearly 60% of the stable, surviving interstitials (~20% of NRT) formed by high energy cascades are found in such clusters for all cascade energies above 1 keV. The largest such cluster observed contained 33 interstitials. Much less in-cascade vacancy cluster formation is observed. Only about 10% of the stable vacancies (~3% of NRT) are clustered if the criterion for defining clusters is vacancies in nearest-neighbor locations. However, analysis of the spatial distribution of vacancies suggests that even the quenched-in cascade vacancies may be correlated out to the 4-nn distance. As confirmed by MC cascade annealing studies, such a collection of vacancies can be considered a nascent vacancy cluster that will coalesce when the vacancies have had time to make a few jumps. Accounting for vacancies out to the 4-nn, the clustering fraction is nearly 50% of the total or 15% of NRT. The vacancy clustering fraction also approaches an apparent asymptotic value at high energies. The effect of irradiation temperature on the in-cascade clustering fractions increases as the cascade energy increases.

The size distribution of both vacancy and interstitial clusters shifts to larger sizes as the cascade energy increases. The nature of the temperature effect depends on the defect type. Increasing temperature leads to a further increase in number of large interstitial clusters, but a decrease in the number of large vacancy clusters. As with the total in-cascade clustering fractions, the effect of temperature on the size distribution increases at higher cascade energies. Because of low homogeneous nucleation rates for extended defect production, in-cascade clustering may play a significant role in defect nucleation in irradiated materials, thereby promoting microstructural evolution and concomitant mechanical property changes. In addition, the formation of larger clusters at higher cascade energies could lead to a greater effect of neutron or PKA energy spectrum than is predicted simply on the basis of total defect survival or the total clustering fraction [20].

Acknowledgements

The author appreciates helpful discussions with Dr S.J. Zinkle of ORNL. Research sponsored by the Office of Nuclear Regulatory Research, US Nuclear Regulatory Commission under interagency agreement DOE 1886-N695-3W with the US Department of Energy, and by the Division of Materials Sciences and the Office of Fusion Energy Sciences, US Department of Energy under contract DE-AC05-84OR21400 with Lockheed Martin Energy Research Corp.

References

[1] R.S. Averback, T. Diaz de la Rubia, R. Benedek, Nucl. Instrum. and Meth. B 33 (1988) 693.
[2] A.J.E. Foreman, W.J. Phythian, C.A. English, Philos. Mag. A 66 (1992) 571.
[3] A.F. Calder, D.J. Bacon, J. Nucl. Mater. 207 (1993) 25.
[4] D.J. Bacon, T. Diaz de la Rubia, J. Nucl. Mater. 216 (1994) 275.
[5] W.J. Phythian, R.E. Stoller, A.J.E. Foreman, A.F. Calder, D.J. Bacon, J. Nucl. Mater. 223 (1995) 245.
[6] R.E. Stoller, in: I.M. Robertson, L.E. Rehn, S.J. Zinkle, W.J. Phythian (Eds.), Proceedings of Symposium on Microstructure of Irradiated Materials, vol. 373, Materials Research Society, Warrendale, PA, 1995, p. 21.
[7] D.J. Bacon, A.F. Calder, F. Gao, V.G. Kapinos, S.J. Wooding, Nucl. Instrum. and Meth. B 102 (1995) 37.
[8] R.E. Stoller, J. Nucl. Mater. 233–237 (1996) 999.
[9] R.E. Stoller, JOM (formerly J. Met.) 48 (1996) 23.
[10] R.E. Stoller, G.R. Odette, B.D. Wirth, J. Nucl. Mater. 251 (1997) 49.
[11] M.J. Norgett, M.T. Robinson, I.M. Torrens, Nucl. Eng. Des. 33 (1975) 50.
[12] ASTM E521, Standard Practice for Neutron Radiation Damage Simulation by Charged-Particle Irradiation, Annual Book of ASTM Standards, vol. 12.02, American Society of Testing and Materials, West Conshohocken, PA.
[13] Yu.N. Osetsky, A. Serra, V. Priego, in: Y. Mishin, G. Vogl, N. Cowern, R. Catlow, D. Farkas (Eds.), Proceedings of Symposium on Diffusion Mechanisms in Crystalline Materials, vol. 527, Materials Research Society, Pittsburgh, PA, 1999, p. 59.
[14] M.W. Finnis, MOLDY6 – A molecular dynamics program for simulation of pure metals, AERE R-13182, AEA Harwell Laboratory, UK (1988).
[15] M.W. Finnis, J.E. Sinclair, Philos. Mag. A 50 (1984) 45.
[16] M.W. Finnis, J.E. Sinclair, Erratum, Philos. Mag. A 53 (1986) 161.
[17] F. Gao, D.J. Bacon, P.E.J. Flewitt, T.A. Lewis, J. Nucl. Mater. 249 (1997) 77.
[18] A. Stuart, J.K. Ord, Kendall's Advanced Theory of Statistics, vol. 1, Griffin, London, 1987.
[19] Manual on Presentation of Data and Chart Control Analysis, sixth edition, American Society for Testing and Materials, West Conshohocken, PA, 1992.
[20] R.E. Stoller, L.R. Greenwood, in: V.V. Butalov, T. Diaz de la Rubia, P. Phillips, E. Kaxiras, N. Ghoniem, (Eds.), Proceedings of Symposium on Multiscale Modeling of Materials, vol. 527, Materials Research Society, Pittsburg, PA, 1999, p. 203.
[21] D.J. Bacon, A.F. Calder, J.M. Harder, S.J. Wooding, J. Nucl. Mater. 205 (1993) 52.

[22] P. Jung, Phys. Rev. B 23 (1981) 664.
[23] G. Wallner, M.S. Anand, L.R. Greenwood, M.A. Kirk, W. Mansel, W. Waschkowski, J. Nucl. Mater. 152 (1988) 146.
[24] B.D. Wirth, PhD dissertation, University of California, Santa Barbara, 1998.
[25] S. Sato, A. Kohyama, N. Igata, Appl. Surf. Sci. 76&77 (1994) 285.
[26] B.D. Wirth, G.R. Odette, in: V.V. Butalov, T. Diaz de la Rubia, P. Phillips, E. Kaxiras, N. Ghoniem (Eds.), Proceedings of Symposium on Multiscale Modeling of Materials, vol. 527, Materials Research Society, Pittsburgh, PA, 1999, p. 211.
[27] M. Valo, R. Krause, K. Saarinen, P. Hautojarvi, J.R. Hawthorne, in: R.E. Stoller, A.S. Kumar, D.S. Gelles (Eds.), Effects of Radiation on Materials, ASTM STP 1125, American Society for Testing and Materials, West Conshohocken, PA, 1992, p. 172.
[28] K.L. Merkle, W.E. King, A.C. Bailey, K. Haga, M. Meshii, J. Nucl. Mater. 117 (1983) 4.

ELSEVIER

Nuclear Engineering and Design 195 (2000) 129–136

Nuclear Engineering and Design

www.elsevier.com/locate/nucengdes

Evaluation of neutron energy spectrum effects and RPV thru-wall attenuation based on molecular dynamics cascade simulations

R.E. Stoller [a,*]

[a] *Metals and Ceramics Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831-6376, USA*

Accepted 16 September 1999

Abstract

Displacement cascade formation in iron has been investigated by the method of molecular dynamics (MD) for cascade energies up to 40 keV. The results of these simulations have been used in the SPECOMP code to obtain effective, energy-dependent cross sections for two measures of primary damage production: (1) the number of surviving point defects expressed as a fraction of the those predicted by the standard secondary displacement model by Norgett et al., 1975 [M.J. Norgett, M.T. Robinson, and I.M. Torrens, Nucl. Eng. Des. 33, (1975) 50–54] and (2) the fraction of the surviving interstitials contained in clusters that formed during the cascade event. The primary knock-on atom (PKA) spectra for iron obtained from the SPECTER code have been used to weight these MD-based damage production cross sections in order to obtain spectrally-averaged values for several locations in commercial fission reactors and test reactors. An evaluation of these results indicates that neutron energy spectrum differences between the various environments do not lead to significant differences between the average primary damage formation parameters. In particular, spectrum-averaged defect production cross sections obtained for PWR and BWR neutron spectra were not significantly different. A representative application of the new defect production cross sections is provided by examining how they vary as a function of depth into the reactor pressure vessel wall. A slight difference was noted between the damage attenuation in a PWR vessel and a BWR vessel. This observation could be explained by a subtle difference in the energy dependence of the neutron spectra. Overall, the simulations indicate that spectrum-averaged defect production cross sections do not vary much among the various environments in light-water moderated fission reactors. As such, the results support the use of atomic displacements per atom (dpa) as a damage correlation parameter and provide guidance for choosing the primary damage source term in kinetic embrittlement models. © 2000 Elsevier Science S.A. All rights reserved.

1. Introduction

Radiation-induced embrittlement in reactor pressure vessel (RPV) steels has traditionally been correlated with exposure parameters such as the

* Tel.: +1-423-576-788; fax: +1-423-574-0641.
E-mail address: rkn@ornl.gov (R.E. Stoller)

0029-5493/00/$ - see front matter © 2000 Elsevier Science S.A. All rights reserved.
PII: S0029-5493(99)00241-1

neutron fluence above 1 MeV or atomic displacements per atom (dpa) (IAEA, 1975; Stoller and Odette, 1992; ASTM, E693, 521). One advantage of dpa over fast fluence is that it explicitly accounts for atomic displacements produced by the entire neutron energy spectrum. Therefore, it provides improved correlation of data obtained from irradiation facilities with different neutron energy spectra. Differences in neutron energy spectra are manifested as differences in the energy spectra of the primary knock-on atoms (PKA) that are produced in elastic collisions with these neutrons. Low energy PKA, such as those produced by thermal neutrons, are somewhat more efficient at net defect production than are high-energy PKA. This suggests that these displacements should be more heavily weighted in dosimetry. Conversely, the lower energy PKA produce fewer point defect clusters that can promote the formation of extended defects responsible for mechanical property changes.

The issue of PKA energy effects can be addressed through the use of displacement cascade simulations using the method of molecular dynamics (MD). Although MD simulations can provide a detailed picture of the formation and evolution of displacement cascades, they impose a substantial computational burden. However, recent advances in computing equipment permit the simulation of high energy displacement events involving more than one-million atoms (Stoller et al., 1996; Stoller, 1996a,b); the results presented below will encompass MD cascade simulation energies from near the displacement threshold to as high as 40 keV. The computing time with the molecular dynamics (MOLDY) code is almost linearly proportional to the number of atoms in the simulation and higher energy events require a larger atom block as listed in Table 1. Two to three weeks of cpu time is required to complete the highest energy 40 keV cascade simulations with 1 024 000 atoms for 15 ps on a modern high-speed workstation.

Two parameters have been extracted from the MD simulations: the number of point defects that remain after the displacement event is completed and the fraction of the surviving interstitials that are contained in clusters. For the purpose of comparison with standard dosimetry, these values have been normalized to the number of atomic displacements calculated with the secondary displacement model by Norgett et al. (1975). The energy dependence of the two MD defect parameters was used to evaluate the effects of neutron energy spectrum. Simple, energy-dependent functional fits to the MD results were obtained and the SPECOMP code (Greenwood, 1989) was used to compute effective cross sections for point defect survival and point defect clustering. PKA spectra for iron obtained from SPECTER (Greenwood and Smither, 1985) were then used to weight these effective cross sections in order to calculate spectrum-averaged values for various neutron irradiation environments. These include several locations through the wall of RPVs of representative commercial pressurized and boiling water reactors

Table 1
Typical MD cascade parameters and required atom block sizes

Neutron energy (MeV)	Average PKA energy (keV)	Corresponding E_{MD} (keV)	NRT displacements	Atoms in simulation
0.0034	0.116	0.1	1	3456
0.0058	0.236	0.2	2	6750
0.014	0.605	0.5	5	6750
0.036	1.24	1.0	10	54 000
0.074	2.54	2.0	20	54 000
0.19	6.60	5.0	50	128 000
0.40	13.7	10.0	100	250 000
0.83	28.8	20.0	200	250 000
1.8	61.3	40.0	400	1 024 000

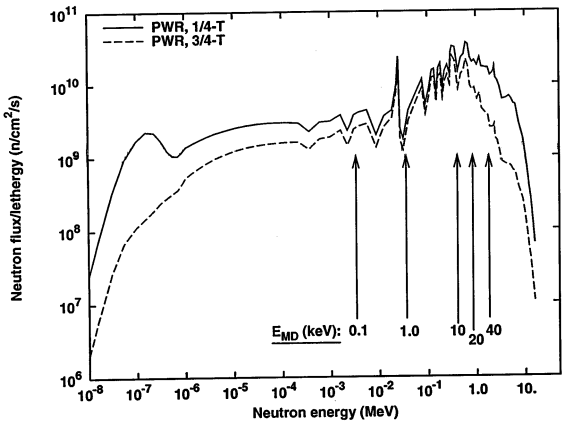


Fig. 1. Representative PWR neutron spectra at the 1/4-T and 3/4-T RPV locations and the corresponding average cascade energy.

(PWR and BWR) and positions in both water and sodium-cooled materials test reactors.

2. MD cascade simulations

The molecular dynamics code, MOLDY, and the interatomic potential for iron that was used are described in detail in Finnis and Sinclair (1984, 1986), Finnis (1988), Calder and Bacon (1993). Additional details on the results obtained by this method can be found in Phythian et al. (1995), Stoller et al. (1996), Stoller (1996a,b). Briefly described, the process of conducting a cascade simulation requires two steps. First, a block of atoms of the desired size is thermally equilibrated, this process permits the lattice thermal vibrations (phonon waves) to be established for the simulated temperature, and typically requires a simulation: equivalent to approximately 10 ps. This atom block can be saved and used as the starting point for several subsequent cascade simulations. Then, the cascade simulations are initiated by giving one of the atoms a defined amount of kinetic energy. E_{MD}, in a specified direction. This atom is equivalent to the PKA following a collision with a neutron. Statistical variability can be introduced by either further equilibration of the starting block or by choosing either a different primary knock-on atom or PKA direction. Typically, at least six different cascades are required to obtain statistically representative average parameters at any one cascade energy and temperature.

The MOLDY code describes only elastic collisions between atoms; it does not account for energy loss mechanisms such as electronic excitation and ionization. Thus, the initial energy E_{MD} given to the simulated MD PKA is analogous to the damage energy (T_{dam}) in the NRT model (Norgett et al., 1975). Using the values of E_{MD} in Table 1 the corresponding E_{PKA} and the NRT defects in iron have been calculated using the procedure described in Norgett et al., (1975) with the recommended 40 eV displacement threshold (ASTM, E521). These values are also listed in Table 1. Note that the difference between the MD simulation, or damage energy and the PKA energy increases as the PKA energy increases.

The cascade simulations are continued until in-cascade recombination of vacancies and interstitials is complete and the atom block has returned to near thermal equilibrium. The required simulation time varies from about 5 ps for the low-energy cascades to 15–20 ps for the 40 keV cascades at 100 K. The range of neutron energies covered by these simulations is also listed in Table 1 and illustrated in Fig. 1. The neutron spectra obtained at the 1/4-T and 3/4-T RPV positions for a typical PWR are shown and the corresponding MD simulation energies are indicated.

Two parameters are of primary interest to this work: the number of point defects that survive after in-cascade recombination is complete and the fraction of the surviving interstitials contained in clusters rather than as isolated defects. The former is important because it is only the surviving point defects that can contribute to radiation-induced microstructural evolution. The latter is significant because these small clusters provide nuclei for the growth of larger defects which can give rise to mechanical property changes. The formation of these small clusters directly within the cascade means that the extended defects can evolve more quickly than if the clusters could only be formed by the much slower process of classical nucleation. For purposes of this work, interstitials were considered clustered if they were within the

nearest-neighbor lattice distance of another interstitial. The size distribution of the interstitial clusters produced is a function of the cascade energy (Phythian et al., 1995) and the interstitial clustering fraction is calculated by summing that distribution.

The surviving MD defects can be conveniently described as a fraction of the NRT displacements (Norgett et al., 1975) and the number of clustered interstitials as a fraction of the surviving MD defects. The energy dependence of the surviving defect fraction (η) and the interstitial clustering fraction (f_{icl}) from the MD simulations is shown in Figs. 2 and 3, respectively. Since the MD results did not exhibit a strong dependence on irradiation temperature, all of the results obtained at 100, 600, and 900 K are shown (Phythian et al., 1995; Stoller et al., 1996; Stoller, 1996a,b). The line drawn through the data in each figure is a nonlinear least-squares fit to the data using the following functions:

$$\eta = 0.5608 \cdot E_{MD}^{-0.3029} + 3.227 \times 10^{-3} \cdot E_{MD} \quad (1)$$

$$f_{icl} = [0.097 \cdot \ln(E_{MD} + 0.9)]^{0.3859} - 7 \times 10^{-6} \cdot E_{MD}^{2.5} \quad (2)$$

where E_{MD} is in keV.

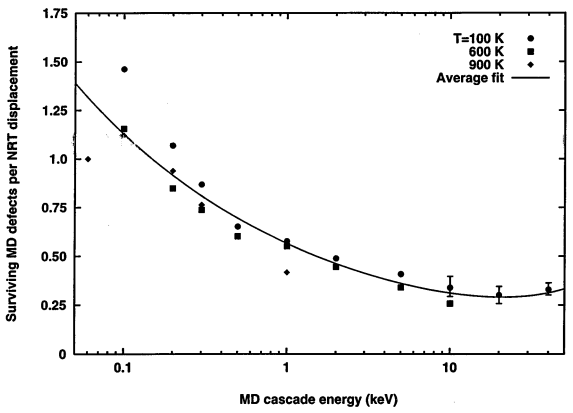


Fig. 2. Average MD point defect survival fraction as a function of cascade energy; results of MD simulations at 100, 600 and 900 K.

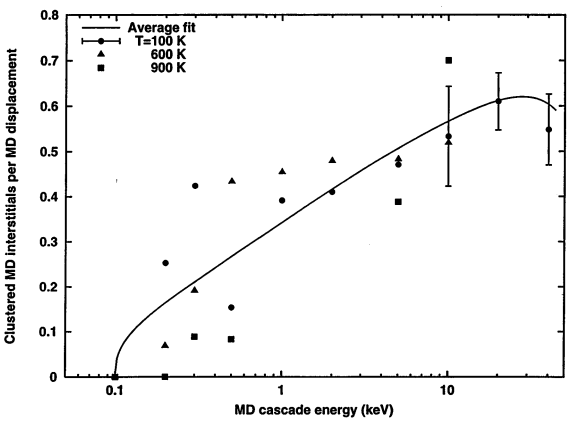


Fig. 3. Average in-cascade interstitial clustering fraction as a function of cascade energy; results of MD simulations at 100, 600 and 900 K.

In both Eqs. (1) and (2), the first term in the function dominates the energy dependence up to about 20 keV. The second term accounts for the effect of subcascade formation at the highest energies (Stoller et al., 1996; Stoller, 1996a). Mathematically, this term is responsible for the minimum in the defect survival curve and the maximum in the interstitial clustering curve at about 20 keV. This change in the energy dependence occurs because subcascade formation makes a high single energy cascade appear to be the equivalent of several lower energy cascades. Thus, the defect survival fraction is slightly higher and the interstitial clustering fraction slightly lower at 40 than at 20 keV. The phenomenon of subcascade formation is illustrated in Fig. 4, in which typical 10, 20, and 40 keV cascades are shown.

The increase in the defect survival fraction between 20 and 40 keV is slight, but it appears to be statistically significant from the magnitude of the standard deviations that are shown as error bars in Fig. 2. Since the standard deviations on the average interstitial clustering fractions are much larger, the significance of the maximum shown in Fig. 3 is less clear. No cascades with energies higher than 40 keV have been completed at this time. However, based on the degree of subcascade formation observed in the 40 keV cascade simulations, it appears unlikely that η and f_{icl} will change significantly at higher cascade energies. Therefore, the values calculated from Eqs. (1) and (2) for 40 keV were applied for all the higher energy PKA in the SPECOMP and SPECTER calculations.

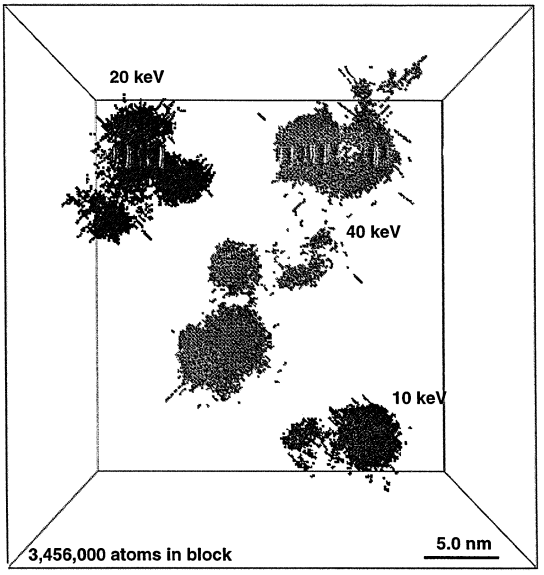


Fig. 4. Illustration of increasing subcascade formation in iron at 100 K as the MD cascade energy increases from 10 to 40 keV.

3. Defect production calculations using SPECOMP and SPECTER

Energy-dependent defect production cross sections for surviving MD defects and clustered interstitials and spectrum-averaged values for several irradiation environments were generated by modifying the SPECOMP (Greenwood, 1989) and SPECTER (Greenwood and Smither, 1985) computer codes. SPECOMP normally calculates displacement cross sections for compounds using the primary knock-on atomic recoil energy distributions contained in a 100-neutron-energy by 100-recoil-energy grid for each of 40 different elements. For the present defect and clustered interstitial calculations, these new functions were used as a factor multiplying the standard displacement cross section equations as a function of the damage energy, T_{dam}. SPECOMP thus produced surviving defect and clustered interstitial cross sections on a 100 point neutron energy grid.

The SPECTER computer code contains libraries of calculated cross sections for displacements, gas production and total energy distribution, as well as atomic recoil energy distributions for more than 40 elements and various compounds. For a given neutron energy spectrum and irradiation time, the code can be used to calculate the net radiation damage effects, as listed above. In the present case, the SPECOMP calculations for the surviving defect and clustered interstitial cross sections were added to the SPECTER libraries. SPECTER runs for various neutron spectra thereby producing spectrum-averaged values for the point defect and interstitial clustering fractions.

4. Results and discussion

The primary results of these calculations are summarized in Fig. 5. The PKA-spectrum-averaged defect survival fraction is shown in Fig. 5a and the interstitial clustering fraction in Fig. 5b. In both cases, the effective production cross sec-

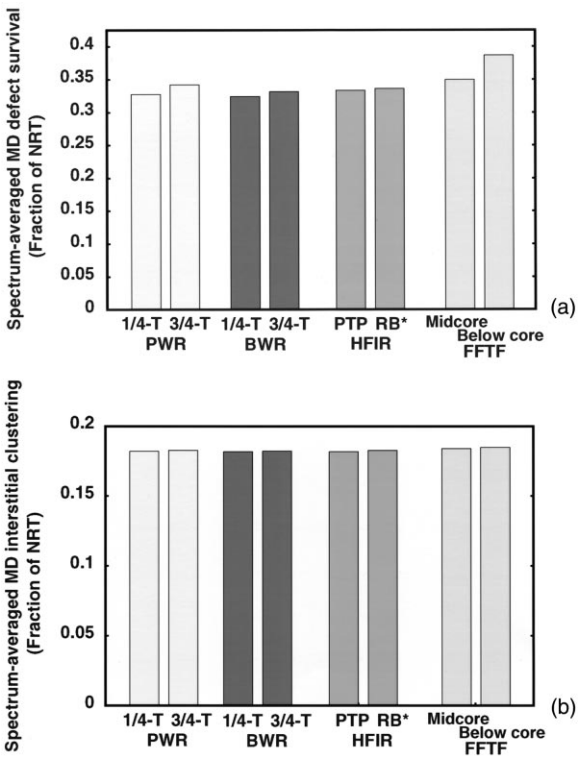


Fig. 5. Comparison of spectrally-averaged damage production cross sections (per NRT dpa) for various irradiation environments; defect survival ratio is shown in (a) and the interstitial clustering fraction is shown in (b).

tion has been divided by the NRT dpa cross section. Values are shown for the 1/4-T and 3/4-T RPV positions from representative PWR and BWR neutron spectra. Four additional irradiation sites are also shown. These are: positions located in the peripheral target position (PTP) and removable beryllium reflector (RB*) of the high flux isotope reactor (HFIR) at ORNL and the mid-core and below-core (BC) positions in the fast flux test facility (FFTF) at the US DOE Hanford Reservation.

Including the HFIR and FFTF positions provides a broad comparison of both hard and soft neutron spectra. The former is a very high-flux (up to $\sim 10^{15}$ n cm^{-2}), water-moderated test reactor and the latter is a liquid metal (Na) cooled fast reactor. In spite of the differences between the neutron energy spectra, only relatively small differences are observed in the two spectrum-averaged damage cross sections. This is consistent with the normalized iron PKA spectra shown in Fig. 6. In the region where the PKA probability is highest, i.e. for PKA energies between 1 keV and 0.1 MeV, the spectra are quite similar.

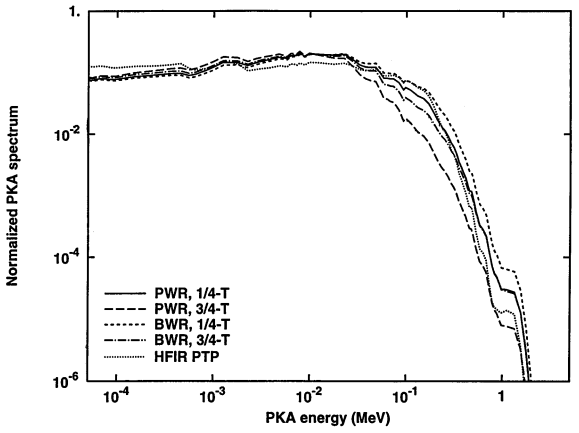


Fig. 6. Normalized iron PKA spectra for PWR and BWR 1/4-T and 3/4-T positions and the HFIR PTP position.

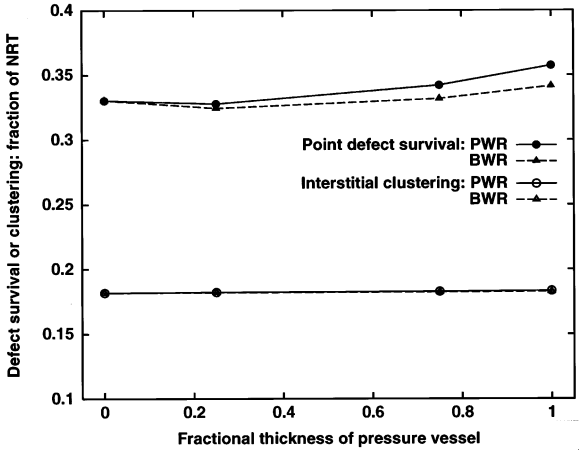


Fig. 7. Variation of spectrally-averaged damage production cross sections (per NRT dpa) through the RPV for PWR and BWR.

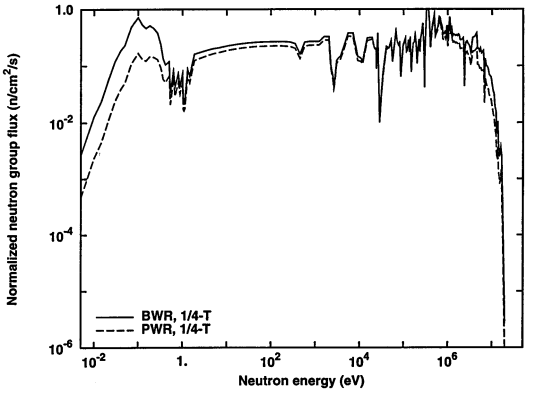


Fig. 8. Comparison of normalized neutron spectra for PWR and BWR at the 1-4-T RPV location. Note that the BWR has both higher relative thermal and fast fluxes.

A slight difference between the PWR and BWR spectra can be seen in the way the spectra change as a function of thickness through the pressure vessel. This is illustrated in Fig. 7, where the spectrum-averaged defect production cross sections are shown for four positions: the last water node point before the RPV, the 1/4 and 3/4 RPV locations and the first node point in the cavity beyond the RPV. The values for the two reactor types are most similar at the pressure vessel wall and diverge somewhat at greater depths. The reason for this modest divergence is a subtle difference in the neutron energy spectra and the way in which neutron attenuation occurs. The BWR spectrum is generally considered to be softer than the PWR and would be expected to give a higher defect survival fraction based on the results shown in Fig. 2. However, the average PKA energy at the 1/4-T and deeper positions is actually some-

what higher for the BWR. This can be rationalized by a comparison of the1/4-T neutron spectra in Fig. 8, where the energy-dependent fluxes arc normalized using their respective peak values. The BWR spectrum shows both a higher relative thermal flux and a higher relative fast flux than the PWR spectrum. Since most of the displacements are generated by the higher energy neutrons, the effective average energy for the BWR spectrum is greater than that of the PWR.

5. Summary

Although it is not yet possible to simulate the highest energy displacement cascades generated in the materials used in fission reactors, the analysis of MD cascade simulations in iron for energies up to 40 keV provides considerable insight. The primary damage parameters derived from the MD results exhibit a strong dependence on cascade energy up to 10 keV; however, this dependence is diminished and slightly reversed between 20 and 40 keV. This reversal is due to the formation of well defined subcascades in this energy region. Analysis of the cascades indicates that little further change should occur at higher energies, which suggests that the results reported here should be relevant to an evaluation of neutron energy spectrum effects.

Notably, the spectrum-averaged defect production cress sections calculated for several fission reactor neutron spectra were all quite similar. This included locations within the pressure vessels of commercial PWRs and BWRs, as well as locations in light water and liquid metal cooled materials test reactors. As a result of the degree of similarity between the cross sections being so high, it appears unlikely that spectral differences will significantly influence comparisons of PWR and BWR data In addition, since the effective damage production cross sections could be described by similar fractions of the NRT displacements, these results support the use dpa as a damage correlation parameter. Finally, the spectrum-averaged defect production cross sections obtained from the MD results can be used to guide the selection of appropriate radiation damage source terms in the kinetic models used to investigate radiation-induced microstructural evolution and embrittlement.

Acknowledgements

The assistance of Dr L.R. Greenwood from the Materials and Chemistry Division of the Pacific Northwest National Laboratory in carrying out the SPECTER and SPECOMP calculations is gratefully acknowledged. Research sponsored by the Office of Nuclear Regulatory Research, US Nuclear Regulatory Commission under inter-agency agreement DOE 1886-8199-8L with the US Department of Energy and by the Division of Materials Sciences, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

References

ASTM E521, 1999, Standard practice for neutron radiation damage simulation by charged-particle irradiation, Annual Book of ASTM Standards, vol. 12.02, American Society of Testing and Materials, Philadelphia.

ASTM E693, 1999, Standard practice for characterizing neutron exposures in iron and low alloy steels in terms of displacements per atom (dpa), Annual Book of ASTM Standards, vol. 12.02, American Society of Testing and Materials, Philadelphia.

Calder, A.F., Bacon, D.J., 1993. J. Nucl. Mater. 207, 25–45.

Finnis, M.W., 1988. MOLDY6-A: Molecular Dynamics Program for Simulation of Pure Metals; AERE R-13182, UK AEA. Harwell Laboratory.

Finnis, M.W., Sinclair, J.E, 1984. Phil. Mag. A50, 45–55.

Finnis, M.W., Sinclair, J.E, 1986. Erratum. Phil. Mag. A53, 161.

Greenwood, L.R., 1989. SPECOMP: Calculations of Radiation Damage in Compounds. In: Farrar, H., Lippincott, E.P. (Eds.), Reactor dosimetry: methods, applications and standardization; ASTM STP 1001. American Society of Testing and Materials, West Conshohocken, PA, pp. 598–602.

Greenwood, L.R., Smither, R.K., January, 1985. SPECTER: Neutron Damage Calculations for Materials Irradiations: ANL/FPP/TM-197. Argonne National Laboratory, Argonne, IL. Recommendations of IAEA specialists meeting

on radiation damage units in graphite and ferritic and austenitic steels. Nucl. Eng. Des. 33, (1975), 48.

Norgett, M.J., Robinson, M.T., Torrens, I.M., 1975. Nucl. Eng. Des. 33, 50–54.

Phythian, W.J., Stoller, R.E., Foreman, A.J.E., Calder, A.F., Bacon, D.J., 1995. J. Nucl. Mater. 223, 245–261.

Stoller, R.E., 1996a. J. Met. 48, 23–27.

Stoller, R.E., 1996b. J. Nucl. Mater. 233–237, 999–1003.

Stoller, R.E., Odette, G.R., 1992. J. Nucl. Mater. 186, 203–205.

Stoller, R.E., Odette, G.R., Wirth, B.D., 1996. Primary defect formation in bcc iron: presented at the International Workshop on defect production, accumulation, and materials performance in irradiation environments, Davos, Switzerland, October 2–8, 1996. Accepted for publication in J. Nuclear Mater. 251 (1997) 49–60.

ERKCM30 (Laves Phase Intermetallics)

Refereed Journal Publications

*Hong, S., C. L. Fu, and M. H. Yoo, "First-Principles Calculation of Stacking Fault and Twin Boundary Energies of Cr2Nb, "Phil. Mag. A 80, 871 (2000).

†*Kuman, K. S., L. Pang, C. T. Liu, J. A. Horton, and E. Kenik, "Structural Stability of the Laves Phase Cr2Ta in a Two-phase Cr-Cr2Ta Alloy," Acta Metall. 48, 911-23 (2000).

*Liu, C. T., J. H. Zhu, M. B. Brady, C. G. McKamey, and L. M. Pike, "Physical Metallurgy and Mechanical Properties of Transition-metal Laves Phase Alloys," Intermetallics 8, 1119-29 (2000).

*Zhu, J. H., and C. T. Liu, "Defect Structures in ZrCo2 Laves Phase," Acta Mater. 48, 2339-47 (2000).

*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

PHILOSOPHICAL MAGAZINE A, 2000, VOL. 80, NO. 4, 871–880

First-principles calculation of stacking fault and twin boundary energies of Cr_2Nb

SUKLYUN HONG†, C. L. FU and M. H. YOO

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6114, USA

[*Received 15 July 1998 and accepted 7 June 1999*]

ABSTRACT

The stacking fault and twin boundary energies of C15 Cr_2Nb are calculated by the first-principles local-density-functional approach. It is found that the intrinsic and extrinsic stacking fault energies are 116 and 94 $mJ\,m^{-2}$, respectively, and the twin boundary energy is 39 $mJ\,m^{-2}$. The lower extrinsic stacking fault energy is consistent with the fact that the C36 structure has a lower energy than the C14 structure. The calculated stacking fault energies at 0 K are larger than the experimental values available in the literature. The equilibrium separations between Shockley partials based on the calculated elastic constants and stacking fault energies are also calculated.

§ 1. INTRODUCTION

For a wide variety of structural applications, Laves phases have some unique properties such as high melting temperature, low density, and high oxidation resistance. Unfortunately, this potential has not been well exploited, largely because of low temperature brittleness due to the lack of plastic deformation. Cr_2Nb, one of the most studied Laves phase compounds, has either the cubic C15 or hexagonal C14 (or C36) structure, which is topologically akin to the face-centred cubic (fcc) (A1) or the hexagonal close-packed (hcp) (A3) structure, or any one of the polytypic phases. The topologically close-packed (TCP) plane of the C15 structure is of the $\{111\}$ type, and the potential modes of plastic deformation in Cr_2Nb are twelve $\{111\}\langle 110\rangle$ slip systems and twelve $\{111\}\langle 112\rangle$ twin systems. Since each of the TCP units consists of the quadruple atomic layers, the slip, twinning, or stress-induced polytypic transformation will require a coordinated process of atomic motions, such as synchroshear, in order to effect the motion of Shockley partial dislocations.

Among other quantities, the knowledge of stacking fault energy (SFE), γ_{SF}, is necessary to understand the deformation mechanism, since SFE will play an important role in processes such as dislocation dissociation, cross-slip and twinning. First-principles calculations are useful to estimate the energetics of stacking faults (SFs) and understand the interaction between Shockley partials.

Using the linear muffin-tin orbital (LMTO) method, Chu *et al.* (1995a) estimated the intrinsic stacking fault energy of Cr_2Nb from the structural energy difference between the C15 and C14 structures, and obtained $\gamma_{SF} = 90\,mJ\,m^{-2}$. On the other

† Present address: Department of Physics, Sejong University, Seoul 143-747, Korea.

Philosophical Magazine A ISSN 0141–8610 print/ISSN 1460-6992 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journ als/tf/01418610.html

hand, Yoshida *et al.* (1995) obtained a value of $8\,\mathrm{mJ\,m^{-2}}$ for SFE from transmission electron microscopy (TEM) observations of extended dislocation nodes at 1623°C. More recently, Kazantzis *et al.* (1996) obtained a value of $25\,\mathrm{mJ\,m^{-2}}$ from TEM observations of extended nodes in Cr_2Nb at 1400 and 1500°C.

In a previous paper (Hong and Fu 1999), we investigated the phase stability of three Laves phases (C15, C14 and C36) of Cr_2Nb. It was found that the C15 phase is the ground-state structure with the lowest energy and the C36 phase is an intermediate state between C15 and C14. These three phases, however, are very close in energy, i.e. within a range of about 60 meV/formula unit (Hong and Fu 1999) indicating the possibility of low stacking fault energies in this system. In this paper, we report the calculation of SFE using supercell geometry, and an evaluation of the equilibrium separation between Shockley partials using the anisotropic elastic theory.

§ 2. Laves structures

In describing the structures of Laves phases and their stacking faults and twin boundary, we follow the notations of Hazzledine (1994).

Laves phases have ideal chemical compositions S_2L; they contain smaller atoms S and larger atoms L in alternate sheets parallel to the TCP planes (i.e. (111) plane for the C15 structure and (0001) plane for the C14 and C36 structures). The main geometric characteristic of Laves phases S_2L is that they consist of two types of atomic stacking sequence, $\alpha A\alpha$ ($\beta B\beta$ and $\gamma C\gamma$) and $\alpha c\beta$ ($\beta a\gamma$ and $\gamma b\alpha$). Here, the Greek letters (α, β, γ) denote the L atoms, while lower case Latin letters (a, b, c) and capital letters (A, B, C) denote the type-1 and type-2 S atoms, respectively. Note that the $\alpha c\beta$-type stackings are more closely spaced (in terms of the interlayer spacings). For Cr_2Nb, Latin and Greek letters represent Cr and Nb, respectively.

Figure 1 represents an atomic layer of A atoms based on the hexagonal unit cell on the TCP plane (a Kagome net) which is determined by two lattice vectors

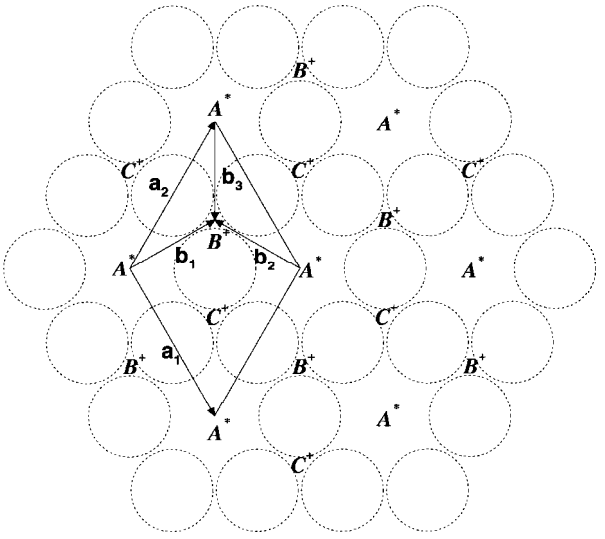


Figure 1. The hexagonal unit cell on the basal planes which is determined by two lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. See the text for details.

$\mathbf{a}_1 = \frac{1}{2}[0\bar{1}1]$ and $\mathbf{a}_2 = \frac{1}{2}[\bar{1}10]$. The vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are Shockley partial Burgers vectors of the type $\frac{1}{6}\langle\bar{1}12\rangle(\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 0)$. The dotted circles in figure 1 denote A atoms, one of the type-2 S atoms, projected on the (111) plane. Let the centres of hexagonal rings consisting of A atoms be denoted by A^*, and thus the A atoms are placed midway between the lattice points A^*. Similarly, the lattice points B^* and C^*, which determine B and C atoms respectively, can be obtained by the displacement of A^* by $\mathbf{b}_1$ and $-\mathbf{b}_1$, respectively, on higher or lower planes. Also, we denote the projected points of B^* and C^*onto the plane given in figure 1 as B^+ and C^+, respectively, which are the saddle points for larger α atoms. Then, the projected atomic positions of α (or a) onto the given plane are at the sites of A^* in figure 1, and the projected positions of β (or b) and γ (or c) are at those of B^+ and C^+, respectively. Consequently, the TCP C15 structure is defined by repeated XYZ stacking:

$$\text{C15}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{1}$$

where the TCP plane unit X (Y or Z) consists of one single layer of A (B or C) atoms and one triple layer of αcβ (βaγ or γbα). A synchroshear mechanism was first introduced by Kronberg (1957), and was used to explain the deformation twinning process by several groups (Livingston and Hall 1990, Chu and Pope 1993, Hazzledine 1994). Synchroshearing of the X unit (Aαcβ) creates the X′ unit (Aαbγ), and similarly for the Y′ (Bβcα) and Z′ (Cγaβ) units. The C14 and C36 Laves phases have repeated X′Z and XYZ′Y′ stackings, respectively, as follows:

$$\text{C14}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{b}\gamma}^{\mathbf{X'}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{2}$$

$$\text{C36}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{a}\beta}^{\mathbf{Z'}}\,\overbrace{\mathrm{B}\beta\mathrm{c}\alpha}^{\mathbf{Y'}}\cdots. \tag{3}$$

It can be seen that the primed X′, Y′, Z′ units are introduced by synchroshear in order to make sure that TCP structure is maintained in C14 and C36. For example, c and β of the X unit in C15 are synchrosheared by $-\mathbf{b}_2$ and $-\mathbf{b}_1$, respectively, to become b and γ of X′ unit. This displacement preserves a close packing between X and Z units in C14.

The stacking sequences for an intrinsic stacking fault (ISF) and an extrinsic stacking fault (ESF) and a twin boundary are given in table 1 in terms of X, Y, Z units. Note that the ISF contains a local C14-like structure in C15, while the ESF contains a local C36-like structure in C15. Figure 2 shows schematic illustrations of an ISF and an ESF obtained from C15 through the synchroshear mechanism: (*a*) C15 → ISF, and (*b*) C15 → ESF. The C15 stacking has a repeated XYZ sequence (no stacking fault). For the ISF, one unit, namely Y, is missing from the C15 XYZ sequence, and X becomes X′ by synchroshear (shifts by $-\mathbf{b}_2$ and $-\mathbf{b}_1$ shown in figure 2 (*a*)), while for an ESF a unit Y′ is added into the original C15 sequence. For the ESF, two successive synchroshears (two synchro-shifts shown in figure 2 (*b*)) are operative in the αcβ type to maintain a TCP structure by introducing two primed units, Z′ and Y′. Twinned structure is equivalent to a mirror reflection about the X′ unit (see table 1).

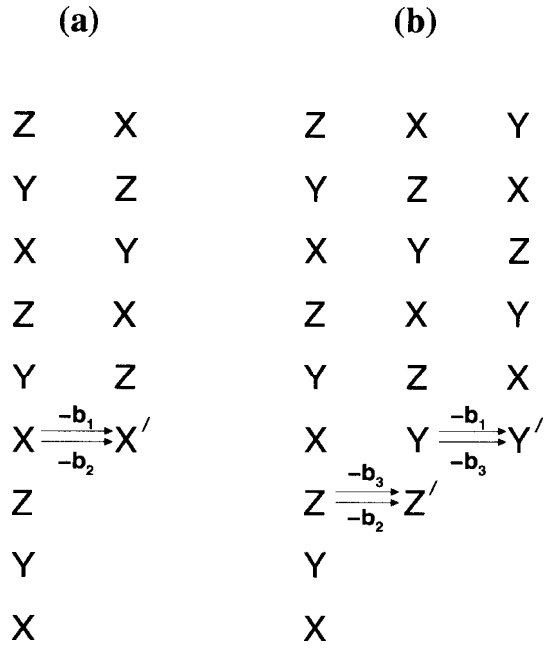


Figure 2. Schematic illustrations of an ISF and an ESF obtained from C15 → ISF, and (*b*) C15 → ESF.

Table 1. Stacking sequences of the C15 structure (no stacking fault), intrinsic stacking fault, extrinsic stacking fault, and twin (C15 + C15T) structure.

	Sequence of faults
C15	···XYZXYZXYZ···
Intrinsic stacking fault	···XYZX′ZXYZ···
Extrinsic stacking fault	···XYZXYZ′Y′XYZ···
Twin (C15 + C15T)	···XYZX′Z′Y′X′Z′Y′···

§ 3. Stacking fault and twin boundary energies

Total-energy calculations for the stacking fault energies of C15 Cr_2Nb are performed using the full-potential linearized augmented plane-wave (FLAPW) method (Wimmer *et al.* 1981) within the local-density approximation. The FLAPW method solves the local-density-functional equations without any shape approximation to the potential or charge density. The atomic positions are relaxed by calculating Hellmann–Feynman forces acting on the atoms.

The supercell geometry is used to obtain the energies of ISF, ESF and twin boundary. In the supercell calculation, we use the experimental lattice constant (6.991 Å) of C15 Cr_2Nb and spacings along $\mathbf{a}_3$ axis corresponding to the ideal hcp c/a ratio, which gives ~4. 04 Å for the thickness of each X, Y, Z unit. We consider a supercell containing XYZX′Z for an ISF and XYZXYZ′Y′ for an ESF. In these supercells, the separations between ISF and ESF planes are about 20 Å and 28 Å,

respectively. Also, a supercell containing XYZX′Z′Y′ (~24 Å) is considered for twin boundary. The supercell for twinning contains two twin boundaries in it, thus the separation between twin boundaries is about 12 Å.

The internal coordinates of each structure are fully relaxed from the ideal positions (defined by the atomic positions of the C15 lattice) by calculating Hellmann–Feynman forces acting on the atoms. The relaxation energies of ISF, ESF, and twin boundary are 41, 20, and 17 mJ m^{-2}, respectively.

The results for the stacking fault and twin boundary energies are given in table 2. It is found that the fault energies are γ_{ISF} = 116 and γ_{ESF} = 94 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is γ_T = 39 mJ m^{-2}. Since ESFs and ISFs contain local C36-like and C14-like structures, respectively, a lower ESF energy (γ_{ESF}) compared to the ISF energy (γ_{ISF}) is consistent with the fact that the C36 structure has a lower energy than the C14 structure (Hong and Fu 1999).

For comparison, we also calculated the stacking fault and twin boundary energies using the theoretical lattice constant (6.822 Å). The relaxation energies of ISF, ESF, and twin boundary are 48, 40, and 21 mJ m^{-2}, respectively. It is also found that the fault energies are 140 and 108 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is 52 mJ m^{-2}. The results using the theoretical lattice constant are slightly larger than those using the experimental one.

Although the intrinsic faults are expected to prevail in fcc crystals rather than the extrinsic faults (Hirth and Lothe 1982), extrinsic stacking faults were observed to be dominant in Laves phase Co_2Ti (Allen *et al.* 1972), and Nb-doped HfV_2 (Chu *et al.* 1998). Since the calculated γ_{ESF} is smaller than γ_{ISF}, it is likely that the observed SFs in Cr_2Nb are also of extrinsic type.

We compare our results with others. Chu *et al.* (1995a) obtained γ_{ISF} = 90 mJ m^{-2} by an estimation from the structural energy difference between the bulk C14 and C15 structures. This value is close to our value of 116 mJ m^{-2}. On the other hand, Yoshida *et al.* (1995) observed extended dislocation nodes in C15 Cr_2Nb deformed at 1623 K, in which the SFs are bounded by three Shockley partials (of the type $\frac{1}{6}\langle 112\rangle$) with a radius of curvature of R. They obtained a smaller value of 8 mJ m^{-2}. Note that Chu *et al.* (1995a) re-estimated SFE from the data of Yoshida *et al.* (1995) to obtain γ_{SF} = 15–60 mJ m^{-2}. More recently, Kazantzis *et al.* (1996) obtained a value of 25 mJ m^{-2} from TEM observations of extended triple-junction nodes in Cr_2Nb at 1400 and 1500°C. These results are tabulated in table 2.

Our results for SFEs are higher than the experimental results. Certainly, SFE obtained at high temperatures can be expected to be lower than the calculated value at 0 K, which has also been suggested by Kazantzis *et al.* (1996). The unknown image

Table 2. Stacking fault and twin boundary energies for C15 Cr_2Nb, using the experimental lattice constant.

Method	γ_{SF} (mJ m^{-2})		γ_T (mJ m^{-2})
This work	116 (intrinsic)	94 (extrinsic)	39
LMTO[a]	90 (intrinsic)		—
Exp[b]	25		—
Exp[c]	>8		—

[a] Chu *et al.* (1995a): estimated from energy difference between C14 and C15.
[b] Kazantzis *et al.* (1996): estimated from dislocation triple-junction.
[c] Yoshida *et al.* (1995): estimated from dislocation triple-junction.

shift of Shockley partial dislocations can also give an uncertainty in the SFE determination; for example, for SFE in silicon, $\gamma_{SF} = 30\,\mathrm{erg\,cm^{-2}}$ has an uncertainty in the range $15 < \gamma_{SF} < 75\,\mathrm{erg\,cm^{-2}}$, after taking into account the image shift (Aerts *et al.* 1962). From these points of view, the temperature and image shift effects may be responsible for the difference between our calculated value and experimental results.

Our theoretical calculations on Cr_2Nb were performed for the stoichiometric composition. In experiments, however, some localized variations in chemical composition are possible in polycrystalline compounds. For instance, the two Cr_2Nb alloys investigated by Yoshida *et al.* (1995) were Cr–32. 2% Nb and Cr-34. 0% Nb, in which the second phase particles observed are Cr solid solution in the former and Nb solid solution in the latter. Also, polycrystals used in experiments may have compositions that deviate from the ideal stoichiometry ratio, since the phase region of the C15 Cr_2Nb is relatively large. For example, as mentioned by Hong and Fu (1999), the calculated elastic moduli for the stoichiometry alloy at the experimental lattice constant are very different from the experimental values (Chu *et al.* 1995b) obtained from polycrystals. Therefore, this composition effect can also be partially responsible for the discrepancy between theory and experiment.

The calculated twin boundary energy, $\gamma_T = 39\,\mathrm{mJ\,m^{-2}}$, at the experimental lattice constant for Cr_2Nb is relatively low, for instance, in comparison to $\gamma_T = 60\,\mathrm{mJ\,m^{-2}}$ for TiAl of the Ll_0 structure (Fu and Yoo 1990). In view of the energetics of twin nucleation, such a low twin boundary energy suggests a high propensity of twinning in Cr_2Nb. This is consistent with the experimental observations of twinned microstructures in Cr_2Nb, formed due to plastic deformation at elevated temperatures (Yoshida *et al.* 1995) and to the internal stresses resulting from the C14–C15 transformation as well as the thermal contraction differences between Cr solid solution and Cr_2Nb in the two-phase alloy (Kumar and Liu 1997). However, the absence of deformation twinning in Cr_2Nb at low temperatures is not understood. Kinetic aspects of the motion of synchro-Shockley partials need to be elucidated in order to better understand twin formation in the C15 Laves phase.

§ 4. Interaction between Shockley partials

From the anisotropic elasticity theory (Stroh 1958, Hirth and Lothe 1982) we calculate the equilibrium separation between Shockley partials, using the calculated elastic constants (Hong and Fu 1999) and stacking fault energies.

First, we consider the following case for the ISF:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \mathrm{ISF} + \frac{1}{6}[2\bar{1}\bar{1}], \tag{4}$$

where $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$, $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$ correspond to $-\mathbf{a}_2$, $\mathbf{b}_3$, and $-\mathbf{b}_1$ in figure 1, respectively. This expression for the formation of an ISF is approximate since the Shockley partial $\mathbf{b}_3 = \frac{1}{6}[121]$ is the sum of two synchro-Shockley partial vectors $-\mathbf{b}_2$ and $-\mathbf{b}_1$, lying in the two successive atomic planes, shown in figure 2 (*a*). The equilibrium separation between Shockley partials can be obtained through the balance of the attractive force (the surface tension due to the ISF) and the repulsive force (due to elastic interaction between the partials).

The radial and tangential components of the interaction force (per unit length) between two parallel dislocations can be calculated by

$$F_r = f_r/(2\pi r), \tag{5}$$

$$F_\theta = f_\theta/(2\pi r). \tag{6}$$

Here, r is the separation between the two partials of Burgers vectors $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$, and f_r and f_θ are the radial (in the $\{111\}$ plane) and tangential (out-of-the plane) components of the interaction force constants, which are determined from the anisotropic elasticity theory of dislocations (Stroh 1958). For the dissociation of equation (4), the interaction force constants f_r and f_θ are obtained numerically (Yoo 1987), by (i) letting $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$, (ii) using the experimental lattice constant $a_0 = 6.991$ Å for C15, and (iii) the calculated elastic constants and γ_{ISF} at experimental lattice constant.

As shown in figure 3, the radial component f_r increases monotonically from $\phi = 0$ (screw) to $\phi = 90°$ (edge), where ϕ is the angle between a dislocation line and the Burgers vector **B**. The tangential component f_θ is maximum in magnitude at $\phi = 31°$ and zero at the edge orientation. At $\phi \sim 30°$, the two parallel Shockley partials are inclined at about 60° and 0° to their respective Burgers vectors. This implies that while the repulsive F_r balances the surface tension γ_{SF} (see equation (7) below), the out-of-plane force ($F_\theta = 0.5F_r$ at $\phi = 30°$) promotes cross-slip of the screw Shockley partial and climb of the 60° Shockley partial, most likely onto the $(\bar{1}01)$ plane. This Shockley partial dissociation configuration of the cross-slip and climb combination may lead to a possible mechanism for the thickening process of a twin embryo originating from a mixed ($\phi = 30°$) $\frac{1}{2}[1\bar{1}0]$ dislocation.

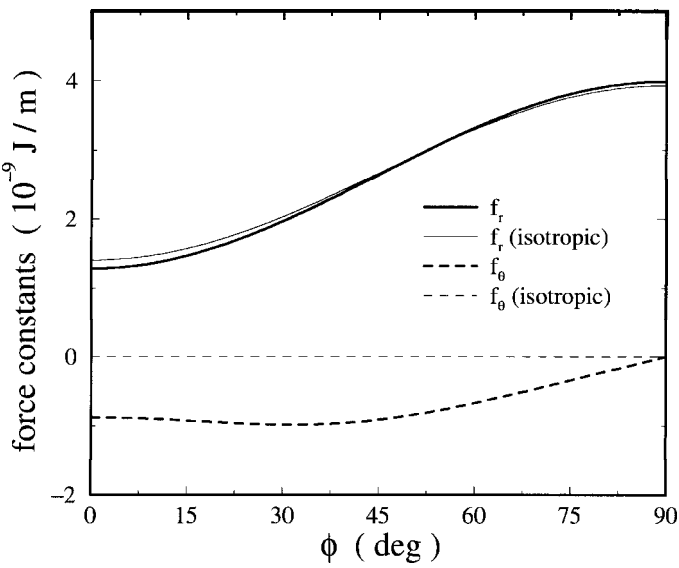


Figure 3. Orientation dependence of the interaction force constants in Cr_2Nb for the ISF-type dissociation, using the results obtained at the experimental lattice constant. ϕ is an angle between a dislocation line and the Burgers vector $\frac{1}{2}[1\bar{1}0]$.

The width of the equilibrium separation, w, can be obtained from

$$\gamma_{\mathrm{SF}} = F_r = f_r/(2\pi w), \tag{7}$$

$$w = f_r/(2\pi\gamma_{\mathrm{SF}}). \tag{8}$$

The separation is $w_{\mathrm{s}} = 17.6$ Å for the $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$ screw dislocation, and $w_{\mathrm{e}} = 54.7$ Å for the edge dislocation. These results are shown in table 3, together with those in the isotropic case (discussed below). Also, the results using γ_{ESF} and the experimental result are tabulated.

For the isotropic case, the equilibrium separation w is given by a simple formula (Hirth and Lothe 1982):

$$w = \frac{Gb^2}{8\pi\gamma_{\mathrm{SF}}}\frac{2-\nu}{1-\nu}\left(1 - \frac{2\nu\cos 2\phi}{2-\nu}\right). \tag{9}$$

Here, $b = |\mathbf{b}^{(1)}| = |\mathbf{b}^{(2)}| = a_0/6^{1/2}$. The Hill's average values of shear modulus and the Poisson's ratio are $G = 50.0$ GPa and $\nu = 0.383$, respectively. Using γ_{ISF}, it is found that $w_{\mathrm{s}} = 19.3$ Å and $w_{\mathrm{e}} = 53.9$ Å. As given in table 3, the results in anisotropic and isotropic cases are rather close, because the shear anisotropy of C15 Cr_2Nb is moderate, $A = 1.45$ (Hong and Fu 1999), compared to the isotropic case ($A = 1$).

The radial component f_r of the isotropic case is very close to that of the anisotropic one, whereas its tangential component f_θ is exactly zero (figure 3). While the anisotropic corrections to f_r and hence to w in equation (9) are very small, the anisotropic tangential component of the elastic interaction force is quite large; for instance, $f_\theta/f_r \sim 0.7$ at $\phi = 0$. This indicates that cross-slip of $\frac{1}{2}[1\bar{1}0]$ dislocation is difficult in Cr_2Nb because of a large constriction energy for the Shockley partials that include the non-radial component of the interaction energy.

Next, let us consider the ESF case:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \frac{1}{6}[11\bar{2}] + \mathrm{ESF} + \frac{1}{6}[2\bar{1}\bar{1}] + \frac{1}{6}[\bar{1}\bar{1}2] \tag{10}$$

$$\approx \frac{1}{6}[2\bar{1}\bar{1}] + \mathrm{ESF} + \frac{1}{6}[1\bar{2}1]\,. \tag{11}$$

To calculate the width of the ESF, we approximate two Shockley partials on two successive TCP units on either side of the ESF by a single Shockley partial (as in the ISF case), and use the same formulas for the ISF. In other words, the only difference

Table 3. Equilibrium separation between Shockley particles bounding an ISF and an ESF, using the results obtained at experimental lattice constant. w_{s} and w_{e} are the width for screw and edge dislocations, respectively. The numbers are in angstroms.

		w_{s} (screw)		w_{e} (edge)	
		ISF	ESF	ISF	ESF
Theory ($T = 0$ K)	Anisotropic	18	22	55	68
	Isotropic	19	24	54	67
Exp[a] ($T = 1400$°C)		99		—	

[a] Kazantzis *et al.* (1996): measured from the ribbon after the correction for image shift and projection effects.

between ISF and ESF in the calculation of separation w is that the stacking fault energy used in the formulas is of intrinsic or extrinsic type. Using γ_{ESF}, it is found that $w_s = 21.7$ Å and $w_e = 67.5$ Å for the anisotropic case (table 3). For the isotropic case, they are $w_s = 24.4$ and $w_e = 66.5$ Å. As mentioned, the results of both anisotropic and isotropic cases are very close. Since γ_{ESF} is not much different from γ_{ISF}, the force constants f_r and f_θ using γ_{ESF} are expected to be similar to those in figure 3 using γ_{ISF}.

On the other hand, we consider the separation between partials by using the theoretical lattice constant. For the ISF, the separation is $w_s = 17.1$ Å and $w_e = 53.5$ Å for the anisotropic case, while $w_s = 18.7$ Å and $w_e = 52.8$ Å for the isotropic case. For the ESF, the separation is $w_s = 22.2$ Å and $w_e = 69.4$ Å for the anisotropic case, while $w_s = 24.2$Å and $w_e = 68.4$ Å for the isotropic case. It can be seen that the separations are very close when using both experimental and theoretical lattice constants.

Kazantzis *et al.* (1996) measured, at $T = 1400°C$, the width of the ribbon (separation between two parallel partial dislocations) as 99 Å after correction for image shift and projection effects. Note that they also obtained a different value of $w_s = 82$ Å for the width using equation (9) with $\gamma_{SF} = 25\,mJ\,m^{-2}$, which was estimated from the measured curvature of R in extended dislocation nodes and the temperature-corrected shear modulus. Compared with the experimental result at high temperatures, our results for the separation between partials are very small.

§ 5. Summary

We performed first-principles total-energy calculations to obtain stacking fault and twin boundary energies. The intrinsic and extrinsic stacking fault energies were calculated to be 116 and 94 mJ m^{-2}, respectively, and the twin boundary energy was 39 mJ m^{-2}. The calculated stacking fault energies are larger than the available experimental data, measured at high temperatures. We also calculated the equilibrium separations between Shockley partials using the calculated elastic constants and stacking fault energies. Our results of the equilibrium separations are very small compared with the experimental results reported at high temperatures. This discrepancy may be due to temperature and/or composition and image shift effects.

Acknowledgements

The research was sponsored by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

References

Aerts, E., Delavignette, P., Siems, R., and Amelinckx, S., 1962, *J. appl. Phys.*, **33**, 3078.
Allen, C. W., Delavignette, P., and Amelinckx, S., 1972, *Phys. status solidi* (a), **9**, 237.
Chu, F., Ormeci, A. H., Mitchell, T. E., Wills, J. M., Thoma, D. J., Albers, R. C., and Chen, S. P., 1995a, *Phil. Mag. Lett.*, **72**, 147.
Chu, F., He, Y., Thoma, D. J., and Mitchell, T. E., 1995b, *Scripta metall. mater.*, **33**, 1295.
Chu, F., Lu, Y.-C., Kotula, P. G., Mitchell, T. E., and Thoma, D. J., 1998, *Phil. Mag. A*, **77**, 941.

CHU, F., and POPE, D. P., 1993, *High-Temperature Ordered Intermetallic Alloys V*, *Materials Research Society Symposium Proceedings*, Vol. 288, edited by I. Baker, R. Darolia, J. D. Whittenberger and M. H. Yoo (Materials Research Society), p. 561.
FU, C. L., and YOO, M. H., 1990, *Phil. Mag. Lett.*, **62**, 159.
HAZZLEDINE, P. M., 1994, *Twinning in Advanced Materials*, edited by M. H. Yoo and M. Wuttig (The Minerals, Metals & Materials Society), p. 403.
HAZZLEDINE, P. M., and PIROUZ, P., 1993, *Scripta metall. mater.,* **28**, 1277.
HIRTH, J. P., and LOTHE, J., 1982, *Theory of Dislocations* (New York: Wiley).
HONG, S., and FU, C. L., 1999, *Intermetallics*, **7**, 5.
KAZANTZIS, A. V., AINDOW, M., and JONES, I. P., 1996, *Phil. Mag. Lett.*, **74**, 129.
KRONBERG, M. L., 1957, *Acta metall.,* **5**, 507.
KUMAR, K. S., and LIU, C. T., 1997, *Acta mater.,* **45**, 3671.
LIVINGSTON, J. D., and HALL, E. L., 1990, *J. mater. Res.,* **5**, 5.
STROH, A. N., 1958, *Phil. Mag.,* **3**, 625.
WIMMER, E., KRAKAUER, H., WEINERT, M., and FREEMAN, A. J., 1981, *Phys. Rev.* B, **24**, 864.
YOO, M. H., 1987, *Acta metall.,* **35**, 1559.
YOSHIDA, M., TAKASUGI, T., and HANADA, S., 1995, *High-Temperature Ordered Intermetallic Alloys VI*, *Materials Research Society Symposium Proceedings*, Vol. 364, edited by J. Horton, I. Baker, S. Hanada, R. D. Noebe and D. S. Schwartz (Materials Research Society), p. 1395.

ELSEVIER

Intermetallics 8 (2000) 1119–1129

Intermetallics

www.elsevier.com/locate/intermet

Physical metallurgy and mechanical properties of transition-metal Laves phase alloys

C.T. Liu *, J.H. Zhu [1], M.P. Brady, C.G. McKamey, L.M. Pike [2]

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6115, USA

Received 1 March 2000; accepted 15 August 2000

Abstract

This paper provides a comprehensive review of the recent research on the phase stability, point defects, and fracture toughness of AB_2 Laves phases, and on the alloy design of dual-phase alloys based on a soft Cr solid solution reinforced with hard XCr_2 second phases (where X = Nb, Ta and Zr). Anti-site defects were detected on both sides of the stoichiometric composition of $NbCr_2$, $NbCo_2$, and $NbFe_2$, while they were observed only on the Co-rich side of $ZrCo_2$. Only thermal vacancies were detected in the Laves phase alloys quenched from high temperatures. The room-temperature fracture toughness cannot be effectively improved by increasing thermal vacancy or reducing stacking fault energy through control of phase stability. Microstructures, mechanical properties, and oxidation resistance of dual-phase alloys based on Cr–$NbCr_2$, Cr–$TaCr_2$, and Cr–$ZrCr_2$ were studied as functions of heat treatment and test temperature at temperatures to 1200°C. Among the three alloy systems, Cr–$TaCr_2$ alloys possess the best combination of mechanical and metallurgical properties for structural use at elevated temperatures. © 2000 Published by Elsevier Science Ltd.

Keywords: A. Laves phases; B. Mechanical properties at high temperatures; B. Mechanical properties at ambient temperature; D. Defects: point defects

1. Introduction

Laves phases with AB_2 compositions are a common type of topologically close-packed (TCP) structures [1–4]. There are over 360 binary Laves phases reported in the literature. Furthermore, Laves phases are capable of dissolving considerable amounts of ternary alloying additions. A total of over 900 combined binary and ternary Laves phases have been documented. Although Laves structures are the most abundant among intermetallic compounds, they are the least investigated. So far, most previous studies on intermetallic alloys have concentrated on simple ordered structures derived from face-centered cubic (f.c.c.), body-centered cubic (b.c.c.), or hexagonal-close packed (h.c.p.) structures. Thus, there is a great need for the study of Laves phase alloys.

There are three common polytypes of Laves phases most frequently observed: cubic C15, hexagonal C14 and dihexagonal C36 [1–4]. These polytypes are related to each other as the basic unit layer of these phases is the same, while the stacking sequence of the unit layer is different in each structure. One unit layer in each Laves phase is composed of four sub-layers. In the C15 structure, the packing sequence of the unit layer is ABCABC, similar to the sequence in conventional cubic f.c.c. metals; in the hexagonal C14, the packing layer is ABABABAB, similar to the sequence in conventional h.c.p. metals. The dihexagonal C36 structure has a packing sequence of ABACABAC.

Laves phases are generally stabilized by the size-factor principles, that is, the atomic size ratio, R_A/R_B, is ideally 1.225, with a range of 1.05–1.68 usually observed [1–4]. The stability of each crystalline structure is also influenced by the electron concentration factor (e/a) [5,6]. In fact, the electron concentration factor becomes clearly important when the atomic size factors are favorable. The classic work by Laves and Witte [5,6] showed that for several quasi-binary alloy systems involving $MgCu_2$ and $MgZn_2$, with increasing valence electron concentration, the three Laves types, $MgCu_2$, $MgNi_2$, and $MgZn_2$, exist in that order.

* Corresponding author.

[1] Present address: Tennessee Technology University, TTU Box 5014, Cookeville, TN 38505-0001, USA.

[2] Present address: Haynes International, Kokomo, IN 46904, USA.

0966-9795/00/$ - see front matter © 2000 Published by Elsevier Science Ltd.
PII: S0966-9795(00)00109-6

Laves phases are an important and attractive type of material both scientifically and technologically. They are being seriously considered for many practical applications. For example, $(Tb,Dy)Fe_2$ has been used as a magnetoelastic transducer because of its great magnetorestriction [7]. Compounds of $(Hf,Zr)V_2$ have demonstrated superconducting properties with a combination of high critical temperature, high current density and good magnetic strength [8,9]. Laves phases, such as $Zr(Cr,Fe)_2$, have been considered for hydrogen storage applications due to favorable hydriding-dehydriding kinetics and high hydrogen-absorbing capabilities [10]. The $Mo(Co,Si)_2$ Laves phase contributes to the wear-resistance of "Tribaloy" materials [11]. More recently, $HfCr_2$-, $NbCr_2$- and $TiCr_2$-based two-phase alloys are being developed for high-temperature structural uses because of their high melting points and good retention of mechanical properties at elevated temperatures [12–20].

Despite their useful properties, the low ductility and brittle fracture characteristics at ambient temperatures are the main limitations for structural applications of this large class of alloys. The high hardness and brittle fracture of Laves phases are due to the complex atomic configuration of each unit layer, interplanar locking and lack of operating slip systems at ambient temperatures. Recently, a synchroshear mechanism [21] has been proposed to facilitate plastic deformation in Laves phases, which raises the hope of toughening the Laves phases, even though so far no experimental support for this mechanism has been furnished. All these facts indicate that more research needs to be undertaken on Laves phases in order to further understand the basic phase stability, defect structure, deformation mechanism and fracture behavior. Moreover, alloy design efforts should be devoted to overcoming the brittleness in these phases.

This paper summarizes our recent studies of single-phase and dual-phase alloys containing transition-metal elements. The phase stability and defect structures in Laves phases $NbCr_2$, $NbFe_2$, $NbCo_2$ and $ZrCo_2$ have been studied for the purposes of (1) promotion of synchroshear deformation by increasing vacancy concentration and (2) promotion of mechanical twinning by controlling phase stability and reducing stacking fault energy. At present only limited progress has been achieved in improving the fracture toughness of these Laves phase alloys. In view of this problem, dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard $NbCr_2$, $ZrCr_2$, or $TaCr_2$ second phases have been selected for alloy development [17,22–26]. Among the three Cr–XCr_2 (X = Nb, Zr, and Ta) systems, Cr–$TaCr_2$ alloys are more resistant to air oxidation and thermally-induced cracking. As a result, our current alloy development effort has been focusing on the Cr–$TaCr_2$ alloy system, and significant advances have been made in enhancing the high-temperature strength, creep resistance and fracture toughness of the dual-phase alloys through compositional adjustment, microstructural control and alloy additions.

2. Point defect and fracture toughness of binary Laves phases

Laves phases are generally considered to be line compounds with a strict AB_2 composition. However, solubility ranges exist for about 25% of the known binary Laves phases [27]. For non-stoichiometric compositions, constitutional defects are incorporated into intermetallic compounds. However, for the Laves phases, little is known about the possible defect structures that may be associated with the deviations from stoichiometry. The excess atoms in off-stoichiometric compounds can stay on their own sublattice (leading to the formation of constitutional vacancies on the other sublattice), insert into interstitial sites, or occupy sites on the other sublattice (anti-site substitution). Since Laves phases have TCP structures and space-filling is relatively high, there are no interstitial sites with a size comparable to that of the component atoms. Therefore, the insertion of the excess atoms into interstitial sites can be excluded, and the possible defect mechanisms in binary Laves phases reduce to either constitutional vacancy or anti-site substitution.

The Laves phases are known to be size compounds, i.e. the atomic size ratio, R_A/R_B is ideally 1.225, with a range of 1.05–1.67 typically observed. Since the A atom is much larger than the B atom, the excess A atoms in A-rich compositions would tend to stay on their own sublattice sites, thus creating vacancies on the B atom sublattice. On the other hand, the excess B atoms on B-rich compositions would be able to occupy the A atom sublattice sites, leading to the formation of anti-site defects. Based on these geometric arguments, it has been postulated that the A-rich side is accommodated by vacancies, while the B-rich compositions result from the anti-site substitution [27]. However, no systematic experimental verification has been undertaken so far.

The types of constitutional defects may affect physical, mechanical and functional properties of Laves phases. Vacancies have been proposed to assist the movement of synchro-Shockley dislocations [21], thus possibly facilitating the synchroshear deformation mechanism and increasing the toughness of Laves phases [28]. Off-stoichiometry was also found to affect the hydride stability of $ZrMn_2$ [29] and the hydrogen storage capability of $TiMn_2$ alloys [30]. The defect structures of the cubic C15 $NbCr_2$, $NbCo_2$ and $ZrCo_2$ and hexagonal C14 $NbFe_2$ Laves phases were studied to clarify the point defect mechanisms on both sides of stoichiometry in these binary systems. This was done by measuring the lattice parameters, bulk densities, and therefore, vacancy concentrations of these alloys of various compositions

[31,32]. The hardness of the above alloys with various point defect concentrations was determined in an attempt to explore the solid solution hardening behavior in the Laves phases. The fracture toughness was also measured to investigate whether the point defects affect the toughness of the alloys, and, in particular, to see if vacancies facilitate synchroshear deformation in the vacancy-containing alloys.

Figs. 1 and 2 show the vacancy concentration as a function of the Nb content for $NbCr_2$ and $NbFe_2$ alloys, respectively, after quenching from different temperatures [31]. It is clear that after quenching the $NbCr_2$ alloys from 1000°C, the vacancy concentration is essentially zero, while the vacancy concentration significantly increases after quenching from higher temperatures, i.e. 1400°C. For the $NbFe_2$ alloy, the vancancy concentration is very low even after quenching from 1300°C. Fig. 3 summarizes all the measured vacancy concentrations for all single-phase alloys after quenching from 1000°C, which are essentially zero. The calculated vacancy concentrations from both the constitutional vacancy model and anti-site substitution model are also plotted in Fig. 3 [31]. The experimental results indicate that there are no constitutional vacancies in these Laves phases on either side of stoichiometry. This is consistent with the anti-site substitution model and contrary to the constitutional vacancy model, which demands much higher vacancy concentrations for off-stoichiometric compositions. Note that the constitutional vacancy model is based on the assumption that all the excess atoms exclusively stay on their own sublattice, thus constitutional vacancies are created on the sublattice of the other element. The anti-site substitutional model assumes that the excess atoms occupy the sublattice sites of the other species. As a result, anti-site defects are created, and no constitutional vacancies are needed for the balance of the lattice sites. Thus, the constitutional defects on the Cr-, Co-, or Fe-rich, as well as the Nb-rich sides of stoichiometry, are of anti-site type for the $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves phases.

Modder and Bakker pointed out that it is very possible that a vacancy type defect, termed quadruple defect, occurs in C15 compounds [33]. Such a quadruple defect consists of one anti-site B atom on the α sublattice and three vacancies on the β sublattice. Ball milling was found to introduce quadruple defects in $GdAl_2$, $GdPt_2$, $GdIr_2$ and $GdRh_2$, while anti-site defects were created in $GdMg_2$ after ball-milling [34]. Since ballistic actions such as ball milling usually lead to the formation of the same type of atomic defects as heating, thermal defects such as quadruple defects or anti-site defects may be created upon quenching the Laves compounds from elevated temperatures. Thermal vacancies are expected

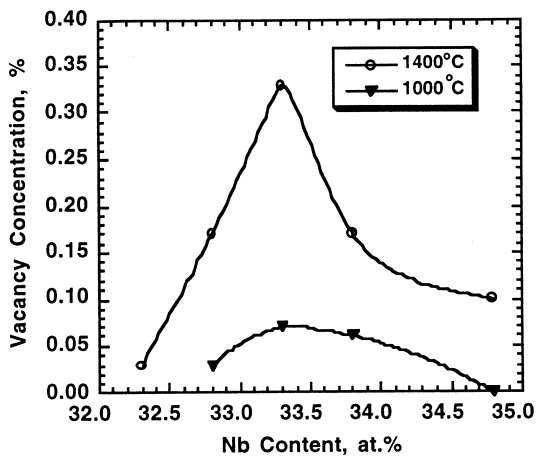


Fig. 1. Vacancy concentration vs Nb content for $NbCr_2$ Laves phase quenched from 1000 and 1400°C [31].

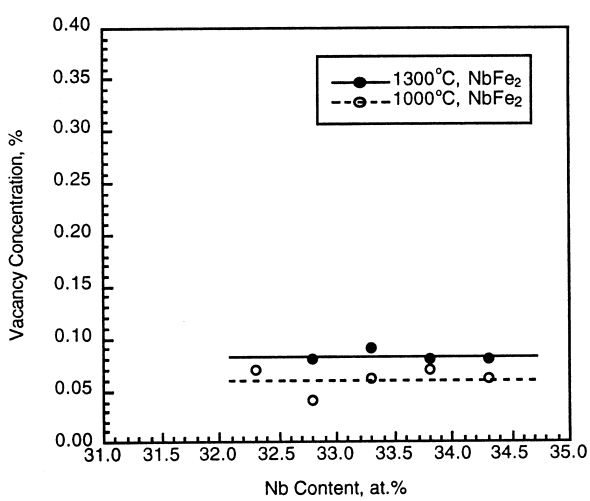


Fig. 2. Vacancy concentration vs Nb content for $NbFe_2$, Laves phase quenched from 1300 and 1000°C [31].

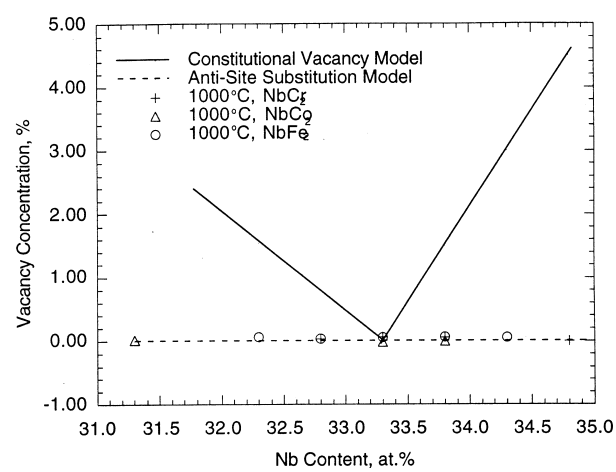


Fig. 3. Measured vacancy concentration vs Nb content in the binary $NbCr_2$, $NbCo_2$, and $NbFe_2$, Laves phases after quenching from 1000°C. Also shown are the calculated data for the constitutional vacancy model and the anti-site substitution model [31].

if quadruple defects are created, as each quadruple defect contains three vacancies.

It is interesting to note that appreciable vacancies were detected in the $NbCr_2$ alloy after quenching from 1400°C (Fig. 1). Since these vacancies were essentially annealed out at 1000°C (Fig. 1), and since the constitutional defects in this compound have been established to be of the anti-site type, it is concluded that these defects are *thermal* vacancies, indicating that quadruple defects may be formed in this compound by heating. Furthermore, the measured vacancy concentration exhibits a maximum at the stoichiometric composition and decreases on both sides of stoichiometry for the $NbCr_2$ alloys when quenched from 1400°C (Fig. 1). A maximum in thermal vacancy concentration, and thus thermal disorder, at the stoichiometric composition can be rationalized by considering the fact that the increase in entropy associated with the introduction of vacancies is the greatest at the stoichiometric composition where the degree of order is necessarily the highest.

For the C14 $NbFe_2$ compounds, no thermal vacancies were detected after quenching from 1300°C (Fig. 2), indicating that anti-site defects, instead of quadruple defects, may be the form of atomic disorder in this compound upon heating. According to the model developed by Modder et al. [34], the type of atomic disorder is related to the relative magnitudes of formation enthalpies of anti-site and quadruple-defect disorder for the Laves phases. Therefore, the differences in thermal defects in $NbCr_2$ and $NbFe_2$ imply that the formation enthalpies of anti-site disorder are larger in magnitude than that of the quadruple defect disorder for $NbCr_2$, while the opposite is true for $NbFe_2$.

A point defect in a crystalline lattice is well known to cause hardening. This phenomenon is referred to as solid solution hardening. Solid solution hardening has been extensively studied in B2 compounds [35]. Generally, it is found that in B2 compounds with anti-site defects, hardness exhibits a minimum at the stoichiometric composition, and deviation from stoichiometry causes hardening of the compounds, due to the presence of constitutional anti-site defects. For triple-defect B2 compounds, the shape of the hardness versus composition curves was found to be not so simple and depended on the degree of thermal disorder. Also, vacancies were found to be a more potent hardener than the anti-site defects.

In the $NbCr_2$ and $NbFe_2$ Laves phases studied, the hardness was also found to have a characteristic V-shape with a minimum occurring at the stoichiometric compositions (Figs. 4 and 5) [31]. Obviously, similar to anti-site B2 compounds, such off-stoichiometric hardening can be attributed to the presence of constitutional anti-site defects, which is the defect mechanism on both sides of stoichiometry in these Laves phases. The anti-site hardening mechanism in Laves phases, however, is not clear so far, and it may be different from that in B2 compounds. In B2 compounds, the anti-site hardening is often correlated to an interaction between the stress field of a moving dislocation and that of the anti-site

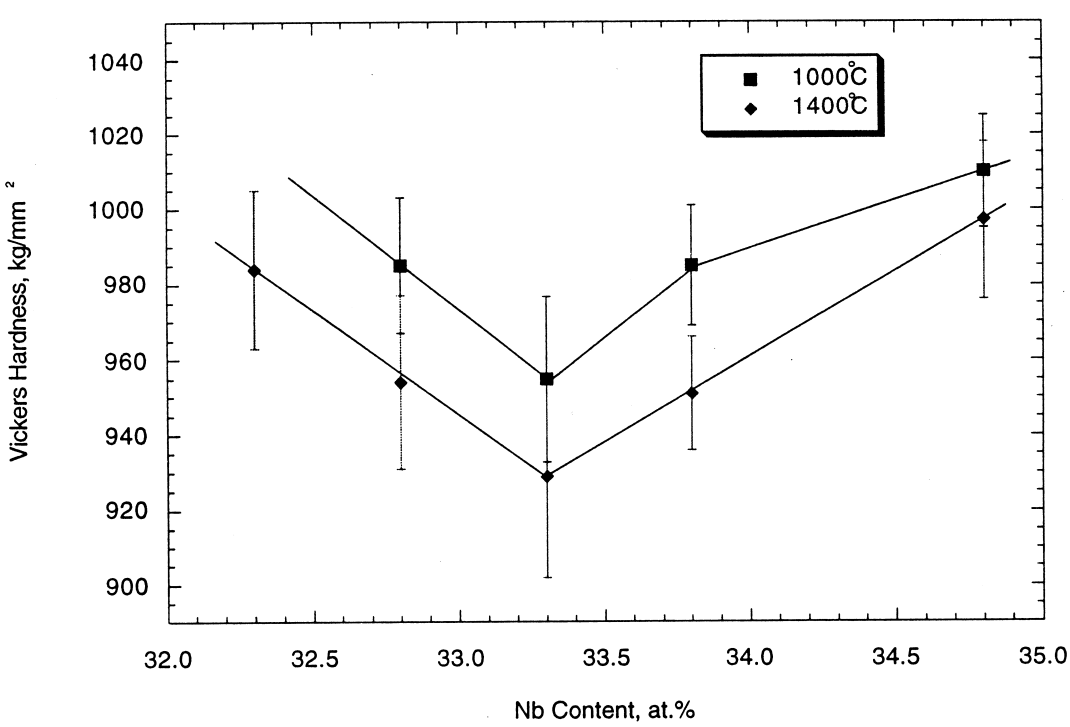


Fig. 4. Vickers hardness vs Nb content for $NbCr_2$ alloys after quenching from 1400 and 1000°C [31].

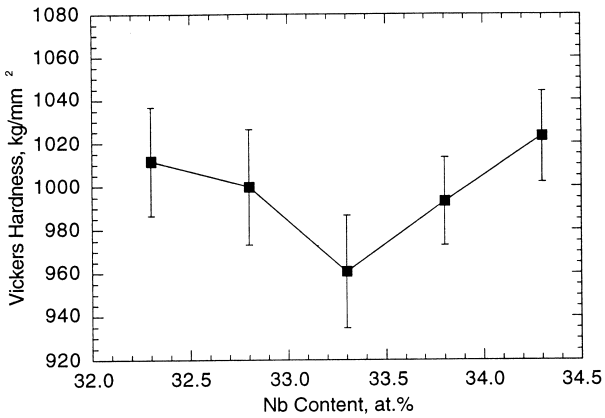


Fig. 5. Vickers hardness vs Nb content for $NbFe_2$ alloys after quenching from 1000°C [31].

defect; however, it is very difficult for ordinary dislocations to move in Laves phases at ambient temperatures.

The hardness values of all $NbCr_2$ alloys after quenching from 1400°C appear to be lower than those after quenching from 1000°C (Fig. 4). Since thermal vacancies are incorporated into the compound after quenching from 1400°C (see Fig. 1), the above results may suggest that vacancies cause softening, instead of hardening of the $NbCr_2$ Laves phases. Such vacancy-induced softening has not been observed previously in other systems. Furthermore, this is in direct contrast to the effect of vacancies on the hardness in B2 compounds, where vacancies are a potent hardener. The vacancy hardening in metals and B2 compounds is often correlated to an interaction between the stress field of a moving dislocation and that of the vacancy. The effect of vacancies in Laves phases may be different from their effects in metals and B2 compounds. This is again likely to be due to the fact that Laves phases have a topologically close-packed structure and it is very difficult for ordinary dislocations to move. The presence of vacancies in Laves phases would make the structure less closely packed and could possibly assist the synchroshear deformation of the alloy [21,36], thus possibly leading to the softening of the alloys.

The fracture toughness values as a function of the Nb content in $NbCr_2$ Laves phases after quenching from 1400 and 1000°C as well as $NbCo_2$ and $NbFe_2$ Laves phases after quenching from 1000°C are plotted in Fig. 6. Basically, the fracture toughness of the $NbCr_2$ and $NbCo_2$ alloys is close to 1 MPa $m^{1/2}$. No effect of stoichiometry on the fracture toughness can be discerned. Even though the anti-site defects harden the Laves phases (Figs. 4 and 5), they are not detrimental to the fracture toughness of the alloys. Furthermore, it is noted that the fracture toughness of the $NbCr_2$ alloys after quenching from 1400°C is similar to that of the alloys after quenching from 1000°C, i.e. the presence of thermal vacancies (<0.4%) does not lead to an improved toughness of the Laves phases, even though the vacancies slightly soften the Laves phases (see Fig. 4). Similar results were observed for $ZrCo_2$ alloys quenched from different temperatures [32]. The fracture toughness of the C14 $NbFe_2$ alloys is lower than that of C15 $NbCr_2$ and $NbCo_2$. This trend can be explained by the fact that the C15 structure is cubic, and is possibly more deformable than the hexagonal C14 structure.

It has been suggested that vacancies in Laves phases may assist the synchroshear mechanism, thus leading to improvement in the deformability and toughness of the Laves alloys [21,36]. Chen et al. [28] attributed the toughness improvement in the off-stoichiometric $TiCr_2$ alloys to the presence of vacancies in the alloys. The quenched-in thermal vacancies do not affect the crack propagation and fracture toughness behavior of the $NbCr_2$ alloys (Fig. 6). This is possibly due to the fact that the level of vacancies is too low in the alloys to enhance toughness significantly. It is also possible that the high residual thermal stress after quenching from 1400°C may also embrittle the $NbCr_2$ Laves phases, and thus could mask the effect of vacancy-assisted synchroshear deformation and its associated toughening.

The atomic size ratio, R_A/R_B, of the AB_2 Laves phases studied can be calculated to be 1.145, 1.173, and 1.152 for $NbCr_2$, $NbCo_2$, and $NbFe_2$, respectively. Here, R_A and R_B are the atomic radii of the A and B atoms with a coordination number of 12. Obviously, all the Laves phases studied have R_A/R_B ratios smaller than the ideal ratio of 1.225 for Laves-phase formation. It will be interesting to know the defect mechanism for Laves phases with R_A/R_B greater than 1.225. It may be easier to obtain constitutional vacancies for Laves phases with $R_A/R_B > 1.225$, due to geometric considerations, i.e. it is more difficult for A atoms to stay on the sublattice sites of B atoms as the atomic size ratio increases. On the other hand, it has been found that for the B2 phases, the

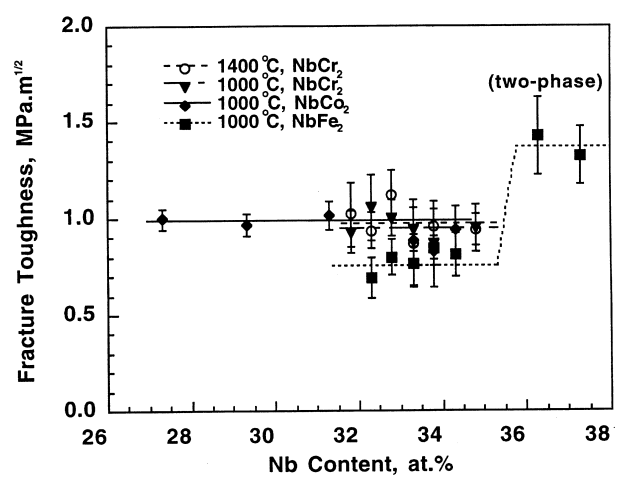


Fig. 6. Fracture toughness vs Nb content for binary $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves-phase alloys. Note two of the $NbFe_2$ alloys are not single Laves phase [31].

defect type is closely related to the enthalpies of formation (ΔH) of the compounds. Neumann has demonstrated that B2 compounds with negative ΔH smaller than 75–90 kJ/mol generally exhibit the anti-site defect structure, while compounds with a greater value generally exhibit the triple defect structure and have constitutional vacancies on the large-atom side of stoichiometry [37]. In other words, constitutional vacancies are the preferred defect mechanism in compounds that are more strongly ordered, i.e. with higher absolute ΔH values on the large-atom side. This empirical observation may be extended to the binary Laves phases. The ΔH values for $NbCr_2$, $NbCo_2$, and $NbFe_2$ are −21, −51, and −63 kJ/mol, respectively [38]. These values are in the same range as those of B2 compounds with anti-site defects. It is not surprising that only anti-site defects are observed in these Laves phases. Based on such geometric and thermodynamic considerations, future work should be concentrated on Laves phases with high R_A/R_B ratios and/or large negative enthalpies of formation. One such Laves phase is $ZrCo_2$, which has a large difference in atomic size between component atoms ($R_A/R_B = 1.28$) and an enthalpy-of-formation value of −123 kJ/mol [32]. Recent results indicate that a second phase, instead of constitutional vacancies, is formed in $ZrCo_2$ alloys with Zr > 33.3%. Interestingly, thermal vacancy concentrations in $ZrCo_2$ alloys of near stoichiometric composition are as high as 1%, see Fig. 7 [32]. These vacancies do not affect fracture toughness.

3. Phase stability and fracture toughness of transition-metal Laves phases

As mentioned above, the different Laves phase polytypes are a result of different stacking sequences. The stacking fault energy is expected to be lower at phase boundaries (e.g. C14/C15 boundary), where mechanical twinning becomes more easily triggered.

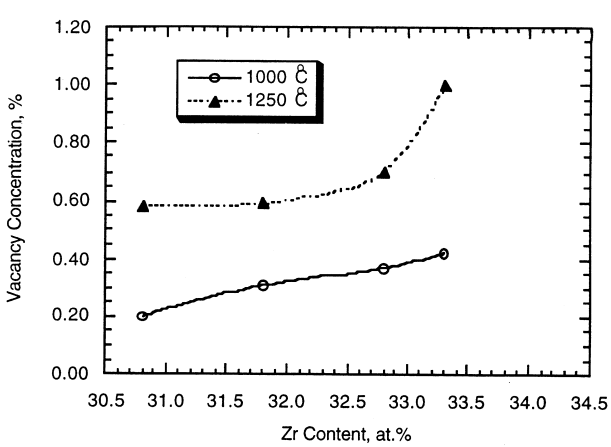


Fig. 7. Vacancy concentration vs Zr content for $ZrCo_2$ Laves phase quenched from 1000 and 1250°C [32].

The stability ranges of the Laves phases in both binary and ternary systems were obtained from phase diagram information, and the e/a effect on the phase stability (C14/C15) in $NbCr_2$-based transition-metal Laves phases is shown in Fig. 8 [39,40]. In the $NbCr_2$-based Laves phases, the e/a ratio for the C15/C14 phase boundaries is quite precise; the critical e/a values, corresponding to the C15/C14/C15 phase transitions from the phase diagram information (as indicated in Fig. 8), are as follows: at e/a values lower than 5.76, the C15 structure is stabilized for both binary and ternary Laves alloys; by increasing e/a to 5.88, the C14 structure is stabilized; over the e/a range of 5.88–7.53, the C14 structure is more stable than the C15 structure; the C15 structure is stabilized again over the C14 structure when the e/a ratio is increased further to 7.65 for the Nb–Cr–Co system. In the ternary Nb–Cr–Fe system, such a C14→C15 transition was not observed, since the highest e/a ratio was 6.69 in this system (Fig. 8). No $NbNi_2$ (with e/a = 8.34) or $NbCu_2$ (with e/a = 9) Laves phases exist in binary Nb–Ni and Nb–Cu systems (the e/a values for imaginary "$NbNi_2$" and "$NbCu_2$" were shown in Fig. 8), consistent with the observation by Bardos et al. [41] that at e/a > 8, a disordered structure is stabilized over the Laves phase in transition-metal systems.

The phase stability in pseudo-binary $Nb(Cr,Fe)_2$ and $Nb(Cr,Co)_2$ systems was also studied experimentally and good agreement between the experimental results and surveyed data based on phase diagram information was observed, as shown in Fig. 8. It is clear that the electron concentration rule is obeyed in these two pseudo-binary systems: the phase transitions between C15 and C14 are determined by the average electron concentration in these alloys. The experimental critical e/a ratios for the C15/C14/C15 transitions are very close to those from phase diagrams. These results were the first experimental determination of the critical e/a values for the different Laves phases in the pseudo-binary $Nb(Cr,X)_2$ systems. Such good agreement may be associated with the fact that all the components in the two

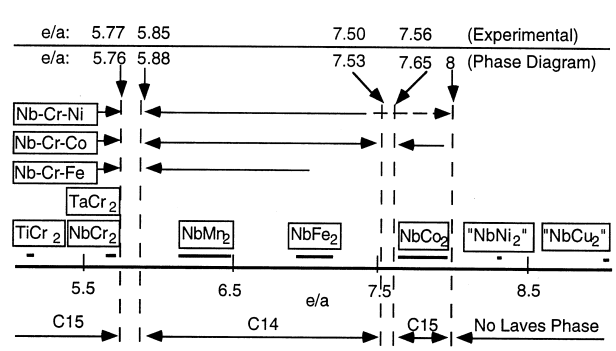


Fig. 8. Effect of average electron concentration (e/a) on phase stability in $NbCr_2$-based systems, with the critical e/a values for phase transitions both from phase diagram information and experimental data [40].

pseudo-binary systems are transition metals. Furthermore, it is determined that the region between C14 and C15 is a two-phase mixture of C14 and C15, with no intermediate structures such as C36 observed.

The effect of Fe and Co contents on the fracture toughness of $Nb(Cr,X)_2$ Laves phase is shown in Figs. 9 and 10, respectively [40]. Basically, the fracture toughness of the monolithic Laves phase is very low, around 1 MPa$\sqrt{m}$, which is a consequence of the topologically close packing manner of the structure and the subsequent difficulty in activating dislocation sources in the Laves phase. The cubic C15 phase has relatively higher fracture toughness than that of C14 structure for the ternary systems, regardless of the Fe or Co addition. This can be explained by the fact that the C15 structure is cubic, and is thus more deformable than the hexagonal C14 structure. The toughness value near the phase boundary could be higher, in view of the fact that the associated structure is metastable with a lower stacking fault energy, thus possibly making stress-assisted phase transformation and/or mechanical twinning easily triggered during deformation. However, such a trend was not observed in the present study. None of the alloys within the two-phase regions or near the phase boundaries exhibits any sign of increase in toughness. The toughness of the two-phase (C14 + C15) alloys is similar to that of the C15 phase. The toughness enhancement through stress-assisted phase transformation or mechanical twinning is not significant enough to be noted in the indentation test.

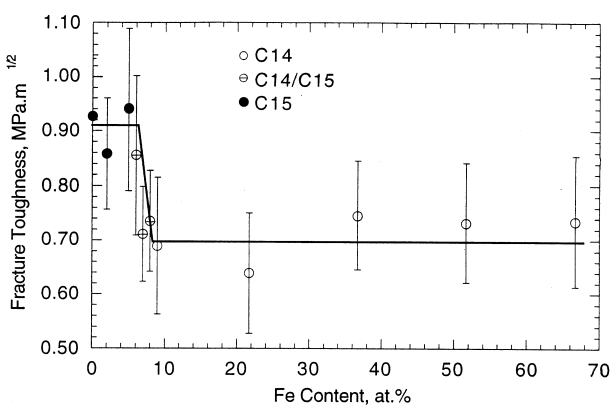


Fig. 9. Fracture toughness values of $Nb(Cr,Fe)_2$ as a function of the Fe content [40].

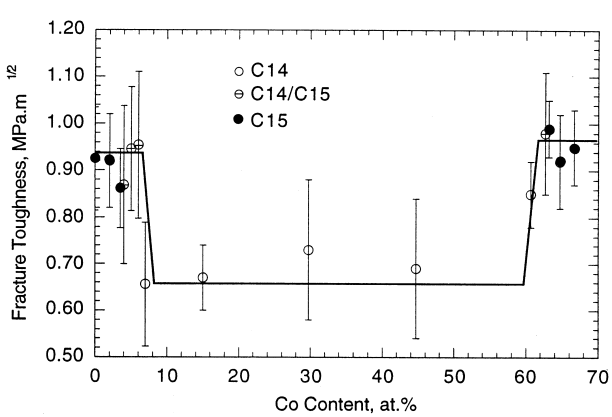


Fig. 10. Fracture toughness values of $Nb(Cr,Co)_2$ as a function of the Co content [40].

4. Alloy design of dual-phase Cr-XCr_2 alloys

As indicated in the forgoing two sections, single-phase Laves phase alloys with cubic C15 and hexagonal C14 are hard and brittle, and their fracture toughness at ambient temperatures can not be effectively improved using physical metallurgy principles, including control of point defect and phase stability. In view of this, dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard Laves phase second phases have been selected for alloy development [17,22–26] for potential structural use at elevated temperatures. This section summarizes the progress made so far in the development of dual-phase Cr–XCr_2 alloys, where X = Nb, Ta, or Zr.

Three eutectic alloy systems, Cr–$NbCr_2$, Cr–$TaCr_2$, and Cr–$ZrCr_2$, were initially selected for alloy development study. The selection is based on consideration of the different eutectic phase relationships in these alloy systems [16,25,26,42], as indicated in Table 1. Fig. 11 shows the Cr–$TaCr_2$ phase diagram which is typical for the three eutectic systems. Among these systems, the Cr–Nb system has the highest maximum solute solubility (5.6% Nb) in the Cr solid solution and the highest solute concentration in the eutectic composition (18.5% Nb), whereas the Cr–Zr system has the least maximum solute solubility (< 0.6% Zr) in Cr and the Cr–Ta system has the lowest solute concentration in the eutectic composition (9.8% Ta). These alloy parameters are expected to strongly affect the hardening behavior of both the Cr solid solution phase and the eutectic structure and thus the degree of brittleness of these phases.

Alloy ingots containing up to 20 at. % X element were prepared by arc melting and drop casting, using pure metal elements. In order to control precipitation reactions, alloy ingots were sectioned and heat treated between 1000 and 1400°C. The first interesting observation was that alloy ingots were quite often cracked for Cr–$NbCr_2$ alloys but not for Cr–$TaCr_2$ and Cr–$ZrCr_2$

Table 1
Alloy parameters existing in the three eutectic systems Cr–Nb, Cr–Ta and Cr–Zr

Alloy parameter	Cr–Nb	Cr–Ta	Cr–Zr
Eutectic temperature (C°)	1703	1760	1592
Max. solute solubility in Cr (at%)	5.6 Nb	~4.0 Ta	< 0.6 Zr
Eutectic composition (at%)	18.5 Nb	9.8 Ta	~18.0 Zr

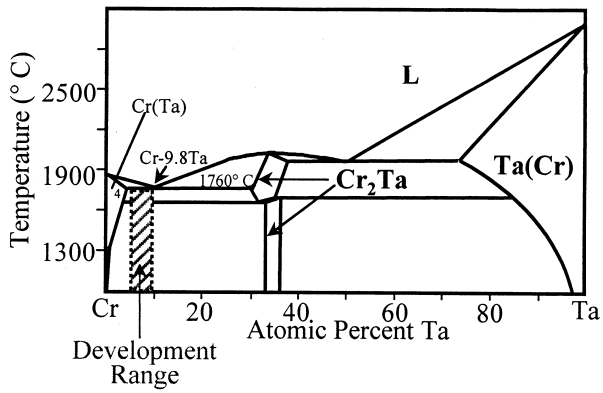


Fig. 11. Schematic binary Cr–Ta phase diagram.

alloys during drop casting and cooling from heat treatment at high temperatures. This cracking behavior cannot be simply related to the difference in coefficients in thermal expansion (CTE) because both experimental measurements and first principle calculations [43–45] indicate no major difference in CTE among the Laves phases $NbCr_2$, $TaCr_2$ and $ZrCr_2$. The cracking behavior, on the other hand, can be better rationalized by considering alloy parameters in the three eutectic systems. As indicated in Table 1, the maximum solute solubility in Cr and the solute concentration in the eutectic structures (which increases the proportion of the hard Laves phases in the eutectic microstructure) are the highest in the Cr–$NbCr_2$ system, resulting in significant hardening and embrittlement of the primary Cr solid solution phase and the eutectic composition. In this case, tremendous hardening is induced from fast cooling from the maximum-solubility temperature to lower temperatures where the primary Cr solid solution is in supersaturated states. As a result, thermal induced stresses generated by CTE mismatch will high enough to crack these phases during drop casting or cooling from high temperatures [46,47], quite often resulting in catastrophic ingot failure. This study clearly indicates the importance of phase relationship that may provide key information governing the mechanical and metallurgical behaviors of individual alloy phases.

In comparison with the dual-phase Cr–$TaCr_2$ alloys, Cr–$ZrCr_2$ alloys suffer from two major problems. One is the poor oxidation resistance of Cr–$ZrCr_2$ alloys, due, in part, to extensive internal oxidation of Zr-rich phases, and the other is the relative low melting point of these alloys (see Table 1). Thus, our alloy development has been concentrated on the Cr–$TaCr_2$ alloy system, and significant progress has been made in improving the mechanical and metallurgical properties of these alloys [26]. The Cr–$TaCr_2$ alloys with enhanced properties have the general compositions (at%) below:

$$\mathrm{Cr}-(6-10)\mathrm{Ta}-(3-6)\mathrm{Mo}-(0.2-1.0)\mathrm{Ti}-(0.5-3.0)\mathrm{Si}-(0.05-0.20)\mathrm{La/Ce}.$$

In these alloys, Mo is added mainly for solid solution hardening the Cr-rich phase, Ti for scavenging interstitials, and Si for both improving oxidation resistance and fracture toughness. Small amounts of reactive elements such as La and Ce are also added for decreased scaling and enhancing the adhesion of oxide scales.

No micro- and macro-cracks were observed in the Cr–$TaCr_2$ alloy ingots prepared by arc-melting and drop casting. In order to control microstructures, some dual-phase alloys were canned in Mo billets and fabricated by hot extrusion at 1480°C. The alloy CN-151 (Cr-9.75Ta-5Mo-0.5Ti-0.01Ce, at. %) with its composition similar to the eutectic composition shows a full lamellar structure in the as-cast condition and a $TaCr_2$-dispersed microstructure in the hot extruded condition (Fig. 12). The mechanical properties with different microstructures were determined in tension as a function of test tem-

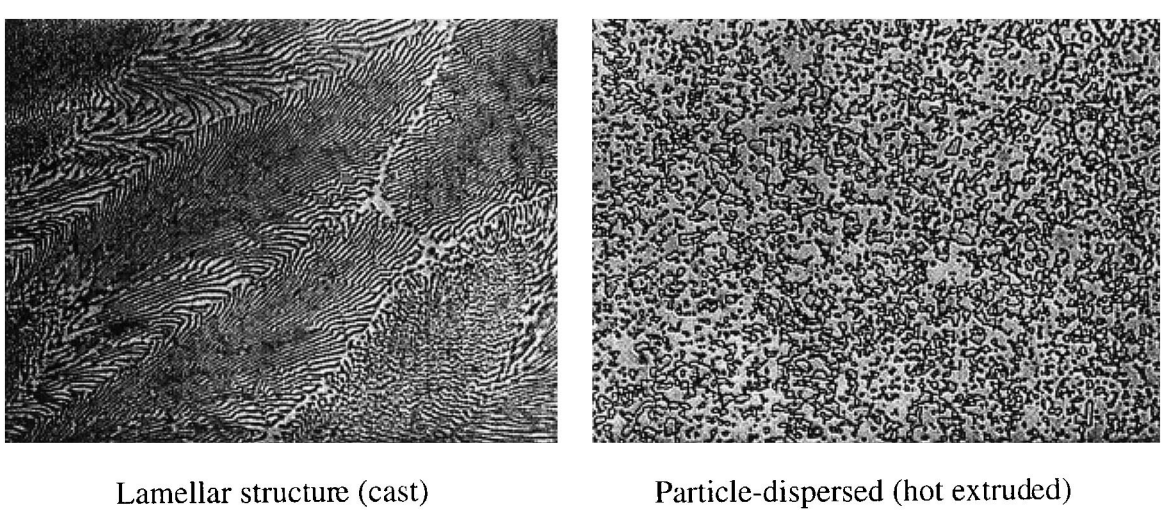


Fig. 12. Comparison of lamellar structure in the cast condition and particle-dispersed microstructure in the hot extruded condition for alloy CN-151.

perature. All tensile specimens were given a final heat treatment of 1 d/1200°C + 1 d/1000°C for stress relief. Note that the heat treatment does not significantly affect the lamellar structure observed in cast alloys.

Table 2 compares the tensile properties of CN-144 alloy with the lamellar and $TaCr_2$-dispersed microstructures. With the dispersed microstructure, the alloy showed a fracture strength of more than 100 ksi (700 MPa), even though no appreciable plastic deformation was measured at room temperature. Plastic deformation was detected at 800°C, and the strain went up to as high as 40% at 1200°C. The alloy is very strong at elevated temperatures, with the yield strength of 37 ksi and the ultimate tensile strength of 46 ksi at 1200°C. Note that the strengths of most Ni-base superalloys vanish at 1200°C. In comparison with the particle-dispersed microstructure, the alloy with the lamellar structure showed the ductile-to-brittle temperature around 1000°C, instead of 800°C. The lamellar structure is extremely strong at 1200°C, with its strengths close to double that for the particle-dispersed microstructure.

The fracture toughness of Cr–$TaCr_2$ alloys with lamellar structures was determined at room temperature, using subsized notched Chevron specimens [48] in accordance with reference 49. The results are summarized in Table 3. The alloy CN-144 with a hypoeutectic composition showed a low fracture toughness of 8.3 MPa $m^{1/2}$. A significant increase in fracture toughness is achieved by increasing the Ta concentration from hypo- to hypereutectic compositions. Alloying with 1.0% Si further increases toughness from 11.5 to 14.3 MPa $m^{-1/2}$. Examination of fracture surfaces reveals that the interfaces between the Cr solid solution and lamellar $TaCr_2$ phases are strong, and no secondary cracks are observed along these interfaces. It is worth noting that hot extrusion and change in microstructures from lamellar to dispersed particles gives no significant increase in fracture toughness.

Table 2
Comparison of tensile properties[a] of CN-144 with lamellar and particle-dispersed microstructures

Test temperature (C°)	Strength (ksi) Yield	Fracture	Elongation (%)
Particle-dispersed microstructure (hot extruded)			
RT	–	104	0.3
800	130	133	1.0
1000	92	103	1.7
1200	37	46	39.0
Lamellar structure (cast)			
RT	–	76	0.2
800	–	–	–
1000	–	80	0.2
1200	68	85	5.3

[a] Tested at a strain rate of 3.3×10^{-3}/s in air.

The creep properties of a Cr–$TaCr_2$ alloy, CN-144 (Cr-8Ta-5Mo-0.5Ti-0.01Ce. at%), were determined in tension at a stress of 20 ksi (138 MPa) at 1000°C in air [26]. Fig. 13 compares the creep curve of CN-144 with the lamellar and dispersed microstructures. The $TaCr_2$ dispersed microstructure exhibited ductile rupture with a rupture life of 695 h at 1000°C. In comparison, the specimen with the lamellar structure is much stronger in creep, with a steady state creep rate lower than that of the particle-dispersed microstructure by an order of magnitude. The test of the lamellar specimen had to be stopped after 1438 h because of grip rod failure. The creep resistance of this alloy in the polycrystalline form with a lamellar structure is stronger than Mar M-200 and DS Rene'80 and is comparable to CMSX-11B in the single crystal form.

The oxidation properties of Cr–XCr_2 alloys were characterized by exposing alloy coupons to air at 900–1100°C. The oxidation studies by Brady et al. [26,50,51] demonstrates that the oxidation resistance of Cr–$TaCr_2$ alloys is much superior to Cr–$NbCr_2$ or Cr–$ZrCr_2$ alloys. Fig. 14 compares the cyclic oxidation behavior of Cr-8Ta alloys with commercial Cr-base alloys tested at

Table 3
Fracture toughness of Cr–Ta alloys at room temperature

Alloy number	Composition[a] (at%)	Fracture toughness (MPa $m^{1/2}$)
CN-144	8.0 Ta–0.01 Ce	8.3
CN-161	9.8 Ta–0.1 La	11.5
CN-145	10.0 Ta–0.01 Ce	12.8
CN-162	9.8 Ta–0.1 La–0.5 Si	11.4
CN-163	9.8 Ta–0.1 La–1.0 Si	14.3

[a] Base composition: Cr-5 Mo-0.5 Ti, at%.

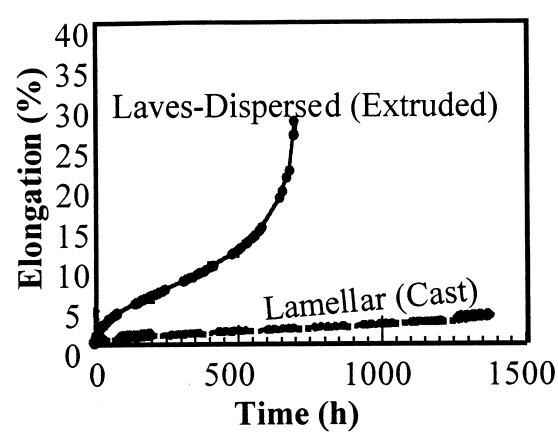


Fig. 13. Comparison of creep properties of CN-144 (Cr-8Ta-5Mo-0.5Ti-0.OlCe, at%) with Laves phase dispersed or lamel lar microstructures tested at 1000°C and 20 ksi in air. Note that the lamellar microstructuie test was stopped due to failure of the grip rods and plate.

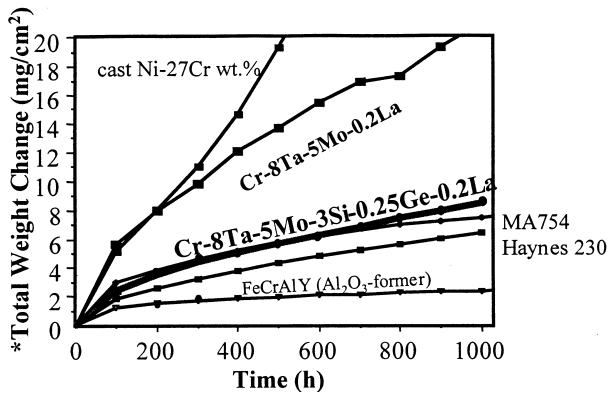


Fig. 14. Comparison of cyclic oxidation of Cr-8Ta alloys with cast Ni-27Cr alloy (wt%) and commercial alloys MA754 and Haynes 230 air oxidized at 1100°C.

1100°C in air [26]. The oxidation rate of Cr-8Ta-5Mo-0.2La is comparable to that of cast Ni-27Cr (wt%), and alloying with additions of Si and Ge significantly lowers the oxidation rate of the Cr–8Ta alloy. As indicated in Fig. 14, the oxidation resistance of Cr-8Ta-5Mo-3Si-0.25Ge-0.2La is comparable to several commercial Cr_2O_3-forming alloys, such as MA-754 and Haynes 230.

5. Conclusions

This paper provides a comprehensive review of phase stability, point defects, and fracture toughness of binary AB2 Laves phase alloys based on $NbCr_2$, $NbFe_2$, $NbCo_2$, and $ZrCo_2$ with on- and off-stoichiometric compositions, and pseudo-binary alloys based on $Nb(Cr,Co)_2$ and $Nb(Cr,Fe)_2$ compositions. The alloy design effort has been concentrated mainly on dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard $TaCr_2$ second phases. Conclusions deduced from these studies are given below:

1. Anti-site defects were observed to form on both sides of the stoichiometry in $NbCr_2$, $NbCo_2$, and $NbFe_2$ with $R_A/R_B < 1.225$. On the other hand, for $ZrCo_2$ with $R_A/R_B > 1.225$, anti-site defects form only on the Co-rich side and second phase particles form on the Zr-rich side.
2. No constitutional vacancies but only thermal vacancies were detected in the Laves phase alloys quenched from high temperatures. The peak vacancy concentration is always located at the stoichiometric composition of each Laves phase alloy.
3. Room-temperature fracture toughness can not be effectively improved either by increasing thermal vacancy concentration at levels below 1% or by reducing stacking fault energy through control of the phase stability in $Nb(Cr,Co)_2$ and $Nb(Cr,Fe)_2$ alloys.
4. Drop casting and heat treatments induce micro- and macro-cracks in dual-phase Cr–$NbCr_2$ alloys but not in Cr–$TaCr_2$, and Cr–$ZrCr_2$ alloys. The cracking behavior can not be simply explained from CTE mismatch; instead, it can be better rationalized from the consideration of eutectic composition and phase relationship.
5. Among the three dual-phase Cr–XCr_2 alloys (where X = Nb, Ta, or Zr), Cr–$TaCr_2$ alloys have the best combination of mechanical and metallurgical properties for structural use at elevated temperatures. These alloys have excellent tensile strength, creep properties, and oxidation resistance at elevated temperatures in air.
6. The alloy design effort has led to the identification of alloy compositions (at%) with enhanced mechanical and metallurgical properties:

$$\mathrm{Cr} - (6-10)\mathrm{Ta} - (3-6)\mathrm{Mo} - (0.2-1.0)\mathrm{Ti} - (0.5-3.0)\mathrm{Si} - (0.05-0.20)\mathrm{La/Ce}.$$

7. The current Cr–$TaCr_2$ alloys showed fracture toughness up to 14.3 M Pa $m^{-1/2}$ at room temperature. Additional development is required to further improve the fracture resistance of these alloys by interfacial design and microalloying.

Acknowledgements

This research was sponsored jointly by the Division of Materials Sciences and Engineering, and the Fossil Energy AR&TD Materials Program, US Department of Energy under contract DE-AC05-00OR22725 with UT-Battelle, LLC. The authors would like to thank L. Heatherly and D. S. Easton for hot extrusion, J. L. Wright for mechanical testing, and C. A. Carmichael for alloy preparation.

References

[1] Laves F. Theory of alloy phases. Metals Park, OH: American Society for Metals, 1956. p. 124
[2] Wernick JH. In: Westbrook JH, editor. Intermetallic compounds. New York: Wiley, 1967. p. 197.
[3] Sinha AK. Progr Mater Sci 1972;15(2):79.
[4] Massalski TB. In: Cahn RW, Hassen P, editors. Physical metallurgy part 1. New York: North-Holland, 1983. p. 190.
[5] Laves F, Witte H. Metallwirschaft 1935;14:645.
[6] Laves F, Witte H. Metallwirschaft 1936;15:840.
[7] Moffett MB, Clark AE, Wun-Fogle M, Lindberg J, Teter JB, McLaughlin EA. J Acoust Soc Am 1991;89:1448.
[8] Inoue K, Tachikawa K. IEEE Trans Magnetics 1979;15:635.
[9] Olzi E, Matacotta FC, Setina P. J Less-Common Met 1988;139:123.
[10] Ivey D, Northwood D. J Less-Common Met 1986;115:23.
[11] Halstead A, Rawlings RD. J. Mater. Sci. 1985;20:1248 & 1693
[12] Anton DL, Shah DM. Mater Res Soc Symp Proc 1990;194:45.

[13] Livingston JD. Phys Status Solidi (a) 1992;131:415.
[14] Chu F, Pope DP. Mater Sci Engng 1993;A170:39.
[15] Kumar KS, Miracle DN. Intermetallics 1994;2:257.
[16] Thoma DJ, Perepezko JH. Mater Sci Engng 1992;A156:97.
[17] Takeyama M, Liu CT. Mater Sci Engng 1991;A132:61.
[18] Chan KS, Davidson DL. JOM 1996;48:62.
[19] Fleischer RL, Zabala RJ. Metall Trans 1990;21A:2149.
[20] Chen KC, Allen SM, Livingston JD. J Mater Res 1997;12:1472.
[21] Hazzledine PM, Kumar KS, Miracle DB, Jackson AG. MRS Symp Proc 1992;288:591.
[22] Liu CT, Tortorelli PF, Horton JA, Carmichael CA. Mater Sci Engng 1996;A214:23.
[23] Kumar KS, Liu CT. Acta Mater 1997;45:3671.
[24] Takasuqi T, Kumar KS, Liu CT, Lee EH. Mater Sci Engng 1999;A260:108.
[25] Kumar KS, Pang L, Liu CT, Hoton JA, Kenik EA. Acta Mater 2000;48:911.
[26] Brady MP, Zhu JH, Liu CT, Tortorelli PF, Walker LR, McKamey CG, Wright JL, Carmichael CA, Larson DJ, Miller MK, Porter WD. Materials at High Temperatures 1999;16(4):189–93.
[27] Thoma DJ, Perepezko JH. J Alloys and Compounds 1995;224:330.
[28] Chen KC, Allen SM, Livingston JD. MRS Symp Proc 1994;364:1406.
[29] Ivey DG, Northwood DO. Z Phys Chem NF 1986;147:191.
[30] Sicking G, Magomedbekov E, Hempelmann R. Ber Bunsenges Phys Chem 1981;85:686.
[31] Zhu JH, Pike LM, Liu CT, Liaw PK. Acta Mater 1999;47:2003.
[32] Zhu JH, Liu CT. Defect structures in $ZrCo_2$ Laves phase, Acta Mater. 2000 (in press).
[33] Modder IW, Bakker H. Phys Stat Sol (b) 1997;199:369.
[34] Modder IW, Kuin MJ, Bakker H. Intermetallics 1998;6:537.
[35] Pike LM, Chang YA, Liu CT. Acta Mater 1997;45:3709.
[36] Hazzledine PM, Pirouz P. Scripta Metall Mater 1993;28:1277.
[37] Neumann JP. Acta Metall 1980;28:1165.
[38] de Boer FR, Boom R, Mattens WCM, Miedema AR, Niessen AK. Cohesion in metals and transition metal alloys, ed. de Boer FR, Pettifor DG. Vol. 1, Second Corrected Printing, North Holland, Amsterdam, 1989.
[39] Zhu JH, Liaw PK, Liu CT. Mater Sci Engng 1997;A239(240):260.
[40] Zhu JH, Liu CT, Liaw PK. Intermetallics 1999;7:1011.
[41] Bardos DI, Gupta KP, Beck PA. Trans Met Soc AIME 1961;221:1087.
[42] Massalski TB, ed. Binary alloy phase diagrams, vol. 1, ASM Publication, 1986.
[43] Thoma DJ. Unpublished results, Lawrence Livermore National Laboratory, 1998.
[44] Liu CT. Unpublished results, Oak Ridge National Laboratory, 1999.
[45] Suklyun Hong, Fu CL. in MRS Proc. vol. 552, MRS Publication, 1999:643–647.
[46] Lee S, Liaw PK, Liu CT, Chou YT. Mater Sci Engng A 1999;268:184–92.
[47] Lee S, Liaw PK, Liu CT, Chou YT. Mater Chem Phys 1999;61:207–13.
[48] Jenkins MG, Kobayashi AS, White KW, Bradt RC. Inst J Fract 1987;34:281.
[49] Schneibel JH, Carmichael CA, Specht ED, Subramanian R. Intermetallics 1997;5:61–7.
[50] Brady MP, Zhu JH, Liu CT, Tortorelli PF, Wright JL, Carmichael CA, Walker LR. Ultrahigh temperature intermetallic alloys. Proceedings of the Eleventh Annual Conference on Fossil Energy Materials, Knoxville, TN, 20–22 May, R.R. Judkins, Compiler, 1997. pp. 195–205.
[51] Brady MP, Tortorelli PF, Walker LR. Materials at high temperatures (in press).

PERGAMON

Acta mater. 48 (2000) 2339–2347

Acta MATERIALIA

www.elsevier.com/locate/actamat

DEFECT STRUCTURES IN $ZrCo_2$ LAVES PHASE

J. H. ZHU† and C. T. LIU

Oak Ridge National Laboratory, Metals and Ceramics Division, Bldg. 4500S-MS6115, PO Box 2008, Oak Ridge, TN 37831-6115, USA

(Received 6 December 1999; accepted 27 December 1999)

Abstract—Point defect structures in the C15 $ZrCo_2$ alloys were studied by bulk density and X-ray lattice parameter measurements. It was found that, for the $ZrCo_2$ alloys quenched from 1000°C, the lattice parameter increases linearly as the Zr content increases up to 33.3 at.% Zr. The lattice parameter of the Laves phase remains constant for the alloys with Zr content higher than 33.3 at.%, indicating that the solubility range of Zr in $ZrCo_2$ on the Zr-rich side is essentially zero. The constitutional defects were found to be of the anti-site type. Thermal vacancies exhibiting a maximum at the stoichiometric composition were observed in the $ZrCo_2$ Laves phase alloys after quenching from 1250 and 1000°C, with higher thermal vacancies obtained from 1250°C. The defect structures in the $ZrCo_2$ phase may be correlated to the relative magnitude of formation enthalpies for anti-site and quadruple defects in this compound. Thermal vacancy concentration at a level of 1% in $ZrCo_2$ does not affect fracture toughness at room temperature. *Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.*

Keywords: Intermetallic compound; Casting; X-ray diffraction; Lattice defects; Mechanical properties (fracture & fracture toughness)

1. INTRODUCTION

Topologically close packed Laves phases with AB_2 composition are known to be size compounds, i.e., the atomic size ratio, R_A/R_B, is ideally 1.225, with a range of 1.05–1.67 typically observed [1, 2]. Laves phases are conventionally considered to be line compounds with a strict AB_2 composition. However, about 25% of the binary Laves phases show certain homogeneity ranges on the A- and/or B-rich side of stoichiometry [3]. Non-stoichiometric intermetallic compounds are stabilized by the incorporation of constitutional defects. The excess atoms in off-stoichiometric compounds may prefer to (1) stay on their own sublattice, leading to the formation of constitutional vacancies on the other sublattice, (2) insert into interstitial sites, or (3) occupy the sites on the other sublattice, leading to the formation of substitutional anti-site defects. Since Laves phases have topologically close packed (TCP) structures and space-filling is relatively high, there are no interstitial sites with a size comparable to that of the component atoms. Therefore, the insertion of the excess atoms into the interstitial sites can be excluded, and the possible defect mechanisms in binary Laves phases are thus constitutional vacancy or anti-site substitution.

The defect structures in cubic C15 $NbCr_2$, $NbCo_2$ and hexagonal C14 $NbFe_2$ Laves phases have recently been determined to be of anti-site type on both sides of stoichiometry [4]. Contrary to the geometric considerations which support the constitutional vacancy mechanism on the A-rich side of stoichiometry, anti-site substitution is found to be the constitutional defect mechanism on the Nb-rich side of these Laves phases. These results indicate that except for the geometric factor, other factors such as chemical bonding characteristics may also affect the defect structures in this Laves phases. Based on geometric and thermodynamic considerations, Laves phases with atomic size ratio R_A/R_B larger than 1.225 and/or high negative enthalpy of formation (ΔH) may exhibit constitutional vacancies in the A-rich side, as suggested by Zhu *et al.* [4].

By heating the ordered Laves phases, thermal defects can also be created in the compounds. Modder and his coworkers [5, 6] have systematically analyzed the possible defect type in Laves phases and two types of thermal defects were identified, namely quadruple defects (consisting of one anti-site B atom on the A-atom sublattice and three

† To whom all correspondence should be addressed.

1359-6454/00/$20.00 Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.
PII: S1359-6454(00)00014-8

vacancies on the B-atom sublattice) and anti-site defects. Furthermore, it was found that the defect types in Laves phases can be correlated to the relative magnitudes of the formation enthalpies of anti-site and quadruple defects in the compounds [6].

The point defect structures in the $ZrCo_2$ system have not been investigated previously. The atomic size ratio R_A/R_B of $ZrCo_2$ is 1.28, which is significantly larger than the ideal one, i.e., 1.225. Furthermore, the enthalpy of formation (ΔH) of $ZrCo_2$ is -113 kJ/mole [7], which is much more negative than those of $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves phases. So far, no work has been conducted to carefully determine the homogeneity range and defect structures in the Laves phase. The binary Co–Zr phase diagram shows certain solubilities on both sides of stoichiometry in this Laves phase system [8]. All these facts indicate that the $ZrCo_2$ Laves phase system deserves further attention as a system where constitutional vacancies may be the defect mechanism on the Zr-rich side.

In the present study, we have attempted to determine the solubility range and to model the defect structures in binary $ZrCo_2$ alloys with different stoichiometric compositions, by measuring the lattice parameters, bulk densities, and therefore, vacancy concentrations of binary $ZrCo_2$ alloys of various compositions after quenching from 1000 and 1250°C. Thermal defects that will be introduced by quenching from elevated temperatures have also been determined. The fracture toughness of selected alloys were also measured at room temperature.

2. EXPERIMENTAL PROCEDURES

The alloy compositions investigated in this study are listed in Table 1. A total of seven binary alloys were prepared. The Zr content in the alloy increases with the alloy code number. The alloys ♯1–♯3 are Co-rich and ♯4 is the stoichiometric alloy, while alloys ♯5–♯7 are Zr-rich. Ingots weighing approximately 50 g were prepared by arc-melting techniques using high-purity starting materials. Each ingot was remelted at least five times to ensure compositional homogeneity before being cast into a 12.7 mm diameter cylinder. The weight loss during arc-melting was essentially zero (less than 0.01 g). It should be noted that the ingots of the alloys ♯4–♯7 remained in one piece, while the ingots of the alloys ♯1–♯3 broke into several pieces after drop casting.

After a homogenization anneal at 1200°C for 20 h in vacuum, the alloys were encapsulated in quartz, which was then evacuated and backfilled with helium. The encapsulated alloys were then heat-treated at 1000°C for 7 days and water quenched. Small fractions of the quenched samples were mounted for microstructural observation using an optical microscope and for chemical composition analyses of the phases present using a field-emission-gun scanning electron microscope (SEM) equipped with an energy dispersive X-ray spectrometer (EDS). All the alloys were found to be fully homogenized following the above heat treatments.

The remaining parts of the samples were crushed in an agate mortar into powders with particle sizes smaller than 40 mesh. Powders with a mesh size of $-40+80$ were used for bulk density measurements, and powders smaller than 320 mesh were used for X-ray analyses. The bulk density was measured using a helium pycnometer with an accuracy of approximately 0.01 vol.%. Powdered samples were used to eliminate the micropores generated during solidification, which leads to more accurate density measurements. The vacancy concentration of the Laves-phase alloys, c_V, is defined as the ratio of the total number of vacancies to the total number of atoms. It can be obtained from the following equation:

$$c_V = \frac{\rho_X - \rho_B}{\rho_B} \tag{1}$$

where ρ_X and ρ_B are the X-ray and the bulk densities, respectively.

The powdered samples of the single-phase alloys ♯1–♯4 were then wrapped in Ta sheets and encapsulated in quartz, which was evacuated and backfilled with helium. They were then heat-treated at 1250°C for 2 h and water quenched. Bulk densities and lattice parameters were measured again and the

Table 1. Composition, lattice parameter, X-ray density, bulk density and vacancy concentration of $ZrCo_2$ alloys quenched from 1250 and 1000°C

Alloy code	Zr (at.%)	Lattice parameter (Å)		X-ray density (g/cm^3)		Bulk density (g/cm^3)		Vacancy concentration (%)	
		1250°C	1000°C	1250°C	1000°C	1250°C	1000°C	1250°C	1000°C
#1	30.8	6.9305	6.9311	8.242	8.240	8.194	8.224	0.58	0.20
#2	31.8	6.9419	6.9430	8.240	8.235	8.192	8.210	0.59	0.31
#3	32.8	6.9543	6.9546	8.234	8.233	8.177	8.203	0.70	0.37
#4	33.3	6.9590	6.9608	8.236	8.230	8.155	8.196	1.00	0.42
#5	33.5	–	6.9609[a]	–	–	–	–	–	–
#6	33.8	–	6.9607[a]	–	–	–	–	–	–
#7	34.3	–	6.9609[a]	–	–	–	–	–	–

[a] Two-phase alloys with the lattice parameter of the $ZrCo_2$ Laves phase listed here.

vacancy concentrations were obtained for these alloys quenched from 1250°C.

X-ray diffraction measurements were made using a Scintag XD2000 diffractometer with Cu K_α radiation. The lattice parameter was determined from 13–15 peaks. The X-ray density was calculated using the measured lattice parameter. The accuracy of the vacancy concentration using our method is approximately 0.08%, which is mainly due to the compositional uncertainty of the alloys.

The fracture toughness of the selected alloys was also determined using the indentation method with a load of 500 g and a holding time of 15 s. The crack lengths were measured immediately after the indentation to eliminate the possibility of slow crack propagation following the removal of the indenter. The following equation was used to calculate the fracture toughness of the alloys [9]:

$$K_{IC} = A\left[\frac{E}{H}\right]^n \frac{P}{L^{3/2}} \quad (2)$$

where K_{IC} is the fracture toughness (MPa$\sqrt{m}$), E is Young's modulus (GPa), which is assumed to be 218 GPa for all of the investigated alloys, H is Vickers hardness (GPa), P is the load (N), and L the average length of the four radial cracks from the center of the indent to the crack tip. A and n are constants, which are taken as 0.016 and 0.5, respectively, for relatively brittle materials [9].

3. RESULTS

3.1. Microstructure

Microstructure examinations of the $ZrCo_2$ Laves phase alloys listed in Table 1 after quenching from 1000°C indicate that alloys ♯1–♯4 are in the single-phase region, while alloys ♯5–♯7 are in the two-phase region. Typical microstructures of single-phase and two-phase alloys are shown in Figs 1(a) and (b). Note that Fig. 1(a) was taken from the Co–33.3%Zr sample (alloy ♯4) etched with a solution of 60% glycerine + 20% HNO_3 + 10% HF + 10% H_2O (in vol.%), while Fig. 1(b) was from the Co–34.3%Zr sample (alloy ♯7) in the as-polished condition under polarized light. There are no annealing twins in the alloys, unlike the $NbCr_2$ system [4], indicating relatively high stacking fault energy in this compound. The second phase in alloys ♯5–♯7 was analyzed with EDS in an SEM, and was found to be richer in Zr than the $ZrCo_2$ matrix. According to the binary Co–Zr phase diagram [8], this second phase is a ζ phase with a B2 structure. The maximum content of Zr in $ZrCo_2$ at 1000°C is found to be about 33.3 at.% from this study, which is different from the estimated phase boundary in the binary Co–Zr phase diagram [8, 10], where it was indicated to be about 34.5 at.%. Furthermore, extensive cracking was observed in the matrix around the second phase. These cracks are most likely caused by the thermal residual stress in the alloys after casting and heat treatment, as a result of the mismatch in thermal expansion coefficients between the second phase and the matrix.

3.2. Lattice parameter, density, and vacancy concentration

The lattice parameters of the C15 $ZrCo_2$ alloys listed in Table 1 are plotted in Fig. 2 as a function of stoichiometry for all the samples quenched from 1000°C and for the single-phase alloys quenched from 1250°C. As the Zr content increases, the lattice parameter of the $ZrCo_2$ phase also increases. For the Co-rich alloys quenched from 1250 and 1000°C, the data points fit nicely along two straight lines, while on the Zr-rich side the lattice constants are identical for alloys ♯5–♯7 quenched from

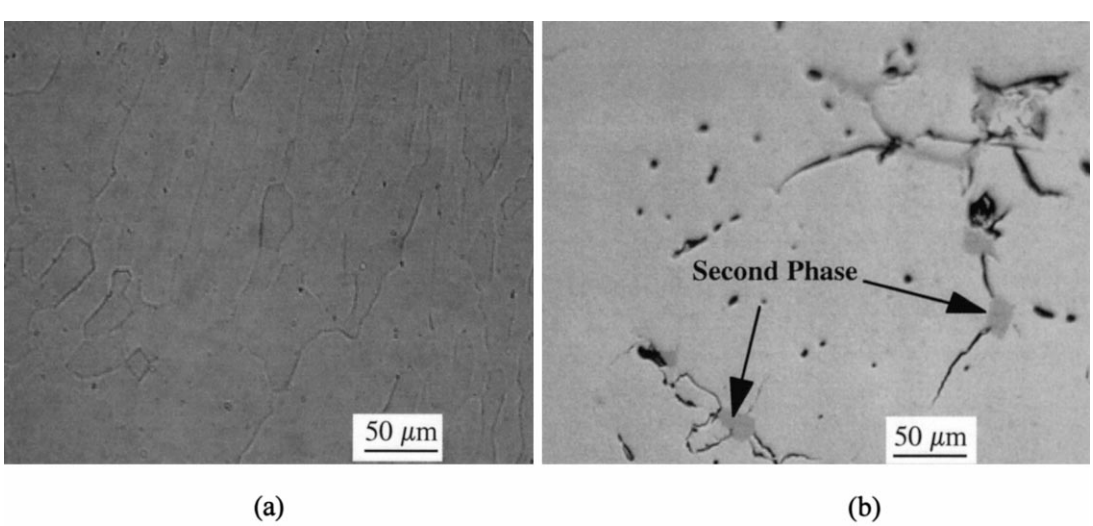


Fig. 1. Optical micrographs of $ZrCo_2$ alloys: (a) alloy ♯4, etched; (b) alloy ♯7, as-polished.

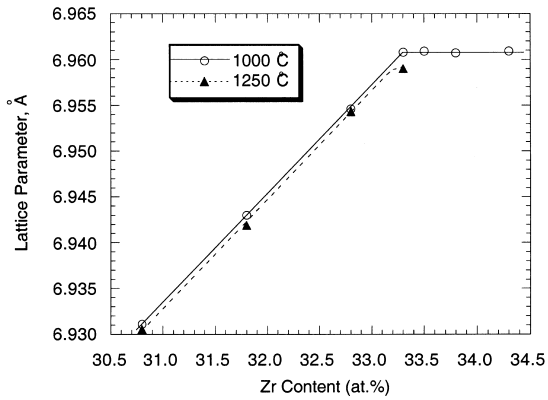


Fig. 2. Lattice parameter as a function of Zr content after quenching from 1250 and 1000°C.

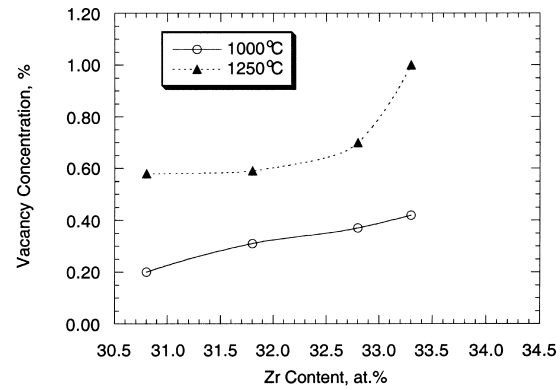


Fig. 4. Vacancy concentration as a function of Zr content after quenching from 1250 and 1000°C.

1000°C, indicating that the composition of the Laves phase is the same in the alloys with more than 33.3% Zr (see Fig. 2). This is expected, because the maximum content of Zr in $ZrCo_2$ is about 33.3 at.% and these three alloys were found to be in the two-phase region. The lattice parameters after quenching from 1250°C are a little bit smaller, compared to those after quenching from 1000°C. This can be due to the presence of vacancies in the alloy.

The X-ray and bulk density results for alloys ♯1–♯4 after quenching from both 1250 and 1000°C are presented in Table 1 and Fig. 3. The density of the alloys decreases as the Zr content increases, and the bulk density after quenching from 1000°C is consistently higher than those after quenching from 1250°C. Using equation (1) and the data presented in Table 1 and Fig. 3, the vacancy concentrations were calculated, and are listed in Table 1. Due to the presence of the second phase, the vacancy concentration could not be determined for alloys ♯5–♯7. Figure 4 is a plot of the vacancy concentration as a function of the Zr content for the single-phase $ZrCo_2$ alloys after quenching from the two different temperatures. It is noted that the vacancy concentration is quite high after quenching from 1250°C. The vacancy concentration reaches a maximum of ~1.0% at the stoichiometric composition and decreases on the Co-rich side of stoichiometry. Annealing and quenching from 1000°C results in vacancy concentrations lower than those after quenching from 1250°C. However, a maximum at the stoichiometric composition can also be recognized.

3.3. Fracture toughness

The fracture toughness values of the single-phase $ZrCo_2$ alloys after quenching from 1000°C were measured as a function of Zr content (see Fig. 5). It was found that all the single-phase $ZrCo_2$ alloys exhibit a fracture toughness value of around 1 MPa

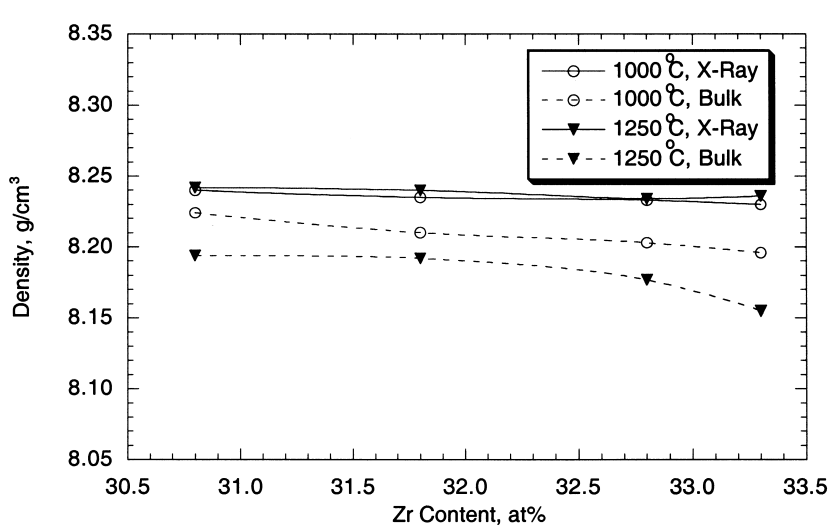


Fig. 3. X-ray and bulk densities as a function of Zr content after quenching from 1250 and 1000°C.

$m^{1/2}$, regardless of the alloy stoichiometry. Furthermore, the fracture toughness value of the stoichiometric $ZrCo_2$ alloy after quenching from 1250°C was found to be close to that of the same alloy after quenching from 1000°C, even though the vacancy concentration in the alloy is much higher after quenching from 1250°C. This indicates that the presence of such vacancies does not lead to toughness improvement in this Laves phase.

4. DISCUSSION

4.1. Constitutional defect mechanisms in $ZrCo_2$ Laves phase

The vacancy concentration as a function of stoichiometry for binary Laves phases has been calculated based on the two constitutional defect models, i.e., the constitutional vacancy model and anti-site substitution model [4]. The constitutional vacancy model is based on the assumption that all the excess atoms exclusively stay on their own sublattice, thus constitutional vacancies are created on the sublattice of the other element. Based on this model, the constitutional vacancy concentrations in Laves phases are zero at the stoichiometric composition and increase on both sides of stoichiometry. The anti-site substitutional model assumes that the excess atoms occupy the sublattice sites of the other species. Thus, anti-site defects are created, and no constitutional vacancies are needed for the balance of the lattice sites. The constitutional vacancy concentrations are zero for both A- and B-rich Laves phases, according to the anti-site substitutional model.

From Fig. 4, it is obvious that the measured vacancy concentrations for all the single-phase $ZrCo_2$ alloys after quenching from 1000°C are relatively low (see also Table 1), indicating that there are no constitutional vacancies in these Laves phases. Furthermore, as the composition deviates further from the stoichiometry, the vacancy concentration on the Co-rich side decreases, while the constitutional vacancy model predicts an opposite trend. Actually, the measured vacancy concentration as a function of stoichiometry can be rationalized by assuming that these vacancies are thermal vacancies instead of constitutional vacancies (see next section for more details). This is consistent with the anti-site substitution model and contrary to the constitutional vacancy model, which demands much higher vacancy concentrations for off-stoichiometric compositions. Thus, the constitutional defects are of anti-site type for the Co-rich $ZrCo_2$ Laves phases.

Point defect mechanisms in binary Laves phases have been the focus of several studies. Fleischer [11] found that anti-site substitution occurs on both sides of stoichiometry for the C15 $ZrCr_2$ Laves phase. Pargeter and Hume-Rothery [12] and Saito and Beck [13] investigated the point defects in Co-rich $NbCo_2$ and dihexagonal C36 $NbCo_3$ Laves phases and found that in both phases, the deviation from stoichiometry resulted from the substitution of Co atoms for the Nb atoms. X-ray and neutron diffraction measurements [14] showed that for both $ZrFe_2$ and $TiFe_2$, the excess Fe atoms are on the Zr or Ti sublattice. The constitutional defect structure of the $NbFe_2$ Laves phase was also investigated by measuring the lattice parameters on both sides of stoichiometry, and again anti-site substitution was found to be the defect mechanism [15]. YAl_2 Laves phase is a line compound with negligible solubility existing under equilibrium conditions; however, Foley *et al.* extended the solubility on both sides of stoichiometry in this alloy system via rapid solidification. From lattice parameter measurements, off-stoichiometry in the YAl_2 Laves phase was postulated to be accompanied by different defect mechan-

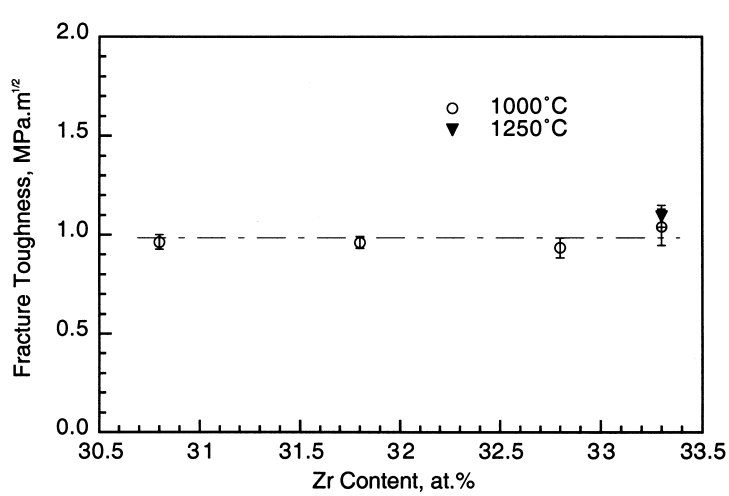


Fig. 5. Fracture toughness vs Zr content for $ZrCo_2$ alloys after quenching from 1250 and 1000°C.

isms: the Y-rich side is accommodated by vacancies, while the Al-rich compositions result from anti-site substitution [16]. Recent studies combining bulk density and X-ray lattice parameter measurements unambiguously indicate that the constitutional defect mechanisms in binary C15 $NbCr_2$, $NbCo_2$, and C14 $NbFe_2$ compounds on both sides of stoichiometry are of anti-site type, with no constitutional vacancies present in the compounds [4].

Table 2 summarizes the atomic size ratio R_A/R_B, enthalpy of formation and point defect structures for a number of binary Laves phases [17]. Based on geometric considerations, the atomic size ratio R_A/R_B of the AB_2 Laves phases should be important in controlling the point defect structures in Laves phases. For Laves phases with R_A/R_B greater than the ideal ratio of 1.225, on the A-rich side, it will be more difficult for A atoms to stay on the sublattice sites of B atoms as the atomic size ratio increases. Therefore, it would be possible to obtain constitutional vacancies for Laves phases with $R_A/R_B > 1.225$ on the A-rich side, even though substitutional anti-site defects are most likely to form under all the other circumstances. In all the AB_2 Laves phases whose point defect structure on the A-rich side has been studied so far and listed in Table 2, only YAl_2 has an R_A/R_B ratio of 1.258, higher than the ideal one, and this alloy system also shows certain indications of constitutional vacancies in the Y-rich compounds [16]. For the $ZrFe_2$ alloys with the R_A/R_B ratio of 1.257, the defect structure on the Zr-rich side has not been studied yet. For the $ZrCo_2$ compound, R_A/R_B is 1.280, much larger than the ideal. Therefore, the possibility is expected to be higher for constitutional vacancies existing in this system.

On the other hand, Neumann [18] found that for B2 compounds their constitutional defect structure is closely related to the magnitude of enthalpy of formation. Constitutional vacancies are the preferred defect mechanism for B2 compounds that are more strongly ordered with more negative enthalpy of formation (ΔH) values on the large atom-rich side of stoichiometry, while B2 compounds with ΔH higher than −75 to −90 kJ/mole generally exhibit the anti-site defect structure. Similar thermodynamic arguments may hold for binary Laves phases. As listed in Table 2, anti-site defects are observed on the large atom-rich side for those Laves phases with ΔH value higher than −75 to −90 kJ/mole such as $NbCr_2$, $NbCo_2$, and $NbFe_2$. For Laves phases with more negative ΔH values, such as YAl_2 with a ΔH value of −160.4 kJ/mole [17], constitutional vacancies are the possible defect mechanism on the Y-rich side. In this regard, $ZrCo_2$ also has a highly negative ΔH value of −123 kJ/mole, and thus constitutional vacancies may form on the Zr-rich side.

Based on both geometric and thermodynamic considerations, therefore, constitutional vacancies may be created on the Zr-rich side of $ZrCo_2$, similar to the case of YAl_2. Unfortunately, this study indicates that the solubility range of Zr is essentially zero in $ZrCo_2$ on the Zr-rich side, which makes it difficult to verify our hypothesis regarding the effects of atomic size ratio and enthalpy of formation on the defect mechanisms in binary Laves phases. Rapid solidification techniques could be utilized in future investigations to extend the solubility range on the Zr-rich side in this compound and defect structures can then be deduced from the lattice parameter measurement, as in the case of the YAl_2 phase [16]. It would also be interesting to study defect structures in other Laves phases with $R_A/R_B > 1.225$ and/or highly negative enthalpy of formation to consider whether it is possible to introduce constitutional vacancies on the A-rich side in these compounds.

The fact that both YAl_2 and $ZrCo_2$ compounds are unstable on the A-rich side of stoichiometry under equilibrium conditions indicates that the creation of constitutional vacancies may make the compound energetically less stable. This may explain the lack of solubility on the A-rich side of these compounds. Furthermore, anti-site defects may also be impossible to form on the A-rich side in these compounds, due to the difficulties involved in A-atoms occupying the much smaller B-atom sublattice sites. Consequently, it may be impossible or at least very difficult to obtain constitutional defects in these Laves phases.

4.2. Thermal defects in $ZrCo_2$ Laves phase

As Modder and Bakker have pointed out, it is very possible that a vacancy-type defect termed

Table 2. Atomic size ratio, enthalpy of formation and point defect structure of a number of binary Laves phases

Compound	R_A/R_B[a]	ΔH (kJ/mole)[a]	Defect structure
$NbCr_2$	1.145	−21	Anti-site substitution on both sides
$NbCo_2$	1.173	−51	Anti-site substitution on both sides
$NbFe_2$	1.152	−62	Anti-site substitution on both sides
$ZrFe_2$	1.257	−75	Fe-rich side: anti-site substitution
$TiFe_2$	1.148	−84	Fe-rich side: anti-site substitution
YAl_2	1.258	−160.4	Al-rich side: anti-site substitution Y-rich side: constitutional vacancy
$ZrCo_2$	1.280	−123	Co-rich side: anti-site substitution Zr-rich side: unknown

[a] The R_A/R_B and ΔH values for these compounds are from Ref. [17].

quadruple defect occurs in C15 compounds [5]. Furthermore, ball milling was found to introduce quadruple defects in $GdAl_2$, $GdPt_2$, $GdIr_2$ and $GdRh_2$, while anti-site defects were created in $GdMg_2$ [6]. Since ball milling usually leads to the formation of the same type of atomic defects as heating, thermal defects such as quadruple defects or anti-site defects may be created upon quenching the Laves phases from elevated temperatures. Since each quadruple defect contains three vacancies, thermal vacancies are expected if quadruple defects are created.

The vacancy concentrations in $ZrCo_2$ alloys are quite high after quenching from 1250°C and the vacancy concentrations are significantly reduced after annealing at lower temperature, i.e., 1000°C (Fig. 4). Since the constitutional defects in this compound have been established to be of the anti-site type, these vacancies are obviously thermal vacancies. The fact that the vacancy concentration increases as the quenching temperature increases is a further indication of the thermal nature of these vacancies. Thus, the quadruple defects may be formed in this compound by heating. According to the quadruple defect model and considering the difficulties involved in Zr sitting on the small atom Co sites, Co will go to the Zr sublattice and thus the thermal vacancies are on the Co sublattice sites. Furthermore, the measured vacancy concentration reaches a peak at the stoichiometric composition and decreases on deviating from stoichiometry for the $ZrCo_2$ alloys when quenched from 1250°C and 1000°C (Fig. 4). This behavior is similar to the $NbCr_2$ system [9], where the thermal vacancy concentration exhibits a maximum at the stoichiometric composition. As in the $NbCr_2$ phase, a maximum in thermal vacancy concentration at the stoichiometric composition is rationalized by considering that the increase in entropy associated with the introduction of vacancies will be the greatest at the stoichiometric composition, where the degree of order is necessarily the highest.

Modder *et al.* [6] have extended Miedema's model to estimate the formation enthalpy of quadruple and anti-site defects in C15 Laves phases. According to this model, the type of atomic disorder is related to the relative magnitudes of formation enthalpies of anti-site and quadruple-defect disorder for the Laves phases. Therefore, the presence of thermal vacancies in $ZrCo_2$ may imply that the formation enthalpies of anti-site disorder is larger in magnitude than that of quadruple defect disorder for $ZrCo_2$. Furthermore, the higher vacancy concentration of $ZrCo_2$ after quenching, compared to $NbCr_2$, suggests that the formation enthalpy of quadruple defects in $ZrCo_2$ should be lower than that in $NbCr_2$, which leads to a higher amount of thermal defects in $ZrCo_2$ phase.

It has been suggested that vacancies in Laves phases may assist the synchroshear deformation mechanism and thus increase the toughness of the Laves alloys [19]. Even though thermal vacancies are relatively high (~1%) in the $ZrCo_2$ alloy after quenching from 1250°C, it was found that such quenched-in thermal vacancies do not affect the fracture toughness of the alloy. It will be of interest to determine the effect of constitutional vacancies on crack propagation and fracture toughness behavior of Laves phases if certain Laves phases with constitutional vacancies can be identified.

4.3. Reconstruction of the binary Co–Zr phase diagram

The existing Co–Zr phase diagram determined around 1964 by Pechin *et al.* [10] has not been updated so far. Those authors found that a homogeneity range of 28.4–34.6 at.% Zr exists for the $ZrCo_2$ Laves phase after rapidly quenching from 1400°C. In other words, there is 1.3 at.% Zr solubility on the Zr-rich side. Based on the present study, the solubility of Zr on the Zr-rich side is zero at 1000°C. This discrepancy may possibly be a result of the high impurity level in the materials used in the study by Pechin *et al.*, since they used the powder metallurgy (PM) method to produce the Laves phases, while in our study the arc-melting technique was employed. It is well known that impurities such as oxygen can easily be picked up during PM processing. The reaction of oxygen with Zr and the formation of ZrO_2 lower the Zr concentration in the Laves phase. Indeed, many voids and impurity particles (possibly ZrO_2) could be observed in the micrograph of the PM Laves phases [10]. Another possibility for the discrepancy may result from the fact that the quenching temperature is different for the two studies, with a much higher quenching temperature (1400°C) used in Pechin's study compared to 1000°C in the present study. The higher equilibrium temperature may increase the solubility range on the Zr-rich side of the compound.

In our view, this discrepancy is most likely due to the inaccurate composition control in Pechin *et al.*'s study, as is clear from the following argument. The slope of the lattice parameter as a function of stoichiometry for the Laves phase should usually change at the stoichiometric composition, as has been commonly observed for other Laves phases (for example, see Ref. [4]). According to Pechin *et al.* [10], the slope of the lattice parameter is the same on both sides of stoichiometry for $ZrCo_2$, suggesting that all the alloy compositions they prepared could be on one (Co-rich) side of stoichiometry, as a result of problems in material processing such as impurity pick-up. As has been demonstrated in the present study for the alloys quenched at 1000°C, negligible Zr solubility is observed for this Laves phase. Based on these arguments, we have refined the binary Co–Zr phase dia-

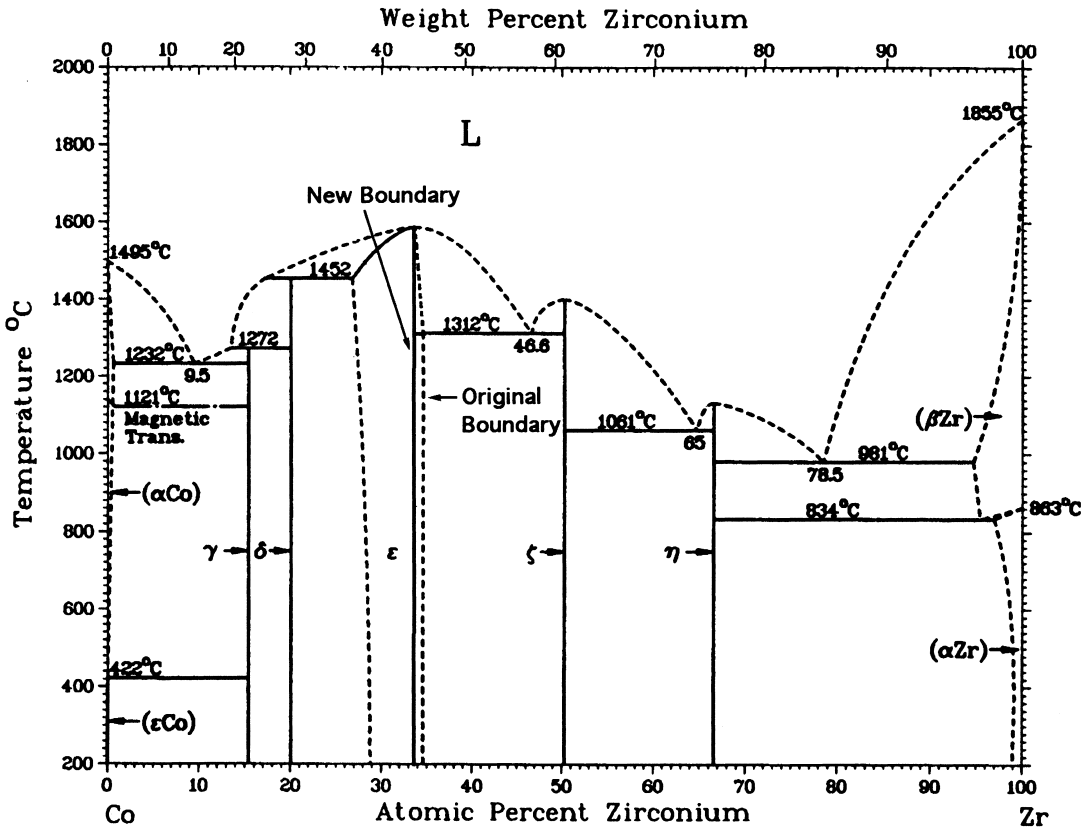


Fig. 6. The modified binary Co–Zr phase diagram.

gram, as shown in Fig. 6, where the solubility range of the Laves phase $ZrCo_2$ has been modified according to this study.

5. CONCLUSIONS

Point defects in the binary $ZrCo_2$ alloys have been clarified, combining both bulk density and X-ray lattice parameter measurements. The constitutional defects on the Co-rich side of stoichiometry for this system are of the anti-site type. As the Zr content in the Laves phase increases, the lattice parameters on the Co-rich side increase. The maximum content of Zr in the $ZrCo_2$ compound is the stoichiometric composition, 33.3 at.%. The binary Co–Zr phase diagram has been modified accordingly.

Thermal vacancies exhibiting a maximum at the stoichiometric composition were noted in $ZrCo_2$ Laves phase alloys after quenching from both 1000 and 1250°C, indicating that quadruple defects may be the type of thermal disorder in this compound. These thermal vacancies do not affect the fracture toughness of the Laves phase.

Acknowledgements—This research was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. This research was supported in part by an appointment of J. H. Zhu to the Oak Ridge National Laboratory Postdoctoral Research Associates Program administered jointly by the Oak Ridge Institute for Science and Education and Oak Ridge National Laboratory (ORNL). Thanks are also due to S. R. Agnew and C. G. McKamey of ORNL for reviewing this manuscript.

REFERENCES

1. Laves, F., in *Theory of Alloy Phases*. American Society for Metals, Metals Park, OH, 1956, p. 124.
2. Zhu, J. H., Liu, C. T., Pike, L. M. and Liaw, L. M., *Metall. Mater. Trans.*, 1999, **30A**, 1449.
3. Thoma, D. J. and Perepezko, J. H., *J. Alloys Compounds*, 1995, **224**, 330.
4. Zhu, J. H., Pike, L. M., Liu, C. T. and Liaw, L. M., *Acta mater.*, 1999, **47**, 2003.
5. Modder, I. W. and Bakker, H., *Phys. Stat. Sol. (b)*, 1997, **199**, 369.
6. Modder, I. W., Kuin, M. J. and Bakker, H., *Intermetallics*, 1998, **6**, 537.
7. Klein, R., O'Hare, P. A. G. and Jacob, I., *J. Alloys Compounds*, 1997, **261**, 204.
8. Massalski, T. B., Murray, J. L., Bennett, L. H. and Baker, H., *Binary Alloy Phase Diagrams*. American Society for Metals, Metals Park, OH, 1986.
9. Sakai, M. and Bradt, R. C., *Int. Mater. Rev.*, 1993, **38**(2), 53.

10. Pechin, W. H., Williams, D. E. and Larsen, W. L., *Trans. ASM*, 1964, **57**, 464.
11. Fleischer, R. L., *Scripta metall. mater.*, 1992, **27**, 799.
12. Pargeter, J. K. and Hume-Ruthery, W., *J. Less-Common Metals*, 1967, **12**, 366.
13. Saito, S. and Beck, P. A., *Trans. Metall. Soc. AIME*, 1960, **218**, 670.
14. Brückner, W., Kleinstück, K. and Schulze, G. E. R., *Phys. Stat. Sol.*, 1967, **23**, 475.
15. Smith, A. W., Rogers, J. A. and Rawlings, R. D., *Phys. Stat. Sol. (a)*, 1973, **15**, K119.
16. Foley, J. C., Thoma, D. J. and Perepezko, J. H., *Metall. Trans.*, 1994, **25A**, 230.
17. Zhu, J. H., Metallurgical factors controlling the phase stability and mechanical behavior of transition-metal Laves phase alloy system. Ph.D. Dissertation, University of Tennessee, Knoxville, TN, August, 1998.
18. Neumann, J. P., *Acta metall.*, 1980, **28**, 1165.
19. Hazzledine, P. M. and Pirouz, P., *Scripta metall. mater.*, 1993, **28**, 1277.

KC 02 01 05 0
ERKCM05 (Science of Non-Equilibrium Processing of Materials)

Refereed Journal Publications

†*Babu, S. S., S. A. David, and M. K. Miller, "Application of Atom Probe Microanalysis for Understanding Microstructure Evolution in Nickel Base Superalloy Welds," *Micros. & Microanal.* **6**, 350-51 (2000).

†*David, S. A., J. M. Vitek, S. S. Babu, L. A. Boatner, and R. W. Reed, "Welding of Nickel-Base Superalloy Single Crystals," *Welding in the World* **44**(2), 38-49 (2000).

†*Hofer, P., M. K. Miller, S. S. Babu, S. A. David, and H. Cerjak, "Atom Probe Field Ion Microscopy Investigation of Boron Containing Martensitic 9% Chromium Steel," *Metallurgical and Materials Transactions A* **31A**, 975-84 (2000).

†Hong, T., T. DebRoy, S. S. Babu, and S. A. David, "Modeling of Inclusion Growth and Dissolution in the Weld Pool," *Metallurgical and Materials Transactions B* **31B**, 161-69 (2000).

†Kozeschnik, E., "A Discussion of Phase Transformations in Fe-C-Mn as Affected by Paraequilibrium Constraints," *J. Phase Equilib.* **21**(4), 336-41 (2000).

†Kozeschnik, E., "Assessment of Diffusional Mobilities of Copper and Zinc in Alpha-Brass," *Z Metallkd* **91**(1), 57-63 (2000).

†Kozeschnik, E., "Multicomponent Diffusion Simulation Based on Finite Elements," *Metallurgical and Materials Transactions A* **30**(10), 2575-82 (1999).

Sarma, G., B. Radhakrishnan, and T. Zacharia, "Modeling the Deformation of Face Centered Cubic Crystals to Study the Effect of Slip on {110} Planes," *Modeling Simul. Mater. Sci. Eng*. **7**, 1025 (1999).

†Vitek, J. M., Y. S. Iskander, and E. M. Oblow, "Improved Ferrite Number Prediction in Stainless Steel Arc Welds Using Artificial Neural Networks—Part 1: Neural Network Development," *Weld. J.* **79**(1), 33s-40s (2000).

†Vitek, J. M., Y. S. Iskander, and E. M. Oblow, "Improved Ferrite Number Prediction in Stainless Steel Arc Welds Using Artificial Neural Networks—Part 2: Neural Network Results," *Weld. J.* **79**(1), 41s-50s (2000).

*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

Modelling Simul. Mater. Sci. Eng. **7** (1999) 1025–1043. Printed in the UK PII: S0965-0393(99)06695-4

Modelling the deformation of face centred cubic crystals to study the effect of slip on {110} planes

G Sarma, B Radhakrishnan and T Zacharia

Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6359, USA

Received 4 August 1999

Abstract. The deformations of single crystals and polycrystals of aluminium have been modelled using the finite-element method. The constitutive behaviour is modelled using crystal plasticity to account for the plastic deformation by slip, and to track the hardening and reorientation of the material. By discretizing each crystal with a large number of elements, the non-uniform deformations due to local inhomogeneities and interactions with neighbouring crystals are modelled. Simulations of plane strain compression of (001)[110] oriented single crystals are used to demonstrate the ability of the model to capture shearing in the rolling-normal plane, and the consequent reorientation of the crystal to the $\{112\}\langle 111\rangle$ copper components. The simulations are used to examine the effect of including $\{110\}\langle 110\rangle$ slip systems in addition to the usual $\{111\}\langle 110\rangle$ systems for face centred cubic metals on the stability of the cubic orientation. The results indicate that slip on $\{110\}$ planes greatly enhances the stability of the cubic orientation, while having little contribution in the deformation of most other orientations.

1. Introduction

Deformation processing is an important component in the manufacturing of metal products. The changes in texture and microstructure that accompany the deformation process control the material behaviour during subsequent thermomechanical processing. Understanding and predicting the texture and microstructure evolution during processing of metals remains a subject for extensive research. In this regard, the merging of the finite-element method (FEM) as a numerical tool for solving the boundary value problem associated with material motion with crystal plasticity theory as a constitutive description for plastic deformation by slip has developed into a significant modelling capability. Early efforts in this area utilized crystal plasticity as a state variable theory in place of simpler continuum models in finite-element simulations of industrial metal working operations such as rolling, forging, extrusion, etc (e.g. Mathur and Dawson 1989, Mathur *et al* 1990, Kalidindi *et al* 1992, Beaudoin *et al* 1994). Some recent efforts at combining FEM and crystal plasticity to model deformation processing of metals have focused at the mesoscale, by shifting the domain used for discretization from the entire workpiece to a single polycrystalline aggregate. These simulations are aimed at using the modelling capabilities to gain better insights into the role of grain interactions on the non-uniform deformations of metal polycrystals. While some early efforts in this area were restricted to two-dimensional (2D) domains (e.g. Becker 1991, Anand and Kalidindi 1994, Becker and Panchanadeeswaran 1995), more recent efforts have been able to treat fairly large three-dimensional (3D) meshes by taking advantage of parallel computing resources (Beaudoin *et al* 1996, Sarma *et al* 1998b, Mika and Dawson 1998).

© 1999 US Government

Applications of FEM and crystal plasticity to deformation process modelling at the mesoscale proceed by discretizing each grain of a polycrystal with a large number of elements, so that the non-uniform deformations of individual grains can be modelled. Unlike in continuum models, where grains are merely abstract entities with no physical description associated with them, here the grains and their neighbours are explicitly defined by the finite-element model. This feature enables the model to incorporate effects of neighbour interactions on the inhomogeneous deformations of individual grains, and the effect of such inhomogeneities on the texture and microstructure evolution. Additionally, these simulations can provide quantitative information on the distribution of orientations and stored energy in the deformed microstructure, which provides a means to simulate the texture and microstructure evolution during recovery and recrystallization (Radhakrishnan *et al* 1998a, b).

The objective of the work described in this paper is to apply the mesoscale deformation modelling approach to single crystals and polycrystals of aluminium deformed in plane strain compression to study their non-uniform deformations. First, some results are presented on modelling the plane strain compression of a single crystal with initial orientation (001)[110]. This particular orientation has been studied extensively, both experimentally and through modelling (Butler and Hu 1989, Becker *et al* 1991, Akef and Driver 1991, Cizek *et al* 1996) and serves as a good test case to validate the mesoscale modelling approach. It is shown that the results are quite consistent with experimental observations, in terms of changes in both orientation and geometry. These simulations are also used to illustrate that use of 2D models may not always be adequate for plane strain deformations.

As a novel application of the mesoscale modelling approach, simulations have been carried out to examine the effect of slip on $\{110\}\langle 110\rangle$ systems on the stability of the cubic orientation, which is an important component in textures produced during hot rolling and also in recrystallized microstructures. Optimization of cubic texture is one of the key issues in the hot rolling of commercial aluminium alloys for beverage-can applications. Subsequent cold rolling of the hot-band is designed to provide, in addition to the required mechanical properties, an optimum combination of cubic and other deformation components that can minimize material waste due to earing. Therefore, considerable effort has been made in the past to understand texture evolution during thermomechanical processing of aluminium alloys. It is well known that plastic deformation in face centred cubic (fcc) metals is accommodated by shearing of $\{111\}$ planes in $\langle 110\rangle$ directions. This knowledge has been utilized extensively in the application of crystal plasticity models to study the texture development during cold working of aluminium and its alloys. It has been postulated that when aluminium is deformed at elevated temperatures, slip occurs on other planes, notably the $\{100\}$, $\{110\}$ and $\{112\}$ (Bacroix and Jonas 1988, Maurice and Driver 1997a). Slip activity on the non-octahedral systems alters the relative strength of various texture components during deformation in plane strain compression. In particular, the $\{011\}\langle 211\rangle$ brass component is enhanced by slip on the $\{112\}$ planes, and the $\{001\}\langle 100\rangle$ cubic component is stabilized by slip on $\{110\}$ planes.

Previous efforts on modelling the effect of slip on the $\{110\}$ planes (Bacroix and Jonas 1988, Maurice and Driver 1993, 1997b, Raabe 1995) have been through analytical models based on some simplifying assumptions or through numerical simulations using an aggregate of grains. Simulations with an aggregate of grains have relied on various assumptions to partition the deformation among the individual grains, such as the full constraints Taylor model (Taylor 1938), the relaxed constraints model (Honneff and Mecking 1978) or the self-consistent model (Lebensohn and Tomé 1993, Molinari and Tóth 1994). This approach is not suited for the examination of the heterogeneous deformation of single crystals or individual grains in a polycrystal, since they assume the deformation of each grain to be homogeneous. The modelling work described in this paper is believed to be the first attempt at simulating the

deformation of single crystals in three dimensions using the FEM. No assumptions regarding the partitioning of deformation among the grains are required. Due to the use of a rate-dependent constitutive model, assumptions are also not needed regarding the choice of active slip systems.

In addition to single crystals initially at and near the cubic orientation, deformations of polycrystals containing cubic grains have been simulated to study the effect of constraints from neighbours. Unlike experiments with polycrystals, where it may be difficult to monitor the behaviour of interior grains during deformation, the mesoscale modelling approach provides information at all locations in the discretized domain. In this regard, it is possible to examine the effect of neighbour interactions in altering the deformation of a grain from that specified by the boundary conditions. Such deviations from the overall deformation are implicit in the mesoscale finite-element model.

2. Finite-element formulation

The finite-element formulation used for the present study is based on the hybrid formulation described by Beaudoin *et al* (1995). It is assumed that elastic deformations are negligibly small, and that deformation occurs by slip dominated plastic flow of the material. Due to the limited modes of deformation available through slip, the crystal must rotate to accommodate arbitrary deformations. The resulting texture development is modelled using a constitutive law based on crystal plasticity.

Details regarding the constitutive model and the finite-element formulation are described elsewhere (Beaudoin *et al* 1995) and only a brief outline is given here. A viscoplastic constitutive law relates the rate of shearing $\dot{\gamma}^{(\alpha)}$ to the resolved shear stress $\tau^{(\alpha)}$ on each slip system α:

$$\dot{\gamma}^{(\alpha)} = \dot{\gamma}_0 \left| \frac{\tau^{(\alpha)}}{\hat{\tau}} \right|^{1/m} \operatorname{sign}(\tau^{(\alpha)}) \tag{1}$$

where m is the rate sensitivity parameter and $\dot{\gamma}_0$ is a reference rate of shearing. $\hat{\tau}$ is a hardness parameter which represents resistance to plastic deformation due to interactions among dislocations. The symmetric part of the Schmid tensor $\boldsymbol{P}^{(\alpha)}$ (dyadic product of the slip direction and slip plane normal vectors) relates the crystal deviatoric Cauchy stress $\boldsymbol{\sigma}'_c$ to the resolved shear stress and the rate of deformation $\boldsymbol{D}_c$ to the shearing rates:

$$\tau^{(\alpha)} = \boldsymbol{\sigma}'_c \cdot \boldsymbol{P}^{(\alpha)} \tag{2}$$

$$\boldsymbol{D}_c = \sum_{\alpha} \dot{\gamma}^{(\alpha)} \boldsymbol{P}^{(\alpha)}. \tag{3}$$

Eliminating $\dot{\gamma}^{(\alpha)}$ between (1) and (3), and substituting (2) for $\tau^{(\alpha)}$ leads to an expression for the crystal deformation rate in terms of the deviatoric stress,

$$\boldsymbol{D}_c = \left[\sum_{\alpha} \frac{\dot{\gamma}_0}{\hat{\tau}} \left| \frac{\tau^{(\alpha)}}{\hat{\tau}} \right|^{(1/m)-1} \boldsymbol{P}^{(\alpha)} \otimes \boldsymbol{P}^{(\alpha)} \right] \boldsymbol{\sigma}'_c = \mathcal{S}_c \boldsymbol{\sigma}'_c \tag{4}$$

where $\mathcal{S}_c$ is the fourth-order crystal 'compliance' tensor. The rate dependence of (1) permits inversion of (4), and the expression of the crystal deviatoric stress under a given deformation rate as

$$\boldsymbol{\sigma}'_c = \mathcal{S}_c^{-1} \boldsymbol{D}_c. \tag{5}$$

The nonlinear nature of the crystal constitutive equation (5) requires an iterative method to compute the deviatoric stress for a given rate of deformation. The anisotropic response due to the crystal orientation is reflected in the crystal compliance.

The crystal rate of deformation is the symmetric part of the crystal velocity gradient and prescribes the rate of shearing of slip planes. The skew-symmetric part of the crystal velocity gradient $\boldsymbol{W}_c$ controls the rotation of the crystal, and contains contributions from both the spin associated with the plastic flow, and the rigid rotation $\boldsymbol{R}^*$ of the crystal lattice necessary to maintain compatibility with neighbouring crystals. The difference between the applied spin and the plastic spin due to slip provides an expression for the crystal reorientation rate

$$\dot{\boldsymbol{R}}^* = \left(\boldsymbol{W}_c - \sum_\alpha \dot{\gamma}^{(\alpha)} \boldsymbol{Q}^{(\alpha)}\right) \boldsymbol{R}^* \tag{6}$$

where $\boldsymbol{Q}^{(\alpha)}$ is the skew-symmetric part of the Schmid tensor.

The plastic deformation of the material is modelled in incremental fashion, by solving the boundary value problem for material motion at each strain increment. Balance laws for equilibrium and mass conservation are applied in conjunction with the constitutive assumptions discussed above. Following the approach of Beaudoin *et al* (1995), a hybrid finite-element formulation is employed for this purpose. Instead of developing the equilibrium statement from the balance of momentum at the global level, here it is written as a balance of tractions at the inter-element boundaries. Weighted residuals are formed on the equilibrium statement and the constitutive relation. A third residual on the conservation of mass (which for the case of incompressible plastic deformation reduces to a divergence-free velocity field) completes the formulation. Interpolation functions are introduced for the nodal velocities, element stress components and the pressure. A proper choice of the shape functions for the stress permits elimination of the stress degrees of freedom at the element level. The result is a system of equations for the discretized velocity field, which is solved assuming fixed material state and geometry.

Upon obtaining a converged velocity solution, the material state and geometry are updated. The material is characterized by the orientation of the crystal and the hardness parameter. The reorientation rate given by (6) is used to update the orientation, while the hardness is evolved using a modified Voce type law,

$$\dot{\hat{\tau}} = H_0 \left(\frac{\hat{\tau}_s - \hat{\tau}}{\hat{\tau}_s - \hat{\tau}_i}\right) \dot{\gamma}^* \tag{7}$$

where hardening rate H_0 and initial hardness $\hat{\tau}_i$ are material parameters. $\dot{\gamma}^*$ is a measure of the net shearing rate on all of the slip systems,

$$\dot{\gamma}^* = \sum_\alpha \left|\dot{\gamma}^{(\alpha)}\right| . \tag{8}$$

The saturation hardness $\hat{\tau}_s$ based on the current slip system state is given by

$$\hat{\tau}_s = \hat{\tau}_{s_0} \left(\frac{\dot{\gamma}^*}{\dot{\gamma}_s}\right)^{m'} \tag{9}$$

where $\hat{\tau}_{s_0}$, $\dot{\gamma}_s$ and m' are material parameters.

Development of the material response entails solution of the nonlinear crystal constitutive relation for each element, and must be performed during each iteration for the velocity field at a given strain increment. In this respect, the methodology described above proves to be computationally demanding. The use of the hybrid approach leads to the introduction of additional degrees of freedom for the crystal stresses, thereby adding to the computational burden associated with the stiffness calculations. The advantage of using this approach is the smoothness in the stress field, due to the enforcement of traction balance at the element interfaces in an approximate sense (Beaudoin *et al* 1995). In the finite-element context, the numerical integration required for computing the stiffness matrix can be performed

concurrently for all elements. The choice of piecewise discontinuous interpolation functions for the stress is a key feature of the formulation, which enables computation of the stiffness matrices in concurrent fashion for all elements. In turn, this feature permits exploitation of parallel computing technologies in order to greatly improve the feasibility of treating large three-dimensional discretizations.

While the stiffness computations are relatively straightforward to implement in a parallel environment, the solution of the resulting system of equations poses a greater challenge. Since direct solvers are difficult to optimize on a parallel machine, it is advantageous to use an iterative procedure, such as the conjugate-gradient method. In this context, enforcing the incompressibility constraint requires special attention, since it degrades the numerical condition of the resulting system of equations. In the current formulation, incompressibility is enforced using a modified consistent-penalty approach, which seeks to decouple the solution for the pressure field from the conjugate-gradient method. Details of the parallel implementation of the formulation on the Intel Paragon may be found elsewhere (Sarma *et al* 1998a).

3. Results and discussion

The formulation described above has been used to simulate the deformation of single crystals of aluminium with different initial orientations, as well as polycrystalline aggregates. For the single crystal simulations, all elements of the mesh were assigned the same orientation. For the polycrystals, all elements of a grain were initialized with the orientation of that grain. Deformation corresponding to idealized plane strain compression was imposed through specification of boundary velocities. One of the sides of the parallelepiped shaped domain was assumed to be a symmetry plane and the other side was unconstrained. Nodes on the front, back and bottom faces were constrained to remain in plane. Velocities of nodes on the top surface were prescribed to maintain a constant rate of deformation. Compressive strain increments of 1% were used to model the deformation. Since the deformation approximates the conditions at the centre of a rolled sheet, the coordinate axes X, Y and Z are identified with the rolling (RD), transverse (TD) and normal directions (ND), respectively.

For the polycrystal deformation simulations, differences in orientations of grains lead to inhomogeneous deformation. For the single crystal simulations, since the initial orientations are the same for all elements, ideally the mesh should deform in a homogeneous fashion. However, due to small numerical differences in the element calculations which accumulate with increasing time steps, random deviations from homogeneous deformation arise. These deviations propagate with strain, eventually causing the single crystal to 'break up', depending on the orientation of the crystal. It is also possible to introduce the inhomogeneity by randomly perturbing the orientations of the elements from the perfect orientation (Becker *et al* 1991). The effect is to accelerate the development of non-uniform deformation, so that deviations are observed at smaller strain values.

3.1. Deformation of single crystal with orientation (001)[110]

The deformation of a single crystal with initial orientation (001)[110] was simulated using material parameters listed in table 1. Various experimental and modelling studies have shown this orientation to be unstable during plane strain compression, with the crystal rotating to the complementary copper components at large strains (Butler and Hu 1989, Becker *et al* 1991, Akef and Driver 1991, Cizek *et al* 1996). For this reason, this particular orientation was considered a good test case to examine the performance of the model in predicting the

Table 1. Material parameters for simulations of (001)[110] oriented single crystal.

m	$\dot{\gamma}_0$	H_0	$\hat{\tau}_i$	τ_{s0}	$\dot{\gamma}_s$	m'
0.016	0.01 s^{-1}	58.41 MPa	20.0 MPa	75.0 MPa	5.0×10^{10} s^{-1}	0.0

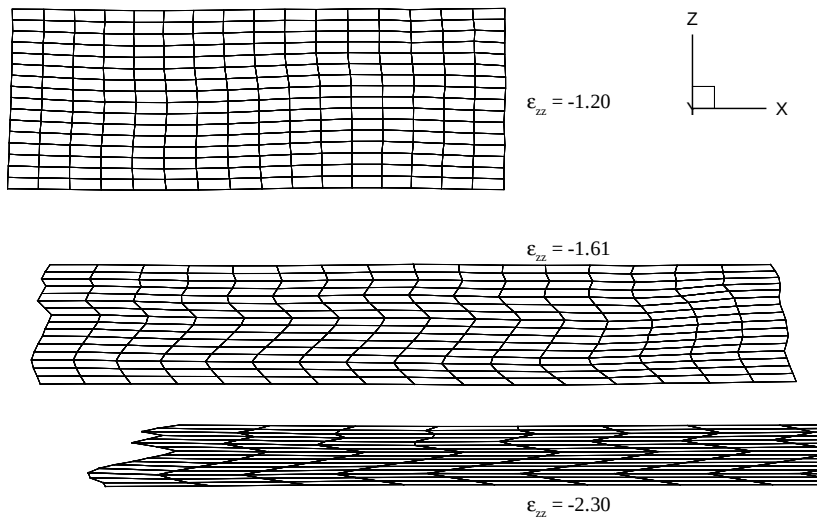


Figure 1. Deformed meshes of the inner 16 × 16 elements for section 10 (out of 20) from the 3D simulation of plane strain compression of a crystal initially oriented at (001)[110].

development of inhomogeneity in the crystal. The simulations were also used to compare the use of a 3D against a planar discretization. The 3D simulations were carried out using a mesh of 20 × 20 × 20 eight-node brick elements, with initial aspect ratio of elements along X, Y and Z of 1:2:4. The deformation is uniform initially, but becomes non-uniform with increasing strain as seen from figure 1, which shows the inner 16 × 16 elements in the RD–ND plane from the mid-section along the TD. There is considerable RD–ND shear in opposite directions in different layers of elements. Examination of the entire 3D mesh shows that no other types of shear develop, so that material lines viewed along the RD and ND remain straight and parallel throughout the deformation. This is quite consistent with experimental observations using optical micrographs from deformation of (001)[110] oriented single crystals (Becker *et al* 1991, Akef and Driver 1991). Formation of transition bands which separate regions of opposite shear is captured by the model. In addition, the experiments reveal that these bands become high angle boundaries which run parallel to the RD and TD, as seen at high strains in the simulation results. While the formation of the shear layers is somewhat evident from the 2D simulations by Becker *et al* (1991), 3D simulations show that the high-angle boundaries separating the layers of opposite shear are indeed in the RD–TD plane.

An interesting aspect of the deformation of the (001)[110] single crystal is that it is essentially 2D in behaviour. The different sections along the TD are identical to each other in their behaviour, so that a 2D simulation is adequate to simulate the deformation. This has been verified by conducting a simulation using a 20 × 20 mesh of elements in the RD–ND plane, which is only one element thick along the TD. The results are identical to the 3D case, indicating that for this particular orientation there is no variation in the deformation along TD.

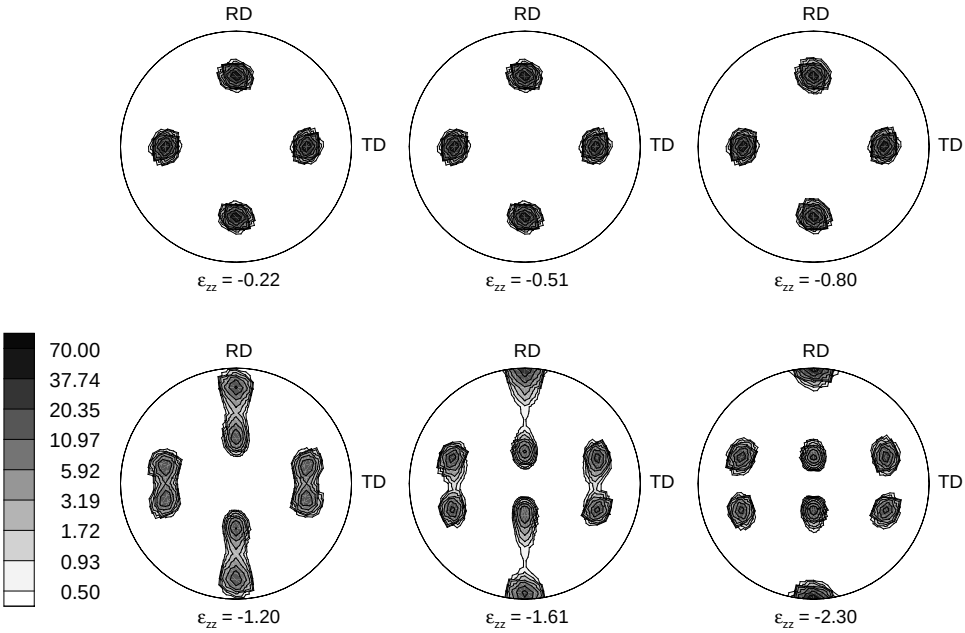


Figure 2. ⟨111⟩ pole figures in equal area projection showing texture evolution during plane strain compression of a crystal initially oriented at (001)[110].

The shearing causes the material to rotate in opposite directions about the TD to the two copper components, $(112)[\bar{1}\bar{1}1]$ and $(\bar{1}\bar{1}2)[\bar{1}\bar{1}\bar{1}]$. Calculation of the misorientation between the initial and final orientations of the elements indicates rotation about the [010] axis (i.e. TD) by $\pm 35^\circ$. The texture evolution is shown in figure 2 in the form of ⟨111⟩ pole figures, generated using popLA (Kallend *et al* 1991) assuming cubic crystal symmetry and a 5° Gaussian smoothing.

The average rotation of the lattice, computed by taking the average of the misorientation values of the inner 4096 elements with the initial (001)[110] orientation, is plotted against the compressive strain in figure 3. Similar plots are presented by both Becker *et al* (1991) and Akef and Driver (1991) to compare their models with experimental data. The data show that the lattice rotation starts at a much smaller strain compared to the value predicted by the present simulations. This is due to the assumption of a perfect crystal with all elements initially at the same orientation, which leads to a strain of about 0.7 before deviations among the elements become significant. However, if the origin is shifted appropriately, then the increase in rotation angle is quite consistent with the experimental data. The analytical expression derived by Akef and Driver predicts a faster rotation to the copper orientation than is observed from the measurements. The results of the current simulations capture quite well the decrease in rotation rate at large strains as the copper component is approached.

Simulations using a fully 3D mesh and a mesh with a single layer of elements in the RD–ND plane were also carried out for an initial orientation rotated 3.5° about ND from (001)[110]. The deviation from the symmetry of the (001)[110] orientation leads to differences between the 3D and planar meshes. As seen from figure 4, which shows the inner 16×16 elements for the planar case, the deformation is no longer restricted to XZ shear. The 3D mesh, on the other hand,

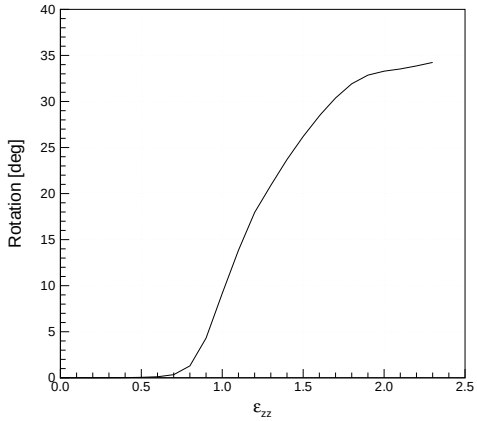


Figure 3. Variation of the average rotation of the crystal with strain, from an initial (001)[110] orientation.

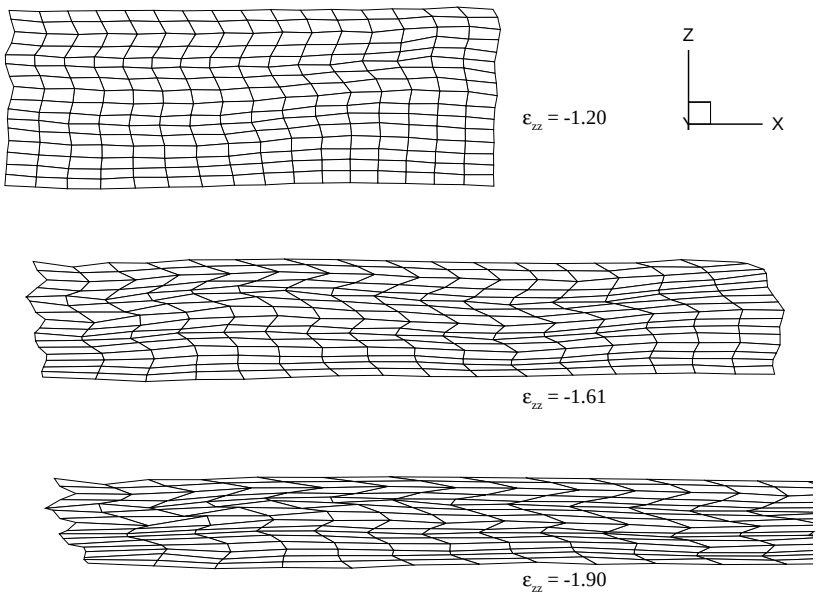


Figure 4. Deformed meshes of the inner 16 × 16 elements from the planar simulation of plane strain compression of a crystal initially rotated 3.5° about ND from (001)[110].

deforms more like the exact (001)[110] orientation, as shown in figure 5. The texture results reflect the influence of the deformation. The 3D simulation, with predominant XZ shear, leads to rotation about the TD towards the copper components, but with a superimposed rotation about the ND due to the initial orientation. This trend is also observed in the experimental results from channel die compression of a (001)[110] crystal initially misoriented by a 3.5° rotation about the compression axis (Becker *et al* 1991). Examination of the pole figures from the experiments shows a slight rotation about the ND, though it is not as pronounced as in the simulation results shown in figure 6. Texture results from the planar simulation are presented in figure 7. A greater amount of element shearing in different directions leads to more inhomogeneity and greater spread in the textures with lower peak intensities. These results indicate that except for special cases where symmetry between the orientation of the crystal

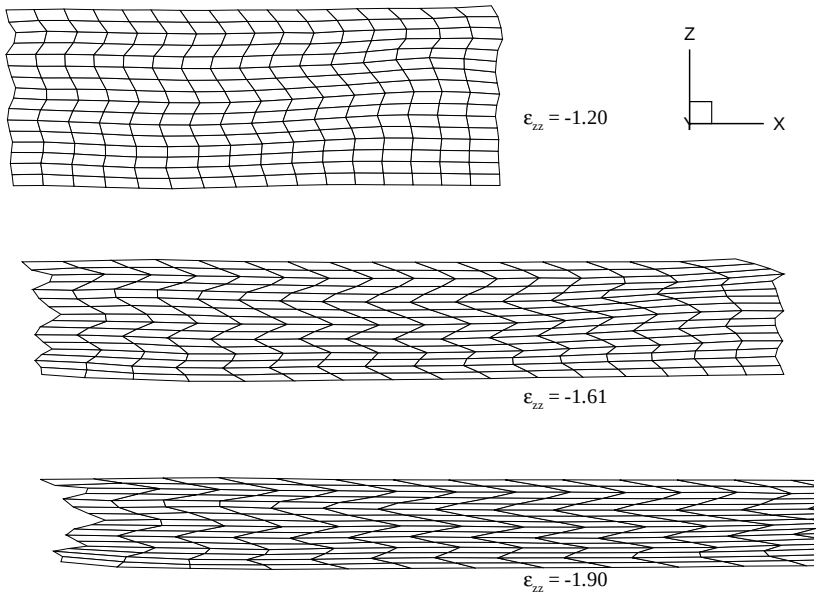


Figure 5. Deformed meshes of the inner 16 × 16 elements for section 10 (out of 20) from the 3D simulation of plane strain compression of a crystal initially rotated 3.5° about ND from (001)[110].

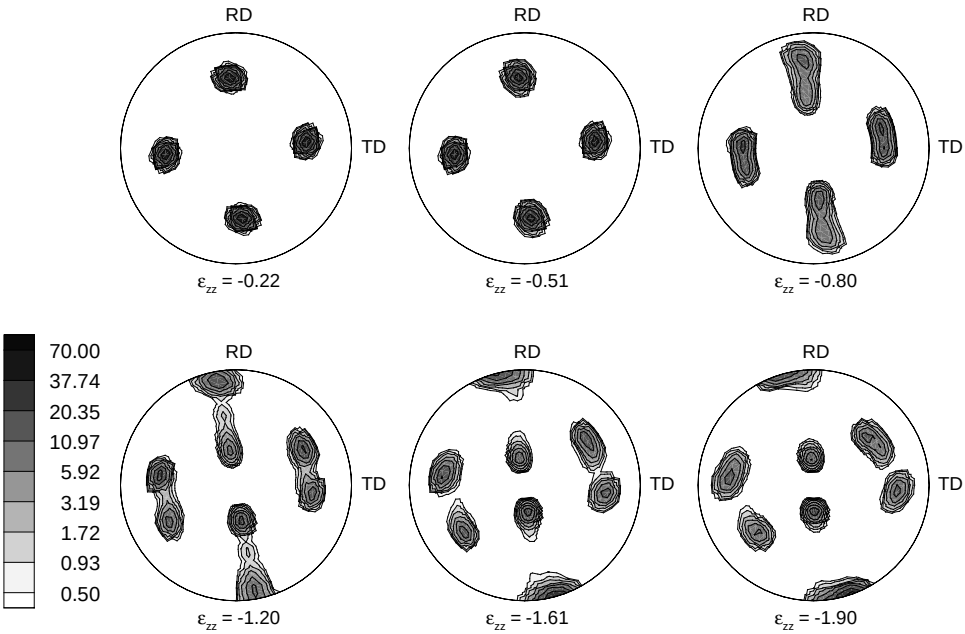


Figure 6. ⟨111⟩ pole figures in equal area projection showing texture evolution during plane strain compression of a crystal initially rotated 3.5° about ND from (001)[110] simulated using a 3D mesh.

and the applied deformation plays a role, constraints imposed by using only a single layer of elements or a 2D approximation can lead to considerable differences from a full 3D analysis.

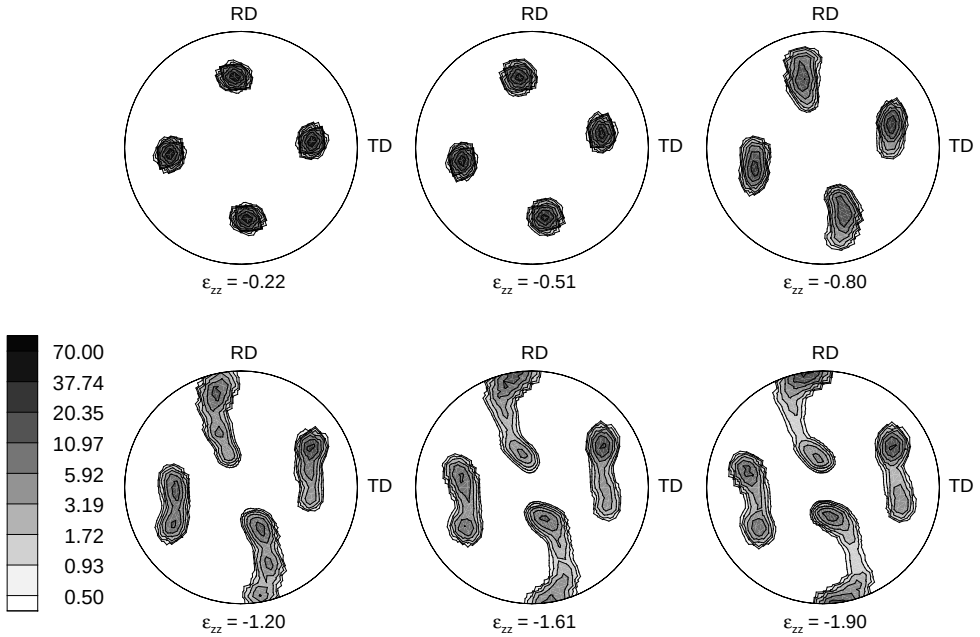


Figure 7. ⟨111⟩ pole figures in equal area projection showing texture evolution during plane strain compression of a crystal initially rotated 3.5° about ND from (001)[110] simulated using a planar mesh.

3.2. Effect of slip on {110} planes in cubic single crystals

The approach discussed above for simulating the deformation of single crystals has been applied to cubic and near-cubic orientations, to study the effect of slip on {110} planes on the stability of the cubic orientation. These simulations may be viewed as numerical experiments involving the deformation of single crystals. As mentioned earlier, previous modelling work on the effect of {110} slip has focused on the texture evolution of an aggregate of grains, using various assumptions to relate the overall deformation to the deformation of each grain. This approach is not suited to examine the inhomogeneous deformation of individual grains. The modelling approach using FEM provides the capability to study the non-uniform deformation of single crystals, as illustrated above for (001)[110] oriented crystals.

For each simulation, the crystal was discretized with 4096 elements arranged in the form of a cube with 16 elements along each side. The inner 1000 elements were used for examining the texture development. Material parameters used for the simulations, based on mechanical test data for aluminium (Mathur and Dawson 1989) are listed in table 2. Maurice and Driver (1997a) have examined the relative values of the critical resolved shear stress on the {111} and {110} planes, and conclude that they are almost identical at elevated temperatures. Based on their results, the critical resolved shear stress was assumed to be the same for all slip systems (for both {111} and {110} planes).

The deformation of a crystal at the cubic orientation (001)[100] was modelled first, using the standard 12 ({111}⟨110⟩) slip systems. The results showed the orientation to be stable to fairly large strains, with the ability to deform in plane strain compression with no change in orientation. The initial and deformed meshes are shown in figure 8, indicating the uniform deformation of the cubic-oriented crystal. Experiments on plane strain compression

Table 2. Material parameters for the deformation simulations to study effect of slip systems.

m	$\dot{\gamma}_0$	H_0	$\hat{\tau}_i$	τ_{s0}	$\dot{\gamma}_s$	m'
0.05	1.0 s^{-1}	58.41 MPa	27.17 MPa	61.80 MPa	5.0×10^{10} s^{-1}	0.005

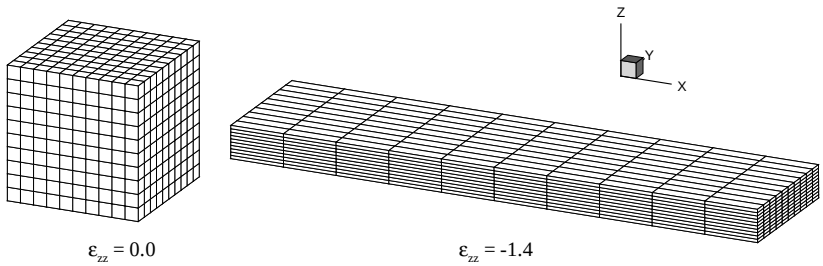


Figure 8. Initial and deformed meshes for the inner 1000 elements from plane strain compression of a cubic-oriented crystal using 12 slip systems.

of cubic-oriented single crystals have shown that it is not stable at low temperatures, but is at high temperatures (e.g. Maurice and Driver 1997a). Examination of the slip system activity revealed that the deformation was accommodated by equal slip on the complementary systems $(11\bar{1})[101]$, $(1\bar{1}\bar{1})[101]$, $(1\bar{1}1)[10\bar{1}]$ and $(111)[10\bar{1}]$, so that the crystal continued to deform with no rotation. Deviation from the mathematically exact cubic orientation was considered necessary to cause variations in the slip system activity, and hence development of non-uniform deformation.

The deformation of a crystal misoriented by a small amount from the exact cubic orientation was modelled without and with the six $\{110\}\langle 110\rangle$ slip systems. The initial orientation of the crystal was taken to be $(0^\circ,\ 5^\circ,\ 5^\circ)$ in terms of Euler angles using Kocks convention† (Kocks 1988). The choice of the amount of deviation from an exact cube was somewhat arbitrary. A somewhat large misorientation was chosen to verify the rotation of the elements towards an exact cube in the presence of $\{110\}$ slip, as indicated by some analytical results (Maurice and Driver 1993). For the simulation with 12 slip systems, the crystal initially deforms with maximum shear rates on the $(11\bar{1})[101]$ and $(1\bar{1}1)[10\bar{1}]$ systems, and with smaller shears on $(1\bar{1}\bar{1})[101]$ and $(111)[10\bar{1}]$. Some of the other slip systems also show limited activity. In the presence of the $\{110\}\langle 110\rangle$ slip systems, the maximum shear rates occur on the complementary slip systems $(101)[10\bar{1}]$ and $(10\bar{1})[101]$, with smaller values on the slip systems with the same directions on the $(11\bar{1})$ and (111) planes, respectively.

The texture evolution computed using the orientations of the inner 1000 elements is shown in figure 9 for the two cases. After a 50% reduction ($\varepsilon_{zz} = -0.7$), there is no noticeable difference between the two textures. With further deformation, the crystal deforming with 12 slip systems begins to show deviation from its initial orientation. After about 83% reduction ($\varepsilon_{zz} = -1.8$), the texture consists of two main peaks, but shows appreciable spread about them. The crystal deformed using 18 slip systems shows some inhomogeneity, but in this case the initial misorientation from cubic gradually disappears, with most of the elements rotating towards cubic orientation. The initial orientation represented by an axis and angle pair (Becker and Panchanadeeswaran 1989) has a misorientation of 7° from cubic. After deformation to $\varepsilon_{zz} = -1.8$, more than 85% of the elements are misoriented by less than 5° from cubic for the 18 slip system case, whereas with 12 slip systems that number is less than 6%.

† $(90^\circ,\ 5^\circ,\ 85^\circ)$ in Bunge notation.

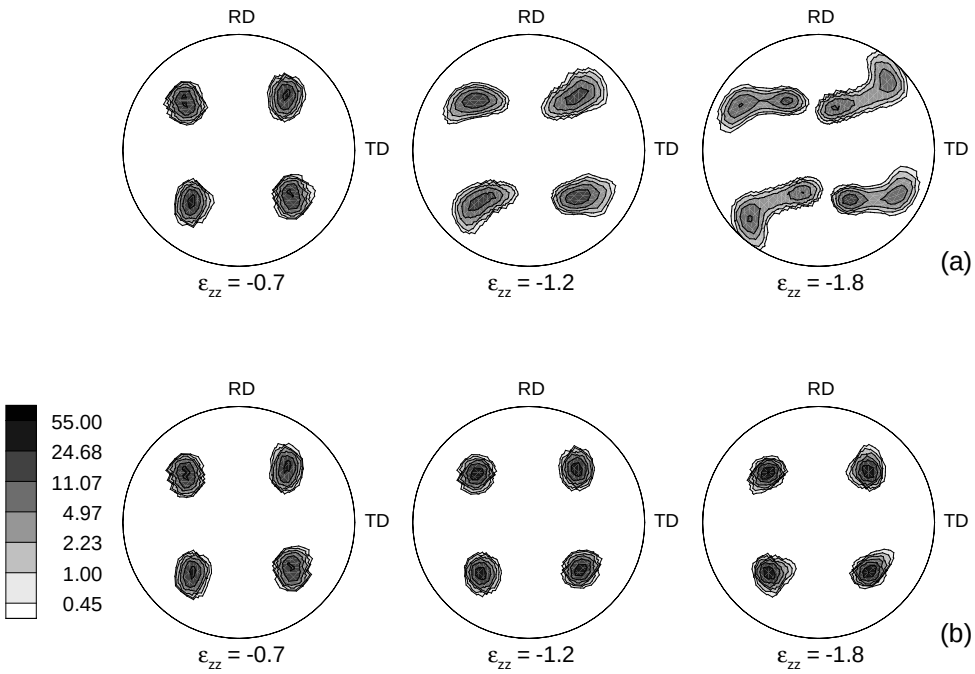


Figure 9. ⟨111⟩ pole figures in equal area projection showing texture evolution of a crystal slightly misoriented from cubic, deformed in plane strain compression using (*a*) 12 slip systems and (*b*) 18 slip systems.

Examination of the slip system activity after deformation to $\varepsilon_{zz} = -1.8$ shows that in the presence of 18 slip systems almost all of the elements deform with maximum and equal shear rates on the $(101)[10\bar{1}]$ and $(10\bar{1})[101]$ systems. Only a few elements have maximum shear values on either the $(11\bar{1})[101]$ or the $(1\bar{1}1)[10\bar{1}]$ systems. These elements are also the ones with high misorientation values from cubic orientation. Thus, the effect of the presence of $\{110\}$ planes is only felt for orientations within about 14° of cubic. With increasing misorientation, the contribution of shear on the $\{110\}$ planes diminishes.

The deformed meshes for the crystal misoriented from cubic after 70% reduction are shown in figure 10 for both 12 and 18 slip systems. Both cases show significant XZ shear deformation in the same direction, and smaller degree of XY shear in opposite directions. Deformation with 12 slip systems also leads to some YZ shear, which is not as pronounced with 18 slip systems. The deformation is far more uniform with 18 slip systems, leading to the sharp cubic texture, whereas the greater inhomogeneity for the 12 slip system case leads to more diffuse texture. These results show that the addition of the $\{110\}\langle 110\rangle$ slip systems helps to stabilize the cubic orientation during plane strain compression.

3.3. Effect of slip on {110} planes in polycrystals

In the previous subsection, the deformations of single crystals at and near cubic orientation were considered with and without $\{110\}\langle 110\rangle$ slip systems. The crystals were not influenced by the presence of neighbouring grains, as would be the case in a microstructure. In this subsection we consider the deformation of an aggregate of grains, each of which is discretized with many elements to model its non-uniform deformation in the presence of constraints from

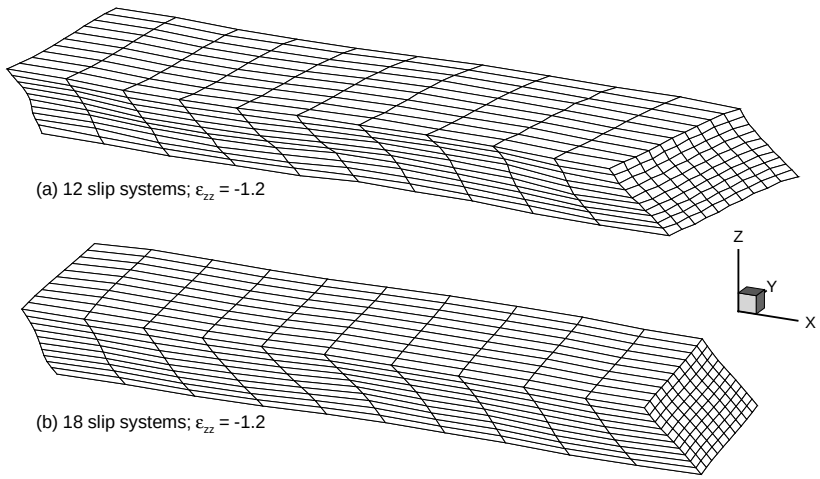
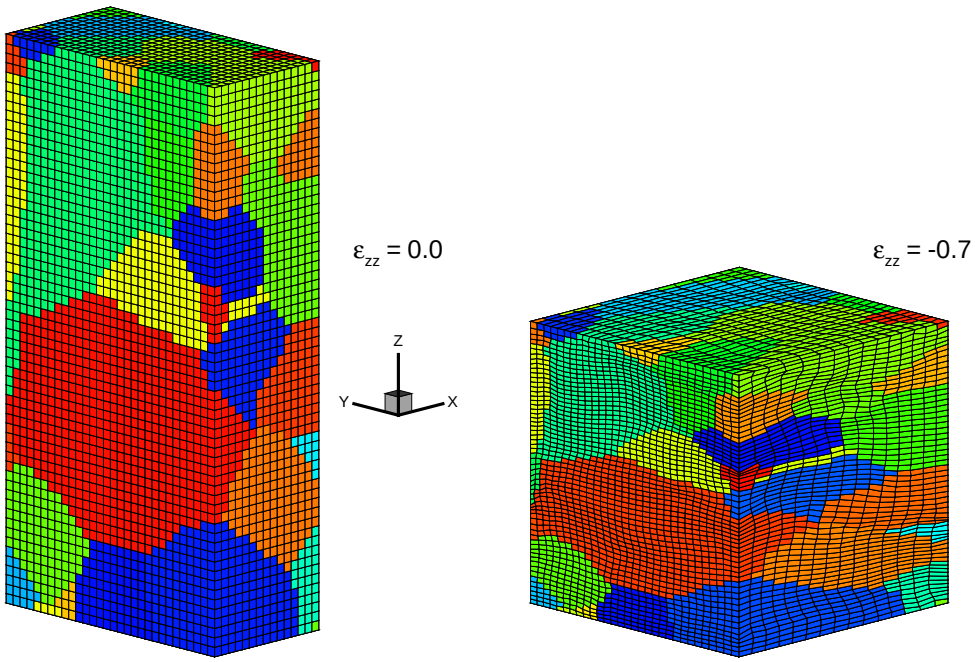


Figure 10. Deformed meshes for the inner 1000 elements from plane strain compression of a crystal slightly misoriented from cubic using (*a*) 12 slip systems and (*b*) 18 slip systems.

the surrounding grains. The simulation procedure was applied to a microstructure consisting of 53 grains (Sarma *et al* 1998b), discretized using a total of 27 000 elements arranged in a $15 \times 30 \times 60$ mesh. The initial mesh and microstructure are shown in figure 11(*a*), where the different colours in the microstructure indicate different grains. The initial orientations of the grains were chosen from a random sampling of Euler space, but since each grain is weighted by the number of elements used to discretize it, the initial texture is not random. The orientation of the largest grain was taken to be cubic, and the microstructure was deformed in plane strain compression to a compressive strain of $\varepsilon_{zz} = 0.7$ (50% reduction in height) using 12 and 18 slip systems. The deformed mesh from the simulation using 12 slip systems is shown in figure 11(*b*), and depicts the inhomogeneous deformations of the grains. While the overall deformation has no shear components, the individual elements undergo shear deformations to varying extent. Introduction of additional $\{110\}$ slip planes does not lead to significant differences in the deformed mesh, indicating that for most grains (orientations), the additional slip systems do not play a major role.

Figure 12 shows the texture after 50% reduction for deformation with 12 and 18 slip systems. The pole figures were constructed from the 27 000 element orientations after applying cubic crystal symmetry and orthotropic sample symmetry. The two textures are very similar except for the peak intensity values at the cubic orientation. The other features are almost identical, resembling typical rolling texture components. The influence of the initial texture leads to a strong $\{123\}\langle 634\rangle$ S component and a weaker $\{112\}\langle 111\rangle$ copper component. The simulation using 18 slip systems leads to a much stronger cubic component, indicating once again the stabilizing effect of slip on $\{110\}$ planes.

While the textures as a whole are very similar, there are some variations in the individual element reorientations, with the misorientation among corresponding elements exceeding 50° in some cases. For most of the grains the difference between deformation with 12 and 18 slip systems is minor, with only a small fraction of the elements comprising the grains being misoriented by more than 10°. However, a few of the grains show significant misorientations for a large fraction of elements after deformation with 12 and 18 slip systems.



(a) (b)

Figure 11. Initial and deformed mesh and microstructure for polycrystal simulations using 12 slip systems. (*a*) Discretized initial microstructure, (*b*) deformed microstructure.

The influence of the slip systems on the deformation of the cube-oriented grain is examined through the spread of the orientations of the elements comprising the grain after a 50% reduction. The orientations were converted from Euler angles to axis-angle pairs†, and the resulting histogram for the distribution of angles is shown in figure 13. Since the cubic orientation coincides with the reference axes, the angle of rotation also represents the deviation from cubic. The histogram for 18 slip systems shows that most of the elements are misoriented from cubic by less than 10°, with very few elements rotating away by more than 15°. The opposite is true for the case of 12 slip systems, with a far greater number of elements misoriented by high angles. The influence of the neighbouring grains in the microstructure causes the initially cube-oriented grain to rotate away, but slip on $\{110\}$ planes retards the rate of rotation.

In the single crystal simulations of cubic and near-cubic orientations, the overall deformation of the crystal was determined by the applied boundary conditions. Therefore, most elements of the crystal were subjected to plane strain compression. However, in the case of the polycrystal, the deformation of the grain is strongly influenced by its neighbours, so that the deformation can deviate appreciably from plane strain compression, especially at regions close to the grain boundaries. In order to study the influence of the proximity to the grain boundary on the elements with initial cubic orientation, they were classified as interior or boundary elements, depending on whether the adjacent elements belonged to the same or different grains in the initial microstructure. The number of adjacent elements was varied to

† Cubic crystal symmetry was used to determine the symmetrically equivalent axis and angle with the smallest angle of rotation.

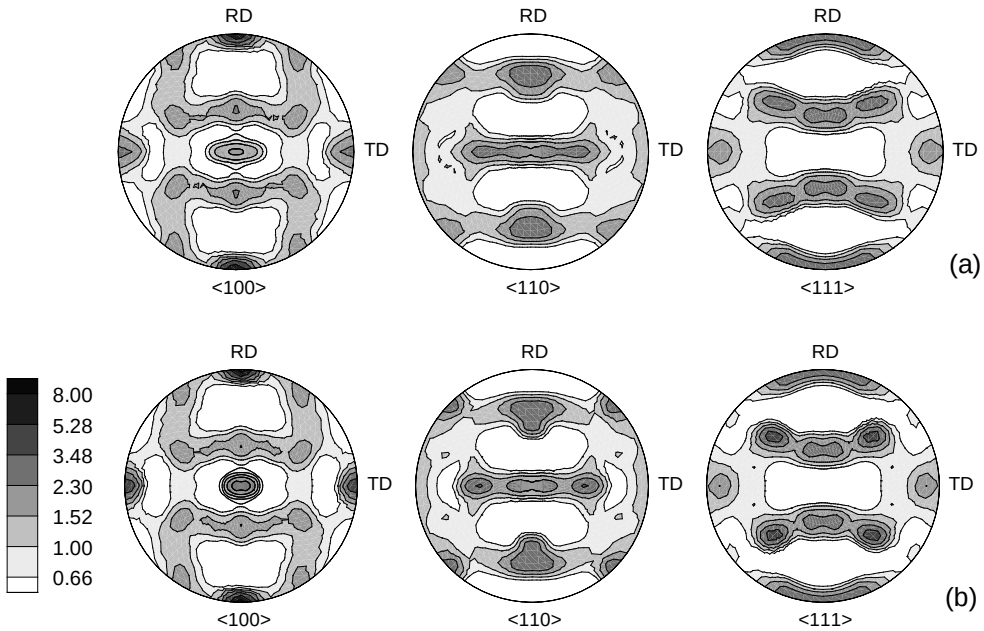


Figure 12. Pole figures in equal area projection showing the texture after deformation to $\varepsilon_{zz} = -0.7$ using (*a*) 12 slip systems and (*b*) 18 slip systems for the microstructure with initial cubic grain.

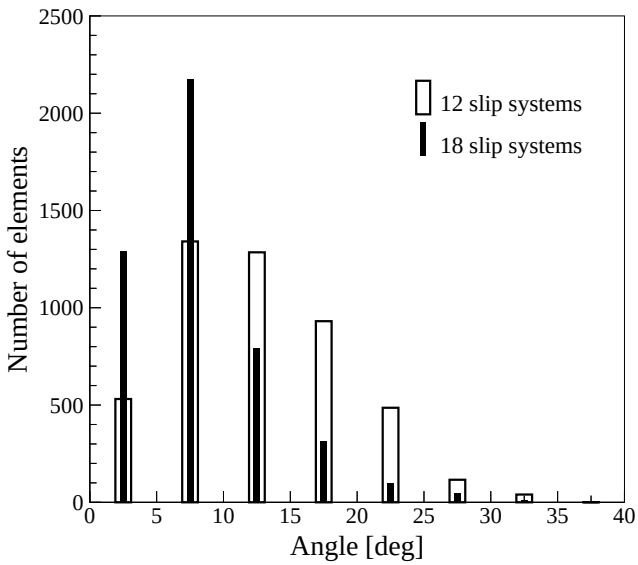


Figure 13. Histograms showing distribution of angles of rotation for the elements comprising the initially cube-oriented grain.

get sets of boundary elements of different thicknesses. Table 3 shows the comparison between the 18 and 12 slip system cases for different thickness values for the boundary elements. As the boundary thickness increases, the maximum angle of rotation for the interior elements

Table 3. Comparison of interior and boundary regions of an initially cube-oriented grain.

	Maximum θ among interior elements		Interior elements with $\theta < 10^\circ$ (%)		Boundary elements with $\theta < 10^\circ$ (%)	
Thickness of boundary elements	18 slip systems	12 slip systems	18 slip systems	12 slip systems	18 slip systems	12 slip systems
1	31.5°	34.2°	76.0	38.6	65.0	42.1
2	23.0°	30.4°	78.4	36.7	67.1	42.9
3	18.4°	28.6°	78.7	34.2	69.9	42.8
4	17.9°	27.2°	77.9	31.2	71.6	42.5

decreases for both cases. However, the decrease is far greater for 18 slip systems, with a maximum deviation from cubic of only $\approx 18^\circ$ when the boundary is four elements thick, compared to $\approx 27^\circ$ for 12 slip systems. The increase in the maximum angle from the interior to the boundary indicates the effect of neighbouring grains on the deformation. The elements near the boundary have greater deviations from plane strain deformation compared to elements in the grain interior, and as a consequence, may be expected to be less affected by the stabilizing presence of the $\{110\}$ slip systems.

Another indication of the difference between 12 and 18 slip systems comes from the examination of the percentage of interior and boundary elements that remain close to cubic after deformation. As seen from table 3, the fraction of elements which rotate by less than 10° away from initial cubic is greater for 18 slip systems for both interior and boundary elements. The fraction of interior elements with small misorientation from cubic remains close to 78% with different boundary thickness values. The fraction of boundary elements with $\theta < 10^\circ$ goes down as the boundary thickness reduces, indicating the increasing effect of neighbouring grains as one gets closer to the boundary. The heterogeneity in deformation also makes the $\{110\}$ slip systems less effective than in the case of more homogeneous plane strain deformation conditions.

The boundary effect is also seen for the case of 12 slip systems, although the trend is reversed. As the boundary thickness increases, causing fewer elements to be considered as part of the interior, the fraction of interior elements with small misorientation from cubic decreases. This means that as one goes towards the interior of the grain, one finds fewer grains with small misorientation from cubic. Near the boundary, the effect of neighbouring grains is in fact to retard the rotation away from cubic, so that the fraction of boundary elements with small misorientation from cubic remains fairly constant.

It must be mentioned here that the results presented on the boundary and interior elements are intended to highlight the differences between the absence and presence of $\{110\}$ slip planes on the cubic grain, and these results are insufficient to draw conclusions regarding the size of the zone of influence near the grain boundaries. Determining the zone of influence would require additional simulations to ensure that mesh resolution effects are accounted for, which is beyond the scope of this work. Even if such simulations are performed, the zone of influence can only be determined relative to the grain size and not as an absolute measure, since there is no size scale implicit in the model.

Simulations without and with the $\{110\}\langle 110\rangle$ slip systems were also carried out using the 3D microstructure where there was no initial cubic grain. In this case, there was no appreciable difference between the two resulting textures as a whole, leading to typical rolling components (Sarma *et al* 1998b). However, when individual grains were examined, a few of them showed significant differences either due to their location or their orientation. The presence of the

initial cubic grain leads to some differences in the reorientation of individual elements of various grains, but the deformations are by and large quite similar. This indicates that slip on $\{110\}$ planes is only significant for select orientations such as the cubic orientation, and is not a factor for most of the other orientations.

Finally, it must be mentioned that for the deformation of the single crystal near cubic, only at fairly large strains is there an appreciable difference in textures between 12 and 18 slip systems (figure 9). However, for the case of a polycrystal, a 50% reduction is enough to bring about significant differences between deformation with 12 and 18 slip systems. This is another indication of the effect of grain interactions, which lead to a more heterogeneous deformation state in the microstructure compared to the more homogeneous conditions of plane strain compression in the single crystal.

4. Summary and conclusions

The combination of crystal plasticity and the FEM provides a powerful approach for simulating the deformation of metals at the mesoscale. By discretizing individual grains, the method allows for modelling the deviations from applied overall deformation, not only from grain to grain but also within the individual grains. The non-uniform deformations of individual grains and their interactions with neighbours can be accounted for using this approach. The methodology has been applied to study deformations of single crystals and polycrystals of aluminium, and to examine the effect of including slip on $\{110\}\langle 110\rangle$ systems. Unlike prior modelling efforts using aggregates of grain orientations, the present approach does not require assumptions to partition the applied deformation among the grains, and is hence suitable to study even single crystal deformations.

Simulations with (001)[110] oriented crystals were used to show the ability of the formulation to model the development of shears in the RD–ND plane and the consequent rotation of the crystal about TD to the copper orientation, as has been observed in plane strain compression experiments. These simulations also show that due to symmetry of the crystal with respect to the applied deformation, the (001)[110] orientation deforms with only RD–ND shear. However, a slight rotation from the exact (001)[110] orientation leads to development of other localized shears, so that 2D and 3D simulations no longer lead to the same result.

Simulations with single crystals initially at orientations close to cubic show that slip on $\{110\}$ planes stabilizes the cubic orientation, with the crystal rotating towards the cube during plane strain compression. The same trend is also observed in a cube-oriented crystal which is part of a larger polycrystalline aggregate. The results presented here are in good agreement with experimental data on hot deformed aluminium alloys. In particular, slip on $\{110\}$ planes favouring stability of cube-oriented grains has been observed in both single crystals and polycrystals. The addition of $\{110\}$ planes does not radically change the overall texture, but merely improves the strength of the cubic component. This would indicate that the effect of the $\{110\}$ slip planes is restricted to a few special orientations like cubic. The influence of neighbour constraints is also evident from the polycrystal simulations. Elements close to grain boundaries are affected to a greater extent by the neighbouring grains, leading to greater deviation from plane strain deformation and a smaller influence of the $\{110\}$ planes.

Due to the difficulty in performing experiments with a small aggregate of grains and monitoring all the grains throughout the deformation process, it is not easy to compare the polycrystal simulations directly with experiments. On the other hand, it is also not trivial to carry out such detailed simulations at the mesoscale with large numbers of grains. Therefore, it becomes necessary to perform the comparisons with single crystal experiments. That is one reason for the simulations with the cubic crystals and with the (001)[110] oriented crystal,

which has been the focus of several experiments. The crystal with slight misorientation from cubic when deformed with the usual octahedral slip systems rotates away to other orientations. Various experiments with cubic crystals at room temperature have shown that it is not stable, and its reorientation is quite complex and not restricted to rotation about a particular axis. The simulations also show that different elements rotate about different axes, and that the final texture is somewhat diffuse. The choice of the initial orientation may be expected to play a role in the texture evolution, and further study is required to determine the reorientation path for near-cubic orientations.

Based on observations of slip traces in single crystals deformed in plane strain compression at elevated temperatures, Maurice and Driver (1997a) conclude that non-octahedral slip systems are active at higher temperatures. The results of their numerical simulations (Maurice and Driver 1997b) show that addition of {112}, {110} and {001} slip planes leads to increased strength of certain texture components such as brass and cubic, as commonly observed in hot rolled aluminium alloys. The results of the present simulations would indicate that specific non-octahedral slip systems favour stability of specific orientations. The introduction of these additional slip systems offers one possible means to apply the mesoscale modelling approach to hot deformation of aluminium alloys and other fcc metals.

Acknowledgments

This research was sponsored by the Office of Basic Energy Sciences, US Department of Energy, under contract DE-AC05-96OR22464 with the Lockheed Martin Energy Research Corporation. The research was supported in part by an appointment to the Oak Ridge National Laboratory Postdoctoral Research Associates Program administered jointly by the Oak Ridge National Laboratory and the Oak Ridge Institute for Science and Education. The authors acknowledge the use of the Intel Paragon XP/S 35 computer, located in the Oak Ridge National Laboratory Center for Computational Sciences (CCS), funded by the Department of Energy's Mathematical, Information, and Computational Sciences (MICS) Division of the Office of Computational and Technology Research.

References

Akef A and Driver J H 1991 *Mater. Sci. Engng* A **132** 245–55
Anand L and Kalidindi S R 1994 *Mech. Mater.* **17** 223–43
Bacroix B and Jonas J J 1988 *Textures Microstruct.* **8/9** 267–311
Beaudoin A J, Dawson P R, Mathur K K and Kocks U F 1995 *Int. J. Plast.* **11** 501–21
Beaudoin A J, Dawson P R, Mathur K K, Kocks U F and Korzekwa D A 1994 *Comput. Methods Appl. Mech. Engng* **117** 49–70
Beaudoin A J Jr, Mecking H and Kocks U F 1996 *Phil. Mag.* A **73** 1503–17
Becker R 1991 *Acta Metall.* **39** 1211–30
Becker R, Butler J F Jr, Hu H and Lalli L A 1991 *Met. Trans.* A **22** 45–58
Becker R and Panchanadeeswaran S 1989 *Textures Microstruct.* **10** 167–94
——1995 *Acta Mater.* **43** 2701–19
Butler J F Jr and Hu H 1989 *Mater. Sci. Engng* A **114** L29–33
Cizek P, Wynne B P, Lu H and Parker B A 1996 *Mater. Sci. Engng* A **219** 44–55
Honneff H and Mecking H 1978 *Proc. Fifth Int. Conf. on Textures of Materials (ICOTOM 5)* ed G Gottstein and K Lücke (Berlin: Springer) pp 265–75
Kalidindi S R, Bronkhorst C A and Anand L 1992 *J. Mech. Phys. Solids* **40** 537–69
Kallend J S, Kocks U F, Rollett A D and Wenk H-R 1991 *Mater. Sci. Engng* A **132** 1–11
Kocks U F 1988 *Proc. Eighth Int. Conf. on Textures of Materials (ICOTOM 8)* ed J S Kallend and G Gottstein (Warrendale, PA: The Metallurgical Society) pp 31–6
Lebensohn R A and Tomé C N 1993 *Acta Metall. Mater.* **41** 2611–24

Mathur K K and Dawson P R 1989 *Int. J. Plast.* **5** 67–94
Mathur K K, Dawson P R and Kocks U F 1990 *Mech. Mater.* **10** 183–202
Maurice Cl and Driver J H 1993 *Acta Mater.* **41** 1653–64
——1997a *Acta Mater.* **45** 4627–38
——1997b *Acta Mater.* **45** 4639–49
Mika D P and Dawson P R 1998 *Mater. Sci. Engng* A **257** 62–76
Molinari A and Tóth L S 1994 *Acta Metall. Mater.* **42** 2453–8
Raabe D 1995 *Acta Metall. Mater.* **43** 1023–8
Radhakrishnan B, Sarma G B and Zacharia T 1998a *Acta Mater.* **46** 4415–33
——1998b *Scr. Mater.* **39** 1639–45
Sarma G, Zacharia T and Miles D 1998a *Comput. Math. Appl.* **35** 41–57
Sarma G B, Radhakrishnan B and Zacharia T 1998b *Comput. Mater. Sci.* **12** 105–23
Taylor G I 1938 *J. Inst. Metals* **62** 307–24

ERKCM06 (High-Temperature Alloy Design)

Refereed Journal Publications

†Cairney, J. M., P. R. Munroe, and J. H. Schneibel, "Examination of Fracture Surfaces Using Focused Ion Beam Milling," *Scr. Mater.* **42**, 473-78 (2000).

*Hong, S.-L., C. L. Fu, and M. H. Yoo, "First-Principles Calculation of Stacking Fault and Twin Boundary Energies of Cr_2Nb," *Phil. Mag. A* **80**, 871-80 (2000).

*Kumar, K. S., L. Pang, C. T. Liu, J.A. Horton, and E. Kenik, "Structural Stability of the Laves Phase Cr_2Ta in a Two-Phase Cr-Cr_2Ta Alloy," *Acta Metall.* **48**, 911-23, 2000.

Liu, C. T., J. H. Zhu, M. P. Brady, C. G. McKamey, and L. M. Pike, "Physical Metallurgy and Mechanical Properties of Transition-metal Laves Phase Alloys," *Intermetallics* **8**, 1119-29 (2000).

†Luzzi, D. E., D. Imamura, H. Inui, E. P. George, L. Heatherly, and M. Yamaguchi, "Auger Spectroscopy of TiAl Interface Diffusion Specimens," *Mater. Sci. Forum* **294**(2), 389-92 (1999).

Miura, S., Y. Terada, T. Suzuki, C. T. Liu, and Y. Mishima, "Thermal Conductivity of Ni-Al Powder Compacts for Reaction Synthesis," *Intermetallics* **8**, 151-55 (2000).

Morris, J. R., K. M. Ho, K. Y. Chen, G. Rengarajan, and M. H. Yoo, "Large Scale Atomic Study of Core Structures and Energies of <**c**+**a**>{11-22} Dislocations in hcp Metals," *Modeling Simul. Mater. Sci. Eng.* **8**, 25-35 (2000).

Rosales, I. and J. H. Schneibel, "Stoichiometry and Mechanical Properties of Mo_3Si," *Intermetallics* **8**, 885-89 (2000).

*Somieski, B., J. H. Schneibel, and L. D. Hulett, "Defects in Iron Aluminides Probed with Positrons," *Phil. Mag. Lett.* **79**, 115-23 (1999).

Stoloff, N. S., C. T. Liu, and S. C. Deevi, "Emerging Applications of Intermetallics," *Intermetallics* **8**, 1313-20 (2000).

Wang, J. G., W. J. Chia, Y. W. Chung, and C. T. Liu, "Interaction of H_2O and O_2 with Ni_3Fe and Their Effects on Ductility," *Intermetallics* **8**, 353-57 (2000).

Wang, J. G., B. W. Choi, T. G. Nieh, and C. T. Liu, "Crystallization and Nanoindentation Behavior of a Bulk Zr-Al-Ti-Cu-Ni Amorphous Alloy," *J. Mater. Res.* **15**(3), 798-807 (2000).

*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

Wang, J. G., B. W. Choi, T. G. Nieh, and C. T. Liu, "Nano-Scratch Behavior of a Zr-Based Bulk Amorphous Alloy," *J. Mater. Res.* **15**(4), 913-22 (2000).

Yoo, M. H. and K. Yoshimi, "An Empirical Model for Ideal Work of Adhesion: Transition-metal Aluminides and Silicides," *Intermetalllics* **8**, 1215-24 (2000).

Zhu, J. H. and C. T. Liu, "Defect Structures in $ZrCo_2$ Laves Phase," *Acta Mater*. **48**, 2339-47 (2000).

*Indicates publications involving two or more BES FWP's
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

PHILOSOPHICAL MAGAZINE A, 2000, VOL. 80, NO. 4, 871–880

First-principles calculation of stacking fault and twin boundary energies of Cr_2Nb

SUKLYUN HONG†, C. L. FU and M. H. YOO

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6114, USA

[*Received 15 July 1998 and accepted 7 June 1999*]

ABSTRACT

The stacking fault and twin boundary energies of C15 Cr_2Nb are calculated by the first-principles local-density-functional approach. It is found that the intrinsic and extrinsic stacking fault energies are 116 and 94 $mJ\,m^{-2}$, respectively, and the twin boundary energy is 39 $mJ\,m^{-2}$. The lower extrinsic stacking fault energy is consistent with the fact that the C36 structure has a lower energy than the C14 structure. The calculated stacking fault energies at 0 K are larger than the experimental values available in the literature. The equilibrium separations between Shockley partials based on the calculated elastic constants and stacking fault energies are also calculated.

§ 1. INTRODUCTION

For a wide variety of structural applications, Laves phases have some unique properties such as high melting temperature, low density, and high oxidation resistance. Unfortunately, this potential has not been well exploited, largely because of low temperature brittleness due to the lack of plastic deformation. Cr_2Nb, one of the most studied Laves phase compounds, has either the cubic C15 or hexagonal C14 (or C36) structure, which is topologically akin to the face-centred cubic (fcc) (A1) or the hexagonal close-packed (hcp) (A3) structure, or any one of the polytypic phases. The topologically close-packed (TCP) plane of the C15 structure is of the $\{111\}$ type, and the potential modes of plastic deformation in Cr_2Nb are twelve $\{111\}\langle 110\rangle$ slip systems and twelve $\{111\}\langle 112\rangle$ twin systems. Since each of the TCP units consists of the quadruple atomic layers, the slip, twinning, or stress-induced polytypic transformation will require a coordinated process of atomic motions, such as synchroshear, in order to effect the motion of Shockley partial dislocations.

Among other quantities, the knowledge of stacking fault energy (SFE), γ_{SF}, is necessary to understand the deformation mechanism, since SFE will play an important role in processes such as dislocation dissociation, cross-slip and twinning. First-principles calculations are useful to estimate the energetics of stacking faults (SFs) and understand the interaction between Shockley partials.

Using the linear muffin-tin orbital (LMTO) method, Chu *et al.* (1995a) estimated the intrinsic stacking fault energy of Cr_2Nb from the structural energy difference between the C15 and C14 structures, and obtained $\gamma_{SF} = 90\,mJ\,m^{-2}$. On the other

† Present address: Department of Physics, Sejong University, Seoul 143-747, Korea.

Philosophical Magazine A ISSN 0141–8610 print/ISSN 1460-6992 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journ als/tf/01418610.html

hand, Yoshida *et al.* (1995) obtained a value of $8\,\mathrm{mJ\,m^{-2}}$ for SFE from transmission electron microscopy (TEM) observations of extended dislocation nodes at 1623°C. More recently, Kazantzis *et al.* (1996) obtained a value of $25\,\mathrm{mJ\,m^{-2}}$ from TEM observations of extended nodes in Cr_2Nb at 1400 and 1500°C.

In a previous paper (Hong and Fu 1999), we investigated the phase stability of three Laves phases (C15, C14 and C36) of Cr_2Nb. It was found that the C15 phase is the ground-state structure with the lowest energy and the C36 phase is an intermediate state between C15 and C14. These three phases, however, are very close in energy, i.e. within a range of about 60 meV/formula unit (Hong and Fu 1999) indicating the possibility of low stacking fault energies in this system. In this paper, we report the calculation of SFE using supercell geometry, and an evaluation of the equilibrium separation between Shockley partials using the anisotropic elastic theory.

§ 2. Laves structures

In describing the structures of Laves phases and their stacking faults and twin boundary, we follow the notations of Hazzledine (1994).

Laves phases have ideal chemical compositions S_2L; they contain smaller atoms S and larger atoms L in alternate sheets parallel to the TCP planes (i.e. (111) plane for the C15 structure and (0001) plane for the C14 and C36 structures). The main geometric characteristic of Laves phases S_2L is that they consist of two types of atomic stacking sequence, $\alpha A\alpha$ ($\beta B\beta$ and $\gamma C\gamma$) and $\alpha c\beta$ ($\beta a\gamma$ and $\gamma b\alpha$). Here, the Greek letters (α, β, γ) denote the L atoms, while lower case Latin letters (a, b, c) and capital letters (A, B, C) denote the type-1 and type-2 S atoms, respectively. Note that the $\alpha c\beta$-type stackings are more closely spaced (in terms of the interlayer spacings). For Cr_2Nb, Latin and Greek letters represent Cr and Nb, respectively.

Figure 1 represents an atomic layer of A atoms based on the hexagonal unit cell on the TCP plane (a Kagome net) which is determined by two lattice vectors

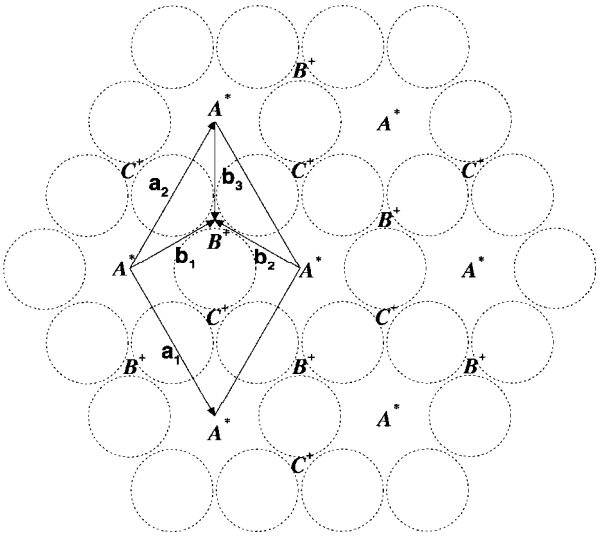


Figure 1. The hexagonal unit cell on the basal planes which is determined by two lattice vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. See the text for details.

$\mathbf{a}_1 = \frac{1}{2}[0\bar{1}1]$ and $\mathbf{a}_2 = \frac{1}{2}[\bar{1}10]$. The vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ are Shockley partial Burgers vectors of the type $\frac{1}{6}\langle\bar{1}12\rangle(\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 0)$. The dotted circles in figure 1 denote A atoms, one of the type-2 S atoms, projected on the (111) plane. Let the centres of hexagonal rings consisting of A atoms be denoted by A^*, and thus the A atoms are placed midway between the lattice points A^*. Similarly, the lattice points B^* and C^*, which determine B and C atoms respectively, can be obtained by the displacement of A^* by $\mathbf{b}_1$ and $-\mathbf{b}_1$, respectively, on higher or lower planes. Also, we denote the projected points of B^* and C^*onto the plane given in figure 1 as B^+ and C^+, respectively, which are the saddle points for larger α atoms. Then, the projected atomic positions of α (or a) onto the given plane are at the sites of A^* in figure 1, and the projected positions of β (or b) and γ (or c) are at those of B^+ and C^+, respectively. Consequently, the TCP C15 structure is defined by repeated XYZ stacking:

$$\text{C15}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{1}$$

where the TCP plane unit X (Y or Z) consists of one single layer of A (B or C) atoms and one triple layer of αcβ (βaγ or γbα). A synchroshear mechanism was first introduced by Kronberg (1957), and was used to explain the deformation twinning process by several groups (Livingston and Hall 1990, Chu and Pope 1993, Hazzledine 1994). Synchroshearing of the X unit (Aαcβ) creates the X′ unit (Aαbγ), and similarly for the Y′ (Bβcα) and Z′ (Cγaβ) units. The C14 and C36 Laves phases have repeated X′Z and XYZ′Y′ stackings, respectively, as follows:

$$\text{C14}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{b}\gamma}^{\mathbf{X'}}\,\overbrace{\mathrm{C}\gamma\mathrm{b}\alpha}^{\mathbf{Z}}\cdots, \tag{2}$$

$$\text{C36}: \quad \cdots \overbrace{\mathrm{A}\alpha\mathrm{c}\beta}^{\mathbf{X}}\,\overbrace{\mathrm{B}\beta\mathrm{a}\gamma}^{\mathbf{Y}}\,\overbrace{\mathrm{C}\gamma\mathrm{a}\beta}^{\mathbf{Z'}}\,\overbrace{\mathrm{B}\beta\mathrm{c}\alpha}^{\mathbf{Y'}}\cdots. \tag{3}$$

It can be seen that the primed X′, Y′, Z′ units are introduced by synchroshear in order to make sure that TCP structure is maintained in C14 and C36. For example, c and β of the X unit in C15 are synchrosheared by $-\mathbf{b}_2$ and $-\mathbf{b}_1$, respectively, to become b and γ of X′ unit. This displacement preserves a close packing between X and Z units in C14.

The stacking sequences for an intrinsic stacking fault (ISF) and an extrinsic stacking fault (ESF) and a twin boundary are given in table 1 in terms of X, Y, Z units. Note that the ISF contains a local C14-like structure in C15, while the ESF contains a local C36-like structure in C15. Figure 2 shows schematic illustrations of an ISF and an ESF obtained from C15 through the synchroshear mechanism: (*a*) C15 → ISF, and (*b*) C15 → ESF. The C15 stacking has a repeated XYZ sequence (no stacking fault). For the ISF, one unit, namely Y, is missing from the C15 XYZ sequence, and X becomes X′ by synchroshear (shifts by $-\mathbf{b}_2$ and $-\mathbf{b}_1$ shown in figure 2 (*a*)), while for an ESF a unit Y′ is added into the original C15 sequence. For the ESF, two successive synchroshears (two synchro-shifts shown in figure 2 (*b*)) are operative in the αcβ type to maintain a TCP structure by introducing two primed units, Z′ and Y′. Twinned structure is equivalent to a mirror reflection about the X′ unit (see table 1).

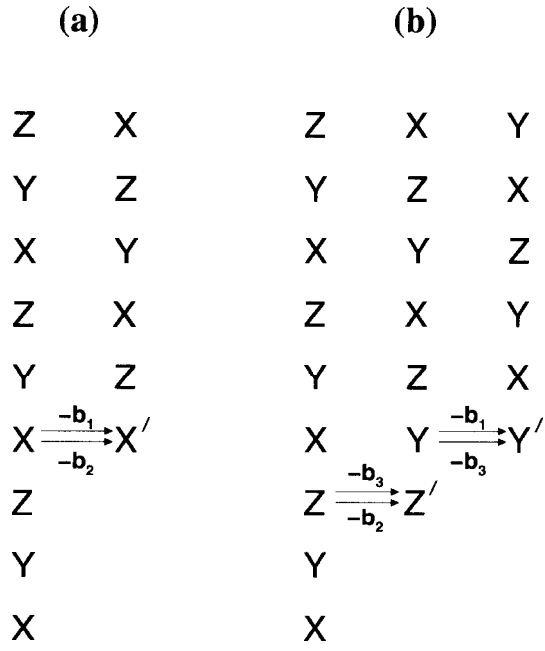


Figure 2. Schematic illustrations of an ISF and an ESF obtained from C15 → ISF, and (*b*) C15 → ESF.

Table 1. Stacking sequences of the C15 structure (no stacking fault), intrinsic stacking fault, extrinsic stacking fault, and twin (C15 + C15T) structure.

	Sequence of faults
C15	···XYZXYZXYZ···
Intrinsic stacking fault	···XYZX′ZXYZ···
Extrinsic stacking fault	···XYZXYZ′Y′XYZ···
Twin (C15 + C15T)	···XYZX′Z′Y′X′Z′Y′···

§ 3. Stacking fault and twin boundary energies

Total-energy calculations for the stacking fault energies of C15 Cr_2Nb are performed using the full-potential linearized augmented plane-wave (FLAPW) method (Wimmer *et al.* 1981) within the local-density approximation. The FLAPW method solves the local-density-functional equations without any shape approximation to the potential or charge density. The atomic positions are relaxed by calculating Hellmann–Feynman forces acting on the atoms.

The supercell geometry is used to obtain the energies of ISF, ESF and twin boundary. In the supercell calculation, we use the experimental lattice constant (6.991 Å) of C15 Cr_2Nb and spacings along $\mathbf{a}_3$ axis corresponding to the ideal hcp c/a ratio, which gives ~4. 04 Å for the thickness of each X, Y, Z unit. We consider a supercell containing XYZX′Z for an ISF and XYZXYZ′Y′ for an ESF. In these supercells, the separations between ISF and ESF planes are about 20 Å and 28 Å,

respectively. Also, a supercell containing XYZX′Z′Y′ (~24 Å) is considered for twin boundary. The supercell for twinning contains two twin boundaries in it, thus the separation between twin boundaries is about 12 Å.

The internal coordinates of each structure are fully relaxed from the ideal positions (defined by the atomic positions of the C15 lattice) by calculating Hellmann–Feynman forces acting on the atoms. The relaxation energies of ISF, ESF, and twin boundary are 41, 20, and 17 mJ m^{-2}, respectively.

The results for the stacking fault and twin boundary energies are given in table 2. It is found that the fault energies are γ_{ISF} = 116 and γ_{ESF} = 94 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is γ_T = 39 mJ m^{-2}. Since ESFs and ISFs contain local C36-like and C14-like structures, respectively, a lower ESF energy (γ_{ESF}) compared to the ISF energy (γ_{ISF}) is consistent with the fact that the C36 structure has a lower energy than the C14 structure (Hong and Fu 1999).

For comparison, we also calculated the stacking fault and twin boundary energies using the theoretical lattice constant (6.822 Å). The relaxation energies of ISF, ESF, and twin boundary are 48, 40, and 21 mJ m^{-2}, respectively. It is also found that the fault energies are 140 and 108 mJ m^{-2} for ISF and ESF, respectively, and the twin boundary energy is 52 mJ m^{-2}. The results using the theoretical lattice constant are slightly larger than those using the experimental one.

Although the intrinsic faults are expected to prevail in fcc crystals rather than the extrinsic faults (Hirth and Lothe 1982), extrinsic stacking faults were observed to be dominant in Laves phase Co_2Ti (Allen *et al.* 1972), and Nb-doped HfV_2 (Chu *et al.* 1998). Since the calculated γ_{ESF} is smaller than γ_{ISF}, it is likely that the observed SFs in Cr_2Nb are also of extrinsic type.

We compare our results with others. Chu *et al.* (1995a) obtained γ_{ISF} = 90 mJ m^{-2} by an estimation from the structural energy difference between the bulk C14 and C15 structures. This value is close to our value of 116 mJ m^{-2}. On the other hand, Yoshida *et al.* (1995) observed extended dislocation nodes in C15 Cr_2Nb deformed at 1623 K, in which the SFs are bounded by three Shockley partials (of the type $\frac{1}{6}\langle 112\rangle$) with a radius of curvature of R. They obtained a smaller value of 8 mJ m^{-2}. Note that Chu *et al.* (1995a) re-estimated SFE from the data of Yoshida *et al.* (1995) to obtain γ_{SF} = 15–60 mJ m^{-2}. More recently, Kazantzis *et al.* (1996) obtained a value of 25 mJ m^{-2} from TEM observations of extended triple-junction nodes in Cr_2Nb at 1400 and 1500°C. These results are tabulated in table 2.

Our results for SFEs are higher than the experimental results. Certainly, SFE obtained at high temperatures can be expected to be lower than the calculated value at 0 K, which has also been suggested by Kazantzis *et al.* (1996). The unknown image

Table 2. Stacking fault and twin boundary energies for C15 Cr_2Nb, using the experimental lattice constant.

Method	γ_{SF} (mJ m^{-2})		γ_T (mJ m^{-2})
This work	116 (intrinsic)	94 (extrinsic)	39
LMTO[a]	90 (intrinsic)		—
Exp[b]	25		—
Exp[c]	>8		—

[a] Chu *et al.* (1995a): estimated from energy difference between C14 and C15.
[b] Kazantzis *et al.* (1996): estimated from dislocation triple-junction.
[c] Yoshida *et al.* (1995): estimated from dislocation triple-junction.

shift of Shockley partial dislocations can also give an uncertainty in the SFE determination; for example, for SFE in silicon, $\gamma_{SF} = 30\,\mathrm{erg\,cm^{-2}}$ has an uncertainty in the range $15 < \gamma_{SF} < 75\,\mathrm{erg\,cm^{-2}}$, after taking into account the image shift (Aerts *et al.* 1962). From these points of view, the temperature and image shift effects may be responsible for the difference between our calculated value and experimental results.

Our theoretical calculations on Cr_2Nb were performed for the stoichiometric composition. In experiments, however, some localized variations in chemical composition are possible in polycrystalline compounds. For instance, the two Cr_2Nb alloys investigated by Yoshida *et al.* (1995) were Cr–32. 2% Nb and Cr-34. 0% Nb, in which the second phase particles observed are Cr solid solution in the former and Nb solid solution in the latter. Also, polycrystals used in experiments may have compositions that deviate from the ideal stoichiometry ratio, since the phase region of the C15 Cr_2Nb is relatively large. For example, as mentioned by Hong and Fu (1999), the calculated elastic moduli for the stoichiometry alloy at the experimental lattice constant are very different from the experimental values (Chu *et al.* 1995b) obtained from polycrystals. Therefore, this composition effect can also be partially responsible for the discrepancy between theory and experiment.

The calculated twin boundary energy, $\gamma_T = 39\,\mathrm{mJ\,m^{-2}}$, at the experimental lattice constant for Cr_2Nb is relatively low, for instance, in comparison to $\gamma_T = 60\,\mathrm{mJ\,m^{-2}}$ for TiAl of the Ll_0 structure (Fu and Yoo 1990). In view of the energetics of twin nucleation, such a low twin boundary energy suggests a high propensity of twinning in Cr_2Nb. This is consistent with the experimental observations of twinned microstructures in Cr_2Nb, formed due to plastic deformation at elevated temperatures (Yoshida *et al.* 1995) and to the internal stresses resulting from the C14–C15 transformation as well as the thermal contraction differences between Cr solid solution and Cr_2Nb in the two-phase alloy (Kumar and Liu 1997). However, the absence of deformation twinning in Cr_2Nb at low temperatures is not understood. Kinetic aspects of the motion of synchro-Shockley partials need to be elucidated in order to better understand twin formation in the C15 Laves phase.

§ 4. Interaction between Shockley partials

From the anisotropic elasticity theory (Stroh 1958, Hirth and Lothe 1982) we calculate the equilibrium separation between Shockley partials, using the calculated elastic constants (Hong and Fu 1999) and stacking fault energies.

First, we consider the following case for the ISF:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \mathrm{ISF} + \frac{1}{6}[2\bar{1}\bar{1}], \tag{4}$$

where $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$, $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$ correspond to $-\mathbf{a}_2$, $\mathbf{b}_3$, and $-\mathbf{b}_1$ in figure 1, respectively. This expression for the formation of an ISF is approximate since the Shockley partial $\mathbf{b}_3 = \frac{1}{6}[121]$ is the sum of two synchro-Shockley partial vectors $-\mathbf{b}_2$ and $-\mathbf{b}_1$, lying in the two successive atomic planes, shown in figure 2 (*a*). The equilibrium separation between Shockley partials can be obtained through the balance of the attractive force (the surface tension due to the ISF) and the repulsive force (due to elastic interaction between the partials).

The radial and tangential components of the interaction force (per unit length) between two parallel dislocations can be calculated by

$$F_r = f_r/(2\pi r), \tag{5}$$

$$F_\theta = f_\theta/(2\pi r). \tag{6}$$

Here, r is the separation between the two partials of Burgers vectors $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$, and f_r and f_θ are the radial (in the $\{111\}$ plane) and tangential (out-of-the plane) components of the interaction force constants, which are determined from the anisotropic elasticity theory of dislocations (Stroh 1958). For the dissociation of equation (4), the interaction force constants f_r and f_θ are obtained numerically (Yoo 1987), by (i) letting $\mathbf{b}^{(1)} = \frac{1}{6}[1\bar{2}1]$ and $\mathbf{b}^{(2)} = \frac{1}{6}[2\bar{1}\bar{1}]$, (ii) using the experimental lattice constant $a_0 = 6.991$ Å for C15, and (iii) the calculated elastic constants and γ_{ISF} at experimental lattice constant.

As shown in figure 3, the radial component f_r increases monotonically from $\phi = 0$ (screw) to $\phi = 90°$ (edge), where ϕ is the angle between a dislocation line and the Burgers vector **B**. The tangential component f_θ is maximum in magnitude at $\phi = 31°$ and zero at the edge orientation. At $\phi \sim 30°$, the two parallel Shockley partials are inclined at about 60° and 0° to their respective Burgers vectors. This implies that while the repulsive F_r balances the surface tension γ_{SF} (see equation (7) below), the out-of-plane force ($F_\theta = 0.5F_r$ at $\phi = 30°$) promotes cross-slip of the screw Shockley partial and climb of the 60° Shockley partial, most likely onto the $(\bar{1}01)$ plane. This Shockley partial dissociation configuration of the cross-slip and climb combination may lead to a possible mechanism for the thickening process of a twin embryo originating from a mixed ($\phi = 30°$) $\frac{1}{2}[1\bar{1}0]$ dislocation.

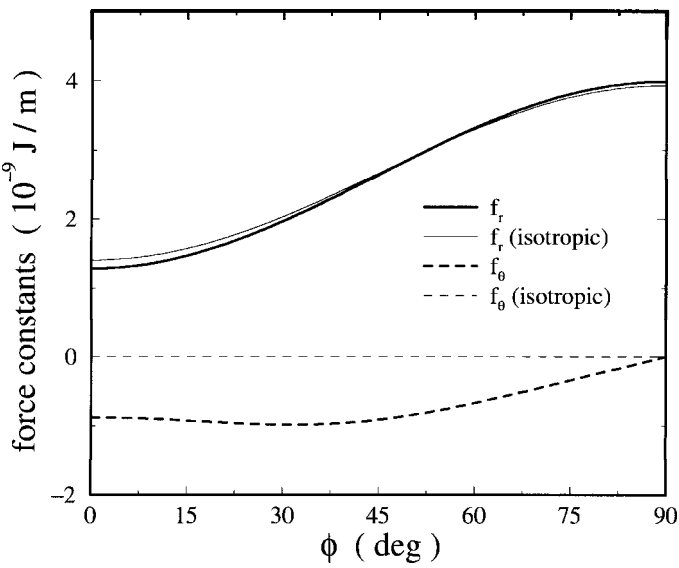


Figure 3. Orientation dependence of the interaction force constants in Cr_2Nb for the ISF-type dissociation, using the results obtained at the experimental lattice constant. ϕ is an angle between a dislocation line and the Burgers vector $\frac{1}{2}[1\bar{1}0]$.

The width of the equilibrium separation, w, can be obtained from

$$\gamma_{\mathrm{SF}} = F_r = f_r/(2\pi w), \tag{7}$$

$$w = f_r/(2\pi\gamma_{\mathrm{SF}}). \tag{8}$$

The separation is $w_{\mathrm{s}} = 17.6$ Å for the $\mathbf{B} = \frac{1}{2}[1\bar{1}0]$ screw dislocation, and $w_{\mathrm{e}} = 54.7$ Å for the edge dislocation. These results are shown in table 3, together with those in the isotropic case (discussed below). Also, the results using γ_{ESF} and the experimental result are tabulated.

For the isotropic case, the equilibrium separation w is given by a simple formula (Hirth and Lothe 1982):

$$w = \frac{Gb^2}{8\pi\gamma_{\mathrm{SF}}}\frac{2-\nu}{1-\nu}\left(1 - \frac{2\nu\cos 2\phi}{2-\nu}\right). \tag{9}$$

Here, $b = |\mathbf{b}^{(1)}| = |\mathbf{b}^{(2)}| = a_0/6^{1/2}$. The Hill's average values of shear modulus and the Poisson's ratio are $G = 50.0$ GPa and $\nu = 0.383$, respectively. Using γ_{ISF}, it is found that $w_{\mathrm{s}} = 19.3$ Å and $w_{\mathrm{e}} = 53.9$ Å. As given in table 3, the results in anisotropic and isotropic cases are rather close, because the shear anisotropy of C15 Cr_2Nb is moderate, $A = 1.45$ (Hong and Fu 1999), compared to the isotropic case ($A = 1$).

The radial component f_r of the isotropic case is very close to that of the anisotropic one, whereas its tangential component f_θ is exactly zero (figure 3). While the anisotropic corrections to f_r and hence to w in equation (9) are very small, the anisotropic tangential component of the elastic interaction force is quite large; for instance, $f_\theta/f_r \sim 0.7$ at $\phi = 0$. This indicates that cross-slip of $\frac{1}{2}[1\bar{1}0]$ dislocation is difficult in Cr_2Nb because of a large constriction energy for the Shockley partials that include the non-radial component of the interaction energy.

Next, let us consider the ESF case:

$$\frac{1}{2}[1\bar{1}0] \rightarrow \frac{1}{6}[1\bar{2}1] + \frac{1}{6}[11\bar{2}] + \mathrm{ESF} + \frac{1}{6}[2\bar{1}\bar{1}] + \frac{1}{6}[\bar{1}\bar{1}2] \tag{10}$$

$$\approx \frac{1}{6}[2\bar{1}\bar{1}] + \mathrm{ESF} + \frac{1}{6}[1\bar{2}1]\,. \tag{11}$$

To calculate the width of the ESF, we approximate two Shockley partials on two successive TCP units on either side of the ESF by a single Shockley partial (as in the ISF case), and use the same formulas for the ISF. In other words, the only difference

Table 3. Equilibrium separation between Shockley particles bounding an ISF and an ESF, using the results obtained at experimental lattice constant. w_{s} and w_{e} are the width for screw and edge dislocations, respectively. The numbers are in angstroms.

		w_{s} (screw)		w_{e} (edge)	
		ISF	ESF	ISF	ESF
Theory ($T = 0$ K)	Anisotropic	18	22	55	68
	Isotropic	19	24	54	67
Exp[a] ($T = 1400$°C)		99		—	

[a] Kazantzis *et al.* (1996): measured from the ribbon after the correction for image shift and projection effects.

between ISF and ESF in the calculation of separation w is that the stacking fault energy used in the formulas is of intrinsic or extrinsic type. Using γ_{ESF}, it is found that $w_s = 21.7$ Å and $w_e = 67.5$ Å for the anisotropic case (table 3). For the isotropic case, they are $w_s = 24.4$ and $w_e = 66.5$ Å. As mentioned, the results of both anisotropic and isotropic cases are very close. Since γ_{ESF} is not much different from γ_{ISF}, the force constants f_r and f_θ using γ_{ESF} are expected to be similar to those in figure 3 using γ_{ISF}.

On the other hand, we consider the separation between partials by using the theoretical lattice constant. For the ISF, the separation is $w_s = 17.1$ Å and $w_e = 53.5$ Å for the anisotropic case, while $w_s = 18.7$ Å and $w_e = 52.8$ Å for the isotropic case. For the ESF, the separation is $w_s = 22.2$ Å and $w_e = 69.4$ Å for the anisotropic case, while $w_s = 24.2$Å and $w_e = 68.4$ Å for the isotropic case. It can be seen that the separations are very close when using both experimental and theoretical lattice constants.

Kazantzis *et al.* (1996) measured, at $T = 1400°C$, the width of the ribbon (separation between two parallel partial dislocations) as 99 Å after correction for image shift and projection effects. Note that they also obtained a different value of $w_s = 82$ Å for the width using equation (9) with $\gamma_{SF} = 25\,mJ\,m^{-2}$, which was estimated from the measured curvature of R in extended dislocation nodes and the temperature-corrected shear modulus. Compared with the experimental result at high temperatures, our results for the separation between partials are very small.

§ 5. Summary

We performed first-principles total-energy calculations to obtain stacking fault and twin boundary energies. The intrinsic and extrinsic stacking fault energies were calculated to be 116 and 94 mJ m^{-2}, respectively, and the twin boundary energy was 39 mJ m^{-2}. The calculated stacking fault energies are larger than the available experimental data, measured at high temperatures. We also calculated the equilibrium separations between Shockley partials using the calculated elastic constants and stacking fault energies. Our results of the equilibrium separations are very small compared with the experimental results reported at high temperatures. This discrepancy may be due to temperature and/or composition and image shift effects.

Acknowledgements

The research was sponsored by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

References

Aerts, E., Delavignette, P., Siems, R., and Amelinckx, S., 1962, *J. appl. Phys.*, **33**, 3078.
Allen, C. W., Delavignette, P., and Amelinckx, S., 1972, *Phys. status solidi* (a), **9**, 237.
Chu, F., Ormeci, A. H., Mitchell, T. E., Wills, J. M., Thoma, D. J., Albers, R. C., and Chen, S. P., 1995a, *Phil. Mag. Lett.*, **72**, 147.
Chu, F., He, Y., Thoma, D. J., and Mitchell, T. E., 1995b, *Scripta metall. mater.*, **33**, 1295.
Chu, F., Lu, Y.-C., Kotula, P. G., Mitchell, T. E., and Thoma, D. J., 1998, *Phil. Mag. A*, **77**, 941.

CHU, F., and POPE, D. P., 1993, *High-Temperature Ordered Intermetallic Alloys V*, *Materials Research Society Symposium Proceedings*, Vol. 288, edited by I. Baker, R. Darolia, J. D. Whittenberger and M. H. Yoo (Materials Research Society), p. 561.
FU, C. L., and YOO, M. H., 1990, *Phil. Mag. Lett.*, **62**, 159.
HAZZLEDINE, P. M., 1994, *Twinning in Advanced Materials*, edited by M. H. Yoo and M. Wuttig (The Minerals, Metals & Materials Society), p. 403.
HAZZLEDINE, P. M., and PIROUZ, P., 1993, *Scripta metall. mater.,* **28**, 1277.
HIRTH, J. P., and LOTHE, J., 1982, *Theory of Dislocations* (New York: Wiley).
HONG, S., and FU, C. L., 1999, *Intermetallics*, **7**, 5.
KAZANTZIS, A. V., AINDOW, M., and JONES, I. P., 1996, *Phil. Mag. Lett.*, **74**, 129.
KRONBERG, M. L., 1957, *Acta metall.,* **5**, 507.
KUMAR, K. S., and LIU, C. T., 1997, *Acta mater.,* **45**, 3671.
LIVINGSTON, J. D., and HALL, E. L., 1990, *J. mater. Res.,* **5**, 5.
STROH, A. N., 1958, *Phil. Mag.,* **3**, 625.
WIMMER, E., KRAKAUER, H., WEINERT, M., and FREEMAN, A. J., 1981, *Phys. Rev.* B, **24**, 864.
YOO, M. H., 1987, *Acta metall.,* **35**, 1559.
YOSHIDA, M., TAKASUGI, T., and HANADA, S., 1995, *High-Temperature Ordered Intermetallic Alloys VI*, *Materials Research Society Symposium Proceedings*, Vol. 364, edited by J. Horton, I. Baker, S. Hanada, R. D. Noebe and D. S. Schwartz (Materials Research Society), p. 1395.

PERGAMON

Acta mater. 48 (2000) 911–923

Acta MATERIALIA

www.elsevier.com/locate/actamat

STRUCTURAL STABILITY OF THE LAVES PHASE Cr_2Ta IN A TWO-PHASE Cr–Cr_2Ta ALLOY

K. S. KUMAR[1]†, L. PANG[1], C. T. LIU[2], J. HORTON[2] and E. A. KENIK[2]

[1]Division of Engineering, Brown University, Providence, RI 02912, USA and [2]Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

(Received 17 August 1999; accepted 4 October 1999)

Abstract—Detailed microstructural analysis of a two-phase alloy of composition Cr–9.8 at.% Ta and lying in the Cr–Cr_2Ta region of the Cr–Ta binary system confirmed that the existing phase diagram is inaccurate; in the cast and annealed condition (1273 K/24 h), blocky primary Cr_2Ta precipitates were observed although the phase diagram indicates the eutectic composition to be ~13 at.% Ta. The eutectic structure is composed of Cr solid solution and the Laves phase Cr_2Ta; the morphology is primarily lamellar although the rod morphology was occasionally observed. The Laves phase eutectic microconstituent exhibits the C14 (2H) hexagonal structure with a low stacking fault (basal faults) density and an average composition corresponding to 28.5 at.% Ta. After a prolonged high-temperature anneal (1573 K/168 h), the morphology breaks down to form discrete particles of Cr_2Ta; the C14, C36 and C15 structures were all recognized in this annealed condition, often more than one form being present in a single precipitate. The C15 structure was not twinned but contained some stacking faults on the {111} planes. Composition measurements confirmed that these structural transformations were accompanied by composition changes, the precipitates becoming more Ta-rich as they transitioned from the C14 via the C36 to the C15 phase. These observations are coupled with the results from earlier studies to present a discussion on factors that influence the stability and C14/C36/C15 transformation kinetics. © *2000 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.*

Keywords: Intermetallic compounds; Laves phase; Phase transformation; Electron diffraction; Energy dispersive X-ray analysis (EDS)

1. INTRODUCTION

Binary alloys that include the chromium-containing Laves phase (i.e. Cr_2X) have been the subject of several recent investigations [1–7]. The Laves phase, Cr_2X, has a compositional range of existence in the respective binary systems (Cr–X, where X = Ti, Zr, Hf, Nb or Ta) and is in equilibrium with the two terminal solid solutions. In all these systems, the published binary phase diagrams [8] show the Laves phase to exhibit the hexagonal C14 (2H) or C36 (4H) structure at high temperatures and the cubic C15 (3C) structure at lower temperatures. At least in the Cr–Nb, and Cr–Ta systems, the C14-to-C15 transformation is shown to occur through a eutectoid or peritectoid transformation, depending on whether a selected two-phase alloy lies in the Cr solid solution + Cr_2X two-phase region or in the X solid solution + Cr_2X two-phase region. The published Cr–Ta phase diagram shown in Fig. 1 demonstrates some of these features.

Under normal solidification conditions, it is likely that these eutectoid and peritectoid reactions will not go to completion, and in these situations, the C14 structure can be metastably retained in the as-cast microstructure of two-phase alloys. This is indeed the case in two-phase alloys of the type Cr–Cr_2Hf [9] and Cr–Cr_2Ti [5], whereas a twinned C15 structure is observed in as-cast Cr–Cr_2Nb alloys [2–4, 10]. It is not clear why the transformation proceeds to completion so readily in the Cr–Nb alloy but not in the Cr–Ti or Cr–Hf alloy. Further, long-term annealing at elevated temperatures has confirmed the sluggish nature of the metastable C14-to-C15 transformation in the Cr–Ti and Cr–Hf alloys. Likewise, as-cast single-phase Cr_2Nb also exhibits the twinned C15 structure [2], whereas as-cast single-phase Cr_2Ti has been shown to maintain the C14 (or C14 + C36) structure at room temperature [5, 6] and to transform to the C15 form via the C36 structure following extended exposures to 1473 K.

The effect of a third element on the structural stability of the Cr_2Nb Laves phase and on the C14/C36/C15 transformation characteristics has been

† To whom all correspondence should be addressed.

1359-6454/00/$20.00 © 2000 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.
PII: S1359-6454(99)00377-8

recently investigated. Takasugi and Yoshida [11] found that when Mo or Ti was substituted for Nb in single-phase Cr_2Nb, the C14 (or C36) structure resulted in addition to the C15 structure in the as-cast condition, which upon extended annealing transformed to a twinned C15 structure. Similar observations were made when W was substituted either for Cr alone, or for Cr and Nb. The hexagonal-to-cubic transformation kinetics at 1673 K appeared to depend on the substitutional species, decreasing in the order Mo, Ti and W. An additional interesting observation that was made was that the boundary between the C14 (or C36) and C15 phase in a ternary Cr–Nb–Mo alloy was incoherent; this prompted the authors to argue that the transformation was one that involved both diffusion and shear. These authors [11] have considered a variety of factors that could affect the thermodynamic stability and kinetics of transformations in these Laves phases but most of the discussion is speculative.

Laves phase transformations are thought to occur by shear, and shear transformation kinetics are believed to be dictated by temperature dependence of dislocation mobility [1]. This may well be the case for the transformation of a metastable C14 structure to the stable C15 structure (often via the C36 structure) upon annealing. However, in two-phase alloys containing the metastable C14 structure, it may be necessary for the composition of the C14 phase to first adjust via diffusion to that necessary to stabilize the C15 structure before it can be sheared into the product phase. If the diffusional requirements are minimal and diffusion is adequately rapid, the metastable C14 structure resulting from the suppression of the eutectoid or peritectoid reaction could perhaps shear to the C15 structure *while cooling down*, making it possible to obtain the latter structure in the as-cast material.

In this paper, we report on the microstructure of the Laves phase in a cast binary two-phase Cr–Ta alloy and how it evolves with subsequent annealing at high temperature. Factors that could influence the C14/C15 transformation are discussed based on these observations and results from earlier studies.

2. EXPERIMENTAL PROCEDURE

The alloy used in this investigation (Cr–9.8 at.% Ta) was produced by drop casting into a copper mold after arc melting the constituents. The as-cast Cr–Ta ingot was then annealed at 1273 K/24 h to relieve the stresses. A Cr–Ta specimen was also annealed for 168 h at 1573 K and then character-

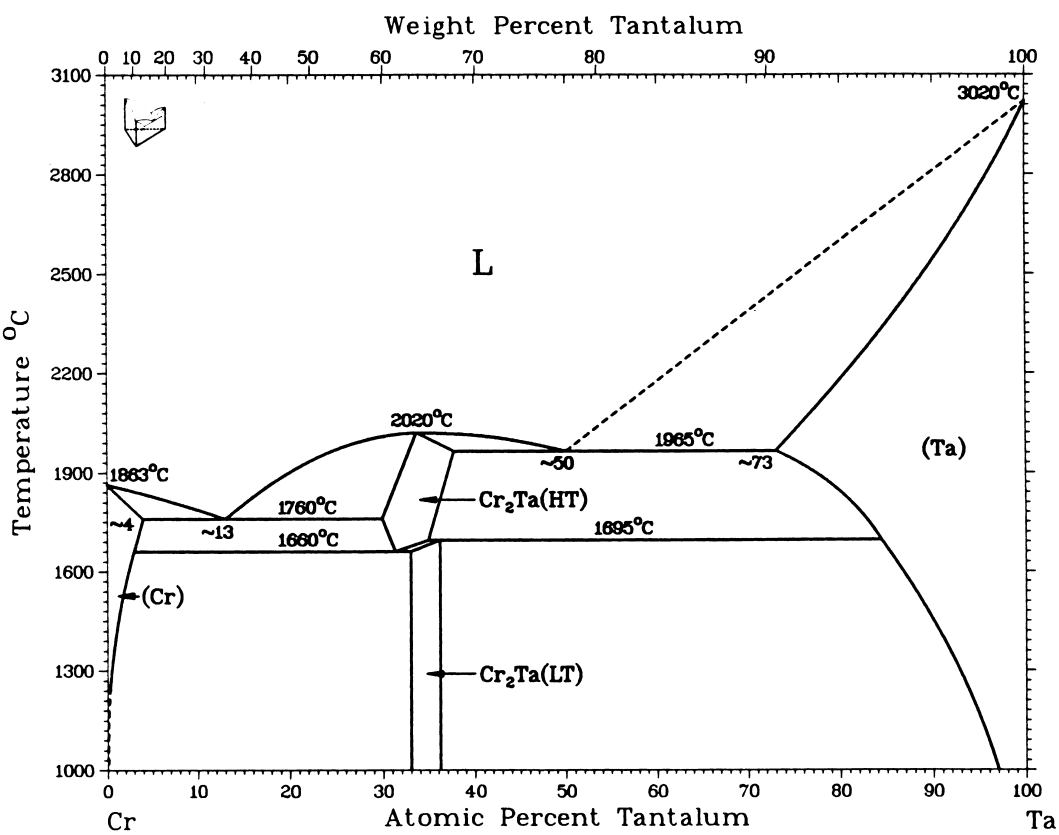


Fig. 1. The binary Cr–Ta phase diagram from Ref. [8].

ized. The microstructure of the alloys was characterized using optical microscopy, scanning electron microscopy (SEM) and transmission electron microscopy (TEM). Thin foil specimens for TEM observations were prepared by slicing sections of appropriate thickness, grinding them down and twin-jet polishing them to perforation using a solution of 10% perchloric acid and 90% methanol at 233 K and 15 V. When adequate electron-transparent area was not obtained using this approach, the perforated foils were further ion-milled for short lengths of time (typically 15–30 min). These specimens were examined using either a JEOL 100CX, a Phillips EM420 or a CM30 (300 kV) transmission electron microscope. Bright field imaging and selected area diffraction (SAD) methods were used to characterize the morphology and crystal structures of the phases present. The compositions of the phases were obtained using energy dispersive X-ray (EDX) techniques in an XL30/FEG SEM and in a CM12 (120 kV) transmission electron microscope. It should be noted that the SEM/EDX quantification was performed using pure Cr AND PURE Ta STANDARDS WHEREAS THE TEM/EDX ANALYSIS WAS STANDARDLESS.

3. RESULTS

According to the binary Cr–Ta phase diagram (Fig. 1), alloys whose composition participate in the eutectic reaction $L \Leftrightarrow Cr(\text{solid solution}) + Cr_2Ta$, should produce the C14 phase either only as a eutectic microconstituent (in hypoeutectic alloys) or as a primary phase from the liquid and as a eutectic microconstituent (in hypereutectic alloys). The eutectic microconstituent with the C14 structure (and the primary phase for relevant alloy compositions) must then undergo minor compositional changes (reject Cr with decreasing temperature) before decomposing via a eutectoid reaction at a lower temperature to form a two-phase mixture of Cr solid solution and C15 Cr_2Ta. The Cr solid solution in such alloys (primary phase and/or eutectic microconstituent) experiences a decrease in Ta solubility between the eutectic and eutectoid temperatures, and accommodates this decrease in solubility by precipitating C14 Cr_2Ta which also undergoes eutectoid decomposition (with accompanying compositional changes) to produce the C15 Cr_2Ta phase. Since these solid-state reactions involve significant compositional adjustments, it is likely that they will not go to completion under conventional solidification conditions. Further, if the eutectoid temperature is significantly lower than the eutectic temperature and the eutectoid composition is significantly different from the composition of the Laves phase at the eutectic temperature, then the compositional adjustments are even less likely to be accommodated during non-equilibrium freezing. The net result is that it is possible to metastably retain the C14 structure at low temperatures.

The as-cast Cr–9.8 at.% Ta alloy was annealed at 1273 K for 24 h. According to the existing binary phase diagram (Fig. 1), this alloy should correspond to a hypoeutectic composition. Optical microscopy, however, confirmed the microstructure of this alloy to be almost completely eutectic with some isolated coarse primary precipitates [Fig. 2(a)]. Both, lamellar and rod eutectic morphologies were observed. Following the prolonged high-temperature anneal (1573 K/168 h), the lamellar (and rod) morphology breaks down and the precipitates exist as discrete particles [Fig. 2(b)] that are coarser than the average width of the lamellae prior to the high-temperature anneal.

A low-magnification transmission electron micrograph [Fig. 3(a)] confirms the presence of three generations of the Cr_2Ta phase in the cast and 1273 K/24 h annealed specimen: a blocky primary phase with a hexagonal cross-section that was occasionally noted, the lamellar eutectic microconstituent, and solid-state precipitation in the previously denuded

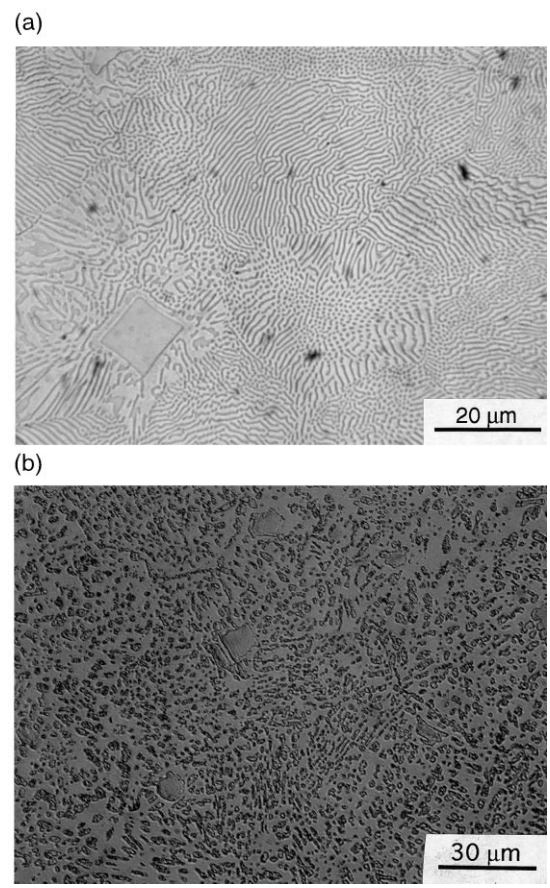


Fig. 2. An optical micrograph of the Cr–9.8Ta alloy: (a) in the cast and annealed condition confirming the presence of primary Cr_2Ta and eutectic Cr_2Ta with rod and lamellar morphology; (b) after annealing for 168 h at 1573 K showing the breakdown of the lamellar structure.

zone between the primary phase and the lamellar eutectic microconstituents. This microstructure is suggestive of an off-eutectic alloy that is marginally hypereutectic. This observation is consistent with the eutectic composition reported by Brady *et al.* [12], but is in conflict with the phase diagram shown in Fig. 1 where the eutectic composition is shown to correspond to ~13 at.% Ta. An examination of the eutectic microstructure at higher magnifications reveals that whereas in some grains, the lamellae are uniform and fairly evenly spaced [Fig. 3(b)], in others, the lamellar morphology breaks down, resulting in faceted rods and rod-to-plate transitions [Fig. 3(c)]. Such transitions in morphology are known to occur as a consequence of local changes in heat-flow patterns and/or composition fluctuations [13, 14]. In some locations where the Cr_2Ta lamellae terminate or form "Y" junctions, the effective spacing between adjacent lamellae changes and in these locations, solid-state precipitation of Cr_2Ta in the Cr solid solution matrix is often noted [Fig. 3(d)]. The composition of the blocky primary Laves phase precipitates was measured using SEM/EDX and, on average, corresponded to 29.7 at.% Ta (Table 1). In contrast, a similar analysis of the eutectic Laves phase microconstituent yielded an average value of 28.5 at.% Ta. This composition is leaner in Ta than that predicted by the phase diagram in Fig. 1. This difference may arise either as a consequence of non-equilibrium cooling of the alloy during drop casting or as a consequence of an inaccurate phase diagram (where the eutectic Laves phase composition is presented as ~30 at.% Ta).

A bright field image of a "Y" junction is shown in Fig. 4(a). This junction is located at the edge of the hole in the specimen. This region (marked X) was used to determine the crystal structure of the Cr_2Ta phase by selected area diffraction. Several zone axes were imaged and six of these are shown in Figs 4(b)–(g). These were consistently indexed to confirm that the Cr_2Ta phase that is a eutectic microconstituent exhibits a hexagonal structure. The relevant portion of a standard (0001) projection for a close-packed hexagonal structure is shown in Fig. 4(h) to illustrate the relationship between the different diffraction patterns shown in Figs 4(b)–(g). When the diffraction patterns are suitably indexed,

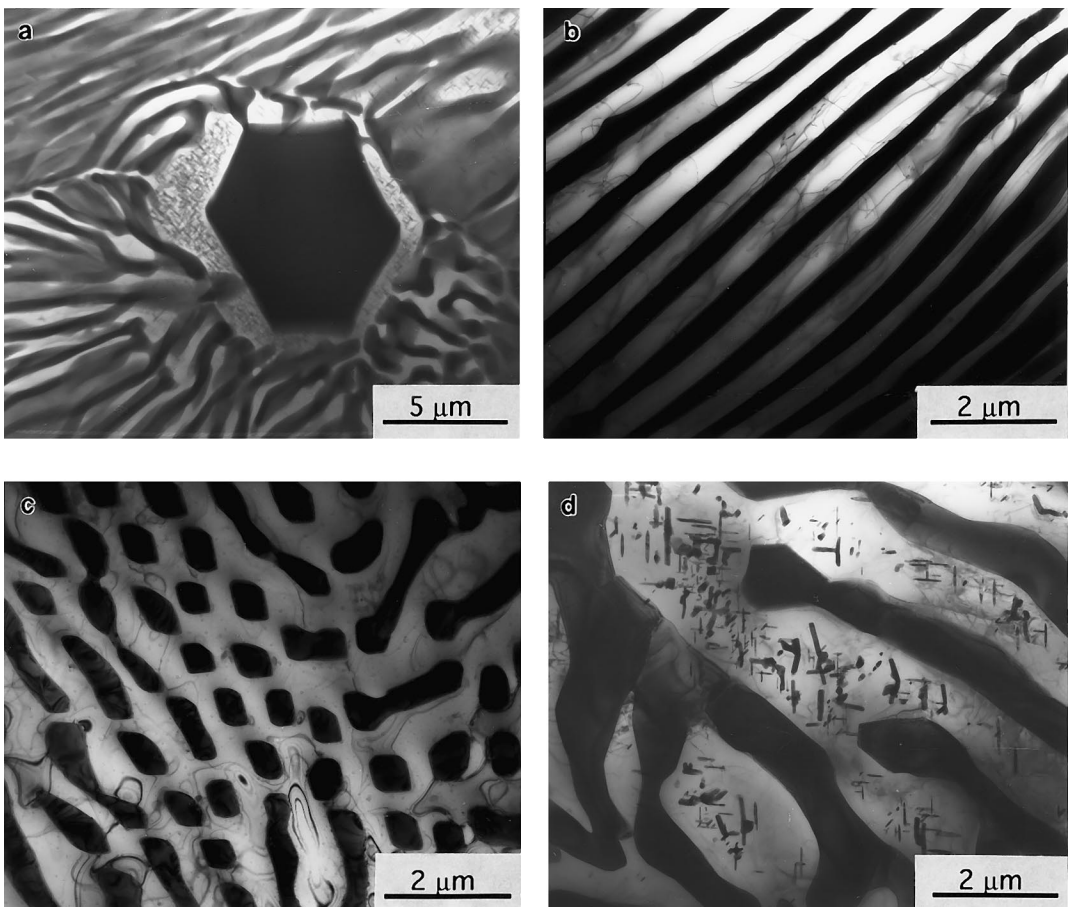


Fig. 3. Bright field images of: (a) primary Cr_2Ta with a precipitate denuded zone that subsequently filled out through solid-state precipitation; (b) lamellar Cr_2Ta eutectic microconstituent; (c) local breakdown of the lamellar morphology; (d) solid-state precipitation of Cr_2Ta between the lamellae.

it is concluded that the Cr_2Ta retains the C14 structure in the cast and annealed (1273 K/24 h) condition. The matrix was confirmed to have the b.c.c. structure expected for Cr solid solution.

A careful examination of some of the Cr_2Ta lamellae at high magnification reveals the presence of several evenly spaced, parallel planar features that exhibit fringe contrast [Fig. 5(a)]. By orienting these features "edge-on" [Fig. 5(b)] and from the associated selected area diffraction pattern corresponding to the $[11\bar{2}0]$ zone axis, we confirm that these features (believed to be stacking faults) lie on the basal plane of the C14 structure. This is analogous to previous observations in the Cr–Hf system [9] and in the Cr–Ti [5] systems where such faults were observed to occur on the basal planes of the C14 structure. The density of such faults on average however appears to be significantly lower in Cr_2Ta as compared to Cr_2Hf. Even though the as-cast microstructure was not specifically examined, it is likely that it was composed of Cr solid solution and C14 Cr_2Ta since the C14 structure is still metastably retained following annealing at 1273 K for 24 h.

Following the extended high-temperature anneal (1573 K/168 h), as shown in the optical micrograph in Fig. 2(b), the lamellar structure breaks down completely to yield discrete particles, although the blocky primary precipitates are still present. The average composition of the big blocky primary precipitates after annealing was 30.7 at.% Ta as determined by SEM/EDX (Table 1). Thus, some Ta enrichment (from 29.7% Ta prior to this anneal) of these precipitates is evident as a consequence of the prolonged high-temperature annealing. The discrete particles in the annealed structure resulting from the breakdown of the lamellar structure were also enriched in Ta with an average of 30.4 at.% Ta (compared with 28.5 at.% Ta before the anneal).

The microstructure after the extended high-temperature anneal (1573 K/168 h) was also examined by TEM. A bright field image of one such particle is shown in Fig. 6(a). This particle consists of two distinctly different regions, one of which is heavily faulted (region X) and the other which exhibits a low fault density (region Y). A $[11\bar{2}0]$ SAD pattern [Fig. 6(b)] confirms the region X to have the C36 (4H) hexagonal structure (or a higher order polytype such as 6H) while region Y was found to have the C15 structure as confirmed by the set of diffraction patterns in Figs 6(c)–(e) corresponding to the [101], [121] and [111] zone axes, respectively. A bright field image of another precipitate in the annealed specimen is shown in Fig. 7(a). Here, the two ends of the precipitate (region X for example) are free of stacking faults whereas the central region (region Y) is heavily faulted. A $[11\bar{2}0]$ SAD pattern from region X [Fig. 7(b)] confirms it to have the C14 (2H) hexagonal structure whereas a diffraction pattern from region Y [Fig. 7(c)] shows that this region has the C36 (4H) structure. A bright field image in Fig. 7(d) of a third precipitate shows the presence of multiple variants of stacking faults. This would not be possible for hexagonal structures with basal faults and a set of diffraction patterns corresponding to the [010], [111] and [121] zone axes in Figs 7(e)–(g) confirms this precipitate to have the cubic C15 (3C) structure. Thus it is evident that the 1573 K anneal for 168 h is instrumental in transforming the metastable C14 Laves phase observed in the cast + 1273 K/24 h specimen to C36 and C15 structures.

Since an individual precipitate contained within it regions with different crystal structures (indicative of transformations in progress) and SEM/EDX could not be used to track precipitate composition as a function of crystal structure, we used a standardless EDX technique in the TEM to measure composition within a precipitate as a function of crystal structure. Several individual precipitates were examined; heavily faulted regions, such as region Y in Fig. 7(a) were assumed to correspond to the C36 structures (or higher order polytypes such as a 6H structure) whereas regions containing multiple variants of faults were assumed to be C15 in nature [e.g. Fig. 7(d)]. Those regions that exhibited virtually no faults in the precipitate were thought to be of the C14 type [e.g. region X in Fig. 7(a)]. Although the compositions of the precipitates measured using the EDX attachment in conjunction with the TEM were on average richer in Cr compared with the values obtained using the SEM/EDX (Table 1), an internal comparison of compo-

Table 1. Measured compositions of the Laves phase Cr_2Ta in a binary Cr–9.8 at.% Ta alloy

Precipitate type	SEM/EDX[a] (at.% Ta)	TEM/EDX[b] (at.% Ta)
Primary (blocky) precipitate—cast + (1273 K/24 h) anneal	29.7 ± 0.4	–
Primary (blocky) precipitate—after prolonged anneal (1573 K/1 week)	30.7 ± 0.1[c]	–
Eutectic Cr_2Ta cast + (1273 K/24 h) anneal	28.5 ± 1.2	–
Globular Cr_2Ta after prolonged anneal (1573 K/1 week)	30.4 ± 1.0	25.8 ± 0.15 (C14 region[d])[c] 27.3 ± 0.15 (C36 region[d]) 28.6 ± 0.8 (C15 region[d])

[a] Pure Cr and pure Ta standards were used.
[b] Standardless analysis.
[c] Only two readings obtained; in all other cases eight to ten readings were obtained.
[d] See text for how these regions were structurally isolated.

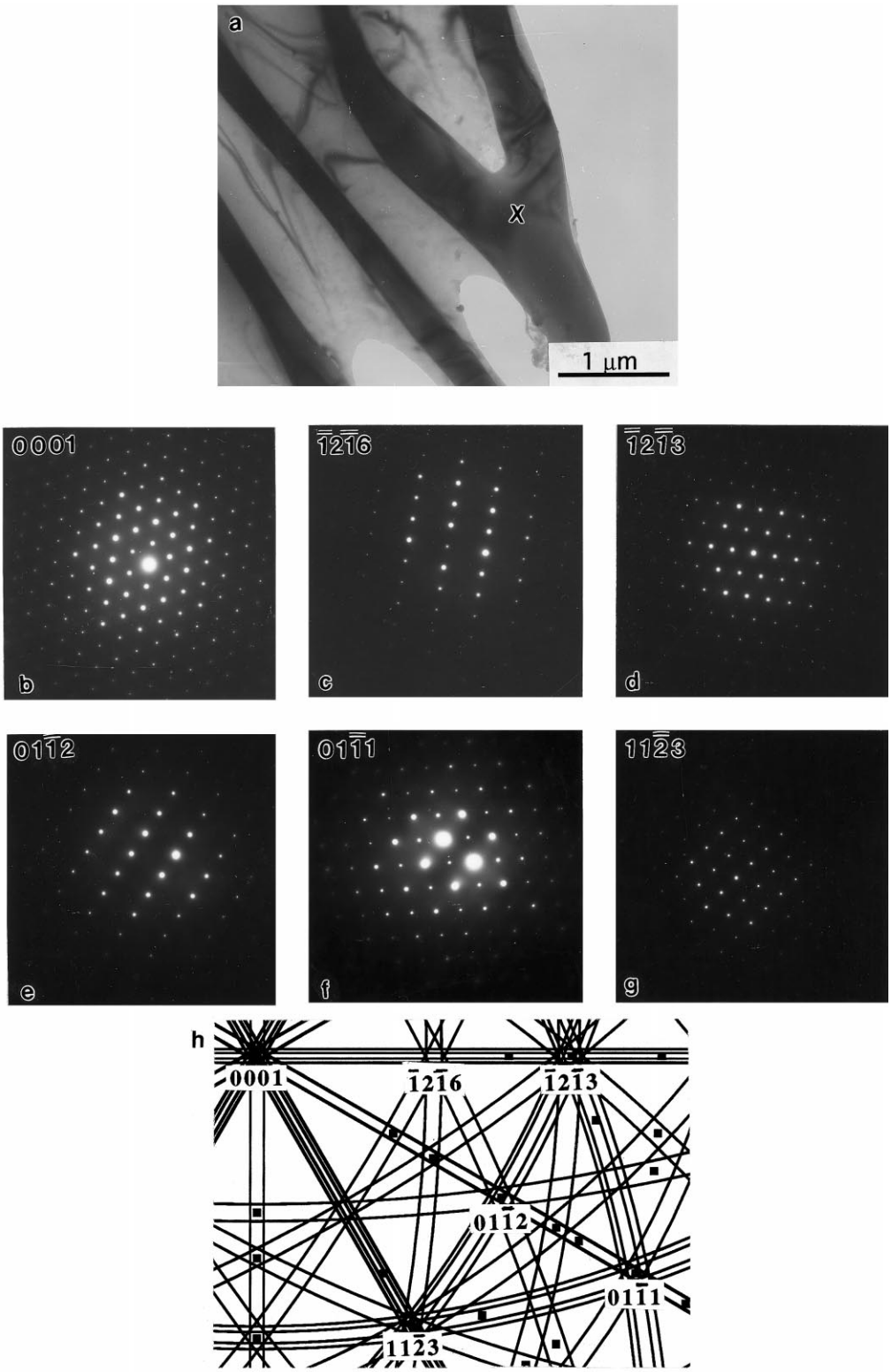


Fig. 4. (a) Bright field image of a "Y" junction in lamellar Cr_2Ta [X marks the location from where selected area patterns (b)–(g) were obtained]; (b)–(g) correspond to [0001], $[\bar{1}2\bar{1}6]$, $[\bar{1}2\bar{1}3]$, $[01\bar{1}2]$, $[01\bar{1}1]$ and $[11\bar{2}3]$ zone axes, respectively. In (h), the relevant portion of a standard (0001) projection is provided to illustrate the relationship between the diffraction patterns shown in (b)–(g).

sition variation with structure could still be made. Thus, regions thought to have the C14 structure, C36 and C15 structures corresponded to average compositions of 25.8, 27.3 and 28.6 at.% Ta, respectively. From these data, it can be seen that the structural transformation from C14 to C15 is accompanied by a gradual increase in Ta content.

4. DISCUSSION

Laves phases nominally have the formula AB_2 and their occurrence is dictated by the radius ratio of the two components A and B. Frequently, these compounds exhibit a stoichiometric range of existence, with the range being skewed to the smaller atom side and the deviation from stoichiometry being accommodated by anti-site defects [15, 16]. These compounds are known to occur with a hexagonal (C14 and C36) or cubic (C15) structure and the stability regime of the C14, C36 and C15 structures is believed to be governed by the electron concentration-to-atom ratio (e/a). Komura and Kitano [17] have demonstrated this electron concentration dependence of structure in Mg-based binary and ternary Laves phases. More recently, Zhu *et al.* [18] have presented a similar correlation for Cr_2Nb-based Laves phases. It is important to recognize however, that such regimes that are identified are empirical and based on structural information collected from existing phase diagrams; their accuracy is therefore limited by the accuracy of the phase diagrams. At elevated temperatures, these compositional regimes may be extended as a consequence of increased solubility, and in fact, Laves phase structures that are not stable at room temperature in some compositional regimes may be present at elevated temperatures. Examples can be seen in the Cr–Nb, Cr–Hf and Cr–Ta phase diagrams where the C14 structure is present at elevated temperature over a compositional range but not at all at lower temperatures. A second feature to note is that in these phase diagrams, at lower temperatures, the C15 structure is stable over a significant compositional range which is not necessarily entirely coincident with the range over which the C14 structure is stable at elevated temperatures, although some compositional overlap exists. However, it should be pointed out that the observed compositions of the C14, C36 and C15 Cr_2Ta in this investigation are consistent with the reported e/a criterion used for predicting the stability regimes of Laves phase types; that is, the decrease in e/a as a consequence of increasing Ta concentration results in the increase in stability of the cubic C15 phase.

A schematic phase diagram that captures the essential features of the Cr–X (X = Nb, Ta and Hf) is shown in Fig. 8. In the Cr–Ti and Cr–Zr systems, a stable C36 phase field is present and

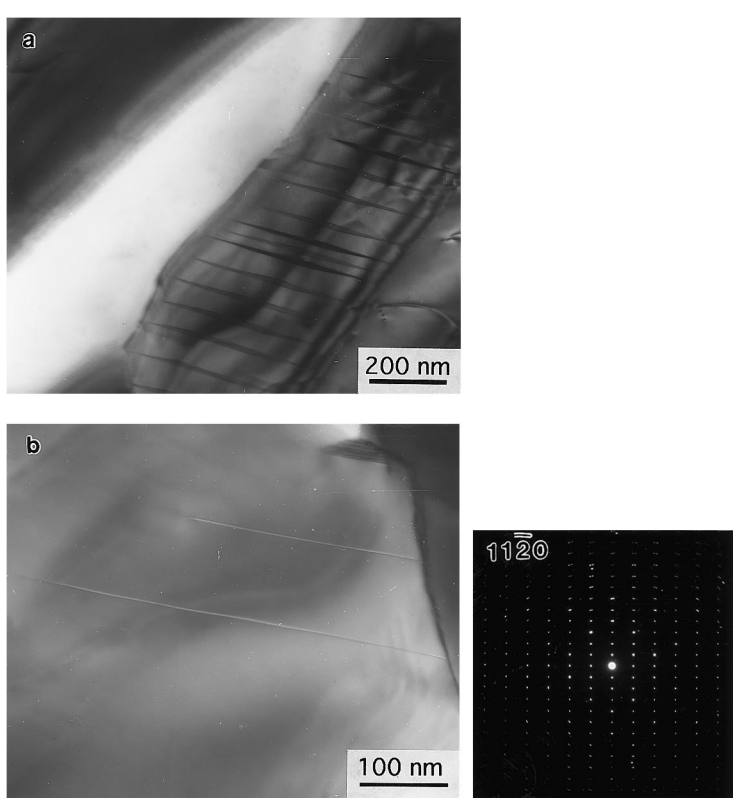


Fig. 5. Bright field images confirming: (a) the presence of planar defects in the lamellar Cr_2Ta; (b) that these faults lie on the basal plane in the C14 structure.

for the purpose of clarity, this feature is omitted from Fig. 8. Nevertheless, the following concepts can also be applied to the Cr–Ti and Cr–Zr systems with some modifications. We will use Fig. 8 to develop our discussions on the structural stability and transformation kinetics of Laves phases and draw from our own results and those of others to support some of the arguments.

We can start by considering the alloy compositions that fall within the area "pqsr"; the range

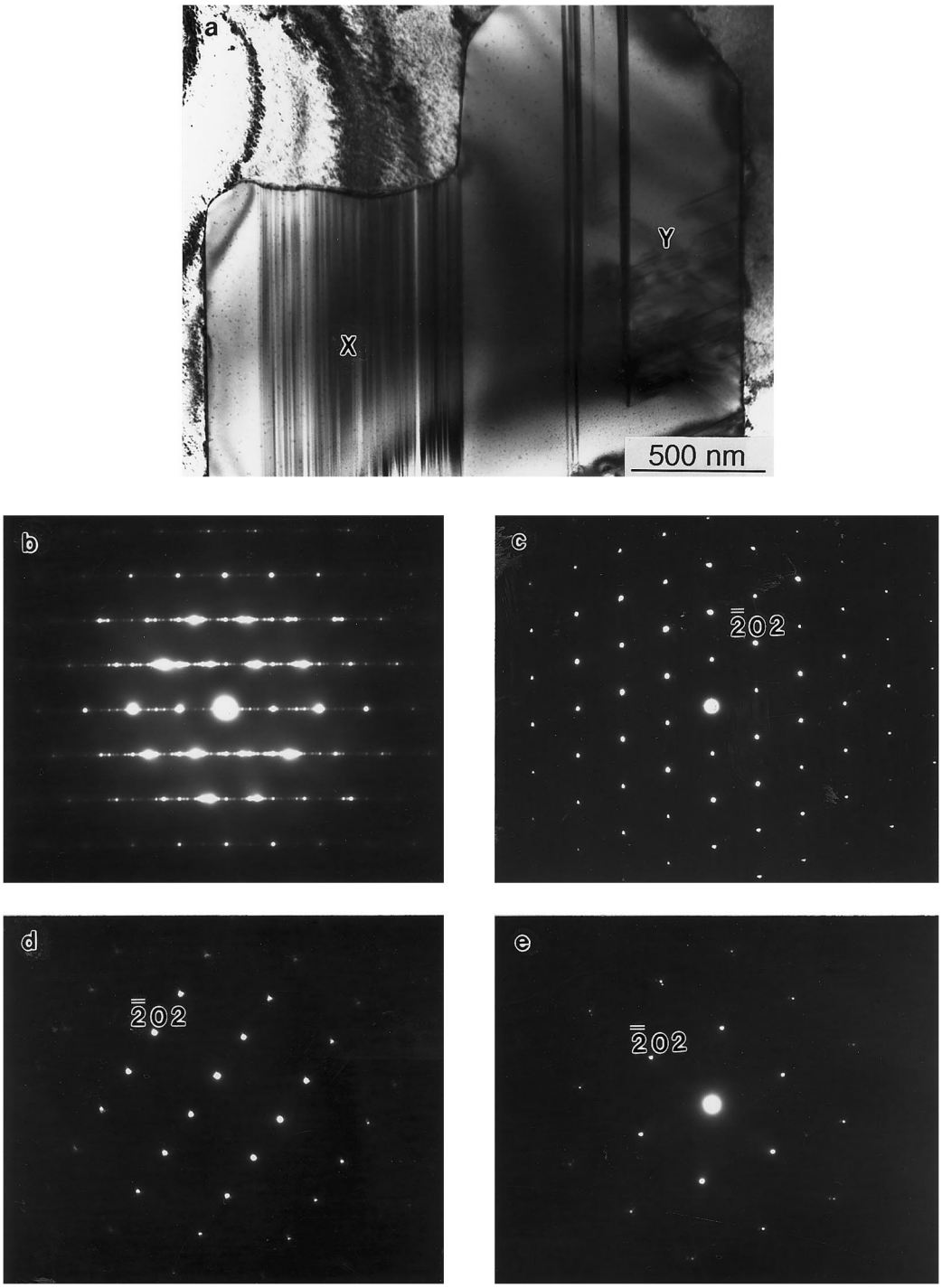


Fig. 6. (a) Bright field image of a discrete Laves phase particle in the Cr–9.8 at.% Ta alloy that was annealed at 1573 K for 168 h showing a heavily faulted region (X) and a region Y containing a low density of faults. Selected area diffraction pattern from region X confirms a hexagonal structure (C36, or a higher order polytype) using the $[11\bar{2}0]$ zone axis, whereas a cubic structure is recognized in region Y using (c) [101], (d) [121] and (e) [111] zone axes.

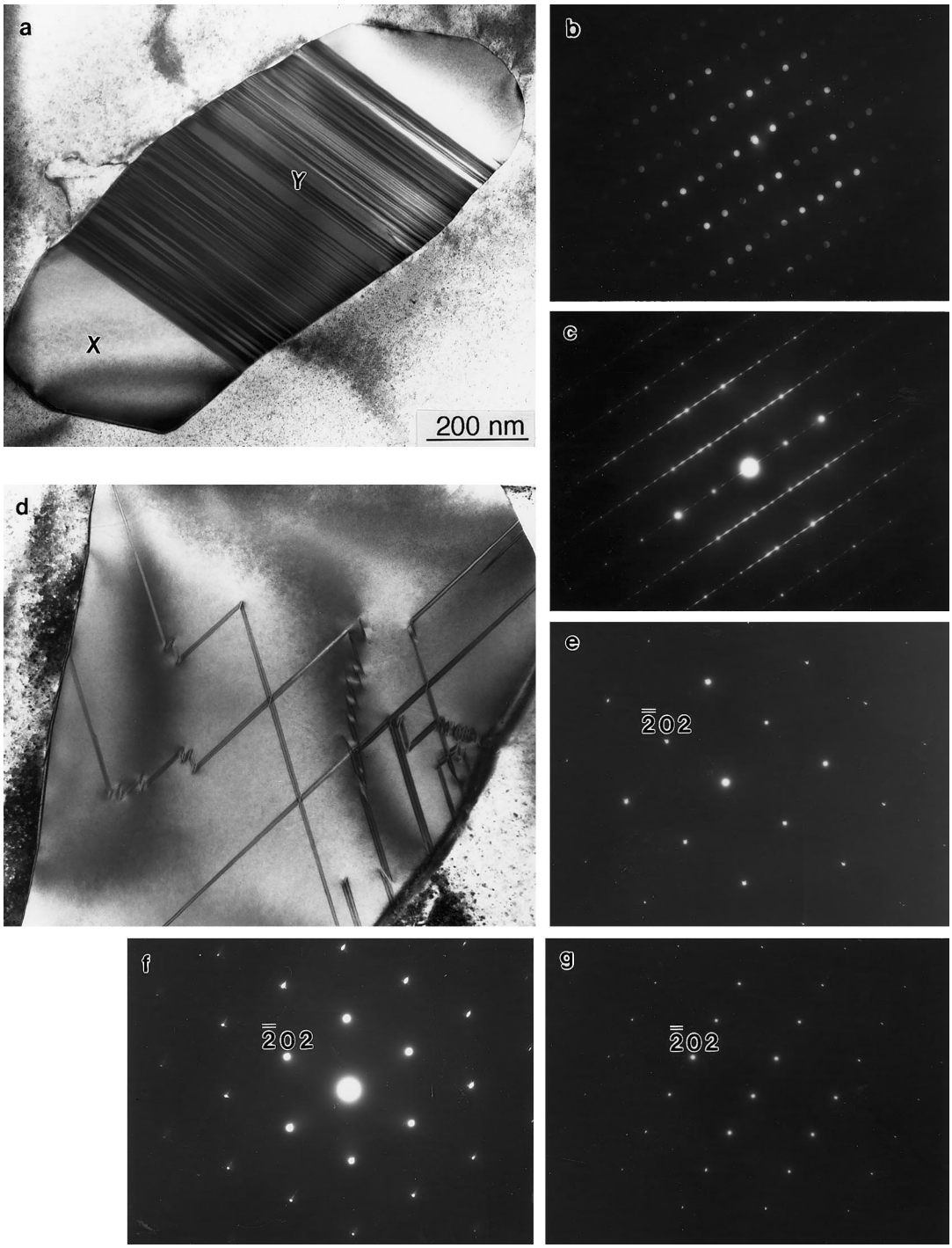


Fig. 7. (a) Bright field of a discrete Laves phase particle in the Cr–9.8 at.% Ta alloy that was annealed at 1573 K for 168 h showing a heavily faulted region (Y) and a region X containing no faults. Selected area diffraction patterns at the $[11\bar{2}0]$ zone axis confirms (b) region X to be C14 and (c) region Y to be C36. (d) A bright field image of another precipitate shows multiple variants of stacking faults on the {111} planes of the cubic C15 structure. Associated diffraction patterns (e)–(g) at the [010], [111] and [121] zone axes, respectively, confirm the cubic structure.

denotes the room temperature composition range for the C15 phase. An alloy composition lying in this range, under equilibrium conditions, will form the C14 structure that upon cooling will transform to the C15 structure through a two-phase region of C14 + C15 where substantial compositional adjustments are required. Such a solid–solid transformation tends to be sluggish and is often suppressed under normal solidification conditions; in addition, in the Cr–Nb and Cr–Ta systems and possibly in the Cr–Hf system, the range of temperature over which the two-phase field is stable for a given composition is small. As cooling continues, the C14 structure is retained metastably but the chemical free energy for forming the C15 structure increases with increased undercooling. Eventually, the C14 structure can shear into the C15 structure often via a metastable intermediate C36 structure [1, 5]. For these alloys, there are no compositional adjustments required for the transformation to proceed and thus it is anticipated that the transformation kinetics will be temperature-dependent to the extent that dislocation mobility in the metastable C14 (and/or C36) phase is affected by temperature [19]. Such a transformation can proceed either during cooling down or subsequently during isothermal annealing. Peculiar to Laves phases is also the fact that synchroshockley dislocation mobility needs to be considered and in this regard, Hazzledine [20] has invoked diffusional requirements at the core of these dislocations for their motion. He has proposed a Zener interchange mechanism at the core in order for the dislocations to move and this is facilitated if vacancies were trapped in the core and a vacancy moved with a kink in the dislocation and enabled the interchange.

For compositions within this range "pqsr", the transformation kinetics is anticipated to vary with composition depending on (1) the slope of the C14 + C15 two-phase region (to a first approximation the slope of the line BC assuming the region to be fairly narrow and bound by straight lines on the top and bottom), (2) the nature of defect structure (such as vacancies) present, and (3) the variation in the intrinsic diffusion coefficients with composition. The second and third factor mentioned above may not be as important for a shear transformation as it would be for a diffusional transformation. For example, in the case of the Cr–Ti system, where the slope of the line BC is steep (B = 1073 K, C = 1493 K over 2 at.% Cr), transformation kinetics has been shown to be a function of stoichiometry [1]. In contrast, in the Cr–Ta and Cr–Nb systems, the slope of the line BC is relatively shallow (B = 1933 K, C = 1968 K over 3 at.% Cr in the Cr–Ta system and B = 1858 K and C = 1898 K over ~7 at.% Cr in the Cr–Nb system [10]) and therefore a substantial difference in transformation kinetics with stoichiometry is not expected. It is worth noting in passing that the eutectoid and peritectoid temperatures are significantly higher in the Cr–Ta and Cr–Nb systems compared with the Cr–Ti and Cr–Hf systems. As mentioned earlier, intrinsic diffusion coefficients can also vary as a function of stoichiometry. Diffusion data on Laves phases are scarce but recently, self-diffusion data for Co and Nb in Co_2Nb in the approximate temperature range 1140–1550 K [21] have been reported. The diffusivity of Nb in $Nb_{31}Co_{69}$ was shown to be considerably slower than that of Co as would be intuitively expected; it was also shown that the diffusivity of Co increased with increasing Co content in the phase. Thus, Co diffusivity in $Nb_{29}Co_{71}$ was found to be significantly higher than in $Nb_{31}Co_{69}$. Thus if the transformation mechanism proposed by Hazzledine [20] is correct, then transformation kinetics at a given temperature may be influenced by stoichiometry as a consequence of diffusion.

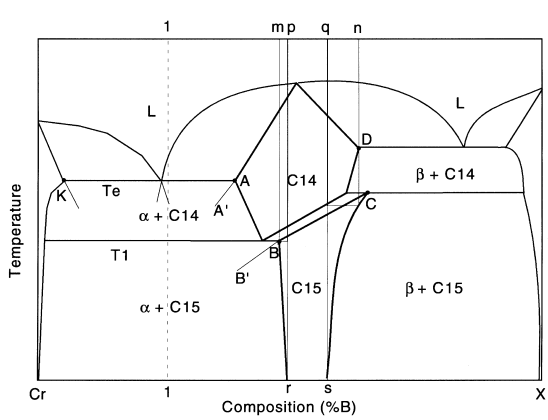


Fig. 8. A schematic Cr–X (X = Nb, Ta, Hf and Zr) phase diagram showing metastable extensions AA′ and BB′. Lines mB, pr, qs and nD denote compositions. The dashed line "1" corresponds to an alloy composition similar to the Cr–9.8 at.% Ta used in this investigation.

Returning to Fig. 8, we next consider transformation from C14 to C15 in those alloys that lie between "mB" and "pr" and between "qs" and "nD". In Fig. 8 we assume that point D is less X-rich than point C (if point D was more X-rich than point C, an additional region between C and D would be present and the situation in that region would be similar to the region between A and B which will be discussed later). Compositions lying in the region bound by "qs" and "nD" and closer to "qs" can metastably retain the C14 structure during non-equilibrium cooling, which can transform by shear into the C15 structure at lower temperatures during cool down without compositional changes provided there is adequate driving force from undercooling. The C15 phase in these alloys will, however, have a non-equilibrium composition at room temperature. For those alloys lying closer to the line "nD" but within the region "qsnD", the C15 phase formation will require minor adjustments

in composition and such adjustments could occur by diffusion at elevated temperatures during cool down. If the appropriate composition that dictates C15 stability at a given temperature is reached and there is simultaneously adequate driving force for the transformation, then shear transformation proceeds and the C15 structure can form. If this does not happen, subsequent isothermal annealing at elevated temperature would produce the required composition change by diffusion and then the structure can shear into the C15 form. Whether these two processes occur simultaneously or sequentially remains to be determined. Similar arguments can be extended for alloys lying in the composition range "mB–pr". It is worth noting that the change in solubility of Cr (or X) with temperature in the C15 Laves phase (i.e. curves Br and sC) influences the transformation response of alloys in these regimes.

Finally, we consider alloy compositions that lie between the compositions corresponding to points A and B in Fig. 8. The analysis presented for this case will also apply to two-phase alloys in the range KA and as typified by alloy 1 marked by dashed lines in Fig. 8. Alloys with compositions lying between points A and B are capable of existing under equilibrium conditions as a single-phase material with the C14 structure at elevated temperatures; a C15 counterpart is however not stable in this composition regime at lower temperatures. Under equilibrium conditions, the phase diagram dictates that such alloys undergo a eutectoid reaction and produce the C15 structure as a eutectoid microconstituent. This requires extensive diffusion and such solid-state reactions involving two substitutional elements are usually sluggish. As a consequence, under normal solidification conditions, it is expected that the C14 structure will be metastably retained at low temperatures. If the composition of point A is fairly close to point B and the eutectoid temperature T1 in Fig. 8 is relatively high, then it is conceivable that minor compositional adjustments can occur during cool down so that a transformed C15 structure is produced in the as-cast alloy, or fairly rapidly upon subsequent isothermal annealing at an elevated temperature. If however, the composition of point A is significantly higher in Cr than that of point B then, extensive diffusion is required for the metastable Laves phase to achieve a composition that is within the structural stability envelope of the C15 structure and consequently, long times at elevated temperatures are required for the transformation to proceed to completion. As mentioned before, these arguments hold equally well for two-phase alloys in the regime defined by KA and of the composition Cr–9.8 at.% Ta studied in this investigation. In considering these two-phase alloys, the metastable C36 structure is frequently encountered as an intermediate product during the transformation of the metastable C14 structure to the stable C15 structure. Metastable solvii for these phase diagrams are not available but it is possible to think in terms of extending the line CB in Fig. 8 to lower temperatures as denoted by the line BB′. In this case, if a two-phase alloy such as alloy 1 that retains the C14 structure to room temperature is isothermally annealed at a temperature below the eutectoid temperature T1 but sufficiently high to permit diffusion to occur in reasonable times, the metastable C14 structure will adjust its composition till it intersects the line BB′ at which point it is capable of transforming to the C15 structure. If a metastable solvus dome is incorporated for the C36 structure that intersects the line BB′ then the intermediate C36 structure will form as the compositional adjustment proceeds before the C15 structure forms. When the required compositional adjustment has occurred, the transformation will proceed by shear. This continuous shift in composition from a Cr-rich C14 structure to a less Cr-rich C15 structure can also be viewed as a progressive change in the e/a ratio which dictates the stability of the various Laves phase polytypes. Finally, it should be noted that while point A provides the equilibrium composition of the eutectic C14 Laves phase microconstituent in two-phase alloys lying in the composition region KA at the eutectic temperature, non-equilibrium freezing processes such as drop casting into a cold copper mold can produce adequate undercooling such that the metastable extension AA′ must be considered to determine the eutectic C14 composition. This implies an even larger deviation in composition that must be adjusted by diffusion during subsequent annealing to permit the C14-to-C15 transformation.

We can now review the results from previous studies as well as from this investigation in light of the above discussion. In the Cr–Nb system, the as-cast *single-phase* alloy (Cr_2Nb phase) exhibits the twinned C15 structure [2]. The C14 structure is still believed to be the stable high-temperature phase but the transformation from the C14 to C15 is thought to be rapid [2] and to proceed during cooling to room temperature. Twinning is thought to be a consequence of the C14-to-C15 transformation. Similar observations have also been made by Takasugi *et al.* [4] in an almost single-phase alloy with the important difference that Bewlay *et al.* [2] made their material by directional solidification at a rate of ~5 mm/min whereas Takasugi *et al.* [4] solidified their "button-melts" on a water-cooled copper hearth. Clearly, the cooling rates in these two cases are different but the transformation nevertheless appears to proceed to completion. In contrast, Thoma and Perepezko [10] reported that in their arc-melted ingot of monolithic Cr_2Nb, they observed both, the C14 and C15 phases and implied that the amount of C14 phase that was retained was a function of specimen location in the ingot, being greater in regions that experienced faster

cooling; subsequent annealing, however, produced a single-phase C15 structure.

In the *two-phase* Cr–Nb binary alloys, the as-cast microstructure still consists of the twinned C15 phase. Thoma and Perepezko [10] have reported a composition of 69.17 ± 0.8 at.% Cr for the C15 Laves phase in the as arc-melted condition in a Cr–Cr_2Nb two-phase alloy. This is in reasonable agreement with the composition of 70.0 at.% Cr reported by Takeyama and Liu [22] and by Takasugi *et al.* [23]. Following a 100 h anneal at 1673 K, which is a temperature well below the eutectoid temperature, the composition of the C15 phase changes to 68.1 at.% Cr [10]. This implies that for the Cr–Nb phase diagram, the compositions corresponding to points A and B in Fig. 8 are very close to each other. Further, the C14/C15 transformation temperature (the eutectoid temperature) for two-phase Cr–Cr_2Nb alloys in the Cr–Nb system is ~1858 K; it would thus be possible without significant compositional adjustments and the aid of rapid diffusion for the metastable C14 structure to shear into the C15 structure at fairly high temperatures (for example, 50–100 K below the eutectoid temperature). The compositional range of existence of the C15 phase below 1200 K, however, is quite limited compared with the range at ~1750 K. It follows that although enough diffusional adjustments occurred at the elevated temperatures for the C15 structure to be present in the as-cast condition, the composition of the C15 phase still does not correspond to the equilibrium composition at the lower temperatures, and upon subsequent annealing, further compositional adjustments occur. Similar arguments can also be extended to the Nb–Cr_2Nb alloys. In effect, due to the size and shape of the Laves phase fields (composition and temperature regimes) in the Cr–Nb system, the transformation kinetics of two-phase alloys resemble those of the single-phase alloys.

In the Cr–Ti system, *the single-phase* Cr_2Ti composition has been shown to maintain the C14 (or C14 + C36) structure at room temperature and to transform to the C15 form via the C36 structure following extended exposures to 1473 K [5, 6]. The C14/C15 transformation temperature regime for the Cr–Ti system is considerably lower than the corresponding range for the Cr–Nb system. The same argument may also explain the prolonged annealing times required to transform the metastable hexagonal phase to the cubic C15 form, since the annealing temperature must necessarily be below the C14/C15 transformation temperature. The transformation kinetics in the two-phase alloys in the Cr–Cr_2Ti regime as well as the Ti–Cr_2Ti regime are extremely sluggish for similar reasons, with the Ti-rich alloys being somewhat better than the Cr-rich alloys [6].

Transformations in single-phase alloys have not been characterized in the Cr–Ta, Cr–Zr and Cr–Hf systems. The transformation in a two-phase Cr–Cr_2Hf alloy was found to be sluggish with the C14 structure being metastably retained in the as-cast condition and transforming to the C36 structure after extended isothermal annealing; the C15 structure was not observed [9]. The Cr–Hf phase diagram does not appear to be well-developed and therefore it is not possible to discuss the reasons for the sluggish nature of the reaction in the context of compositions and temperatures.

In this study, the composition of the Laves phase eutectic microconstituent in the Cr–9.8 at.% Ta alloy in the cast and annealed (1273 K/24 h) condition corresponds to 28.5 at.% Ta; this is different from the Cr–Nb phase diagram, where the composition of the Laves phase (in the as-cast condition) in two-phase Cr–Cr_2X alloys at the eutectic temperature corresponds to 30.8 at.% Nb. It is interesting to note that after long-term annealing of the Cr–Ta alloy, the average Laves phase composition corresponds to 30.4 at.% Ta which is close to the composition of the Cr_2Nb phase in the Cr–Cr_2Nb two-phase alloys in the as-cast condition. In the Cr–Ta system, in contrast to the Cr–Ti system, the eutectoid temperature is high (1933 K) and this allows isothermal annealing to be performed at high temperatures (e.g. 1573 K/168 h) and consequently, diffusion is fairly rapid and the C14-to-C15 transformation proceeds adequately to evidence the presence of the C15 phase even though the C14 phase has a Cr-rich composition in the reference condition. Composition measurements before and after the prolonged high-temperature anneal (1573 K/168 h) confirm that the C14-to-C15 transformation in the two-phase alloys is accompanied by compositional changes and hence diffusion plays a primary role in affecting transformation kinetics.

5. CONCLUSIONS

1. The C14 Cr_2Ta phase is metastably retained as a eutectic microconstituent in a cast and annealed (1273 K/24 h) two-phase Cr–Cr_2Ta alloy. The density of stacking faults in the C14 Cr_2Ta is low, confirming its stability.
2. The observed microstructure in the Cr–Ta alloy confirms that the published Cr–Ta phase diagram is inaccurate in terms of the eutectic composition and the composition of the Laves phase at the eutectic temperature.
3. Following a long-term, high-temperature anneal (1573 K/168 h), the lamellar morphology breaks down and discrete particles of Cr_2Ta are present in a Cr solid solution matrix. These particles were confirmed to be present and exhibited the C14, C36 and C15 structures. Often multiple structures were present in a single particle confirming that they are in the process of transform-

ing from the metastable C14 to the stable C15 structures.

4. The average composition of the Laves phase particles following prolonged high-temperature annealing is richer in Ta compared with the composition of the lamellar eutectic Laves phase microconstituent. This confirms that annealing below the eutectoid temperature leads to the transformation of the C14 structure to more stable configurations and that these transformations are accompanied by compositional changes. Therefore, at least in these two-phase alloys the transformation of the metastable C14 structure to the C36 and C15 structures can be thought of as occurring due to a combination of diffusion and shear. This may in part account for the sluggishness of these transformations.
5. Compositional differences were recognized within individual particles and could be associated with the C14, C36 and C15 structures. Such compositional differences can be thought of in terms of different e/a regimes in which these different structures are stabilized.

Acknowledgements—This research was supported by the Division of Materials Science, United States Department of Energy under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation; L.P. and K.S.K. used the Materials Research Science and Engineering Center (MRSEC) Shared Experimental Facilities at Brown University for a part of this investigation. The MRSEC is supported by the National Science Foundation under Award No. DMR-9632524. Microscopy and microanalysis performed at ORNL was done at the Shared Research Equipment (SHaRE) User Facility.

REFERENCES

1. Allen, C. W., in *High Temperature Ordered Intermetallic Alloys*, ed. C. C. Koch, C. T. Liu and N. S. Stoloff. Materials Research Society, Pittsburgh, PA, 1985, p. 141.
2. Bewlay, B. P., Sutliff, J. A., Jackson, M. R. and Lipsitt, H. A., *Acta metall. mater.*, 1994, **42**, 2869.
3. Kumar, K. S. and Liu, C. T., *Acta metall. mater.*, 1997, **45**, 3671.
4. Takasugi, T., Yoshida, M. and Hanada, S., *Mater. Sci. Engng*, 1995, **A192–193**, 805.
5. Allen, C. W., Delavignette, P. and Amelinckx, S., *Physica status solidi (a)*, 1972, **9**, 237.
6. Chen, K. C., Allen, S. M. and Livingston, J. D., in *High Temperature Ordered Intermetallic Alloys VI*, Vol. 364, ed. J. Horton, *et al.* Materials Research Society, Pittsburgh, PA, 1995, p. 1401.
7. Liu, C. T., Tortorelli, P. F., Horton, J. A. and Carmichael, C. A., *Mater. Sci. Engng*, 1996, **A214**, 23.
8. Venkatraman, M. and Neumann, J. P., in *Binary Alloy Phase Diagrams*, Vol. 1, ed.-in-chief T. B. Massalski, ed. J. L. Murray, L. H. Bennett and H. Baker. American Society for Metals, Metals Park, OH, 1986, p. 867.
9. Kumar, K. S. and Miracle, D. B., *Intermetallics*, 1994, **2**, 257.
10. Thoma, D. J. and Perepezko, J. H., *Mater. Sci. Engng*, 1992, **A156**, 97.
11. Takasugi, T. and Yoshida, M., *J. Mater. Res.*, 1998, **13**, 2505.
12. Brady, M. P., Zhu, J. H., Liu, C. T., Totorelli, P. F. and Walker, L. R., *Intermetallics*, in press.
13. Hunt, J. D. and Chilton, J. P., *J. Inst. Metals*, 1962–63, **91**, 338.
14. Jackson, K. A. and Hunt, J. D., *Trans. TMS–AIME*, 1966, **236**, 1129.
15. Smith, A. W., Rogers, J. A. and Rawlings, R. D., *Physica status solidi (a)*, 1973, **15**, K119.
16. Bruckner, W., Kleinstuck, K. and Schulze, G. E. R., *Physica status solidi*, 1967, **23**, 475.
17. Komura, Y. and Kitano, Y., *Acta crystallogr.*, 1977, **B33**, 2496.
18. Zhu, J. H., Liaw, P. K. and Liu, C. T., *Mater. Sci. Engng*, 1997, **A239–240**, 260.
19. Kubsch, H., Paufler, P. and Schulze, G. E. R., *Physica status solidi (b)*, 1973, **56**, 231.
20. Hazzledine, P. M., in *Twinning in Advanced Materials*, ed. M. H. Yoo and M. Wuttig. The Minerals, Metals and Materials Society, Warrendale, PA, 1994, p. 403.
21. Denkinger, M., Mehrer, H., in E. P. George, M. J. Mills and M. Yamaguchi. *High Temperature Ordered Intermetallic Alloys VIII*, Vol. 552, ed. E. P. George, M. J. Mills and M. Yamaguchi, 1999. MRS, Warrendale, PA, P. KK 7.6.1.
22. Takeyama, M. and Liu, C. T., *Mater. Sci. Engng*, 1991, **A132**, 61.
23. Takasugi, T., Kumar, K. S., Liu, C. T. and Lee, E. H., *Mater. Sci. Engng*, 1999, **A260**, 108.

ELSEVIER

Intermetallics 8 (2000) 1119–1129

Intermetallics

www.elsevier.com/locate/intermet

Physical metallurgy and mechanical properties of transition-metal Laves phase alloys

C.T. Liu *, J.H. Zhu [1], M.P. Brady, C.G. McKamey, L.M. Pike [2]

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6115, USA

Received 1 March 2000; accepted 15 August 2000

Abstract

This paper provides a comprehensive review of the recent research on the phase stability, point defects, and fracture toughness of AB_2 Laves phases, and on the alloy design of dual-phase alloys based on a soft Cr solid solution reinforced with hard XCr_2 second phases (where X = Nb, Ta and Zr). Anti-site defects were detected on both sides of the stoichiometric composition of $NbCr_2$, $NbCo_2$, and $NbFe_2$, while they were observed only on the Co-rich side of $ZrCo_2$. Only thermal vacancies were detected in the Laves phase alloys quenched from high temperatures. The room-temperature fracture toughness cannot be effectively improved by increasing thermal vacancy or reducing stacking fault energy through control of phase stability. Microstructures, mechanical properties, and oxidation resistance of dual-phase alloys based on Cr–$NbCr_2$, Cr–$TaCr_2$, and Cr–$ZrCr_2$ were studied as functions of heat treatment and test temperature at temperatures to 1200°C. Among the three alloy systems, Cr–$TaCr_2$ alloys possess the best combination of mechanical and metallurgical properties for structural use at elevated temperatures. © 2000 Published by Elsevier Science Ltd.

Keywords: A. Laves phases; B. Mechanical properties at high temperatures; B. Mechanical properties at ambient temperature; D. Defects: point defects

1. Introduction

Laves phases with AB_2 compositions are a common type of topologically close-packed (TCP) structures [1–4]. There are over 360 binary Laves phases reported in the literature. Furthermore, Laves phases are capable of dissolving considerable amounts of ternary alloying additions. A total of over 900 combined binary and ternary Laves phases have been documented. Although Laves structures are the most abundant among intermetallic compounds, they are the least investigated. So far, most previous studies on intermetallic alloys have concentrated on simple ordered structures derived from face-centered cubic (f.c.c.), body-centered cubic (b.c.c.), or hexagonal-close packed (h.c.p.) structures. Thus, there is a great need for the study of Laves phase alloys.

There are three common polytypes of Laves phases most frequently observed: cubic C15, hexagonal C14 and dihexagonal C36 [1–4]. These polytypes are related to each other as the basic unit layer of these phases is the same, while the stacking sequence of the unit layer is different in each structure. One unit layer in each Laves phase is composed of four sub-layers. In the C15 structure, the packing sequence of the unit layer is ABCABC, similar to the sequence in conventional cubic f.c.c. metals; in the hexagonal C14, the packing layer is ABABABAB, similar to the sequence in conventional h.c.p. metals. The dihexagonal C36 structure has a packing sequence of ABACABAC.

Laves phases are generally stabilized by the size-factor principles, that is, the atomic size ratio, R_A/R_B, is ideally 1.225, with a range of 1.05–1.68 usually observed [1–4]. The stability of each crystalline structure is also influenced by the electron concentration factor (e/a) [5,6]. In fact, the electron concentration factor becomes clearly important when the atomic size factors are favorable. The classic work by Laves and Witte [5,6] showed that for several quasi-binary alloy systems involving $MgCu_2$ and $MgZn_2$, with increasing valence electron concentration, the three Laves types, $MgCu_2$, $MgNi_2$, and $MgZn_2$, exist in that order.

* Corresponding author.

[1] Present address: Tennessee Technology University, TTU Box 5014, Cookeville, TN 38505-0001, USA.

[2] Present address: Haynes International, Kokomo, IN 46904, USA.

0966-9795/00/$ - see front matter © 2000 Published by Elsevier Science Ltd.
PII: S0966-9795(00)00109-6

Laves phases are an important and attractive type of material both scientifically and technologically. They are being seriously considered for many practical applications. For example, $(Tb,Dy)Fe_2$ has been used as a magnetoelastic transducer because of its great magnetorestriction [7]. Compounds of $(Hf,Zr)V_2$ have demonstrated superconducting properties with a combination of high critical temperature, high current density and good magnetic strength [8,9]. Laves phases, such as $Zr(Cr,Fe)_2$, have been considered for hydrogen storage applications due to favorable hydriding-dehydriding kinetics and high hydrogen-absorbing capabilities [10]. The $Mo(Co,Si)_2$ Laves phase contributes to the wear-resistance of "Tribaloy" materials [11]. More recently, $HfCr_2$-, $NbCr_2$- and $TiCr_2$-based two-phase alloys are being developed for high-temperature structural uses because of their high melting points and good retention of mechanical properties at elevated temperatures [12–20].

Despite their useful properties, the low ductility and brittle fracture characteristics at ambient temperatures are the main limitations for structural applications of this large class of alloys. The high hardness and brittle fracture of Laves phases are due to the complex atomic configuration of each unit layer, interplanar locking and lack of operating slip systems at ambient temperatures. Recently, a synchroshear mechanism [21] has been proposed to facilitate plastic deformation in Laves phases, which raises the hope of toughening the Laves phases, even though so far no experimental support for this mechanism has been furnished. All these facts indicate that more research needs to be undertaken on Laves phases in order to further understand the basic phase stability, defect structure, deformation mechanism and fracture behavior. Moreover, alloy design efforts should be devoted to overcoming the brittleness in these phases.

This paper summarizes our recent studies of single-phase and dual-phase alloys containing transition-metal elements. The phase stability and defect structures in Laves phases $NbCr_2$, $NbFe_2$, $NbCo_2$ and $ZrCo_2$ have been studied for the purposes of (1) promotion of synchroshear deformation by increasing vacancy concentration and (2) promotion of mechanical twinning by controlling phase stability and reducing stacking fault energy. At present only limited progress has been achieved in improving the fracture toughness of these Laves phase alloys. In view of this problem, dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard $NbCr_2$, $ZrCr_2$, or $TaCr_2$ second phases have been selected for alloy development [17,22–26]. Among the three Cr–XCr_2 (X = Nb, Zr, and Ta) systems, Cr–$TaCr_2$ alloys are more resistant to air oxidation and thermally-induced cracking. As a result, our current alloy development effort has been focusing on the Cr–$TaCr_2$ alloy system, and significant advances have been made in enhancing the high-temperature strength, creep resistance and fracture toughness of the dual-phase alloys through compositional adjustment, microstructural control and alloy additions.

2. Point defect and fracture toughness of binary Laves phases

Laves phases are generally considered to be line compounds with a strict AB_2 composition. However, solubility ranges exist for about 25% of the known binary Laves phases [27]. For non-stoichiometric compositions, constitutional defects are incorporated into intermetallic compounds. However, for the Laves phases, little is known about the possible defect structures that may be associated with the deviations from stoichiometry. The excess atoms in off-stoichiometric compounds can stay on their own sublattice (leading to the formation of constitutional vacancies on the other sublattice), insert into interstitial sites, or occupy sites on the other sublattice (anti-site substitution). Since Laves phases have TCP structures and space-filling is relatively high, there are no interstitial sites with a size comparable to that of the component atoms. Therefore, the insertion of the excess atoms into interstitial sites can be excluded, and the possible defect mechanisms in binary Laves phases reduce to either constitutional vacancy or anti-site substitution.

The Laves phases are known to be size compounds, i.e. the atomic size ratio, R_A/R_B is ideally 1.225, with a range of 1.05–1.67 typically observed. Since the A atom is much larger than the B atom, the excess A atoms in A-rich compositions would tend to stay on their own sublattice sites, thus creating vacancies on the B atom sublattice. On the other hand, the excess B atoms on B-rich compositions would be able to occupy the A atom sublattice sites, leading to the formation of anti-site defects. Based on these geometric arguments, it has been postulated that the A-rich side is accommodated by vacancies, while the B-rich compositions result from the anti-site substitution [27]. However, no systematic experimental verification has been undertaken so far.

The types of constitutional defects may affect physical, mechanical and functional properties of Laves phases. Vacancies have been proposed to assist the movement of synchro-Shockley dislocations [21], thus possibly facilitating the synchroshear deformation mechanism and increasing the toughness of Laves phases [28]. Off-stoichiometry was also found to affect the hydride stability of $ZrMn_2$ [29] and the hydrogen storage capability of $TiMn_2$ alloys [30]. The defect structures of the cubic C15 $NbCr_2$, $NbCo_2$ and $ZrCo_2$ and hexagonal C14 $NbFe_2$ Laves phases were studied to clarify the point defect mechanisms on both sides of stoichiometry in these binary systems. This was done by measuring the lattice parameters, bulk densities, and therefore, vacancy concentrations of these alloys of various compositions

[31,32]. The hardness of the above alloys with various point defect concentrations was determined in an attempt to explore the solid solution hardening behavior in the Laves phases. The fracture toughness was also measured to investigate whether the point defects affect the toughness of the alloys, and, in particular, to see if vacancies facilitate synchroshear deformation in the vacancy-containing alloys.

Figs. 1 and 2 show the vacancy concentration as a function of the Nb content for $NbCr_2$ and $NbFe_2$ alloys, respectively, after quenching from different temperatures [31]. It is clear that after quenching the $NbCr_2$ alloys from 1000°C, the vacancy concentration is essentially zero, while the vacancy concentration significantly increases after quenching from higher temperatures, i.e. 1400°C. For the $NbFe_2$ alloy, the vancancy concentration is very low even after quenching from 1300°C. Fig. 3 summarizes all the measured vacancy concentrations for all single-phase alloys after quenching from 1000°C, which are essentially zero. The calculated vacancy concentrations from both the constitutional vacancy model and anti-site substitution model are also plotted in Fig. 3 [31]. The experimental results indicate that there are no constitutional vacancies in these Laves phases on either side of stoichiometry. This is consistent with the anti-site substitution model and contrary to the constitutional vacancy model, which demands much higher vacancy concentrations for off-stoichiometric compositions. Note that the constitutional vacancy model is based on the assumption that all the excess atoms exclusively stay on their own sublattice, thus constitutional vacancies are created on the sublattice of the other element. The anti-site substitutional model assumes that the excess atoms occupy the sublattice sites of the other species. As a result, anti-site defects are created, and no constitutional vacancies are needed for the balance of the lattice sites. Thus, the constitutional defects on the Cr-, Co-, or Fe-rich, as well as the Nb-rich sides of stoichiometry, are of anti-site type for the $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves phases.

Modder and Bakker pointed out that it is very possible that a vacancy type defect, termed quadruple defect, occurs in C15 compounds [33]. Such a quadruple defect consists of one anti-site B atom on the α sublattice and three vacancies on the β sublattice. Ball milling was found to introduce quadruple defects in $GdAl_2$, $GdPt_2$, $GdIr_2$ and $GdRh_2$, while anti-site defects were created in $GdMg_2$ after ball-milling [34]. Since ballistic actions such as ball milling usually lead to the formation of the same type of atomic defects as heating, thermal defects such as quadruple defects or anti-site defects may be created upon quenching the Laves compounds from elevated temperatures. Thermal vacancies are expected

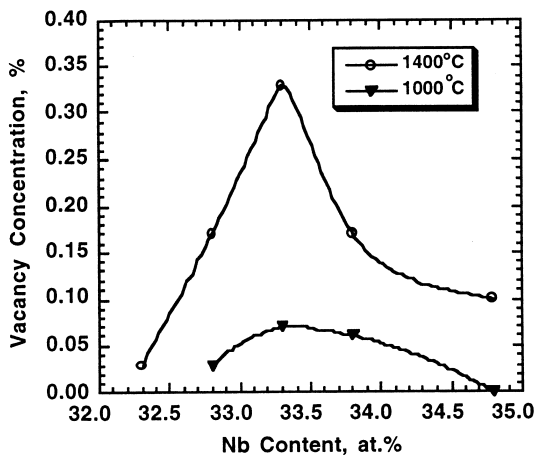


Fig. 1. Vacancy concentration vs Nb content for $NbCr_2$ Laves phase quenched from 1000 and 1400°C [31].

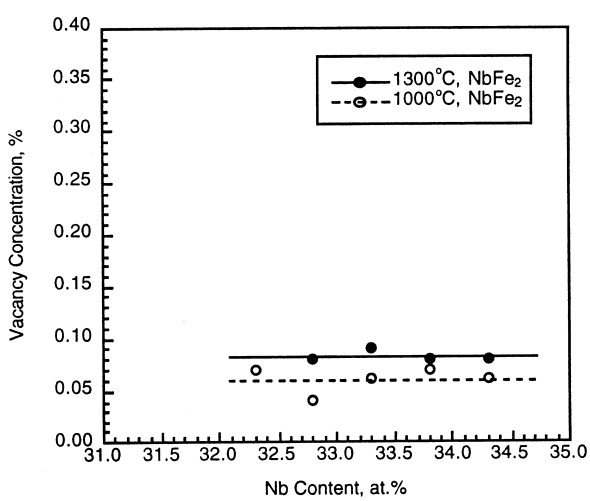


Fig. 2. Vacancy concentration vs Nb content for $NbFe_2$, Laves phase quenched from 1300 and 1000°C [31].

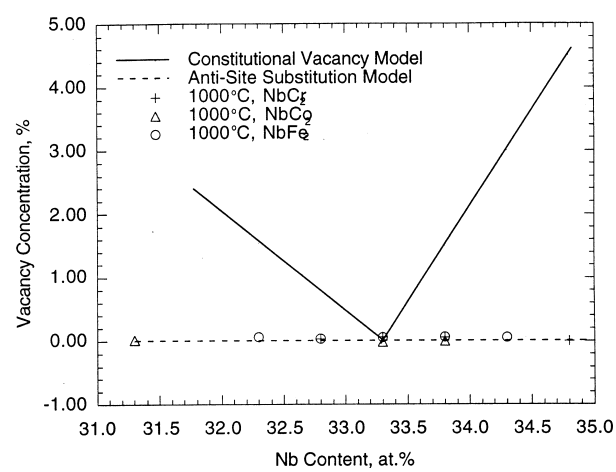


Fig. 3. Measured vacancy concentration vs Nb content in the binary $NbCr_2$, $NbCo_2$, and $NbFe_2$, Laves phases after quenching from 1000°C. Also shown are the calculated data for the constitutional vacancy model and the anti-site substitution model [31].

if quadruple defects are created, as each quadruple defect contains three vacancies.

It is interesting to note that appreciable vacancies were detected in the $NbCr_2$ alloy after quenching from 1400°C (Fig. 1). Since these vacancies were essentially annealed out at 1000°C (Fig. 1), and since the constitutional defects in this compound have been established to be of the anti-site type, it is concluded that these defects are *thermal* vacancies, indicating that quadruple defects may be formed in this compound by heating. Furthermore, the measured vacancy concentration exhibits a maximum at the stoichiometric composition and decreases on both sides of stoichiometry for the $NbCr_2$ alloys when quenched from 1400°C (Fig. 1). A maximum in thermal vacancy concentration, and thus thermal disorder, at the stoichiometric composition can be rationalized by considering the fact that the increase in entropy associated with the introduction of vacancies is the greatest at the stoichiometric composition where the degree of order is necessarily the highest.

For the C14 $NbFe_2$ compounds, no thermal vacancies were detected after quenching from 1300°C (Fig. 2), indicating that anti-site defects, instead of quadruple defects, may be the form of atomic disorder in this compound upon heating. According to the model developed by Modder et al. [34], the type of atomic disorder is related to the relative magnitudes of formation enthalpies of anti-site and quadruple-defect disorder for the Laves phases. Therefore, the differences in thermal defects in $NbCr_2$ and $NbFe_2$ imply that the formation enthalpies of anti-site disorder are larger in magnitude than that of the quadruple defect disorder for $NbCr_2$, while the opposite is true for $NbFe_2$.

A point defect in a crystalline lattice is well known to cause hardening. This phenomenon is referred to as solid solution hardening. Solid solution hardening has been extensively studied in B2 compounds [35]. Generally, it is found that in B2 compounds with anti-site defects, hardness exhibits a minimum at the stoichiometric composition, and deviation from stoichiometry causes hardening of the compounds, due to the presence of constitutional anti-site defects. For triple-defect B2 compounds, the shape of the hardness versus composition curves was found to be not so simple and depended on the degree of thermal disorder. Also, vacancies were found to be a more potent hardener than the anti-site defects.

In the $NbCr_2$ and $NbFe_2$ Laves phases studied, the hardness was also found to have a characteristic V-shape with a minimum occurring at the stoichiometric compositions (Figs. 4 and 5) [31]. Obviously, similar to anti-site B2 compounds, such off-stoichiometric hardening can be attributed to the presence of constitutional anti-site defects, which is the defect mechanism on both sides of stoichiometry in these Laves phases. The anti-site hardening mechanism in Laves phases, however, is not clear so far, and it may be different from that in B2 compounds. In B2 compounds, the anti-site hardening is often correlated to an interaction between the stress field of a moving dislocation and that of the anti-site

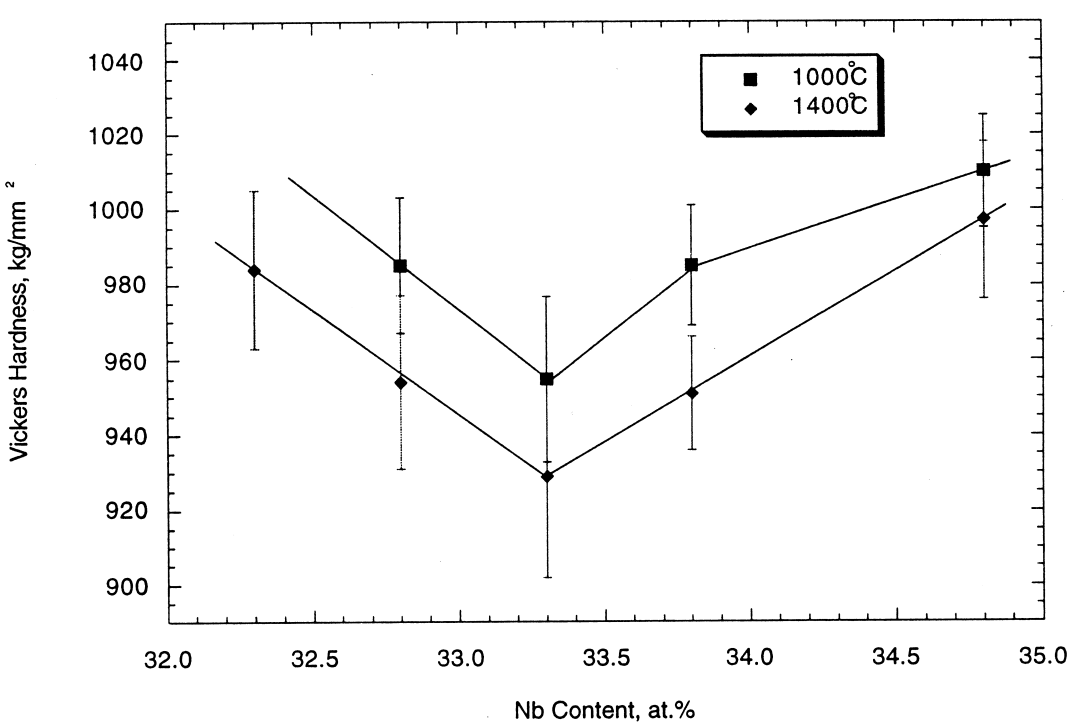


Fig. 4. Vickers hardness vs Nb content for $NbCr_2$ alloys after quenching from 1400 and 1000°C [31].

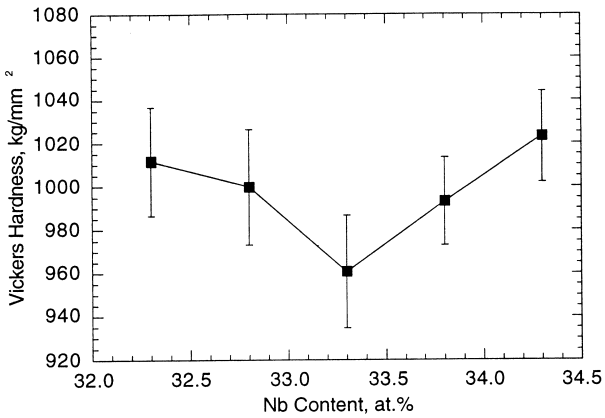


Fig. 5. Vickers hardness vs Nb content for $NbFe_2$ alloys after quenching from 1000°C [31].

defect; however, it is very difficult for ordinary dislocations to move in Laves phases at ambient temperatures.

The hardness values of all $NbCr_2$ alloys after quenching from 1400°C appear to be lower than those after quenching from 1000°C (Fig. 4). Since thermal vacancies are incorporated into the compound after quenching from 1400°C (see Fig. 1), the above results may suggest that vacancies cause softening, instead of hardening of the $NbCr_2$ Laves phases. Such vacancy-induced softening has not been observed previously in other systems. Furthermore, this is in direct contrast to the effect of vacancies on the hardness in B2 compounds, where vacancies are a potent hardener. The vacancy hardening in metals and B2 compounds is often correlated to an interaction between the stress field of a moving dislocation and that of the vacancy. The effect of vacancies in Laves phases may be different from their effects in metals and B2 compounds. This is again likely to be due to the fact that Laves phases have a topologically close-packed structure and it is very difficult for ordinary dislocations to move. The presence of vacancies in Laves phases would make the structure less closely packed and could possibly assist the synchroshear deformation of the alloy [21,36], thus possibly leading to the softening of the alloys.

The fracture toughness values as a function of the Nb content in $NbCr_2$ Laves phases after quenching from 1400 and 1000°C as well as $NbCo_2$ and $NbFe_2$ Laves phases after quenching from 1000°C are plotted in Fig. 6. Basically, the fracture toughness of the $NbCr_2$ and $NbCo_2$ alloys is close to 1 MPa $m^{1/2}$. No effect of stoichiometry on the fracture toughness can be discerned. Even though the anti-site defects harden the Laves phases (Figs. 4 and 5), they are not detrimental to the fracture toughness of the alloys. Furthermore, it is noted that the fracture toughness of the $NbCr_2$ alloys after quenching from 1400°C is similar to that of the alloys after quenching from 1000°C, i.e. the presence of thermal vacancies (<0.4%) does not lead to an improved toughness of the Laves phases, even though the vacancies slightly soften the Laves phases (see Fig. 4). Similar results were observed for $ZrCo_2$ alloys quenched from different temperatures [32]. The fracture toughness of the C14 $NbFe_2$ alloys is lower than that of C15 $NbCr_2$ and $NbCo_2$. This trend can be explained by the fact that the C15 structure is cubic, and is possibly more deformable than the hexagonal C14 structure.

It has been suggested that vacancies in Laves phases may assist the synchroshear mechanism, thus leading to improvement in the deformability and toughness of the Laves alloys [21,36]. Chen et al. [28] attributed the toughness improvement in the off-stoichiometric $TiCr_2$ alloys to the presence of vacancies in the alloys. The quenched-in thermal vacancies do not affect the crack propagation and fracture toughness behavior of the $NbCr_2$ alloys (Fig. 6). This is possibly due to the fact that the level of vacancies is too low in the alloys to enhance toughness significantly. It is also possible that the high residual thermal stress after quenching from 1400°C may also embrittle the $NbCr_2$ Laves phases, and thus could mask the effect of vacancy-assisted synchroshear deformation and its associated toughening.

The atomic size ratio, R_A/R_B, of the AB_2 Laves phases studied can be calculated to be 1.145, 1.173, and 1.152 for $NbCr_2$, $NbCo_2$, and $NbFe_2$, respectively. Here, R_A and R_B are the atomic radii of the A and B atoms with a coordination number of 12. Obviously, all the Laves phases studied have R_A/R_B ratios smaller than the ideal ratio of 1.225 for Laves-phase formation. It will be interesting to know the defect mechanism for Laves phases with R_A/R_B greater than 1.225. It may be easier to obtain constitutional vacancies for Laves phases with $R_A/R_B > 1.225$, due to geometric considerations, i.e. it is more difficult for A atoms to stay on the sublattice sites of B atoms as the atomic size ratio increases. On the other hand, it has been found that for the B2 phases, the

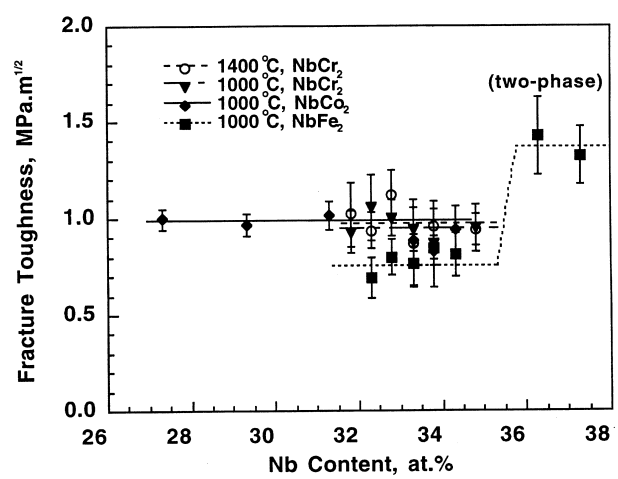


Fig. 6. Fracture toughness vs Nb content for binary $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves-phase alloys. Note two of the $NbFe_2$ alloys are not single Laves phase [31].

defect type is closely related to the enthalpies of formation (ΔH) of the compounds. Neumann has demonstrated that B2 compounds with negative ΔH smaller than 75–90 kJ/mol generally exhibit the anti-site defect structure, while compounds with a greater value generally exhibit the triple defect structure and have constitutional vacancies on the large-atom side of stoichiometry [37]. In other words, constitutional vacancies are the preferred defect mechanism in compounds that are more strongly ordered, i.e. with higher absolute ΔH values on the large-atom side. This empirical observation may be extended to the binary Laves phases. The ΔH values for $NbCr_2$, $NbCo_2$, and $NbFe_2$ are −21, −51, and −63 kJ/mol, respectively [38]. These values are in the same range as those of B2 compounds with anti-site defects. It is not surprising that only anti-site defects are observed in these Laves phases. Based on such geometric and thermodynamic considerations, future work should be concentrated on Laves phases with high R_A/R_B ratios and/or large negative enthalpies of formation. One such Laves phase is $ZrCo_2$, which has a large difference in atomic size between component atoms ($R_A/R_B = 1.28$) and an enthalpy-of-formation value of −123 kJ/mol [32]. Recent results indicate that a second phase, instead of constitutional vacancies, is formed in $ZrCo_2$ alloys with Zr > 33.3%. Interestingly, thermal vacancy concentrations in $ZrCo_2$ alloys of near stoichiometric composition are as high as 1%, see Fig. 7 [32]. These vacancies do not affect fracture toughness.

3. Phase stability and fracture toughness of transition-metal Laves phases

As mentioned above, the different Laves phase polytypes are a result of different stacking sequences. The stacking fault energy is expected to be lower at phase boundaries (e.g. C14/C15 boundary), where mechanical twinning becomes more easily triggered.

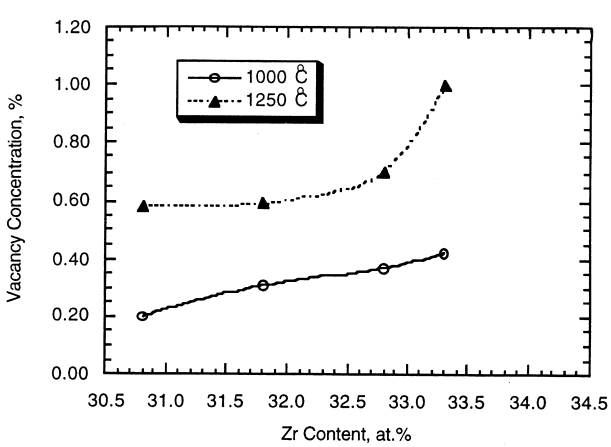


Fig. 7. Vacancy concentration vs Zr content for $ZrCo_2$ Laves phase quenched from 1000 and 1250°C [32].

The stability ranges of the Laves phases in both binary and ternary systems were obtained from phase diagram information, and the e/a effect on the phase stability (C14/C15) in $NbCr_2$-based transition-metal Laves phases is shown in Fig. 8 [39,40]. In the $NbCr_2$-based Laves phases, the e/a ratio for the C15/C14 phase boundaries is quite precise; the critical e/a values, corresponding to the C15/C14/C15 phase transitions from the phase diagram information (as indicated in Fig. 8), are as follows: at e/a values lower than 5.76, the C15 structure is stabilized for both binary and ternary Laves alloys; by increasing e/a to 5.88, the C14 structure is stabilized; over the e/a range of 5.88–7.53, the C14 structure is more stable than the C15 structure; the C15 structure is stabilized again over the C14 structure when the e/a ratio is increased further to 7.65 for the Nb–Cr–Co system. In the ternary Nb–Cr–Fe system, such a C14→C15 transition was not observed, since the highest e/a ratio was 6.69 in this system (Fig. 8). No $NbNi_2$ (with e/a = 8.34) or $NbCu_2$ (with e/a = 9) Laves phases exist in binary Nb–Ni and Nb–Cu systems (the e/a values for imaginary "$NbNi_2$" and "$NbCu_2$" were shown in Fig. 8), consistent with the observation by Bardos et al. [41] that at e/a > 8, a disordered structure is stabilized over the Laves phase in transition-metal systems.

The phase stability in pseudo-binary $Nb(Cr,Fe)_2$ and $Nb(Cr,Co)_2$ systems was also studied experimentally and good agreement between the experimental results and surveyed data based on phase diagram information was observed, as shown in Fig. 8. It is clear that the electron concentration rule is obeyed in these two pseudo-binary systems: the phase transitions between C15 and C14 are determined by the average electron concentration in these alloys. The experimental critical e/a ratios for the C15/C14/C15 transitions are very close to those from phase diagrams. These results were the first experimental determination of the critical e/a values for the different Laves phases in the pseudo-binary $Nb(Cr,X)_2$ systems. Such good agreement may be associated with the fact that all the components in the two

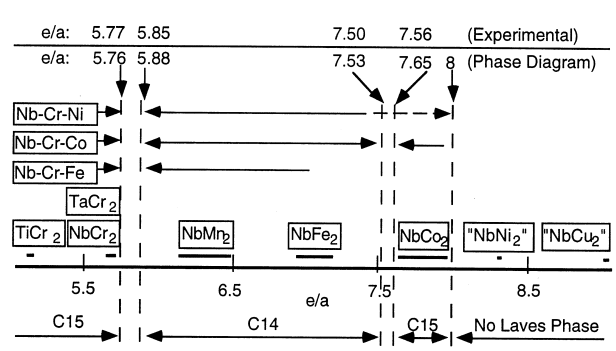


Fig. 8. Effect of average electron concentration (e/a) on phase stability in $NbCr_2$-based systems, with the critical e/a values for phase transitions both from phase diagram information and experimental data [40].

pseudo-binary systems are transition metals. Furthermore, it is determined that the region between C14 and C15 is a two-phase mixture of C14 and C15, with no intermediate structures such as C36 observed.

The effect of Fe and Co contents on the fracture toughness of $Nb(Cr,X)_2$ Laves phase is shown in Figs. 9 and 10, respectively [40]. Basically, the fracture toughness of the monolithic Laves phase is very low, around 1 MPa$\sqrt{m}$, which is a consequence of the topologically close packing manner of the structure and the subsequent difficulty in activating dislocation sources in the Laves phase. The cubic C15 phase has relatively higher fracture toughness than that of C14 structure for the ternary systems, regardless of the Fe or Co addition. This can be explained by the fact that the C15 structure is cubic, and is thus more deformable than the hexagonal C14 structure. The toughness value near the phase boundary could be higher, in view of the fact that the associated structure is metastable with a lower stacking fault energy, thus possibly making stress-assisted phase transformation and/or mechanical twinning easily triggered during deformation. However, such a trend was not observed in the present study. None of the alloys within the two-phase regions or near the phase boundaries exhibits any sign of increase in toughness. The toughness of the two-phase (C14 + C15) alloys is similar to that of the C15 phase. The toughness enhancement through stress-assisted phase transformation or mechanical twinning is not significant enough to be noted in the indentation test.

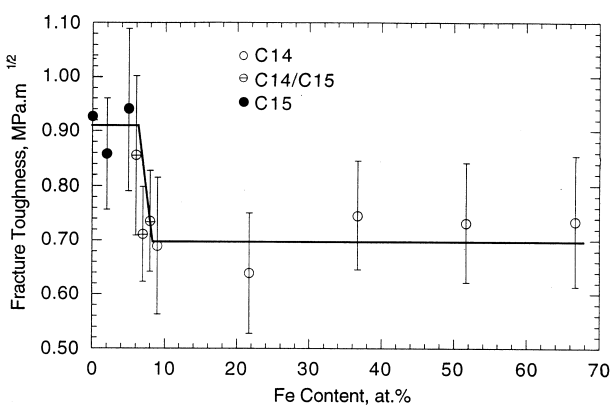


Fig. 9. Fracture toughness values of $Nb(Cr,Fe)_2$ as a function of the Fe content [40].

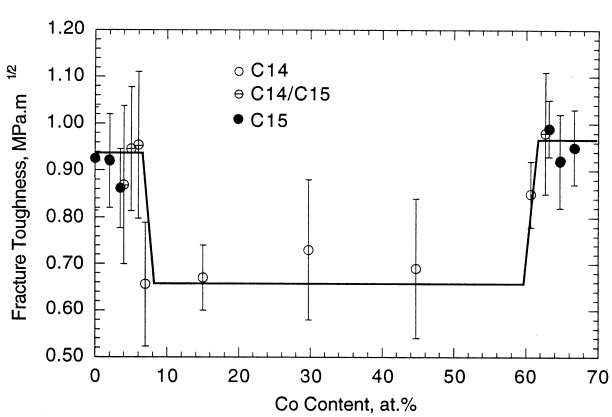


Fig. 10. Fracture toughness values of $Nb(Cr,Co)_2$ as a function of the Co content [40].

4. Alloy design of dual-phase Cr-XCr_2 alloys

As indicated in the forgoing two sections, single-phase Laves phase alloys with cubic C15 and hexagonal C14 are hard and brittle, and their fracture toughness at ambient temperatures can not be effectively improved using physical metallurgy principles, including control of point defect and phase stability. In view of this, dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard Laves phase second phases have been selected for alloy development [17,22–26] for potential structural use at elevated temperatures. This section summarizes the progress made so far in the development of dual-phase Cr–XCr_2 alloys, where X = Nb, Ta, or Zr.

Three eutectic alloy systems, Cr–$NbCr_2$, Cr–$TaCr_2$, and Cr–$ZrCr_2$, were initially selected for alloy development study. The selection is based on consideration of the different eutectic phase relationships in these alloy systems [16,25,26,42], as indicated in Table 1. Fig. 11 shows the Cr–$TaCr_2$ phase diagram which is typical for the three eutectic systems. Among these systems, the Cr–Nb system has the highest maximum solute solubility (5.6% Nb) in the Cr solid solution and the highest solute concentration in the eutectic composition (18.5% Nb), whereas the Cr–Zr system has the least maximum solute solubility (< 0.6% Zr) in Cr and the Cr–Ta system has the lowest solute concentration in the eutectic composition (9.8% Ta). These alloy parameters are expected to strongly affect the hardening behavior of both the Cr solid solution phase and the eutectic structure and thus the degree of brittleness of these phases.

Alloy ingots containing up to 20 at. % X element were prepared by arc melting and drop casting, using pure metal elements. In order to control precipitation reactions, alloy ingots were sectioned and heat treated between 1000 and 1400°C. The first interesting observation was that alloy ingots were quite often cracked for Cr–$NbCr_2$ alloys but not for Cr–$TaCr_2$ and Cr–$ZrCr_2$

Table 1
Alloy parameters existing in the three eutectic systems Cr–Nb, Cr–Ta and Cr–Zr

Alloy parameter	Cr–Nb	Cr–Ta	Cr–Zr
Eutectic temperature (C°)	1703	1760	1592
Max. solute solubility in Cr (at%)	5.6 Nb	~4.0 Ta	< 0.6 Zr
Eutectic composition (at%)	18.5 Nb	9.8 Ta	~18.0 Zr

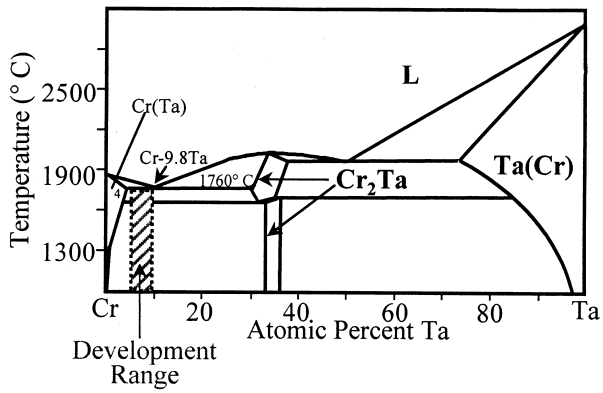


Fig. 11. Schematic binary Cr–Ta phase diagram.

alloys during drop casting and cooling from heat treatment at high temperatures. This cracking behavior cannot be simply related to the difference in coefficients in thermal expansion (CTE) because both experimental measurements and first principle calculations [43–45] indicate no major difference in CTE among the Laves phases $NbCr_2$, $TaCr_2$ and $ZrCr_2$. The cracking behavior, on the other hand, can be better rationalized by considering alloy parameters in the three eutectic systems. As indicated in Table 1, the maximum solute solubility in Cr and the solute concentration in the eutectic structures (which increases the proportion of the hard Laves phases in the eutectic microstructure) are the highest in the Cr–$NbCr_2$ system, resulting in significant hardening and embrittlement of the primary Cr solid solution phase and the eutectic composition. In this case, tremendous hardening is induced from fast cooling from the maximum-solubility temperature to lower temperatures where the primary Cr solid solution is in supersaturated states. As a result, thermal induced stresses generated by CTE mismatch will high enough to crack these phases during drop casting or cooling from high temperatures [46,47], quite often resulting in catastrophic ingot failure. This study clearly indicates the importance of phase relationship that may provide key information governing the mechanical and metallurgical behaviors of individual alloy phases.

In comparison with the dual-phase Cr–$TaCr_2$ alloys, Cr–$ZrCr_2$ alloys suffer from two major problems. One is the poor oxidation resistance of Cr–$ZrCr_2$ alloys, due, in part, to extensive internal oxidation of Zr-rich phases, and the other is the relative low melting point of these alloys (see Table 1). Thus, our alloy development has been concentrated on the Cr–$TaCr_2$ alloy system, and significant progress has been made in improving the mechanical and metallurgical properties of these alloys [26]. The Cr–$TaCr_2$ alloys with enhanced properties have the general compositions (at%) below:

$$\mathrm{Cr}-(6-10)\mathrm{Ta}-(3-6)\mathrm{Mo}-(0.2-1.0)\mathrm{Ti}-(0.5-3.0)\mathrm{Si}-(0.05-0.20)\mathrm{La/Ce}.$$

In these alloys, Mo is added mainly for solid solution hardening the Cr-rich phase, Ti for scavenging interstitials, and Si for both improving oxidation resistance and fracture toughness. Small amounts of reactive elements such as La and Ce are also added for decreased scaling and enhancing the adhesion of oxide scales.

No micro- and macro-cracks were observed in the Cr–$TaCr_2$ alloy ingots prepared by arc-melting and drop casting. In order to control microstructures, some dual-phase alloys were canned in Mo billets and fabricated by hot extrusion at 1480°C. The alloy CN-151 (Cr-9.75Ta-5Mo-0.5Ti-0.01Ce, at. %) with its composition similar to the eutectic composition shows a full lamellar structure in the as-cast condition and a $TaCr_2$-dispersed microstructure in the hot extruded condition (Fig. 12). The mechanical properties with different microstructures were determined in tension as a function of test tem-

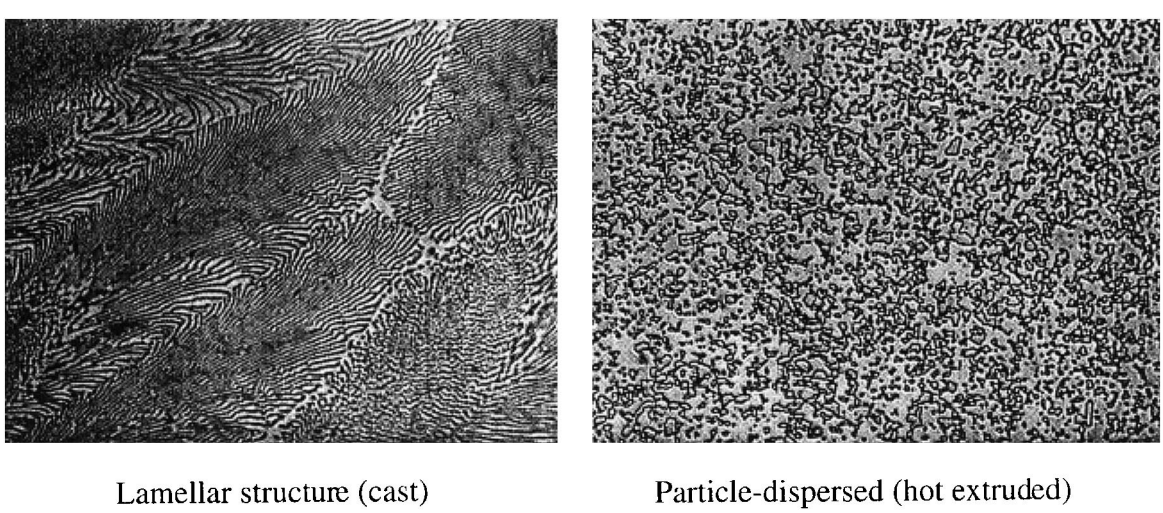


Fig. 12. Comparison of lamellar structure in the cast condition and particle-dispersed microstructure in the hot extruded condition for alloy CN-151.

perature. All tensile specimens were given a final heat treatment of 1 d/1200°C + 1 d/1000°C for stress relief. Note that the heat treatment does not significantly affect the lamellar structure observed in cast alloys.

Table 2 compares the tensile properties of CN-144 alloy with the lamellar and $TaCr_2$-dispersed microstructures. With the dispersed microstructure, the alloy showed a fracture strength of more than 100 ksi (700 MPa), even though no appreciable plastic deformation was measured at room temperature. Plastic deformation was detected at 800°C, and the strain went up to as high as 40% at 1200°C. The alloy is very strong at elevated temperatures, with the yield strength of 37 ksi and the ultimate tensile strength of 46 ksi at 1200°C. Note that the strengths of most Ni-base superalloys vanish at 1200°C. In comparison with the particle-dispersed microstructure, the alloy with the lamellar structure showed the ductile-to-brittle temperature around 1000°C, instead of 800°C. The lamellar structure is extremely strong at 1200°C, with its strengths close to double that for the particle-dispersed microstructure.

The fracture toughness of Cr–$TaCr_2$ alloys with lamellar structures was determined at room temperature, using subsized notched Chevron specimens [48] in accordance with reference 49. The results are summarized in Table 3. The alloy CN-144 with a hypoeutectic composition showed a low fracture toughness of 8.3 MPa $m^{1/2}$. A significant increase in fracture toughness is achieved by increasing the Ta concentration from hypo- to hypereutectic compositions. Alloying with 1.0% Si further increases toughness from 11.5 to 14.3 MPa $m^{-1/2}$. Examination of fracture surfaces reveals that the interfaces between the Cr solid solution and lamellar $TaCr_2$ phases are strong, and no secondary cracks are observed along these interfaces. It is worth noting that hot extrusion and change in microstructures from lamellar to dispersed particles gives no significant increase in fracture toughness.

Table 2
Comparison of tensile properties[a] of CN-144 with lamellar and particle-dispersed microstructures

Test temperature (C°)	Strength (ksi) Yield	Fracture	Elongation (%)
Particle-dispersed microstructure (hot extruded)			
RT	–	104	0.3
800	130	133	1.0
1000	92	103	1.7
1200	37	46	39.0
Lamellar structure (cast)			
RT	–	76	0.2
800	–	–	–
1000	–	80	0.2
1200	68	85	5.3

[a] Tested at a strain rate of 3.3×10^{-3}/s in air.

The creep properties of a Cr–$TaCr_2$ alloy, CN-144 (Cr-8Ta-5Mo-0.5Ti-0.01Ce. at%), were determined in tension at a stress of 20 ksi (138 MPa) at 1000°C in air [26]. Fig. 13 compares the creep curve of CN-144 with the lamellar and dispersed microstructures. The $TaCr_2$ dispersed microstructure exhibited ductile rupture with a rupture life of 695 h at 1000°C. In comparison, the specimen with the lamellar structure is much stronger in creep, with a steady state creep rate lower than that of the particle-dispersed microstructure by an order of magnitude. The test of the lamellar specimen had to be stopped after 1438 h because of grip rod failure. The creep resistance of this alloy in the polycrystalline form with a lamellar structure is stronger than Mar M-200 and DS Rene'80 and is comparable to CMSX-11B in the single crystal form.

The oxidation properties of Cr–XCr_2 alloys were characterized by exposing alloy coupons to air at 900–1100°C. The oxidation studies by Brady et al. [26,50,51] demonstrates that the oxidation resistance of Cr–$TaCr_2$ alloys is much superior to Cr–$NbCr_2$ or Cr–$ZrCr_2$ alloys. Fig. 14 compares the cyclic oxidation behavior of Cr-8Ta alloys with commercial Cr-base alloys tested at

Table 3
Fracture toughness of Cr–Ta alloys at room temperature

Alloy number	Composition[a] (at%)	Fracture toughness (MPa $m^{1/2}$)
CN-144	8.0 Ta–0.01 Ce	8.3
CN-161	9.8 Ta–0.1 La	11.5
CN-145	10.0 Ta–0.01 Ce	12.8
CN-162	9.8 Ta–0.1 La–0.5 Si	11.4
CN-163	9.8 Ta–0.1 La–1.0 Si	14.3

[a] Base composition: Cr-5 Mo-0.5 Ti, at%.

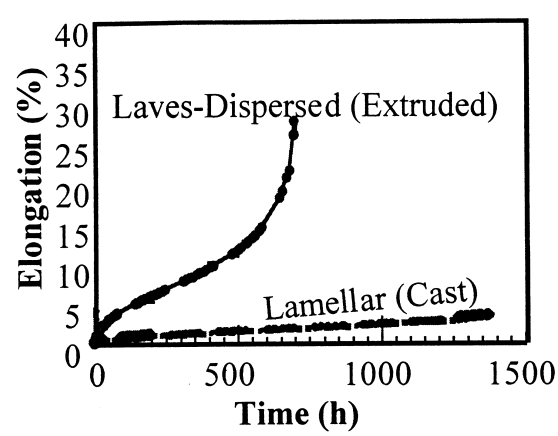


Fig. 13. Comparison of creep properties of CN-144 (Cr-8Ta-5Mo-0.5Ti-0.OlCe, at%) with Laves phase dispersed or lamel lar microstructures tested at 1000°C and 20 ksi in air. Note that the lamellar microstructuie test was stopped due to failure of the grip rods and plate.

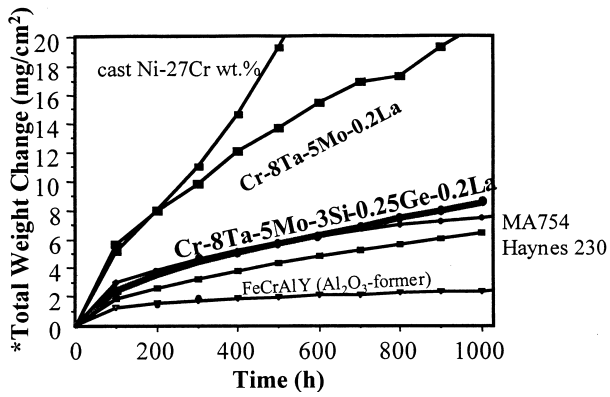


Fig. 14. Comparison of cyclic oxidation of Cr-8Ta alloys with cast Ni-27Cr alloy (wt%) and commercial alloys MA754 and Haynes 230 air oxidized at 1100°C.

1100°C in air [26]. The oxidation rate of Cr-8Ta-5Mo-0.2La is comparable to that of cast Ni-27Cr (wt%), and alloying with additions of Si and Ge significantly lowers the oxidation rate of the Cr–8Ta alloy. As indicated in Fig. 14, the oxidation resistance of Cr-8Ta-5Mo-3Si-0.25Ge-0.2La is comparable to several commercial Cr_2O_3-forming alloys, such as MA-754 and Haynes 230.

5. Conclusions

This paper provides a comprehensive review of phase stability, point defects, and fracture toughness of binary AB2 Laves phase alloys based on $NbCr_2$, $NbFe_2$, $NbCo_2$, and $ZrCo_2$ with on- and off-stoichiometric compositions, and pseudo-binary alloys based on $Nb(Cr,Co)_2$ and $Nb(Cr,Fe)_2$ compositions. The alloy design effort has been concentrated mainly on dual-phase alloys based on a soft Cr-rich solid solution reinforced with hard $TaCr_2$ second phases. Conclusions deduced from these studies are given below:

1. Anti-site defects were observed to form on both sides of the stoichiometry in $NbCr_2$, $NbCo_2$, and $NbFe_2$ with $R_A/R_B < 1.225$. On the other hand, for $ZrCo_2$ with $R_A/R_B > 1.225$, anti-site defects form only on the Co-rich side and second phase particles form on the Zr-rich side.
2. No constitutional vacancies but only thermal vacancies were detected in the Laves phase alloys quenched from high temperatures. The peak vacancy concentration is always located at the stoichiometric composition of each Laves phase alloy.
3. Room-temperature fracture toughness can not be effectively improved either by increasing thermal vacancy concentration at levels below 1% or by reducing stacking fault energy through control of the phase stability in $Nb(Cr,Co)_2$ and $Nb(Cr,Fe)_2$ alloys.
4. Drop casting and heat treatments induce micro- and macro-cracks in dual-phase Cr–$NbCr_2$ alloys but not in Cr–$TaCr_2$, and Cr–$ZrCr_2$ alloys. The cracking behavior can not be simply explained from CTE mismatch; instead, it can be better rationalized from the consideration of eutectic composition and phase relationship.
5. Among the three dual-phase Cr–XCr_2 alloys (where X = Nb, Ta, or Zr), Cr–$TaCr_2$ alloys have the best combination of mechanical and metallurgical properties for structural use at elevated temperatures. These alloys have excellent tensile strength, creep properties, and oxidation resistance at elevated temperatures in air.
6. The alloy design effort has led to the identification of alloy compositions (at%) with enhanced mechanical and metallurgical properties:

$$\mathrm{Cr} - (6-10)\mathrm{Ta} - (3-6)\mathrm{Mo} - (0.2-1.0)\mathrm{Ti} - (0.5-3.0)\mathrm{Si} - (0.05-0.20)\mathrm{La/Ce}.$$

7. The current Cr–$TaCr_2$ alloys showed fracture toughness up to 14.3 M Pa $m^{-1/2}$ at room temperature. Additional development is required to further improve the fracture resistance of these alloys by interfacial design and microalloying.

Acknowledgements

This research was sponsored jointly by the Division of Materials Sciences and Engineering, and the Fossil Energy AR&TD Materials Program, US Department of Energy under contract DE-AC05-00OR22725 with UT-Battelle, LLC. The authors would like to thank L. Heatherly and D. S. Easton for hot extrusion, J. L. Wright for mechanical testing, and C. A. Carmichael for alloy preparation.

References

[1] Laves F. Theory of alloy phases. Metals Park, OH: American Society for Metals, 1956. p. 124
[2] Wernick JH. In: Westbrook JH, editor. Intermetallic compounds. New York: Wiley, 1967. p. 197.
[3] Sinha AK. Progr Mater Sci 1972;15(2):79.
[4] Massalski TB. In: Cahn RW, Hassen P, editors. Physical metallurgy part 1. New York: North-Holland, 1983. p. 190.
[5] Laves F, Witte H. Metallwirschaft 1935;14:645.
[6] Laves F, Witte H. Metallwirschaft 1936;15:840.
[7] Moffett MB, Clark AE, Wun-Fogle M, Lindberg J, Teter JB, McLaughlin EA. J Acoust Soc Am 1991;89:1448.
[8] Inoue K, Tachikawa K. IEEE Trans Magnetics 1979;15:635.
[9] Olzi E, Matacotta FC, Setina P. J Less-Common Met 1988;139:123.
[10] Ivey D, Northwood D. J Less-Common Met 1986;115:23.
[11] Halstead A, Rawlings RD. J. Mater. Sci. 1985;20:1248 & 1693
[12] Anton DL, Shah DM. Mater Res Soc Symp Proc 1990;194:45.

[13] Livingston JD. Phys Status Solidi (a) 1992;131:415.
[14] Chu F, Pope DP. Mater Sci Engng 1993;A170:39.
[15] Kumar KS, Miracle DN. Intermetallics 1994;2:257.
[16] Thoma DJ, Perepezko JH. Mater Sci Engng 1992;A156:97.
[17] Takeyama M, Liu CT. Mater Sci Engng 1991;A132:61.
[18] Chan KS, Davidson DL. JOM 1996;48:62.
[19] Fleischer RL, Zabala RJ. Metall Trans 1990;21A:2149.
[20] Chen KC, Allen SM, Livingston JD. J Mater Res 1997;12:1472.
[21] Hazzledine PM, Kumar KS, Miracle DB, Jackson AG. MRS Symp Proc 1992;288:591.
[22] Liu CT, Tortorelli PF, Horton JA, Carmichael CA. Mater Sci Engng 1996;A214:23.
[23] Kumar KS, Liu CT. Acta Mater 1997;45:3671.
[24] Takasuqi T, Kumar KS, Liu CT, Lee EH. Mater Sci Engng 1999;A260:108.
[25] Kumar KS, Pang L, Liu CT, Hoton JA, Kenik EA. Acta Mater 2000;48:911.
[26] Brady MP, Zhu JH, Liu CT, Tortorelli PF, Walker LR, McKamey CG, Wright JL, Carmichael CA, Larson DJ, Miller MK, Porter WD. Materials at High Temperatures 1999;16(4):189–93.
[27] Thoma DJ, Perepezko JH. J Alloys and Compounds 1995;224:330.
[28] Chen KC, Allen SM, Livingston JD. MRS Symp Proc 1994;364:1406.
[29] Ivey DG, Northwood DO. Z Phys Chem NF 1986;147:191.
[30] Sicking G, Magomedbekov E, Hempelmann R. Ber Bunsenges Phys Chem 1981;85:686.
[31] Zhu JH, Pike LM, Liu CT, Liaw PK. Acta Mater 1999;47:2003.
[32] Zhu JH, Liu CT. Defect structures in $ZrCo_2$ Laves phase, Acta Mater. 2000 (in press).
[33] Modder IW, Bakker H. Phys Stat Sol (b) 1997;199:369.
[34] Modder IW, Kuin MJ, Bakker H. Intermetallics 1998;6:537.
[35] Pike LM, Chang YA, Liu CT. Acta Mater 1997;45:3709.
[36] Hazzledine PM, Pirouz P. Scripta Metall Mater 1993;28:1277.
[37] Neumann JP. Acta Metall 1980;28:1165.
[38] de Boer FR, Boom R, Mattens WCM, Miedema AR, Niessen AK. Cohesion in metals and transition metal alloys, ed. de Boer FR, Pettifor DG. Vol. 1, Second Corrected Printing, North Holland, Amsterdam, 1989.
[39] Zhu JH, Liaw PK, Liu CT. Mater Sci Engng 1997;A239(240):260.
[40] Zhu JH, Liu CT, Liaw PK. Intermetallics 1999;7:1011.
[41] Bardos DI, Gupta KP, Beck PA. Trans Met Soc AIME 1961;221:1087.
[42] Massalski TB, ed. Binary alloy phase diagrams, vol. 1, ASM Publication, 1986.
[43] Thoma DJ. Unpublished results, Lawrence Livermore National Laboratory, 1998.
[44] Liu CT. Unpublished results, Oak Ridge National Laboratory, 1999.
[45] Suklyun Hong, Fu CL. in MRS Proc. vol. 552, MRS Publication, 1999:643–647.
[46] Lee S, Liaw PK, Liu CT, Chou YT. Mater Sci Engng A 1999;268:184–92.
[47] Lee S, Liaw PK, Liu CT, Chou YT. Mater Chem Phys 1999;61:207–13.
[48] Jenkins MG, Kobayashi AS, White KW, Bradt RC. Inst J Fract 1987;34:281.
[49] Schneibel JH, Carmichael CA, Specht ED, Subramanian R. Intermetallics 1997;5:61–7.
[50] Brady MP, Zhu JH, Liu CT, Tortorelli PF, Wright JL, Carmichael CA, Walker LR. Ultrahigh temperature intermetallic alloys. Proceedings of the Eleventh Annual Conference on Fossil Energy Materials, Knoxville, TN, 20–22 May, R.R. Judkins, Compiler, 1997. pp. 195–205.
[51] Brady MP, Tortorelli PF, Walker LR. Materials at high temperatures (in press).

ELSEVIER

Intermetallics 8 (2000) 151–155

Intermetallics

Thermal conductivity of Ni-Al powder compacts for reaction synthesis

Seiji Miura [a,*], Yoshihiro Terada [b], Tomoo Suzuki [c], C.T. Liu [d], Yoshinao Mishima [e]

[a] *Division of Materials Science and Engineering, Graduate School of Engineering, Hokkaido University, Kita-13, Nishi-8, Kita-ku, Sapporo 060-8628, Japan*
[b] *Department of Metallurgy and Ceramics Science, Tokyo Institute of Technology, 2-12-1 O-okayama, Meguro-ku, Tokyo 152-8552, Japan*
[c] *Kochi University of Technology, Tosayamada-cho, Kochi 782-8520, Japan*
[d] *Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN37831-6115, USA*
[e] *Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan*

Received 7 July 1999; accepted 23 August 1999

Abstract

Thermal conductivity of 3Ni–1Al elemental powder compacts was measured using a laser-flash method as a function of temperature, compact density and heat treatment. Temperature dependence of thermal conductivity of the compacts is similar to that of pure Ni and Al metals. Heat treatment has a significant effect on the thermal conductivity of the compacts as particles are joined together to form Ni solid solution and intermetallic phases on powder surfaces. Fast diffusion on powder surfaces during heat treatment results in the formation of intermetallic phase zones in the interior of the former Ni particles. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Nickel-aluminides, based on TiAl; B. Thermal properties; C. Reaction synthesis; C. Powder metallurgy

1. Introduction

Reaction synthesis has recently been studied by many researchers as a novel processing technique for fabrication of high temperature ordered intermetallic alloys from elemental powders [1–3]. This technique has advantages over conventional ingot metallurgy in achieving alloy homogeneity, near net shape, and no contamination from crucibles, resulting in cost and energy savings. The reaction sustains itself with heat generated by forming intermediate phases through interdiffusion processes, and it is, therefore, called self-sustaining high-temperature synthesis (SHS). The heat generation and transfer are key factors controlling the degree of reaction evolution.

In a previous study, it was shown that the initiation of reaction synthesis in 3Ni–1Al mixed powder compacts is strongly affected by initial compact thickness [4]. Full or partial reaction was observed in the specimens having a thickness of 7–6 mm within a few minutes when heated to 620°C; however, such SHS was not observed in a 5 mm thick Ni-Al mixed powder compact even after 1.8 ks at 620°C. Note that all the tests were under a uni-axial compressive stress of about 10 MPa. It is believed that the accumulation and transfer of the heat generated by the formation of intermetallic phases are sensitive to the specimen thickness and can be predicted by modeling.

For the modeling the rate of heat generation can be estimated through the formation rate of intermediate phases in terms of bulk diffusion, whereas the heat transfer is controlled by thermal conductivity [5,6]. Some of physical properties required for understanding and predicting the SHS process have been compiled and listed in the literature, but other information is quite limited [7–13]. Among them available data on the thermal conductivity of intermetallics and those of mixed powder compacts are very limited, which are important for understanding the evolution of reaction synthesis. Although a few equations have been proposed to estimate the thermal conductivity of porous materials, it would be much better if the thermal conductivity can be directly measured from experiments. The thermal conductivity of powder compacts is expected to be affected

* Corresponding author.
E-mail address: miura@eng.hokudai.ac.jp (S. Miura).

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(99)00080-1

Modelling Simul. Mater. Sci. Eng. **8** (2000) 25–35. Printed in the UK PII: S0965-0393(00)08126-2

Large-scale atomistic study of core structures and energetics of $\langle c + a \rangle$ $\{11\bar{2}2\}$ dislocations in hexagonal close packed metals

J R Morris†, K M Ho†, K Y Chen†, G Rengarajan‡‖ and M H Yoo§

† Ames Laboratory, US Department of Energy, Ames, IA 50011–3020, USA
‡ Department of Mechanical Engineering, Texas A & M University, College Station, TX 77843–3123, USA
§ Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831–6115, USA

Received 7 October 1998, accepted for publication 15 September 1999

Abstract. Using two different kinds of many-body potentials as well as the Lennard-Jones potential for hexagonal close packed metals, we have found that $\langle c + a \rangle$ edge dislocations with dislocation lines along $\langle 1\bar{1}00 \rangle$ can split onto the basal plane, forming a non-planar sessile structure. The 'type I' undissociated dislocation core, observed in previous papers, is shown to be stable only for small simulations. The observed dissociated core structure has a large distorted region that we interpret as a $(11\bar{2}1)$ twin nucleus, which may help the formation of $(11\bar{2}1)$ tension twins. We also find that this core structure is lower in energy than the previously observed glissile 'type II' configuration, dissociated on the $(11\bar{2}2)$ plane. The sessile splitting of the core gives rise to non-Schmid behaviour, with the twin nucleus expanding under c-axis tension, and contracting under compression.

1. Introduction

The study of dislocation core structures and the energetics associated with slip and twinning provides valuable information for the understanding of strength and ductility of materials. It is now well established that the core effects play a major role in the plastic deformation of metals [1], and could explain such behaviour as the temperature dependence of yield stress. It is also known that complex dissociations of the core significantly influence mechanical properties [2]. Atomistic simulations of dislocations are becoming viable tools to study the associated core structures, due to the increase in computational power and improvements in modelling materials at the atomistic level.

While numerous studies in the past have dealt with face centred cubic (fcc) and body centred cubic (bcc) material systems, relatively less focus has been placed on hexagonal close packed (hcp) materials. It has been known that basal slip, non-basal slip and twinning are the primary modes of plastic deformation in hcp metals (for a review, see [3]). Dislocations with a Burgers vector in the basal plane, $\boldsymbol{b} = \frac{1}{3}\langle 11\bar{2}0 \rangle$ (denoted as $\langle \boldsymbol{a} \rangle$ in this paper) may split in a manner similar to $\langle 10\bar{1} \rangle$ fcc dislocations. This dissociation may be understood by considering simple hard-sphere models. However, dislocations with non-basal Burgers vectors (e.g. $\langle \boldsymbol{c} + \boldsymbol{a} \rangle$) also significantly affect the mechanical properties of polycrystalline hcp metals, and are considerably more complicated to study than the basal slip mode [3].

‖ Current address: General Electric Corporate Research and Development, Schenectady, NY, USA.

 © 2000 IOP Publishing Ltd

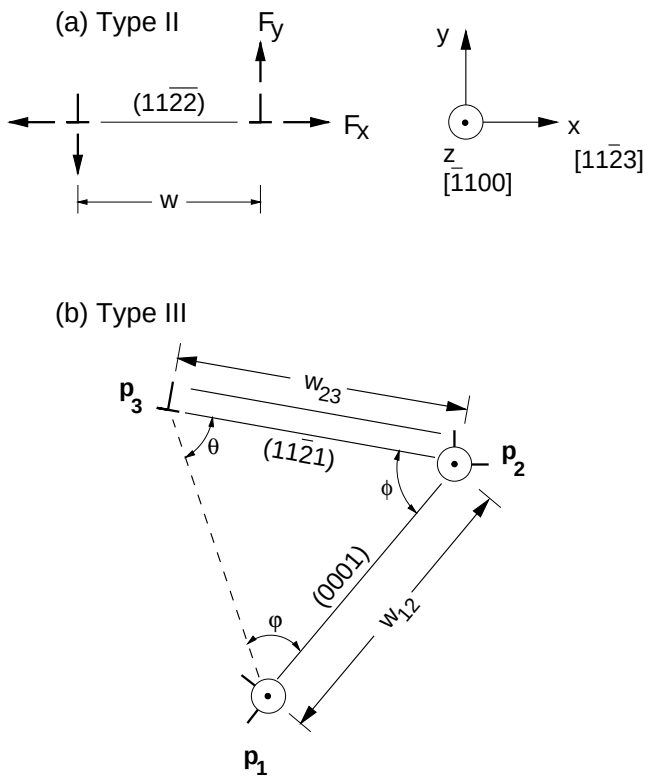


Figure 1. Schematic diagrams (*a*) and (*b*) of the type II and type III dislocation core structures, respectively. In the type III dislocation, (*b*), the dislocation has undergone a splitting into a Shockley partial $\boldsymbol{p}_1$ connected by a basal-plane stacking fault to a sessile partial $\boldsymbol{p}_2$. The twinned region extends between $\boldsymbol{p}_2$ along the (11$\bar{2}$1) plane to the twinning partials located at $\boldsymbol{p}_3$. This dissociation is given in (7). The distances and angles are given in table 3.

When a single crystal of hcp metal is deformed in compression or tension along the c-axis, plastic deformation can occur by pyramidal slip systems, e.g. [11$\bar{2}\bar{3}$]/3 slip on either the first-order (10$\bar{1}$1) or the second-order (11$\bar{2}$2) plane, and/or by compression twins including the (11$\bar{2}$2)[11$\bar{2}\bar{3}$] system or tension twin systems [3]. Because of the large magnitude of the Burgers vector, the $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ dislocation may be either decomposed to $\boldsymbol{c}$ and $\langle \boldsymbol{a}\rangle$ dislocations, or dissociated into partial dislocations that may have some relationships with the atomic movements involved in twinning. Therefore, detailed information on the source and mobility of $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ dislocations is important in understanding the mechanical behaviour of polycrystalline hcp metals and alloys.

In this paper, we re-examine previous simulations of the $\{11\bar{2}2\}\frac{1}{3}\langle 11\bar{2}\bar{3}\rangle$ edge dislocations. The previous simulations, performed by Minonishi *et al* [4, 5] and Liang and Bacon [6], produced two types of core structures: type I—an undissociated core, and type II—a dissociated core as shown in figure 1(*a*):

$$\boldsymbol{b} \longrightarrow \tfrac{1}{2}\boldsymbol{b} + \tfrac{1}{2}\boldsymbol{b}. \tag{1}$$

While this structure is expected to be lower in energy than the undissociated structure, none of the previous studies found either core to be significantly lower in energy than the other.

Our primary result is that the undissociated, 'type I' dislocation is not observed in our simulations, and that the previous simulations produced this result only due to the small size of

their simulations. An undissociated core is not observed, except in very small-size simulations. For our larger simulations, we observe either the type II core or a basal-plane dissociation after relaxation of an initially undissociated core, depending upon the initial conditions. The basal plane dissociated core, which we refer to as a type III core, is similar to that seen in recent simulations of the (Lennard-Jones) lj56 $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ edge dislocation [7]. In order to test the reliability of our results, we have used two different many-body potentials, developed for hcp metals, and also the lj56 potential used both by Minonishi [4, 5] and by Liang and Bacon [6]. In all cases, we see only the type II and type III dislocations. We believe that these two types of dislocations can account for the difference in behaviour between the brittle systems with limited twinning (Mg and Be), and ductile systems such as Ti and Zr.

2. Simulation methods and potential

As indicated above, we have used several interatomic potentials, to test the sensitivity of the results to the potentials. We believe that none of the potentials are sufficiently accurate for quantitative descriptions of the real materials and the corresponding energetics. However, structures that are potential independent are likely to represent (qualitatively) the structures that can be observed in real materials. As we shall show in subsequent sections, the observed core dissociations are independent of potential (although specific partial separations are different).

One of the potentials that we have chosen to use is the embedded atom method (EAM) potential developed by Zhang *et al* [8] (see details in [9]). This potential predicts that the hcp phase is stable relative to the fcc and bcc phases, with reasonable elastic constants and phonons for both the hcp and bcc phases. The elastic constants and other parameters are shown in table 1. This potential has been used to model the behaviour of the high-temperature bcc phase of Zr [8] and also the structure of twin boundaries in Zr [9]. More recently, it has been used to explore stacking faults on the $\{11\bar{2}2\}$ plane [10], related to the type II dislocation core structure described below. While EAM models are not ideal for modelling materials with directional bonding [11], including Ti and Zr, we feel that this model serves as a model hcp material, and may be more realistic than the simple pair potentials that have been used previously. However, it has an unphysically low basal-plane, stacking-fault energy, compared with *ab initio* potentials (see table 2).

Table 1. Equilibrium values calculated for EAM and FS models of hcp Zr, compared with experimental values.

Parameter	EAM [10]	FS [13]	Experiment	Reference
E_{coh} (eV)	6.253 79	6.25	6.25	Brewer [25]
a (Å)	3.21	3.25	3.23	Goldak *et al* [26]
c/a	1.63	1.595	1.592	Goldak *et al* [26]
C_{11} (Mbar)	1.63	1.5	1.554	Fisher and Renken [27]
C_{12} (Mbar)	0.93	0.85	0.672	
C_{13} (Mbar)	0.78	0.67	0.646	
C_{33} (Mbar)	1.78	1.75	1.725	
C_{44} (Mbar)	0.27	0.36	0.363	

Recognizing that a single potential may have unphysical aspects, it is important to identify which results are dependent upon the potential, and which are independent of the potential. To this end, we have also used a Finnis–Sinclair (FS) potential [12] for Zr developed by Ackland *et al* [13]. These potentials are similar to the EAM model, with the same drawbacks. This particular potential more accurately reproduces the experimental c/a ratio than the Zr EAM

Table 2. Basal-plane, stacking-fault energies and $\boldsymbol{p}_1$–$\boldsymbol{p}_2$ separations from atomistic simulations and from estimations from linear anisotropic elastic theory. Separations are given in units of the lattice spacing a_0. For comparison, *ab initio* calculations of stacking-fault energies for Mg [19] and Zr [20] have been included.

				ab initio	
	Zr EAM	Zr FS	lj56	Zr	Mg
Stacking-fault energy (mJ m^{-2})	3.3	50	0.138	102	44
Simulated $\boldsymbol{p}_1$–$\boldsymbol{p}_2$ separation (a_0)	12	5	5	—	—
Estimated $\boldsymbol{p}_1$–$\boldsymbol{p}_2$ separation (a_0)	167	12	—	5.9	5.3

potential described above. It also has a significantly higher (and more accurate) basal-plane, stacking-fault, energy. Most of the figures in this paper were generated using this potential.

In addition, we have also repeated some of the calculations using the lj56 potential [14]. This was chosen primarily to compare with the results of Minonishi *et al* [4, 5] and Liang and Bacon [15] who performed simulations of the same dislocation considered here using the lj56 potential. Previous work differs from ours primarily in the size of the simulation. As we shall show below, this directly affects the observed dislocation core structures.

The initial conditions were generated as in previous studies [4, 6]. A displacement field was calculated from anisotropic elasticity [16, 17], corresponding to a perfect dislocation with a Burgers vector $\boldsymbol{b} = \langle \boldsymbol{c} + \boldsymbol{a} \rangle$. This displacement field was imposed on a perfect lattice, with a number of different starting positions of the dislocation. From this structure, a cylindrical region was chosen, including all atoms within a distance R of the dislocation line. In all of these studies, we set R to either 120 Å or 180 Å, to check the effects of the system size. No differences were found between these different simulation sizes. The repeat distance along the line of the dislocation was chosen to be four times $[1\bar{1}00]$. Thus, the simulation is essentially two dimensional.

Once this region was selected, an outer 'frozen zone' was selected. In this outer zone, the atoms were fixed during the simulations. This zone was chosen to be all of the atoms within 10 Å from the outer surface of the simulation, somewhat larger than the interaction distance for the EAM or FS potentials. The size of the simulation region plays an important role in determining the core structure. As shown later, a small simulation region can prevent the core from splitting, and hence, may produce a stable undissociated core. In comparable simulations, Minonishi *et al* [5] used a total of approximately 8000 atoms, and Liang and Bacon [6] used a total of 9000 atoms. However, the number of atoms in the inner 'relaxable' region was about 1000 in both papers, which may not be large enough to give the splitting of dislocations. In order to overcome this limitation, we have performed large-scale simulations, using approximately 97 000 atoms, of which about 86 000 atoms are free to relax, and have thus minimized the size effects.

For the lj56 potential, we chose a rectangular box shape, with an active region ranging from $-40d$ to $40d$ in the x-direction (along the Burgers vector) and from $-35d$ to $35d$ in the y-direction. This left a total of 50 000 atoms free to move. We chose this size to explore size effects (see discussion), in order to compare with the lj56 simulations of Minonishi *et al* [4, 5].

After generating the initial coordinates, we used a classical molecular dynamics (MD) code to perform a rapid quench of the system, in order to find a local minimum in the energy. This was performed using the 'dynamic' technique [18]. For the EAM model, the atomic relaxation was also carried out independently using a simple steepest-descent technique, producing nearly identical results. Following this, we performed MD simulations of the system at $T = 500$ K for

approximately 40 ps, in order to 'anneal' the structure. Finally, another quench was performed. In all cases, the dislocation structure after the final quench was qualitatively the same as after the initial quench. Based upon these results, we only used the dynamic technique for atomic relaxation, for the FS and the lj56 systems.

3. Non-planar dislocation core structures

To fully explore the possible dislocation structures, we tried a series of 12 initial conditions. For each potential, we found that certain initial conditions resulted in structures corresponding to a type II dislocation observed in previous simulations [4, 6]. In the remaining cases, a new structure, which we have denoted as a type III dissociation, was observed. We did not observe an undissociated (type I) structure after the initial quench. We note that there did not seem to be any correlation between the initial conditions that produced the type III dislocations in one potential, and those that produced these configurations in other potentials.

As the type II structure has been thoroughly studied [4, 6], we will only describe the type III structure. The core structure for the FS potential is shown in figure 2. The most interesting feature is a large distorted region extending to the left and upwards from the dislocation. This contains a small area that is twinned on the $(11\bar{2}1)$ planes, as indicated by the traces on the figure. This is very similar to the twinned regions near type I cores seen in [4, 6]. However, there is an important difference: the dislocation dissociates according to

$$\boldsymbol{b} \rightarrow \tfrac{1}{3}[10\bar{1}0] + \tfrac{1}{3}[01\bar{1}\bar{3}] \tag{2}$$

with the first of these terminating a basal-plane stacking fault. This non-planar splitting is a sessile configuration, unlikely to slip on the $(11\bar{2}2)$ plane.

We will denote this dissociation by

$$\boldsymbol{b} \rightarrow \boldsymbol{p}_1 + \boldsymbol{p}_2 \tag{3}$$

with the identification

$$\boldsymbol{p}_1 \equiv \tfrac{1}{3}[10\bar{1}0] \tag{4}$$

$$\boldsymbol{p}_2 \equiv \tfrac{1}{3}[01\bar{1}\bar{3}]. \tag{5}$$

The partial $\boldsymbol{p}_1$ also occurs in the usual splitting of the $\langle \boldsymbol{a}\rangle$ Burgers vector,

$$\tfrac{1}{3}[11\bar{2}0] \rightarrow \tfrac{1}{3}[10\bar{1}0] + \tfrac{1}{3}[01\bar{1}0]. \tag{6}$$

To identify the partial $\boldsymbol{p}_1$, we examined the relative displacements around the dislocation. In figure 3, the relative displacement of the basal planes immediately above and below the partial $\boldsymbol{p}_1$ is shown. There are relative motions both parallel and perpendicular to the line, indicating the mixed nature of the $\boldsymbol{p}_1$ partial dislocation. The edge component is dominant, with the magnitude of the screw component smaller by a factor of $\sqrt{3}$. This indicates that the Burgers vector is at an angle of 60° to the line of the dislocation. This, combined with the calculated total magnitude of the Burgers vector, is consistent with the dissociation given in (2).

The separation of the partials along the basal plane (equation (2)) is significant, approximately 22 Å for the FS potential. The separations for each potential are given in table 2. From this table, we see that there is a correlation between the basal-plane, stacking-fault energies and the separation: a larger value of the stacking-fault energy reduces the separation. However, the separation is significantly less than what would be predicted from anisotropic linear elasticity theory, based upon (2).

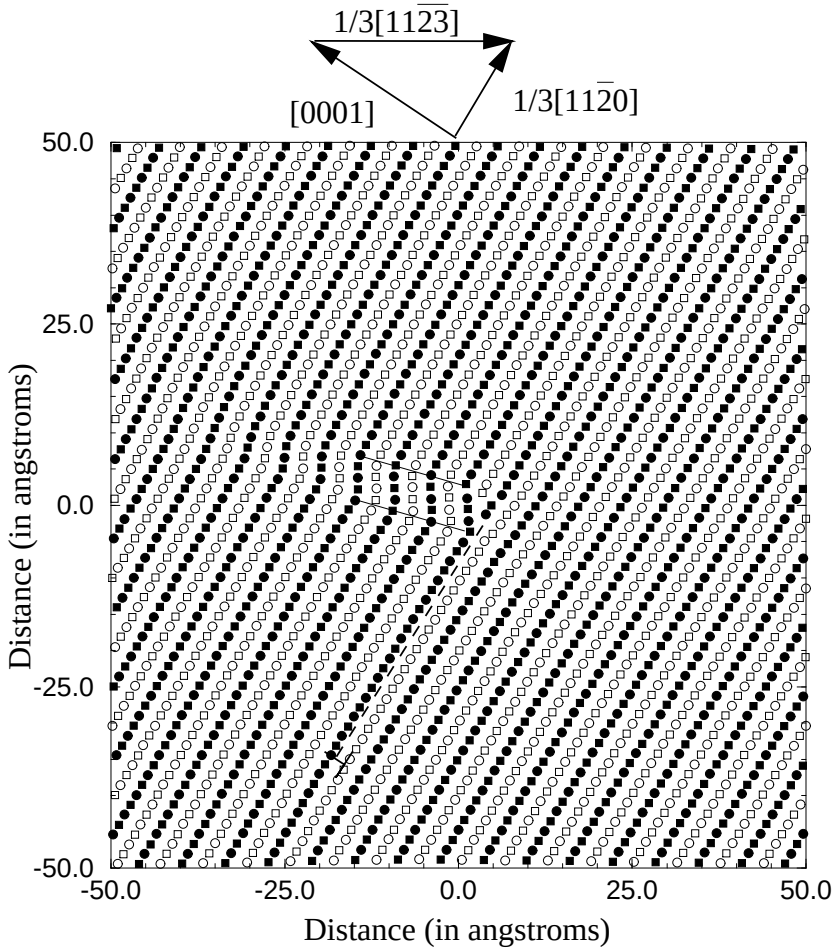


Figure 2. The structure of the type III dislocation. The dislocation has split according to $b \rightarrow \frac{1}{3}[10\bar{1}0] + \frac{1}{3}[01\bar{1}3]$, with the partials separated by a basal-plane stacking fault.

Type III dissociations found for the EAM and lj56 potentials were similar. The separations for both the FS and EAM potentials are shown in table 3, along with the stacking-fault energy for each potential. The stacking-fault energy is particularly low for the EAM potential, producing the largest splitting. To provide a comparison, we note that the *ab initio* calculations of the stacking-fault energy for Mg is 44 mJ m^{-2} [19] and 102 mJ m^{-2} for Zr [20]. Bond-order potential calculations for Ti produce stacking-fault energies similar to the *ab initio* Zr results [11]. Although the empirical potentials generally produce lower stacking-fault energies than the more accurate *ab initio* results, the separation is significant in all cases.

Table 3 shows a qualitative correlation between the basal-plane, stacking-fault energy Γ_{b} and the $\boldsymbol{p}_1 - \boldsymbol{p}_2$ splitting w_{12}. Including only the interaction between the $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ dislocations is overly simplistic for the type III dislocation core, because there are additional interactions with the twin nucleus. While difficult to account for in an accurate manner, we present a schematic approach in figure 1(*b*). In this figure, we have added twinning dislocations at either side of the twin nucleus, sufficient to produce a four-layer twin. For the $\frac{1}{3}(11\bar{2}1)[\bar{1}\bar{1}26]$ tension twin system, twinning partials have Burgers vectors of $\boldsymbol{b}_{\mathrm{t}} = (e/2)\eta$, where $\eta = \frac{1}{3}[\bar{1}\bar{1}26]$ and

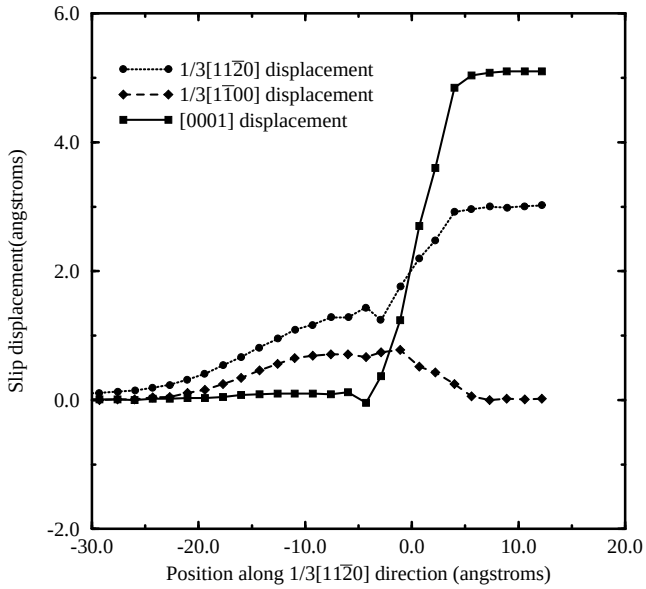


Figure 3. This figure shows the relative displacement of (0001) basal planes passing around the dislocation in the type III dissociation. From this figure, we see that the partial $\boldsymbol{p}_1$ has both a $[11\bar{2}0]$ edge components, and a $[1\bar{1}00]$ screw component. The Burgers vector of the partial $\boldsymbol{p}_1$ forms an angle of 60° with the line of the dislocation.

Table 3. Geometry of the type III dislocation core dissociations, along with basal-plane, stacking-fault energies and $(11\bar{2}1)$ twin boundary energies derived from the separations. The widths w_{12} and w_{23} are the $\boldsymbol{p}_1$–$\boldsymbol{p}_2$ separation and the twin nucleus width, respectively. For definitions of θ and ϕ, see figure 1(*b*).

	w_{12} (nm)	w_{23} (nm)	ϕ (degrees)	θ (degrees)	Γ (mJ m^{-2}) (0001)	Γ (mJ m^{-2}) $(11\bar{2}1)$
EAM	4.2	3.6	50.6	56.5	26	99
FS	2.2	4.2	76.4	30.7	140	47

$e = 2/(4\gamma^2+1) \approx \frac{1}{6}$ for Zr with $\gamma \equiv c/a = 1.63$ [21]. Therefore, the total strength of four twinning partials is $\boldsymbol{p}_3 = [\bar{1}\bar{1}26]/9$. Consequently, the Burgers vector of the sessile partial located at the corner of this non-planar dissociation is $\boldsymbol{p}_2 = [14\bar{5}3]/9$. Thus, the splitting given in (2) has been modified to

$$\boldsymbol{b} \rightarrow \tfrac{1}{3}[10\bar{1}0] + \tfrac{1}{9}[14\bar{5}3] + \tfrac{1}{9}[\bar{1}\bar{1}26]. \tag{7}$$

The widths of the basal stacking fault, w_{12}, and the tension twin, w_{23}, are given in table 3. The angle between the twin habit plane and the basal plane is $\Phi = \tan^{-1}(2\gamma) = 73.0°$. The other two angles, θ and ϕ, defined in figure 1(*b*) are also given in table 3. Using anisotropic elasticity theory, we calculated the forces on the dislocations. Both F_x and F_y force components (along the $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ direction and perpendicular to the $(11\bar{2}2)$ plane, respectively) of the pair-wise interaction were calculated first. The net results of F_x in the fault planes were determined at $\boldsymbol{p}_1$ and $\boldsymbol{p}_3$. By requiring that the total glide forces balance to zero, we calculate the basal-plane, stacking-fault energy, Γ_{b}, and twice the $(11\bar{2}1)$ twin boundary energy, $2\Gamma_{\mathrm{t}}$. These results with the two sets of elastic constants are listed in table 3.

As compared to $\Gamma_b = 3.3$ mJ m^{-2} and $\Gamma_t = 285$ mJ m^{-2} [22] calculated using the EAM potential, the basal-plane, stacking-fault energy is higher by a factor of 7.9, while the twin boundary energy is lower by a factor of 5.8. Based upon the FS potential, we obtain $\Gamma_b = 50$ mJ m^{-2}, and Serra and Bacon [23] reported $\Gamma_t = 278$ mJ m^{-2}. As compared to these results for Γ_b and Γ_t, the values listed in table 3 are higher by a factor of 2.8 and lower by a factor of 11.8, respectively. This comparison suggests that nonlinear elastic interactions, ignored in this analysis, assist in the twin nucleation, creating a larger than expected twinned region. Simultaneously, the complementary splitting on the basal plane is suppressed.

We note that we have not observed a separation into separate $\boldsymbol{c}$ and $\langle \boldsymbol{a} \rangle$ dislocations. Within isotropic, linear elasticity theory, there should be no reason for these to interact, and we would expect that they could separate. However, this neglects two points. First, anisotropy can introduce an interaction between these dislocations. However, this can be expected to be weak. Second, they may lower their energy through an interaction of the dislocation cores. This effect is beyond the assumptions of linear elasticity theory. We have applied a strain on the basal plane, coupling to the $\langle \boldsymbol{a} \rangle$ portion of the dislocation, to attempt to separate the dislocations. We find that under reasonably strong strains (2%), this has the effect of increasing the $\langle \boldsymbol{p}_1 \rangle$ partial separation, without unbinding the dislocations.

4. Energy calculations

Given that several possible dislocation structures exist, it is important to consider the energy of the dislocations. While the energy of a single dislocation in an infinite crystal is not defined, we may compare the core energies of different core structures, to find which structure has the lowest energy.

To do this, we consider the energy contained within a distance R from the dislocation. The total energy E_T per unit length of the dislocation in the region defined by $(0, R)$ may be written as

$$E_T = E_c(r_c) + E_e(r_c, R) \tag{8}$$

where $E_c(r_c)$ is the energy in the 'core' region, where linear elasticity is not applicable. (Note that the choice of r_c also defines the value of E_c.) The elastic energy outside the core region, $E_e(r_c, R)$, is given by [24]

$$E_e = \frac{K_e}{4\pi} b^2 \ln \frac{R}{r_c} \tag{9}$$

where K_e is the energy factor determined from anisotropic elasticity theory. From the geometry of the $\langle \boldsymbol{c} + \boldsymbol{a} \rangle$ dislocation and the elastic constants given in table 1, we find the value $K_e = 0.385$ eV Å^{-3} for the FS potential.

We may also calculate the energy from the distortions of the dislocation from atomistic simulations. If the equilibrium bulk energy per atom in the pure hcp structure is E_0, then the excess energy due to the atomic displacements within a distance R of the dislocation core is

$$E_T = \frac{1}{d} \sum_i (E_i - E_0) \tag{10}$$

where E_i is the total potential energy of atom i, d is the repeat distance along the direction of the dislocation line (for this geometry, $d = \sqrt{3}a_0$) and the sum extends for atoms within a distance R of the dislocation core. At long distances, the behaviour should follow that of (9).

Figure 4 shows the variation of calculated elastic energies with the logarithm of the radius from the core, from the results for the FS potential. From this, we see that the basal plane splitting (type III) is significantly lower in energy than that of the type II dislocation.

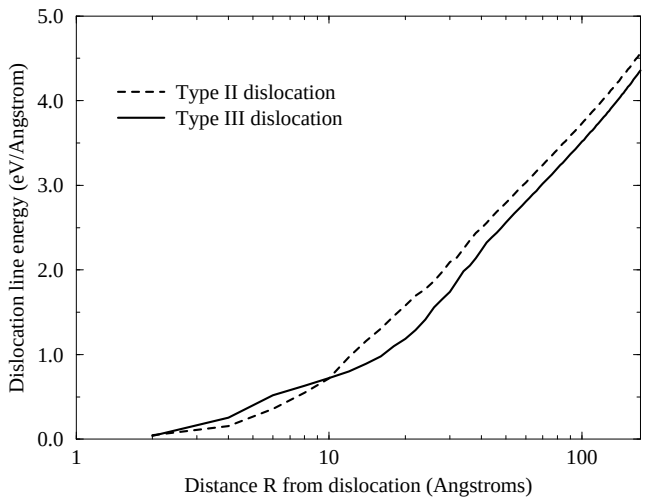


Figure 4. Energy $E_{\mathrm{T}}(R)$ per unit length of the dislocation, for the two types of dislocation found in the simulations, as a function of R. The dislocation energy is given in units of eV Å^{-1}.

A similar calculation for both EAM and the lj56 potentials again shows that the type III dislocation is lower in energy. At large distances (greater than approximately 40 Å), the graphs are essentially parallel. This indicates that for $R > 40$ Å, the energies found in the simulations behave as predicted by elasticity theory. At large distances, the energy difference between the two structures corresponds to the difference in the total energy of the dislocations.

5. Discussion

In this paper, we have used atomistic simulations to explore possible core structures of the $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ edge dislocation. We expect that this dislocation core plays an important (and complicated) role in determining the behaviour of hcp metals under c-axis compression and tension. For Mg and Be, the dominant deformation mode is the $(11\bar{2}2)$ $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ slip system; the compression twinning mode in Ti and Zr has the same elements. Therefore, it is expected that $\langle \boldsymbol{c}+\boldsymbol{a}\rangle$ dislocations may determine the non-basal deformation mode.

Our simulation techniques are similar to those performed by Minonishi *et al* [4, 5] and by Liang and Bacon [6]. However, they differ in several respects. First, in addition to simple pair potentials, we have also used more sophisticated EAM and FS potentials fitted to equilibrium properties. Second, and most importantly, we have used significantly larger system sizes. Finally, we have not only performed local minimization, but we have also annealed the dislocation cores in an effort to find the lowest-energy structures.

Unlike previous investigations [4–6] we do not observe a stable, undissociated core. Instead, we see a new core dissociation, which we label as 'type III', where the core emits a $\boldsymbol{p}_1 \equiv \frac{1}{3}[10\bar{1}0]$ partial dislocation along the basal plane. The remaining partial has a core structure containing a $(11\bar{2}1)$ twin nucleus, similar to the undissociated 'type I' core found in previous simulations [4–6]. We observe the type II dislocation found previously [5, 6], in which the dislocation splits along the $(11\bar{2}2)$ slip plane, with a stable stacking fault on the slip plane. We have found that the type III dislocation has a core energy that is lower than that of the type II, by analysing the energy distribution in our simulations.

We believe that previous simulations [4–6] were severely affected by the small system sizes. In these previous simulations, only a small region (about 1000 atoms) is allowed to relax. This prevented the type III core from forming. Having only a small region to dissociate, and having to match the boundary conditions, prevents the partial $\boldsymbol{p}_1$ from separating from the original dislocation. We have tested size effects for the type III dislocation for the EAM, FS and lj56 potentials. For these calculations, we chose a rectangular system, and varied the size of the system. For the EAM potential, which has a large splitting, the system size must be made very small ($\sim$15 Å normal to the $(11\bar{2}2)$ plane) in order to suppress the dissociation. For the FS and lj56 systems, however, the splitting for the large system is still not too large, and is suppressed rather easily. Once the system size perpendicular to the $(11\bar{2}2)$ plane is smaller than $\sim 15a_0$, then the core remains in an undissociated (type I) configuration.

The splittings are very sensitive to the stacking-fault energy, and are not likely to be accurately modelled for any given material by any of these potentials. In particular, Zr and Ti have large basal-plane, stacking-fault energies, close to 100 mJ m^{-2} [11, 20]. Basal-plane slip is therefore not usually observed in these materials. However, the stacking fault is stable [11, 20], and so a type III dissociation may occur for these materials.

We expect that the type III core will be more sessile than the type II, as the partials for the latter core will slip on the $(11\bar{2}2)$ plane. For the type III core, the partial $\boldsymbol{p}_1$ moves on the basal plane, while the $\boldsymbol{p}_2$ partial should slip on the $(11\bar{2}1)$ plane. Thus, the two partials cannot move together. We also expect that shear on the $(11\bar{2}2)$ plane will have a different effect on the type III dislocation than compression/tension along the c-axis, as the latter will not couple to the motion of the $\boldsymbol{p}_1$ partial along its easy-slip direction.

It is possible that different types of cores occur in different materials, giving rise to different behaviour under c-axis compression and tension. It is plausible that in a material where the dislocation has a type II structure, that the $(11\bar{2}2)$ slip will be the deformation mode. If the type III structure occurs, then the lack of mobility may hinder slip, and twinning may be favoured. We also note that the presence of a type III configuration may assist in the nucleation of $(11\bar{2}1)$ twins. This suggests that there may be a correlation between metals that exhibit compression twinning on the $(11\bar{2}2)$ plane, and those that have tension twinning on the $(11\bar{2}1)$ plane. This is borne out by experiments: Ti and Zr show both types of twin modes, while Be and Mg show neither [3].

In conclusion, we have observed two different core structures for the $\langle \boldsymbol{c} + \boldsymbol{a} \rangle$ edge dislocation in simulations of hcp materials. Both of these structures are dissociated, with the type III structure (not observed in previous simulations) having a sessile geometry, splitting along the basal plane. For all potentials considered here, this latter type is lower in energy.

Acknowledgments

This research used resources of the Center for Computational Sciences at Oak Ridge National Laboratory (ORNL) and the Scalable Computing Laboratory, funded by Iowa State University and Ames Laboratory. This work was supported by the High Performance Computing and Communications Initiative (HPCCI). Ames Laboratory is operated for the US Department of Energy by Iowa State University under contract no W-7405-Eng-82. The work at ORNL was sponsored by the Division of Materials Sciences, Office of Basic Energy Sciences, US Department of Energy under contract no DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

References

[1] Vitek V 1992 *Prog. Mater. Sci.* **36** 1
[2] Duesbery M S 1997 *Dislocations in Solids* vol 10, ed F R N Nabarro (Amsterdam: North-Holland) p 67
[3] Yoo M H 1981 *Metall. Trans.* A **12** 409
[4] Minonishi Y, Ishioka S, Koiwa M, Morozumi S and Yamaguchi M 1981 *Phil. Mag.* A **44** 1225
[5] Minonishi Y, Ishioka S, Koiwa M, Morozumi S and Yamaguchi M 1982 *Phil. Mag.* A **45** 835
[6] Liang M H and Bacon D J 1986 *Phil. Mag.* A **53** 181
[7] Ando S, Takashima K and Tonda H 1996 *Mater. Trans. Japan. Inst. Metals* **37** 319
[8] Zhang B L, Wang C Z, Ho K M, Turner D E and Ye Y Y 1995 *Phys. Rev. Lett.* **74** 1375
[9] Morris J R, Ye Y Y, Ho K M, Chan C T and Yoo M H 1995 *Phil. Mag.* A **72** 751
[10] Morris J R, Scharff J, Ho K M, Ye Y Y and Yoo M H 1997 *Phil. Mag.* A **76** 1065
[11] Girshick A, Pettifor D G and Vitek V 1998 *Phil. Mag.* A **77** 999
[12] Finnis M W and Sinclair J E 1984 *Phil. Mag.* A **50** 45
[13] Ackland G J, Wooding S J and Bacon D J 1995 *Phil. Mag.* A **71** 553
[14] Bacon D J and Martin J W 1981 *Phil. Mag.* A **43** 883
[15] Liang M H and Bacon D J 1986 *Phil. Mag.* A **53** 205
[16] Yoo M H and Loh B T M 1970 *J. Appl. Phys.* **41** 2805
[17] Stroh A N 1958 *Phil. Mag.* A **3** 625
[18] Gehlen P C 1970 *J. Appl. Phys.* **41** 5165
[19] Chetty N and Weinert M 1997 *Phys. Rev.* B **56** 10 844
[20] Morris J R and Ye Y Y 1999 in preparation
[21] Yoo M H 1969 *Trans. AIME* **245** 2051
[22] Morris J R, Ye Y Y, Ho K M, Chan C T and Yoo M H 1994 *Twinning in Advanced Materials* ed M H Yoo and M Wuttig (Warrendale, PA: Minerals, Metals and Materials Society) p 43
[23] Serra A and Bacon D J 1993 *Mater. Sci. Forum* **126–128** 69
[24] Hirth J P and Lothe J 1968 *Theory of Dislocations* (New York: McGraw-Hill)
[25] Brewer L 1975 *Lawrence Berkeley Laboratory, Report No 3720* unpublished
[26] Goldak J, Lloyd L T and Barrett C S 1966 *Phys. Rev.* **144** 474
[27] Fisher E S and Renken C J 1964 *Phys. Rev.* A **135** A482

by many factors such as powder size and their distributions, mole fraction of substances, material purity, particle surface condition, particle shape, mixing condition, and deformation of powders during compaction [14]. Moreover, a formation of low-thermal conductivity intermediate phases during reaction synthesis may also affect the thermal conductivity of compacts.

In this study, the effects of heat treatment, test temperature, material density, and formation of intermediate phases on the thermal conductivity of elemental powder compacts with a nominal composition of 75Ni–25Al in mole fraction are investigated using a laser-flash method.

2. Experimental procedures

A powder mixture which contains 75Ni–25Al in mole fraction was prepared from these elemental powders by milling with Al_2O_3 balls. Ni and Al elemental powders used previously [4,15] had an average particle size of approximately 5 and 15 μm, respectively.

In order to investigate the effect of relative density and temperature on the thermal conductivity, two groups of powder compacts were prepared. Disc specimens of 25 mm in diameter and 2 mm in thickness were shaped by applying a uni-axial stress of about 140 or 350 MPa. Loading time of a few minutes was used in order to achieve a preset compact density. The powder samples were then compacted to a density, relative to the bulk Ni_3Al, of 60% at an applied stress of 140 MPa, and 72% at 350 MPa. These values will be referred to as "nominal" relative density for each group of compact samples.

The effect of heat treatment on the thermal conductivity of compact samples with a relative density of 60%, produced by 140 MPa, was studied. In the previous work, powder compacts were heated to a temperature of 620°C, just below the melting point of Al, for several minutes for SHS initiation [4]. In the present study heat treatments were conducted at 400°C in order to avoid SHS for 48 or 1000 h in an evacuated quartz tube.

The microstructure of heat-treated specimen was observed using a scanning electron microscopy (SEM) operated at 20 kV and compositional analysis was conducted by an energy dispersive X-ray spectroscopy (EDS) equipped on the SEM. Thermal conductivities of these specimens were measured by a laser-flash method. Details of the method were described elsewhere [16].

3. Results and discussion

3.1. Thermal conductivity of powder compacts

Fig. 1 compares the temperature dependence of the thermal conductivity of powder compacts with 72 or 60% relative density with that of pure Ni and Al metals obtained from the literature [10]. The thermal conductivity for each group of compact samples is lower than those of the metals, however, their temperature dependences are quite similar. The temperature dependence of the thermal conductivity of the compacts can thus be attributed to those of elemental materials. This similarity suggests that the thermal conductivity at high temperature region can be estimated from the data at ambient temperature.

The effect of relative density on the thermal conductivity of powder compacts without heat treatment is shown as open symbols in Fig. 2. As can be expected, the compact with a higher density has a higher thermal conductivity. In this figure, several curves are calculated based on the equation proposed by Russell [17] to predict the thermal conductivity of pure Ni or Al powder compacts under different assumptions. In one case where voids with the same size are separated by solid walls, referred as the isolated-void model, the thermal conductivity of such compact can be given as,

$$\frac{k_c}{k_s} = \frac{P^{2/3} + \frac{k_s}{k_a}(1-P)}{P^{2/3} - P + \frac{k_s}{k_a}\left(1 - P^{2/3} + P\right)} \tag{1}$$

where k_c is thermal conductivity of a compact, k_s for solid, k_a for air space, and P for porosity fraction. The

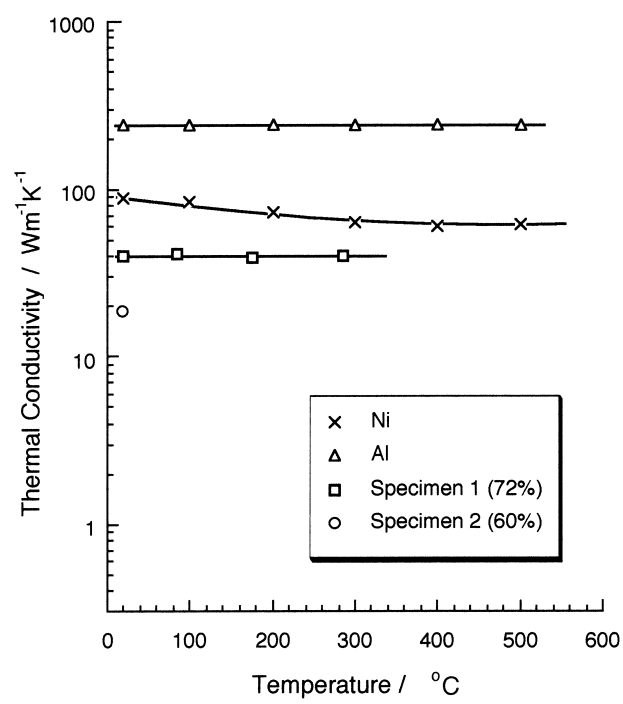


Fig. 1. The temperature dependence of the thermal conductivity of a 3Ni–1Al mixed powder compact (specimen 1) with 72% relative density, Ni and Al elemental bulk specimens from the literature [10]. The room temperature thermal conductivity of a powder compact (specimen 2) with 60% relative density (without heat treatment) is also plotted.

thermal conductivity of air is used for k_a. Both the calculated thermal conductivities of pure Ni and Al powder compacts are shown in Fig. 2 for comparison.

In another case where solid particles of the same size are surrounded by air or vacuum, a reverse of the isolated-void model, referred to as the isolated-particle model, the conductivity is still given in the same type of equation, except that P is replaced with $(1-P)$ and k_s is changed with k_a. In this case, there is no difference in the calculated thermal conductivity of Ni and Al compacts with the relative density lower than 98% as shown in Fig. 2. Note that the assumption of the isolated-particle is not applicable to 3Ni–1Al powder compacts with much higher thermal conductivities measured experimentally.

The values measured experimentally are rather lower than those predicted by the isolated-void model using the thermal conductivity of bulk Ni metal with "nominal" relative densities. In view of the volume fraction and particle sizes used in the present study, it is reasonable to assume that large Al powder particles are surrounded by small Ni powder particles. Al oxides are expected to form on Al powder surfaces, which will serve as a thermal barriers. The 3Ni–1Al powder compacts can be thus considered as Ni network separated by "Al" voids (Al oxide coated Al particles). This configuration of powder compacts will effectively reduce the

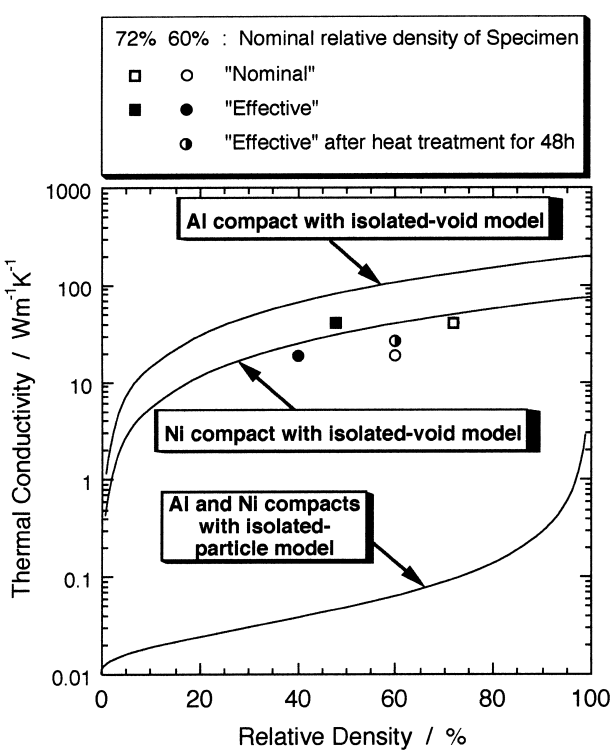


Fig. 2. Effect of relative density on the thermal conductivity of compacts with and without heat treatment. In order to evaluate the Ni powder network as a heat conducting path, "effective" relative densities of both powder mixtures are calculated based on the amount of Ni particles in each compact.

relative density from 60% (normal density) to 40% (effective density), and from 72 to 48%. As indicated in Fig. 2, the solid data points based on the effective density agree well with the theoretical prediction for Ni powder compact. The higher thermal conductivity of the compact sample with a higher relative density than the theoretical value may be achieved by considering that Al_2O_3 layer on Al particles are broken mechanically during pressing.

3.2. Formation of intermediate phases and their effect on thermal conductivity

Before ignition or during combustion reaction, intermediate phases are formed, and these phases may affect the thermal conductivity of powder compacts. To simulate microstructural evolution produced by annealing at 620°C for a few minutes, typical duration prior to ignition, specimens were heat-treated for 48 or 1000 h at 400°C. These heat treatment durations will result in the formation of Al_3Ni and Ni_2Al_3 in Ni–Al diffusion couple [7,9,12]. Fig. 3 shows the back scattering electron-image of the compact by the SEM, which was heat-treated for 1000 h at 400°C, corresponding to heat treatment for about 300 s at 620°C in terms of bulk diffusion rate [7,9,12]. Observed in the figure are only intermetallics such as Ni_2Al_3, NiAl and Ni_3Al, and Ni-solid solution with no residual Al powder particles. Through the analysis by EDS, it is found that darker regions contain a higher amount of Al. It is noteworthy to mention that intermetallics are formed not only between former Ni and Al elemental particles but also between former Ni particles. As shown in Fig. 3, phase contour lines indicating the boundaries of intermediate phases are visible in concentric configurations in the interior of the former Ni particles. In view that no liquid phase formation in the specimen, this configuration can be attributed to a transport of Al across Ni particle boundaries by fast diffusion on free and/or contacting surfaces, followed by slower bulk diffusion into former Ni particles.

Fig. 4 shows the effect of heat treatment duration on the thermal conductivity. The thermal conductivity of the compact specimen without heat treatment is 18.6 W/m/K, and it increases to 26.4 W/m/K after heat treatment for 48 h at 400°C. This increase is considered to be caused by the destruction of the oxide layer on Al particles and the joining of Al–Ni particles by interdiffusion. If all the oxide layers on Al particles are broken, Al particles should contribute to the thermal conduction thoroughly and not "effective" but "nominal" relative density should be taken. The Al particle, which has a higher thermal conductivity than the Ni particle, is expected to raise the thermal conductivity of the compact to the value higher than the theoretical one predicted by the Ni-powder compact curve shown in Fig. 2. The

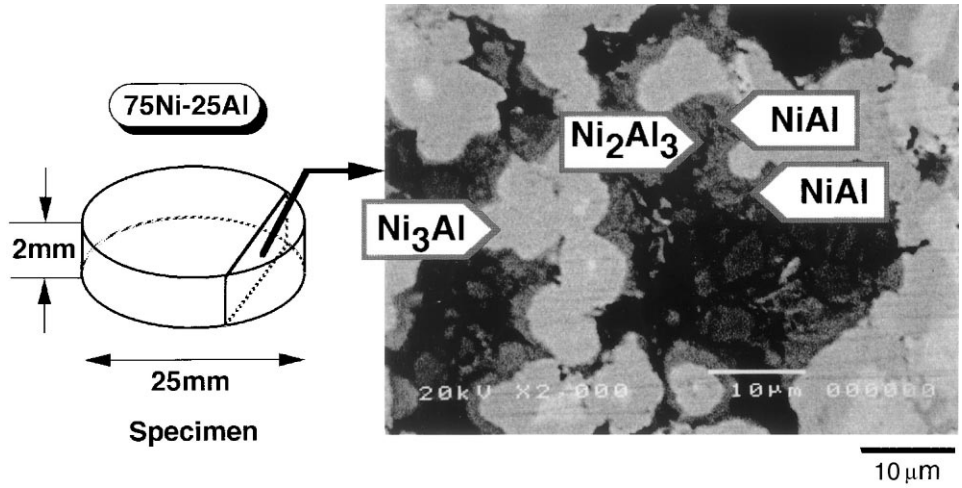


Fig. 3. The back scattering electron-image of compact by SEM, which was heat-treated for 1000 h at 400°C.

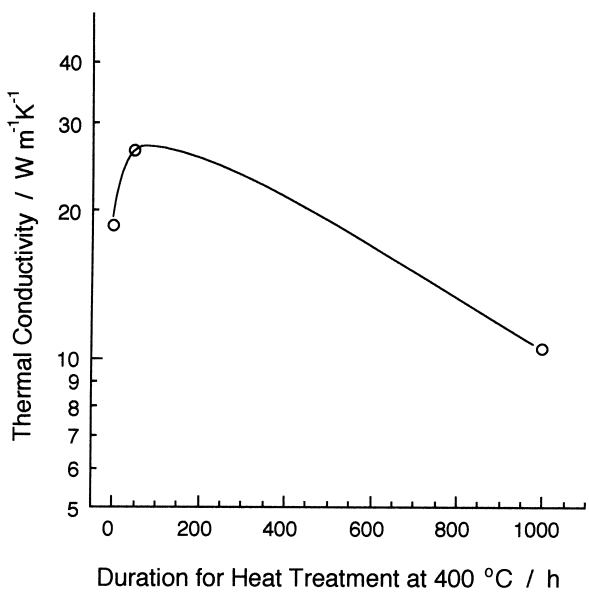


Fig. 4. Effect of heat treatment duration at 400°C on the thermal conductivity of Ni–Al powder mixture compact with a relative green density of 60%.

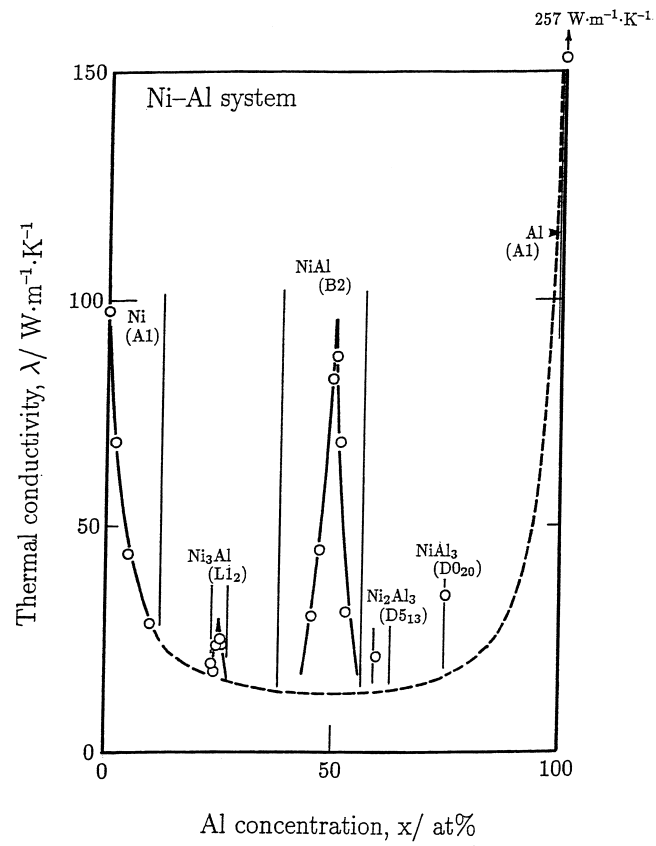


Fig. 5. Thermal conductivities of various phases in the Ni–Al alloy system. The U-shaped line represents the conductivity of disordered

thermal conductivity after heat treatment for 48 h (a half-filled symbol) becomes close to but is still lower than the predicted value from the relative density of 60%, as shown in Fig. 2. This comparison suggests that not all oxide layers on Al and Ni particles are broken at this moment.

Thermal conductivity decreases to 10.5 W/m/K after heat treatment for 1000 h at 400°C, which must be associated with the formation of intermetallics and Ni-solid solution layers on particle surfaces as shown in Fig. 3. Fig. 5 shows that the thermal conductivity of a solid solution alloy is much lower than that of pure elements [16,18]. It is found that thermal conductivities of intermetallic phases are rather higher than interpolated values for disordered states with corresponding compositions; however, they are still lower than those of pure elements that compose the alloys. These intermetallic phases and Ni-solid solution formed between and inside of particles, govern the thermal conductivity of the long-term heat-treated compact. The formation of those phases through the interdiffusion process is also a key nature of the reaction synthesis. This implies that the thermal

conductivity of compact changes during reaction synthesis and the measurement of the thermal conductivity of initial mixed powder compact itself is not enough for precise modeling of the reaction synthesis, unless all the effects of the formation of intermetallic phases during heat treatment or pre-ignition duration are taken into account.

As mentioned above, intermediate phases and Ni-solid solution are formed not only between former Ni and Al particles but also between former Ni particles, because of fast diffusion on free and/or contacting surfaces. Nevertheless, the data on such diffusion processes are quite limited [10]. As the activation energies for diffusion on free and/or contacting surface are generally smaller than that for bulk diffusion, the effect of such diffusion on the thermal conductivity and microstructural evolution during the reaction synthesis becomes dominant over a bulk diffusion at low temperatures. It also suggests that such diffusion processes will play a key role in simulating the reaction synthesis.

4. Conclusions

Thermal conductivity of 3Ni–1Al elemental powder compacts was investigated as a function of temperature, compact density and heat treatment. The following conclusions can be drawn from this work.

1. Thermal conductivity of the compacts is much lower than those of elemental bulk specimens, although temperature dependences are quite similar. Highly densified specimen virtually has higher thermal conductivity.
2. Not only between former Ni and Al elemental particles but also between former Ni particles, various intermetallic phases are formed after long-term heat-treatment. The contour lines are observed in former-Ni particles as a result of the transport of Al into Ni particle boundaries by fast diffusion on a free and/or contacting surface followed by slower bulk diffusion.
3. Thermal conductivity of a specimen after heat treatment for 48 h at 400°C is 40% higher than that without heat treatment. This increment is caused by the contribution of Al particles to the thermal conduction as a result of destruction of surface oxide layers on Al particles and joining of them together.
4. Thermal conductivity after heat treatment for 1000 h at 400°C decreases to 10.5 W/m/K (being about 40% of that after heat treatment for 48 h). This decrease is caused by the formation of layers of low thermal conductivity materials such as intermetallics and Ni-solid solution between and inside of former Ni particles.

Acknowledgements

The authors would like to thank Mr. Elmer H. Lee at ORNL and Mr. Kenji Ohkubo at Hokkaido Univ. for their technical assistance. This research was partially supported by the Division of Matrials Science, U.S. Department of Energy under contract DE-ACO5-86OR22464 with Lockheed Martin Energy Research Corporation.

References

[1] Bose A, Rabin BH, German RM. Powder Metall Int 1988;20(3):25.
[2] German RM, Bose A, Stoloff NS. High temperature ordered intermetallic alloys III. MRS Symp Proc 1989;133:403.
[3] Alman DE, Stoloff NS. High temperature ordered intermetallic alloys IV. MRS Symp Proc 1991;213:989.
[4] Miura S, Liu CT. Intermetallics 1994;2:297.
[5] Bhattcharya AK. J Mat Sci 1992;27:3050.
[6] Lakshmikantha MG, Bhattacharya A, Sekhar JA. Metallurgical Transactions A 1992;23A:23.
[7] Janssen MMP, Rieck GD. Trans Met Soc AIME 1967;239:1372.
[8] Hultgren R et al. Selected values of the thermodynamic properties of binary alloys. Metals Park, Ohio, USA: ASM, 1973.
[9] Sarkhel AK, Seigle LL. Met Trans A 1982;13A:1313.
[10] Brandes EA. Smithells metals reference book. 6th ed. London, UK: Butterworth & Co, 1983.
[11] Massalski TB et al. Binary alloy phase diagram. 2nd. ed. Materials Park, Ohio, USA: ASM International, 1990.
[12] Jung SB, Minamino Y, Yamane T, Saji S. J Mat Sci Letter 1993;12:1384.
[13] Kubaschewski O, Alcock CB, Spencer PJ. Materials Thermochemistry. 6th ed. Oxford, UK: Pergamon Press, 1993.
[14] Dawson DM, Briggs A. J Mater Sci 1981;16:3346.
[15] Nishimura C, Liu CT. Scripta Metall et Mater 1992;26:381.
[16] Terada Y, Mohri T, Suzuki T. High temperature ordered intermetallic alloys VI. MRS Symp Proc 1995;364:201.
[17] Russell HW. J Amer Ceram Soc 1935;18:1.
[18] Suzuki T, Terada Y, Mohri T In: Deevi SC et al, editors. Nickel and Iron Aluminides: Processing, Properties and Applications. Materials Park, Ohio, USA: ASM International, 1997. p. 223.

ELSEVIER

Intermetallics 8 (2000) 885–889

Intermetallics

www.elsevier.com/locate/intermet

Stoichiometry and mechanical properties of Mo_3Si

I. Rosales, J.H. Schneibel *

Metals and Ceramics Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831-6115, USA

Received 18 September 1999; accepted 23 March 2000

Abstract

The A15 phase Mo_3Si is an important constituent of a new class of silicides based on Mo–Si–B. In this research it will be shown that, contrary to published results, single-phase Mo_3Si is slightly off-stoichiometric. In addition, it remains single phase in a small composition range. Its room temperature fracture toughness is on the order of 3 MPa $m^{1/2}$. The compressive strength at 1400°C in argon decreases with decreasing strain rate and increasing Si concentration. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Molybdenum silicides; B. Fracture toughness; B. Yield stress

1. Introduction

Silicide intermetallics based on Mo, Si and B are currently of interest as potentially oxidation resistant, high temperature structural alloys [1–9]. One type consists of Mo_5SiB_2, Mo_3Si, and a toughening α-Mo phase [5–9], another consists of Mo_3Si, Mo_5Si_3, and Mo_5SiB_2 [1–4]. The phase Mo_5Si_3 has been studied in considerable detail. The elastic constants of Mo_5Si_3 determined by Fu et al. [10] by ab-initio calculations are in good agreement with Chu et al.'s experiments [11]. Meyer et al. [1] investigated the high temperature creep properties of powder metallurgical Mo_5Si_3 at compressive stresses of 120–180 MPa. Some research has also been performed for Mo_3Si. Fig. 1 shows a section of the Mo–Si binary phase diagram. Mo_3Si forms by a peritectic reaction at 2025°C. Christensen [12] prepared single crystals of Mo_3Si and investigated them by neutron scattering. Consistent with the phase diagram in Fig. 1, Christensen found Mo_3 Si to be a line compound. Meyer et al. [1] found a high density of dislocations in the Mo_3Si phase in a three-phase Mo_5Si_3–Mo_5SiB_2–Mo_3Si alloy after creep testing at temperatures near 1300°C. The slip system was consistent with (001) < 100 >. Misra et al. investigated the microstructure and mechanical properties of a Mo_3Si–Mo_5Si_3 eutectic composite [13]. Using an indentation cracking technique, they determined the room temperature fracture toughness to be 1–1.5 MPa $m^{1/2}$. Raj [14] investigated the oxidation and mechanical properties of a ternary $(Cr,Mo)_3Si$ compound. Using a single-edge precracked beam technique, he found room temperature fracture toughness values on the order of 2–3 MPa $m^{1/2}$. The purpose of the present work is to add to our knowledge of binary Mo_3Si. We will verify its single-phase composition, determine its room temperature fracture toughness and hardness, and assess its high temperature strength.

2. Experimental procedure

Alloys with nominal silicon concentrations ranging from 22 to 28 at.% were prepared by arc-melting of nominally pure elements in a partial pressure of argon (99.999% purity). Prior to melting, a Zr getter was melted to remove residual oxygen from the pumping chamber. The alloys were drop-cast into water-cooled copper molds with a diameter of 12.5 mm. Table 1 shows the nominal compositions and the compositions calculated by assuming that the weight loss during melting and casting were due to evaporation of either Si or Mo (since Si has a higher vapor pressure than Mo, it is assumed that the weight loss is due to loss of Si). A comparison with the nominal compositions shows that the actual Si concentrations of the cast specimens were lower than the nominal ones by typically 0.1–0.9 at.% (all discussions in this report will refer to nominal composition). Inductively coupled plasma (ICP) analysis by a commercial vendor resulted in substantially lower Si contents. As

*Corresponding author. Tel.: +1-865-576-4644; fax: +1-865-574-7659
E-mail address: schneibel@ornl.gov (J.H. Schneibel)

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(00)00058-3

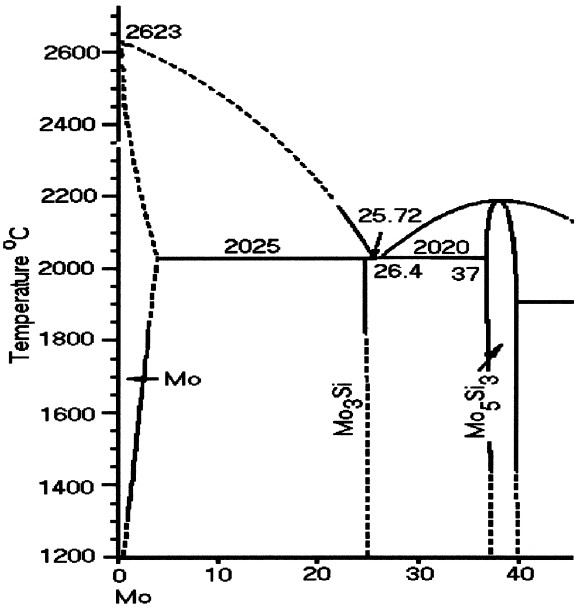


Fig. 1. A section of the binary Mo–Si phase diagram showing the compositions of the Mo_3Si and Mo_5Si_3 phases [14].

pointed out by Huebsch et al. [16], the calculations from the weight losses were the best option for determining the compositions. The specimens were annealed in a vacuum of 10^{-4} Pa for 24 h at 1600°C, and then cooled using a cooling rate of 2.5°C s^{-1} between 1600 and 1000°C. After metallographic polishing, the specimens were etched with Murakami's reagent for 1–2 s. The etched specimens were observed in an optical microscope as well as a scanning electron microscope equipped with an energy dispersive spectroscopy (EDS) system for determining the Mo:Si ratios in the alloys. The lattice parameters were determined by X-ray diffraction of powders with a size of <45 μm and an internal silicon standard. The experimental accuracy of the lattice parameter measurements was estimated to be 3×10^{-4} Å. Pycnometric densities of coarse 40–80 mesh powder ground from the castings were determined in a He pycnometer. Microhardnesses were measured with a Buehler microhardness tester at a load of 500 g and a holding time of 15 s. Standard deviations were calculated from 10 measurements each. Fracture toughnesses were measured at room temperature using three-point flexure bars with cross-sections of 3×4 mm, a span of 10 mm, and an electro discharge machined single-edge notch with a depth of 2 mm. The fracture toughnesses were calculated from the maximum load according to the corresponding equation in ASTM standard E399-90 [17]. Compression specimens with dimensions of 2×2×4 mm were also machined. They were compressed in an Instron 4501 testing machine at 1400°C in argon. In order to know the dependence of the yield stress on the strain rate, different initial strain rates of 10^{-3}, 1.8×10^{-4} and 10^{-5} s^{-1} were chosen.

3. Results and discussion

3.1. Microstructure

Fig. 2 shows optical micrographs of Mo_3Si specimens with different Si concentrations. As expected from the phase diagram, the specimen with 22 at.% Si contains α-Mo precipitates in a Mo_3Si matrix (Fig. 3a). These precipitates were identified by EDS as well as X-ray powder diffraction. They are at the grain boundaries suggesting that they controlled grain growth during the heat treatment. The specimen with 25 at.% Si (Fig. 2c) contains small particles of Mo_5Si_3 in a Mo_3Si matrix. Its grain size (not visible in Fig. 2c) was much larger than that of Mo–22 at.% Si. The specimen with 24 at.% Si (Fig. 2b) has also a large grain size because it does not contain second phases impeding its grain growth. Consistent with X-ray analysis, this alloy is single phase. Calculations of the single phase composition of "Mo_3Si" were also made based on the measured precipitate volume fractions in Mo–22 at.% Si and Mo–25 at.% Si. In both cases a single phase composition near 24 at.% Si was predicted from the overall compositions and the measured volume fractions of α-Mo or Mo_5Si_3, respectively. In disagreement with the published Mo–Si phase diagram, the results of this study suggest that the single phase composition of Mo_3Si after cooling from 1600°C is not Mo–25 at.% Si, but near Mo–24 at.% Si.

3.2. Lattice parameter

Fig. 3 shows the room temperature lattice parameter as a function of silicon concentration. If Mo_3Si were a true

Table 1
Nominal compositions of Mo_3Si and the compositions calculated by assuming the weight loss during melting and casting to be due either Si or Mo evaportation

Nominal Si concentration (at.%)	22.0	23.0	23.5	24.0	24.5	25.0	28.0
Initial mass of Si (g)	4.58	3.22	3.30	5.92	3.47	2.67	3.07
Initial mass of Mo (g)	55.42	36.78	36.70	64.08	36.53	27.33	26.93
Weight loss after melting (g)	0.25	0.068	0.105	0.04	0.036	0.06	0.002
Concentration of Si calculated from weight loss (at.% due to Si evaporation)	21.06	22.64	22.92	23.86	24.30	24.59	28.00
Concentration of Si calculated from weight loss (at% due to Mo evaporation)	22.09	23.05	23.55	23.88	24.52	25.07	28.03

line compound, one would not expect a change in the lattice parameter as the Si concentration crosses the single phase composition. Fig. 3 shows, however, a lattice parameter change between 23.5 and 24 at.% Si This indicates a limited compositional range of the single phase, on the order of 0.5 at.% or less. Fig. 3 is consistent with the metallographic observation that the single phase composition is near 24 at.% Si.

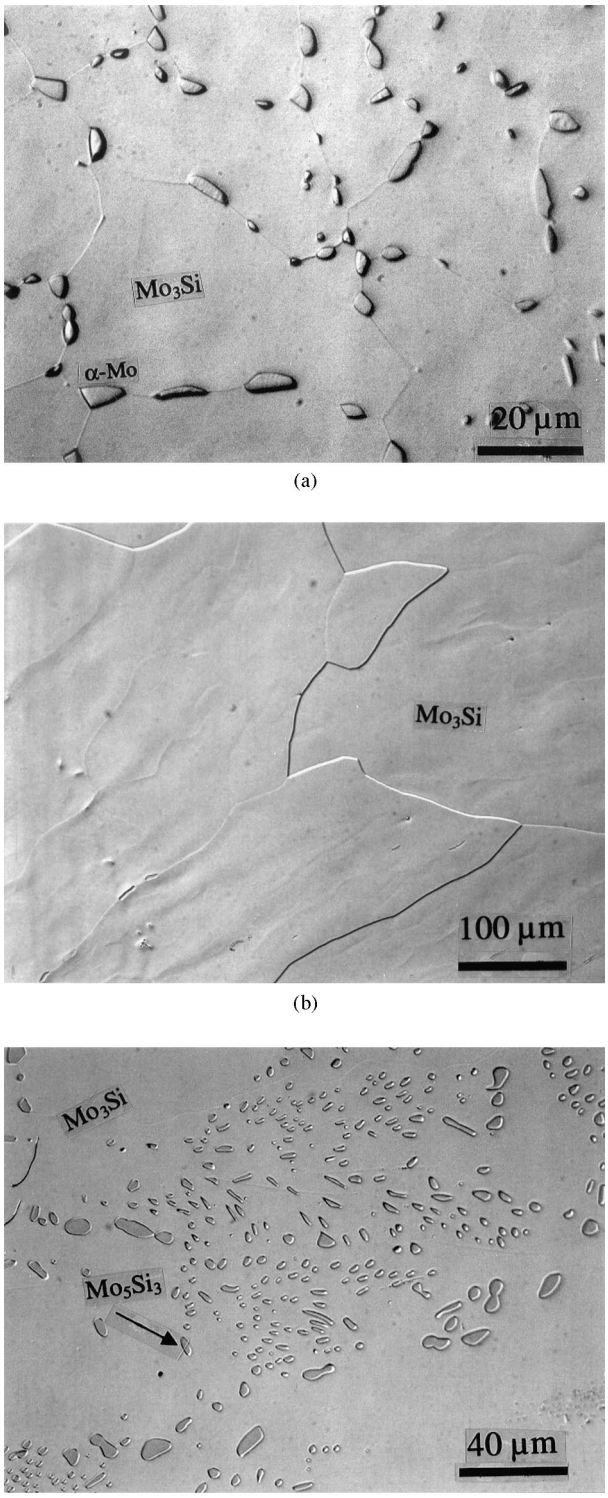


Fig. 2. Optical micrographs of etched specimens with (a) Mo–22 at.% Si with α-Mo precipitates; (b) single-phase Mo–24 at.% Si; (c) Mo–25 at.% Si with Mo_5Si_3 precipitates.

Between 23.5 and 24 at.% Si, the lattice parameter in Fig. 3 decreases with a slope of −0.007 Å/at.% Si. Since Si is a smaller atom than Mo, such a decrease is expected. A hard sphere model was used to estimate the change in lattice parameter with Si concentration. The Goldschmidt radii for Si and Mo (1.34 and 1.40 Å, respectively [18]) were used, and adjacent Mo atoms in Mo_3Si were allowed to overlap. Substitution of the Si atoms by Mo resulted in a lattice parameter increase of 0.1 Å, corresponding to a slope of −0.004Å/at.% Si. The similarity between the measured and predicted slopes shows that the slope in Fig. 3 is physically reasonable. Following Zhu et al. [19] the vacancy concentrations determined from X-ray and pycnometric densities were found to be on the order of 0.1 at.%. These values are on the order of the experimental error. This means that, for the heat treatment used in the current study, Mo_3Si contains less than 0.1 at.% thermal vacancies.

3.3. Microhardness

Table 2 shows the hardnesses of Mo_3Si alloys after furnace cooling from 1600°C or quenching from 1200°C

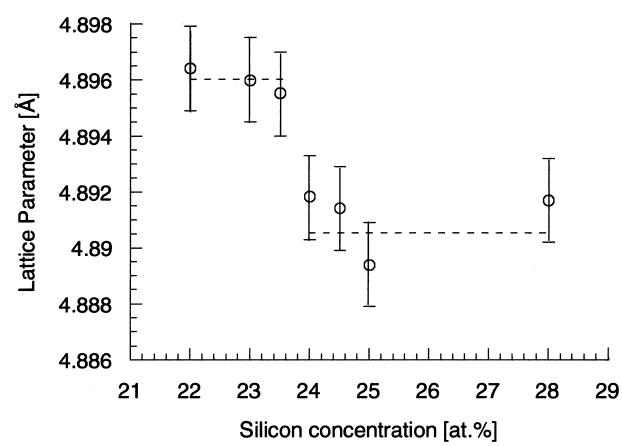


Fig. 3. Lattice parameters of cast and annealed (24 h/1600°C) Mo–Si alloys as a function of Si concentration.

Table 2
Microhardness of Mo–Si alloys for furnace cooled as well as quenched specimens

Silicon concentration (at.%)	Microhardness HVN (0.5)[a]	Microhardness HVN (0.5)[b]
22	1280±39	1295±43
24	1318±26	1322±22
25	1304±41	1322±35

[a] Annealed 24 hr/1600°C/furnace cooled.
[b] Annealed 2 hr/1200°C/oil quenched.

into oil. The indentations for these measurements were placed only in the matrix phase. In all cases, microcracks originated from the corners of the indent. Within the scatter of the measurements there was no distinct trend in the hardness as a function of Si concentration. Consistent with the absence of significant vacancy concentrations, the hardness of quenched Mo_3Si was similar to that of furnace cooled Mo_3Si.

3.4. Fracture toughness

Table 3 shows the fracture toughnesses of three Mo–Si alloys. For each alloy, two measurements were obtained. Since the Mo–22 at.% Si alloy contains α-Mo particles, some ductile phase toughening would be expected. However, the scatter in the measurements was too high to observe this trend. It should also be noted that the α-Mo volume fraction in Mo–22 at.% Si is small (see Fig. 3a). Quantitative metallography showed it to be only 8.0 vol.% and therefore not much toughening would be expected. The data in Table 3 indicate an average value on the order of 3.0 MPa $m^{1/2}$. The low value of one Mo–24 at.% Si specimen probably indicates the presence of a crack below the surface of the notch.

3.5. Compression tests

Fig. 4 shows the 0.2% offset yield stresses determined from compressive engineering stress–strain curves for Mo–22 at.% Si, Mo–24 at.% Si, and Mo–25 at.% Si deformed at 1400°C with initial strain rates of 10^{-3}, 1.8×10^{-4} and 10^{-5} s^{-1}, respectively (the stresses were evaluated from the load and the initial cross sectional area of the specimens). As would be expected, the yield stress decreases with decreasing strain rate. Although Mo–22 at.% Si contains α-Mo precipitates, its yield stress is higher than that of Mo–24 at.% Si. This is somewhat unexpected since α-Mo at 1400°C has a yield strength of 83 MPa [20], which is lower than that of Mo_3Si (see Fig. 4 for $10^{-3}s^{-1}$). One possible explanation for this finding is that the α-Mo particles may reduce internal cracking during compression. For high Si concentrations (25 at.%), Mo_5Si_3 precipitates form. Because of their high thermal expansion anisotropy [10], they may enhance internal cracking, and this may result in lower observed stress. Also, the defect populations in Mo_3Si may be different depending on whether the compound is Si-poor or rich, and this might give rise to different strengths.

Table 3
Room temperature fracture toughnesses of Mo–Si alloys

at.% Si	22	24	25
K_Q (MPa $m^{1/2}$)	3.05	3.73	3.49
K_Q (MPa $m^{1/2}$)	3.39	2.09	2.9

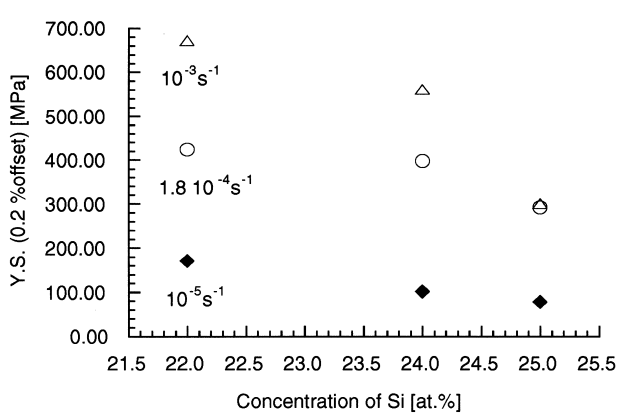


Fig. 4. Plot of 1400°C 0.2% offset yield strength vs. silicon concentration with initial strain rate of 10^{-3} s^{-1}, 1.8×10^{-4} s^{-1} and 10^{-5} s^{-1}.

4. Summary

We have shown that the A15 silicide Mo_3Si is not precisely stoichiometric, nor is it a true line compound. Its single phase composition is near Mo-24 at.% Si. After furnace cooling from 1600°C, the width of the single phase region is on the order of 0.5%. Within experimental accuracy (0.1 at%), furnace cooled Mo_3Si does not contain thermal vacancies. Its room temperature fracture toughness is on the order of 3.0 MPa $m^{1/2}$. Finally, the strength at 1400°C depends on the Si concentration and the strain rate.

Acknowledgements

This research was sponsored by the Division of Materials Sciences, US Department of Energy under contract number DE-AC05-00OR22725 with Oak Ridge National Laboratory, managed by UT-Battelle, LLC. The authors are grateful to C.A. Carmichael, J. Wright and E. Lee. Thanks go to J.A. Horton and C.G. McKamey of ORNL for reviewing this manuscript. I. Rosales acknowledges partial support by the Universidad Nacional Autonoma de México (UNAM-DGEP).

References

[1] Meyer MK, Kramer MJ, Akinc M. Intermetallics 1996;4:273–81.
[2] Meyer ML, Kramer MJ, Akinc M. Adv Mater 1996;8:85–8.
[3] Schneibel JH, Liu CT, Heatherly L, Kramer MJ. Scr Mater 1998;38:1169–76.
[4] Schneibel JH, Liu CT, Heatherly L, Kramer MJ. Scr Mater 1998;39:831.
[5] Berczik DM. US patent no. 5,595,616, 21 January 1997.
[6] Berczik DM. US patent no. 5,693,156, 2 December, 1997.
[7] Nunes CA, Sakidja R, Perepezko JH. In: Nathal MV et al., editors. Structural intermetallics 1997, Warrendale, (PA): TMS, 1997. 831–9
[8] Schneibel JH, Liu CT, Easton DS, Carmichael CA. Mater Sci Eng 1999;A26l:78–83.

[9] Liu CT, Schneibel JH, Heatherly L. In: High-temperature ordered intermetallic alloys VII. MRS Symposium Proceedings, vol 552, l999. p. KK6.2.1–7.
[10] Fu CL, Xindong W, Ye YY, Ho KM. Intermetallics 1999;7:179–84.
[11] Chu F, Thoma DJ, McClennan K, Peralta P, He Y. Intermetallics 1999;7:611–20.
[12] Christensen NA. Acta Chemica Scandinavica 1983;A37:519–22.
[13] Misra A, Petrovic JJ, Mitchell TE. Scripta Mater 1999;40:191–6.
[14] Raj SV. Mater Sci Eng l995;A201:229–41.
[15] Gokhale AB, Baschian GJ. In: Binary alloys phase diagrams. Metals Park (OH): ASM, 1986. p. 1632.
[16] Huebsch JH, Kramer MJ, Zhao HL, Akinc M. Intermetallics 2000;8:143–50.
[17] Anon. Standard test method for plane-strain fracture toughness of metallic materials (E399-90). ASTM, 1990.
[18] Laves F. Crystal struct. and atomic size. In: Theory of alloy phases. ASM, 1956 p. 131.
[19] Zhu JH. Acta Mater 1999;47:2003–18.
[20] Tietz TE, Wilson JW. Behavior and properties of refractory metals. Stanford (CA): Stanford University Press, 1965. p. 162.

PHILOSOPHICAL MAGAZINE LETTERS, 1999, VOL. 79, NO.3, 115–123

Defects in iron aluminides probed with positrons

BERTRAM SOMIESKI†§, JOACHIM H. SCHNEIBEL† and LESTER D. HULETT‡

† Metals & Ceramics Division and ‡ Chemical & Analytical Sciences Division, Oak Ridge National Laboratory, PO Box 208, Oak Ridge, Tennessee 37831, USA

[*Received 15 August 1998 and accepted 18 November 1998*]

ABSTRACT

Positron lifetime spectroscopy and shape analysis of the Doppler broadened positron annihilation peak were performed on iron aluminides containing 28 at.% , 40 at.% and 50 at.% Al. The electronic structure of the open volume defects, probably double vacancies, as seen by positron spectroscopy was independent of the heat treatment. Slight increases of the average positron lifetime as a function of quenching temperature indicate densities of double vacancies sufficiently high to cause the positron wavefunction to overlap with more than one defect. Positron lifetimes corresponding to trapping of vacancy-type defects increased with aluminium content. Changes in the shape of the annihilation peak suggest that the addition of aluminium causes a chemical shift of the iron 3d electrons to lower momenta.

§ 1. INTRODUCTION

The system $Fe_{1-x}Al_x$ is one of the most interesting transition metal aluminide systems with respect to high-temperature application (Rivière and Grilhé 1975, Munroe 1997). Besides the property changes due to different chemical compositions, it is found that vacancies play an important role in the strengthening of FeAl; see for example Chang *et al.* (1993) or Schneibel *et al.* (1996, 1997).

Positron spectroscopy is a very sensitive probe for open volume defects such as vacancies, vacancy clusters or dislocations. In fact, vacancy formation and migration enthalpies have been measured with high precision by means of positron lifetime spectroscopy in a broad variety of metals and alloys, such as Fe_3Al, cf. Schaefer *et al.* (1990) or Wolff *et al.* (1997a). It is known that vacancies in this system occur primarily on the iron sublattice (Paris and Lesbats 1978, Fu *et al.* 1993). Recently, it was shown by internal friction measurements that the thermal vacancies in the stochiometric compound FeAl are most likely double vacancies of second next neighbours on the iron sublattice $[V_{Fe}]_2$ and triple defects, consisting of two iron vacancies and one antisite iron atom $2[V_{Fe}] - Fe_{Al}$ (Damson 1998). The purpose of this work is to explore the influence of the aluminium content on the nature of the thermal vacancies.

§ 2. EXPERIMENTS AND ANALYSIS

The reference samples of 99.999 at.% iron and 99.99 at.% aluminium were polycrystalline material annealed for 1 h in vacuum at 1100°C and 350°C, respectively. The alloy specimens were prepared from (99.99 at.%) iron and aluminium metals, by

§ Corresponding author: email: somieskib@ornl.gov.

Philosophical Magazine Letters ISSN 0950–0839 print/ISSN 1362–3036 online © 1999 Taylor & Francis Ltd
http://www.tandf.co.uk/JNLS/phl.htm
http://www.taylorandfrancis.com/JNLS/phl.htm

Table 1. Sample specifications and preparations.

Sample	Material	Thermal treatment
FA28-7	$Fe_{72}Al_{28}$	1h @ 1000°C oil quenched
FA40-7	$Fe_{60}Al_{40}$	3d @ 700°C oil quenched
FA50-5	$Fe_{50}Al_{50}$	3d @ 700°C oil quenched 3.7d @ 500°C oil quenched
FA50-6	$Fe_{50}Al_{50}$	3d @ 700°C oil quenched 1d @ 500°C oil quenched
FA50-7	$Fe_{50}Al_{50}$	3d @ 700°C oil quenched
FA50-8	$Fe_{50}Al_{50}$	3d @ 700°C oil quenched 1d @ 800°C oil quenched

arc melting in argon. To establish a thermodynamic equilibrium of vacancies the samples were annealed and quenched as listed in table 1. The thin scale that appeared after oil quenching was removed by electrochemical etching.

The samples were analysed by standard positron lifetime spectroscopy. A fast-fast coincidence positron lifetime spectrometer with Pilot-U scintillators having a resolution function of 220 ps full width of half maximum (FWHM) was used. The source material was 14 μCi ^{22}NaCl solution deposited and dried on a 2 μm thick Al foil, which was placed between two identical specimens (sample–source–sample sandwich). Each spectrum contained 5–7 million single events. Source correction was made as described by Staab *et al.* (1996). The lifetime spectra were numerically analysed using the program LIFESPECFIT (Puska 1978). Repeated measurements confirmed the measured lifetimes.

Doppler broadening measurements were made by recording the energy shifts of both annihilation photons in coincidence. This caused an increased peak–background ratio of $\geq 10^5$ compared with a single-detector set-up (≈ 200), see, for example, Lynn and Goland (1976) or Szpala *et al.* (1996). The spectrometer consisted of two pure germanium detectors facing each other with a distance of 30 cm each to the centre where the sample sandwich was located. The same ^{22}Na source was used for both the lifetime and Doppler broadening measurements. Each 2γ coincidence spectrum contained 2^{24} ($\approx 1.67 \times 10^7$) single events, roughly 96% of which remained after discarding the Compton scattered gamma quanta. For that we added together the events along the lines $E_1 = E_2 + const.$ in a strip of $2m_0c^2 - 1.74\,\text{keV} < E_1 + E_2 < 2m_0c^2 + 1.74\,\text{keV}$, cf. the description in Asoka-Kumar *et al.* (1996). The distribution of Doppler shifts in energy was approximately symmetrical about the channel corresponding to 511 keV. The two sides of the distribution were summed and smoothed, with channel averaging over the FWHM of the spectrometer resolution function.

The S parameter and the W parameter are empirical measures of the degree of broadening. The S parameter is defined as the ratio of the area of the mid portion of the annihilation peak to the area of the entire peak. The W parameter is the ratio of the area of the wings of the peak to its entire area. The width of the regions for both parameters is user-chosen, frequently so that the defect free material has $S = 0.5$ and $W = 0.05$. S parameters are larger when positrons annihilate mostly

with low-momentum electrons. W parameters are larger, and S parameters are smaller, when annihilation occurs with core electrons of atoms, cf. West (1979).

To delineate differences in the summed and averaged spectra, they were normalized to a reference spectrum of pure silicon, Czochralsky-grown and oriented in the [100] direction. We normalized the spectra to the same area, i.e. overall annihilation probability $= 1$, instead of normalization to the same peak annihilation rate as performed by Myler and Simpson (1997), before they were divided channel-by-channel by the reference spectrum. We assume the normalization to the same overall annihilation probability to have the most physical meaning. Prior to the normalization, both reference as well as measured spectrum, were smoothed by adjacent averaging over a number of channels equal to the resolution function of the spectrometer to decrease stochastic noise.

The positron age-momentum correlated spectrum was obtained by using the output of the time–amplitude converter of the positron lifetime spectrometer as second channel in the 2γ coincidence set-up. For that the sample sandwich was placed right in front of the STOP probe facing the germanium detector at 30 cm distance. The START probe was located at an angle of $\approx 120^\circ$ from the STOP probe.

§ 3. Results and discussion

We measured in all samples positron lifetime spectra that consisted of a single component, the value of which increased as a function of the aluminium content, cf. figure 1. The lifetime values give evidence that all positrons are trapped by open volume defects. Our results are in reasonable agreement with those reported by Würschum *et al.* (1995) for thermal vacancies in $Fe_{61}Al_{39}$ ($\tau_1 = 187$ ps), but differ somewhat from those reported by Wolff *et al.* (1997a). They found positron lifetimes of 182 ps and 187 ps for $Fe_{57}Al_{43}$ and $Fe_{52}Al_{48}$, respectively, at 700°C, which are distinctly lower than those measured in this study.

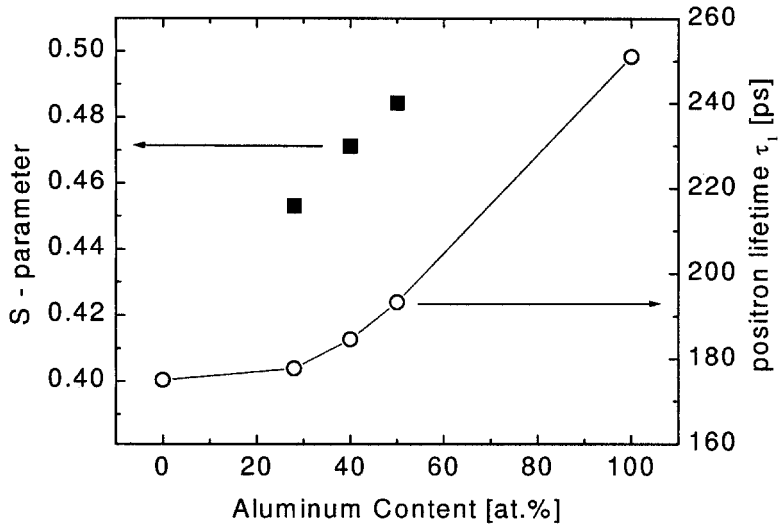


Figure 1. S parameter and defect lifetime τ_1 as a function of aluminium content in samples of $Fe_{1-x}Al_x$ oil-quenched from 700°C; τ_1 values for Fe and Al are from Vehanen *et al.* (1982) and Schaefer *et al.* (1987) respectively; errors are smaller than symbol size, lines are to guide the eye.

Recent experiments with positron spectroscopy, cf. Wolff *et al.* (1997a), and internal friction measurements, cf. Damson (1998), gave good reasons to assume that the thermal defects in stoichiometric $Fe_{50}Al_{50}$ above a temperature of $\approx 700^{\circ}C$ are primarily antisites as well as double vacancies on the iron sublattice, whereas the positron traps in the disordered A2 phase are apparently single vacancies, cf. the positron spectroscopy results of Wolff *et al.* (1997b). Neutron scattering experiments, however, were interpreted as to show vacancies on the Al sublattice, too (Kogachi *et al.* 1998).

Figure 1 shows that the measured lifetime values are close to that for single vacancies in pure iron, $\tau_{1v} = 175$ ps, cf. Vehanen *et al.* (1982). The observed increases in defect lifetime and S parameter with aluminium content agree with the measurements of Wolff *et al.* (1997a) and suggest that the average volume of the trapping sites increases as a function of aluminium content.

The increase in $\overline{\tau}$ and S parameter with the annealing temperature, cf. figure 2, might be interpreted as a change from double vacancies to vacancy aggregates of higher order. Owing to stochastic fluctuations it is possible, on the other hand, that

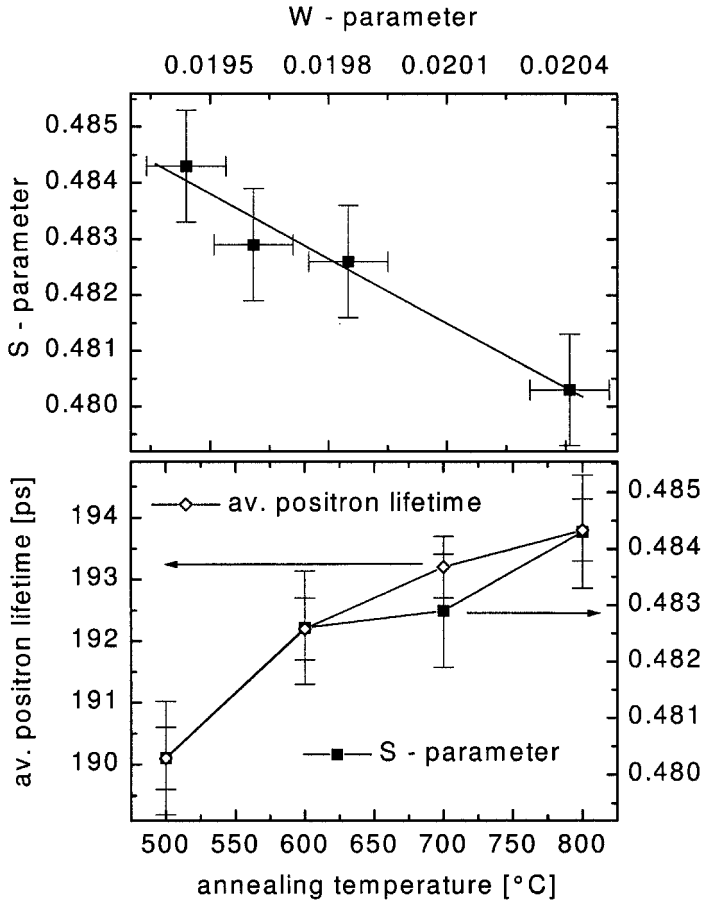


Figure 2. Average positron lifetime and S parameter in stoichiometric FeAl as a function of annealing temperature prior to quenching and S versus W plot; lines are to guide the eye.

after quenching a sufficiently large number of double vacancies, it is close enough for the positron wavefunctions to overlap two of them and, hence, be more confined. We agree with other workers (X. G. Zhang 1998, private communication) that this effect should slightly extend the lifetime of the positrons as well as increase the S parameter. The overall density of open volume defects, on the order of 1 at.% , cf. Chang *et al.* (1993), is high enough to be considered in this case. This is also supported by the S versus W plot in figure 2 which shows a rather straight line, indicating a change in the electronic structure of the trapping sites.

As can be seen from figure 3, the positron annihilation lines obtained from the material containing 40 at.% and 50 at.% Al, respectively, resemble almost completely the line of pure Al. The annihilation line from $Fe_{72}Al_{28}$, however, exhibits roughly the same overall shape as the line from pure iron, although the ratio to the reference is always lower in the entire high energy range $|E_\gamma - m_0c^2| \geq 7.6$ keV. We tried to decompose this peak as a linear combination of the iron and aluminium peaks, but due to the chemical shift, cf. figure 4, this linear combination did not match the line of $Fe_{72}Al_{28}$. The fact that the positron annihilation lines in $Fe_{50}Al_{50}$ and $Fe_{60}Al_{40}$ are only slightly different from that of pure Al seems to indicate that the positrons are trapped at vacancies in the iron sublattice which are surrounded by Al atoms. This would be in agreement with the conclusions of many other authors that vacancies in the system Fe–Al are mainly located on the iron sublattice, e.g. Damson (1998), even if Mössbauer spectroscopy shows that at high temperature iron atoms diffuse partially via vacancies on the aluminium sublattice, cf. Vogl and Sepiol (1994) and Feldwisch *et al.* (1995). On the other hand, it might be possible that, due to electron transfer, the positron is attracted more by the aluminium atoms and, therefore, its wavefunction is biased towards the aluminium sites. To address this question, *ab initio* calculations are proposed. The shape of the quotient spectrum of $Fe_{72}Al_{28}$,

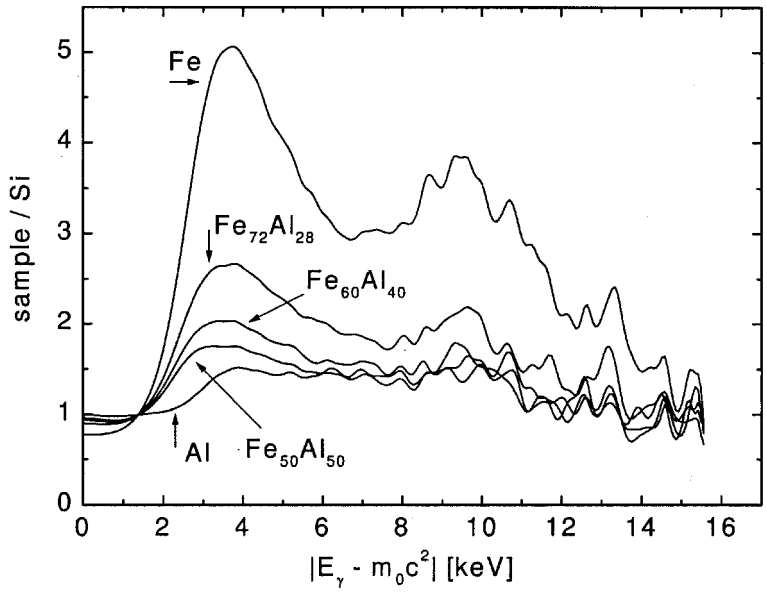


Figure 3. Doppler broadened positron annihilation line in $Fe_{1-x}Al_x$ alloys oil-quenched from 700°C; the lines are normalized to the one of pure silicon.

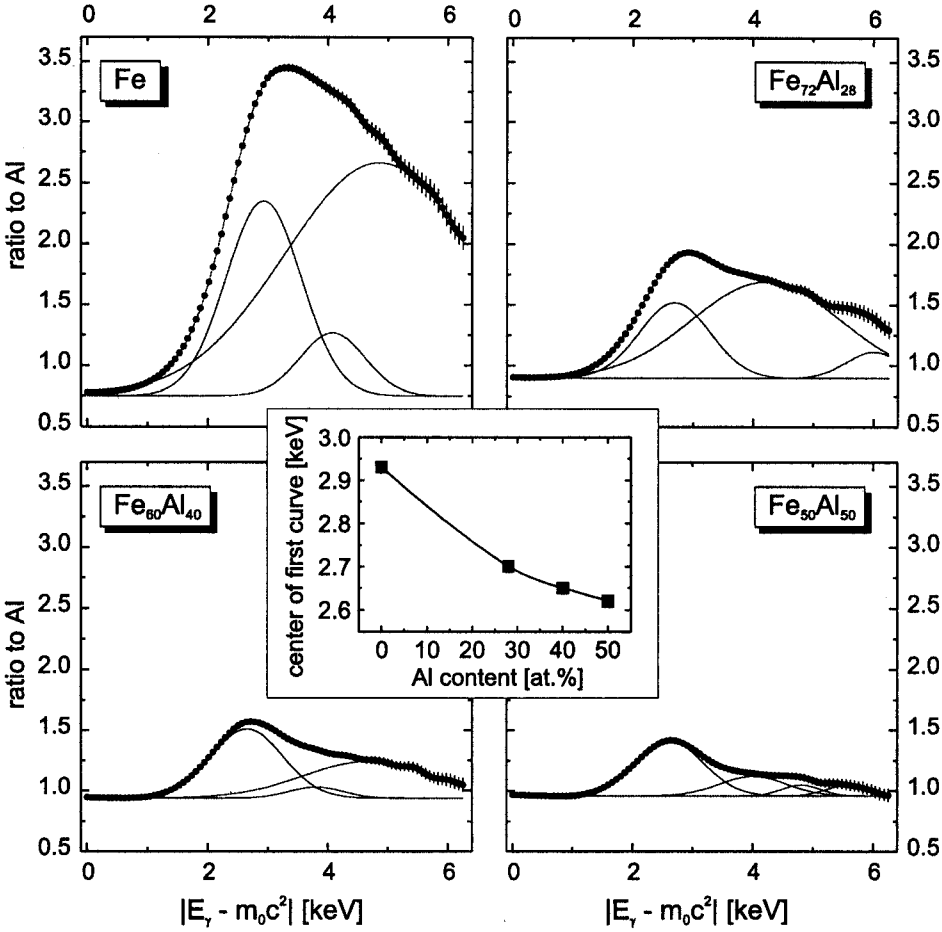


Figure 4. Fit of three Gaussian curves to the positron annihilation line of $Fe_{1-x}Al_x$ normalized to pure Al. The inset shows the centre position of the first fitted curve as a function of the aluminium content; the line is to guide the eye.

however, is similar to that of pure iron, attributed to annihilation with iron electrons. It indicates that the positrons are attracted not only by the aluminium atoms. The environment of the vacancies consists of iron as well as aluminium atoms in the $A2$ structure.

The shape of the normalized positron annihilation peak of pure iron agreed qualitatively with that given by Myler and Simpson (1997). The main peak of our spectrum was located at the same energy $|E_\gamma - m_0c^2| \approx 3.8$ keV as that of Myler and Simpson, but there were distinct differences in the peak shapes for $|E_\gamma - m_0c^2| \geq 8$ keV. It is unknown at present what caused these differences. Different maximum quotients can be attributed to the different normalization, cf. § 2, but not differences in shape.

The shape of the positron annihilation peak for $Fe_{50}Al_{50}$ was independent of the annealing temperature, consistent with the positron lifetime results. The quotient spectra derived from the peaks measured after the different quenches would all fit a curve slightly broader than that shown in figure 3 for the 700°C specimen. These

results indicate that the chemical environment of the trapping sites remained basically the same after the different heat treatments. Distinct differences, however, were found in the very sensitive S parameters and W parameters, as shown in figure 2, which is attributed to different defect densities only.

The normalized spectra in figure 3 indicate that the addition of aluminium causes the main peak, corresponding to iron, to be shifted to different energies. To better illustrate this, the data in figure 4 are normalized to pure aluminium rather than to pure silicon as in figure 3 to bring out the differences in the electronic structure. We concentrate on the energy range $|E_\gamma - m_0c^2| \leq 6.25$ keV. Clearly, the maximum of the peak shifts to lower energies as the aluminium content increases. By an empirical approach similar to that of Myler and Simpson (1997), we fitted three Gaussian curves to the quotient spectra. The inset in figure 4 shows that the centroid of the main Gaussian (lowest energy) shifts from a difference of 2.93 keV with respect to m_0c^2, to a difference of 2.63 keV. Assuming the effective mass to be similar to the rest mass, $m_{\text{eff}}/m_0 \approx 1$, the energy shift of the γ quanta would arise from a shift in the kinetic energy of the involved electrons from 33.6 eV to 27 eV. We tentatively interpret this effect as due to the incorporation of the iron 3d electrons into the band states of the alloy. The delocalization, perhaps, caused the momenta of the iron electrons to be lowered.

It is questionable whether or not the high density of defects will significantly alter or even destroy the band structure of the intermetallic alloys. This gives rise to the possibility that prethermal or pretrapping annihilation causes the observed changes in the electron momenta. To exclude this possibility we measured an age-momentum correlated spectrum (AMOC), cf. the article of Siegle *et al.* (1997) for the $Fe_{50}Al_{50}$ sample, which had the highest density of defects. As can be seen from figure 5, the S

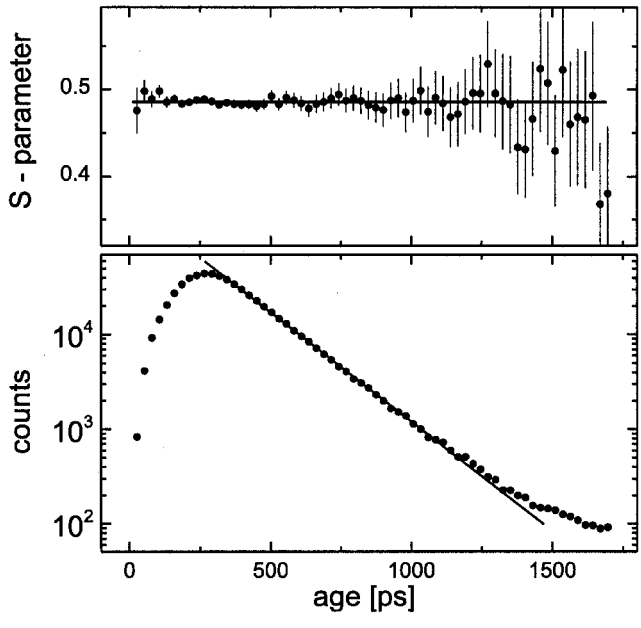


Figure 5. S parameter as a function of positron age in $Fe_{50}Al_{50}$ (sample FA50-7, cf. table 1) and corresponding positron lifetime spectrum. Slight deviations from the straight line are due to contributions from the positron source.

parameter, and, therefore most likely, the entire electron environment is independent of the age of the positron. Hence, the observed shift of the annihilation gamma quanta is really due to lowered electron momenta. Owing to the low count rate of $\approx 2.6\,s^{-1}$ we were limited to $\approx 6 \times 10^5$ counts per spectrum, which prevented more detailed studies of the annihilation line as a function of the positron age.

§ 4. Conclusions

Positron lifetime and the Doppler broadened positron annihilation line were measured in oil quenched iron aluminium samples of different stoichiometry. In all samples the density of the quenched-in defects was sufficient to trap all positrons. We found single lifetimes only, indicating one major kind of defect. The lifetime values found were consistent with conclusions in the literature that at aluminium contents of 40 at.% and greater the defects are predominantly double vacancies.

Slight increases in lifetime and S parameter as a function of quenching temperature are interpreted as due to the very high density of double vacancies, causing the positron wavefunctions to overlap more than one site.

The annihilation peak arising from positrons trapped at double vacancies in $Fe_{60}Al_{40}$ and $Fe_{50}Al_{50}$ resembles the peak of pure aluminium, supporting the theory that thermal vacancies in those alloys are mainly located in the iron sublattice.

By means of 2γ coincidence Doppler broadening spectroscopy it was found that the momenta of some of the iron electrons, probably the 3d, with which the positrons annihilate decrease with an increasing amount of aluminium in the alloy.

Acknowledgements

We are indebted to X. G. Zhang and to D. M. C. Nicholson, Oak Ridge National Laboratory, Metals & Ceramics Division for valuable discussions.

This research was supported in part by an appointment to the Oak Ridge National Laboratory Postdoctoral Research Associates Program administered jointly by the Oak Ridge National Laboratory and the Oak Ridge Institute for Science and Education.

This research was sponsored by the US Department of Energy, Office of Fossil Energy, Divison of Materials Sciences under contract DE-AC05-96OR 22464 with Lockheed Martin Energy Research Corporation.

References

Asoka-Kumar, P., Alatalo, M., Ghosh, V. J., Kruseman, A. C., Nielsen, B., and Lynn, K. G., 1996, *Phys. Rev. Lett.*, **77**(10), 2097.

Chang, Y. A., Pike, L. M., Liu, C. T., Bilbrey, A. R., and Stone, D. S., 1993, *Intermetallics*, **1**, 107.

Damson, B., 1998, PhD thesis, Universität Stuttgart, Stuttgart, Germany.

Feldwisch, R., Sepiol, B., and Vogl, G., 1995, *Acta metall. mater.*, **43**(5), 2033.

Fu, C. L., Ye, Y.-Y., Yoo, M. H., and Ho, K. M., 1993, *Phys. Rev.* B, **48**(9), 6712.

Kogachi, M., Haraguchi, T., and Kim, S. M., 1998, *Intermetallics*, **6**(6), 499.

Lynn, K. G., and Goland, A. N., 1976, *Solid State Commun.*, **18**, 1549.

Munroe, P. R., 1997, *Proceedings of the International Symposium on Nickel and Iron Aluminides: Processing, Properties, and Applications* (Materials Park, Ohio: ASM International), pp. 95–113.

Myler, U., and Simpson, P. J., 1997, *Phys. Rev.* B, **56**(22), 14303.

Paris, D., and Lesbats, S., 1978, *J. Nucl. Mater.*, **69&70**, 628.

Puska, M. J., 1978, program LIFESPECFIT, Teknillinen Korkeakoulu, Otaniemi, Finland.

RIVIÈRE, J. P., and GRILHÈ, J., 1975, *Scripta metall.*, **9,** 967.

SCHAEFER, H.-E., GUGELMEIER, R., SCHMOLZ, M., and SEEGER, A., 1987, *Mater. Sci. Forum*, **15–18**(1), 111.

SCHAEFER, H.-E., WÜRSCHUM, R., ŠOB, M., ŽAK, T., YU, W. Z., ECKERT, W., and BANHARD, F., 1990, *Phys. Rev. B*, **41**(17), 11869.

SCHNEIBEL, J. H., SPECHT, E. D., and MUNROE, P. R., 1996, *Intermetallics*, **4,** 581.

SCHNEIBEL, J. H., SPECHT, E. D., and SIMPSON, W. A., 1997, *Mater. Sci. Eng.*, **A239–240**, 245.

SIEGLE, A., STOLL, H., CASTELLAZ, P., MAJOR, J., SCHNEIDER, H., and SEEGER, A., 1997, *Appl. Surf. Sci.*, **116**, 140.

STAAB, T. E. M., SOMIESKI, B., and KRAUSE-REHBERG, R., 1996, *Nucl. Instr. Methods*, **A381,** 141.

SZPALA, S., ASOKA-KUMAR, P., NIELSEN, B., PENG, J. P., HAYAKAWA, S., LYNN, K. G., and GOSSMAN, H.-J., 1996, *Phys. Rev.* B, **54**(7), 4722.

VEHANEN, A., HAUTOJÄRVI, P., JOHANSSON, J., YLI-KAUPPILA, J., and MOSER, P., 1982, *Phys. Rev.* B, **25**(2), 762.

VOGL, G., and SEPIOL, B., 1994, *Acta metall. mater.*, **42**(9), 3175.

WEST, R. N., 1979, *Positrons in Solids*, Vol. 12, *Topics in Current Physics*, edited by P. Hautojärvi (Berlin: Springer), pp. 89–144.

WOLFF, J., BROSKA, A., FRANZ, M., KÖHLER, B., AND HEHENKAMP, T., 1997a, *Mater. Sci. Forum*, **255–257,** 593.

WOLFF, J., FRANZ, M., BROSKA, A., and HEHENKAMP, T., 1997b, *Defect Diffus. Forum*, **143–147** 239.

WÜRSCHUM, R., GRUPP, C., and SCHAEFER, H.-E., 1995, *Phys. Rev. Lett.*, **74**(1), 97.

ELSEVIER

Intermetallics 8 (2000) 1313–1320

Intermetallics

www.elsevier.com/locate/intermet

Emerging applications of intermetallics

N.S. Stoloff [a,*], C.T. Liu [b], S.C. Deevi [c]

[a] Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA
[b] Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37380, USA
[c] Research Center, Philip Morris Inc, Richmond, VA 23234, USA

Received 1 March 2000; accepted 31 May 2000

Abstract

Many intermetallic compounds display an attractive combination of physical and mechanical properties, including high melting point, low density and good oxidation or corrosion resistance. This has led to their utilization in many non-structural applications, but success in structural applications has, to date, been limited. This paper reviews the current status of intermetallic applications, with emphasis on new uses that are in place or pending. Most of the paper deals with aluminides and silicides, but there are several more complex intermetallics that are being employed in battery and magnetic applications. Research on improved processing and studies of the role of environment in mechanical behavior are shown to be key to developing practical alloys. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. iron aluminides (based on Fe_3Al); A. Magnetic intermetallics; A. Nickel aluminides, based on Ni_3Al; A. Molybdenum silicides; B. Superplastic behavior

1. Introduction

Intensive studies of the mechanical and physical properties of intermetallic compounds have led to many suggestions for potential structural and non-structural applications. These include high temperature gas turbine hardware, corrosion resistant materials, heat treatment fixtures, magnetic materials and hydrogen storage materials. It is the objective of this paper to describe in detail these and other applications, and to point out research that needs to be done to insure improved mechanical and/or physical properties.

2. Characteristics of intermetallics

For at least the past four decades research on intermetallic compounds has largely focussed upon mechanical properties, especially involving low temperature ductility and high temperature strength. Unfortunately, many intermetallics are brittle or semi-brittle at room temperature, rendering them difficult to fabricate or utilize in structural applications. In recent years it has been found that brittleness in several aluminides and silicides arises from an extrinsic environmental effect which can be overcome by alloying, microstructural control or coating. Examples of such compounds that are adversely affected by moist environments are FeAl, Fe_3Al, Ni_3Al and Ni_3Si. Intermetallics, which are of interest for elevated temperature applications, such as the titanium aluminides, $MoSi_2$ and several Laves phases, often display inadequate creep resistance. In some cases such resistance has been improved by alloying or compositing. Other key issues for many intermetallics are the ability to process them in a cost-effective manner and the provision of adequate corrosion or oxidation resistance. Finally, for most compounds an insufficient database exists for designers to have confidence in replacing existing materials with intermetallics.

Other compounds have been studied for their interesting physical properties, such as superconductivity, hydrogen storage capability or magnetic properties. Examples of intermetallics in each of these categories, together with their prospects for industrial application, will be discussed in subsequent sections of this paper. An overall summary of applications for intermetallics appears in Table 1.

* Corresponding author. Tel.: +1-518-276-6436; fax: +1-518-276-8554.

E-mail address: stolon@rpi.edu (N.S. Stoloff).

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(00)00077-7

3. Intermetallics for structural, heat-resistant and corrosion-resistant applications

The most widely studied intermetallics in this class include aluminides of titanium, nickel, iron and niobium, silicides of nickel, molybdenum and niobium and Laves phases such as Cr_2Nb. The physical properties of these compounds are summarized in Table 2. Many of these compounds have excellent corrosion and oxidation properties because of the high content of elements that form protective oxides. Therefore, applications for these compounds sometimes extend far beyond their strength and ductility.

3.1. Ni_3Al

This compound has been the object of much research to understand the factors controlling low temperature ductility. It is now known that the brittleness of unalloyed Ni_3Al stems from an environmental effect, and boron serves to suppress the embrittlement. Numerous alloys based upon Ni_3Al have been developed with the aim of improving high temperature creep resistance. Although the mechanical properties are attractive, most current usage is mainly in corrosion-related structural applications. This is because oxidation and carburization resistance is high, as are resistance to wear and cavitation-erosion. A summary of current and potential applications appears in Table 3. These applications range from furnace rolls and radiant burner tubes for steel production to heat treating fixtures, forging dies and corrosion-resistant fixtures for chemical industries. A photograph of various sizes of heavy wall, centrifugally cast tubes of Ni_3Al alloy IC–221M (Ni–8wt.%Al–7.7 Cr-1.43Mo–1.7Zr–0.008B) appears in Fig. 1 [1]. Other applications for this alloy include radiant burner tubes, center posts for pit-carburizing furnaces and guide rolls for continuous casters. A stronger, more creep resistant Ni_3Al alloy, IC-438, (Ni–8.1wt.%Al–5.23Cr–7.02Mo–0.13Zr–0.005B) has been identified; this alloy allows potential users to extend the maximum use temperature of 1100°C for IC-221M to 1250–

Table 1
Applications of intermetallics

Structural
automotive
aerospace
Magnetic
Energy storage
batteries
hydrogen storage
Heating elements
Tools and dies
Furnace hardware
Corrosion-resistant
piping for chemical industries
cladding
coatings
Electronic devices

Table 2
Properties of high temperature intermetallics

Compound	Structure	T_m °K	ρ g/cc	$E_{25°C}$ GPa	Comment
Cr_3Si	A15	2043	6.5	357	Solubility range
$MoSi_2$	C11	2303	6.24	425	Line compound
Cr_2Nb	C15	2043	7.7	218	C14 above 1858°K; line compound
Nb_3Al	A15	2233	7.29	–	Line compound
TiAl	$L1_0$	1723	3.9	173	Solubility range
Ti_3Al	DO_{19}	1873	4.2	147	Solubility range
Nb_5Si_3	D8	2793	7.16	–	Line compound
Nb_3Si	$L1_2$	2153	7.4	–	Limited range of stability

Table 3
Applications of Ni_3Al

Steel
As furnace rolls
Casting rolls
Radiant burner tubes
Heat treating
Fixtures for carburizing, furnaces, and air
Link belts for heating treating furnaces
Furnace muffles
Radiant burner tubes
Chemical
Reaction vessels for higher temperatures
Tube hangers
Pallet tips for phosphate ore calcination
Pump impellers for slurries
Forging
Forging dies
Die repair as weld overlay
Chemical
For ethylene crackers
Air deflectors for burning of high sulfur fuel

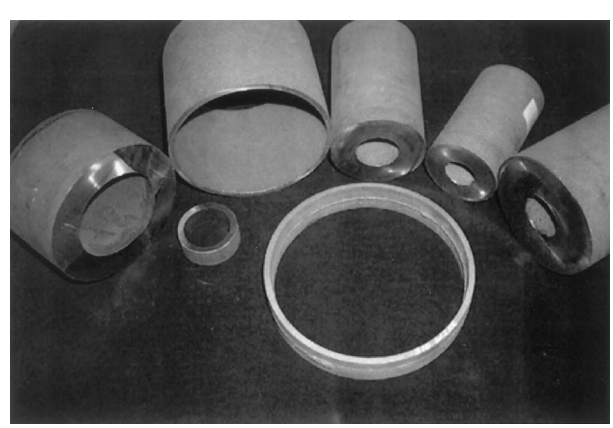


Fig. 1. Photograph of various sizes of heavy-wall centrifugally cast tubes of IC-221 M [1].

1300°C for IC-438. Creep rupture strengths of the two alloys, based upon the Larson–Miller parameter, are compared in Fig. 2 [1,2].

3.2. NiAl

Much effort has been invested in attempts to utilize NiAl alloys for gas turbine hardware. Unfortunately, while NiAl single crystals were developed with creep strengths comparable to those of Ni-base superalloy single crystals, other mechanical properties of NiAl were found to be inadequate. Ductility and fracture surfaces remain low. In addition, Walston and Darolia [3] showed in 1997 that impact resistance of high strength single crystals is inadequate for turbine blades, but might be sufficient for stationary parts such as vanes and combustor liners. A directionally solidified NiAl alloy was more resistant to impact. Physical vapor deposited thermal barrier coatings did not improve impact resistance. It appears unlikely that these shortcomings can be overcome in the near future.

3.3. FeAl and Fe_3Al

The iron aluminides, FeAl and Fe_3Al, are notable for their low cost, ease of fabrication and corrosion and oxidation resistance. In addition, FeAl is characterized by good resistance to catalytic coking, carburization, sulfidation and wear. As a result, FeAl has seen application as transfer rolls for hot rolled steel strip, ethylene crackers and air deflector for burning high sulfur coal. Structural applications for these compounds have been limited by low ambient ductility, due largely to embrittlement by moisture in air. However, several methods to combat environmental embrittlement have been developed. These include control of grain size and shape, use of alloying elements such as Cr for Fe_3Al and B for FeAl, and the application of oxide or copper coatings [4]. These developments, combined with improved creep and impact resistance provided by alloying, have improved the likelihood that monolithic iron aluminides may be utilized for structural applications. Alternatively, the excellent corrosion and oxidation resistance of iron aluminides suggests their possible usefulness as coatings. For example, steels have been successfully coated by $(Fe,Cr)_3Al$ by a two-step pack cementation process, as shown in Fig. 3 [5].

Another development that favors the near-term utilization of iron aluminides is the development of the Exo-Melt™ process, see Fig. 4 [6]. This low cost, easily controllable process is useful also for nickel aluminides, as it exploits the exothermic heat of reaction between aluminum and other elements to reduce the need for external power during melting.

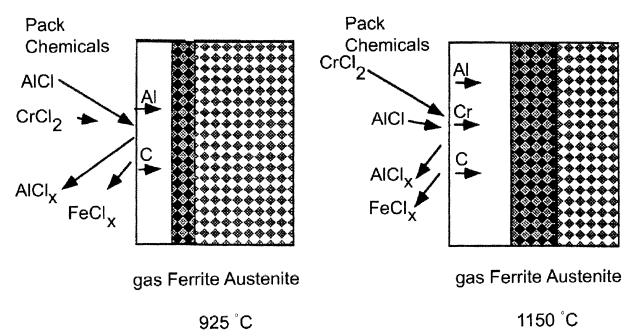


Fig. 3. Schematic of two-step pack cementation coating process. Longer arrows correspond to higher vapour pressures. The darkend area represents a zone of carbon enrichment in austenite. Microstructure of T11 steel coated in a pack with 2AI-18Cr, $2CrCI_2$ at 925°C for 4 h then at 1150°C for 3 h [5].

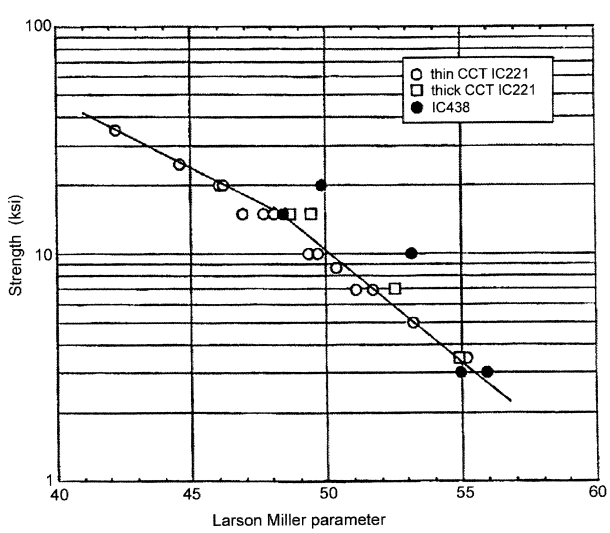


Fig. 2. Comparison of creep-rupture strength of IC-438 with the data for IC-221M [1].

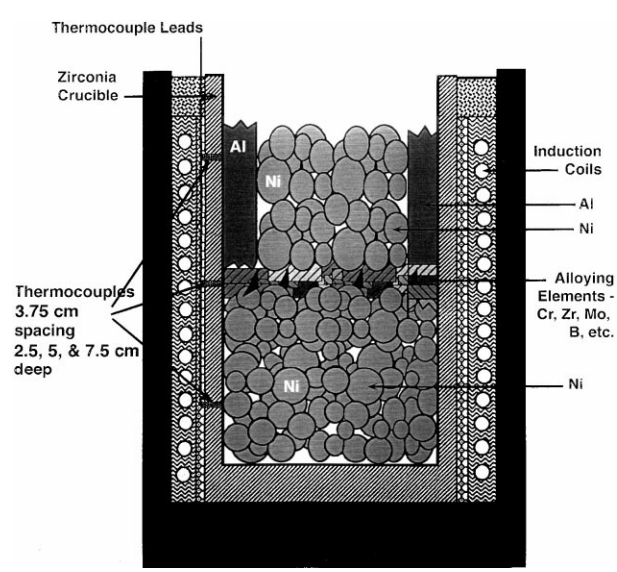


Fig. 4. Furnace-loading scheme of the EXO-Melt™ process for melting of iron and nickel aluminides [6].

Yet another promising processing technique for iron aluminides is the use of superplastic forming. Li et al. [7] and Lin et al. [8] have demonstrated superplasticity for FeAl and Fe_3Al alloys, respectively. However, no applications utilizing this technique have been reported.

3.4. TiAl

One of the most promising intermetallics for turbine applications is TiAl. This compound has a higher melting point, better oxidation resistance and resistance to fires, as well as lower density than conventional titanium alloys, but has suffered from low room temperature ductility and fracture toughness. Nevertheless, numerous potential aircraft applications have been identified, see Table 4 [9]. However, the aircraft industry is extremely demanding in qualifying new alloys, and the difficulty in fabricating these compounds, coupled with the lack of an adequate data base of mechanical properties, have been serious impediments to implementation. An advantage of TiAl is that this compound has a higher melting point than competing alloys that combine the requisite ductility and creep resistance for turbine applications. One such alloy, developed by M.Nazmy and co-workers, is Ti–47at.%Al–2.1W–0.5Si. [10,11]. This alloy has improved oxidation and creep resistance, such that it is being proposed as a marine turbine alloy for a high speed ferry. In this application, the maximum turbine inlet temperature is 610°C. It has been successfully run in tests to 1856 h, when it underwent its first inspection (September, 1999). Additional uses of TiAl alloys may arise from the ability to superplastically deform this compound at relatively low temperatures [12] or to spray form deposits on a substrate, see Fig. 5 [13]. These processing techniques, as well as the ability to cast large shapes, avoid the problems arising from lack of formability by conventional working processes.

3.5. $MoSi_2$

This line compound is particularly noteworthy for its combination of very high melting point, low density and outstanding oxidation resistance. Unfortunately, the very low ductility and fracture toughness of the pure compound has limited usage to mostly non-structural applications. A list of current and proposed applications appears in Table 5. The most important use to date is as heating elements (Kanthal). Recent work by Akinc et al. [14] has shown that improvements in electrical resisivity of $MoSi_2$ can be achieved by alloying with B-containing phases, see Fig. 6. An important advantage of this compound is the ability to utilize a wide range of processing techniques to synthesize. One of the most promising of these techniques is powder processing to produce an alloy or a functionally gradient material (FGM) between $MoSi_2$ and ceramics such as Si_3N_4 [15]. The latter is particularly beneficial as a solute in $MoSi_2$ because increased strength is accompanied by resistance to catastrophic oxidation (pesting). A schematic drawing of a diesel energy combustion chamber, see Fig. 7 [16] illustrates the use of $MoSi_2$–Si_3N_4 glow plugs produced from a FGM process. Another recent application for FGM $MoSi_2$ is in a hybrid direct energy conversion system under development in Japan, as shown in Fig. 8 [17] This device aims to convert solar energy into electricity by combining thermionic and thermoelectric conversion stages in a single device. The objective is increased energy conversion efficiency. Other $MoSi_2$-ceramic alloys

Table 4
Aircraft engine applications for TiAl [9]

Low pressure turbine blades: CF6-80C2, GE 90
Carbon seal supports: F414
Transition duct beams: GE 90
Blade dampers, high pressure turbine
Compressor blades: Allison 14th stage
High speed civil transport components
Diffuser casting: advanced engines

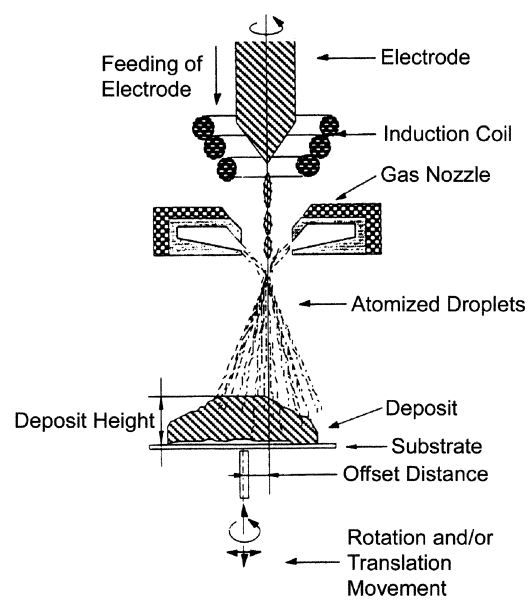


Fig. 5. Schematic drawing of EIGA setup for spray forming [13].

Table 5
Applications for $MoSi_2$

Turbine hardware
vanes
combustor liners
Flame holders
Diesel engine glow plugs
Igniters
Electronic devices
Heating elements
Glass melting
Energy conversion devices

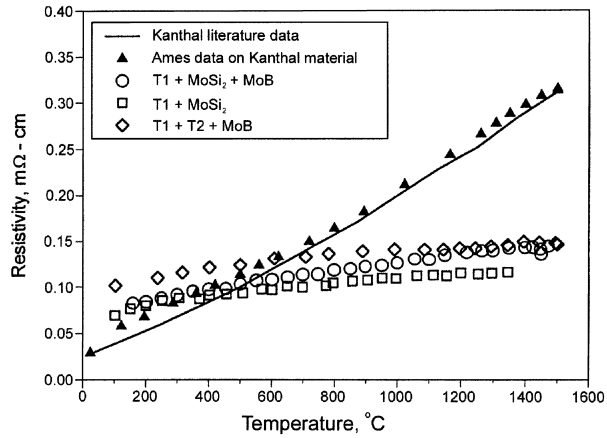


Fig. 6. Temperature dependence of electrical resistivity for boron-doped Mo_5Si_3 with commercial Kanthal® Super $MoSi_2$ heating element material for comparison (T1 : Mo_5Si_3 , T2 : Mo_5SiB_2) [14].

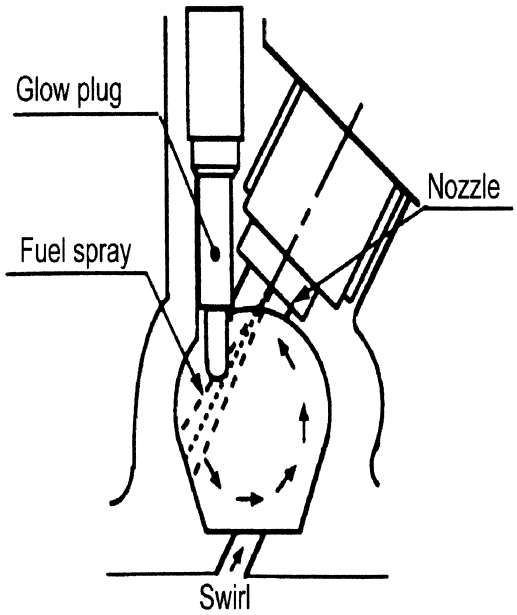


Fig. 7. Combustion chamber (Swirl chamber) for diesel engine [16].

have been patented for use as heating elements, igniters and heat sensors [18,19].

Because of high corrosion resistance, $MoSi_2$ tubes have been tested in closed-circuit monitoring systems for glass melters [20].

Recently, research on molybdenum silicides for high temperature applications has shifted from $MoSi_2$-base to Mo_5Si_3-base compositions. [21]. This is because the latter contain more Mo and demonstrate better creep resistance. Oxidation resistance of Mo_5Si_3 is not as good, but can be improved substantially by the addition of boron [22]. Considerable effort is now being devoted to study of the Mo–Si–B system [23].

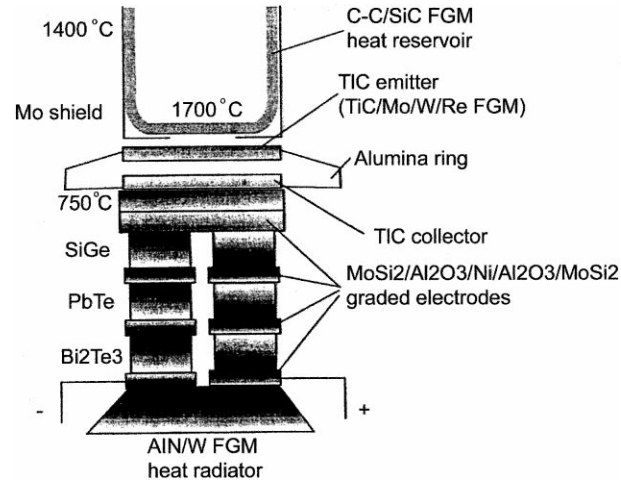


Fig. 8. The concept and the design of HYDECS (based on [17]).

3.6. Nb_3Si

Niobium silicides are among the most recent intermetallic alloys to be studied for possible high temperature structural applications. Bewlay and co-workers [24] have reported outstanding mechanical properties for such alloys, some of which are produced as directionally solidified eutectic composites. A comparison of temperature capability for less than 25% μm of metal loss in 100 hours for initial and more recent niobium silicide-based composites is shown in Fig. 9, together with the estimated goal [24]. Note that the present silicide composites are close to advanced nickel-base superalloy single crystals in their temperature capability.

3.7. Ni_3Si

This $L1_2$ intermetallic resembles Ni_3Al in that it is ductile when tested in inert environments, but is embrittled by contact with moisture. Both alloys are ductilized in air by the addition of small amounts of boron. Researchers at ORNL have demonstrated excellent mechanical properties of Ni_3Si alloys [25,26] at elevated temperatures. Such alloys have the potential for fabrication of complex components for use in chemical process systems. Two phase Ni_3Si–Ni alloys with excellent fabricability and weldability are currently being developed for structural use in acid and oxidizing environments [27].

4. Intermetallics for electronic, magnetic, battery applications

4.1. Electronics and sensors

The widespread use of silicides in the electronics industry has been reviewed by Kumar [28]. These compounds

are used as superconductors, ohmic contacts for integrated circuits, for growth of epitaxial films and as infrared detectors and sensors. Other intermetallics are now being studied for electronic applications. For example, NiAl and Ni_3Al substrates are being used to form an insulating alumina layer by oxidation in air at temperatures between about 900 and 1200°C, prior to applying conductive elements to the alumina layer [29]. These circuit components have improved mechanical properties and higher thermal conductivity compared to alumina substrates sold in Japan and the United States.

The shape-memory alloys typified by TiNi have been examined extensively for possible use in small devices such as microvalves. These devices have the potential to be used in microelectromechanical systems (MEMS) [30].

4.2. Magnets

Several intermetallic compounds, including FeCo and rare earth compounds, have been used as permanent magnets. One of the most interesting compounds is $Nd_2Fe_{14}B$, which has the highest energy product of commercial permanent magnets [31]. Improvements in the fracture stress and toughness of these magnets would allow greater machinability, easier handling and use as a structural element. However, very little work has been done in recent years on the mechanical properties of hard magnets. Improved mechanical properties are needed to fabricate electric vehicle wheel motors. The use of improved permanent magnets in heat pump compressors and fan motors could provide substantial energy savings. Researchers at Oak Ridge National Laboratory have initiated an investigation of mechanical properties of $Nd_2Fe_{14}B$ [32].

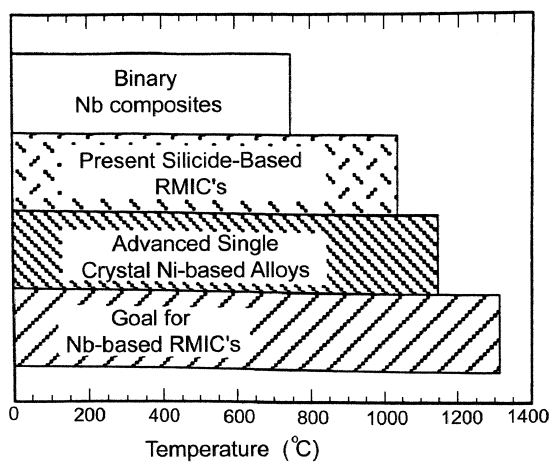


Fig. 9. A comparison of the temperature capability for less than 25μm of metal loss in 100 h for initial and more recent niobium silicide-based RMICs, together with the estimated goal. The temperature capability of an advanced single-crystal nickel based super-alloy is also presented in the figure for comparison [24].

4.3. Batteries

There is a large family of hydride-forming intermetallic compounds that lend themselves to use in Ni-metal

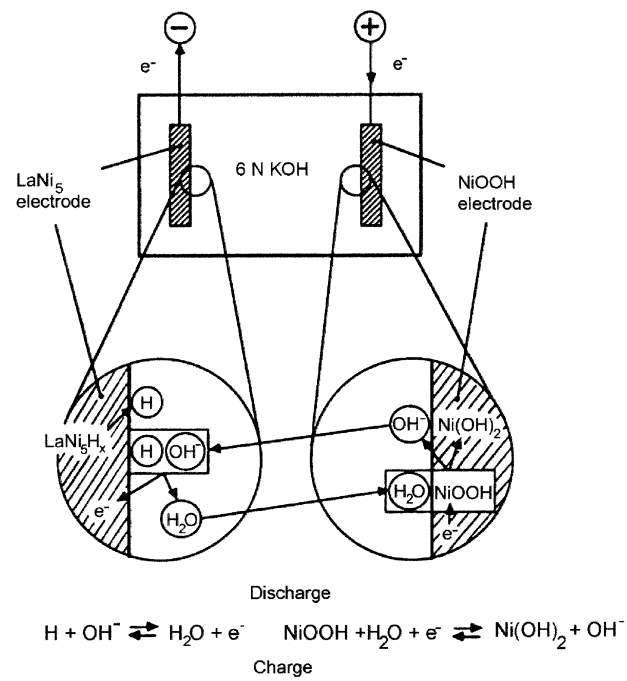


Fig. 10. A reversible battery with $LaNi_5$ metal hydride and NiOOH electrode and KOH electrolite [33].

Table 6
Families of hydride-forming intermetallic compounds for Ni-metal hydride batteries from [33]

Family	Prototype material	Crystal structure type		Hydride/deuteride
AB_5	$LaNi_5$	$CaCu_5$, $D2_d$ (hP6)		$LaNi_5H_{6.5}$
AB_2	$ZrMn_2$, $TiMn_2$	$MgZn_2$, C14 (hP12)	Laves phases	$ZrMn_2D_3$, $TiMn_2D_3$
	$ZrCr_2$, ZrV_2	$MgCu_2$, C15 (cF24)	Laves phases	$ZrCr_2H_{3.8}$, $ZrV_2H_{4.9}$
AB	TiFe	CsCl, B2 (cP2)		TiFeH, $TiFeH_{1.9}$
A_2B	Mg_2Ni	Mg_2Ni (hP18)		Mg_2NiH_4
	Ti_2Ni	Ti_2Ni (cF96)		Ti_2NiH
A_2B–AB	Ti_2Ni–TiNi	Multiphase alloy (cF96 and mP4)		Ti_2NiH, TiNiH

hydride batteries, see Table 6 [33]. A schematic of a reversible battery with $LaNi_5$ metal hydride and NiOOH electrodes in a KOH electrolyte is shown in Fig. 10 [34]. NiMH batteries comprise more than 30% of a $6 billion market for rechargeable batteries used in many portable electronic devices such as cell phones and laptop computers [34–36]. Advantages of NiMH batteries include higher storage capacity than Pb-acid and Ni-Cd batteries, less toxicity than lead and cadmium and lower cost than Li-ion batteries. However, NiMH batteries have lower energy density than Li-ion and high initial costs than Pb-acid and Ni-Cd batteries. New intermetallic alloys: $(Zr,Ti)(Ni,Cr,V,Mn)_2$ have been developed as hydrogen storage materials for possible use in electric cars. The capacity of batteries based upon the new compounds is 50% higher than for $LaNi_5$.

5. Summary

This paper has described a wide range of industrial applications of intermetallic compounds. Although many alloys with attractive high temperature strength and ductility have been developed, applications in aerospace have been sparse. This arises in part from the inherent conservatism of this industry when new materials are considered, especially in view of the lack of a large database for most intermetallics. Recent developments in processing (e. g. the use of powders to produce functionally gradient materials) suggests that intermetallics may be useful in structural applications where ceramics are now contemplated for use. Fortunately, numerous non-structural applications that exploit the electrical, thermal, magnetic and corrosion properties of intermetallics have been identified. Even in these applications, continuing research to improve strength, ductility and toughness will add to the usefulness of intermetallics. A summary of new alloys described in this paper, together with their major attributes, appears in Table 6. While alloy design principles are well understood, better computational techniques to aid in alloy development are needed. The major impact of processing techniques on mechanical properties demonstrates a need to optimize processing parameters to provide the best balance of strength and toughness or ductility. Further work on suppressing environmental embrittlement at low temperatures is desirable, as is the improvement of high temperature oxidation resistance for alloys slated for structural applications. The establishment of property databases and the ability to demonstrate scale-up of laboratory research to production are prerequisites for widespread application of intermetallic compounds. Finally, while not addressed in this paper, there is continuing necessity to hold down costs of finished products, notwithstanding the advantages of using intermetallics with abundant, low-cost elements such as aluminum and silicon.

Acknowledgements

This research was sponsored by the Division of Materials Science and Engineering, U.S. Department of Energy under contract DE-AC05-00OR22725 with UT-Battelle, LLC.

References

[1] Sikka VK, Santella ML, Seindeman RW, Aramayo G. Intermetallic alloy development and technology transfer, pp. 89–107 Advanced Industrial Materials (AIM) Program Annual Progress Report for FY 1998, May 1999, ORNL/TM-1999/83, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1999. p. 89–107.
[2] Deevi SC, Sikka VK. Intermetallics 1996;4:357–75.
[3] Walston WS, Darolia R. In: Nathal MV et al. editors. Structural intermetallics 1997. Warrendale, USA: TMS, 1997. p. 613.
[4] Stoloff NS, Liu C.T. In: Stoloff NS, Sikka VK, editors. Physical metallurgy and processing of intermetallic compounds. New York: Chapman and Hall, 1996. p.159.
[5] Zheng M, He Y, Rapp RA. Proceedings of the 11th Annual Conference on Fossil Energy Materials, ORNL/FMP-97/1, May 1997, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1997.
[6] Sikka VK. In Intl. Symp. on Nickel and Iron Aluminides; Materials Park, USA, ASM, 1997. p. 361.
[7] Li D., Shan A., Liu Y., Lin D. Scripta Metall Mater.1995;33:681.
[8] Lin D, Shan A, Li D. Scripta Metall Mater 1995;31:1455.
[9] Austin CM, Kelly TJ, Mcallister KJ, Chesnutt JC. Structural intermetallics. TMS Warrendale, USA.
[10] Nazmy M., Noseda C., Staubli M., Phillipsen B., Processing and design issues in high temperature materials. Warrendale, USA: TMS 1997, p. 159.
[11] Tomasi A., Noseda C., Nazmy M., Gialanella S. MRS Symp. Proc. 460, Pittsburgh USA, 1997, p. 225.
[12] Nieh TG, Wadsworth J. Intl Mater Reviews 1999;44(2):59.
[13] Schimansky FP, Meyer MK, Gerling R. Intermetallics 1999;7:1275.
[14] Akinc M, Meyer MK, Kramer MJ, Thom AJ, Hoebsch JJ, Cook B. Mater Sci and Engng 1999;A261:16.
[15] Sadananda K, Feng CR, Mitra R, Deevi SC. Mater Sci Engng 1999;A261:223.
[16] Yamada K, Kamiya N. Mater Sci Eng 1999;A261:270.
[17] FGM-II, Research Activity Reports. FGM News 1996;30:16–25 (in Japanese).
[18] Washburn ME, Patent No. 5,045,237, 3 September 1991.
[19] Crandall WB, Shipley LE, US Patent No. 3,875,476, 1 April 1975.
[20] Bartlett AH, Castro RG, Buff DP, Kung H, Petrovic JJ, Zurceki Z. Industrial Heating, January 1996.
[21] Nowotny H, Kimakopoulou E, Kudielka H. Mh Chem 1957;88:180–92.
[22] Meyer MM, Kramer MJ, Akinc M. Intermetallics 1996;4:273.
[23] Perepezko J, Nunez CA, Yi SH, Thoma DJ. High temperature ordered intermetallic compounds. MRS Symp Proc 1997;460:3.
[24] Bewlay BP, Lewandowski JJ, Jackson MR. Journal of Metals 1997;49(8):44.
[25] Oliver WC, Liu CT. Improved mechanical properties of alloys based on Ni_3Si. ORNL/TM-12154, November 1992.
[26] Liu CT, George EP, Oliver WC. Intermetallics 1996;4:77.

[27] Zhu J., Liu C.T. Oak Ridge National Laboratory, 1999, unpublished.
[28] Kumar KS. In: Westbrook JH, Fleischer RL, editors. Intermetallic compounds—principles and practice, Vol. 2. 1995. p. 211.
[29] Deevi S.C., Sikka V.K., US Patent No. 5,965,274, 12 October 1999.
[30] Wolf RH, Heuer AH. J Microelectromech Syst 1995;4:206–12.
[31] Stadelmaier, HH, Reinsch B, In: Westbrook JH, Fleischer RL, editors. Intermetallic compounds—principles and practice, Vol. 2, John Wiley, 1995. p. 303.
[32] Horton JA, Wright JL, Herchenroeder JW. IEEE Trans On Magnetics 1996;32:4374–6.
[33] Schlapbach L, Meli F, Zuttel A. In: Westbrook JH, Fleischer RL, editors. Intermetallic compounds—principles and practice, Vol. 2, John Wiley, 1995. p.475.
[34] GM Ovonic—The NiMH Choice. Troy, MI: G.M. Ovonic L. L.C., 1996.
[35] Ovshinsky SR, Fetcenko MA, Ross J. Science 1993;260:76.
[36] George EP, Unpublished results, Oak Ridge National Laboratory, Oak Ridge, TN, 1999.

ELSEVIER

Intermetallics 8 (2000) 353–357

Intermetallics

Interaction of H_2O and O_2 with Ni_3Fe and their effects on ductility

J. Wang [a,*], W.J. Chia [a,1], Y.W. Chung [a], C.T. Liu [b]

[a] *Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA*
[b] *Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6115, USA*

Received 13 August 1999; accepted 20 September 1999

Abstract

Possible reasons for the high ductility of Ni_3Fe and its insensitivity toward the testing environment have been investigated. Thermal desorption experiments have shown that water dissociates on clean Ni_3Fe surfaces to produce atomic hydrogen. Ductility measurements of cast and cold-rolled polycrystalline Ni_3Fe demonstrated that the reduced ductility was obtained only when the testing was performed with oxygen carefully removed. X-ray photoemission studies indicate that oxygen interacts with water to form hydroxyls, thereby suppressing the production of atomic hydrogen. Hydrogen desorption from the Ni_3Fe surface requires a lower activation energy, resulting in a smaller surface hydrogen concentration at a given temperature. Hence it is possible that there is insufficient hydrogen to cause the nucleation and growth of brittle cracks for severe embrittlement. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Intermetallics, miscellaneous; B. Brittleness and ductility; B. Environmental embrittlement; B. Surface properties

1. Introduction

Recent experimental results have shown that moisture-induced embrittlement is the major cause of the low ductility and grain boundary failure of many $L1_2$ ordered intermetallic alloys in ambient air [1,2]. The proposed embrittlement mechanism involves chemical reactions between the reactive element in the intermetallic alloy (e.g. Al in Ni_3Al) and water vapor in the atmosphere to produce atomic hydrogen, which embrittles the crack tip during tensile testing. Earlier work by Liu and Takeyama showed that Ni_3Fe (24 at% Fe) achieves over 40% ductility in vacuum, air and water [3]. This high ductility and insensitivity toward the testing environment are unexpected, since water dissociation on iron surfaces to produce atomic hydrogen is well known [4–7]. In addition, it was found that Ni_3Fe shows hydrogen embrittlement at room temperature when there is sufficient hydrogen fugacity [8]. In this article, we will explore the possible reasons why Ni_3Fe is ductile in air or water by (i) examining the surface interactions of H_2O and O_2 with Ni_3Fe and (ii) performing ductility measurements under well-controlled environments.

2. Experimental procedure

All surface science experiments were performed in an ultrahigh vacuum chamber with a working base pressure of 5×10^{-8} Pa equipped with sputtering, gas handling, Auger and X-ray photoelectron spectroscopy (XPS) capabilities. Photoelectron spectra were collected with a double-pass cylindrical mirror analyzer at a pass energy of 50 eV using the Al K_α radiation at 1486.6 eV. All binding energies were determined using the Ni $2p_{3/2}$ peak at 852.3 eV as a reference. All gas exposures are expressed in Langmiur units where 1 L = 1.33×10^{-4} Pa-sec (10^{-6} torr-sec). The ionization gauge was turned off during gas exposure to minimize activation of gas species by the hot filament.

Thermal desorption spectroscopy (TDS) experiments were performed at a heating rate of 1∼2 K/sec. We used D_2O in water exposures to avoid interference from background H_2O and H_2. H and D are considered to be equivalent in the following discussion. D_2O was purified by multiple freeze-pump-thaw cycles. It was admitted into the chamber through a needle doser placed 1 cm

* Corresponding author.
[1] Now at Western Digital Corporation, USA.

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(99)00091-6

from the sample. Oxygen was introduced from a gas bottle (99.9999% purity) without further treatment. A glass envelope was mounted in front of the ionizer of the mass spectrometer. During thermal desorption, the sample was placed in front of this envelope to minimize detection of background desorption.

The sample can be heated to 1000 K using a tungsten wire radiation heater mounted behind the sample or cooled to ~100 K via a copper contact with a liquid-nitrogen reservoir. Temperatures were measured using a K-type (chromel-alumel) thermocouple spot-welded to the sample. Surface cleaning was done with several cycles of Ar-ion bombardment at 2 kV, 45° angle of incidence, followed by annealing at ~700 K for 30 min. Auger electron spectroscopy was used to confirm the surface cleanliness. The polycrystalline Ni_3Fe (25 at% Fe) sample of ~1 $cm^2 \times 1$ mm was mechanically polished using standard metallographic techniques, using 0.05 μm alumina for the final step. An ordering heat treatment was carried out in ultrahigh vacuum at 743 K for 100 h. The average grain size was 300–500 μm.

Two types of Ni_3Fe specimens were prepared for tensile testing. The first was a polycrystalline alloy of Ni-25 at% Fe produced by arc-melting and drop casting, using commercially pure nickel and iron. After an ordering treatment at 743 K for 100 h in vacuum, tensile specimens of the appropriate geometry were fabricated by electric-discharge-machining. The second was cold-rolled Ni_3Fe prepared from an Ni_3Fe ingot containing 24 at% Fe by repeated cold forging and rolling at room temperature, with intermediate anneals at 1273 K in vacuum. Tensile specimens with a gauge section of $12.7 \times 3.2 \times 0.7$ mm were stamped from the sheet stock and then given a duplex heat treatment in vacuum: 2 d at 1273 K for recrystallization and grain growth, and 200 h at 743 K for atomic ordering. The average grain size was about 150 μm. Tensile tests were performed on an Instron testing machine. In order to remove oxygen in water vapor, a set of specimens were tested in a vacuum (1×10^{-6} torr) system back-filled with a low pressure of pure water (12 torr). The ductility value for cast specimens is the average of 2–3 measurements, while that for cold-rolled specimens is obtained for one single measurement in each environment. Fractured surfaces were examined in a Hitachi S-3500N scanning electron microscope (SEM) operated at 20 kV.

3. Results and discussion

Table 1 shows that water can reduce the ductility of cast Ni_3Fe only when the strain rate is sufficiently small and water is thoroughly purged of dissolved oxygen. Similarly, as shown in Table 2, when dissolved oxygen is removed progressively via purging water or exposure to pure water vapor in a vacuum system, the ductility of cold-rolled Ni_3Fe is also reduced. Note that the cold-rolled recrystallized samples show higher ductility than the cast samples. There are two possible reasons for this: (a) finer grain size resulting in less stress concentration at grain boundaries; (b) different grain boundary chemistry (such as less segregation of impurities) in these two alloys. It is interesting to note that Takasugi and Hanada [9] also reported high tensile ductility (~50%) for recrystallized Ni_3Fe.

Table 1
Ductility of cast Ni_3Fe measured in different environments and strain rates

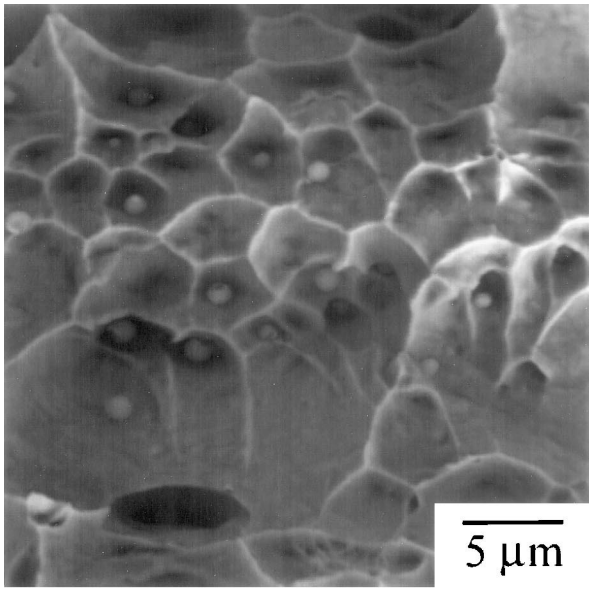
Strain rate (/sec)	Environment	Ductility (%) ±1.5%
1.3×10^{-4}	Air	15.3
1.3×10^{-4}	6.5×10^{-3} Pa	15.0
6.5×10^{-5}	Water, purged with air	14.7
1.3×10^{-3}	Water, purged with He for 1 h	14.2
6.5×10^{-5}	Water, purged with He for 1 h	9.0

Table 2
Tensile ductility of cold-rolled Ni_3Fe measured in different environments with strain rate 3.3×10^{-4}/sec

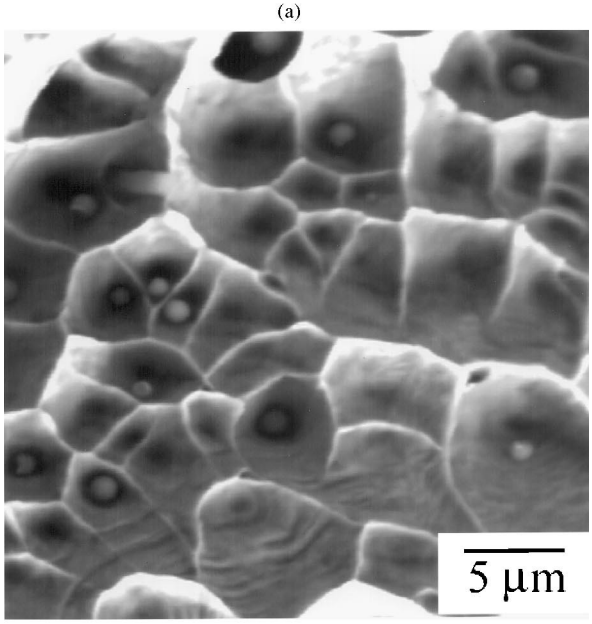
Environment	Ductility (%)
Air	54.8
Water	56.7
Water, purged with He for 2 h	53.6
Water, purged with He for 16 h	52.2
Water vapor (12 torr)	48.1

The SEM fractographs of cast Ni_3Fe tested in both 6.5×10^{-3} Pa and water purged with helium show ductile dimple fracture (Fig. 1). The corrosion dots on the fracture surface are believed to result from the exposure to air for an extended period. It has already been reported that cold-rolled Ni_3Fe exhibits ductile dimple fracture in vacuum and water at room temperature [3].

Temperature-programmed-desorption experiments demonstrate that water readily dissociates on clean polycrystalline Ni_3Fe surfaces, resulting in the evolution of hydrogen peaking at 290 K (Fig. 2). Fig. 3 compares the thermal desorption spectra of D_2 from polycrystalline Ni_3Fe after exposure to 45 L of D_2O and 45 L $D_2O + 5$ L O_2 mixture. Note the strong suppression of D_2 production in the presence of oxygen. This suppression may be partly due to physical site-blocking. That is, adsorbed oxygen blocks some active sites on the Ni_3Fe surface for water adsorption so as to avoid subsequent dissociation. Another possible reason is suggested by the series of O 1*s* X-ray photoelectron spectra obtained after sequential oxygen and D_2O exposure to the Ni_3Fe surface (Fig. 4). Oxygen exposure at 230 K results in an O 1*s* peak at 530.8 eV characteristic of atomic oxygen. Subsequent adsorption of D_2O



(a)



(b)

Fig. 1. SEM fractograph of cast Ni_3Fe tested (a) in 6.5×10^{-3} Pa vacuum at strain rate 1.3×10^{-4}/sec, and (b) in water purged with helium for 1 h at strain rate 6.5×10^{-5}/sec.

removes atomic oxygen, and the O 1*s* peak binding energy shifts to 532.3 eV, which is due to OD formation (Fig. 4c). In light of above X-ray photoemission studies, we believe that adsorbed oxygen promotes the formation of surface hydroxyls, most likely via an abstraction reaction,

$$D_2O_{ads} + O_{ads} \rightarrow 2OD_{ads} \quad (1)$$

Upon further warming, one observes the desorption of water vapor and the formation of atomic oxygen (Fig. 4e), suggesting the following reaction,

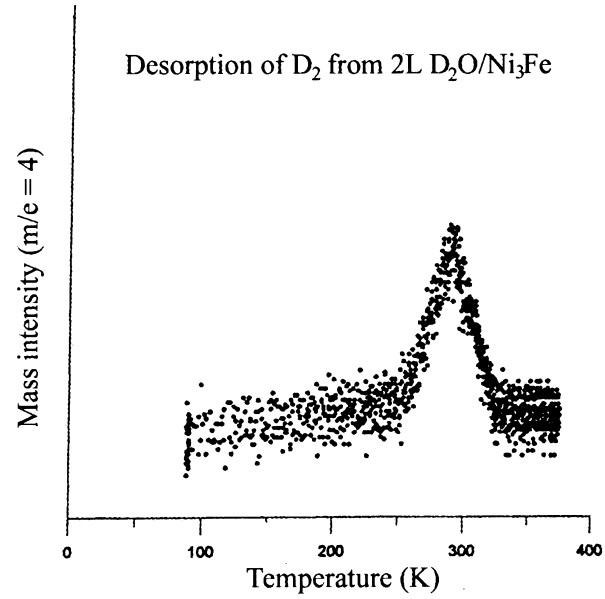


Fig. 2. Thermal desorption spectrum of D_2 (m/e = 4) from polycrystalline Ni_3Fe after 2 L D_2O exposure at 77 K.

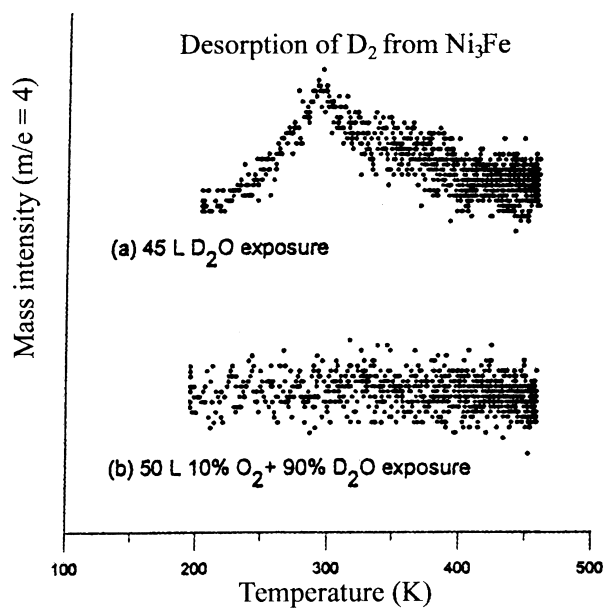


Fig. 3. Thermal desorption spectra of D_2 (m/e = 4) from polycrystalline Ni_3Fe after exposure to (a) 45 L D_2O, and (b) 45 L D_2O + 5 L O_2 at 180 K.

$$2OD_{ads} \rightarrow D_2O \uparrow + O_{ads} \quad (2)$$

An isotope scavenging experiment was carried out to confirm the occurrence of reaction (2). We repeated the above thermal desorption experiments by first exposing the clean Ni_3Fe surface to the oxygen-18 isotope, followed by D_2O^{16}. Thermal desorption was then performed, with the mass spectrometer monitoring the

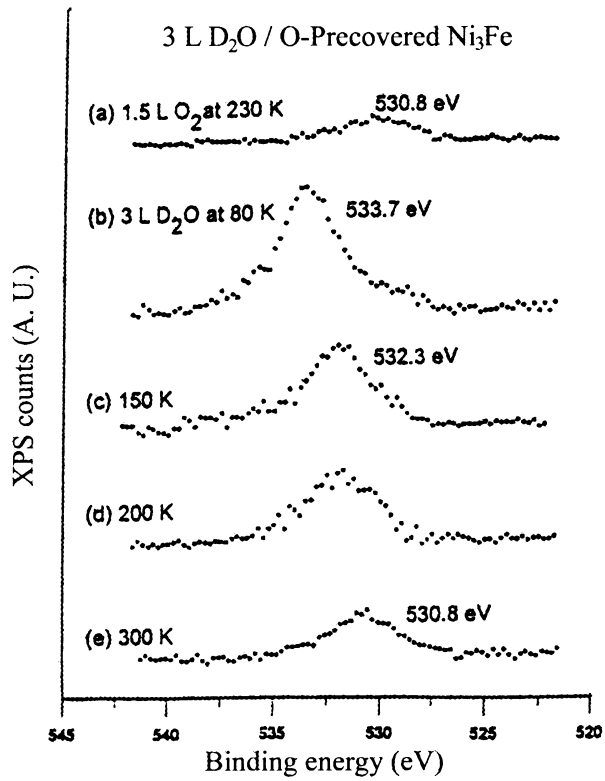


Fig. 4. O 1*s* spectra from Ni_3Fe taken at different temperatures after exposure to 1.5 L oxygen at 230 K, followed by 3 L D_2O at 80 K and then warming to 300 K.

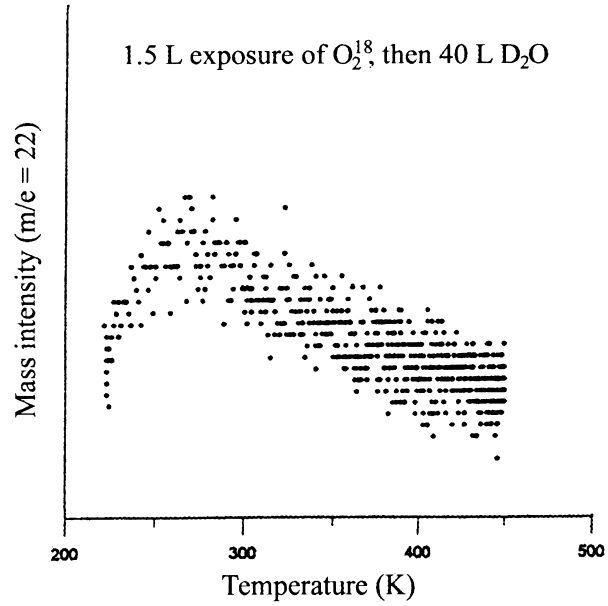


Fig. 5. Thermal desorption spectrum of D_2O^{18} (m/e = 22) from polycrystalline Ni_3Fe after 1.5 L O_2^{18} exposure at 250 K, followed by 40 L D_2O^{16} at 200 K.

signal of D_2O^{18} (m/e = 22). Fig. 5 shows clearly the evolution of D_2O^{18}, which verifies reactions (1) and (2).

Even when oxygen is eliminated, Ni_3Fe exhibits high ductility in water vapor (Tables 1 and 2). One possible reason is insufficient hydrogen concentration to nucleate and propagate brittle cracks for severe embrittlement. This low hydrogen concentration is due to rapid hydrogen desorption. Our previous studies indicate that water dissociates on clean Ni_3(Al, Ti) (100) surface to produce atomic hydrogen [10]. In this case, hydrogen evolution peaks at about 330 K. Fig. 2 shows that the peak hydrogen desorption temperature is only 290 K for Ni_3Fe (actually, the desorption starts at ~250 K). This implies that less thermal energy is needed for hydrogen desorption from Ni_3Fe surfaces, resulting in a smaller surface concentration of hydrogen at a given temperature. Therefore, there may be insufficient hydrogen available for brittle crack nucleation and growth.

It is also possible that grain boundaries of Ni_3Fe are stronger than those of other ordered intermetallic compounds which suffer severe environmental embrittlement. If grain boundaries remain intact during tensile testing, fresh exposed surfaces (cracks) will partly or mostly be transgranular (depends on the grain size), as shown in Fig. 1 and previous report [3]. In this case, the atomic hydrogen concentration at crack tips (produced via interaction of water with clean Ni_3Fe surfaces) is controlled mainly by the much slower bulk diffusion (cf. grain boundary diffusion) and is likely insufficient to cause embrittlement.

4. Conclusions

Careful ductility measurements performed on polycrystalline Ni_3Fe under different environments and strain rates reveal a subtle environmental effect: one obtains reduced ductility only when oxygen is carefully removed from the testing environment. Thermal desorption experiments clearly indicate that water does dissociate on clean polycrystalline Ni_3Fe surfaces. X-ray photoemission studies demonstrate that adsorbed oxygen interacts with water to form hydroxyls on Ni_3Fe surface, thus suppressing the formation of atomic hydrogen. Compared with Ni_3Al, Ni_3Fe has a lower activation energy for hydrogen desorption, resulting in lower surface hydrogen concentration. This provides a natural explanation for the high tensile ductility of Ni_3Fe under pure water vapor environments.

Acknowledgements

This work is jointly supported by the National Science Foundation, grant No. DMR-9713052, and the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Research Corporation.

References

[1] Liu CT. Scr Met 1992;27:25 .
[2] George EP, Liu CT. Mat Res Soc Symp Pro 1995;364:1131.
[3] Liu CT, Takeyama M. Scr Met 1990;24:1583.
[4] Thiel PA, Madey TE. Surf Sci Reports 1987;7:211.
[5] Simmons GW, Dwyer DJ. Surf Sci 1975;48:373.
[6] Dwyer DJ, Simmons GW, Wei RP. Surf Sci 1977;64:617.
[7] Doering DL, Semancik S, Madey TE. Surf Sci 1983;133:49.
[8] Camus GM, Stoloff NS, Duquette DJ. Acta Met 1989;37:1497.
[9] Takasugi T, Hanada S. Intermetallics 1994;2:225.
[10] Chia WJ, Chung YW. Intermetallics 1995;3:505.

Crystallization and nanoindentation behavior of a bulk Zr–Al–Ti–Cu–Ni amorphous alloy

J.G. Wang[a)]
Department of Chemical and Biochemical Engineering and Materials Science, University of California, Irvine, California 92697

B.W. Choi and T.G. Nieh
Chemistry and Materials Science, Lawrence Livermore National Laboratory, P.O. Box 808, L-350, Livermore, California 94551-9900

C.T. Liu
Metals and Ceramics Division, P.O. Box 2008, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6376

(Received 26 May 1999; accepted 7 January 2000)

The crystallization and nanoindentation behavior of a Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) bulk amorphous alloy (BAA) were studied. Resulting from the kinetic nature of phase transformation in multicomponent alloys, the crystallization path is complex. Despite the complexity of different crystallization paths, the main final crystallized product in the Zr-based BAA is Zr_2Cu. Young's modulus and hardness of the BAA were found to increase with an increase in annealing temperature. The observed mechanical properties were correlated with the microstructure of the material. Also, in the present paper, both the observed crystallization and nanoindentation behavior are compared with existing data. Zr-based BAAs exhibit a ratio of hardness to Young's modulus (H/E ratio) of about 1/10, suggesting the interatomic bonding in the alloys is close to being covalent.

I. INTRODUCTION

Bulk amorphous alloys (BAAs) with complex multicomponent chemistry and excellent glass-forming ability have recently been developed.[1,2] The production of these bulk amorphous alloys is the subject of considerable scientific interest because of the promising properties shown by these materials and their possible applications as engineering materials. Several alloy families, for example, La base,[3] Mg base,[4] Zr base,[3,5–8] and Fe base[9] have been developed. Among them, Zr-based BAAs, e.g., Zr–Al–TM,[3,5] Zr–Ti–TM–Be,[6] and Zr–Ti–Al–TM[7,8] are most interesting because of having a broad supercooled liquid region.[10] Zr–10Al–5Ti–17.9Cu–14.6Ni (in at.%, hereafter) is a new BAA composition developed.[11] The effect of oxygen impurity on the undercooling and crystallization kinetics,[11] the influence of testing environment on the room-temperature mechanical properties of this alloy,[12] and the superplastic behavior in the supercooled liquid region[13] have been reported. It was found that the mechanical behavior strongly depends on the microstructure of the material. The presence of a minor amount of crystalline phase can significantly embrittle the alloy. In the present paper, a systematic study of the crystallization behavior and its influence on the nanoindentation behavior of this alloy are carried out.

II. EXPERIMENTAL PROCEDURE

The Zr-based alloy with a nominal composition of Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) was prepared by arc melting and drop casting, as described previously.[12] The as-cast amorphous samples used in the present study are of a cylindrical shape with a diameter of 7 mm. The amorphous nature of the cast samples was verified by Rigagu's RU300 x-ray diffractometry (XRD), with a rotating anode of Cu as a target material, and transmission electron microscopy (TEM). The thermal properties of the alloy were measured using differential scanning calorimetry (DSC) in a Du Pont 9900 Thermal Analysis System with different heating rates. A typical DSC run involves heating a sample from room temperature to 923 K. The crystallization behavior of the annealed samples was studied by using XRD and TEM. Annealing treatments were carried out *in situ* in the DSC cell with a

[a)]Address all correspondence to this author.
e-mail: wangjg@uci.edu

 © 2000 Materials Research Society

heating rate of 20 K/min under flowing N_2. TEM specimens were prepared using standard twin-jet electropolishing in a solution containing one part sulfuric acid plus seven parts methanol at −25 °C and then examined using a JEOL-200CX microscope.

Nanoindentation tests were carried on the as-cast and annealed samples with a Nano Indenter XP®. Hardness and elastic modulus were measured using the continuous stiffness option, which yields elastic modulus and hardness as a function of indentation depth.[14] A Berkovich indenter was used in the experiments. Five types of experiments, hardness test at a load, hardness test at a depth, continuous stiffness test at a depth, continuous stiffness at a load, and a four times load/unload experimental procedure,[14] were performed at different loads and depths. Ten indentations were made for each test on each specimen, with most of the results presented here representing averages for the group.

III. EXPERIMENTAL RESULTS

A. As-received BAA

The XRD and TEM results obtained from the as-cast amorphous sample are presented in Fig. 1. The XRD pattern consists of only a broad amorphous peak with a maximum intensity at the wave vector ($K_p = 4\pi\sin\theta/\lambda$) of 26.2 nm^{-1} and no crystalline phases are discernible. The electron-diffraction pattern reveals the presence of a diffuse amorphous ring without any crystalline feature. It is consistent with that reported previously;[12] namely, the as-cast alloy contains essentially no crystalline particles and cast cavity.

B. DSC measurements

The DSC curves of the Zr-based BAA at various heating rates are shown in Fig. 2. The amorphous alloy shows an endothermic reaction caused by the glass transition, followed by a supercooled liquid region and, then, an exothermic reaction caused by crystallization. In the figure, the onset temperatures for the glass transition and crystallization are denoted by T_g and T_x, respectively. As expected, both T_g and T_x increase with an increase in heating rate. The temperature interval between T_g and T_x, i.e., the supercooled liquid region, $\Delta T_x = (T_x - T_g)$, also increases with increasing heating rate, clearly indicating that glass transition and crystallization of the BAA are kinetic processes.[15] Table I summarizes the thermal parameters of T_g, T_x, ΔT_x at different heating rates. These data are in good agreement with the DSC results of Lin *et al.*[11] who reported that at a heating rate of 20 K/s glass transition occurs at temperatures between 682 and 713 K, corresponding to a reduced glass-transition temperature of about 0.64 to 0.66 (melting temperature 1069 K).

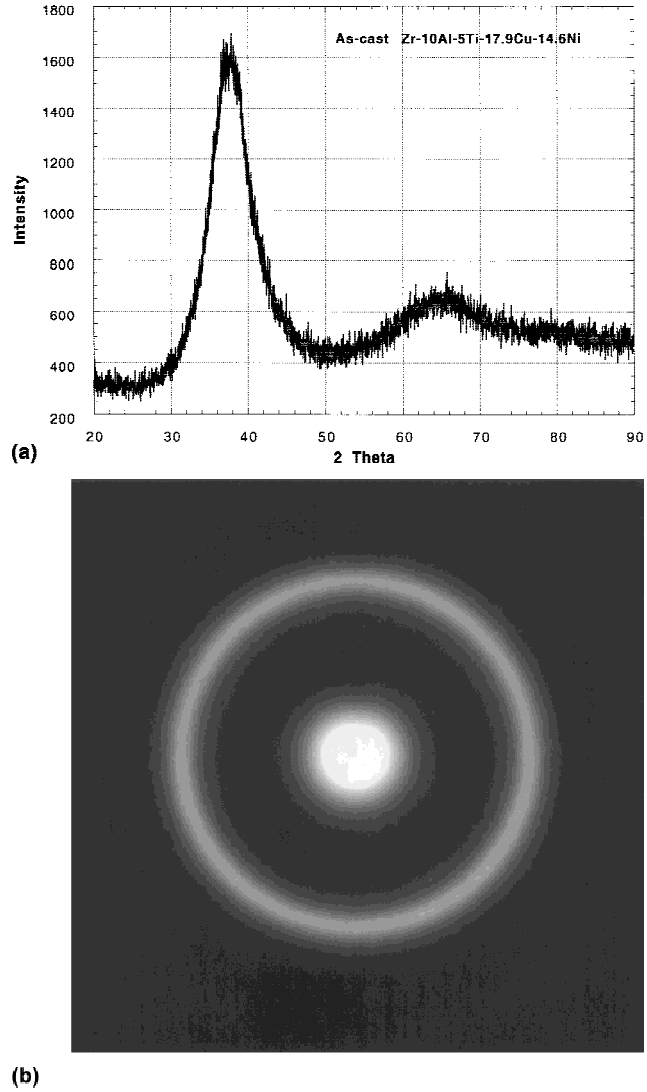


FIG. 1. (a) X-ray and (b) TEM diffraction patterns of the as-cast amorphous $Zr_{52.5}Al_{10}Ti_5Cu_{17.9}Ni_{14.6}$ alloy.

C. Crystallization behavior

Annealing treatment was performed *in situ* in a DSC cell with a heating rate of 0.33 K/s under flowing N_2 to minimize kinetic effects. Annealing temperatures were selected to be 683, 783, and 893 K, corresponding to the supercooled liquid region, after the first crystallization peak, and after the second crystallization peak, in the DSC curves (Fig. 2), respectively. After reaching the predetermined temperatures, the samples were rapidly cooled to room temperature (drops from 923 to 658 K within 600 s after shutting off the electric power). The annealed samples were then examined by XRD. Figure 3 shows the XRD patterns obtained from as-received BAA sample [Fig. 3(a)], samples annealed at 683 K [Fig. 3(b)], 783 K [Fig. 3(c)], and 893 K [Fig. 3(d)], respectively. Also included in Fig. 3(e) is the XRD pattern from a sample annealed at 923 K for 1 h. The 683-K–

annealed sample does not reveal any discernible crystallization, although the diffuse amorphous ring is sharper than that from the as-cast sample. Apparently, the amorphous glass relaxes into a supercooled liquid state. Annealing at 783 K leads to a partial crystallization and the formation of tetragonal Zr_2Ni (space group *I*4/*mcm*, $a = 0.649$ nm, $c = 0.528$ nm) and [Ni,(Zr,Ti)] with a hexagonal structure. It is noted that, at lower cooling rates, e.g., 0.167 and 0.33 K/s, two subpeaks appear and are actually associated with the first crystallization peak in the DSC curve. These two subpeaks may correspond to these two crystalline phases. The formation of primary Zr_2Ni during the crystallization of amorphous $Zr_{60}Al_{10}Cu_{20}Ni_{10}$[8] and $Zr_{55}Al_{10}Cu_{30}Ni_5$[16] has been reported.

Annealing at 893 K results in two additional phases, $Zr(Al,Ni)_2$ and NiZr. $Zr(Al,Ni)_2$ is a metastable phase with a cubic structure (space group *F*d3*m*), and NiZr is an orthorhombic phase with a space group *Cmcm*. Annealing at a higher temperature of 923 K for 1 h produces a fully crystallized sample. The crystallized products include Zr_2Cu (main phase) with a tetragonal structure (space group *I*4/*mmm*, $a = 0.322$ nm, $c = 1.118$ nm), $Ni_{10}Zr_7$ with an orthorhombic structure (space group *A*b2*a*, $a = 0.922$ nm, $b = 1.229$ nm, $c = 0.915$ nm), Al_3Zr_4 with a hexagonal structure (space group *P*6, $a = 0.5432$ nm, $c = 0.5389$ nm), and probably a trace of Zr_2Ni. No appreciable diffraction peaks corresponding to Ti-based compounds are seen. Because Zr and Ti form a complete solid solution over the entire composition range, Ti may have dissolved in a substitutional form in the Zr-based compounds.

A bright-field TEM image and selected-area electron-diffraction pattern of the as-cast sample and samples annealed at different temperature are shown in Fig. 4. Both the as-cast sample and the 683 K–annealed sample show a uniform amorphous microstructure [Fig. 4(a) and Fig. 4(b)]. The halo ring in the diffraction pattern of the 683 K annealed sample is apparently sharper than that of the as-cast sample, as a result of relaxation of the glassy state. In contrast, the samples annealed at 783 and 893 K consist mainly of very fine Zr_2Ni grains with a size of less than 10 nm at 783 K and 40 nm at 893 K [Figs. 4(c) and 4(d)]. The diffraction patterns contain distinct triple rings resulting from the three diffraction peaks of the Zr_2Ni phase. The volume fraction of this nanocrystalline phase is estimated to be 80% at 783 K and 95% at 893 K from TEM images. The microstructure of the sample annealed at 923 K for 1 h, shown in Fig. 4(e), is a polycrystal with an average grain size of about 250 nm. The diffraction pattern consists of spotty reflection rings resulting from the presence of the Zr_2Cu phase; it no longer has the triple ring characteristic of the Zr_2Ni phase.

D. Nanoindentation behavior

Representative load-displacement plots of indentations made at a 6-mN indentation load on the samples annealed at different temperatures are shown in Fig. 5. The experiments were carried out using a typical load-time sequence developed by Oliver and Pharr.[14] The indenter was loaded and unloaded four times in succession at a

TABLE I. Temperatures for the start (T_g^s) and end (T_g^e) of glass transition, crystallization (T_x), and supercooled range (ΔT_x) for amorphous Zr–10Al–5Ti–17.9Cu–14.6Ni alloy.

Heating rates (K/s)	T_g^s, K	T_g^e, K	T_x, K	ΔT_x, K
0.167	628	696	714	86
0.33	631	705	729	98
0.67	637	716	742	105
1.33	640	728	758	118

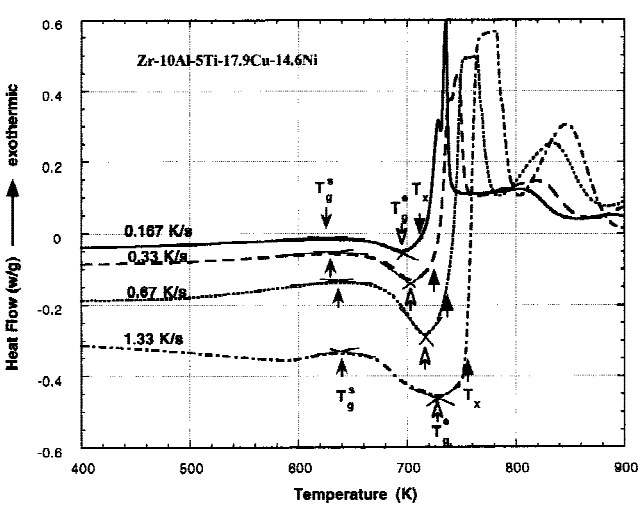


FIG. 2. DSC curves of the Zr-based BAA at various heating rates: (a) 0.167 K/s, (b) 0.33 K/s, (c) 0.67 K/s, and (d) 1.33 K/s.

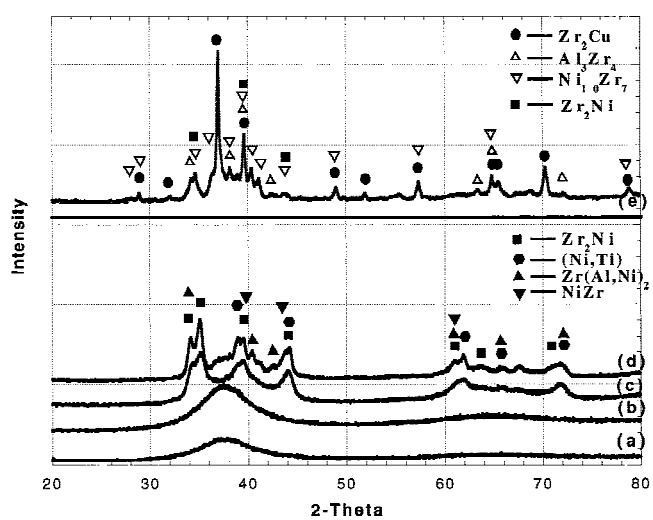


FIG. 3. XRD patterns from as-cast and annealed BAA samples: (a) as received, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h.

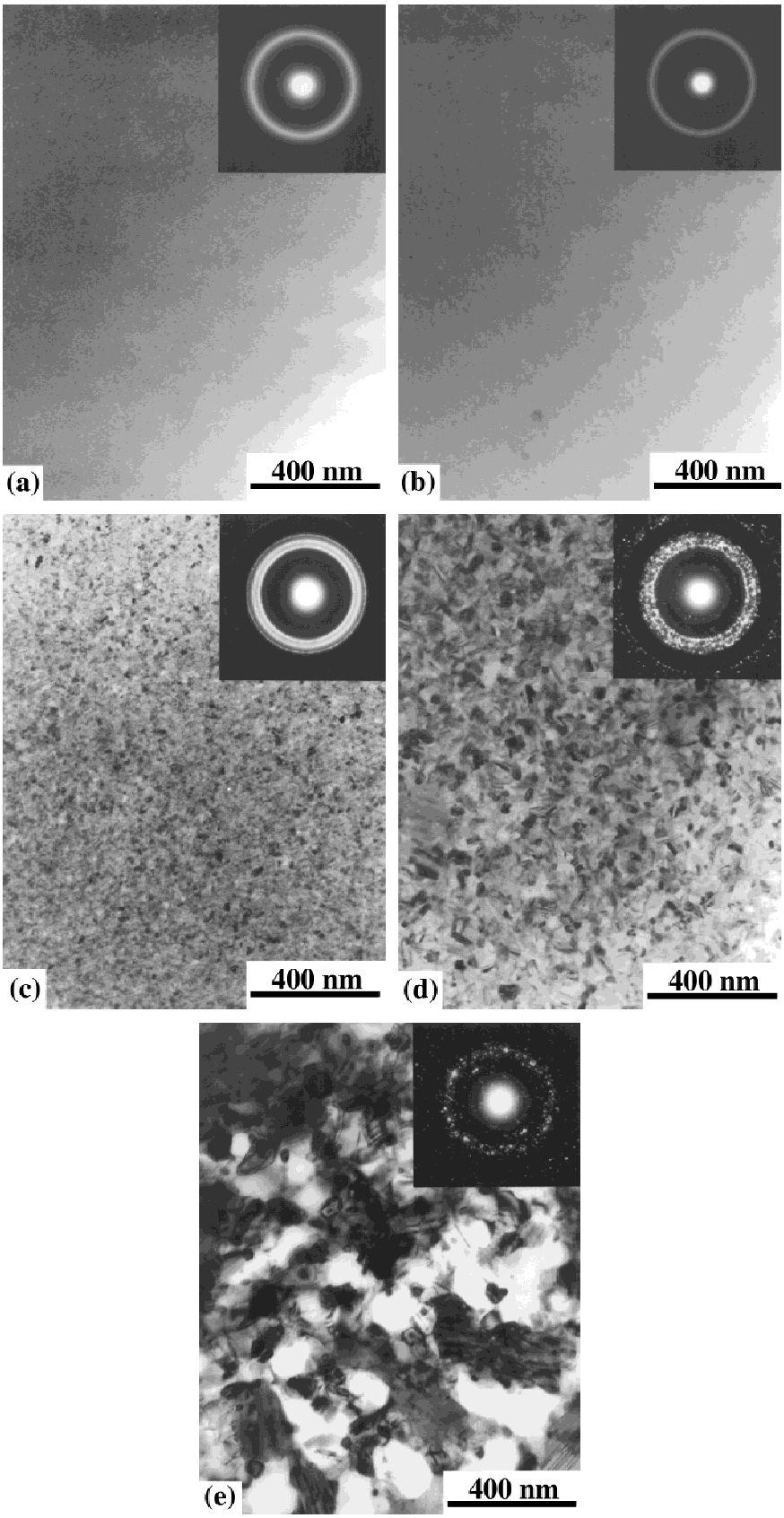


FIG. 4. Bright-field TEM images and selected-area electron-diffraction patterns of the as-cast and annealed BAA samples: (a) as received, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h.

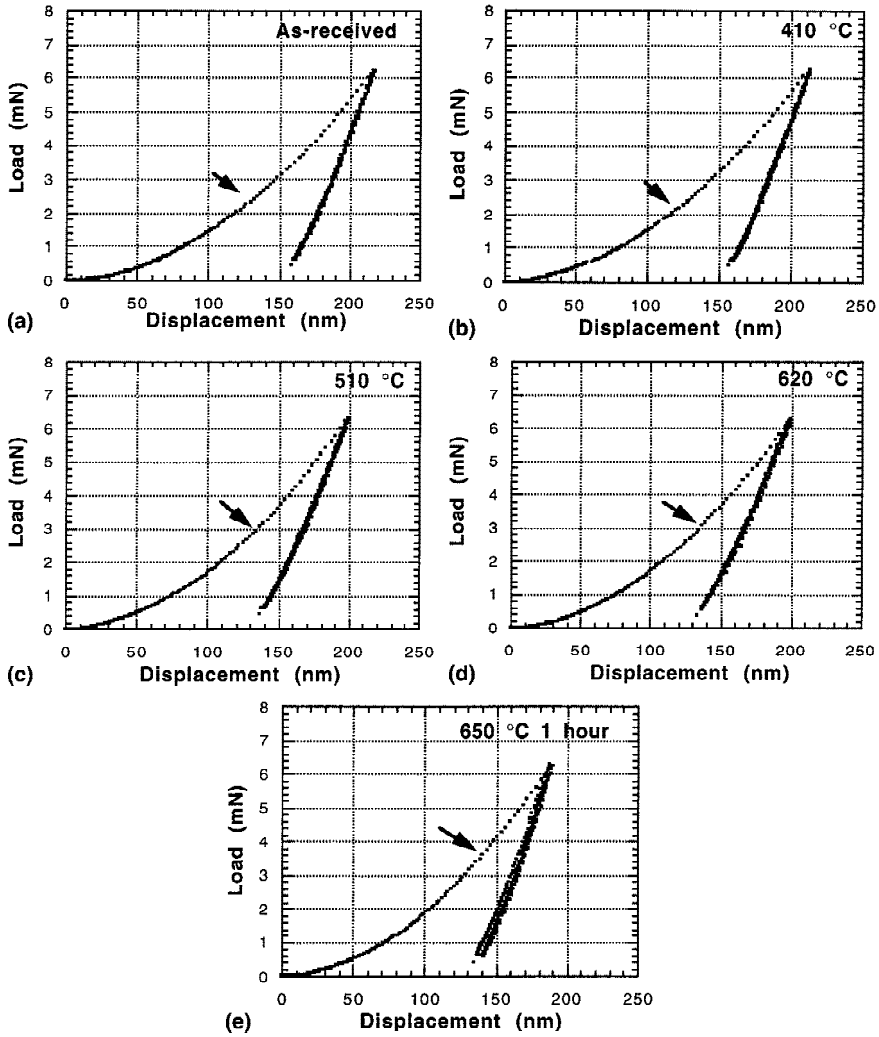


FIG. 5. Representative load-displacement plots for as-received and different-temperature-annealed BAA samples obtained from nanoindentation at 6.0-mN peak indentation load: (a) as received, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h. Locations on the loading curve marked by arrows show a displacement discontinuity or pop-in of the tip into material during the loading step in the indentation process.

constant loading rate/load ($1/P\ \mathrm{d}P/\mathrm{d}t$), because a constant loading rate/load ($1/P\ \mathrm{d}P/\mathrm{d}t$) results in a constant indentation strain rate ($1/\mathrm{h}\ \mathrm{d}h/\mathrm{d}t$).[17] A lower indentation depth indicates a higher hardness and a higher slope of the unloading curve indicates a higher stiffness or elastic modulus. As shown in Fig. 5, the sample annealed at 923 K exhibits the lowest indentation depth, about 186 nm, as compared to other samples. The shape of load-displacement curves and indentation depths obtained from the as-received sample are essentially the same as from the 683 K annealed sample. The load-displacement curve for the 783 K annealed sample is also similar to that for the 893 K annealed sample.

It is noted that load-displacement plots exhibit several displacement discontinuity or pop-in marks during the first loading of indentation. These pop-in marks indicate a sudden penetration of the tip into the sample. The non-uniform penetration is probably a result of either onset of a sudden plastic deformation or formation of cracks. The load value for the onset of pop-in increases with increasing annealing temperature, suggesting the as-cast material has the lowest resistance to plastic deformation or to crack formation.

Hardness and elastic modulus profiles as a function of indentation depth for the as-cast and annealed samples are shown in Fig. 6. These hardness and elastic modulus are also summarized in Table II. Here, data from two types of nanoindentation test are included: test at a fixed load and test at a fixed depth.

The as-cast BAA sample exhibits the lowest hardness of 6.3 GPa and the lowest elastic modulus of 112 GPa (at a load of 60 mN). In contrast, the sample annealed at 923 K/1h exhibits the highest elastic modulus of 150 GPa and the highest hardness of 8.6 GPa. In general, the hardness increases with increasing annealing temperature, and an increase in annealing temperature also

results in an increased elastic modulus. A close examination of the modulus data shows that the data can be divided into three ranges: the data of as-cast sample (~112 MPa); the data of 683, 783 and 893 K annealed samples (~119 MPa); and the data of 923 K/1h annealed sample (~150 MPa). Evidently, the as-cast sample has a uniform amorphous structure, and the 683 K annealed sample is also amorphous, but in a relaxed state (supercooled liquid). In comparison, the 783 and 893 K annealed samples consist of nanocrystals dispersed in a relaxed amorphous matrix. The 923 K/1h annealed sample is fully crystallized (main phase is Zr_2Cu). The hardness data can also be divided into three groups: the data of as-cast alloy (~6.3 MPa), the data of 683 K annealed sample (~6.9 MPa), and the data of 783, 893, and 923 K/1h annealed samples (~8.6–8.9 MPa). The relaxed amorphous material is slightly harder than the as-cast material, whereas the nanocrystal-containing materials have the highest hardness.

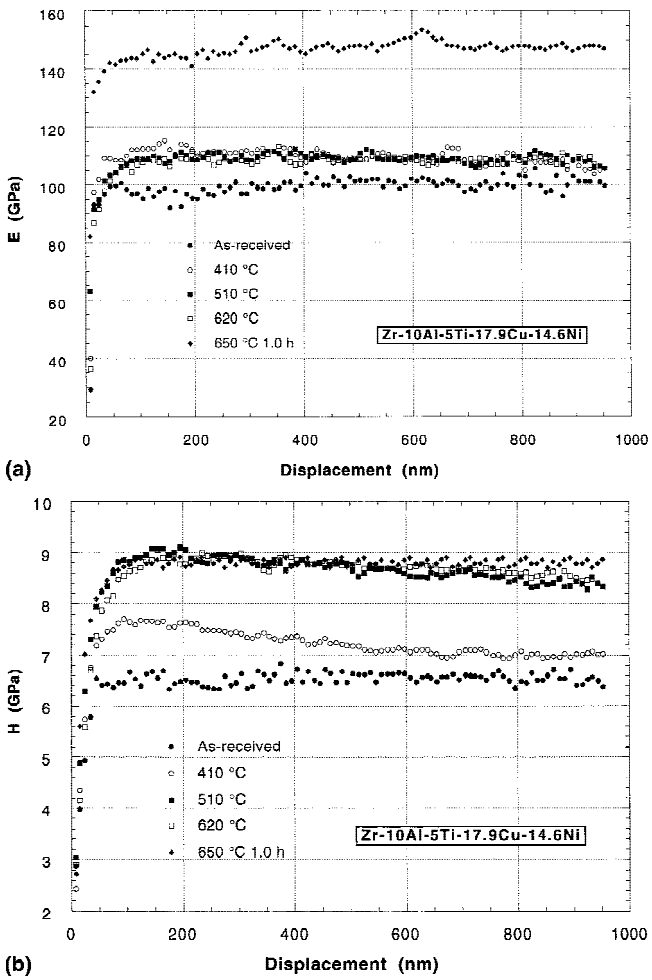


FIG. 6. (a) Elastic modulus and (b) hardness profiles as a function of indentation depth for the as-cast and annealed BAA samples.

TABLE II. Hardness and elastic modulus of as-cast and annealed BAA samples obtained from different nanoindentation test.

Sample	Test[a]	Depth (nm)	Load (mN)	Elastic modulus (GPa)	Hardness (GPa)
As-received	HL	733 ± 4	60	112 ± 6	6.3 ± 0.6
	HD	2000	474 ± 4	92 ± 4	6.2 ± 2
683 K annealed	HL	700 ± 3	60	119 ± 2	6.9 ± 0.1
	HD	2000	496 ± 8	100 ± 5	6.5 ± 0.2
783 K annealed	HL	661 ± 3	60	118 ± 2	8.3 ± 0.1
	HD	2000	590 ± 12	99 ± 4	7.8 ± 0.2
893 K annealed	HL	652 ± 4	60	119 ± 1	8.6 ± 0.1
	HD	2000	604 ± 10	99 ± 4	8.0 ± 0.2
923 K 1 h	HL	627 ± 4	60	150 ± 2	8.6 ± 0.2
	HD	2000	619 ± 17	142 ± 1	8.3 ± 0.1

[a]HL, hardness test at a load; HD, hardness test at a depth.

IV. DISCUSSION

A. Crystallization

Thermal stability and crystallization behavior of several Zr-based BAAs have been examined.[6,8,18,19] Table III lists the reported data of T_g, T_x, and ΔT_x for these Zr-based BAAs. Compared with the base alloys Zr–Al–Cu–Ni,[8,18] any slight alloying addition tends to alter the glass transition and crystallization temperatures. For Ti addition, in particular, T_g, T_x, and ΔT_x all decrease, indicating the bonding among the constituent elements weakens as a result of Ti addition. This is consistent with the experimental results observed in the present study.

Recent results on the crystallization behavior of Zr-based BAAs are summarized in Table IV. As a result of the inclusion of a large number of constituent elements in each alloy, phase formation during crystallization is complex. As expected, the phase formation was strongly affected not only by a specific constituent element, but also by the atomic fraction of the element. For example, the phase structure of the crystallized products can be strongly modified by a small addition of Pd[22] and Ag[18] to the baseline $Zr_{60}Cu_{30}Al_{10}$ or $Zr_{65}Al_{7.5}Ni_{10}Cu_{17.5}$ alloys. Increasing Ti content in a $Zr_{70-x-y}Ti_xAl_yCu_{20}Al_{10}$ alloy can cause the main phase to change from Zr_2Ni to

TABLE III. T_g, T_x and ΔT_x for various Zr-based BAAs.

Composition (at.%)	Heating rate (K/s)	T_g (K)	T_x (K)	ΔT_x (K)	Refs.
$Zr_{41.2}Ti_{13.8}Cu_{12.5}Ni_{10.0}Be_{22.5}$	0.33	625	705	80	6
$Zr_{65}Al_{7.5}Ni_{10}Cu_{17.5}$	0.67	635	746	111	18
$Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Ag_5$		660	723	63	
$Zr_{64.8}Al_{7.7}Cu_{18.2}Ni_{9.3}O_{0.2}$	0.33	637	744	107	19
$Zr_{66.1}Al_{7.2}Cu_{17.4}Ni_{9.3}O_{0.4}$		638	732	94	
$Zr_{66.6}Al_{7.2}Cu_{17.2}Ni_{9.0}O_{0.8}$		647	714	67	
$Zr_{55}Al_{15}Cu_{20}Ni_{10}$	0.67	709	805	96	8
$Zr_{55}Ti_{2.5}Al_{12.5}Cu_{20}Ni_{10}$		690	775	85	
$Zr_{55}Ti_5Al_{10}Cu_{20}Ni_{10}$		669	752	83	

TABLE IV. Crystallization behavior of Zr-Based BAAs.

Alloy composition (at.%)	Annealing condition	Crystallization behaviors	Refs.
$Zr_{41}Ti_{13}Cu_{13}Ni_{10}Be_{23}$	>688 K(*in situ* x-ray)	Am→Be_2Zr+$CuZr_2$+CuZr	20
$Zr_{41}Ti_{15}Cu_{12}Ni_{11}Be_{21}$	DSC at 2 K/min	Am→hex+Am′→hex+Tet	21
$Zr_{60}Al_{10}Cu_{30}$	T_x+50K, 900 s	Am→Zr_2Cu+ZrCu+Zr_3Al	16
$Zr_{55}Al_{10}Cu_{30}Ni_5$		Am→Zr_2Cu+ZrCu+Zr_2Ni+Zr_3Al	
$Zr_{60}Cu_{30}Al_{10}$	705 K, 1.2 ks	Am→Zr_2Cu+Zr_2Al	22
$Zr_{55}Cu_{30}Pd_5Al_{10}$	726 K, 3.6ks; 808 K, 900s	Am→Am′+Zr_2(Cu,Pd)→Zr_2(Cu,Pd)+Zr_3(Al,Pd)$_2$	
$Zr_{65}Al_{7.5}Ni_{10}Cu_{17.5}$		Am→Zr_2(Cu,Ni)+Zr_2(Al,Ni)	18
$Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Ag_5$	730 K, 180s	Am→Am′+Zr_3(Al,Ag)$_2$	
	750 K, 120s	→Zr_3(Al,Ag)$_2$+Zr_2(Cu,Ni)	
$Zr_{70-x-y}Ti_xAl_yCu_{20}Al_{10}$	T_x+100 K, 3.6ks	Am→Zr_2Ni+ZrCu+Zr_3Al(5~15% Al, 0~2.5% Ti)	8
		Am→ZrNi+Zr_2Cu+Zr_3Al(5~15% Al, 2.5~7.5% Ti)	
$(Zr_{0.65}Al_{0.075}Cu_{0.175}Ni_{0.1})_{100-x}O_x$	~698 K, 1.8ks	Am→Zr_2Cu+Zr_6NiAl_2+qc	19
	~723 K, 1.8ks	→Zr_2Cu+Zr_6NiAl_2+fcc−$NiZr_2$	
	~873 K, 1.8ks	→Zr_2Cu+Zr_6NiAl_2	
$Zr_{57}Cu_{20}Al_{10}Ni_8Ti_5$	0.33 K/s to 720 K, 60s	Am→Zr_2Ni+Zr_2Cu+qc	23
	0.33 K/s to 823 K, 60s	→Zr_2Ni+Zr_2Cu	
	0.33 K/s to 960 K, 60s	→Zr_2Cu+Zr_6NiAl_2	

Zr_2Cu.[8] Even for the same $Zr_{60}Cu_{30}Al_{10}$ alloy, different phases have been observed for different crystallization treatments by the same research group.[16,22] Whereas some authors reported the observation of quasi-crystals during the initial stage of crystallization,[19,23] others did not detect the quasi-crystal-like (qc) phase.[8,16,18,22] Apparently, the exact phases formed during crystallization also strongly depend upon the specific crystallization process. Different annealing treatments result in different reaction paths; thus, different kinds of crystallized phases formed. This may be anticipated because of the highly kinetic nature of amorphous alloys. However, despite the complexity of different crystallization paths, it generally agrees that the main product in the equilibrium, crystallized Zr-based BAAs, is Zr_2Cu. This phase is also observed in the present study (annealing treatment at 923 K for 3600 s).

It is worth noting that diffraction patterns change distinctly from samples annealed at 893 K to samples annealed at 923 K for 1 h. As shown in Fig. 1, the structure consists of Zr_2Ni + Ni(Ti,Zr) + NiZr + Zr(AlNi)$_2$ for the 893-K–annealed sample and Zr_2Cu + Zr_2Ni + $Ni_{10}Zr_7$ + Al_3Zr_4 for the 923-K–annealed sample. The 923 K/1h annealing is regarded as an equilibrium annealing treatment; thus, the crystallized phases are thermodynamic equilibrium phases. The phases present in samples annealed at 783 K and 893 K are, however, metastable. In the present study, the main metastable phase obtained was Zr_2Ni. Similar results have been reported by Xing *et al.*[23] in a $Zr_{57}Cu_{20}Al_{10}Ni_8Ti_5$ alloy. Evidently, there exists a solid reaction causing Zr_2Ni to transform into Zr_2Cu either by annealing at 893 K for a longer period of time or at a higher temperature.

In order to examine the difference in phase formation between the equilibrium and nonequilibrium annealing treatment, an *in situ* annealing treatment was performed in the DSC at 923 K followed by rapid cooling. Figure 7 shows the XRD pattern for this sample. The phases present in the 923-K–annealed sample in Fig. 7 are essentially the same as those observed for the 893 K–annealed sample. The main phase is Zr_2Ni. Apparently, there exists a phase transformation from Zr_2Ni to Zr_2Cu by annealing a sample at 923 K for 1 h. Based on the above, schematic continuous cooling transformation (CCT) curves for the crystallization of the Zr–10Al–5Ti–17.9Cu–14.6Ni alloy were constructed as shown in Fig. 8. Only the main phases Zr_2Ni and Zr_2Cu are considered in the curves. When the sample is heated to 923 K at a heating rate of 0.33 K/s, Zr_2Ni is the initial main phase; after a long period (3600 s) of isothermal annealing, Zr_2Cu phase begins to form and Zr_2Ni phase transforms into Zr_2Cu phase. At the present time, it is impossible to predict the partitioning of each element during the crystallization of a multicomponent metallic glass.

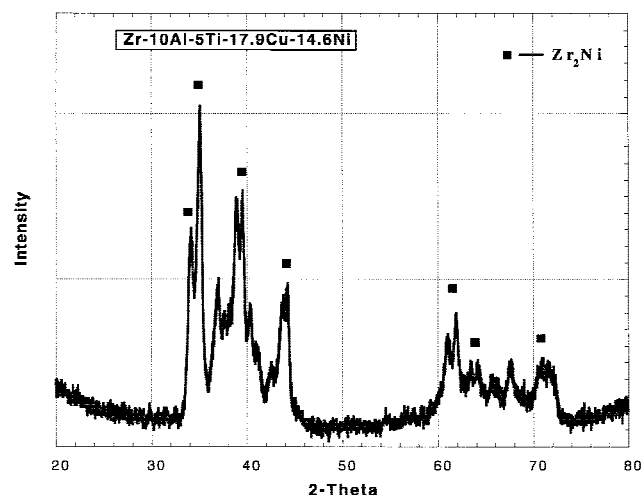


FIG. 7. XRD pattern of a BAA sample annealed at 923 K.

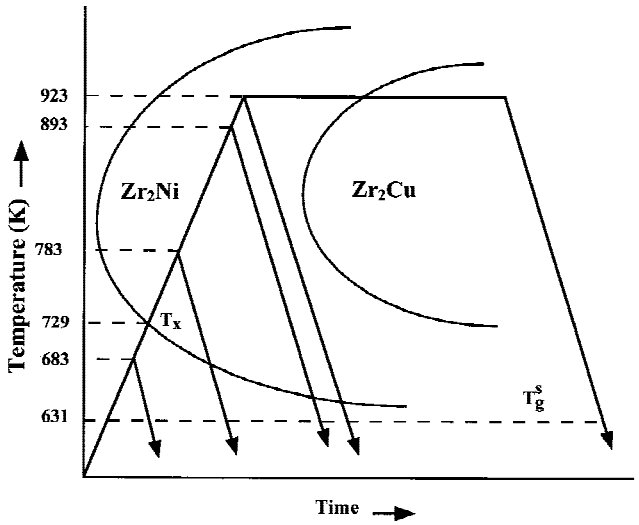


FIG. 8. Schematic CCT curves for the crystallization of the Zr–10Al–5Ti–17.9Cu–14.6Ni alloy.

B. Young's modulus (E) and hardness (H)

Table III summarizes available data on the mechanical properties of Zr-based BAAs, including elastic modulus, hardness, yield strength, fracture strength, and fracture toughness. In comparison with the elastic modulus of 89 MPa determined from the tensile stress-strain curve of $Zr_{52.5}Al_{10}Ti_5Cu_{17.9}Ni_{14.6}$,[12] the data measured by nanoindentation are slightly higher.

The present experimental data (Table II) indicate that material having a crystallized structure has a higher modulus than the material with amorphous structure. Elastic modulus is one of the fundamental properties and is determined essentially by the bonding between atoms. To the first order of approximation (e.g., Lennard–Jones potential), the stiffness or elastic modulus of a material depends upon the interatomic distance. A relaxed structure (e.g., annealed at 683 K) has on an average a shorter interatomic distance (caused by volume decrease) than the as-cast amorphous state. As a result, relaxed structure has a larger modulus than the as-cast amorphous structure. Transient annealing at higher temperatures (783 and 893 K) produces a small volume fraction of metastable crystalline phases dispersed in amorphous matrix, but does not affect appreciably the modulus. After annealing at 923 K for 1 h the entire structure crystallizes and the crystallized alloy exhibits a high modulus. It is known that alloy with a crystallized structure usually has a higher modulus than its amorphous counterpart.[29] This has been demonstrated in $Zr_{41}Ti_{13}Ni_{10}Cu_{13}Be_{23}$,[20] $Zr_{55}Al_{10}Ni_5Cu_{30}$,[10] and $Zr_{65}Al_{10}Ni_{10}Cu_{15}$[27] amorphous alloys.

Hardness is a measure of a material's resistance to plastic deformation. In the case of crystalline solids, hardness is directly related to the yield strength of a material. The yield strength of a crystalline solid is determined by dislocation interaction with lattice defects such as vacancy, solute, dislocation, and boundary interfaces. In the case of amorphous alloys, although the exact atomic mechanisms for plastic deformation are unclear, an empirical relationship exists between hardness H and fracture strength σ_f (i.e., $H/\sigma_f \approx 3$).[30]

For the present BAA, partially and fully crystallized materials have the highest hardness, followed by the material with relaxed structure, and the as-cast material is the softest. The fully crystallized material essentially contains nanocrystalline grains. Grain-size strengthening (e.g., Hall–Petch relationship) is expected to dominate. In the partially crystallized materials, on the other hand, a strengthening mechanism, which is similar to dispersion strengthening occurring in crystalline solids, is expected to take place. In fact, Inoue and colleagues[18,22] recently demonstrated that the dispersion of nanometer crystalline particles in amorphous structure can cause a significant increase in hardness. For example, the hardness of an amorphous Zr–Al–Ni–Cu–Ag alloy is 6.3 MPa but increases to 8.6 MPa when the alloy partially crystallizes and contains nanometer crystals. It is unclear why the relaxed structure is slightly harder than the as-cast amorphous structure. Presumably, the as-cast structure contains a higher defect density and residual stresses, and thus is harder than the relaxed structure. However, the data show the opposite. It could be related to more dense packing after relaxation. Without knowing the exact plastic flow mechanisms in amorphous alloys, it is impossible to offer an explanation for the discrepancy at the present time.

Whang *et al.*[31] demonstrated that there is a linear correlation between hardness and modulus. The ratio of H/E for Zr-BAAs listed in Table V varies from 0.047 to 0.096. It is of interest to know that a covalently bonded solid exhibits a higher H/E than that of a metal; the H/E ratio is about 1/10 for most covalently bonded materials, as compared to about 1/500 for face-centered-cubic metals.[32,33] This difference is a result of the fact that plastic yielding in covalent materials involves actual breakage of atomic bonds, a process that requires high stress. In contrast, the low resistance to plastic flow relative to the modulus in metals is attributed to the ease of dislocation slip in these materials. The Zr-based BAAs exhibit an H/E ratio close to 0.1, thereby suggesting that plastic flow is relatively difficult in these materials. Also, the bonding tendency of Zr-BAAs is probably covalent.

V. CONCLUSIONS

The crystallization and nanoindentation behavior of the as-cast Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) BAA and material annealed at different temperatures has been characterized. The relationship between the microstruc-

TABLE V. Summary of available mechanical properties of Zr-based BAAs.

Alloy composition (at.%) and microstructure	Mechanical properties[a]				Refs.
	E (GPa)	H (GPa)	Other properties	H/E	
$Zr_{41.25}Ti_{13.75}Ni_{10}Cu_{12.5}Be_{22.5}$(Am)	93 ± 1		$\nu = 0.387$, $\sigma_s = 1.87 \sim 2.12$		24
$Zr_{41.25}Ti_{13.75}Ni_{10}Cu_{12.5}Be_{22.5}$(Am)	96	5.2	$\mu = 34.3$, $\nu = 0.36$ $\sigma_s = 1.9$, $K_{IC} = 57$	0.054	25
$Zr_{60}Cu_{20}Pd_{10}Al_{10}$ (Am)	52.2	5.0	$\sigma_f = 1.64$	0.096	22
$Zr_{60}Cu_{20}Pd_{10}Al_{10}$ ($V_f = 75\%$)	62.3	5.9	$\sigma_f = 1.91$	0.095	
$Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Ag_5$ (Am)		4.4	$\sigma_f = 1.15$		18
$Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Ag_5$ ($V_f = 14\%$)		4.7	$\sigma_f = 1.52$		
$Zr_{55}Al_{10}Ni_5Cu_{30}$ (Am)	103	4.8	$\sigma_f = 1.82$	0.047	26
$Zr_{70-x-y}Ti_xAl_yCu_{20}Ni_{10}$		4.7–5.1	$\sigma_f = 1.55$–1.85		8
$Zr_{41}Ti_{13}Ni_{10}Cu_{13}Be_{23}$ (Am)	88	5.5		0.062	20
$Zr_{41}Ti_{13}Ni_{10}Cu_{13}Be_{23}$ (Cryst.)	116	9.5			
$Zr_{55}Al_{10}Ni_5Cu_{30}$ (Am)	78–80		$\sigma_f = 1.42$–1.57		10
$Zr_{55}Al_{10}Ni_5Cu_{30}$ (Cryst. + Am)	117		$\sigma_f = 0.93$		
$Zr_{65}Al_{10}Ni_{10}Cu_{15}$ (Am)	80		$\sigma_f = 1.52$		27
$Zr_{65}Al_{10}Ni_{10}Cu_{15}$ (Cryst. + Am)	117		$\sigma_f = 0.93$		
$Zr_{60}Al_{10}Co_3Ni_9Cu_{18}$ (Am)	97	4.7	$\sigma_f = 1.51$	0.048	28
$Zr_{57}Ti_5Al_{10}Ni_8Cu_{20}$ (Am)			$\sigma_s = 1.54$, $\sigma_f = 1.69$		23
$Zr_{52.5}Al_{10}Ti_5Cu_{17.9}Ni_{14.6}$	89		$\sigma_f = 1.65$–1.88		12
$Zr_{52.5}Al_{10}Ti_5Cu_{17.9}Ni_{14.6}$	92–112	6.2–6.3		0.056~0.067	This study

[a]E, elastic modulus; μ, shear modulus (GPa); ν, Poission's ratio; H, hardness; σ_s, yield strength (GPa); σ_f, fracture strength (GPa); K_{IC}, fracture toughness (Mpa $\sqrt{m}$); V_f, volume percentage of crystalline particles.

ture and mechanical properties has been discussed. This study has led to the following conclusions:

(1) The crystallization sequence of the BAA with increasing temperatures is as follows:

$$\begin{aligned}
&\mathrm{Am} \xrightarrow{683\ \mathrm{K}} \mathrm{Am}' \xrightarrow{783\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} \\
&\quad + (\mathrm{Ni}, (\mathrm{Zr}, \mathrm{Ti})) \xrightarrow{893/923\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} \\
&\quad + (\mathrm{Ni}, (\mathrm{Zr}, \mathrm{Ti})) + \mathrm{Zr(AlNi)_2} \\
&\quad + \mathrm{NiZr} \xrightarrow{923\mathrm{K}\ 1\mathrm{h}} \mathrm{Zr_2Cu} + \mathrm{Al_3Zr_4} + \mathrm{NiZr_2} + \mathrm{Ni_{10}Zr_7} \quad .
\end{aligned}$$

(2) Resulting from the kinetic nature of phase transformation in multicomponent alloys, the crystallization path was complex. Despite the complexity of different crystallization paths, the main final crystallized product in the Zr-based BAA is Zr_2Cu. A phase transformation from Zr_2Ni to Zr_2Cu occurs during long-period isothermal annealing treatment at relatively higher temperature.

(3) Young's modulus and hardness of the BAA increase with an increase in annealing temperature, from Am to supercooled liquid Am′, to partially crystallized, and to fully crystallized, in an ascending order.

(4) Zr-based BAAs exhibit an H/E ratio of about 1/10, suggesting the interatomic bond in the alloys is close to being covalent.

ACKNOWLEDGMENT

This work was performed under the auspices of the United States Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

REFERENCES

1. A. Inoue, Mater. Trans., JIM **36,** 866 (1995).
2. W.L. Johnson, Mater. Sci. Forum **225–227,** 35 (1996).
3. A. Inoue, T. Zhang, and T. Masumoto, Mater. Trans., JIM **31,** 177 (1990).
4. A. Inoue, T. Nakamura, N. Nishiyama, and T. Masumoto, Mater. Trans., JIM **33,** 937 (1992).
5. T. Zhang, A. Inoue, and T. Masumoto, Mater. Trans., JIM **32,** 1005 (1991).
6. A. Peker and W.L. Johnson, Appl. Phys. Lett. **63,** 2342 (1993).
7. A. Inoue, T. Shibata, and T. Zhang, Mater. Trans., JIM **36,** 1420 (1995).
8. T. Zhang and A. Inoue, Mater. Trans., JIM **39,** 857 (1998).
9. A. Inoue and J.S. Gook, Mater. Trans., JIM **36,** 1180 (1995).
10. A. Inoue, Sci. Rpt. RITU, **A42,** 1 (1996).
11. X.H. Lin, W.L. Johnson, and W.K. Rhim, Mater. Trans., JIM **38,** 473 (1997).
12. C.T. Liu, *et al.*, Metall. Mater. Trans. A, **29A,** 1811 (1998).
13. T.G. Nieh, T. Mukai, C.T. Liu, and J. Wadsworth, Scr. Mater. **40,** 1021 (1999).
14. W.C. Oliver and G.M. Pharr, J. Mater. Res. **7,** 1564 (1992).
15. R. Busch, Y.J. Kim, and W.L. Johnson, J. Appl. Phys. **77,** 4039 (1995).
16. A. Inoue, T. Negishi, H.M. Kimura, T. Zhang, and A.R. Yavari, Mater. Trans., JIM **39,** 318 (1998).
17. B.N. Lucas, W.C. Oliver, G.M. Pharr, and J-L. Loubet, in *Thin Films: Stresses and Mechanical Properties VI,* edited by W.W. Gerberich, H. Gao, J-E. Sundgren, and S. Baker (Mater. Res. Soc. Symp. Proc. **436,** Pittsburgh, PA, 1996) pp. 233.
18. A. Inoue, T. Zhang, and Y.H. Kim, Mater. Trans., JIM **38,** 749 (1997).
19. J. Eckert, N. Mattern, M. Zinkevitch, and M. Seidel, Mater. Trans., JIM **39,** 623 (1998).
20. H.J. Fecht, Philos. Mag. B **76,** 495 (1997).
21. S. Spriano, C. Antonione, R. Doglione, L. Battezzati, S. Crdoso, J.C. Soares, and M.F. da Silva, Philos. Mag. B **76,** 529 (1997).
22. C. Fan and A. Inoue, Mater. Trans., JIM **38,** 1040 (1997).

23. L.Q. Xing, J. Eckert, W. Loser, and L. Schultz, Appl. Phys. Lett. **73,** 2110 (1998).
24. H.A. Bruck, T. Christman, A.J. Rosakis, and W.L. Johnson, Scr. Metall. Mater. **30,** 429 (1994).
25. R.D. Conner, A.J. Rosakis, W.L. Johnson, and D.M. Owen, Scr. Mater. **37,** 1373 (1997).
26. H. Kato and A. Inoue, Mater. Trans. JIM **38,** 793 (1997).
27. Y. Kawamura, H. Kato, A. Inoue, and T. Masumoto, Appl. Phys. Lett. **67,** 2008 (1995).
28. A. Inoue, T. Zhang, and T. Masumoto, Mater. Trans. JIM **36,** 391 (1995).
29. T.C. Chou, T.G. Nieh, S.D. McAdams, and G.M. Pharr, Scr. Metall. Mater. **25,** 2203 (1991).
30. H. Kimura and T. Masumoto, in *Amorphous Metallic Alloys*, edited by F.E. Luborsky (Butterworths, London, United Kingdom, 1983), p. 187.
31. S.H. Whang, D.E. Polk, and B.C. Giessen, in *Proceedings of the Fourth International Conference on Rapidly Quenched Metals*, edited by T. Masumoto and K. Suzuki (Japan Institute of Metals, 1982, 1981), p. 1365.
32. J.J. Gilman, in *The Science of Hardness Testing and Its Research Applications*, edited by Westbrook and Conrad (American Society for Metals, Metals Park, OH, 1973), p. 51.
33. J.J. Gilman, J. Appl. Phys. **46,** 1435 (1975).

Nano-scratch behavior of a bulk Zr–10Al–5Ti–17.9Cu–14.6Ni amorphous alloy

J.G. Wang,[a] B.W. Choi, and T.G. Nieh
Chemistry and Materials Science, Lawrence Livermore National Laboratory, P.O. Box 808, L-350, Livermore, California 94551-9900

C.T. Liu
Metals and Ceramics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6376

(Received 27 August 1999; accepted 18 January 2000)

The tribological behavior of a Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) bulk amorphous alloy, in both the as-cast and annealed states, was investigated using nano-scratch tests, including ramping load scratch and multiple sliding wear techniques. The crystallization sequence of the alloy was also characterized. Mechanical properties, such as Young's modulus, hardness, friction coefficient, and tribological wear were measured. These properties were found to vary with microstructure. In general, an increase in annealing temperature results in an increase in hardness, which in turn produces a decrease in friction coefficient but an increase in wear resistance. Samples having a structure consisting of supercooled liquid matrix with dispersed nanocrystalline particles exhibit the best wear performance.

I. INTRODUCTION

Recently, a number of bulk amorphous alloys (BAAs) with critical cooling rates less than 10 K/s have been discovered in many muticomponent alloy systems of Ln–Al–TM,[1,2] Mg–Ln–TM,[3,4] Zr–Al–TM,[5,6] Zr–Ti–Al–TM,[7] Ti–Zr–TM,[8,9] Zr–Ti–TM–Be,[10] Pd–Cu–Ni–P,[11] and Fe–(Al, Ga)–(P, B, C, Si)[12,13] (where Ln = lanthanide metal and TM = transition metal). The high glass-forming ability of these alloys was observed to be dependent on the physical and chemical character of the constituent elements, as well as the number of elements in the alloys.[14–16] Amorphous alloys (or metallic glasses) have many unusual properties such as good corrosion resistance, high elasticity, and high hardness.[17,18] In view of their high hardness and strength, BAAs are promising materials for tribological (friction and wear) applications.

Several studies have been conducted on the wear and friction properties of rapidly quenched ribbons. The results, however, are controversial. Fe- Ni-, and Co-based amorphous alloys containing Si, B, and, sometimes, C, were used in these studies because of the consideration of using these alloys for magnetic recording devices.[19,20] Some metallic glasses showed high wear resistance, better than their crystallized counterpards.[21,22] However, opposite results have also been reported where metallic glasses exhibited wear performance well below expectation.[23–25] Some researchers demonstrated that an amorphous alloy containing a homogeneous dispersion of intermetallic compounds exhibited the best wear performance, and the fully amorphous material exhibited the highest wear rate.[26,27] The causes for these discrepancies are unclear. The purpose of this paper is to use nano-scratch tests to determine the tribological behavior of a Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) amorphous alloy and to shed light on the structure–property relationship.

II. EXPERIMENTAL PROCEDURES

A bulk amorphous alloy with a nominal composition of Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%) was prepared by arc melting and drop casting, as described elsewhere.[28] The as-cast amorphous samples used for the test are of a cylindrical shape with a diameter of 7 mm. The amorphous nature of the cast samples was confirmed by using x-ray diffraction (XRD) and transmission electron microscopy (TEM). The sequential phase transition from glass to supercooled liquid and to crystallization states was studied by using differential scanning calorimetry (DSC) at a heating rate of 0.33 K/s.

To prepare samples for characterizing structures at different crystallization stages, and to minimize kinetic effects, various annealing treatments were performed *in*

[a] Address all correspondence to this author.
Present address: Department of Chemical and Biochemical Engineering and Materials Science, University of California, Irvine, 916 Engineering Tower, Irvine, California 92697-2575.
e-mail: wangjg@uci.edu

 © 2000 Materials Research Society

situ in the DSC cell with a heating rate of 0.33 K/s under flowing N_2. After reaching the desired temperature, the sample was rapidly cooled to room temperature. The heat-treated samples were examined using XRD and TEM.

Nanoindentation and nano-scratch tests were conducted on the as-cast and annealed samples using a Nano Indenter-XP®. Hardness and Young's modulus were measured using the continuous stiffness option, which yields hardness as a continuous function of depth into the sample.[29,30] A quartz sample was used as a standard for the initial calibration. Scratch tests were performed using the lateral force measurement option. A Berkovich indenter was used for the scratch test. The orientation of the indenter was aligned so that the sliding was performed with edge forward and face backward. Ramping load scratch tests were performed on samples over a 700-μm length. During a test, the applied normal load was linearly increased from an initial 20 μN to the maximum load, and the maximum loads selected in the present study were 20, 80, and 150 mN. Multiple sliding wear tests were also conducted. In each experiment, thirty-one passes were made over a 40-μm length using constant normal loads of 20 mN. The speed of the indenter for all sliding segments was about 1.0 μm/s.

III. EXPERIMENTAL RESULTS

A. Crystallization and mechanical properties

The DSC curve from the as-cast Zr–10Al–5Ti–17.9Cu–14.6Ni alloy (referred to hereafter as Zr–Al–Ti–Cu–Ni alloy) with a heating rate of 0.33 K/s is shown in Fig. 1. The amorphous alloy exhibits an endothermic reaction resulting from a glass transition, followed by the supercooled liquid region, and then an exothermic reaction caused by crystallization. The onset temperatures of the glass transition and crystallization, denoted by T_g and T_{x1} and T_{x2}, are determined to be 631, 729, and 818 K, respectively. From these measurements, three annealing temperatures, 683, 783, and 893 K, were selected; these temperatures correspond to the supercooled liquid region, and temperatures after the first and second crystallization peaks, respectively. The annealed samples were examined using XRD and the results are shown in Fig. 2, in which (a) is from the as-cast sample; (b), (c), and (d) are from samples annealed at 683, 783, and 893 K, respectively; and (e) is from a sample annealed at 923 K for 1 h. For discussion purposes, these five samples will be denoted as samples A, B, C, D, and E hereafter. According to a previous study,[31] the crystallization sequence of the amorphous Zr–10Al–5Ti–17.9Cu–14.6Ni alloy can be described by

$$\begin{aligned} \mathrm{Am} &\xrightarrow{683\ \mathrm{K}} \mathrm{Am}' \xrightarrow{783\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} + [\mathrm{Ni},\ (\mathrm{Zr},\ \mathrm{Ti})] \\ &\xrightarrow{893\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} + [\mathrm{Ni},\ (\mathrm{Zr},\ \mathrm{Ti})] + \mathrm{Zr(AlNi)_2} \\ &\quad + \mathrm{NiZr} \xrightarrow{923\ \mathrm{K},\ 1\ \mathrm{h}} \mathrm{Zr_2Cu} + \mathrm{NiZr_2} + \mathrm{Ni_{10}Zr_7} \\ &\quad + \mathrm{Al_3Zr_4} \quad , \end{aligned}$$

where Am and Am′ stand for the alloy in amorphous and relaxed states, respectively. A dark-field TEM image and selected-area electron diffraction pattern from samples A, B, C, D, and E are shown in Fig. 3. Both the as-cast and 683 K–annealed samples show a uniform amorphous structure [Figs. 3(a) and 3(b)]. The halo ring in the diffraction pattern of the 683 K–annealed sample is apparently sharper than that of the as-cast sample, as a result of the relaxation of the glassy state. In contrast, samples annealed at 783 and 893 K mainly consist of extremely fine Zr_2Ni grains with a size of less than 10 nm at 783 K and 40 nm at 893 K [Figs. 3(c) and 3(d)]. Diffraction patterns show distinct triple rings resulting from the three

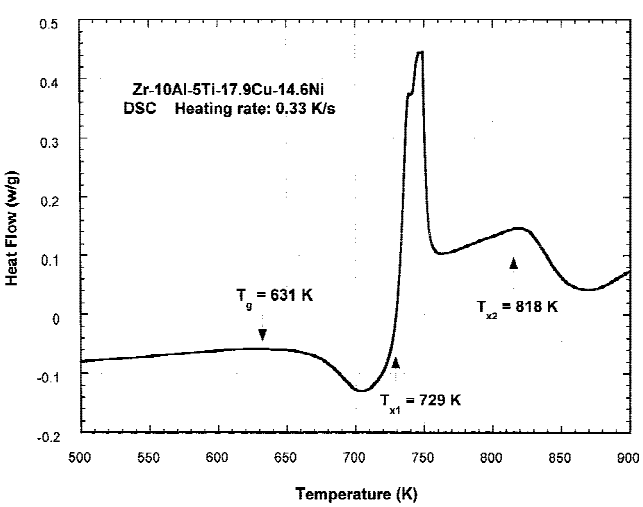


FIG. 1. DSC curve of the amorphous Zr–10Al–5Ti–17.9Cu–14.6Ni alloy with a heating rate of 0.33 K/s.

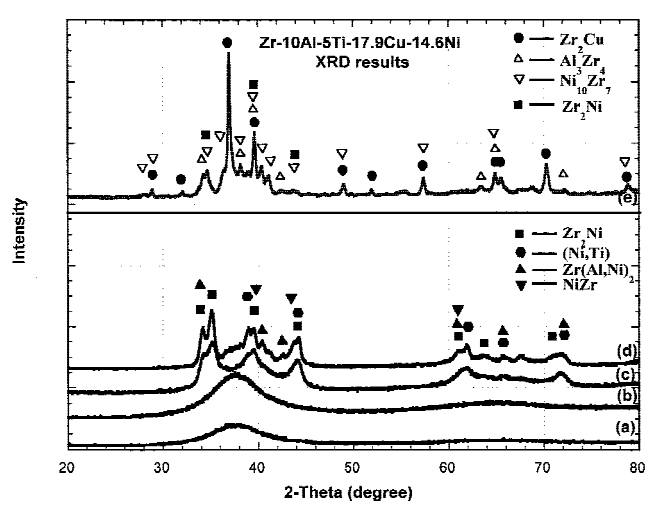


FIG. 2. XRD patterns from as-cast and annealed BAA samples: (a) as-cast, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h.

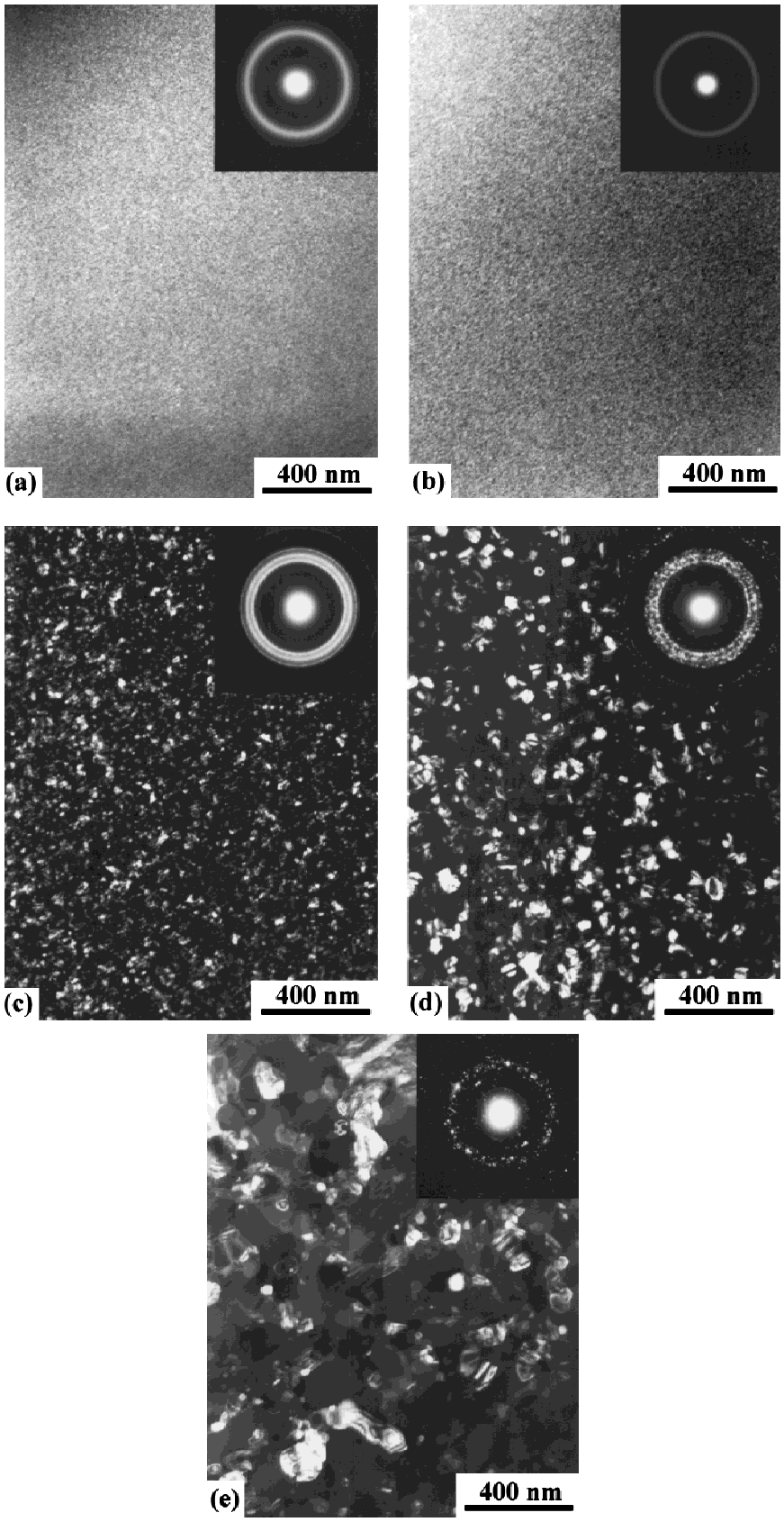


FIG. 3. Dark-field TEM images and selected-area electron diffraction patterns of the as-cast and annealed BAA samples: (a) as-cast, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h.

diffraction peaks of the Zr_2Ni phase. The volume fraction of this nanocrystalline phase is estimated to be about 80% at 783 K and about 95% at 893 K. The structure of the sample annealed at 923 K for 1 h contains polycrystals with a mean grain size of about 250 nm, as shown in Fig. 3(e). The diffraction pattern consists of spotty reflection rings resulting from the presence of the Zr_2Cu phase; it no longer has the triple ring characteristic of the Zr_2Ni phase.

Hardness and Young's modulus measurements were performed on these samples. The results are summarized in Table I. It is readily observed that the as-cast sample has the lowest hardness and modulus values, while the fully crystallized sample exhibits the highest values.

B. Nano-scratch behavior

1. Ramping load scratch

A ramping load scratch (RLS) technique was used to characterize the friction behavior of the Zr–Al–Ti–Cu–Ni alloy. The measured friction coefficient [Figs. 4(a) and 4(c)] and surface profile [Figs. 4(b) and 4(d)] of samples A and B are shown in Fig. 4. The ramping normal load of 0.02–80 mN was used for the scratch tests. For the purpose of discussion, the ramping normal load is also included in the figure (bold solid line). As shown in Fig. 4, sudden increases in the friction coefficient (indicated by arrows) are frequently observed. It is interesting to note that this has also been often observed during nanoindentation of thin film. In the case of thin films, a sudden increase in friction coefficient is associated with microcracking or delamination of the films, which provides a measure of scratch resistance or adhesive strength of a film.[32–34] The load that causes the delamination is denoted as "critical load." In the case of bulk monolithic material, the sudden increase in friction coefficient is attributed to microcracking. Bulk amorphous alloys are known to be brittle; when the local strain exceeds the elastic limit during nanoindentation, microcracking is expected to occur.

TABLE I. Hardness and Young's modulus of as-cast and annealed Zr–10Al–5Ti–17.9Cu–14.6Ni alloy obtained by using different nanoindentation tests.

Sample	Test[a]	Depth (nm)	Load (mN)	Elastic modulus (GPa)	Hardness (GPa)
A: As-cast	HL	733 ± 4	60	112 ± 6	6.3 ± 0.6
	HD	2000	474 ± 4	92 ± 4	6.2 ± 2
B: 683 K annealed	HL	700 ± 3	60	119 ± 2	6.9 ± 0.1
	HD	2000	496 ± 8	100 ± 5	6.5 ± 0.2
C: 783 K annealed	HL	661 ± 3	60	118 ± 2	8.3 ± 0.1
	HD	2000	590 ± 12	99 ± 4	7.8 ± 0.2
D: 893 K annealed	HL	652 ± 4	60	119 ± 1	8.6 ± 0.1
	HD	2000	604 ± 10	99 ± 4	8.0 ± 0.2
E: 923 K annealed	HL	627 ± 4	60	150 ± 2	8.6 ± 0.2
(1 h)	HD	2000	619 ± 17	142 ± 1	8.3 ± 0.1

[a]HL, hardness test at a load; HD, hardness test as a depth.

As shown in Fig. 4(a), the baseline friction coefficient of sample A (i.e., as-cast sample) is about 0.2–0.25. However, at locations where sudden increases in friction coefficient occur, the friction coefficient can jump to about 0.3–0.35. (The scratch proceeds from left to right). To assess the scratch resistance of the sample, surface profiles [Fig. 4(b)] were measured by scanning the sample before and after the scratch test. It is noted that both the prescan and postscan were conducted using an extremely low load of 20 nN. The difference between the prescan and postscan profiles is presumably the damage caused by scratching, and the difference between the scratch and postscan profiles represents the elastic recovery of the sample. It is particularly pointed out that the locations where microcracking occurs (indicated by the sudden changes in friction coefficient) coincide with the depressions in the postscan surface profiles. The critical load can be determined from the first abrupt change in the surface profile and friction coefficient, and specifically it is about 28 mN (indicated in the figure). Other peaks at less than 28 mN on the friction coefficient profile may be caused by the surface roughness or break of the oxide layer during the initial stage of scratching. *In situ* optical microscopic observation revealed the formation of scratched debris at this load, confirming the occurrence of microcracking.

The structure of sample B (683 K annealed) is in a relaxed glass (supercooled liquid) state. The friction coefficient of this sample is averaged around 0.17 but it can sometimes jump to about 0.25–0.30. Sudden changes in friction coefficient are also observed at several locations, as indicated by arrows. The critical load determined by the first spike in both the surface and friction coefficient profiles [Figs. 4(c) and 4(d)] is about 46 mN, which is higher than that observed in sample A [Figs. 4(a) and 4(b)]. The higher critical load may be caused by the fact that a relaxed glass has less free volume, thus a stronger bonding and higher hardness value, than the as-cast glass. Gaskell[35] showed that when glass clusters approach the supercooled liquid state, the average chemical bonding increases.

For samples consisting of crystallized phases (samples C, D, and E), that is, samples annealed respectively at 783, 893, and 923 K, relatively low friction coefficients of about 0.15 were obtained; this is shown in Fig. 5. In contrast to samples A and B, virtually no "critical loads" were detected in these samples. The applied load appears to have no significant effect on the friction coefficient except in the initial scratch stage. Also, no evidence for the formation of debris was observed on the tested samples up to the maximum load of 80 mN. However,

when the maximum applied load was increased to 150 mN, critical loads of approximately 110 and 86 mN were observed in samples C and D, respectively. The baseline friction coefficient increases slightly to about 0.18 (Fig. 6) for both samples at this higher maximum load. The observed higher critical loads in samples C and D (both partially crystallized) are presumably associated with a less defective structure. It is worth noting that the critical load of 110 mN for the 783 K–annealed sample is higher than that for the 893 K–annealed sample (86 mN).

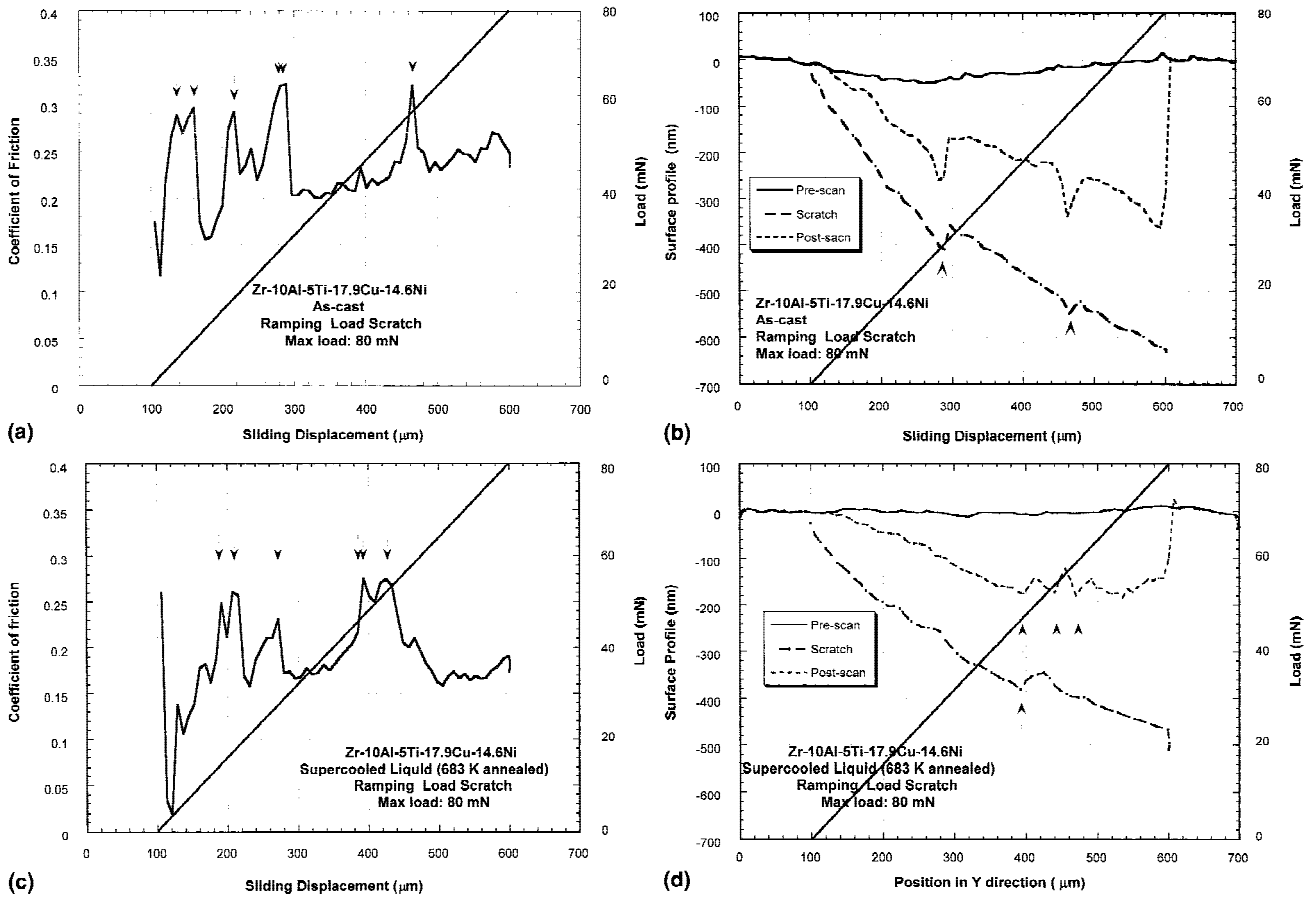


FIG. 4. Friction coefficient and surface profiles of the two samples in amorphous state (as-cast Am and supercooled Am′) scratched with ramping normal loads of 0.02–80 mN: (a) friction coefficient and (b) surface profile of as-cast Am; and (c) friction coefficient and (d) surface profile of supercooled Am′.

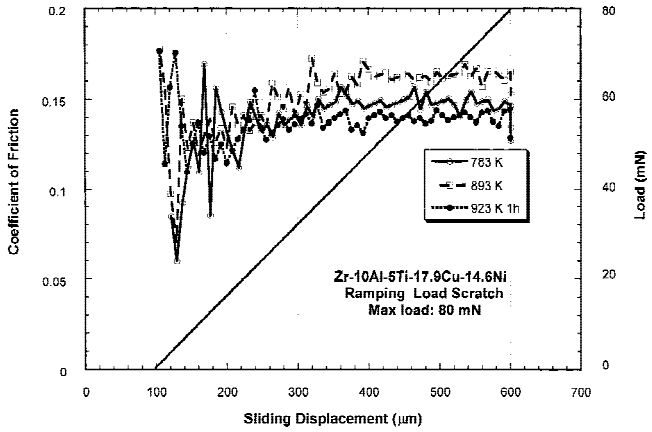


FIG. 5. Friction coefficient profiles of samples consisting of crystallized phases (annealed at 783 and 893 K, and at 923 K for 1 h) scratched with ramping normal loads of 0.02–80 mN.

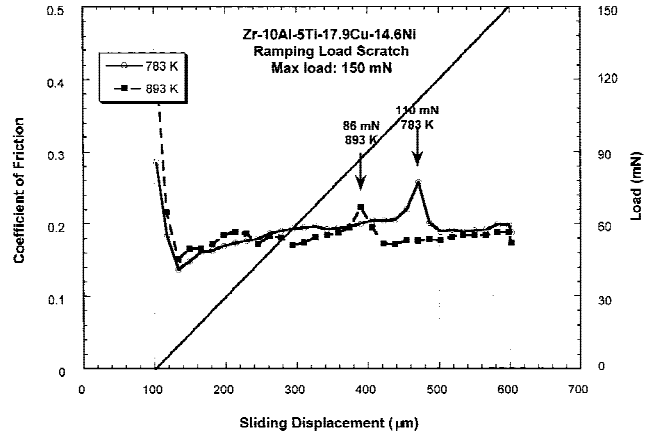


FIG. 6. Friction coefficient profiles of the nanocrystallized microstructure of the two samples (annealed at 783 and 893 K) scratched with ramping normal loads of 0.02–150 mN.

A summary of the baseline friction coefficients ($\overline{\mu}_B$), scratch (d_s), and postscan (d_p) depths at the maximum load of 80 mN for samples A, B, C, D, and E is listed in Table II. Data in the table are the average values from at least five tests. For each scratch experiment, at least five tests were performed to measure the critical loads. It is apparent that the as-cast sample has the highest friction coefficient, and the fully crystallized (nanocrystalline) sample has the lowest friction coefficient. Also, the friction coefficient decreases with increasing hardness, suggesting friction coefficient is dependent on the structure of the alloy.

The values of elatic recovery, represented by the difference between the scratch and postscan profile, $d_s - d_p$, for samples A, B, C, D, and E are 310, 350, 320, 320, and 220 nm, respectively. This result indicates that the allow in both the amorphous (A and B) and partially crystallized (C and D) states exhibits more elastic recovery than that in the fully crystallized state (E). However, the permanent damage caused by scratch is evaluated from the difference between the postscan and prescan profiles.

2. Multiple sliding wear

A multiple sliding wear (MSW) technique was also used for evaluating the wear resistance of Zr–Al–Ti–Cu–Ni alloy. The friction coefficients of the alloy in various annealed states are shown in Figs. 7(a)–7(e). It is noted in these figures that thirty-one sliding scans were made on each sample to ensure the attainment of a steady state. The applied normal load and sliding distance used were 20 mN and 40 μm, respectively. A constant normal load of 20 mN was intentionally chose to be lower than the critical load (28 mN) to avoid microcracking. Data displayed in Fig. 7 include both forward (odd number) and backward (even number) slides. Interestingly, despite the great difference between the edge-forward and face-backward alignment of the Berkovich diamond tip, friction behavior appears to be independent of the sliding direction (forward and backward). The friction coefficient (~0.1–0.2) was generally low in the initial scan, but increases rapidly to about 0.35–0.50 in the subsequent scans. It usually took more than five or six scans for the friction coefficient to reach a steady value.

The friction coefficient, μ, of sample A (as-cast) obtained in the first scan was only 0.15–0.2, increases rapidly, and reaches a steady state with an average value of about 0.42 [Fig. 7(a)]. In a similar manner, sample B (relaxed glass) also exhibits a low initial friction coefficient of about 0.1 and quickly reaches a steady state after five scans, yielding a slightly lower friction coefficient of about 0.34 [Fig. 7(b)]. It is also noted that the friction coefficient appears to increase slightly with sliding distance. A rapid increase in friction coefficient at the end of sliding, especially during the forward sliding, was apparently attributed to the pileup of worn material. As a result, the lateral forces tend to "lift" the indenter during sliding; this was described previously by Deng *et al.*[34]

The friction behavior of sample C (partially crystallized) is different from that of samples A and B. As shown in Fig. 7(c), the starting friction coefficient of sample C is low (~0.12), but increases rapidly to a peak value of about 0.28 after five scans. Instead of having a constant value, the friction coefficient gradually decreases to an average value of 0.24. This decrease may be attributed to work hardening taking place during multiple slides. Recalled that sample C has a microstructure containing nanocrystalline grains. Work hardening is expected to occur in crystalline grains and contribute to an increase in hardness, which, in turn, causes a decrease in friction coefficient. This is further supported by the results obtained from sample D. As shown in Fig. 7(d), a low final friction coefficient of about 0.15 is observed in sample D; this is because higher volume fraction of crystalline phase produces more work hardening. For the fully crystallized sample (sample E), the steady-state friction coefficient is about 0.3, which is higher than that for samples C and D (partially crystallized) but smaller than that for the amorphous samples.

The typical surface profiles of sample E before (prescan) and after (postscan) multiple sliding are shown in Fig. 8. The postscan depth can be used to estimate the wear resistance of a sample, because it represents the material that was worn away. The depth values are 600, 370, 250, 270, and 300 for samples A, B, C, D, and E, respectively. The steady-state friction coefficient for the forward sliding (μ_F), hardness (H), and postscan depth (d_p) under a normal sliding load of 20 mN are summarized in Table III. It is apparent in the table that there

TABLE II. Young's modulus, hardness, average value of friction coefficients ($\overline{\mu}_B$), scratch (d_s) and postscan (d_p) depths of a ramping load scratch test at the load limit of 80 mN.

Sample	Microstructure	E (GPa)	H (GPa)	$\overline{\mu}_B$	d_s (nm)	d_p (nm)
A: As-cast	Am	92–112	6.2–6.3	0.220	610	300
B: 683 K annealed	Am′	100–119	6.5–6.9	0.180	550	200
C: 783 K annealed	Am′ + $NiZr_2$ + [Ni, (Zr, Ti)]	99–118	7.8–8.3	0.155	460	140
D: 893 K annealed	Am′ + $NiZr_2$ + [Ni, (Zr, Ti)] + $Zr(AlNi)_2$ + NiZr	99–119	8.0–8.6	0.150	440	120
E: 923 K annealed (1 h)	Zr_2Cu + $NiZr_2$ + $Ni_{42}Zr_{58}$ + Al_3Zr_4	142–150	8.3–8.6	0.140	330	110

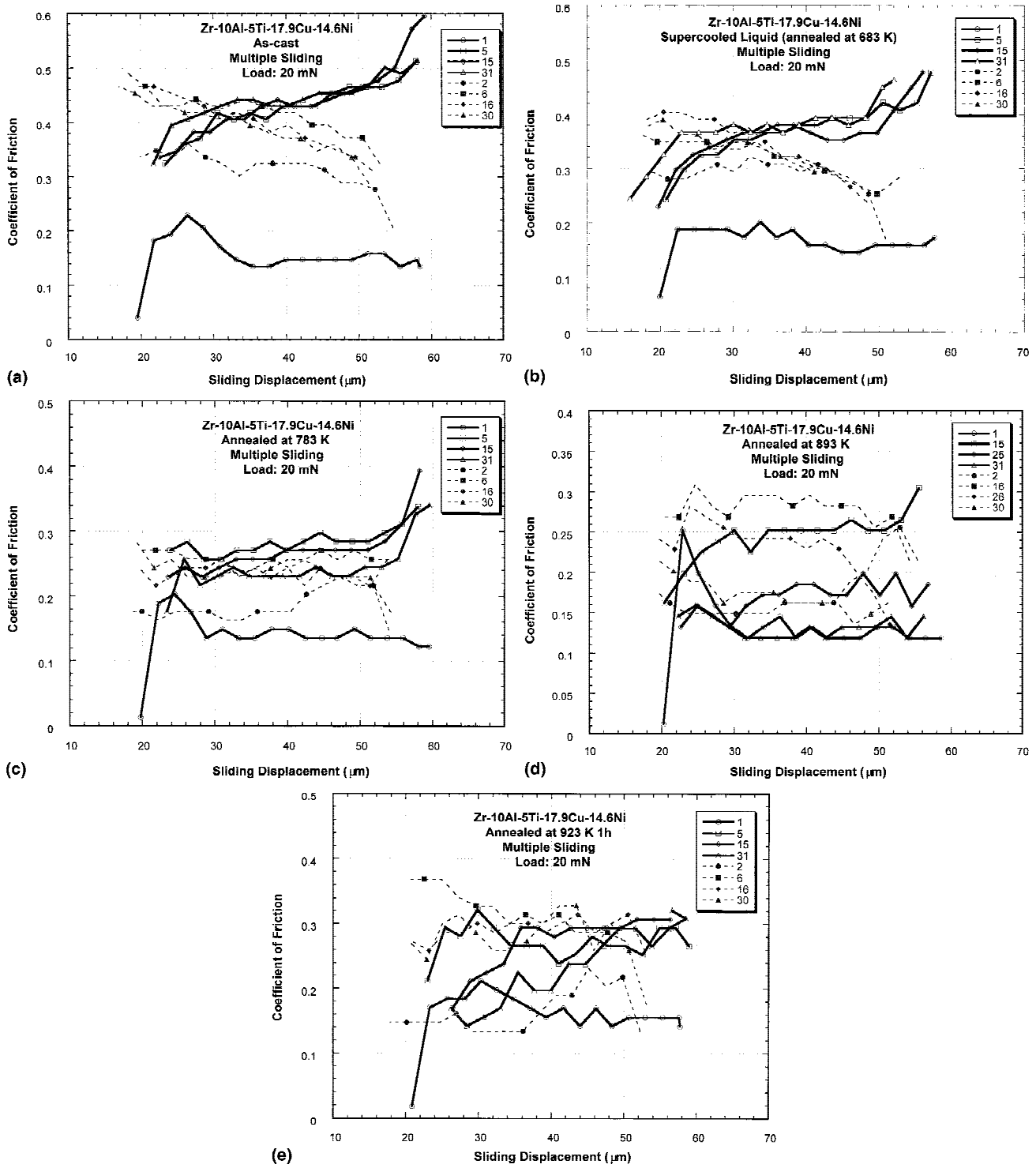


FIG. 7. Friction coefficient profiles of the multiple sliding wear test (thirty-one sliding scans) on all of the five samples under the normal load of 20 mN: (a) as-cast, (b) 683 K, (c) 783 K, (d) 893 K, and (e) 923 K for 1 h.

exists a general trend; namely, the worn depth is shallower for a harder sample. Once again, data indicate that both the partially crystallized (samples C and D) and fully crystallized (sample E) samples are more wear resistant than the as-cast sample (sample A).

IV. DISCUSSION

Friction is known to depend on a number of variables, such as chemical affinity between the two contacting materials, surface roughness, the presence or absence of

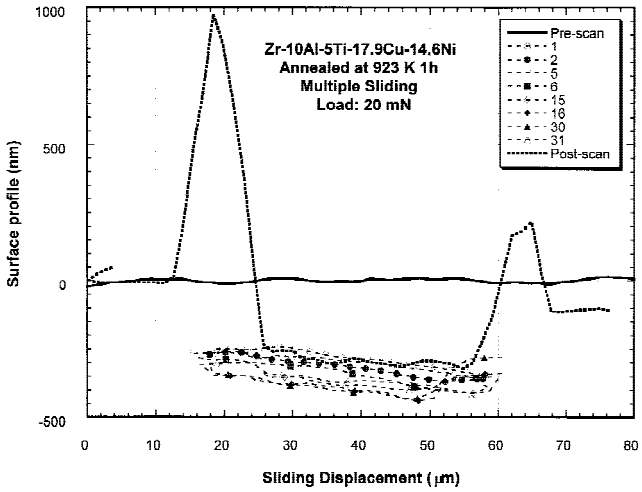


FIG. 8. Typical surface profiles of the multiple sliding wear test (thirty-one sliding scans) on sample E under the normal load of 20 mN.

oxides, absorbed films, and load conditions.[36] Existing data for the friction coefficient of metallic glasses cover a wide range; friction coefficients range from exceptionally low values (~0.015)[37] to medium values (0.2–0.6)[38,39] and to high values (0.8–1.0).[40–42] These data were obtained by using a spherical or cylindrical hard slider moving on sheet (tape or ribbon) metallic glass samples. Limited tribological data produced from ramping load scratch test using nanoindentation technique are available and they are summarized in Table IV. Results obtained in the present study are also included. It is noted in the table that critical loads exist in all materials, except the single-crystal Al and Al_2O_3–TiC. Also noted is the fact that the friction coefficient is slightly higher when the applied loading is above the critical load than when it is below the critical load.

Critical load can be considered as a measure of scratch resistance. As shown in Table IV, the present amorphous alloy apparently has a much higher critical load than the other materials, including ceramics. The critical loads of 110 and 86 mN found in samples C and D, respectively, are much higher than the 28 mN for the fully amorphous sample. This suggests that a material having a mixed amorphous–nanocrystalline microstructure is more scratch resistant than the fully amorphous sample. The postscan depth measurements from the MSW tests also confirmed that samples with an amorphous–nanocrystalline structure have the least wear loss; that is, the best wear resistance. Similar results have been observed in the study of the wear resistance of metallic glass ribbons[24–26] and Al-based bulk amorphous alloys.[27,43]

Kim *et al.*[44] recently demonstrated that a homogeneous dispersion of nanoscale crystalline particles embedded in Al-based amorphous matrix resulted in a significant increase in tensile strength, as compared to the corresponding single-phase, amorphous alloys. Similar results have also been reported in $Zr_{41}Ti_{13}Ni_{10}Cu_{13}Be_{23}$,[45] $Zr_{55}Al_{10}Ni_5Cu_{30}$,[46] and $Zr_{65}Al_{10}Ni_{10}Cu_{15}$[47] amorphous alloys. Apparently, a mixed amorphous–nanocrystalline structure offers high strength and hardness, and thus good wear resistance. This can be rationalized in the following way.

It is well-known that the Hall–Petch (HP) equation can describe the relationship between hardness (or yield strength) and grain size in conventional polycrystalline solids. The HP relation predicts that hardness should increase with decreasing grain size. The basic HP relation, given in Eq. (1), predicts that strength or hardness, H, should increase with decreasing grain size, d.

$$H = H_0 + Kd^{-1/2} \quad , \tag{1}$$

where H_0 if the intrinsic hardness and K is a material constant. Obviously, the equation has limitations because the strength cannot increase indefinitely with decreasing grain size, as shown in Fig. 9. From a practical viewpoint, for example, the strength value cannot exceed the theoretical strength; that is, the strength of a perfect whisker. Also, from another point of view, when the grain size approaches zero, the material essentially becomes amorphous. The grain boundary strengthening effect (HP relationship) will then disappear. There is some evidence indicating that the hardness (or strength) of an amorphous material is lower than its crystalline counterpart. For example, amorphous Al_2O_3 is softer than polycrystalline Al_2O_3,[48,49] and amorphous Si is softer than polycrystalline Si.[50] Thus, there must exist an optimum microstructure that yields the maximum hardness, as shown in Fig. 9. In the present study, it is a mixed amorphous–nanocrystalline structure that yields the maximum hardness. It is worth noting that the strength (or hardness) of a partially crystallized Zr–Al–Cu–Pd alloy was found to exhibit a maximum when the size of Zr_2(Cu, Pd) precipitate particles was ~10 nm and the volume fraction of the crystalline phase was about 75%.[51,52] This is consistent with the microstructures observed in samples C and D in the present study.

Data obtained in the present study indicate that there exists a general correlation between the hardness, friction coefficient, and wear resistance: a lower hardness produces a higher friction coefficient, and a higher friction results in a faster wear. The as-cast amorphous sample exhibits the lowest hardness and, thus, highest friction coefficient and poorest wear resistance. This result is consistent with the observations in some crystallized materials and metallic glass ribbons.[21,42] Comparing Tables II and III, it is noted that friction coefficient meas-

ured from the multiple sliding wear test is normally greater than that measured from the ramping load test. This is probably associated with the formation of a "transfer layer" on a scratched surface as suggested by Kishore *et al.*[17,53] In essence, a transfer layer produced during a slide can be removed by a subsequent slide at a load greater than the "critical load"; this results in a reduced friction coefficient. In the present study, increasing load was used during the ramping load tests. In contrast, the multiple sliding wear tests were always conducted at loads below the "critical load." However, it is pointed out that, despite the difference in the absolute value of friction coefficient, the general trend for these two measurements are consistent; namely, a higher hardness value leads to a lower friction coefficient.

V. CONCLUSIONS

Tribological (friction and wear) behavior of a bulk amorphous alloy Zr–10Al–5Ti–17.9Cu–14.6Ni (at.%), both in the as-cast and annealed states, was characterized by using ramping load scratch and multiple sliding wear tests. This study led to the following conclusions.

(1) The crystallization sequence of the alloy is described by

$$\begin{aligned} \mathrm{Am} &\xrightarrow{683\ \mathrm{K}} \mathrm{Am}' \xrightarrow{783\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} + [\mathrm{Ni, (Zr, Ti)}] \\ &\xrightarrow{893\ \mathrm{K}} \mathrm{Am}' + \mathrm{NiZr_2} + [\mathrm{Ni, (Zr, Ti)}] + \mathrm{Zr(AlNi)_2} \\ &\quad + \mathrm{NiZr} \xrightarrow{923\ \mathrm{K},\ 1\ \mathrm{h}} \mathrm{Zr_2Cu} + \mathrm{NiZr_2} + \mathrm{Ni_{42}Zr_{58}} \\ &\quad + \mathrm{Al_3Zr_4} \quad . \end{aligned}$$

(2) Both Young's modulus and hardness increase with increasing annealing temperature; the modulus and hardness values are, in the ascending order, amorphous, supercooled liquid, partially crystallized, and fully crystallized.

(3) Critical loads associated with microcracking were observed in all test samples. The as-cast material has the lowest critical load, and the critical load increases with increased annealing temperature.

(4) There exists a general correlation between the hardness, friction coefficient, and wear resistance; a lower hardness produces a higher friction coefficient and thus a faster wear.

(5) The present experimental results indicate that wear resistance is (in an ascending order): as-cast amorphous, supercooled liquid, fully crystallized, and

TABLE III. Steady-state friction coefficient for the forward sliding (μ_F) and postscan depth (d_p) under a normal sliding load of 20 mN in multiple sliding wear tests.

Sample	Microstructure	μ_F	H (GPa)	d_p (nm)
A: As-cast	Am	0.42	6.2–6.3	600 ± 20
B: 683 K annealed	Am′	0.30	6.5–6.9	370 ± 10
C: 783 K annealed	Am′ + $NiZr_2$ + [Ni, (Zr, Ti)]	0.25	7.8–8.3	250 ± 20
D: 893 K annealed	Am′ + $NiZr_2$ + [Ni, (Zr, Ti)] + $Zr(AlNi)_2$ + NiZr	0.15	8.0–8.6	270 ± 20
E: 923 K annealed (1 h)	Zr_2Cu + $NiZr_2$ + $Ni_{42}Zr_{58}$ + Al3Zr_4	0.27	8.3–8.6	300 ± 20

TABLE IV. Friction coefficient and critical load for several bulk materials measured from ramping load scratch test.

Material	Ramping load (mN)	Critical load (mN)	Friction coefficient	
			Baseline	Maximum
Single-crystal Si(100)	2–25	16	0.10	0.80
Ni–Zn ferrite	2–25	10	0.10	0.80
Al_2O_3–TiC	2–25	From beginning	0.15	0.30
SiC	2–25	15	0.12	0.75
Single-crystal Al	0.02–15	From beginning	0.35	0.50
A: As-cast	0.02–80	28	0.20	0.32
B: 683 K annealed	0.02–80	46	0.17	0.28
C: 783 K annealed	0.02–80	⋯	0.150	0.155
	0.02–150	110	0.18	0.28
D: 893 K annealed	0.02–80	⋯	0.150	0.15
	0.02–150	86	0.18	0.23
E: 923 K annealed (1 h)	0.02–80	⋯	0.14	0.14

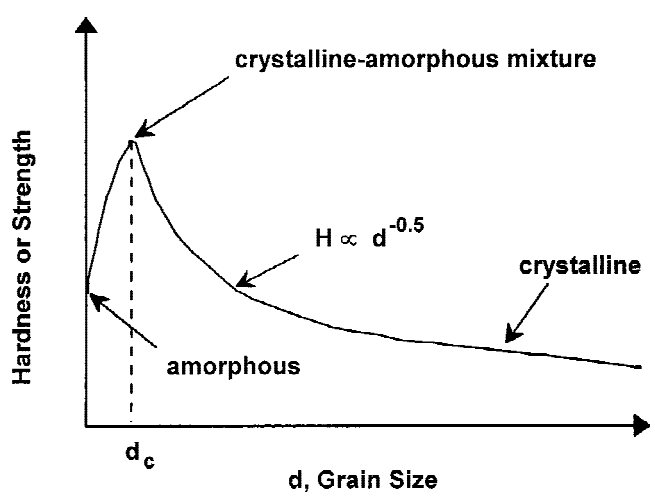


FIG. 9. Schematic illustration of Hall–Petch relationship for the metallic glass from amorphous to crystalline state.

amorphous–nanocrystalline mixture. Samples annealed at 783 and 893 K have a mixed amorphous–nanocrystalline structure and thus are the most wear resistant.

ACKNOWLEDGMENT

This work was performed under the auspices of the United States Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

REFERENCES

1. A. Inoue, T. Zhang, and T. Masumoto, Mater. Trans., JIM **30,** 965 (1989).
2. A. Inoue, H. Yamaguchi, T. Zhang, and T. Masumoto, Mater. Trans., JIM **31,** 104 (1990).
3. A. Inoue, K. Ohtera, K. Kita, and T. Masumoto, Jpn. J. Appl. Phys. **27,** L2248 (1988).
4. S.G. Kim, A. Inoue, and T. Masumoto, Mater. Trans., JIM **31,** 929 (1990).
5. A. Inoue, T. Zhang, and T. Masumoto, Mater. Trans., JIM **31,** 177 (1990).
6. T. Zhang, A. Inoue, and T. Masumoto, Mater. Trans., JIM **32,** 1005 (1991).
7. A. Inoue, T. Shibata, and T. Zhang, Mater. Trans., JIM **36,** 1420 (1995).
8. A. Inoue, N. Nishiyama, K. Amiya, T. Zhang, and T. Masumoto, Mater. Lett. **19,** 131 (1994).
9. X.H. Lin and W.L. Johnson, J. Appl. Phys. **78,** 6514 (1995).
10. A. Peker and W.L. Johnson, Appl. Phys. Lett. **63,** 2342 (1993).
11. A. Inoue, N. Nishiyama, and T. Matsuda, Mater. Trans., JIM **37,** 181 (1996).
12. A. Inoue and J.S. Gook, Mater. Trans., JIM **36,** 1180 (1995).
13. A. Inoue and J.S. Gook, Mater. Trans. JIM **36** 1282 (1995).
14. A. Inoue, Mater. Trans., JIM **36,** 866 (1995).
15. A. Inoue, Mater. Sci. Eng. A **226–228,** 357 (1997).
16. A. Inoue, A. Takeuchi, and T. Zhang, Metall. Mater. Trans. A **29,** 1779 (1998).
17. U. Sudarsan, K. Chattopadhyay, and Kishore, in *Proceedings of the Fifth International Conference on Rapidly Quenched Metals,* edited by S. Steed and H. Warlimont (North-Holland, Amsterdam, The Netherlands, 1985), p. 1439.
18. A.L. Greer, Science **267,** 1947 (1995).
19. K. Ozawa, H. Walkasugi, and K. Tanaka, IEEE Trans. Magn. **20,** 425 (1984).
20. O. Kohmoto, K. Ohya, and T. Ojima, IEEE Trans. Magn. **25,** 4490 (1989).
21. D.G. Morris, in *Proceedings of the Fifth International Conference on Rapidly Quenched Metals,* edited by S. Steed and H. Warlimont (North-Holland, Amsterdam, The Netherlands, 1985), p. 1775.
22. N. Dolezal and G. Hausch, in *Proceedings of the Fifth International Conference on Rapidly Quenched Metals,* edited by S. Steed and H. Warlimont (North-Holland, Amsterdam, The Netherlands, 1985), p. 1767.
23. C. Atzeni, M. Pes, U. Sanna, A. Corrias, G. Paschina, and D. Zedda, Adv. Perform. Mater. **1,** 243 (1994).
24. K.L. Rutherford and I.M. Hutchings, J. Test. Eval. **25,** 250 (1997).
25. C.J. Wong and J.C.M. Li, Wear **98,** 45 (1984).
26. M. Anis, W.M. Rainforth, and H.A. Davies, Wear **172,** 135 (1994).
27. T. Gloriant and A.L. Greer, Nanostruct. Mater. **10,** 389 (1998).
28. C.T. Liu, L. Heatherly, D.S. Easton, C.A. Carmichael, J.H. Schneibel, C.H. Chen, J.L. Wright, M.H. Yoo, J.A. Horton, and A. Inoue, Metall. Mater. Trans. A **29,** 1811 (1998).
29. B.N. Lucas, W.C. Oliver, G.M. Pharr, and J-L. Loubet, in *Thin Films: Stresses and Mechanical Properties VI,* edited by W.W. Gerberich, H. Gao, J-E. Sundgren, and S.P. Baker (Materials Research Society, Pittsburgh, PA, 1997), p. 233.
30. W.C. Oliver and G.M. Pharr, J. Mater. Res. **7,** 1564 (1992).
31. J.G. Wang, B.W. Choi, T.G. Nieh, and C.T. Liu, J. Mater. Res. **15,** 798 (2000).
32. B. Bhushan, B.K. Gupta, and M.H. Azarian, Wear **181–183,** 743 (1995).
33. B.K. Gupta and B. Bhushan, Wear **190,** 110 (1995).
34. H. Deng, T.W. Scharf, and J.A. Barnard, IEEE Trans. Magn. **33,** 3151 (1997).
35. P.H. Gaskell, J. Non-Cryst. Solids **32,** 207 (1979).
36. J. Larsen-Basse, in *Friction, Lubrication, and Wear Technology,* edited by P.J. Blau (ASM International, The Materials Information Society, Materials Park, OH, 1995), p. 24.
37. J.K.A. Amuzu, J. Phys. D: Appl. Phys. **13,** L127 (1980).
38. D-H. Lee and J.E. Evetts, Acta Metall. **32**(7), 1035 (1984).
39. G. Girerd, P. Guiraldenq, and N. Du, Wear **102,** 233 (1985).
40. S.H. Whang and B.C. Giessen, in *Rapidly Solidified Amorphous and Crystalline Alloys,* edited by B.H. Kear, B.C. Giessen, and M. Cohen (New York, North-Holland, 1982), p. 301.
41. R. Klinger and H.G. Feller, Wear **86,** 287 (1983).
42. R. Moreton and J.K. Lancaster, J. Mater. Sci. Lett **4,** 133 (1985).
43. K. Ohtera, A. Inoue, T. Terabayashi, H. Nagahama, and T. Masumoto, Mater. Trans., JIM **33,** 775 (1992).
44. Y.H. Kim, A. Inoue, and T. Masumoto, Mater. Trans., JIM **31,** 747 (1990).
45. H.J. Fecht, Philos. Mag. B **76,** 495 (1997).
46. A. Inoue, Sci. Rep. RITU, **A42,** 1 (1996).
47. Y. Kawamura, H. Kato, A. Inoue, and T. Masumoto, Appl. Phys. Lett. **67,** 2008 (1995).
48. J-E. Sundgren and H.T.G. Hentzell, J. Vac. Sci. Technol., A **4,** 2259 (1986).
49. T.C. Chou, T.G. Nieh, S.D. McAdams, G.M. Pharr, and W.C. Oliver, J. Mater. Res. **7,** 2774 (1992).
50. G.M. Pharr, Rice University (private communication, 1998).
51. C. Fan and A. Inoue, Mater. Trans., JIM **38,** 1040 (1997).
52. C. Fan, A. Takeuchi, and A. Inoue, Mater. Trans., JIM **40,** 42 (1999).
53. Kishore, U. Sudarsan, N. Chandran, and K. Chattopadhyay, Acta Metall. **35,** 1463 (1987).

ELSEVIER

Intermetallics 8 (2000) 1215–1224

Intermetallics

www.elsevier.com/locate/intermet

An empirical model for ideal work of adhesion: transition-metal aluminides and silicides

M.H. Yoo *, K. Yoshimi [1]

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6115, USA

Received 1 November 1999; accepted 31 May 2000

Abstract

The role of atomic bonding in the brittle fracture of intermetallic alloys has been assessed theoretically. A simple empirical model is proposed for the ideal work of adhesion, resulting from a rigid-body separation, in terms of four variables, viz. the elastic stiffness constant, the equilibrium interplanar spacing, and two scaling length parameters. The ratio of these two length parameters is assessed based on the available results of ab initio slab-supercell calculations. Ideal cleavage energies and critical stress intensity factors of transition-metal aluminides and silicides are estimated, and the results are discussed by comparing with the available experimental data of brittle fracture. The difference between the proposed model and other models, temperature dependence of surface energies, and an extension of this model for interfacial adhesion are also discussed. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Aluminides, miscellaneous; A. Silicides, various; B. Anisotropy; B. Fracture stress; E. Physical properties, miscellaneous

1. Introduction

Transition-metal (TM) aluminides and silicides belong to an important class of material systems because of the use of these silicides in microelectronics industry and the aluminides for possible applications as structural components. Though these TM aluminides and silicides are potential candidates for structural applications at ultra-high temperatures because of their high melting points, relatively low density and good oxidation resistance, the major drawback has been their intrinsic brittleness at room temperature [1,2]. In order to better understand the brittle fracture behavior in silicides of the noncubic crystal structures at low temperatures, specific surface energies of low index habit planes are needed. The purpose of this work is to introduce an empirical method of estimating orientation-dependent uniaxial cohesive energies, or ideal cleavage energies, for TM aluminides and silicides. A brief description of this empirical method was introduced in a recent conference [3].

Experimental data on surface energies of TM aluminides and silicides are not available [4]. A number of different empirical methods have been proposed to correlate surface energies with physical properties such as the heat of fusion, Debye temperature, and the melting point [5,6]. These methods, however, cannot furnish information on the dependence of surface energies on crystallographic planes. On the other hand, electronic structure calculations can give quantitative description on the orientation dependence of surface energy in metals [6–11]. By using the so-called slab model, i.e. free-standing thin film, surface energies of a number of low-index planes in some TM aluminides and silicides were determined by ab initio calculations within the framework of the local density functional (LDF) theory, e.g. in Ni_3Al [12], Ni_3Si [13], TiAl [14], Ti_3Al [15], and $CoSi_2$ [16]. In Al_3Sc [17], NiAl and FeAl [18], and $MoSi_2$ [19], the maximum tensile stresses (or theoretical strength) and uniaxial cohesive energies (or ideal work of adhesion) of a few low-index habit planes were determined as functions of the separation distance by using the so-called supercell model. While most of the experimental techniques to measure surface energies are

* Corresponding author. Tel.: +1-865-574-5165; fax: +1-865-574-7659.

E-mail address: yoo@ornl.gov (M.H. Yoo).

[1] Present address: Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan.

0966-9795/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0966-9795(00)00082-0

applicable at relatively high homologous temperatures, such as zero-creep technique, theoretical calculations refer to surface energies at the ground state. Therefore, any information on the temperature dependence of surface energies would be useful in comparing the results from theory and experiment.

In this paper, we first present an empirical method of estimating ideal cleavage energies on the basis of the available results of supercell calculations. Second, ideal cleavage energies of TMs and TM aluminides and silicides of cubic and noncubic crystal structures are estimated, and these are compared with the data available in the literature. Third, we discuss temperature and composition dependence of cleavage energies. Finally, the validity and usefulness of the proposed method are discussed in light of the available experimental data and other methods of approximation for surface energy.

2. Uniaxial cohesive stress function

Consider a crystalline slab as illustrated in Fig. 1(a) where the open arrows indicate the direction of uniaxial tensile stress, σ, as a function of the displacement, v, along the y-axis. The three material-specific length parameters of d_0, a, and D represent the equilibrium interplanar spacing of (hkl) planes, an effective range of interatomic cohesive forces, and a scaling factor related to the nonlinear nature of $\sigma(v)$, respectively. The change in total energy of a slab, ΔE, as it is separated into two halves, is to increase with increasing separation distance, v, as shown in Fig. 2(a). The corresponding tensile stress function is shown in Fig. 2(b) such that the ideal cleavage energy is obtained by

$$G_c = \int_0^\infty \sigma(v)\mathrm{d}v = \gamma_1 + \gamma_2 \tag{1}$$

where γ_1 and γ_2 are the specific energies per unit area of the two surfaces created, as indicated in Fig. 1(b). In

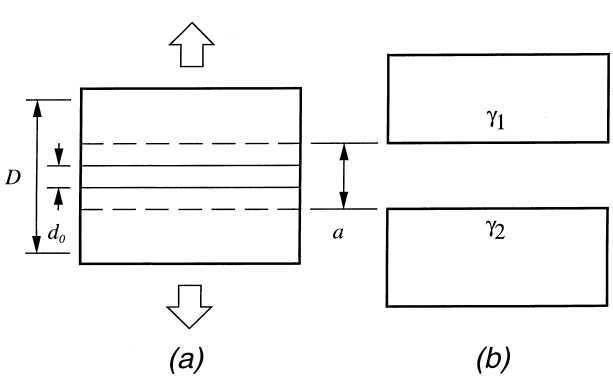


Fig. 1. Schematic illustration of (a) a supercell unit and (b) two parts after separation.

ordered intermetallic compounds, these two surface energies are equal only at some special crystallographic orientations. In addition, for some (hkl) planes in noncubic crystals, there are more than one choice of atomically distinct cleavage plane.

If the cohesive stress function, $\sigma(v)$, is known, then the ideal cleavage energy (work of adhesion) can be obtained from Eq. (1). But, in reality, it is the change in total energy as a function of separation distance, $\Delta E(v)$, that can be calculated by a supercell calculation. This result becomes then the basis for constructing the stress function, $\sigma(v)$. The so-called universal binding energy relation (UBER) [20,21], which is known valid for metallically and covalently bonded solids, can be used to fit the ΔE–v data of supercell calculations by

$$\Delta E(v) = G_c[1 - (1 + d^*)\exp(-d^*)], \tag{2}$$

where d^* is defined in terms of a scaling length (or the electronic screening length [20]), l, as $d^* = v/l$. The uniaxial tensile stress function is given by

$$\sigma(v) = \sigma_m d^* \exp(1 - d^*). \tag{3}$$

The peak tensile strength and the cleavage energy are related to each other simply by [22]

$$G_c/\sigma_m = el, \tag{4}$$

where e is the base of natural logarithm. The interfacial energy results of supercell calculations for (001) planes in Mo, Nb, V, and $MoSi_2$ were found to fall accurately on a single UBER curve when the scaling length is in the range of $l = 0.63$–0.70 Å [19].

In this paper, we are interested in the anisotropy of cleavage energy for a given TM aluminides or silicide. In using a triangular approximation for the stress function in Fig. 2(b), we assume that (i) the maximum slope of the ΔE–v curve (i.e. $\sigma = \sigma_m$) occurs at $v = l$, (ii) at $v < l$, the nonlinear $\sigma(v)$ is approximated by the reduced elastic stiffness of $C_{22}(d_0/D)$ in the direction perpendicular to (hkl) planes, and (iii) the long-range interatomic forces (at $v > l$) diminish effectively to zero at $v \geqslant a$. On the basis of these three assumptions, the tensile cohesive stress may be given simply as

$$\sigma = C_{22}\left(\frac{v}{D}\right) \qquad \text{for } 0 \leqslant v \leqslant l, \tag{5a}$$

$$\sigma = \frac{C_{22}}{D}\left(\frac{l}{a-l}\right)(a - v) \qquad \text{for } l < v \leqslant a, \tag{5b}$$

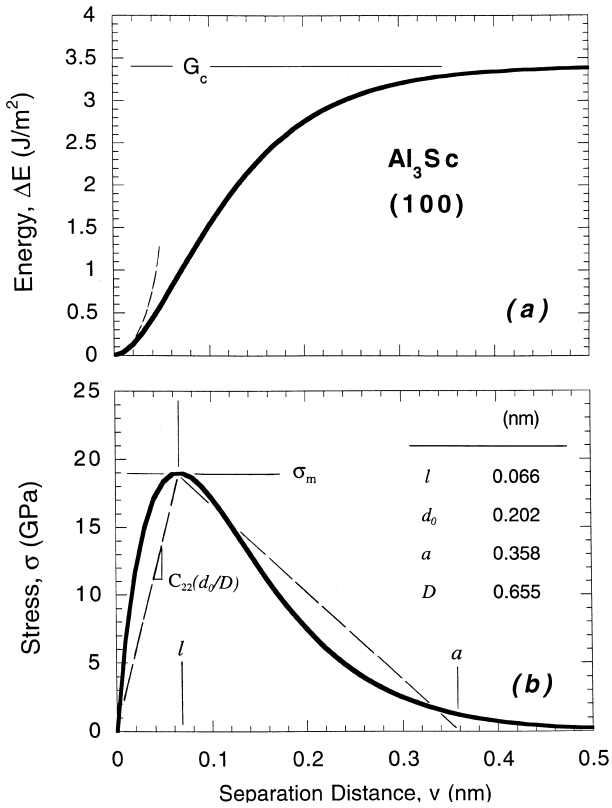


Fig. 2. (a) Interfacial energy and (b) tensile stress for the (100) plane in Al_3Sc based on the ab initio supercell calculations [17].

$$\sigma = 0 \qquad \text{for } v > a, \tag{5c}$$

where C_{22} is the transformed elastic stiffness constant with respect to the tensile y-axis. Use of the above stress function based on the three assumptions and Eq. (1) gives the maximum stress and the cleavage energy as

$$\sigma_m = C_{22}\left(\frac{l}{D}\right), \tag{6}$$

$$G_c = \frac{C_{22}l}{2}\left(\frac{a}{D}\right). \tag{7}$$

Eq. (6) for the ideal cleavage strength is equivalent to a number of such expressions [23–26] based on the original Orowan–Polanyi approximation for the theoretical tensile strength. The elastic stiffness constants in place of Young's modulus were used earlier for the case of hcp metals [27]. The dashed curve of Fig. 2(a) is the harmonic range of the energy increase corresponding to the elastic constant C_{22}. The Orowan–Polanyi approximation usually makes an overestimate of the theoretical strength in metals. For instance, in the case of TM aluminides, the estimated cleavage strength in the range of $\sigma_m = (0.12–0.18)\ C_{22}$ was found to be slightly larger than those determined from the slab-supercell calculations [14]. The proposed triangular approximation is designed to match the maximum stresses determined from supercell calculations by use of the D parameter, which may be related to the anharmonic property at the "bond-stretching" part ($v < l$) of Fig. 2.

The area under a triangular stress function is chosen exactly the same as that under a $\sigma(v)$ curve, as illustrated in Fig. 2(b) by the two dashed lines in compliance with the solid curve. According to the prescription of scaling for the UBER [20,21], σ_m occurs at $v = l$. The supercell calculations for TM aluminides [17,18] showed that ΔE converges to different G_c values for different cleavage habit planes. Therefore, the parameter a, which is related to the "bond-breaking" part ($v > l$) of Fig. 2, is to be correlated not only to the actual magnitude of cleavage energy for a given compound, but also to the anisotropy of cleavage energy in the compound. This parameter has a simple relationship with the scaling length, l,

$$\frac{a}{l} = 2e. \tag{8}$$

Use of this in Eq. (7) gives an alternative expression of cleavage energy,

$$G_c = eC_{22}\left(\frac{l^2}{D}\right). \tag{9}$$

3. Determination of length parameters

In this section, the length parameters involved in the triangular approximation are determined for three bcc metals (Mo, V, and Nb), two B2 aluminides (NiAl and FeAl), two $L1_2$ aluminides (Al_3Sc and Ni_3Al), and one $C11_b$ silicide ($MoSi_2$).

Fig. 3 and Table 1 summarize the supercell calculations for (001) cleavage in Mo, V, Nb, and $MoSi_2$ by Hong et al. [19]. In Tables 1–3, numerical factors of $\lambda = l/d_0$ are listed. $MoSi_2$ has the tetragonal $C11_b$ structure, which consists of three pseudo-bcc unit cell stacked along the [001] direction. The data of $MoSi_2$ are referred to fracture between the Si (001) plane, and hence $G_c = 2\gamma$.

Fig. 4 and Table 2 summarize the results of supercell calculations for NiAl and FeAl reported by Yoo and Fu [18]. In an AB compound of the B2 structure, one of the (100) cleavage surfaces consist of all A atoms and the other with all B atoms, accordingly, $G_c = \gamma_1 + \gamma_2$. In contrast, all {110} surfaces being equal in the A + B composition, we have $G_c = 2\gamma$ for the case of (110) cleavage in stoichiometric B2 compounds.

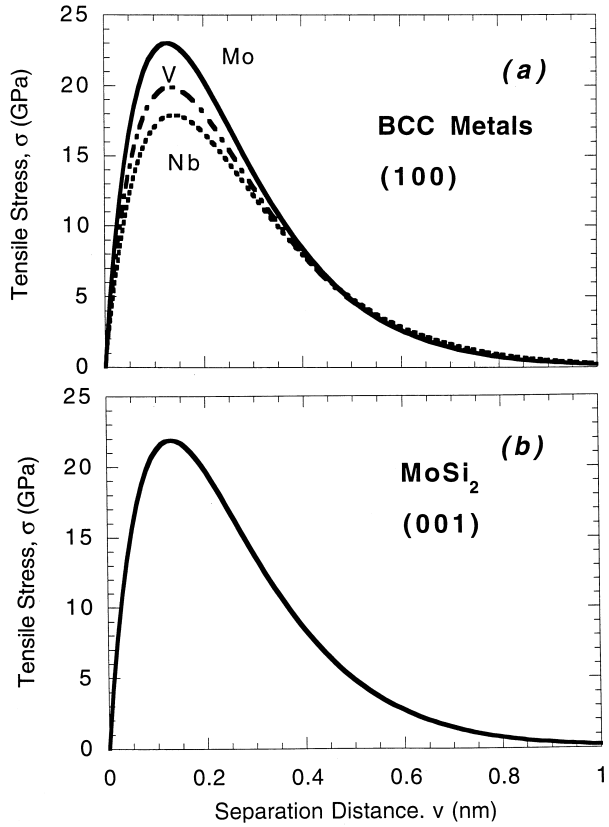


Fig. 3. Tensile stress vs. cleavage displacement in (a) bcc TM metals and (b) $MoSi_2$ based on the ab initio supercell calculations [19].

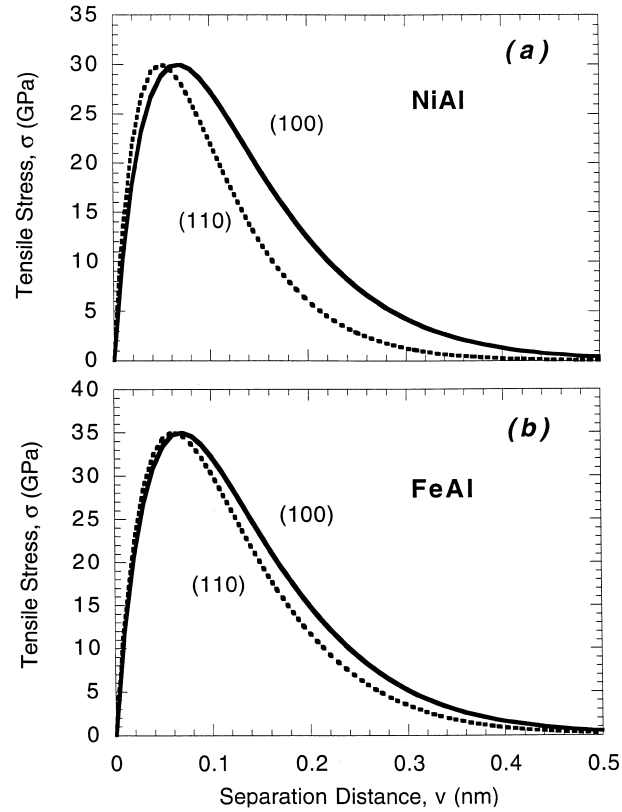


Fig. 4. Tensile stress vs. cleavage displacement in (a) NiAl and (b) FeAl based on the ab initio supercell calculations [18].

Table 3 lists the data for Al_3Sc [17] and Ni_3Al [12] of the $L1_2$ structure. Cleaving on (100) or (110) plane of an A_3B crystal of the $L1_2$ structure creates two surfaces of different chemical composition, one containing all A atoms and the other of A + B atom mixture. Therefore, $G_c = \gamma_1 + \gamma_2$ for both cases. The $\sigma(v)$ curves for Al_3Sc and Ni_3Al are very similar in shape to Fig. 4(b) and 2(b), respectively, with an exception that the (100) curve is slightly lower than the (110) curve in Al_3Sc.

Table 4 summarizes all the length parameters involved in Eqs. (6), (7), and (9). The elastic constants used for metals are from the experimental data either at

Table 1
Ideal work of fracture, theoretical strength, and length parameters for (001) cleavage in the BCC case [19]

Metals	(*hkl*)	G_c (J/m^2)	σ_m (GPa)	d_o (nm)	l	λ
Mo	200	7.94	23.0	0.157	0.063	0.404
V	200	7.20	19.9	0.152	0.067	0.438
Nb	200	6.82	17.9	0.165	0.070	0.425
$MoSi_2$	006	7.72	21.9	0.131	0.065	0.495

Table 2
Ideal work of fracture, theoretical strength, and length parameters in the B2 case [18]

Aluminides	(*hkl*)	G_c (J/m^2)	σ_m (GPa)	d_o (nm)	l	λ
NiAl	200	5.5	30	0.141	0.067	0.478
	110	4.1	30	0.199	0.050	0.253
FeAl	200	6.5	35	0.142	0.068	0.481
	110	5.8	35	0.200	0.061	0.348

Table 3
Ideal work of fracture, theoretical strength, and length parameters in the $L1_2$ case [12,17]

Aluminides	(*hkl*)	G_c (J/m^2)	σ_m (GPa)	d_0 (nm)	l	λ
Al_3Sc	200	3.4	19	0.202	0.066	0.326
	110	3.7	19	0.286	0.072	0.250
Ni_3Al	200	5.8	30	0.175	0.071	0.406

4 K or their extrapolations to 0 K [28]. Unless mentioned otherwise, the elastic constants used for intermetallic compounds are from the tabulation of theoretical results at the ground state [29]. The elastic constants of NiAl and FeAl are from Fu and Yoo [30]. Among eleven cases representing eight different materials, the numerical value of λ ranges from 0.25 to 0.50, where the low bound is represented by Al_3Sc (110) and the high bound by $MoSi_2$ (006). Among the three bcc metals, the peak stress increases ($\sigma_m = 18$, 20, and 23 GPa for Nb, V, and Mo) while the peak position decreases slightly ($l = 0.70$, 0.67, and 0.63 Å). This increase can be correlated to the increase in the reduced elastic stiffness constant, viz., $C_{22}(d_0/D) = 42$, 45, and 57 GPa for Nb, V, and Mo, respectively.

In the B2 case, cleavage energy of the (110) plane is lower than that of the (100) plane. Though only slight, the anisotropy is opposite in the case of Al_3Sc. The rather pronounced anisotropy of G_c (about 30%) in NiAl, while the $\sigma_m = 30$ GPa is the same, can be correlated to both the a and D parameters. That is, a is larger for the (100) plane than for the (110) plane by 34% and $C_{22}(d_0/D)$ is smaller by a factor of two.

4. Ideal cleavage energy and strength

4.1. Transition metals

Table 5 lists estimated values of (100) cleavage energies for four bcc (Ta, Cr, W, and Fe) and four fcc (Ni, Pd, Pt, and Al) metals. The numerical value of l^2/D used for Ta is the algebraic average of the two values for V and Nb. Two different values of l^2/D (corresponding to Nb and Mo) were used for Fe. Because no supercell calculation database is available for fcc metals, $l^2/D = 0.908 \times 10^{-2}$ determined for Ni_3Al was used for Ni, Pd, and Pt. The two l^2/D values determined for (100) and (110) planes of Al_3Sc were used for the two planes of Al.

The estimated results of G_c by this study are compared with the available results of first-principles slab calculations within the local density approximation (LDA), which are listed in Table 5. Also listed in Table 5 are the predicted values from the so-called equivalent-crystal-theory (ECT) [31] and other values obtained using semi-empirical methods, such as the embedded atom method (EAM) [32] and Finnis–Sinclair method [33]. Except for Pd and Pt [34], all other data of the three columns (LDA, ECT, and Other) in Table 5 are obtained from Table III of Smith et al. [31].

When compared to LDA values, the estimated G_c values for (100) plane in Fe and Ni, and (110) plane in Al are larger by about 6 and 12%, respectively, and the estimated G_c value for (100) plane in W is smaller by about 8%. While the results from ECT are generally larger than the LDA results by about 10% on average, the predictions of G_c values by other semi-empirical methods are 45–55% lower than the LDA results. In the empirical method proposed in this paper, transferring of the l^2/D values from Ni_3Al to Ni and from Al_3Sc to Al

Table 4
Elastic constants and length parameters

Material	(*hkl*)	C_{22} (GPa)	D (nm)	a	$(l^2/D)\times 10^2$	$\frac{D}{d_0}$	$\frac{l}{D}\times 10^2$
Mo	200	450	0.621	0.345	0.649	7.89	5.1
V	200	232	0.389	0.362	1.140	5.13	8.6
Nb	200	253	0.495	0.381	0.993	6.00	14.0
$MoSi_2$	006	536	0.794	0.353	0.530	12.13	4.1
NiAl	200	233	0.524	0.367	0.868	3.73	12.9
	110	318	0.533	0.273	0.474	2.68	9.4
FeAl	200	290	0.566	0.371	0.825	4.00	12.1
	110	375	0.653	0.331	0.569	3.26	9.3
Al_3Sc	200	189	0.655	0.358	0.662	3.24	10.5
	110	182	0.686	0.390	0.748	2.40	10.4
Ni_3Al	200	235	0.557	0.387	0.908	3.19	12.8

Table 5
Estimate of (100) ideal cleavage energy for transition metals and Al

Metals	C_{22} (GPa)	$(l^2/D)\times10^2$ (nm)	This study	LDA	ECT	Other
			G_c (J/m^2)			
V[a]	232	1.140	7.20	7.20[a]		
Nb[a]	253	0.993	6.82	6.82[a]		
Ta	266	1.066	7.72			
Cr	391	0.649	6.90			
Mo[a]	450	0.649	7.94	7.94[a]		
W	533	0.649	9.40	10.20	11.76	5.85
Fe	243	0.649	4.29			
		0.993	6.56	6.20	6.98	3.39
Ni	261	0.908	6.45	6.10	6.24	3.16
Pd	234	0.908	5.78			3.82
Pt	358	0.908	8.84			4.88
Al (200)	116	0.662	2.09		2.58	
(110)	121	0.748	2.47	2.20	2.32	

[a] Database from ab initio supercell calculations [19] available for this study.

is a reasonable exercise in view of the good agreement between the estimated values and the LDA results. Also, the reasonable agreement between the estimated value of $G_c = 6.56$ J/m^2 and the LDA result of $G_c = 6.10$ J/m^2 indicates that Fe is more similar to Nb than to Mo as far as the l^2/D numerical factor is concerned.

4.2. TM aluminides and silicides

In addition to the eight cases of five intermetallic compounds for which all the length parameters were determined (Table 4), there are more data for G_c from *ab initio* slab calculations, from which numerical values of l^2/D can be determined. Table 6 lists these additional data for Al_3Sc [17], Ni_3Al [12], Ni_3Si [13,35], $CoSi_2$ [16], TiAl [14,36], and Ti_3Al [15,36], which contains 12 cases for the six alloys. Together with those cases for alloys listed in Table 4, a total number of cases is 20 for nine alloys.

Using appropriate values of l^2/D from Tables 4 and 6, we can estimate ideal cleavage energies for many TM aluminides and silicides of $L1_2$, $D0_{22}$, $D0_{23}$, B2, $D0_3$, and $C11_b$ structures, as listed in Table 7. Transfer of l^2/D values within the same crystal structure or from one pseudo-fcc (or-bcc) building block to another was made in the following ways: $Ni_3Al \rightarrow Pt_3Al$, $Al_3Sc \rightarrow$ all trialuminides, $(NiAl+FeAl)/2 \rightarrow$ B2 aluminides, FeAl $\rightarrow$ Fe-base $D0_3$ and $MoSi_2 \rightarrow WSi_2$. To our best knowledge, there has been no information on numerical values of G_c for the cases listed in Table 7. The elastic constants used in Table 7 are mostly from experimental data at room temperature [29]. Theoretically determined elastic constants at 0 K were used for Pt_3Al [37], Al_3Ti ($L1_2$) [17], and RuAl and CoAl [38].

Though there is no experimental database for G_c values, a comparison can be made between theoretically-determined Griffith strength and experimentally-measured fracture toughness. Table 8 lists such results for a comparison between $MoSi_2$ and WSi_2. The critical stress intensity factor for mode-I (*hkl*)[*uvw*] Griffith crack, k_{IG}, was determined by taking full account of the elastic anisotropy [13]. The fracture toughness values,

Table 6
Cleavage energy and the ratio of length parameters

Alloys	(*hkl*)	G_c (J/m^2)	C_{22} (GPa)	l^2/D (10^{-2})
Ni_3Al	111	4.6	351	0.482
Ni_3Si	200	7.2	375	0.706
	111	5.1	481	0.390
Al_3Sc	111	3.3	180	0.675
$CoSi_2$	400	5.96	277	0.792
	220	4.62	309	0.550
	222	4.34	320	0.499
TiAl	200	4.6	188	0.900
	002	5.6	190	1.084
	110	5.3	243	0.802
	111	4.5	283	0.585
Ti_3Al	002	4.8	238	0.742

Table 7
Estimate of ideal cleavage energy for TM aluminides and silicides

Alloys	(*hkl*)	C_{22} (GPa)	$(l^2/D)\times10^2$ (nm)	G_c (J/m^2)
Pt_3Al	200	436	0.908	10.8
	111	479	0.482	6.3
Al_3Ti[a]	200	177	0.662	3.2
	110	212	0.748	4.3
	111	224	0.676	4.1
Al_3Ti	200	218	0.662	3.9
	110	254	0.748	5.2
	111	250	0.676	4.6
Al_3Zr	200	209	0.662	3.8
	110	242	0.748	4.9
	111	241	0.676	4.4
RuAl	200	308	0.847	7.1
	110	348	0.522	4.9
CoAl	200	251	0.847	5.8
	110	310	0.522	4.4
Fe_3Al	200	171	0.825	3.8
	110	267	0.569	4.1
Fe_3Si	200	232	0.825	5.2
	110	331	0.569	5.1
$MoSi_2$	110	454	0.522	6.4
WSi_2	110	400	0.522	5.7
	006	552	0.530	8.0

[a] $L1_2$ structure (hypothetical).

K_{IC}, were obtained from three-point bending tests of single crystals [39]. The k_{IG} and K_{IC} values listed in Table 8 indicate that the amount of crack-tip plasticity seems to be relatively larger at (001)[110] crack than at a (110)[001] crack in both WSi_2 and $MoSi_2$. This is more so in $MoSi_2$.

5. Temperature, composition, and interface

We have discussed ideal cleavage energy and strength in terms of four variables (C_{22}, d_0, l or a, and D) that are not only material specific, but also orientation dependent. So far, we have dealt with pure metals and stoichiometric compounds at the ground state and/or room temperature and G_c as the net change in interfacial energy by a rigid-body separation. The role of surface relaxation (or reconstruction) is neglected because it is known to be a minor effect, e.g. less than 2% of the W(100) surface energy and less than 4.5% of the V(100) surface energy [40].

There have been many papers during the 1980s that discuss the physical meaning of the scaling length, l, in the UBER. One possible interpretation is that the first-order anharmonic properties (such as thermal expansion) are related in a simple way to the harmonic properties (bulk modulus). In the present study, the anisotropy of this simple relationship is of primary interest.

For a given (*hkl*) plane how the two length parameters, l and D, change with temperature is not known, but the temperature dependence of G_c due to changes in elastic constants can be assessed. Fig. 5 shows temperature dependence of G_c for the (100) and (110) planes in NiAl by using the elastic constants over a temperature range of 300–1100 K [41]. The change in d_0 values on account of thermal expansion is negligibly small compared to that in elastic constants. The anisotropy of G_c is reduced significantly at higher temperatures, the G_c ratio of (110)/(100) reaching 0.93 at 1100 K, as far as the enthalpy term of surface energy is concerned. Contribution of the surface entropy may become an important factor at high temperatures, e.g. $\Delta S\approx1.5\times10^{-3}$ J/m^2 deg above 1000 K for W (001) surface [40], but such information is not available for intermetallic compounds.

Using the elastic constants of Al-rich TiAl [42], we find the temperature dependence of G_c as shown in Fig. 6. Because of the difference in elastic constants between the calculated results for stoichiometric TiAl at 0 K [38] and the experimental data on Ti–56% Al at 300–750 K [42], G_c values of the (110) and (111) planes in Fig. 6 are lower than the corresponding values in Table 6 by 11 and 13%, respectively. Thus, Fig. 6 actually represents a combined effect of temperature and hyperstoichiometry on G_c through elastic constants. The anisotropy of thermal expansion coefficients [42] is too small to make any appreciable change in the axial ratio ($c/a=1.02$). With an exception of the (100) curve, the three curves in Fig. 6 show essentially the same temperature dependence.

In many intermetallic compounds, elastic constants are known to be strongly composition dependent, e.g. Ni–Al and Fe–Al systems [43,44]. In both systems, the elastic anisotropy increases as a binary compound becomes more TM-rich. Consequently, the difference in G_c between (100) and (110) planes in a hypostoichiometric

Table 8
Cleavage energy, Griffith strength, and fracture toughness of $C11_b$ TM disilicides [39]

Disilicides	(*hkl*)	G_c (J/m^2)	[*uvw*]	k_{IG}	K_{IC}
				(MPa$-\sqrt{m}$)	
WSi_2	006	8.0	110	2.0	3.7
	110	5.7	001	1.6	3.1
$MoSi_2$	006	7.7	110	1.9	4.0
	110	6.4	001	1.7	2.4

NiAl alloy (e.g. > 54% Ni) is expected to be much larger than that shown in Fig. 5. It is not known, however, how l and D or l^2/D are affected by a high concentration of intrinsic point defects.

In the case of two-phase TiAl–Ti_3Al alloys containing γ/γ and α_2/γ interfaces, the interfacial cleavage energies can be estimated by

$$G_i = G_c - \Gamma_i - E_m \tag{10}$$

where Γ is the interfacial energy and E_m is the misfit energies based on Frank and van der Merwe method [45]. Because of the approximations involved in determining the interfacial and misfit energies, the interfacial fracture energies obtained earlier [36,46] are only estimates. Nevertheless, the results showed that interfacial fracture is more likely to occur on an 120° rotational γ/γ boundary than on the α_2/γ boundary. Use of $G_c = 4.5$ J/m^2 for TiAl (111) and $G_c = 4.8$ J/m^2 for Ti_3Al from Table 6 brings another degree of approximation because, in reality, both γ and α_2 phases are off-stoichiometric in their chemical compositions.

6. Discussion

On the total of 25 cases for three metals and nine alloys, for which the results of G_c by ab initio slab-supercell calculations are available in the literature, a systematic analysis of correlation between G_c and l^2/D was made based on the proposed triangular approximation for the cohesive stress function. Transfer of l^2/Ds from one bcc TM to another was made directly within the isoelectronic VA and VIA groups. Whereas, for fcc TMs and Al, the l^2/D values of Ni_3Al and Al_3Sc were used, respectively. In the case of TM aluminides and silicides, transfer of l^2/D from one crystal class to another within a given chemical group (mono-aluminides, monosilicides, trialuminides, or disilicides) was made by comparing pseudo-fcc and pseudo-bcc building blocks of cubic and noncubic crystal structures. Then, estimate of G_c was made for a total of eight metals (4 bcc and 4 fcc) and for 22 cases for 10 intermetallic compounds.

The lowest cleavage energy among the total of 43 cases is $G_c = 3.3$ J/m^2 for (111) plane of Al_3Sc (Table 6). In other trialuminides, essentially isotropic C_{22} and a moderate variation of l^2/D in Al_3Sc give the lowest G_c for (100) planes in both $TiAl_3$ and $ZrAl_3$ (Table 7). Relatively low cleavage energies ($G_c \leqslant 5$ J/m^2) of trialuminides stem directly from the high coordination of Al atoms. The highest cleavage energy is the estimated value of $G_c = 10.8$ J/m^2 for Pt_3Al (100), which is comparable to the LDA and ECT values for W listed in Table 5. It can be seen also in Table 5 that Pt has the highest (100) G_c value among the fcc TMs in the VIIIA group. When (100) and (111) G_c values are compared, these are larger for Pt_3Al than for Ni_3Al by factors of 1.9 and 1.4, respectively, which are due to the substantially larger elastic stiffness constants of Pt_3Al than of Ni_3Al.

The magnitude and anisotropy of G_c of RuAl are larger than those of CoAl and NiAl (Table 7). The estimates of (100) G_cs for Fe_3Al and Fe_3Si are lower than the calculated values of FeAl (Table 2) and the estimated value for Fe (Table 5) by 42 and 20%, respectively. Again, according to Eq. (6) or (9), these differences are due to the differences in C_{22} values.

When a cleaving interface creates two different surfaces, the range of interatomic interaction is estimated to be more long-range compared to the case when the two surfaces are chemically the same. A comparison of {100} planes to more close-packed {110} planes in bcc-based B2 aluminides (Table 2) is a case in point. Similarly, we notice such correlation between {100} and {111} in $L1_2$ aluminides and silicide (Tables 3 and 6), and also between (001) and (111) in TiAl (Table 6).

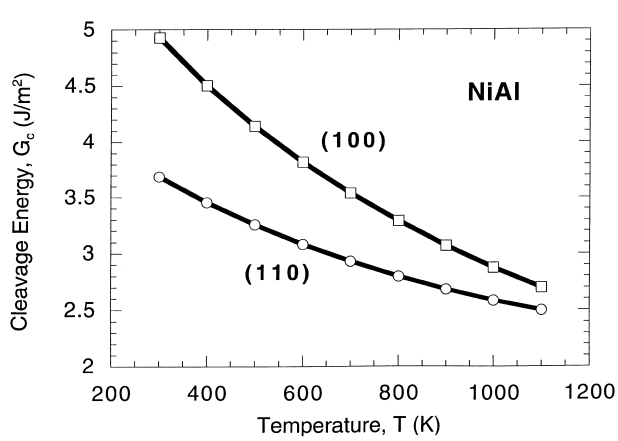


Fig. 5. Temperature dependence of cleavage energies in NiAl based on the change in elastic constants [41].

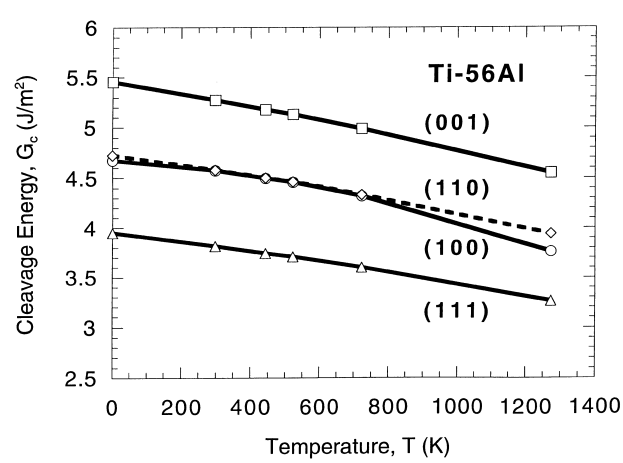


Fig. 6. Temperature dependence of cleavage energies in TiAl based on the change in elastic constants [42].

Along the close-packed (nearest-neighbor) direction, strong directional bonding between TM and Al or Si arises from the combination of charge transfer and strong p (Al or Si)–d (TM) hybridization effect [30,47].

The difference in magnitude and anisotropy of G_cs between Ni_3Si and Ni_3Al can be discussed in terms of an extra p-electrons in Ni_3Si [13], which raises the magnitude of C_{22}. Analogously, the two higher G_cs and the anisotropy in Fe_3Si as compared with Fe_3Al (Table 7) can be attributed to the role of an extra p-electron in increasing the interatomic interaction in Fe_3Si, more across the (100) plane than across the (110) plane. This is consistent with the differences in C_{22} on these two planes between Fe_3Si and Fe_3Al, i.e. C_{22} on (100) increases by 36% while C_{22} on (110) increases by 17%. In terms of G_c alone, both Ni-base and Fe-base silicides are intrinsically less brittle than the corresponding aluminides. The reason for much higher fracture toughness measured in the aluminides is obviously related to the role of crack-tip plasticity.

The present model based on a triangular approximation for cleavage strength and energy has the following relationships with other models. The first and second assumptions of the present model give an expression for the ideal cleavage strength, Eq. (6), where $\varepsilon = l/D$ can be called the "maximum elastic strain" or the "elastic limit." The magnitude of this strain in the range of $4\% < l/D < 14\%$ (Table 4) is somewhat lower than $\epsilon \approx 1/10$ [24] and $\epsilon \approx 1/\pi$ [23] estimated on the basis of the Orowan–Polanyi approximation. As shown by Eq. (8), the a parameter is larger than the l parameter of the UBER by a factor of 5.44. The so-called four point model of UBER [19] is capable of estimating ideal cleavage stresses, for a given surface orientation, without resorting to any more than four supercell calculations. The model proposed here introduces the elastic stiffness constant as one of the four variables in such a way that the combined length parameter (l^2/D), together with C_{22}, can be used to approximately describe the anisotropy of cleavage energy.

As in the case of bcc and fcc TMs discussed earlier (Table 5), the G_c values of B2 and $L1_2$ intermetallics estimated using the EAM interatomic potentials are generally lower than those given by LDA calculations. For instance, as compared to the G_c values listed in Table 2, those for (100) and (110) surfaces in NiAl estimated by Ludwig and Gumbsch [48] are lower by 36 and 33%, respectively, and the G_c values given by Farkas [49] are lower by 28 and 17%. In Ni_3Al, Foiles and Daw [50] estimated $G_c = 3.51$ and 3.30 J/m^2 for (100) and (111) surfaces, respectively, which are lower than the corresponding G_c values in Tables 3 and 6 by 39 and 28%. The reason for these large discrepancies are not known.

Experimental data on cleavage habit planes are scarce, but the available information is generally consistent with the calculated G_c values in Tables 1–3 and 6 as was discussed earlier [13,15,18]. In Al_3Sc, the main fracture surface was observed to be {110} plane [51], which is not in accord with the G_c values listed in Tables 3 and 6. This indicates that the anisotropy in crack-tip plasticity by slip and twinning is the controlling factor, particularly in the so-called "soft and brittle" compounds. In the study of fractured surfaces based on the selected area electron channeling patterns, George et al. [52] concluded that the availability of several low-strength cleavage planes in an Al–23Ti–6Fe–5V polycrystalline alloy apparently exacerbates its brittleness. On the other hand, the estimate of $G_c = 3.7$ J/m^2 being the lowest in $TiAl_3$ (Table 7) is consistent with the transmission electron microscopy observation of {100} habit plane in Al–25Ti–8Cr single crystals [53].

Fracture toughness of $MoSi_2$ at room temperature has been estimated by examining the cracks around a hardness indenter [54,55]. The estimated data are 3 MPa-$\sqrt{m}$ in polycrystals [54] and 4 MPa-$\sqrt{m}$ for (001) < 100 > cracks in single crystals [55]. As was discussed by referring to Table 8, in intrinsically brittle materials such as $MoSi_2$ and WSi_2, anisotropic G_c values and fracture toughness data from single crystals can be used together to assess the relative role of crack-tip plasticity. It may be concluded from the results in Table 8 that in $MoSi_2$ the amount of crack-tip plasticity appears to be larger at a (001)[110] crack, $(K_{IC} - k_{IG})/k_{IG} = 1.1$, than at a (110)[001] crack, 0.3, by a factor of more than three. The transmission electron microscopy investigation by Ito et al. [39] revealed stacking faults of the Frank-type on (001) which are formed probably due to the loss of silicon during high-temperature exposure. These faults, which are more abundant in WSi_2 than in $MoSi_2$, may contribute also to the difference and the anisotropy of fracture toughness in these two TM disilicides.

As we consider tougher and stronger silicides of ultra-high melting temperatures, which have mostly noncubic crystal structures, the anisotropy of G_c remains a very important factor in describing their integrity after processing and handling.

7. Summary

A simple expression for the ideal work of adhesion, resulting from a rigid-body decohesion of a (hkl) plane, is proposed in terms of four variables, viz. the elastic stiffness constant, the interplanar spacing, and two scaling length parameters (D and a or l). By using the available results of ab initio slab-supercell calculations, the anisotropy of l^2/D is assessed. Within a chemical group, the ideal cleavage energies of TM aluminides and silicides were estimated by comparing pseudo-fcc and pseudo-bcc building blocks of cubic and tetragonal crystals.

Acknowledgements

The authors would like to thank C.L. Fu, J.A. Horton, and E.P. George for stimulating discussions and helpful comments on the manuscript. This research was sponsored by the Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, US Department of Energy, under Contract DE-AC05-00OR22725 with UT-Battelle, LLC, and also supported in part by the program of Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad.

References

[1] Bewlay BP et al., editors. High-temperature silicides and refractory alloys. Pittsburgh Symp. Proc., vol. 322, Pittsburgh (PA): MRS, 1994.
[2] Petrovic JJ et al., editors. High temperature structural silicides. Abstract Bulletin of Engineering Foundation Conference, 25–29 May 1998, Hyannis, MA.
[3] Yoo MH, Yoshimi K. In: George EP et al., editors. High-temperature ordered intermetallic alloys VIII. Symp. Proc., vol. 552. Pittsburgh (PA): MRS, 1999. p. KK6.6.1.
[4] Murr LE. Interfacial phenomena in metals and alloys. Reading: Addison-Wesley (MA), 1975.
[5] Grimvall G, Sjödin S. Physica Scripta 1974;10:340.
[6] Lang ND, Kohn W. Phys Rev, B1:4555
[7] Harris J, Jones RO. Phys Lett 1974;46A:407.
[8] Rose JH, Smith JR, Ferrante J. Phys Rev Lett 1981;47:675
[9] Smith JR, Rose JH, Ferrante J, Guinea F. In: Langreth D et al., editors. Many-body phenomena at surfaces. New York: Academic, 1984. p. 159–174.
[10] Methfessel M, Henning D, Scheffler M. Appl Phys 1992;A55:442.
[11] Mehl MJ, Papaconstantopoulos DA. Phys Rev 1996;B54:4519.
[12] Fu CL, Yoo MH. In: Liu CT et al., editors. Ordered intermetallics — physical metallurgy and mechanical behaviour. NATO ASI Series, vol. 213. Applied Sciences, 1992. p. 155.
[13] Yoo MH, Fu CL, Horton JA. Mater Sci Eng 1994;A176:431.
[14] Yoo MH, Fu CL. Mater Sci Eng 1992;A153:470.
[15] Yoo MH, Zou J, Fu, CL. Mater Sci Eng 1995;A192/193:14
[16] Stadler R, Podloucky R, Kresse G, Hafner J. Phys Rev 1998;B57:4088
[17] Fu CL. J Mater Res 1990;5:971.
[18] Yoo MH, Fu CL. Scr Metall Mater 1991;25:2345.
[19] Hong T, Smith JR, Srolovitz DJ, Gay JG, Richter R. Phys Rev B45:8775
[20] Banerjea A, Smith JR. Phys Rev B 1988;37:6632.
[21] Vinet P, Rose JH, Ferrante J, Smith JR. J Phys 1989;1:1941.
[22] Reynolds JE, Smith JR, Srolovitz DJ, Zhao G-L. In: Selinger RLB et al., editors. Fracture-instability dynamics, scaling, and ductile/brittle behavior. Symp. Proc., vol. 409. Pittsburgh (PA): MRS, 1996. p. 177–182.
[23] Gilman JJ. Fracture. Cambridge (MA): The MIT Press, 1959.
[24] Kelly A. Strong solids. Oxford: Clarendon Press, 1966.
[25] Macmillan NH, Kelly A. Mater Sci Eng 1972;10:139.
[26] Cherepanov GP. Mechanics of brittle fracture. New York: McGraw-Hill Int'l Book Co, 1979.
[27] Yoo MH. Scr Metall 1979;13:131.
[28] Simmons G, Wang H. Single crystals elastic constants and calculated aggregate properties: a handbook. 2nd ed. New York: John Wiley & Sons, 1971.
[29] Nakamura M. In: editors Intermetallic compounds, vol. 1, principles. Westbrook JH and Fleischer RL, New York: John Wiley & Sons, 1994. p. 873.
[30] Fu CL, Yoo MH. Acta Metall Mater 1992;40:703.
[31] Smith JR, Perry T, Banerjea A, Ferrante J, Bozzolo G. Phys Rev B 1991;44:6444.
[32] Daw MS. Phys Rev B 1989;39:7441.
[33] Finnis MW, Sinclair JE. Phil Mag A 1984;50:45.
[34] Cox H. Surface Sci 1998;397:374.
[35] Fu CL, Ye Y-Y, Yoo MH. Phil Mag Lett 1993;67:179
[36] Yoo MH, Fu CL. Metall Mater Trans A 1998;29:49.
[37] Fu CL, Yoo MH. Phil Mag Lett 1988;58:199.
[38] Mehl MJ, Klein BM, Papaconstantopoulos DA. In: Westbrook JH, Fleischer RL, editors. Intermetallic Compounds, vol. 1, Principles. John Wiley & Sons, 1994. p. 195.
[39] Ito K, Nakamoto T, Inui H, Yamaguchi M. In: Koch CC et al., editors. High-temperature ordered intermetallic alloys VII. Symp. Proc., vol. 460. Pittsburgh (PA): MRS, 1997. p. 599.
[40] Fu CL, Ohnishi S, Jansen HJF, Freeman AJ. Phys Rev B, 1985;31:1168.
[41] Wasilewski RJ, Trans. TMS-AIME, 1966;236:455.
[42] He Y, Schwarz RB, Darling T, Hundley M, Whang SH, Wang ZM. Mater Sci Eng 1997;A239–240:157.
[43] Rusovic, N. and Warlimont, H., Phys. Stat. Sol (a), 1977, 44, 609.
[44] Leamy HJ, Gibson ED, Kayser FX. Acta Met 1967;15:1827.
[45] Mathew JW, Phil Mag 1974;29:797.
[46] Yoo MH, Fu CL. In: Phillpot SR et al., editors. Microscopic simulation of interfacial phenomena in solids and liquids. Symp. Proc., vol. 492. Pittsburgh (PA): MRS, 1998. p. 345.
[47] Yoo MH, Fu CL. ISIJ International 1991;31:1049.
[48] Ludwig M, Gumbsch P. Acta Mater 1998;46:3135.
[49] Farkas D. In: Koch CC et al. High-temperature ordered intermetallic alloys VII. Symp. Proc., vol. 460. Pittsburgh (PA): MRS, 1997. p. 399.
[50] Foiles, Daw MS. J Mater Res 1967;2:5.
[51] Schneibel JH, Hazzledine PM. J Mater Res 1992;7:868.
[52] George EP, Porter WD, Joy DC. In: Liu CT et al., editors. High-temperature ordered intermetallic alloys III. Symp. Proc., vol. 133. Pittsburgh (PA): MRS, 1989. p. 311.
[53] Wu ZL, Pope DP. In: Darola R et al., editors. Structural intermetallics. Warrendale (PA): TMS, 1993. p. 107.
[54] Wade RK, Petrovic JJ. J Am Cer Soc 1992;75:1682.
[55] Peralta P, Maloy SA, Chu F, Petrovic JJ, Mitchell TE. Scr Mater 997;37:1599.

PERGAMON

Acta mater. 48 (2000) 2339–2347

Acta MATERIALIA

www.elsevier.com/locate/actamat

DEFECT STRUCTURES IN $ZrCo_2$ LAVES PHASE

J. H. ZHU† and C. T. LIU

Oak Ridge National Laboratory, Metals and Ceramics Division, Bldg. 4500S-MS6115, PO Box 2008, Oak Ridge, TN 37831-6115, USA

(Received 6 December 1999; accepted 27 December 1999)

Abstract—Point defect structures in the C15 $ZrCo_2$ alloys were studied by bulk density and X-ray lattice parameter measurements. It was found that, for the $ZrCo_2$ alloys quenched from 1000°C, the lattice parameter increases linearly as the Zr content increases up to 33.3 at.% Zr. The lattice parameter of the Laves phase remains constant for the alloys with Zr content higher than 33.3 at.%, indicating that the solubility range of Zr in $ZrCo_2$ on the Zr-rich side is essentially zero. The constitutional defects were found to be of the anti-site type. Thermal vacancies exhibiting a maximum at the stoichiometric composition were observed in the $ZrCo_2$ Laves phase alloys after quenching from 1250 and 1000°C, with higher thermal vacancies obtained from 1250°C. The defect structures in the $ZrCo_2$ phase may be correlated to the relative magnitude of formation enthalpies for anti-site and quadruple defects in this compound. Thermal vacancy concentration at a level of 1% in $ZrCo_2$ does not affect fracture toughness at room temperature. *Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.*

Keywords: Intermetallic compound; Casting; X-ray diffraction; Lattice defects; Mechanical properties (fracture & fracture toughness)

1. INTRODUCTION

Topologically close packed Laves phases with AB_2 composition are known to be size compounds, i.e., the atomic size ratio, R_A/R_B, is ideally 1.225, with a range of 1.05–1.67 typically observed [1, 2]. Laves phases are conventionally considered to be line compounds with a strict AB_2 composition. However, about 25% of the binary Laves phases show certain homogeneity ranges on the A- and/or B-rich side of stoichiometry [3]. Non-stoichiometric intermetallic compounds are stabilized by the incorporation of constitutional defects. The excess atoms in off-stoichiometric compounds may prefer to (1) stay on their own sublattice, leading to the formation of constitutional vacancies on the other sublattice, (2) insert into interstitial sites, or (3) occupy the sites on the other sublattice, leading to the formation of substitutional anti-site defects. Since Laves phases have topologically close packed (TCP) structures and space-filling is relatively high, there are no interstitial sites with a size comparable to that of the component atoms. Therefore, the insertion of the excess atoms into the interstitial sites can be excluded, and the possible defect mechanisms in binary Laves phases are thus constitutional vacancy or anti-site substitution.

The defect structures in cubic C15 $NbCr_2$, $NbCo_2$ and hexagonal C14 $NbFe_2$ Laves phases have recently been determined to be of anti-site type on both sides of stoichiometry [4]. Contrary to the geometric considerations which support the constitutional vacancy mechanism on the A-rich side of stoichiometry, anti-site substitution is found to be the constitutional defect mechanism on the Nb-rich side of these Laves phases. These results indicate that except for the geometric factor, other factors such as chemical bonding characteristics may also affect the defect structures in this Laves phases. Based on geometric and thermodynamic considerations, Laves phases with atomic size ratio R_A/R_B larger than 1.225 and/or high negative enthalpy of formation (ΔH) may exhibit constitutional vacancies in the A-rich side, as suggested by Zhu *et al.* [4].

By heating the ordered Laves phases, thermal defects can also be created in the compounds. Modder and his coworkers [5, 6] have systematically analyzed the possible defect type in Laves phases and two types of thermal defects were identified, namely quadruple defects (consisting of one anti-site B atom on the A-atom sublattice and three

† To whom all correspondence should be addressed.

1359-6454/00/$20.00 Published by Elsevier Science Ltd on behalf of Acta Metallurgica Inc.
PII: S1359-6454(00)00014-8

vacancies on the B-atom sublattice) and anti-site defects. Furthermore, it was found that the defect types in Laves phases can be correlated to the relative magnitudes of the formation enthalpies of anti-site and quadruple defects in the compounds [6].

The point defect structures in the $ZrCo_2$ system have not been investigated previously. The atomic size ratio R_A/R_B of $ZrCo_2$ is 1.28, which is significantly larger than the ideal one, i.e., 1.225. Furthermore, the enthalpy of formation (ΔH) of $ZrCo_2$ is -113 kJ/mole [7], which is much more negative than those of $NbCr_2$, $NbCo_2$ and $NbFe_2$ Laves phases. So far, no work has been conducted to carefully determine the homogeneity range and defect structures in the Laves phase. The binary Co–Zr phase diagram shows certain solubilities on both sides of stoichiometry in this Laves phase system [8]. All these facts indicate that the $ZrCo_2$ Laves phase system deserves further attention as a system where constitutional vacancies may be the defect mechanism on the Zr-rich side.

In the present study, we have attempted to determine the solubility range and to model the defect structures in binary $ZrCo_2$ alloys with different stoichiometric compositions, by measuring the lattice parameters, bulk densities, and therefore, vacancy concentrations of binary $ZrCo_2$ alloys of various compositions after quenching from 1000 and 1250°C. Thermal defects that will be introduced by quenching from elevated temperatures have also been determined. The fracture toughness of selected alloys were also measured at room temperature.

2. EXPERIMENTAL PROCEDURES

The alloy compositions investigated in this study are listed in Table 1. A total of seven binary alloys were prepared. The Zr content in the alloy increases with the alloy code number. The alloys ♯1–♯3 are Co-rich and ♯4 is the stoichiometric alloy, while alloys ♯5–♯7 are Zr-rich. Ingots weighing approximately 50 g were prepared by arc-melting techniques using high-purity starting materials. Each ingot was remelted at least five times to ensure compositional homogeneity before being cast into a 12.7 mm diameter cylinder. The weight loss during arc-melting was essentially zero (less than 0.01 g). It should be noted that the ingots of the alloys ♯4–♯7 remained in one piece, while the ingots of the alloys ♯1–♯3 broke into several pieces after drop casting.

After a homogenization anneal at 1200°C for 20 h in vacuum, the alloys were encapsulated in quartz, which was then evacuated and backfilled with helium. The encapsulated alloys were then heat-treated at 1000°C for 7 days and water quenched. Small fractions of the quenched samples were mounted for microstructural observation using an optical microscope and for chemical composition analyses of the phases present using a field-emission-gun scanning electron microscope (SEM) equipped with an energy dispersive X-ray spectrometer (EDS). All the alloys were found to be fully homogenized following the above heat treatments.

The remaining parts of the samples were crushed in an agate mortar into powders with particle sizes smaller than 40 mesh. Powders with a mesh size of $-40+80$ were used for bulk density measurements, and powders smaller than 320 mesh were used for X-ray analyses. The bulk density was measured using a helium pycnometer with an accuracy of approximately 0.01 vol.%. Powdered samples were used to eliminate the micropores generated during solidification, which leads to more accurate density measurements. The vacancy concentration of the Laves-phase alloys, c_V, is defined as the ratio of the total number of vacancies to the total number of atoms. It can be obtained from the following equation:

$$c_V = \frac{\rho_X - \rho_B}{\rho_B} \tag{1}$$

where ρ_X and ρ_B are the X-ray and the bulk densities, respectively.

The powdered samples of the single-phase alloys ♯1–♯4 were then wrapped in Ta sheets and encapsulated in quartz, which was evacuated and backfilled with helium. They were then heat-treated at 1250°C for 2 h and water quenched. Bulk densities and lattice parameters were measured again and the

Table 1. Composition, lattice parameter, X-ray density, bulk density and vacancy concentration of $ZrCo_2$ alloys quenched from 1250 and 1000°C

Alloy code	Zr (at.%)	Lattice parameter (Å)		X-ray density (g/cm^3)		Bulk density (g/cm^3)		Vacancy concentration (%)	
		1250°C	1000°C	1250°C	1000°C	1250°C	1000°C	1250°C	1000°C
#1	30.8	6.9305	6.9311	8.242	8.240	8.194	8.224	0.58	0.20
#2	31.8	6.9419	6.9430	8.240	8.235	8.192	8.210	0.59	0.31
#3	32.8	6.9543	6.9546	8.234	8.233	8.177	8.203	0.70	0.37
#4	33.3	6.9590	6.9608	8.236	8.230	8.155	8.196	1.00	0.42
#5	33.5	–	6.9609[a]	–	–	–	–	–	–
#6	33.8	–	6.9607[a]	–	–	–	–	–	–
#7	34.3	–	6.9609[a]	–	–	–	–	–	–

[a] Two-phase alloys with the lattice parameter of the $ZrCo_2$ Laves phase listed here.

vacancy concentrations were obtained for these alloys quenched from 1250°C.

X-ray diffraction measurements were made using a Scintag XD2000 diffractometer with Cu K_α radiation. The lattice parameter was determined from 13–15 peaks. The X-ray density was calculated using the measured lattice parameter. The accuracy of the vacancy concentration using our method is approximately 0.08%, which is mainly due to the compositional uncertainty of the alloys.

The fracture toughness of the selected alloys was also determined using the indentation method with a load of 500 g and a holding time of 15 s. The crack lengths were measured immediately after the indentation to eliminate the possibility of slow crack propagation following the removal of the indenter. The following equation was used to calculate the fracture toughness of the alloys [9]:

$$K_{IC} = A\left[\frac{E}{H}\right]^n \frac{P}{L^{3/2}} \quad (2)$$

where K_{IC} is the fracture toughness (MPa$\sqrt{m}$), E is Young's modulus (GPa), which is assumed to be 218 GPa for all of the investigated alloys, H is Vickers hardness (GPa), P is the load (N), and L the average length of the four radial cracks from the center of the indent to the crack tip. A and n are constants, which are taken as 0.016 and 0.5, respectively, for relatively brittle materials [9].

3. RESULTS

3.1. Microstructure

Microstructure examinations of the $ZrCo_2$ Laves phase alloys listed in Table 1 after quenching from 1000°C indicate that alloys ♯1–♯4 are in the single-phase region, while alloys ♯5–♯7 are in the two-phase region. Typical microstructures of single-phase and two-phase alloys are shown in Figs 1(a) and (b). Note that Fig. 1(a) was taken from the Co–33.3%Zr sample (alloy ♯4) etched with a solution of 60% glycerine + 20% HNO_3 + 10% HF + 10% H_2O (in vol.%), while Fig. 1(b) was from the Co–34.3%Zr sample (alloy ♯7) in the as-polished condition under polarized light. There are no annealing twins in the alloys, unlike the $NbCr_2$ system [4], indicating relatively high stacking fault energy in this compound. The second phase in alloys ♯5–♯7 was analyzed with EDS in an SEM, and was found to be richer in Zr than the $ZrCo_2$ matrix. According to the binary Co–Zr phase diagram [8], this second phase is a ζ phase with a B2 structure. The maximum content of Zr in $ZrCo_2$ at 1000°C is found to be about 33.3 at.% from this study, which is different from the estimated phase boundary in the binary Co–Zr phase diagram [8, 10], where it was indicated to be about 34.5 at.%. Furthermore, extensive cracking was observed in the matrix around the second phase. These cracks are most likely caused by the thermal residual stress in the alloys after casting and heat treatment, as a result of the mismatch in thermal expansion coefficients between the second phase and the matrix.

3.2. Lattice parameter, density, and vacancy concentration

The lattice parameters of the C15 $ZrCo_2$ alloys listed in Table 1 are plotted in Fig. 2 as a function of stoichiometry for all the samples quenched from 1000°C and for the single-phase alloys quenched from 1250°C. As the Zr content increases, the lattice parameter of the $ZrCo_2$ phase also increases. For the Co-rich alloys quenched from 1250 and 1000°C, the data points fit nicely along two straight lines, while on the Zr-rich side the lattice constants are identical for alloys ♯5–♯7 quenched from

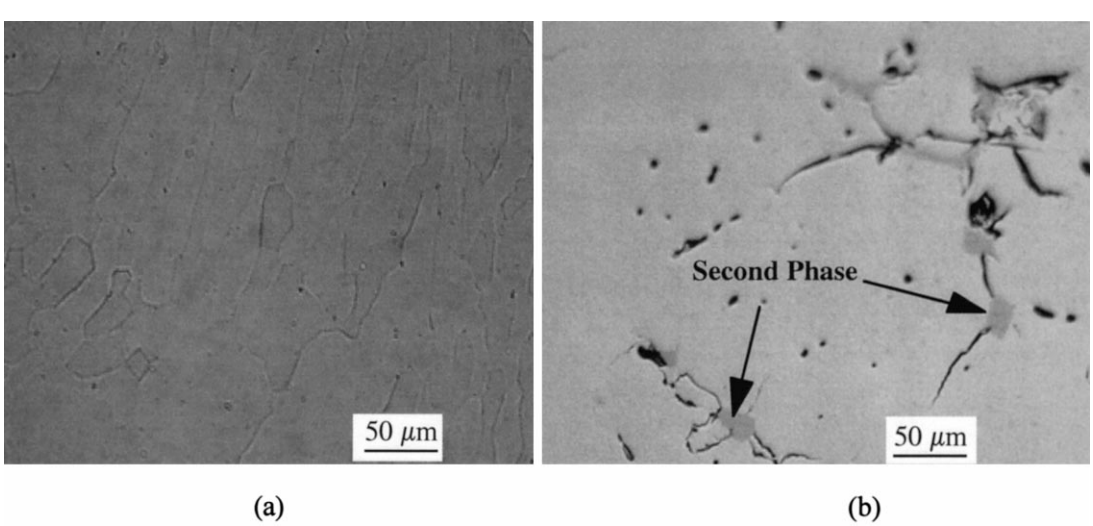


Fig. 1. Optical micrographs of $ZrCo_2$ alloys: (a) alloy ♯4, etched; (b) alloy ♯7, as-polished.

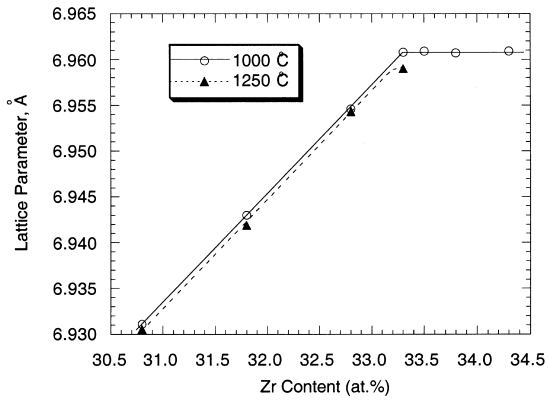


Fig. 2. Lattice parameter as a function of Zr content after quenching from 1250 and 1000°C.

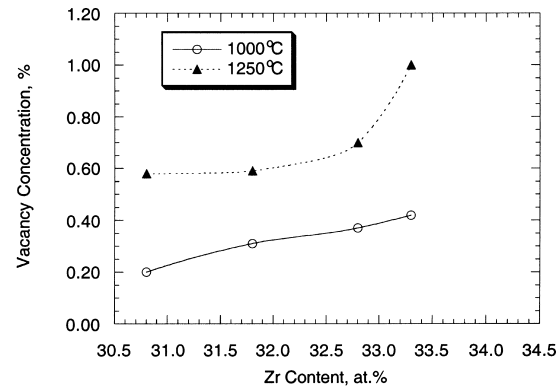


Fig. 4. Vacancy concentration as a function of Zr content after quenching from 1250 and 1000°C.

1000°C, indicating that the composition of the Laves phase is the same in the alloys with more than 33.3% Zr (see Fig. 2). This is expected, because the maximum content of Zr in $ZrCo_2$ is about 33.3 at.% and these three alloys were found to be in the two-phase region. The lattice parameters after quenching from 1250°C are a little bit smaller, compared to those after quenching from 1000°C. This can be due to the presence of vacancies in the alloy.

The X-ray and bulk density results for alloys ♯1–♯4 after quenching from both 1250 and 1000°C are presented in Table 1 and Fig. 3. The density of the alloys decreases as the Zr content increases, and the bulk density after quenching from 1000°C is consistently higher than those after quenching from 1250°C. Using equation (1) and the data presented in Table 1 and Fig. 3, the vacancy concentrations were calculated, and are listed in Table 1. Due to the presence of the second phase, the vacancy concentration could not be determined for alloys ♯5–♯7. Figure 4 is a plot of the vacancy concentration as a function of the Zr content for the single-phase $ZrCo_2$ alloys after quenching from the two different temperatures. It is noted that the vacancy concentration is quite high after quenching from 1250°C. The vacancy concentration reaches a maximum of ~1.0% at the stoichiometric composition and decreases on the Co-rich side of stoichiometry. Annealing and quenching from 1000°C results in vacancy concentrations lower than those after quenching from 1250°C. However, a maximum at the stoichiometric composition can also be recognized.

3.3. Fracture toughness

The fracture toughness values of the single-phase $ZrCo_2$ alloys after quenching from 1000°C were measured as a function of Zr content (see Fig. 5). It was found that all the single-phase $ZrCo_2$ alloys exhibit a fracture toughness value of around 1 MPa

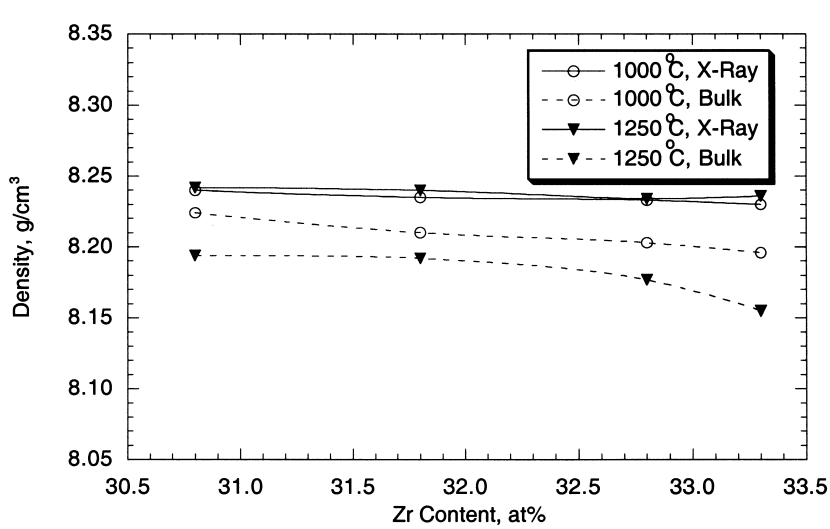


Fig. 3. X-ray and bulk densities as a function of Zr content after quenching from 1250 and 1000°C.

$m^{1/2}$, regardless of the alloy stoichiometry. Furthermore, the fracture toughness value of the stoichiometric $ZrCo_2$ alloy after quenching from 1250°C was found to be close to that of the same alloy after quenching from 1000°C, even though the vacancy concentration in the alloy is much higher after quenching from 1250°C. This indicates that the presence of such vacancies does not lead to toughness improvement in this Laves phase.

4. DISCUSSION

4.1. Constitutional defect mechanisms in $ZrCo_2$ Laves phase

The vacancy concentration as a function of stoichiometry for binary Laves phases has been calculated based on the two constitutional defect models, i.e., the constitutional vacancy model and anti-site substitution model [4]. The constitutional vacancy model is based on the assumption that all the excess atoms exclusively stay on their own sublattice, thus constitutional vacancies are created on the sublattice of the other element. Based on this model, the constitutional vacancy concentrations in Laves phases are zero at the stoichiometric composition and increase on both sides of stoichiometry. The anti-site substitutional model assumes that the excess atoms occupy the sublattice sites of the other species. Thus, anti-site defects are created, and no constitutional vacancies are needed for the balance of the lattice sites. The constitutional vacancy concentrations are zero for both A- and B-rich Laves phases, according to the anti-site substitutional model.

From Fig. 4, it is obvious that the measured vacancy concentrations for all the single-phase $ZrCo_2$ alloys after quenching from 1000°C are relatively low (see also Table 1), indicating that there are no constitutional vacancies in these Laves phases. Furthermore, as the composition deviates further from the stoichiometry, the vacancy concentration on the Co-rich side decreases, while the constitutional vacancy model predicts an opposite trend. Actually, the measured vacancy concentration as a function of stoichiometry can be rationalized by assuming that these vacancies are thermal vacancies instead of constitutional vacancies (see next section for more details). This is consistent with the anti-site substitution model and contrary to the constitutional vacancy model, which demands much higher vacancy concentrations for off-stoichiometric compositions. Thus, the constitutional defects are of anti-site type for the Co-rich $ZrCo_2$ Laves phases.

Point defect mechanisms in binary Laves phases have been the focus of several studies. Fleischer [11] found that anti-site substitution occurs on both sides of stoichiometry for the C15 $ZrCr_2$ Laves phase. Pargeter and Hume-Rothery [12] and Saito and Beck [13] investigated the point defects in Co-rich $NbCo_2$ and dihexagonal C36 $NbCo_3$ Laves phases and found that in both phases, the deviation from stoichiometry resulted from the substitution of Co atoms for the Nb atoms. X-ray and neutron diffraction measurements [14] showed that for both $ZrFe_2$ and $TiFe_2$, the excess Fe atoms are on the Zr or Ti sublattice. The constitutional defect structure of the $NbFe_2$ Laves phase was also investigated by measuring the lattice parameters on both sides of stoichiometry, and again anti-site substitution was found to be the defect mechanism [15]. YAl_2 Laves phase is a line compound with negligible solubility existing under equilibrium conditions; however, Foley *et al.* extended the solubility on both sides of stoichiometry in this alloy system via rapid solidification. From lattice parameter measurements, off-stoichiometry in the YAl_2 Laves phase was postulated to be accompanied by different defect mechan-

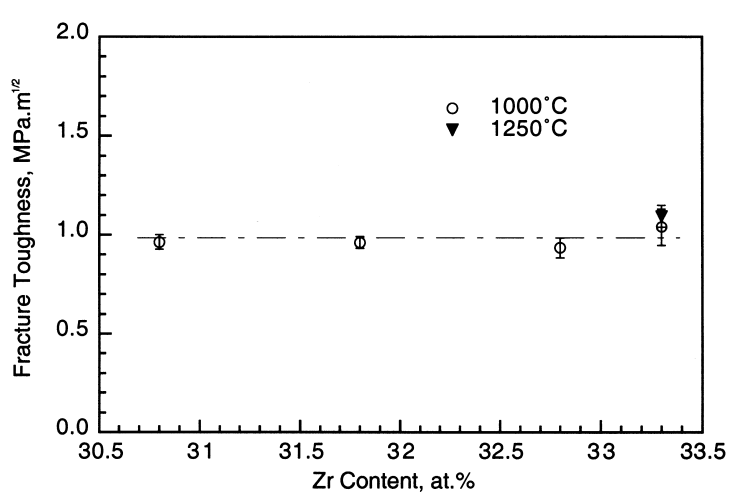


Fig. 5. Fracture toughness vs Zr content for $ZrCo_2$ alloys after quenching from 1250 and 1000°C.

isms: the Y-rich side is accommodated by vacancies, while the Al-rich compositions result from anti-site substitution [16]. Recent studies combining bulk density and X-ray lattice parameter measurements unambiguously indicate that the constitutional defect mechanisms in binary C15 $NbCr_2$, $NbCo_2$, and C14 $NbFe_2$ compounds on both sides of stoichiometry are of anti-site type, with no constitutional vacancies present in the compounds [4].

Table 2 summarizes the atomic size ratio R_A/R_B, enthalpy of formation and point defect structures for a number of binary Laves phases [17]. Based on geometric considerations, the atomic size ratio R_A/R_B of the AB_2 Laves phases should be important in controlling the point defect structures in Laves phases. For Laves phases with R_A/R_B greater than the ideal ratio of 1.225, on the A-rich side, it will be more difficult for A atoms to stay on the sublattice sites of B atoms as the atomic size ratio increases. Therefore, it would be possible to obtain constitutional vacancies for Laves phases with $R_A/R_B > 1.225$ on the A-rich side, even though substitutional anti-site defects are most likely to form under all the other circumstances. In all the AB_2 Laves phases whose point defect structure on the A-rich side has been studied so far and listed in Table 2, only YAl_2 has an R_A/R_B ratio of 1.258, higher than the ideal one, and this alloy system also shows certain indications of constitutional vacancies in the Y-rich compounds [16]. For the $ZrFe_2$ alloys with the R_A/R_B ratio of 1.257, the defect structure on the Zr-rich side has not been studied yet. For the $ZrCo_2$ compound, R_A/R_B is 1.280, much larger than the ideal. Therefore, the possibility is expected to be higher for constitutional vacancies existing in this system.

On the other hand, Neumann [18] found that for B2 compounds their constitutional defect structure is closely related to the magnitude of enthalpy of formation. Constitutional vacancies are the preferred defect mechanism for B2 compounds that are more strongly ordered with more negative enthalpy of formation (ΔH) values on the large atom-rich side of stoichiometry, while B2 compounds with ΔH higher than −75 to −90 kJ/mole generally exhibit the anti-site defect structure. Similar thermodynamic arguments may hold for binary Laves phases. As listed in Table 2, anti-site defects are observed on the large atom-rich side for those Laves phases with ΔH value higher than −75 to −90 kJ/mole such as $NbCr_2$, $NbCo_2$, and $NbFe_2$. For Laves phases with more negative ΔH values, such as YAl_2 with a ΔH value of −160.4 kJ/mole [17], constitutional vacancies are the possible defect mechanism on the Y-rich side. In this regard, $ZrCo_2$ also has a highly negative ΔH value of −123 kJ/mole, and thus constitutional vacancies may form on the Zr-rich side.

Based on both geometric and thermodynamic considerations, therefore, constitutional vacancies may be created on the Zr-rich side of $ZrCo_2$, similar to the case of YAl_2. Unfortunately, this study indicates that the solubility range of Zr is essentially zero in $ZrCo_2$ on the Zr-rich side, which makes it difficult to verify our hypothesis regarding the effects of atomic size ratio and enthalpy of formation on the defect mechanisms in binary Laves phases. Rapid solidification techniques could be utilized in future investigations to extend the solubility range on the Zr-rich side in this compound and defect structures can then be deduced from the lattice parameter measurement, as in the case of the YAl_2 phase [16]. It would also be interesting to study defect structures in other Laves phases with $R_A/R_B > 1.225$ and/or highly negative enthalpy of formation to consider whether it is possible to introduce constitutional vacancies on the A-rich side in these compounds.

The fact that both YAl_2 and $ZrCo_2$ compounds are unstable on the A-rich side of stoichiometry under equilibrium conditions indicates that the creation of constitutional vacancies may make the compound energetically less stable. This may explain the lack of solubility on the A-rich side of these compounds. Furthermore, anti-site defects may also be impossible to form on the A-rich side in these compounds, due to the difficulties involved in A-atoms occupying the much smaller B-atom sublattice sites. Consequently, it may be impossible or at least very difficult to obtain constitutional defects in these Laves phases.

4.2. Thermal defects in $ZrCo_2$ Laves phase

As Modder and Bakker have pointed out, it is very possible that a vacancy-type defect termed

Table 2. Atomic size ratio, enthalpy of formation and point defect structure of a number of binary Laves phases

Compound	R_A/R_B[a]	ΔH (kJ/mole)[a]	Defect structure
$NbCr_2$	1.145	−21	Anti-site substitution on both sides
$NbCo_2$	1.173	−51	Anti-site substitution on both sides
$NbFe_2$	1.152	−62	Anti-site substitution on both sides
$ZrFe_2$	1.257	−75	Fe-rich side: anti-site substitution
$TiFe_2$	1.148	−84	Fe-rich side: anti-site substitution
YAl_2	1.258	−160.4	Al-rich side: anti-site substitution Y-rich side: constitutional vacancy
$ZrCo_2$	1.280	−123	Co-rich side: anti-site substitution Zr-rich side: unknown

[a] The R_A/R_B and ΔH values for these compounds are from Ref. [17].

quadruple defect occurs in C15 compounds [5]. Furthermore, ball milling was found to introduce quadruple defects in $GdAl_2$, $GdPt_2$, $GdIr_2$ and $GdRh_2$, while anti-site defects were created in $GdMg_2$ [6]. Since ball milling usually leads to the formation of the same type of atomic defects as heating, thermal defects such as quadruple defects or anti-site defects may be created upon quenching the Laves phases from elevated temperatures. Since each quadruple defect contains three vacancies, thermal vacancies are expected if quadruple defects are created.

The vacancy concentrations in $ZrCo_2$ alloys are quite high after quenching from 1250°C and the vacancy concentrations are significantly reduced after annealing at lower temperature, i.e., 1000°C (Fig. 4). Since the constitutional defects in this compound have been established to be of the anti-site type, these vacancies are obviously thermal vacancies. The fact that the vacancy concentration increases as the quenching temperature increases is a further indication of the thermal nature of these vacancies. Thus, the quadruple defects may be formed in this compound by heating. According to the quadruple defect model and considering the difficulties involved in Zr sitting on the small atom Co sites, Co will go to the Zr sublattice and thus the thermal vacancies are on the Co sublattice sites. Furthermore, the measured vacancy concentration reaches a peak at the stoichiometric composition and decreases on deviating from stoichiometry for the $ZrCo_2$ alloys when quenched from 1250°C and 1000°C (Fig. 4). This behavior is similar to the $NbCr_2$ system [9], where the thermal vacancy concentration exhibits a maximum at the stoichiometric composition. As in the $NbCr_2$ phase, a maximum in thermal vacancy concentration at the stoichiometric composition is rationalized by considering that the increase in entropy associated with the introduction of vacancies will be the greatest at the stoichiometric composition, where the degree of order is necessarily the highest.

Modder *et al.* [6] have extended Miedema's model to estimate the formation enthalpy of quadruple and anti-site defects in C15 Laves phases. According to this model, the type of atomic disorder is related to the relative magnitudes of formation enthalpies of anti-site and quadruple-defect disorder for the Laves phases. Therefore, the presence of thermal vacancies in $ZrCo_2$ may imply that the formation enthalpies of anti-site disorder is larger in magnitude than that of quadruple defect disorder for $ZrCo_2$. Furthermore, the higher vacancy concentration of $ZrCo_2$ after quenching, compared to $NbCr_2$, suggests that the formation enthalpy of quadruple defects in $ZrCo_2$ should be lower than that in $NbCr_2$, which leads to a higher amount of thermal defects in $ZrCo_2$ phase.

It has been suggested that vacancies in Laves phases may assist the synchroshear deformation mechanism and thus increase the toughness of the Laves alloys [19]. Even though thermal vacancies are relatively high (~1%) in the $ZrCo_2$ alloy after quenching from 1250°C, it was found that such quenched-in thermal vacancies do not affect the fracture toughness of the alloy. It will be of interest to determine the effect of constitutional vacancies on crack propagation and fracture toughness behavior of Laves phases if certain Laves phases with constitutional vacancies can be identified.

4.3. Reconstruction of the binary Co–Zr phase diagram

The existing Co–Zr phase diagram determined around 1964 by Pechin *et al.* [10] has not been updated so far. Those authors found that a homogeneity range of 28.4–34.6 at.% Zr exists for the $ZrCo_2$ Laves phase after rapidly quenching from 1400°C. In other words, there is 1.3 at.% Zr solubility on the Zr-rich side. Based on the present study, the solubility of Zr on the Zr-rich side is zero at 1000°C. This discrepancy may possibly be a result of the high impurity level in the materials used in the study by Pechin *et al.*, since they used the powder metallurgy (PM) method to produce the Laves phases, while in our study the arc-melting technique was employed. It is well known that impurities such as oxygen can easily be picked up during PM processing. The reaction of oxygen with Zr and the formation of ZrO_2 lower the Zr concentration in the Laves phase. Indeed, many voids and impurity particles (possibly ZrO_2) could be observed in the micrograph of the PM Laves phases [10]. Another possibility for the discrepancy may result from the fact that the quenching temperature is different for the two studies, with a much higher quenching temperature (1400°C) used in Pechin's study compared to 1000°C in the present study. The higher equilibrium temperature may increase the solubility range on the Zr-rich side of the compound.

In our view, this discrepancy is most likely due to the inaccurate composition control in Pechin *et al.*'s study, as is clear from the following argument. The slope of the lattice parameter as a function of stoichiometry for the Laves phase should usually change at the stoichiometric composition, as has been commonly observed for other Laves phases (for example, see Ref. [4]). According to Pechin *et al.* [10], the slope of the lattice parameter is the same on both sides of stoichiometry for $ZrCo_2$, suggesting that all the alloy compositions they prepared could be on one (Co-rich) side of stoichiometry, as a result of problems in material processing such as impurity pick-up. As has been demonstrated in the present study for the alloys quenched at 1000°C, negligible Zr solubility is observed for this Laves phase. Based on these arguments, we have refined the binary Co–Zr phase dia-

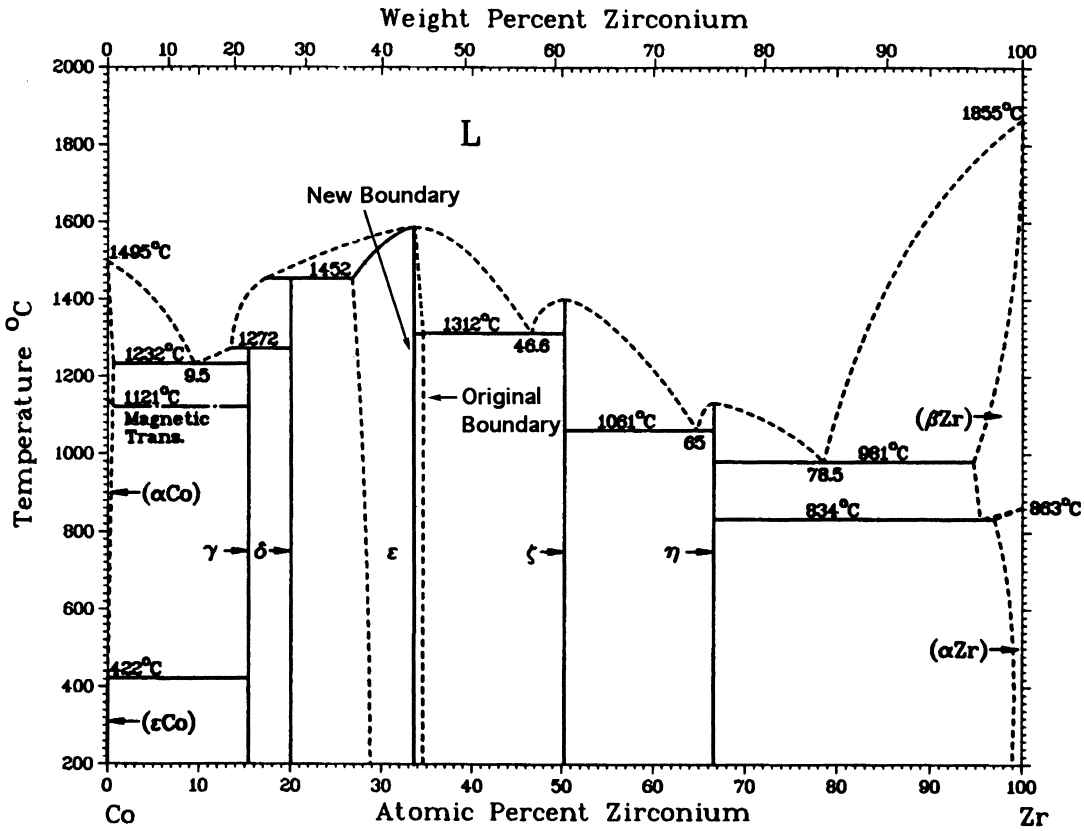


Fig. 6. The modified binary Co–Zr phase diagram.

gram, as shown in Fig. 6, where the solubility range of the Laves phase $ZrCo_2$ has been modified according to this study.

5. CONCLUSIONS

Point defects in the binary $ZrCo_2$ alloys have been clarified, combining both bulk density and X-ray lattice parameter measurements. The constitutional defects on the Co-rich side of stoichiometry for this system are of the anti-site type. As the Zr content in the Laves phase increases, the lattice parameters on the Co-rich side increase. The maximum content of Zr in the $ZrCo_2$ compound is the stoichiometric composition, 33.3 at.%. The binary Co–Zr phase diagram has been modified accordingly.

Thermal vacancies exhibiting a maximum at the stoichiometric composition were noted in $ZrCo_2$ Laves phase alloys after quenching from both 1000 and 1250°C, indicating that quadruple defects may be the type of thermal disorder in this compound. These thermal vacancies do not affect the fracture toughness of the Laves phase.

Acknowledgements—This research was sponsored by the Division of Materials Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. This research was supported in part by an appointment of J. H. Zhu to the Oak Ridge National Laboratory Postdoctoral Research Associates Program administered jointly by the Oak Ridge Institute for Science and Education and Oak Ridge National Laboratory (ORNL). Thanks are also due to S. R. Agnew and C. G. McKamey of ORNL for reviewing this manuscript.

REFERENCES

1. Laves, F., in *Theory of Alloy Phases*. American Society for Metals, Metals Park, OH, 1956, p. 124.
2. Zhu, J. H., Liu, C. T., Pike, L. M. and Liaw, L. M., *Metall. Mater. Trans.*, 1999, **30A**, 1449.
3. Thoma, D. J. and Perepezko, J. H., *J. Alloys Compounds*, 1995, **224**, 330.
4. Zhu, J. H., Pike, L. M., Liu, C. T. and Liaw, L. M., *Acta mater.*, 1999, **47**, 2003.
5. Modder, I. W. and Bakker, H., *Phys. Stat. Sol. (b)*, 1997, **199**, 369.
6. Modder, I. W., Kuin, M. J. and Bakker, H., *Intermetallics*, 1998, **6**, 537.
7. Klein, R., O'Hare, P. A. G. and Jacob, I., *J. Alloys Compounds*, 1997, **261**, 204.
8. Massalski, T. B., Murray, J. L., Bennett, L. H. and Baker, H., *Binary Alloy Phase Diagrams*. American Society for Metals, Metals Park, OH, 1986.
9. Sakai, M. and Bradt, R. C., *Int. Mater. Rev.*, 1993, **38**(2), 53.

10. Pechin, W. H., Williams, D. E. and Larsen, W. L., *Trans. ASM*, 1964, **57**, 464.
11. Fleischer, R. L., *Scripta metall. mater.*, 1992, **27**, 799.
12. Pargeter, J. K. and Hume-Ruthery, W., *J. Less-Common Metals*, 1967, **12**, 366.
13. Saito, S. and Beck, P. A., *Trans. Metall. Soc. AIME*, 1960, **218**, 670.
14. Brückner, W., Kleinstück, K. and Schulze, G. E. R., *Phys. Stat. Sol.*, 1967, **23**, 475.
15. Smith, A. W., Rogers, J. A. and Rawlings, R. D., *Phys. Stat. Sol. (a)*, 1973, **15**, K119.
16. Foley, J. C., Thoma, D. J. and Perepezko, J. H., *Metall. Trans.*, 1994, **25A**, 230.
17. Zhu, J. H., Metallurgical factors controlling the phase stability and mechanical behavior of transition-metal Laves phase alloy system. Ph.D. Dissertation, University of Tennessee, Knoxville, TN, August, 1998.
18. Neumann, J. P., *Acta metall.*, 1980, **28**, 1165.
19. Hazzledine, P. M. and Pirouz, P., *Scripta metall. mater.*, 1993, **28**, 1277.

ERKCM07 (Microstructural Design of Advanced Ceramics)

Refereed Journal Publications

He, M. Y., C. H. Hsueh, and P. F. Becher, "Deflection Versus Penetration of a Wedge-Loaded Crack: Effects of Branch-Crack Length and Penetrated-Layer Width," *Composites: Part B* **31**(4), 299-308 (2000).

(A) Hsueh, C. H. and E. R. Fuller, Jr., "Analytical Modeling of Oxide Thickness Effects on Residual Stresses in Thermal Barrier Coatings," *Scripta Mater*. **42**(8), 781-87 (2000).

(A) Hsueh, C. H. and E. R. Fuller, Jr., "Residual Stresses in Thermal Barrier Coating: Effects of Interface Asperity Curvature/Height and Oxide Thickness," *Mater. Sci. Eng*. **A283**(1-2), 46-55 (2000).

Hsueh, C. H., "Analyses of Edge Effects on Residual Stresses in Film Strip/Substrate Systems," *J. Appl. Phys*. **88**(5), 3022-28 (2000).

Hsueh, C. H., "Young's Modulus of Unidirectional Discontinuous Fibre Composites," *Compos. Sci. Technol*. **60**(14), 2671-80 (2000).

†(A) Ferber, M. K., A. A. Wereszczak, M. J. Lance, J. A. Haynes, and M. A. Antelo, "Application of Infrared Imaging to the Study of Controlled Failure of Thermal Barrier Coatings," *J. Mater. Sci.* **35**(11), 2643-51 (2000).

†(A) Haynes, J. A., M. J. Lance, K. M. Cooley, M. K. Ferber, R. A. Lowden, and D. P. Stinton,"CVD Mullite Coatings in High-Temperature High-Pressure Air-H_2O," *J. Am. Ceram. Soc.* **83**(3), 657-59 (2000).

†(A) Lance, M. J., J. A. Haynes, M. K. Ferber, and W. R. Cannon, "Monoclinic Zirconia Distributions in Plasma Sprayed Thermal Barrier Coatings," *J. Therm. Spray Technol.* **9(**1), 68-72 (2000).

(A)*Wang, X. L., C. M. Hoffmann, C. H. Hsueh, G. Sarma, C. R. Hubbard, and J. R. Keiser, "Influence of Residual Stress on Thermal Expansion Behavior," *Appl. Phys. Lett*. **75**(21), 3294-96 (1999).

*Indicates publications involving two or more BES FWP's,
(A) indicates partial support of ORNL research by programs other than BES
† indicates publication is not available

ELSEVIER

Composites: Part B 31 (2000) 299–308

composites
Part B: engineering

www.elsevier.com/locate/compositesb

Deflection versus penetration of a wedge-loaded crack: effects of branch-crack length and penetrated-layer width ☆

M.Y. He[a], C.H. Hsueh[b,*], P.F. Becher[b]

[a]*Materials Department, University of California, Santa Barbara, CA 93106-5050, USA*
[b]*Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6068, USA*

Received 9 September 1999; accepted 24 December 1999

Abstract

A crack intersecting an interface between two dissimilar material layers may advance by either deflecting along or penetrating through the interface. The criterion of deflection versus penetration can be established by the comparison of two ratios, the energy release rate ratio and the fracture energy ratio, of deflection to penetration. The effects of: (1) a finite length of the branch crack emanating from the main crack tip; and (2) a finite width of the layer subjected to crack penetration were examined in the present study. The results reveal that the above two factors have profound effects on the criterion of deflection versus penetration for a wedge-loaded crack. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Layered structures; B. Debonding; C. Finite element analysis (FEA)

1. Introduction

In an earlier paper by He and Hutchinson [1], a study was made on the tendency of a wedge-loaded crack meeting a bimaterial interface to either deflect along the interface or penetrate through the interface into the next layer. The analysis was conducted in terms of the energy release rate ratio of crack deflecting into the interface, G_d, to crack penetrating through the interface, G_p. The criterion of deflection versus penetration was then established depending on whether G_d/G_p was greater or less than the ratio of the fracture energy of the interface Γ_i to that of the adjoining layer, Γ_f. This criterion has been adopted extensively to predict interfacial debonding versus reinforcement fracture for fiber- (whisker- and self-) reinforced ceramic composites. However, due to the assumptions made in the asymptotic analysis, in which integral equation methods were used, two limitations should be noted before the criterion is applied. First, the results were obtained based on the condition that the branch crack emanating from the main crack tip was very small compared with all other lengths in the problem; including the length of the main crack itself. Second, the analysis was for two semi-infinite elastic materials bonded at the interface. For the crack initiation problem, the branch crack can be treated as an infinitely small crack; therefore, these two assumptions are satisfied. However, in some cases (e.g. for a wedge-loaded crack, or in the presence of residual stresses), the length of the branch crack becomes an important parameter, and these two conditions are not satisfied. The effects of the finite branch-crack length may well be significant.

Recently, the bonding strength at the whisker/glass interface in a model composite of oxynitride glass matrix containing 5 vol% β-Si_3N_4 whiskers has been evaluated by an indentation-induced crack-deflection method [2,3]. A cube-corner indenter was used to generate a crack in the glass. When the crack intersects the whisker, it will either deflect at the interface propagating a finite length and then kinking into the whisker or penetrate the whisker. In this case, the branch crack has a finite length and the whisker subjected to crack penetration has a finite width.

As an initial attempt to analyze the indentation-induced crack-deflection problem in the whisker/glass system and as a complement to the previous study [1], the present study sought to examine effects of: (1) a finite length of the branch crack emanating from the main crack tip; and (2) a finite

☆ This manuscript has been authorized by a contractor of the US Government, under contract no. DE-ACO596OR22464. Accordingly, the US Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

* Corresponding author. Tel.: + 1-423-576-6586; fax: + 1-423-574-8445.

E-mail address: hsuehc@ornl.gov (C.H. Hsueh).

1359-8368/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S1359-8368(00)00007-X

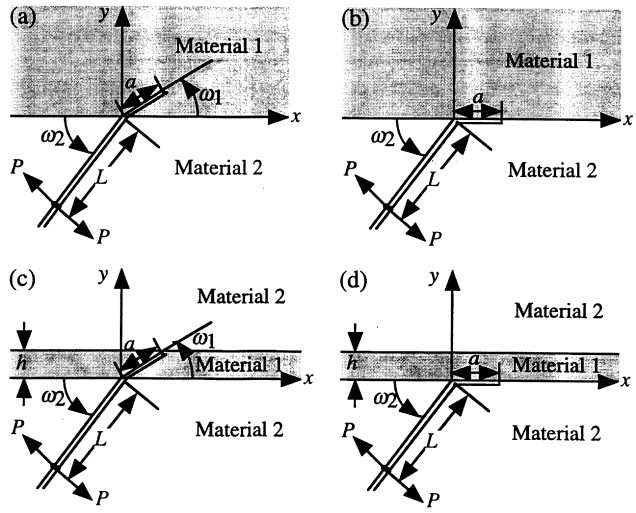


Fig. 1. Schematic drawings showing: (a) a penetrating crack; and (b) a deflected crack for two semi-infinite dissimilar materials bonded at the interface; and (c) a penetrating crack; and (d) a deflected crack for material 1 with a width, h, sandwiched between two semi-infinite material 2.

width of the adjoining layer subjected to crack penetration on the criterion of crack deflection versus penetration. To achieve this, integral equation methods were not feasible because of the complication of the problem and the finite element method was hence adopted. First, to verify the accuracy of the finite element method, solutions were compared to those obtained from Muskhelishvili's method [4] and integral equation methods [1], respectively, for two special cases. Then, using the finite element method, the effects of the branch-crack length and the penetrated-layer width were examined. It is noted that the related problems with different geometries have been examined by a number of researchers. For example, the tendency of a crack to deflect or penetrate at an interface between two dissimilar elastic materials in a double-edged notch specimen [5] and the effect of flaws [6] have been investigated using the boundary element method. Recently Bush [7] studied the effects of a pre-existing flaw on the interface of a particle on the crack propagation.

2. The system

Schematic drawings of crack penetration and deflection for an oblique wedge-loaded crack used in the previous study [1] are shown in Fig. 1a and b, respectively. Two semi-infinite elastic materials, material 1 and material 2, are bonded at the interface. The main crack in material 2 is subjected to wedge opening loads, P, at a distance, L, from the interface along the crack line. The crack intersects the interface at an oblique angle, ω_2, and can either penetrate across the interface into material 1 or deflect at the interface. The resultant branch crack has a length a. For the case of a penetrating crack (Fig. 1a), the angle between the branch crack and the interface is ω_1. To examine the effects of a finite length of the branch crack, the problems were solved for finite values of a/L. To examine the effects of a finite width of material 1, a layer of material 1 with a width, h, sandwiched between two semi-infinite material 2 was considered, and the schematic drawings for crack penetration and deflection are shown, respectively, in Fig. 1c and d.

For the plane-strain bimaterial problem, the solution variables of interest depend on two non-dimensional elastic mismatch parameters; i.e. the Dundurs' parameters [8] which are

$$\alpha = (\bar{E}_1 - \bar{E}_2)/(\bar{E}_1 + \bar{E}_2) \tag{1}$$

$$\beta = \tfrac{1}{2}[\mu_1(1-2\nu_2) - \mu_2(1-2\nu_1)]/[\mu_1(1-\nu_2) + \mu_2(1-\nu_1)] \tag{2}$$

where E, μ and ν are Young's modulus, shear modulus and Poisson's ratio, respectively, $\bar{E} = E/(1-\nu^2)$, and the subscripts 1 and 2 denote materials 1 and 2, respectively. Since experience with related problems suggests that α is the much more important one of the two parameters, the role of α is emphasized and $\beta = 0$ is taken in the present study.

3. Analyses

The finite element method was used in the present study to analyze: (1) the energy release rates for a deflected crack and a penetrating crack; and (2) the stress intensity factors and, hence, the mode mixity for a deflected crack. While the comparison of the energy release rate ratio to the fracture energy ratio of a deflected crack to a penetrating crack defines the criterion of crack deflection versus penetration, the mode mixity of a deflected crack characterizes the tendency for the deflected crack to kink into the adjoining layer.

3.1. The energy release rates and stress intensity factors

The solutions for the stress intensity factors for the problem of a penetrating crack depicted by Fig. 1a can be written as [1]

$$K_{\mathrm{I}} + \mathrm{i}K_{\mathrm{II}} = c(\alpha, \omega_1, \omega_2, a/L)PL^{-1/2} \tag{3}$$

where $\mathrm{i} = \sqrt{-1}$, and c is a dimensionless complex-valued function of the arguments indicated. The corresponding energy release rate, G_{p}, is

$$G_{\mathrm{p}} = \frac{(1-\nu_1)}{2\mu_1}|c|^2\frac{P^2}{L} \tag{4}$$

It is noted that G_{p} is a function of ω_1, and the maximum value of G_{p} with respect to ω_1 for a fixed a/L is denoted by $G_{\mathrm{p}}^{\max}$.

The stress intensity factors for a deflected crack (Fig. 1b)

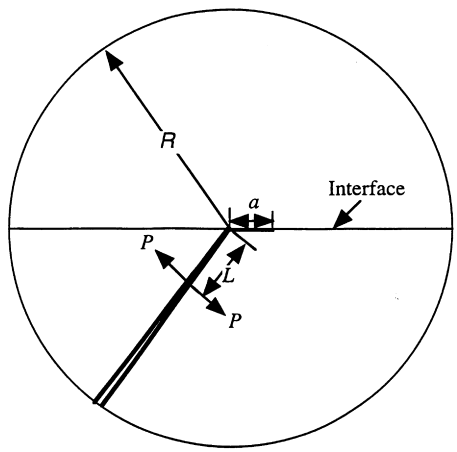


Fig. 2. A schematic drawing of the circular region adopted in the finite element calculation.

can be expressed by [1]

$$K_1 + \mathrm{i}K_2 = d(\alpha, \omega_2, a/L)PL^{-1/2} \tag{5}$$

where d is a dimensionless complex-valued function of the arguments indicated. The corresponding energy release rate of the deflected crack, G_d, is given by

$$G_\mathrm{d} = \frac{1}{\bar{E}^*}|d|^2P^2/L \tag{6}$$

where $\bar{E}^*$ is given by

$$\frac{1}{\bar{E}^*} = \left[\frac{1}{\bar{E}_1} + \frac{1}{\bar{E}_2}\right]\frac{1-\beta^2}{2} \equiv \frac{1-\beta^2}{(1-\alpha)\bar{E}_1} \tag{7}$$

The ratio of the competing energy release rates is hence

$$\frac{G_\mathrm{d}}{G_\mathrm{p}^\mathrm{max}} = \frac{|d|^2}{(1-\alpha)|c|^2} \tag{8}$$

3.2. The finite element method and convergence studies

The numerical results presented in this paper were computed using a finite element code, ABAQUS (Version 5.5) [9], with eight-node isoparametric elements. A quarter-point crack tip element served to model the inverse square root stress singularity at the crack tip. The model employed in the finite element calculation was a circular region with a radius R and its origin located at the intersection of the main crack and the interface (Fig. 2). In order to obtain the asymptotic solution, R should be much larger than both the length of the branch crack, a, and the distance from the wedge load to the crack tip, L. A refined mesh was used, and the detailed discussion of the finite element mesh can be found in the Appendix of Ref. [10]. For example, the mesh for the deflected crack contained 2499 eight-node isoparametric elements and 9339 nodes.

Two techniques were employed to calculate strain energy release rates. In the first, the J-integral was calculated by the domain integral method [11] for ten contours. In the second, the stress intensity factors K_1 and K_2 were obtained from the crack opening displacements, δ_y and δ_x, in accordance with (for $\beta = 0$)

$$K_1 = \lim_{r\to 0}\frac{E_1\sqrt{\pi}\delta_y}{4\sqrt{2r}} \tag{9a}$$

$$K_2 = \lim_{r\to 0}\frac{E_1\sqrt{\pi}\delta_x}{4\sqrt{2r}} \tag{9b}$$

The mode mixity, ψ, was then obtained, such that

$$\psi = \tan^{-1}(K_2/K_1) \tag{10}$$

It was found that the energy release rates obtained by these two methods were in very good agreement. The results for the energy release rate presented in this paper were obtained by J-integral, and the results for the mode mixity were obtained by the stress intensity factor method.

To ensure the convergence of the finite element solutions, results were compared to those obtained from Muskhelishvili's method and integral equation methods, respectively, for two special cases shown as follows.

3.2.1. Comparison to Muskhelishvili's method

For an infinite plane containing a semi-infinite crack, which is subjected to a point load, P, at a distance L from the crack tip, the exact analytical solution for the stress intensity at the crack tip, K_1, has been derived using Muskhelishvili's method, such that [12]

$$K_1 = \frac{p}{\sqrt{2\pi L}} \tag{11a}$$

The corresponding strain energy release rate, G, is hence

$$G = \frac{p^2(1-\nu^2)}{2\pi LE} \tag{11b}$$

Comparison of the finite element analysis using different ratios of R/L to the exact analytical solution (Eq. (11b)) is shown in Table 1. The results indicate that the ratio

Table 1
Errors of the strain energy release rate from the finite element solutions for meshing a circular region with a radius R for the system of an infinite plane containing a semi-infinite crack subjected to a point load at a distance L from the crack tip

R/L	200	400	1000	2500	5000
Error (%)	5.43	2.63	1.06	0.32	0.14

Table 2
Comparison of energy release rate ratio $G_\mathrm{d}/G_\mathrm{p}^\mathrm{max}$ between finite element solutions and integral equation solutions ($a/L = 0.1$ and $\omega_2 = 45°$)

α	$G_\mathrm{d}/G_\mathrm{p}^\mathrm{max}$	
	Finite element	Integral equation
−0.5	0.637	0.639
−0.1	0.686	0.688
0.1	0.754	0.759

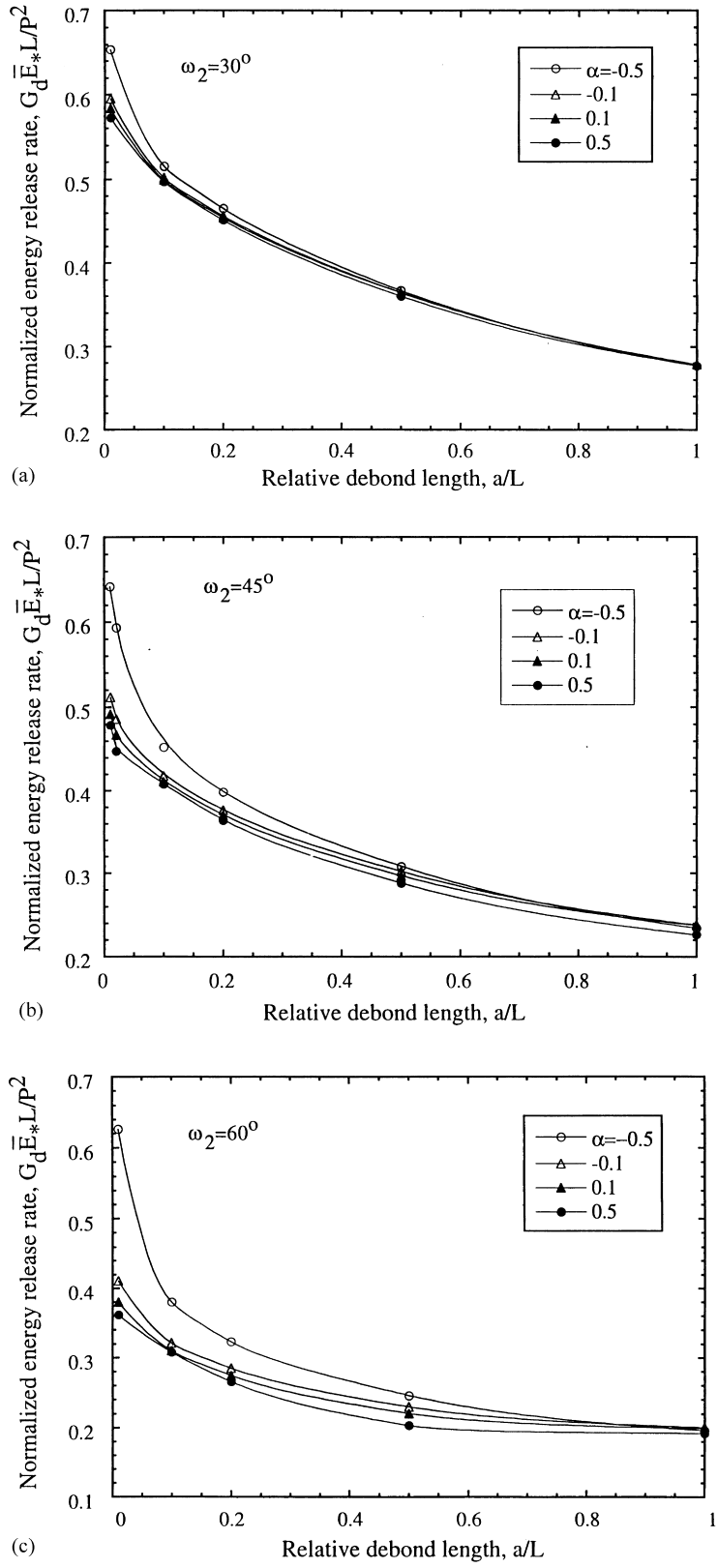


Fig. 3. The normalized energy release rate, $G_d\bar{E}^*L/P^2$, as a function of the relative debonding length, a/L, (a) for $\omega_2 = 30°$, (b) for $\omega_2 = 45°$, (c) for $\omega_2 = 60°$.

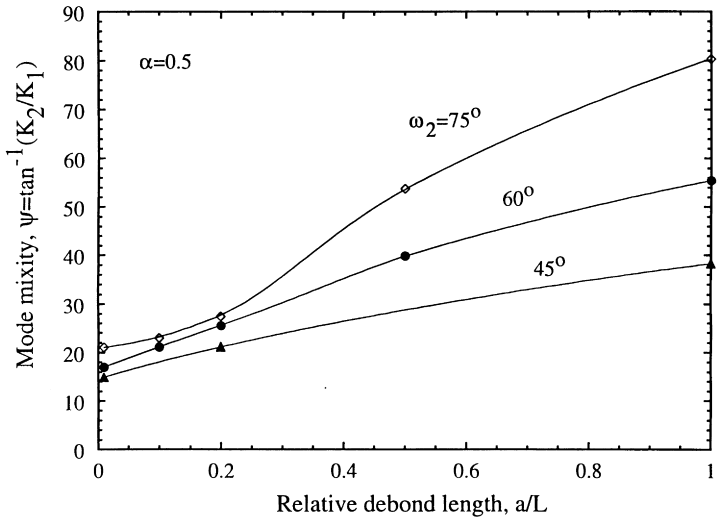


Fig. 4. The mode mixity, ψ, of the deflected crack as a function of the relative debonding length, a/L, for $\alpha = 0.5$.

$R/L = 1000$ is sufficient to obtain the asymptotic solutions. Hence, R/L ratios $\geq$1000 were used in the present finite element analysis.

3.2.2. Comparison to integral equation methods

For a penetrating crack (Fig. 1a), the energy release rate ratio G_d/G_p^{max} as a function of α has been calculated using integral equation methods for $a/L = 0.1$ and $\omega_2 = 45°$ [1]. Using finite element analyses in the present study, results were also obtained and excellent agreement was found when comparing these values to those obtained from integral equation methods (see Table 2).

4. Results

In glass matrix containing whiskers, a deflected crack is found to propagate a finite length along the interface before it kinks into the whisker [2,3]. In the following, the effects of the debond length on the energy release rate and the mode mixity of a deflected crack were examined first. Then, the effects of both a finite length of the branch crack emanating from the main crack tip and a finite width of the adjoining layer subjected to crack penetration on the criterion of crack deflection versus penetration were studied.

4.1. Effects of the debond length on the energy release rate and mode mixity of a deflected crack

For the geometry depicted in Fig. 1a, the normalized energy release rate of a deflected crack, $G_d\bar{E}^*L/P^2 = |d|^2$, as a function of the relative debond length, a/L, is plotted in Fig. 3a–c, respectively, for $\omega_2 = 30$, 45 and 60° at different values of α. The energy release rate, G_d, decreases as the

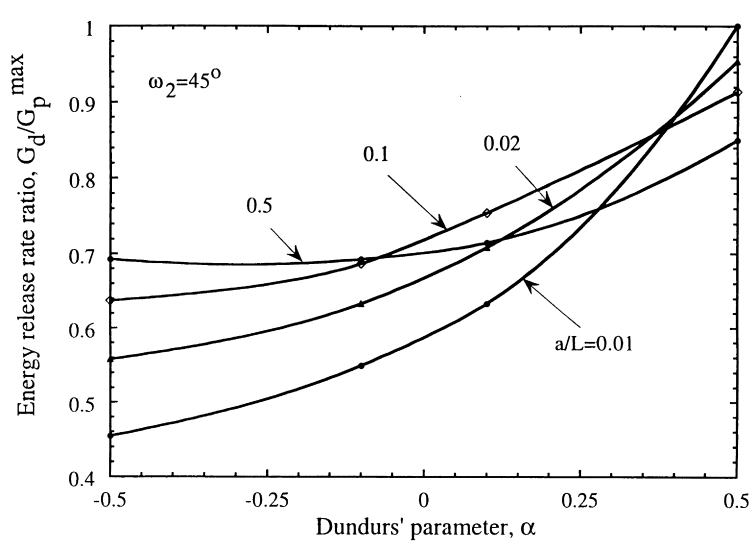


Fig. 5. The energy release rate ratio, G_d/G_p^{max}, as a function of the Dundurs' parameter, α, at different lengths of the branch crack for $\omega_2 = 45°$.

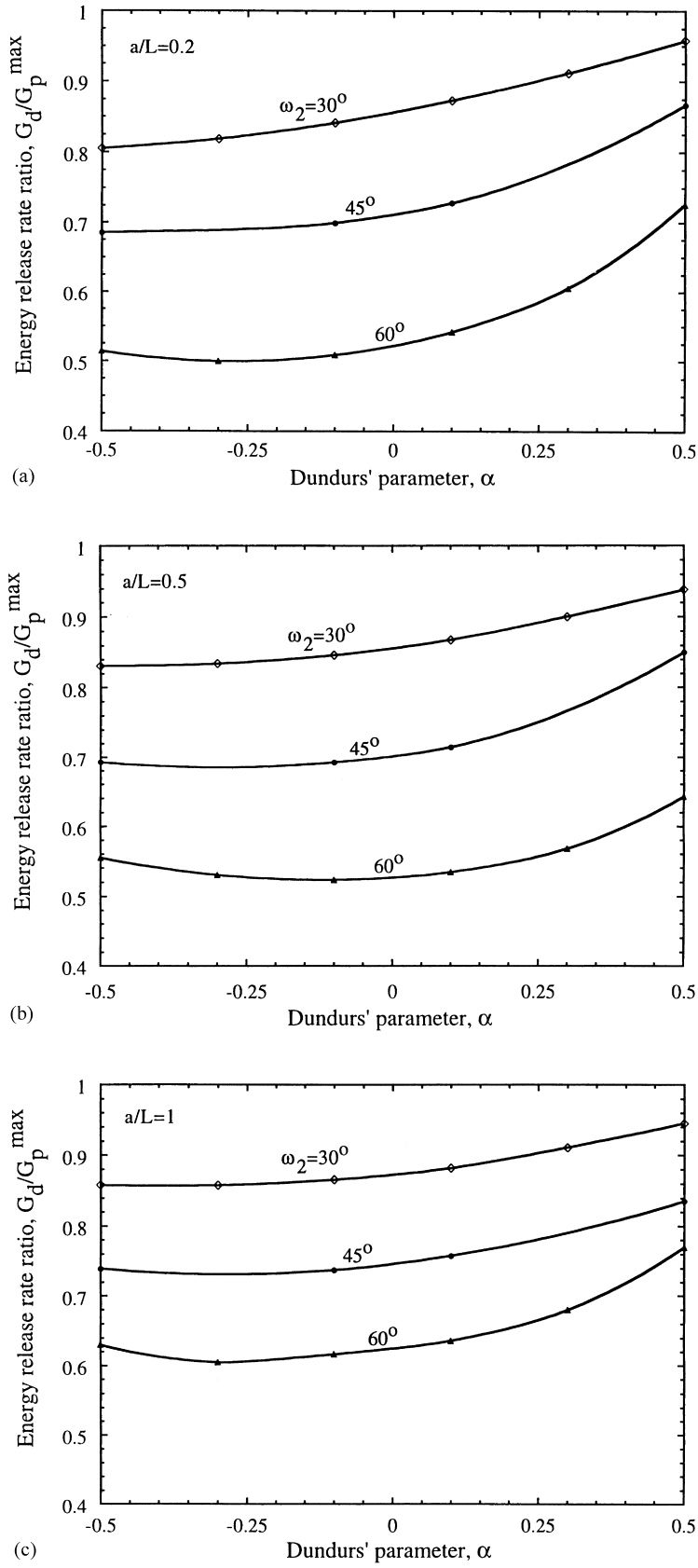


Fig. 6. The energy release rate ratio, G_d/G_p^{max}, as a function of the Dundurs' parameter, α, at different oblique angles, ω_2: (a) for $a/L = 0.2$; (b) for $a/L = 0.5$; (c) for $a/L = 1$.

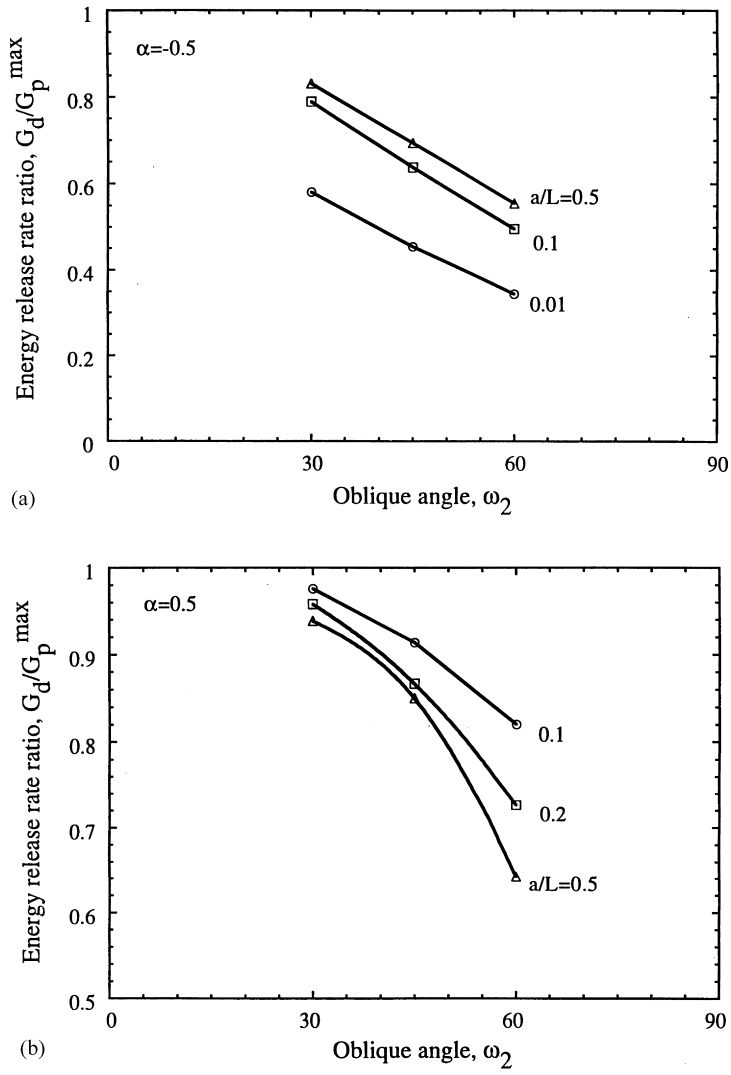


Fig. 7. The energy release rate ratio, G_d/G_p^{max}, as a function of the oblique angle, ω_2: (a) for $\alpha = -0.5$; (b) for $\alpha = 0.5$.

debond length, a, increases. However, the elastic mismatch parameter, α, has only a secondary effect on the normalized energy release rate. The associated mode mixity, ψ, of the deflected crack as a function of the relative debond length, a/L, is given in Fig. 4 for $\alpha = 0.5$ at different oblique angles, ω_2. The results show that the mode mixity, ψ, increases with increase in both the debond length and the oblique angle. Hence, based on Figs. 3 and 4, the energy release rate G_d (and hence G_d/G_p^{max}) decreases and the mode mixity ψ increases as the deflected crack grows which, in turn, make it more likely that the deflected crack in the interface will kink into material 1. These results agree with experimental observations [2,3]. However, to quantitatively characterize the bonding strength at the whisker/glass interface from the deflected crack length, G_p^{max} for a deflected crack to kink into material 1 should also be calculated. This involves two stress singularities in the finite element analysis and will be done in the future.

4.2. Effects of the branch-crack length on the energy release rate ratio, G_d/G_p^{max}

Using the geometries depicted in Fig. 1a and b, the energy release rate ratio, G_d/G_p^{max}, as a function of α is plotted in Fig. 5 at different lengths of the branch crack (from $a/L = 0.01$ to 0.5) for $\omega_2 = 45°$. It can be seen that the relative branch-crack length, a/L, has significant effects on the energy release rate ratio.

While G_d/G_p^{max} as a function of α (i.e. the criterion of crack deflection versus penetration) in an earlier work [1] was established for $a/L = 0.1$, curves of G_d/G_p^{max} as a function of α for $\omega_2 = 30$, 45 and 60° are shown in Fig. 6a–c for $a/L = 0.2$, 0.5 and 1.0, respectively. Curves of G_d/G_p^{max} as a function of ω_2 at different ratios of a/L are shown in Fig. 7a and b, respectively, for $\alpha = -0.5$ and 0.5. The energy release rate ratio, G_d/G_p^{max}, decreases with the increase in the oblique angle, ω_2. It can also be seen that effects of the

branch-crack length on the energy release rate ratio depend on α. For $\alpha = -0.5$ the ratio increases with the increasing branch-crack length (Fig. 7a). On the other hand, for $\alpha = 0.5$, the ratio decreases with the increasing crack length (Fig. 7b).

4.3. Effects of a finite width of material 1

For ceramic composites, the reinforcement embedded in the matrix has a finite width. To examine the effects of the reinforcement width on the criterion of crack deflection versus penetration, simplified geometries depicted by Fig. 1c and d were adopted in the present study. The energy release rate ratio, G_d/G_p^{max}, as a function of α is plotted in Fig. 8 at different widths of the sandwiched layer (i.e. material 1) for $\omega_2 = 30, 45, 60$ and $75°$ and $a/L = 0.01$. The curve for $L/h = 0$ (i.e. $h \rightarrow \infty$) is also included which was the result presented in the earlier work [1]. The results in Fig. 8 show that the curve becomes flatter as the width of material 1 decreases (i.e. L/h increases).

The results for $\omega_2 = 45$ and $60°$ in Fig. 8 are, respectively, replotted in Fig. 9a and b, in which G_d/G_p^{max} as a function of the main crack length to sandwiched layer width ratio, L/h, is plotted at different values of α. When $\alpha = 0$ (i.e. materials 1 and 2 are the same), G_d/G_p^{max} is independent of the width of material 1. When $\alpha > 0$ and $\alpha < 0$, the calculated G_d/G_p^{max} decreases and increases, respectively, as the width of material 1 decreases (i.e. L/h increases).

5. Concluding remarks

By considering two semi-infinite elastic materials bonded at the interface, specific crack propagation problems have been previously analyzed [1]. When a crack reaches the interface, the crack either deflects along the interface or penetrates into the next layer depending upon the ratio of energy release rate due to debonding to that due to crack penetration. This criterion has been used extensively to predict interfacial debonding versus reinforcement fracture for a crack propagating in fiber- (whisker- and self-) reinforced ceramic composites. However, two limitations should be noted before the criterion is applied. First, the results were obtained based on the condition that the branch crack emanating from the main crack tip was very small compared with the main-crack length. Second, the fiber has a finite width and is not semi-infinite in a two-dimensional sense.

The present study examined effects of: (1) a finite branch-crack length; and (2) a finite penetrated-layer width on the criterion of crack deflection versus penetration, and the following results were concluded.

1. For a deflected crack (Fig. 1b), the energy release rate decreases (Fig. 3) and the mode mixity, ψ, increases (Fig. 4) as the deflected crack grows. This would make it more favorable for the deflected crack in the interface to kink into the adjoining layer as the deflected crack grows longer.
2. The branch-crack length has effects on the energy release rate ratio of crack deflection to penetration, G_d/G_p^{max} (Figs. 5 and 6). The energy release rate ratio, G_d/G_p^{max}, can increase or decrease with the increase in the branch-crack length depending on the Dundurs' parameter, α (Fig. 7a and b).
3. The penetrated-layer width has significant effects on the energy release rate ratio of crack deflection to penetration, G_d/G_p^{max}. The diagram of crack deflection versus penetration (i.e. G_d/G_p^{max} as a function of α) is shown in Fig. 8. Specifically, G_d/G_p^{max} is independent of the width of the penetrated-layer only when $\alpha = 0$ (Fig. 9),

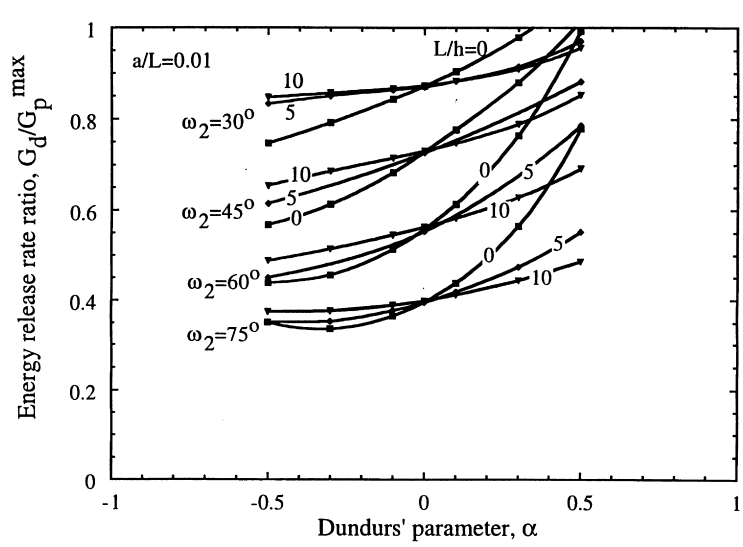


Fig. 8. The energy release rate ratio, G_d/G_p^{max}, as a function of the Dundurs' parameter, α, for $a/L = 0.01$ at different values of L/h and ω_2.

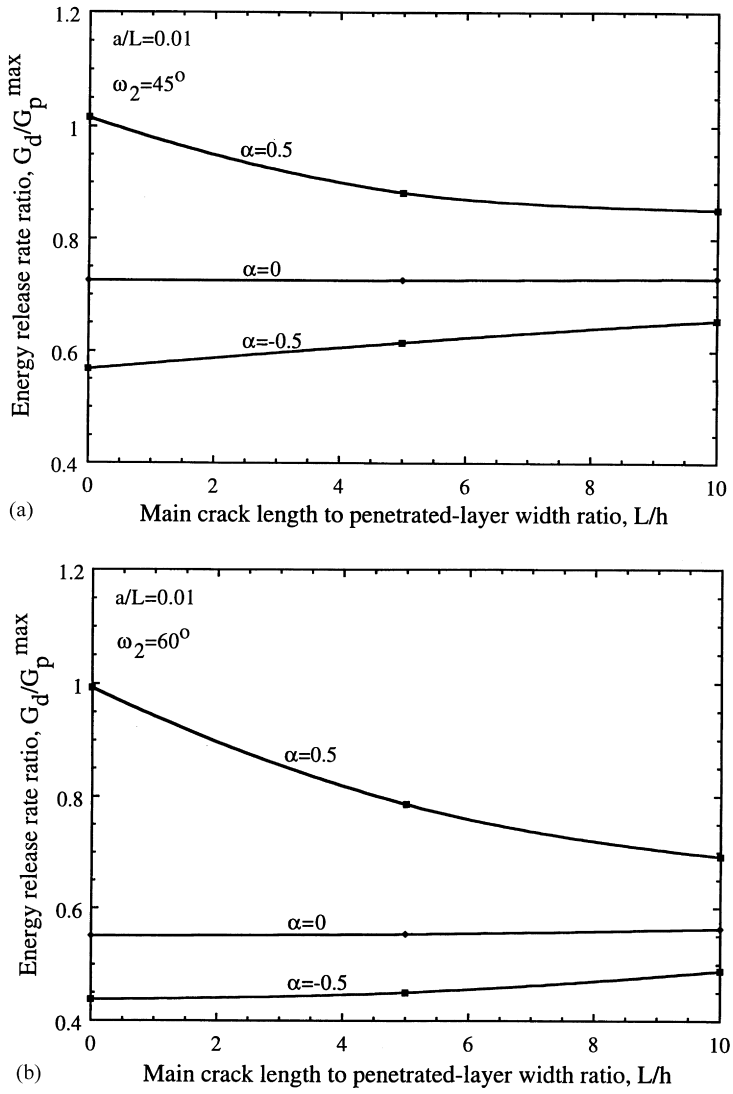


Fig. 9. The energy release rate ratio, G_d/G_p^{max}, as a function of the main crack length to penetrated-layer width ratio, L/h, at different values of α: (a) for $\omega_2 = 45°$; (b) for $\omega_2 = 60°$.

and the curve in Fig. 8 becomes flatter as the width of the penetrated-layer (i.e. material 1 in Fig. 1c and d) decreases.

The present analysis is two-dimensional and is applicable to the layered materials. It should be noted that the crack propagation problem in fiber-reinforced composites is three-dimensional. For an embedded fiber of a finite radius, there are three options when a matrix crack reaches the interface: the interface can debond, the fiber can fracture, or the crack can circumvent the fiber. The present study considered a simplified geometry of material 1 with a finite width sandwiched between two semi-infinite material 2 (Fig. 1c and d), and hence focused on the case that the matrix crack does not circumvent the fiber. However, when the matrix crack circumvents the fiber, the crack is bridged by intact fibers, and the bridging-fiber (or fiber-pullout) geometry [13–15] can be used as a representative volume element for this case. Considering a bridging fiber behind the crack tip, G_d/G_p^{max} as a function of α (for $\omega_2 = 90°$) has been derived elsewhere [16]. It is found that this ratio decreases with the increase in the Dundurs' parameter, α. Also, the curve in the diagram of interfacial debonding versus fiber fracture becomes flatter as the fiber radius decreases [16].

Acknowledgements

The authors thank Dr M.J. Andrews and Dr M.J. Lance for reviewing the manuscript. Research sponsored by the US Department of Energy, Division of Materials Sciences, Office of Basic Energy Sciences, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

References

[1] He MY, Hutchinson JW. Crack deflection at an interface between dissimilar elastic materials. Int J Solids Struct 1989;25(9):1053–67.
[2] Becher PF, Hwang SL, Hsueh CH. Using microstructure to attack the brittle nature of silicon nitride ceramics. MRS Bull Ser: Silicon-based Ceram 1995;20:23–7.
[3] Sun EY, Becher PF, Waters SB, Hsueh CH, Plucknett KP, Hoffmann MJ. Control of interface in silicon nitride ceramics: influence of different rare earth elements. In: Tomsia AP, Glaeser A, editors. Ceramic Microstructures '96: Control at the atomic level, New York: Plenum Press, 1998. p. 779–86.
[4] Muskhelishili NI. Some basic problems of mathematical theory of elasticity. Leiden: P. Noordhoff, 1953.
[5] Tullock DL, Reimanis IE, Graham AL, Petrovic JJ. Deflection and penetration of cracks at an interface between two dissimilar materials. Acta Metall Mater 1994;42:3245–52.
[6] Mammoli AA, Graham AL, Reimanis IE, Tullock DL. The effect of flaws on the propagation of cracks at bi-materials interfaces. Acta Metall Mater 1995;43:1149–56.
[7] Bush MB. The interaction between a crack and a particle cluster. Int J Fracture 1998;88:215–32.
[8] Dundurs J. Edge-bonded dissimilar orthogonal elastic wedge. J Appl Mech 1969;36:650–2.
[9] Abaqus Finite Element Code, Version 5.5, Hibbit, Karlsson and Sorensen, Inc., Pawtucket, RI.
[10] He MY, Evans AG, Hutchinson JW. Crack deflection at an interface between dissimilar elastic materials: role of residual stresses. Int J Solids Struct 1994;31(24):3443–55.
[11] Shih CF, Moran B, Nakamura J. Energy release rate along a three-dimensional crack front in a thermally stressed body. Int J Fracture 1986;30:79–102.
[12] Tada H, Paris PC, Irwin GR. The stress analysis of cracks handbook. Paris Productions, 1985.
[13] Gao YC, Mai YW, Cottrell B. Fracture of fiber-reinforced materials. J Appl Math Phys 1988;39(7):550–72.
[14] McCartney LN. New theoretical model of stress transfer between fiber and matrix in a uniaxially fiber-reinforced composite. Proc R Soc London A 1989;425:215–44.
[15] Hsueh CH. Interfacial debonding and fiber pull-out stresses of fiber-reinforced composites. Mater Sci Engng A 1990;123(1):1–11.
[16] Hsueh CH, Becher PF. Interfacial shear debonding problems in fiber-reinforced ceramic composites. Acta Mater 1998;46(9):3237–45.

PERGAMON

Scripta mater. 42 (2000) 781–787

Scripta MATERIALIA

www.elsevier.com/locate/scriptamat

ANALYTICAL MODELING OF OXIDE THICKNESS EFFECTS ON RESIDUAL STRESSES IN THERMAL BARRIER COATINGS

C.H. Hsueh[1] and E.R. Fuller, Jr.[2]

[1] Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6068

[2] National Institute of Standards and Technology, Gaithersburg, Maryland 20899

(Received September 27, 1999)
(Accepted in revised form December 2, 1999)

Keywords: Modeling; Thermal barrier coating; Thermal expansion; Stress

Introduction

During high temperature operation, an oxide scale forms along the irregular top coat/bond coat interface in the plasma-sprayed thermal barrier coating (TBC) system. The residual stresses in the system are affected by the presence of the thermally grown oxide (TGO) [1–5]. Along the irregular interface, the asperity can be convex or concave. Semicircular convex and concave asperities have been adopted for numerical simulations to examine the effects of the TGO thickness on residual thermal stresses [6]. It was found that in the tip region of a convex asperity, the residual stress normal to the interface, σ_y, in the ceramic top coat is tensile for a thin oxide but becomes compressive for a thick oxide. In the tip region of a concave asperity, σ_y in the ceramic top coat is compressive for a thin oxide and becomes less compressive for a thick oxide.

The purpose of the present study was to explore the physical meaning of the trend of the stress state of σ_y in the ceramic top coat with the variation of the TGO thickness. To achieve this, a simple analytical model of three concentric circles was adopted. First, the residual thermal stresses in the three-concentric-circles model were derived. Then, the results for residual radial stresses at the top coat/TGO and the TGO/bond coat interfaces as functions of the TGO thickness were presented. Also, the physical meaning of the above results was discussed.

Analytical Modeling

A two-concentric-circles model has been used previously to study the residual stresses in a coating formed on a wrinkled substrate [7,8]. A three-concentric-circles model is hence adopted in the present study to simulate the wrinkle geometry in a TBC system. A circular region of phase 1 with a radius a is surrounded by a concentric annulus of phase 2 with an outer radius b which is then surrounded by another concentric annulus of phase 3 with an outer radius c (see Fig. 1). Whereas phases 1, 2, and 3 are designated respectively as the bond coat, the TGO, and the top coat in modeling a convex asperity, they are designated respectively as the top coat, the TGO, and the bond coat in modeling a concave asperity. In both cases, the TGO thickness is dictated by b-a.

Residual thermal stresses develop during the temperature change in the system because of the thermal-mechanical mismatch among the three phases. The stresses can be determined by the procedure

1359-6462/00/$–see front matter. © 2000 Acta Metallurgica Inc. Published by Elsevier Science Ltd. All rights reserved.
PII: S1359-6462(99)00430-3

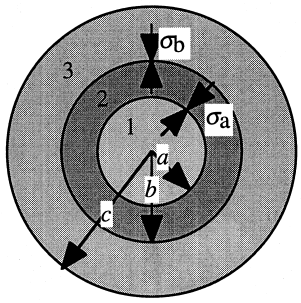


Figure 1. A schematic showing the three-concentric-circles model.

of first allowing the three phases to exhibit unconstrained differential thermal strains during the temperature change. Then, radial stresses, σ_a and σ_b, are placed at the interfaces, $r = a$ and $r = b$, respectively to restore the displacement continuity at the two interfaces. With interfacial radial stresses, σ_a and σ_b, the radial and the tangential stresses in phases 1, 2, and 3 are

$$\sigma_{1,r} = \sigma_a \tag{1a}$$

$$\sigma_{1,t} = \sigma_a \tag{1b}$$

$$\sigma_{2,r} = \frac{a^2b^2(\sigma_a - \sigma_b)}{r^2(b^2 - a^2)} - \frac{a^2\sigma_a - b^2\sigma_b}{b^2 - a^2} \tag{2a}$$

$$\sigma_{2,t} = \frac{-a^2b^2(\sigma_a - \sigma_b)}{r^2(b^2 - a^2)} - \frac{a^2\sigma_a - b^2\sigma_b}{b^2 - a^2} \tag{2b}$$

$$\sigma_{3,r} = \frac{b^2(c^2 - r^2)\sigma_b}{r^2(c^2 - b^2)} \tag{3a}$$

$$\sigma_{3,t} = \frac{-b^2(c^2 + r^2)\sigma_b}{r^2(c^2 - b^2)} \tag{3b}$$

where r is the distance from the center of the circles, the subscripts 1, 2, and 3 denote phases 1, 2, and 3, and the subscripts r and t denote the radial and the tangential components, respectively.

With a temperature change, ΔT, the strain consists of two components: the elastic strain and the unconstrained thermal strain. The tangential strains in phases 1, 2, and 3, $\epsilon_{1,t}$, $\epsilon_{2,t}$, and $\epsilon_{3,t}$, are respectively

$$\epsilon_{1,t} = \frac{\sigma_{1,t} - \nu_1\sigma_{1,r}}{E_1} + \alpha_1\Delta T \tag{4}$$

$$\epsilon_{2,t} = \frac{\sigma_{2,t} - \nu_2\sigma_{2,r}}{E_2} + \alpha_2\Delta T \tag{5}$$

$$\epsilon_{3,t} = \frac{\sigma_{3,t} - \nu_3\sigma_{3,r}}{E_3} + \alpha_3\Delta T \tag{6}$$

where E, ν, and α are Young's modulus, Poisson's ratio, and the coefficient of thermal expansion (CTE), respectively.

The residual stresses in the system are contingent upon the solutions of the interfacial radial stresses, σ_a and σ_b, which can be determined by the continuity condition at the interface. Continuity of the radial displacement at the interface is required. With the tangential strain proportional to the radial displacement, this continuity condition becomes

$$\epsilon_{1,t} = \epsilon_{2,t} \qquad (\text{at } r = a) \tag{7}$$

$$\epsilon_{2,t} = \epsilon_{3,t} \qquad (\text{at } r = b) \tag{8}$$

Combination of Equations (1) through (8), yields

$$\sigma_a = \frac{E_2[P_1(\alpha_2 - \alpha_1) + P_4(\alpha_3 - \alpha_2)]\Delta T}{P_1P_2 - P_3P_4} \tag{9}$$

$$\sigma_b = \frac{E_2[P_3(\alpha_2 - \alpha_1) + P_2(\alpha_3 - \alpha_2)]\Delta T}{P_1P_2 - P_3P_4} \tag{10}$$

where the parameters, P_1, P_2, P_3, and P_4, are given by

$$P_1 = \frac{b^2 + a^2}{b^2 - a^a} - \nu_2 + \frac{E_2}{E_3}\left(\frac{c^2 + b^2}{c^2 - b^2} + \nu_3\right) \tag{11a}$$

$$P_2 = \frac{b^2 + a^2}{b^2 - a^a} + \nu_2 + \frac{E_2(1 - \nu_1)}{E_1} \tag{11b}$$

$$P_3 = \frac{2a^2}{b^2 - a^2} \tag{11c}$$

$$P_4 = \frac{2b^2}{b^2 - a^2} \tag{11d}$$

The stress state at $r = b$ (i.e., σ_b is tensile or compressive) depends on the comparison between α_3 and the effective CTE in the region of $r<b$ which is a function of dimensions, elastic constants, and CTEs of phases 1 and 2. This effective CTE, α_{12}, is dictated by the radial displacement at $r = b$ in the absence of phase 3, and has been derived previously, such that [9]

$$\alpha_{12} = \alpha_2 + \frac{P_3(\alpha_1 - \alpha_2)}{P_2} \tag{12}$$

In the specific case of $E_1 = E_2$ and $\nu_1 = \nu_2$, Equation (12) becomes

$$\alpha_{12} = \frac{a^2\alpha_1 + (b^2 - a^2)\alpha_2}{b^2} \qquad (\text{for } E_1 - E_2 \text{ and } \nu_1 = \nu_2) \tag{13}$$

Hence, when phases 1 and 2 have the same elastic constants, the effective CTE, α_{12}, can be obtained using rule-of-mixtures.

Similarly, the stress state at $r = a$ depends on the comparison between α_1 and the effective CTE in the region of $a<r<c$ which is a function of dimensions, elastic constants, and CTEs of phases 2 and 3. This effective CTE, α_{23}, is dictated by the radial displacement at $r = a$ in the absence of phase 1. Using the similar technique as that for deriving α_{12}, α_{23} can be obtained, such that

$$\alpha_{23} = \alpha_2 + \frac{P_4(\alpha_3 - \alpha_2)}{P_1} \qquad (\text{at } r = a) \tag{14}$$

In the specific case of $E_2 = E_3$ and $\nu_2 = \nu_3$, Equation (14) becomes

$$\alpha_{23} - \frac{(b^2 - a^2)\alpha_2 + (c^2 - b^2)\alpha_3}{c^2 - a^2} \qquad (\text{for } E_2 = E_3 \text{ and } \nu_2 = \nu_3) \tag{15}$$

Hence, when phases 2 and 3 have the same elastic constants, the effective CTE, α_{23}, can also be obtained using rule-of-mixtures. It can also be derived that when the radial displacement at $r = c$ is considered, the effective CTE, α_{23}, becomes

$$\alpha_{23} = \alpha_3 - \frac{2b^2E_2(\alpha_3 - \alpha_2)}{(c^2 - b^2)P_1E_3} \qquad (\text{at } r = c) \tag{16}$$

It is noted that Equation (16) is different from Equation (14). Nevertheless, when phases 2 and 3 have the same elastic constants, Equation (16) also becomes Equation (15). However, to determine the stress state at $r = a$, comparison should be made between α_1 and α_{23} given by Equation (14).

Using the definitions of α_{12} and α_{23} in Equations (12) and (14), α_a and α_b in Equations (9) and (10) become

$$\sigma_a = \frac{P_1E_2(\alpha_{23} - \alpha_1)\Delta T}{P_1P_2 - P_3P_4} \tag{17}$$

$$\sigma_b = \frac{P_2E_2(\alpha_3 - \alpha_{12})\Delta T}{P_1P_2 - P_3P_4} \tag{18}$$

When a three dimensional geometry is considered, a three-concentric-spheres model can be used. The analytical steps for deriving the residual thermal stresses in the three-concentric-spheres model are similar to those for the three-concentric-circles model. The solutions for α_{12}, α_{23}, α_a, and α_b have the same general forms as those for the three-concentric-circles model [i.e., Equations (12), (14), (17), and (18)]. However, the parameters, P_1, P_2, P_3, and P_4, become

$$P_1 = \frac{2b^3(1 - 2\nu_2) + a^3(1 + \nu_2)}{2(b^3 - a^3)} + \frac{E_2}{E_3}\left[\frac{c^3(1 + \nu_3) + 2b^3(1 - 2\nu_3)}{2(c^3 - b^3)}\right] \tag{19a}$$

$$P_2 = \frac{b^3(1 + \nu_2) + 2a^3(1 - 2\nu_2)}{2(b^3 - a^3)} + \frac{E_2(1 - 2\nu_1)}{E_1} \tag{19b}$$

$$P_3 = \frac{3a^3(1 - \nu_2)}{2(b^3 - a^3)} \tag{19c}$$

$$P_4 = \frac{3b^3(1 - \nu_2)}{2(b^3 - a^3)} \tag{19d}$$

Results

The residual stresses calculated from the three-concentric-circles model were presented in the present study. The thermal/mechanical properties pertinent to the air plasma-sprayed TBC system of ZrO_2 top coat/α-Al_2O_3 TGO/vacuum plasma-sprayed Ni-22Cr-10Al-1Y bond coat listed in Table 1 were adopted

TABLE 1
The Thermal/Mechanical Properties of the Plasma Sprayed TBC System

	Young's Modulus, E (GPa)	Poisson's Ratio, ν	CTE ($\times 10^{-6}$/°C)
Top Coat	50	0.1	10
Oxide Scale	360	0.27	8
Bond Coat	200	0.3	15.2

for calculations. The temperature change, ΔT, used in calculating the thermal strain was -1125°C. Young's modulus, E, was replaced by $E/(1-\nu)$ before it was substituted into the analytical solutions due to the biaxial stress state in the TBC system as discussed in Refs. [2–5].

For a convex asperity, phases 1, 2, and 3 correspond to the bond coat, the TGO, and the top coat, respectively. The residual radial stresses at $r = a$ and $r = b$ (i.e., σ_a and σ_b) as functions of the normalized TGO thickness, $(b-a)/a$, are shown in Fig. 2a for $c/a = 5$, 10, and 20. In this case, σ_b corresponds to σ_y in the top coat at the tip of a convex asperity. The stresses are not sensitive to c/a when $c/a \geq 5$. The results for $c/a = 10$ almost overlap with those for $c/a = 20$, and cannot be distinguished in Fig. 2a. While σ_a is tensile and increases with the TGO thickness, σ_b is tensile when $(b-a)/a < \sim 0.75$ and becomes compressive when $(b-a)/a > \sim 0.75$. The effective CTE, α_{12}, as a function of the normalized TGO thickness, $(b-a)/a$, is shown in Fig. 2b. It is noted that α_{12} is independent of c [see Equation (12)]. The CTE of phase 3, α_3 ($=10\times10^{-6}$/°C), is also shown in Fig. 2b. It can be seen that $\alpha_{12} > \alpha_3$ when $(b-a)/a < \sim 0.75$ which, in turn, results in a tensile residual radial stress at $r = b$ during cooling. Conversely, $\alpha_{12} < \alpha_3$ when $(b-a)/a > \sim 0.75$ which, in turn, results in a compressive residual radial stress at $r = b$ during cooling.

For a concave asperity, phases 1, 2, and 3 correspond to the top coat, the TGO, and the bond coat, respectively. The residual radial stresses at $r = a$ and $r = b$ as functions of the normalized TGO thickness, $(b-a)/a$, are shown in Fig. 3a at different c/a ratios. In this case, σ_a corresponds to σ_y in the top coat at the tip of a concave asperity. The stresses are not sensitive to c/a when $c/a \geq 10$. While σ_b is compressive and its magnitude increases with the TGO thickness for $c/a \geq 10$, σ_a becomes less compressive with the increasing TGO thickness and decreasing c/a for the dimension considered in Fig. 3a. The effective CTE, α_{23}, as a function of the normalized TGO thickness, $(b-a)/a$, is shown in Fig. 3b at different c/a ratios. The CTE of phase 1, α_1 ($=10\times10^{-6}$/°C), is also shown. It can be seen that

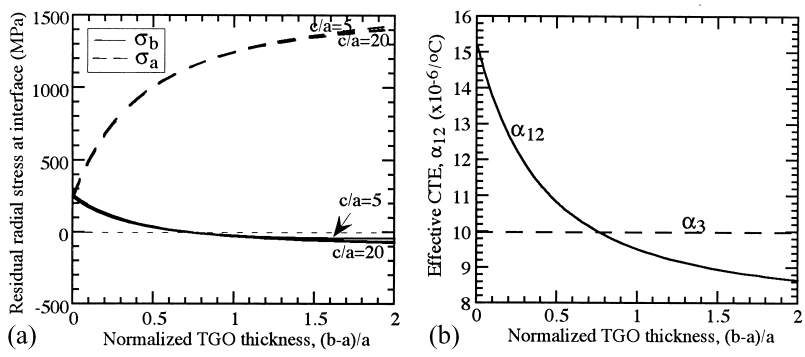


Figure 2. (a) The residual radial stress at the TGO/bond coat interface, σ_a, and at the top coat/TGO interface, σ_b, and (b) the effective CTE, α_{12}, for the region $r<b$, and the CTE of top coat, α_3, as functions of the normalized TGO thickness, $(b-a)/a$, at different c/a ratios for a convex asperity.

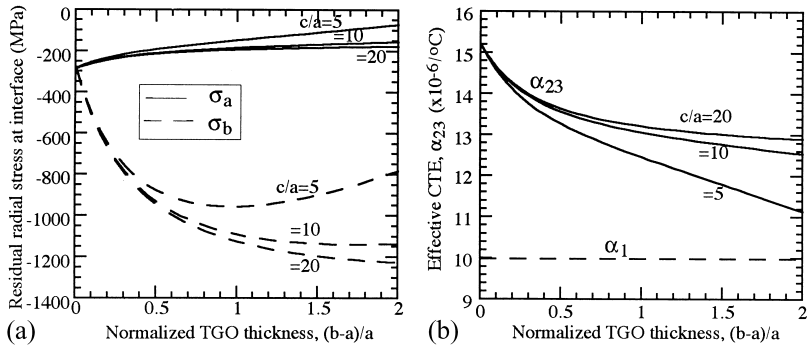


Figure 3. (a) The residual radial stress at the top coat/TGO interface, σ_a, and at the TGO/bond coat interface, σ_b, and (b) the effective CTE, α_{23}, for the region $a<r<c$, and the CTE of top coat, α_1, as functions of the normalized TGO thickness, $(b - a)/a$, at different c/a ratios for a concave asperity.

$\alpha_{23}>\alpha_1$ for the dimension considered in Fig. 3b. However, the difference between α_{23} and α_1 decreases with the increasing TGO thickness and decreasing c/a. Compared to Fig. 2b, Fig. 3b shows that the decrease in α_{23} is much slower than the decrease in α_{12} when the TGO thickness increases. This is due to the fact that the bond coat for a concave asperity (i.e., phase 3 in Fig. 1) has a much greater area than that for a convex asperity (i.e., phase 1 in Fig. 1) in the three-concentric-circles model, and the effect of the TGO thickness on the effective CTE is much weaker for a concave asperity than for a convex asperity.

Conclusions

The residual thermal stress normal to the interface in the ceramic top coat and close to the TGO is of interest, since it could result in cracking and spallation of the plasma-sprayed TBC. The effects of the TGO thickness on this residual stress were qualitatively interpreted in the present study by using an analytical model of three concentric circles (Fig. 1). An effective CTE was defined to illustrate the qualitative trends of the TGO thickness effects. To model a convex interface asperity, phases 1, 2, and 3 in Fig. 1 correspond to bond coat, TGO, and top coat, respectively. The residual radial stress at $r = b$ (i.e., at the top coat/TGO interface), σ_b, is of interest. The stress state of σ_b can be obtained by comparing α_3 with the effective CTE in the region of $r<b$ which is a function of dimensions, elastic constants, and CTEs of phases 1 and 2 (Fig. 2). To model a concave interface asperity, phases 1, 2, and 3 in Fig. 1 correspond to top coat, TGO, and bond coat, respectively. The residual radial stress at $r = a$ (i.e., at the top coat/TGO interface), σ_a, is of interest. The stress state of σ_a can be obtained by comparing α_1 with the effective CTE in the region of $a<r<c$ which is a function of dimensions, elastic constants, and CTEs of phases 2 and 3 (Fig. 3).

Acknowledgments

CHH acknowledges jointly support of the U.S. Department of Energy, Division of Materials Sciences, Office of Basic Energy Sciences, and Assistant Secretary for Energy, Efficiency and Renewable Energy, Office of Industrial Technologies, Advanced Turbine Systems Program, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. ERF acknowledges support of the Alexander von Humboldt Foundation.

References

1. J. A. Haynes, E. D. Rigney, M. K. Ferber, and W. D. Porter, Surf. Coat. Technol. 86–87, 102 (1996).
2. R. J. Christensen, V. K. Tolpygo, and D. R. Clarke, Acta Mater. 45, 1761 (1997).
3. A. G. Evans, M. Y. He, and J. W. Hutchinson, Acta Mater, 45, 3543 (1997).
4. A. M. Freborg, B. L. Ferguson, W. J. Brindley, and G. J. Petrus, Mater. Sci. Eng. A245, 182 (1998).
5. C. H. Hsueh, J. A. Haynes, M. J. Lance, P. F. Becher, M. K. Ferber, E. R. Fuller, S. A. Langer, W. C. Carter, and W. R. Cannon, J. Am. Ceram. Soc. 82, 1073 (1999).
6. C. H. Hsueh, and E. R. Fuller, unpublished work.
7. C. H. Hsueh, and A. G. Evans, J. Appl. Phys. 54, 6672 (1983).
8. X. Y. Gong, and D. R. Clarke, Oxidation Metals. 50, 355 (1998).
9. C. H. Hsueh, P. F. Becher, and E. Y. Sun, J. Mater. Sci. submitted.

ELSEVIER

Materials Science and Engineering A283 (2000) 46–55

MATERIALS SCIENCE & ENGINEERING A

www.elsevier.com/locate/msea

Residual stresses in thermal barrier coatings: effects of interface asperity curvature/height and oxide thickness

Chun-Hway Hsueh [a,*], Edwin R. Fuller Jr. [b]

[a] *Metals and Ceramics Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831-6068, USA*
[b] *National Institute of Standards and Technology, Gaithersburg, MD 20899, USA*

Received 27 September 1999; received in revised form 21 December 1999

Abstract

The effects of curvature and height of the interface asperity on residual thermal stresses in a plasma-sprayed thermal barrier coating were numerically simulated. In the tip region of a convex asperity, the residual stress normal to the interface, σ_y is tensile in the ceramic top coat and increases with both curvature and height of the asperity. However, this residual tensile stress is lower for a periodic array of asperities than for an isolated asperity. The effects of thickness of the thermally grown oxide at the top coat–bond coat interface on residual thermal stresses were also numerically simulated. In the tip region of a convex asperity, σ_y in the ceramic top coat is tensile for a thin oxide but becomes compressive for a thick oxide. In the tip region of a concave asperity, σ_y in the ceramic top coat is compressive for a thin oxide and becomes less compressive for a thick oxide. The physical meaning of the above trend was qualitatively interpreted using an analytical model of three concentric circles. Published by Elsevier Science S.A.

Keywords: Residual stresses; Coating; Numerical simulations; Analytical modeling

1. Introduction

A typical plasma-sprayed thermal barrier coating (TBC) system consists of an oxidation-resistant metallic bond coat overlaid with a porous, thermally-insulating ceramic top coat (see Fig. 1) [1–4]. It has been found that a rough bond coat surface can increase the lifetime of a plasma-sprayed TBC system during thermal cycling in a turbine engine [3,5,6]. However, the residual thermal stresses in the system are affected by the asperities present at the top coat–bond coat interface [7–17]. Specifically, the residual stress normal to the interface, σ_y, would be zero if the interface is perfectly flat, but it becomes non-zero in the presence of interface asperities. When σ_y is tensile in the ceramic top coat, it may result in cracking and spallation of the top coat. The interface asperities can have various geometries with different curvatures and heights. An issue regarding the effects of the curvature and the height of the interface asperity on the residual stresses was hence raised. Also, an oxide scale grows along the interface during high temperature exposure [2–6,18]. This thermally grown oxide (TGO) inevitably modifies the residual stresses in the system. Hence, the second issue is how the TGO thickness affects the residual stresses.

The purpose of the present study was to address the above two issues. To achieve this, the object oriented finite element analysis (OOF) [19] was utilized for numerical simulations. First, the effects of the asperity curvature were examined by simulating a convex semicircular-top asperity and a convex rectangular-top asperity with the same height, and the effects of the asperity height were examined by simulating asperities with a semicircular top (or a rectangular top) and different heights. Second, a convex and concave semicircular asperities with different TGO thicknesses were, respectively, simulated to examine effects of the TGO thickness on residual stresses. Third, a periodic array of convex and concave semicircular (or rectangular) asperities was simulated to examine effects of interaction between asperities on residual stresses. Finally, a simple analytical model of three concentric circles was used to explore the physical meaning of the trend of the TGO thickness effects on residual stresses.

* Corresponding author. Tel.: +1-865-5766586; fax: +1-865-5748445.

0921-5093/00/$ - see front matter Published by Elsevier Science S.A.
PII: S0921-5093(00)00743-7

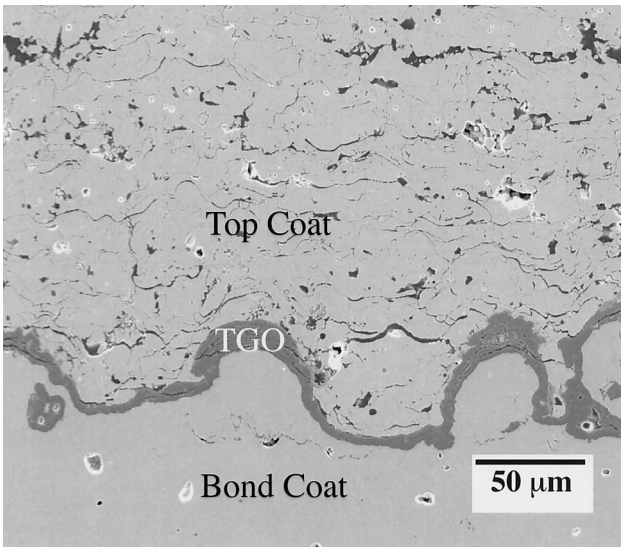


Fig. 1. The microstructure of a plasma sprayed ZrO_2 top coat–Al_2O_3 TGO–NiCrAlY bond coat system [4]. The René N5 substrate is not shown.

2. Numerical simulations and results

Numerical simulations were performed using OOF, which was uniquely designed to operate on microstructural images and has been run previously on both an actual cross-sectional microstructure and model microstructures of a plasma-sprayed TBC to obtain full-field residual stresses [4,20]. The thermal–mechanical properties pertinent to the air plasma-sprayed ZrO_2 top coat–α-Al_2O_3 TGO–vacuum plasma-sprayed Ni–22Cr–10Al–1Y bond coat–René N5 substrate system are listed in Table 1 [2,21]. The temperature change used in calculating the residual stresses was -1125°C. It is noted that the TBC system is subjected to biaxial residual stresses in the plane parallel to the interface when interface asperities are not considered. The residual stress normal to the interface induced by interface asperities is expected to be much smaller than that parallel to the interface. Hence, while the simulation is a two-dimensional analysis, E should be replaced by $E/(1-\nu)$ in applying the plane-stress analysis, OOF, to the cross-section of the TBC system to account for the biaxial-stress in the TBC system [8,10,20,22] where E is Young's modulus and ν is Poisson's ratio.

2.1. Effects of curvature and height of interface asperity

Two sets of model microstructures were used in the present study to examine the effects of the curvature and the height of the interface asperity. Since these effects can be obscured by the presence of TGO at the interface, the TGO is not included in these two sets of model microstructures. In the first set, the top of the asperity is a semicircle with a radius a, and the height of asperity is controlled by adding a block at the base of the semicircle. The semicircular asperities with heights a, $2a$, and $3a$ are shown in Fig. 2(a–c), respectively. In the second set, the top of the asperity is flat such that the asperity has a rectangular shape, and the width of the asperity is $2a$. The rectangular asperities with heights a, $2a$, and $3a$ are shown in Fig. 2(d–f), respectively. In this case, the curvatures at the asperity top are $1/a$ and 0, respectively, for semicircular and rectangular asperities.

The model microstructure was digitized to a portable pixel map (ppm) format. Using a uniform grid in OOF, each pixel in the ppm format corresponded to a grid in OOF, and each grid consisted of two triangular elements. The processed images shown in Fig. 2 consist of 64 082 elements each. Since the substrate (not shown in Figs. 1 and 2) was much thicker than the TBC, the thermal strain of the substrate (i.e. -0.0163) was used as the boundary condition at the right and the left edges of the image in Fig. 2 to constrain its distortion in the x-direction.

Using OOF, the predicted full-field residual stresses normal to the interface, σ_y, are also shown in Fig. 2(a–f). While effects of the asperity curvature can be obtained by comparing the results from the semicircular asperity with those from the rectangular asperity having the same height, effects of the asperity height can be obtained by examining the results from the semicircular (or rectangular) asperities with different heights. It can be seen that the curvature and the height of the asperity affect the residual stress field in the region of the asperity. Within the asperity, σ_y is tensile and it increases with both the curvature and the height of the asperity. The simulated results of σ_y within the top coat at the tip of the semicircular asperity and at the middle of the top of the rectangular asperity as functions of the

Table 1
The thermal–mechanical properties of the plasma sprayed ZrO_2 top coat–Al_2O_3 TGO–NiCrAlY bond coat–René N5 substrate system

	Young's modulus, E (GPa)	Poisson's ratio, ν	CTE, α ($\times 10^{-6}$/°C)
Porous top coat	50	0.1	10
Oxide scale	360	0.27	8
Bond coat	200	0.3	15.2
Substrate	213	0.25	14.5

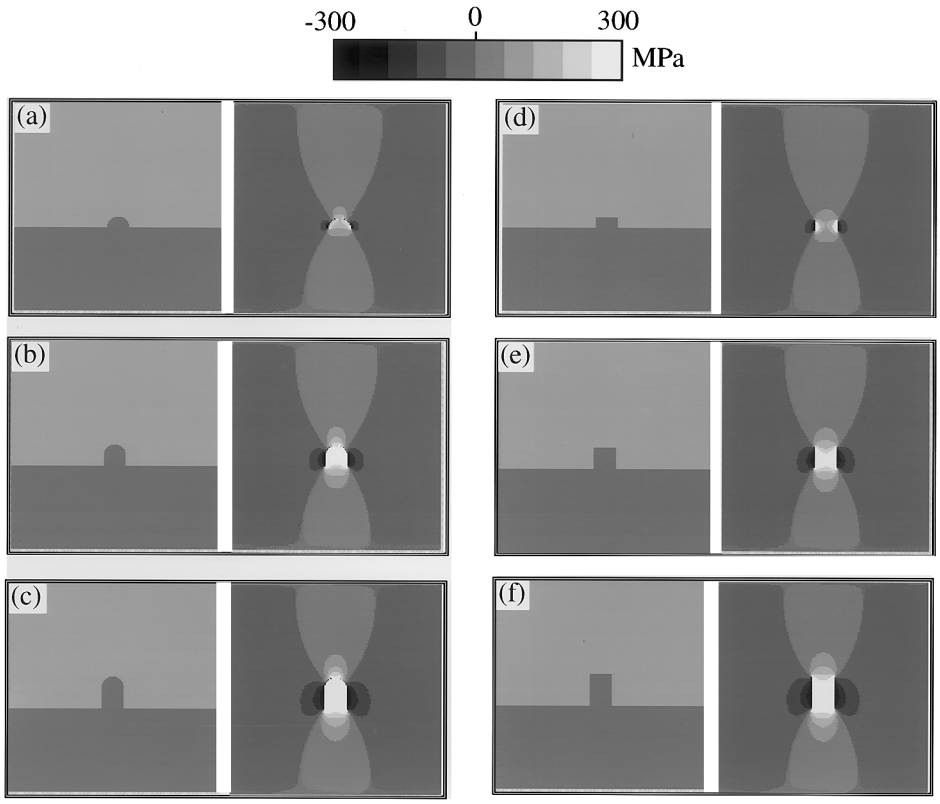


Fig. 2. The model microstructures and the simulated residual stress, σ_y, for the plasma-sprayed TBC system with a semicircular interface asperity having a height (a) a, (b) $2a$, and (c) $3a$, and with a rectangular interface asperity having a height (d) a, (e) $2a$, and (f) $3a$.

asperity height are shown in Fig. 3. For convenience, a subscript, o, is added to define σ_y at the above mentioned positions. It can be seen that σ_{yo} increases with both the curvature and the height of the asperity. The thickness ratio of the top coat to the bond coat for the model microstructures adopted in the present study is 1.5. Since the thermal strain of the substrate was used to constraint the distortion of the TBC in the x-direction, the calculated σ_{yo} is independent of the thickness ratio of the top coat to the bond coat when the asperity is remote from the free surface of the top coat.

It is noted that the effects of curvature and height of interface asperities have also been studied by assuming a sinusoidal interface with varying wavelength and amplitude [17]. The general trend exhibits that the magnitude of the residual stress normal to the interface increases with increasing wrinkling of the interface [17] which is in agreement with the present result. It is also noted that corners were introduced by using the rectangular asperity in the present study which would result in singularities. However, the stresses used for comparison are not at corners, and different sizes of mesh have been used in the OOF analysis to ensure convergence of the simulated result for σ_{yo}.

2.2. Effects of the TGO thickness

The residual stress states associated with a convex and a concave interface asperities are different. To examine the effects of the TGO thickness on residual stresses, convex and concave semicircular asperities are considered, respectively. The semicircular asperity has a

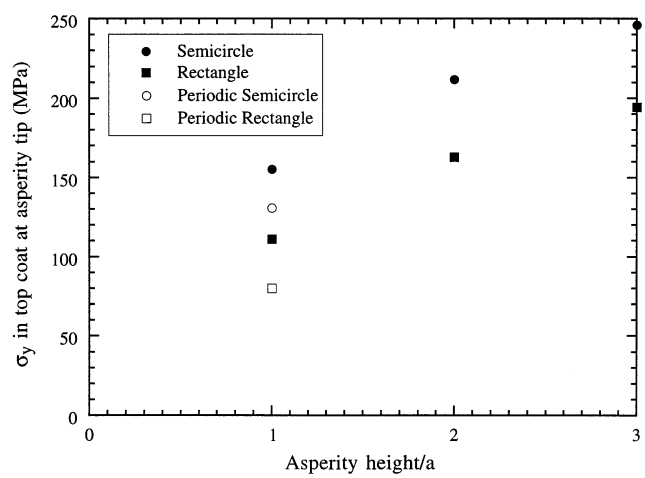


Fig. 3. Simulated results of σ_y within the top coat at the tip of the semicircular asperity and at the middle of the top of rectangular asperity as functions of the asperity height. The stress reduction for a periodic array of asperities is also shown at the asperity height of a.

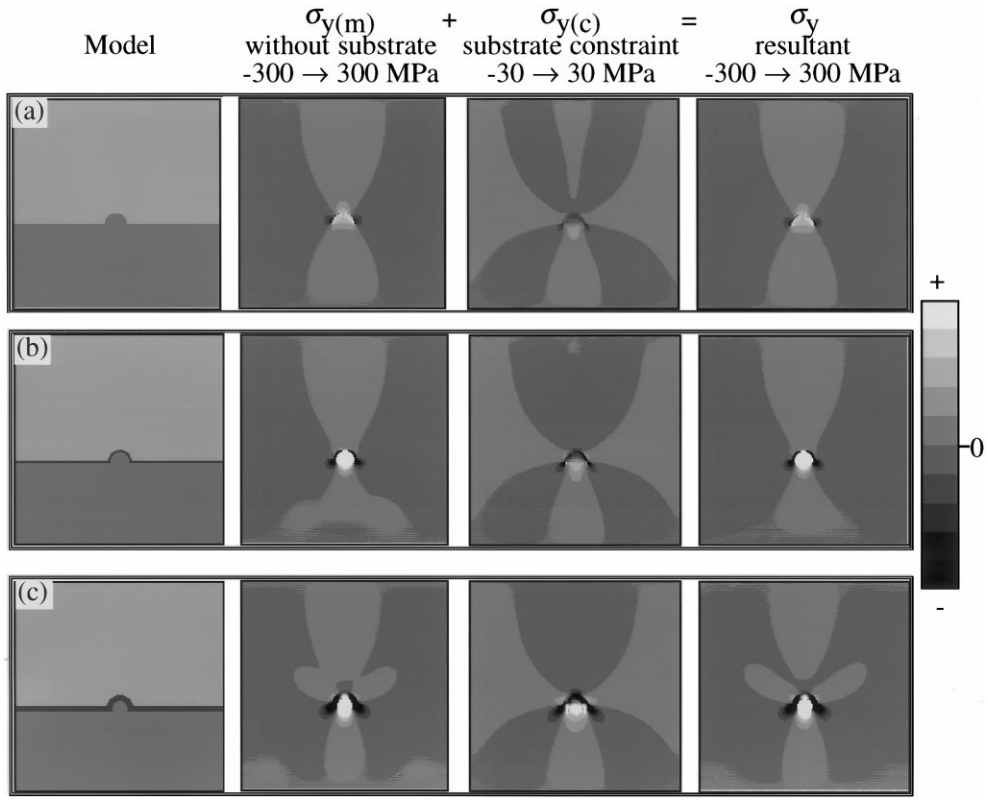


Fig. 4. The model microstructures and the simulated residual stresses, $\sigma_{y(m)}$, $\sigma_{y(c)}$, and σ_y, for the plasma-sprayed TBC system with a convex semicircular interfacial asperity having a radius, a, and a TGO thickness of (a) 0, (b) 0.23a, and (c) 0.72a.

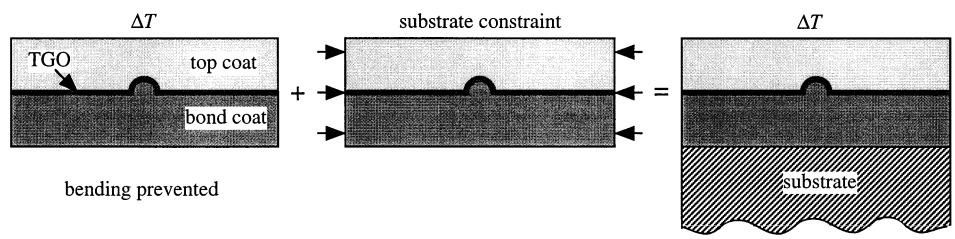


Fig. 5. Schematic showing the residual stress in the TBC system can be decomposed into two components: the one due to the thermal-mechanical mismatch among the top coat, the TGO, and the bond coat, and the one due to the substrate constraint imposed on the coating system.

radius a and the TGO has a variable thickness t. It is noted that oxidation generally induces a volume expansion. It can be readily incorporated in the OOF analysis by replacing the coefficient of thermal expansion (CTE) of the oxide scale by an effective CTE, such that the thermal strain due to the effective CTE equals the sum of the thermal strain due to the actual CTE and the oxidation strain. However, this was not done here because the oxidation strain for converting the bond coat to the oxide scale is not well documented.

2.2.1. Convex asperities

A convex semicircular asperity without TGO ($t/a = 0$), with a relatively thin TGO ($t/a = 0.23$), and with a relatively thick TGO ($t/a = 0.72$) are shown in Fig. 4(a–c), respectively. To explore the physical insights of the TGO thickness effects on residual stresses, it is instructive to decompose the residual stress, σ_y, into two components: (1) $\sigma_{y(m)}$, the one due to the thermal–mechanical mismatch among the top coat, the TGO, and the bond coat, and (2) $\sigma_{y(c)}$, the one due to the substrate constraint imposed on the coating system (see Fig. 5). Running OOF, these two components can be obtained using the following procedures.

First, using the procedures described in Section 2.1, the residual stress in the y-direction, σ_y, can be obtained. In this case, the thermal strain of the substrate, -0.0163, was imposed as the boundary condition at the right and the left edges of the image of the coating system to constrain its distortion in the x-direction. After equilibrating under $\Delta T = -1125$°C, the top coat, the TGO, and the bond coat are, respectively, subjected

to uniform elastic strains in the x-direction at the right and the left edges of the image. For example, these uniform elastic strains are -0.00502, -0.00727, and 0.00082, respectively, for the top coat, the TGO, and the bond coat when $t/a = 0.23$ (Fig. 4(b)). Second, the right and the left edges of the image are, respectively, set to have the same displacement in the x-direction (i.e. no bending in the coating system). This is achieved by using the 'enslave' boundary condition in the OOF menus. Using $\Delta T = -1125°C$, the residual stress obtained in the y-direction is $\sigma_{y(m)}$. In this case, unlike the above, the coating system is free of the substrate constraint; however, the top coat, the TGO, and the bond coat are, respectively, subjected to uniform elastic strains in the x-direction at the right and the left edges of the image because of the CTE mismatch. For example, these uniform elastic strains are -0.00426, -0.00651, and 0.00158, respectively, for the top coat, the TGO, and the bond coat when $t/a = 0.23$ (Fig. 4(b)). Third, the difference between the two sets of elastic stain is due to the substrate constraint effect. For example, the elastic stain in the x-direction due to substrate constraint is -0.00076 when $t/a = 0.23$ (Fig. 4(b)) which is then imposed as the boundary condition at the right and the left edges of the image to constrain its distortion in the x-direction. Using $\Delta T = 0$, the residual stress thus obtained in the y-direction is $\sigma_{y(c)}$.

Using the above procedures, the corresponding $\sigma_{y(m)}$, $\sigma_{y(c)}$, and σ_y are also shown in Fig. 4(a–c). It can be seen that the full-field residual stress of $\sigma_{y(m)}$ is similar to that of σ_y. Hence, compared to $\sigma_{y(c)}$, $\sigma_{y(m)}$ dominates the resultant σ_y (see Fig. 4). This is due to the fact that the substrate and the bond coat have the similar CTEs (see Table 1), and the substrate constraint effect imposed on the coating system is insignificant. However, the contribution of $\sigma_{y(c)}$ to the resultant σ_y increases with the increase in the TGO thickness. The simulated

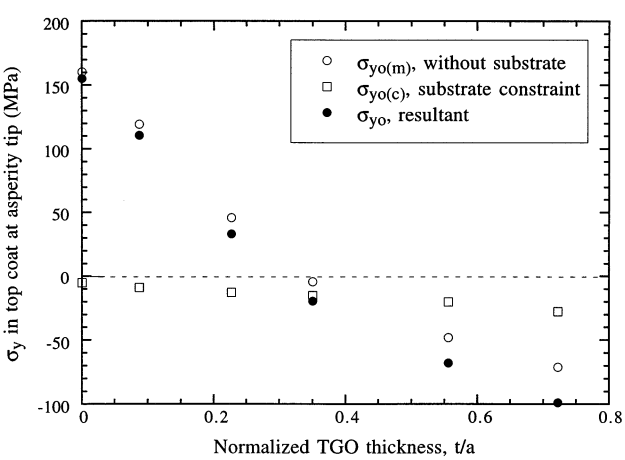


Fig. 6. The residual stresses within the top coat at the tip of a convex semicircular asperity, $\sigma_{yo(m)}$, $\sigma_{yo(c)}$, and σ_{yo}, as functions of the TGO thickness.

residual stresses within the top coat at the tip of the semicircular asperity (i.e. $\sigma_{yo(m)}$, $\sigma_{yo(c)}$, and σ_{yo}) are plotted in Fig. 6 as functions of the TGO thickness. Whereas $\sigma_{yo(m)}$ is tensile for thin TGOs and becomes compressive for thick TGOs, $\sigma_{yo(c)}$ is always compressive and its magnitude increases with the TGO thickness. Two general features can be observed in Fig. 4. First, the tensile stress state, $\sigma_{y(m)}$ (and σ_y), moves from the tip to the side of the asperity as the TGO grows. Second, the stress state, $\sigma_{y(m)}$, far above the asperity persists to be tensile even when σ_{yo} becomes compressive for thick TGOs. These two features are explained as follows.

First, the polar coordinates with its origin located at the center of the semicircular asperity can be used to facilitate visualizing the stress state in the top coat close to the asperity. The bond coat has a greater CTE than the top coat. Hence, in the absence of the TGO, the residual stress in the top coat around the asperity is tensile in the radial direction and compressive in the tangential direction. When σ_y is considered, it becomes the radial stress at the tip of the asperity and becomes the tangential stress at the side of the asperity. Hence, σ_y is tensile and compressive, respectively, at the tip and the side of the asperity. The TGO has the lowest CTE in the coating system. Hence, when the TGO is sufficiently thick, the residual stress in the top coat around the asperity becomes compressive in the radial direction and tensile in the tangential direction. This would cause σ_y to be compressive and tensile, respectively, at the tip and the side of the asperity. As a result, the tensile stress state, $\sigma_{y(m)}$ (and σ_y), moves from the tip to the side of the asperity as the TGO grows.

Second, the stress state, $\sigma_{y(m)}$, far above the asperity can be visualized by dividing the model microstructure into vertical thin strips. Compared to other strips, the strip passing the convex asperity has a greater bond coat area and a smaller top coat area. Since the bond coat has a greater CTE than the top coat, the strip passing the convex asperity has a greater effective CTE than other strips. As a result, in the strip passing the convex asperity, the far field residual stress, $\sigma_{y(m)}$, above the asperity is always tensile during cooling despite the fact that the near field stress $\sigma_{yo(m)}$ becomes compressive for thick TGOs.

2.2.2. Concave asperities

A concave semicircular asperity without TGO ($t/a = 0$), with a relatively thin TGO ($t/a = 0.24$), and with a relatively thick TGO ($t/a = 0.72$) are shown in Fig. 7(a–c), respectively. Like the residual stress, σ_y, in the case of convex semicircular asperity, σ_y in the case of concave asperity can also be decomposed into two components: $\sigma_{y(m)}$ and $\sigma_{y(c)}$. Using OOF and the procedures described in Section 2.2.1, the corresponding simulated results for $\sigma_{y(m)}$, $\sigma_{y(c)}$, and σ_y are also

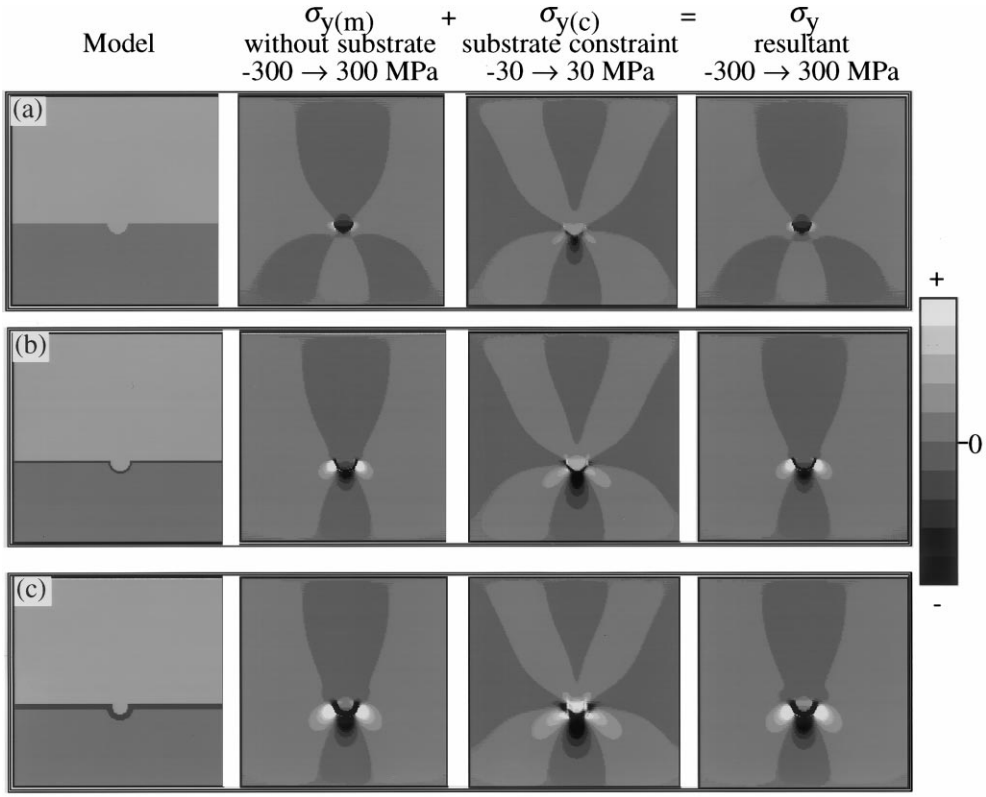


Fig. 7. The model microstructures and the simulated residual stresses, $\sigma_{y(m)}$, $\sigma_{y(c)}$, and σ_y, for the plasma-sprayed TBC system with a concave semicircular interfacial asperity having a radius, a, and a TGO thickness of (a) 0, (b) 0.24a, and (c) 0.72a.

shown in Fig. 7(a–c). The resultant σ_y is dominated by $\sigma_{y(m)}$ (see Fig. 7). The residual stresses within the top coat at the tip of the semicircular asperity (i.e. $\sigma_{yo(m)}$, $\sigma_{yo(c)}$, and σ_{yo}) are plotted in Fig. 8 as functions of the TGO thickness. For the range of the TGO thickness considered in the present study, $\sigma_{yo(m)}$ is always compressive and $\sigma_{yo(c)}$ is always tensile. As the TGO thickness increases, the magnitudes of $\sigma_{yo(m)}$ and $\sigma_{yo(c)}$ decreases and increases, respectively. It is noted that in the top coat, a tensile σ_y region above the asperity is observed for thick oxides (see σ_y in Fig. 7(c)) although σ_y remains compressive at the asperity tip. Specifically, this tensile σ_y reaches 100 MPa for the case considered in Fig. 7(c) which could result in cracking in the ceramic top coat.

2.3. Effects of interaction between asperities

To examine the effects of interaction between asperities on residual stresses, periodical arrays of convex and concave semicircular asperities with radius a (Fig. 9(a)) and of square asperities with length 2a and separation 2a (Fig. 9(b)) were considered. In this case, the asperity amplitude is a from the centerline of the interface for both arrays of asperities.

A uniform grid was also used to simulate the model microstructures in Fig. 9, and the processed images consisted of 64 082 elements each. The corresponding full-field residual stress, σ_y, are also shown in Fig. 9. The simulated results of σ_{yo} are 130.8 and −240.6 MPa, respectively, for the semicircular convex and concave asperities, and are 80 and −185 MPa, respectively, for the rectangular convex and concave asperities. These tensile stresses σ_{yo} (i.e. 130.8 and 80 MPa) for convex asperities are also shown in Fig. 3 at

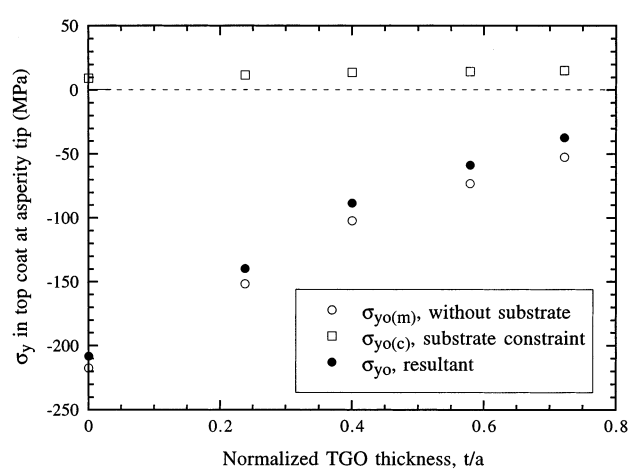


Fig. 8. The residual stresses within the top coat at the tip of a concave semicircular asperity, $\sigma_{yo(m)}$, $\sigma_{yo(c)}$, and σ_{yo}, as functions of the TGO thickness.

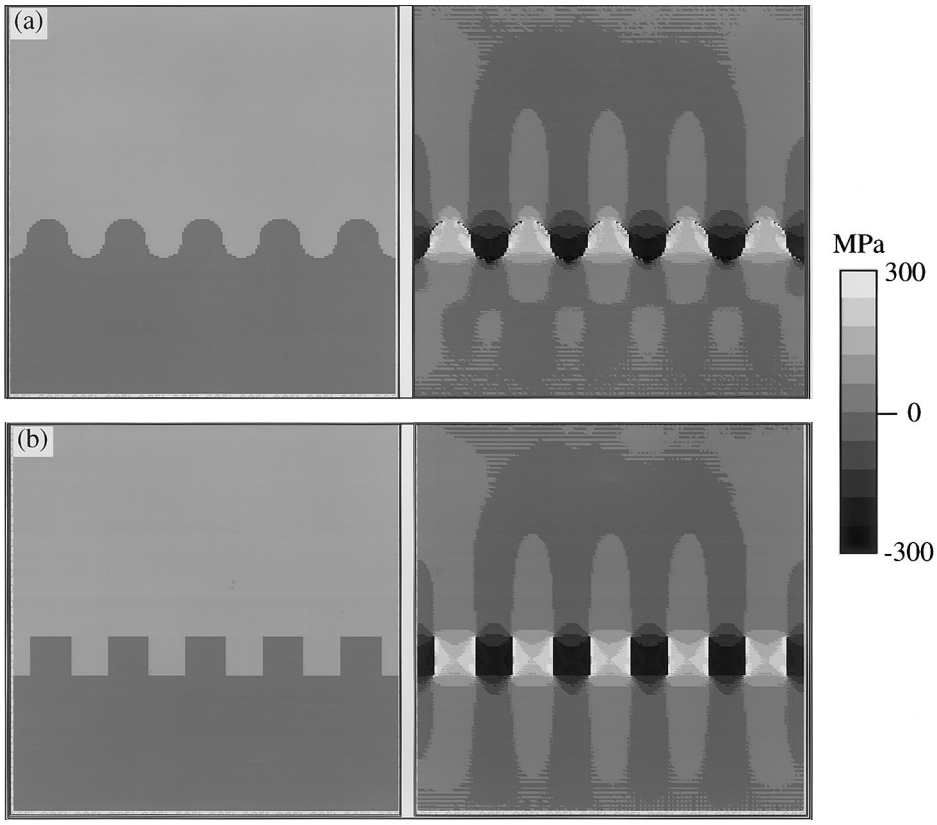


Fig. 9. The model microstructures and the simulated residual stress, σ_y, for the plasma-sprayed TBC system with a periodic array of (a) convex and concave semicircular asperities with radius a, and (b) square asperities with length $2a$ and separation 2a.

an asperity height value of a to compare with those due to an isolated asperity.

3. Analytical modeling and results

A two-concentric-circles model has been used previously to study the effects of the surface curvature of the substrate on oxidation behavior [7,17]. To explore the physical meaning of the numerically simulated (or experimental) results, a simple analytical model bearing the essential features of the simulated geometry (or specimen) is required. The essential features in the present case are the top coat–TGO interface, the TGO–bond coat interface, convex and concave asperities, and the TGO thickness. To include the above features, a simple three-concentric-circles model is adopted in the present study which is described as follows.

A circular region of phase 1 with a radius a is surrounded by a concentric annulus of phase 2 with an outer radius b which is then surrounded by another concentric annulus of phase 3 with an outer radius c (see Fig. 10). Whereas phases 1, 2, and 3 are designated respectively as the bond coat, the TGO, and the top coat in modeling a convex asperity, they are designated respectively as the top coat, the TGO, and the bond coat in modeling a concave asperity. In both cases, the two interfaces are located at $r = a$ and b, and the TGO thickness is dictated by $b - a$. It has been shown in Section 2.2 that the substrate constraint effect on the coating system is insignificant because the substrate and the bond coat have the similar CTEs. Hence, excluding the substrate from the three-concentric-circles model is appropriate in analyzing the residual thermal stresses in the TBC system.

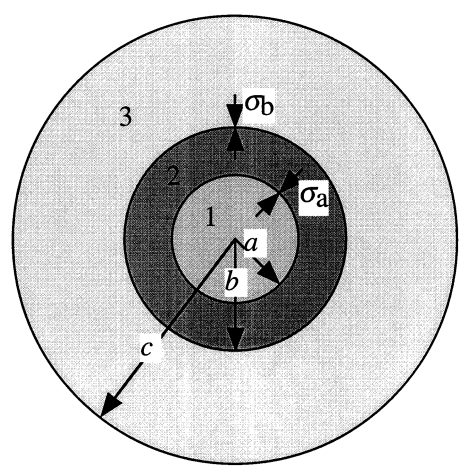


Fig. 10. A schematic showing the three-concentric-circles model.

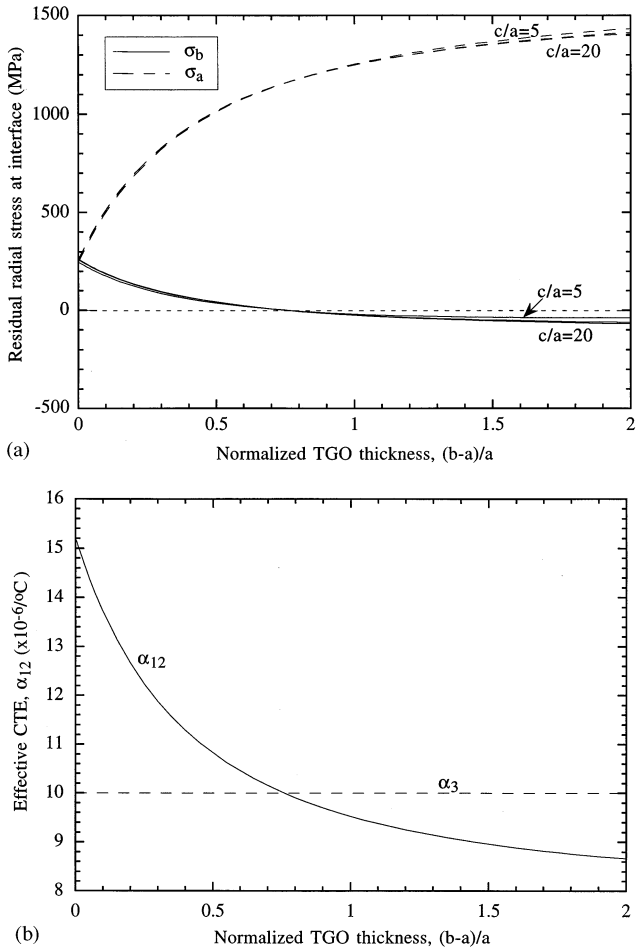


Fig. 11. (a) The residual radial stress at the TGO/bond coat interface, σ_a, and at the top coat–TGO interface, σ_b, and (b) the effective CTE, α_{12}, for the region $r < b$, and the CTE of top coat, α_3, as functions of the normalized TGO thickness, $(b-a)/a$, at different c/a ratios for a convex asperity.

3.1. Analyses

Residual thermal stresses develop during the temperature change in the system because of the thermal–mechanical mismatch among the three phases. The stresses can be determined by the procedure of first allowing the three phases to exhibit unconstrained differential thermal strains during the temperature change, ΔT. Then, radial stresses, σ_a and σ_b, are placed at the interfaces, $r = a$ and b, respectively, to restore the displacement continuity at the two interfaces. It can be derived that the interfacial radial stresses, σ_a and σ_b, are [23]

$$\sigma_a = \frac{P_1 E_2(\alpha_{23}\text{-}\alpha_1)\Delta T}{P_1 P_2\text{-}P_3 P_4} \tag{1}$$

$$\sigma_b = \frac{P_2 E_2(\alpha_3\text{-}\alpha_{12})\Delta T}{P_1 P_2\text{-}P_3 P_4} \tag{2}$$

where the parameters, P_1, P_2, P_3, and P_4, are given by

$$P_1 = \frac{b^2+a^2}{b^2-a^2} - \nu_2 + \frac{E_2}{E_3}\left(\frac{c^2+b^2}{c^2-b^2} + \nu_3\right) \tag{3a}$$

$$P_2 = \frac{b^2+a^2}{b^2-a^2} + \nu_2 + \frac{E_2(1-\nu_1)}{E_1} \tag{3b}$$

$$P_3 = \frac{2a^2}{b^2-a^2} \tag{3c}$$

$$P_4 = \frac{2b^2}{b^2-a^2} \tag{3d}$$

subscripts 1, 2, and 3 for the elastic constants denote phases 1, 2, and 3, respectively, and where α_1 and α_3 are the CTEs of phases 1 and 3, α_{12} has the physical meaning of the effective CTE in the region of $r < b$ which is a function of properties of phases 1 and 2, and α_{23} has the physical meaning of the effective CTE in the region of $a < r < c$ which is a function of properties of phases 2 and 3, such that

$$\alpha_{12} = \alpha_2 + \frac{P_3(\alpha_1 - \alpha_2)}{P_2} \tag{4}$$

$$\alpha_{23} = \alpha_2 + \frac{P_4(\alpha_3 - \alpha_2)}{P_1} \tag{5}$$

3.2. Results

The thermal–mechanical properties of the TBC system listed in Table 1 and $\Delta T = -1125$°C were used, and E was replaced by $E/(1-\nu)$ before it was substituted into the analytical solutions due to the biaxial stress state in the TBC system as discussed in section 2.

3.2.1. Convex asperities

For a convex asperity, phases 1, 2, and 3 correspond to the bond coat, the TGO, and the top coat, respectively. The residual radial stresses at $r = a$ and b (i.e. σ_a and σ_b) as functions of the normalized TGO thickness, $(b-a)/a$, are shown in Fig. 11(a) for $c/a = 5$, 10, and 20. In this case, σ_b corresponds to σ_y in the top coat at the tip of a convex asperity. The stresses are not sensitive to c/a when $c/a \geq 5$. The results for $c/a = 10$ almost overlap with those for $c/a = 20$, and can't be distinguished in Fig. 11(a). While σ_a is tensile and increases with the TGO thickness (which agrees with the trend in Fig. 4), σ_b is tensile when $(b-a)/a < \sim 0.75$ and becomes compressive when $(b-a)/a > \sim 0.75$. The effective CTE, α_{12}, as a function of the normalized TGO thickness, $(b-a)/a$, is shown in Fig. 11(b). It is noted that α_{12} is independent of c (see Eq. (4)). The CTE of phase 3, α_3 ($= 10 \times 10^{-6}$/°C), is also shown in Fig. 11(b). It can be seen that $\alpha_{12} > \alpha_3$ when $(b-a)/a < \sim 0.75$ which, in turn, results in a tensile residual radial stress at $r = b$ during cooling. Conversely, $\alpha_{12} < \alpha_3$ when $(b-a)/a > \sim 0.75$ which, in turn, results in a compressive residual radial stress at $r = b$ during cooling.

3.2.2. Concave asperities

For a concave asperity, phases 1, 2, and 3 correspond to the top coat, the TGO, and the bond coat, respectively. The residual radial stresses at $r=a$ and b as functions of the normalized TGO thickness, $(b-a)/a$, are shown in Fig. 12(a) at different c/a ratios. In this case, σ_a corresponds to σ_y in the top coat at the tip of a concave asperity. The stresses are not sensitive to c/a when $c/a \geq 10$. While σ_b is compressive and its magnitude increases with the TGO thickness for $c/a \geq 10$ (which agrees with the trend in Fig. 7), σ_a becomes less compressive with the increasing TGO thickness and the decreasing c/a for the dimension considered in Fig. 12(a). The effective CTE, α_{23}, as a function of the normalized TGO thickness, $(b-a)/a$, is shown in Fig. 12(b) at different c/a ratios. The CTE of phase 1, α_1 ($=10\times10^{-6}$/°C), is also shown. It can be seen that $\alpha_{23} > \alpha_1$ for the dimension considered in Fig. 12(b). However, the difference between α_{23} and α_1 decreases with the increasing TGO thickness and the decreasing c/a. Compared to Fig. 11(b), Fig. 12(b) shows that the decrease in α_{23} is much slower than the decrease in α_{12} when the TGO thickness increases. This is due to the fact that the bond coat for a concave asperity (i.e. phase 3 in Fig. 10) has a much greater area than that for a convex asperity (i.e. phase 1 in Fig. 10) in the three-concentric-circles model, and the effect of the TGO thickness on the effective CTE is much weaker for a concave asperity than for a convex asperity.

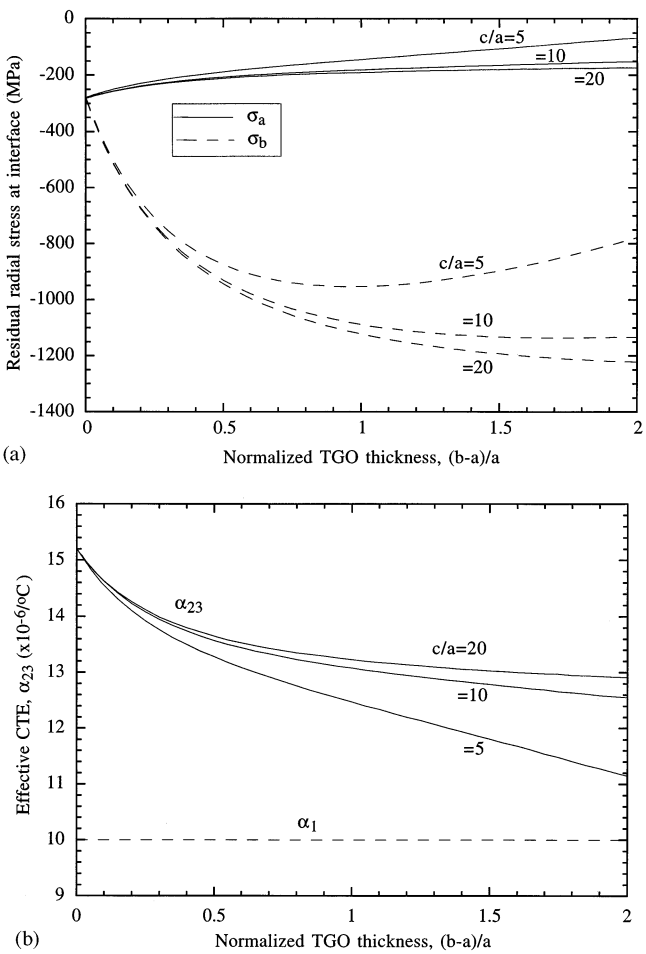


Fig. 12. (a) The residual radial stress at the top coat–TGO interface, σ_a, and at the TGO–bond coat interface, σ_b, and (b) the effective CTE, α_{23}, for the region $a<r<c$, and the CTE of top coat, α_1, as functions of the normalized TGO thickness, $(b-a)/a$, at different c/a ratios for a concave asperity.

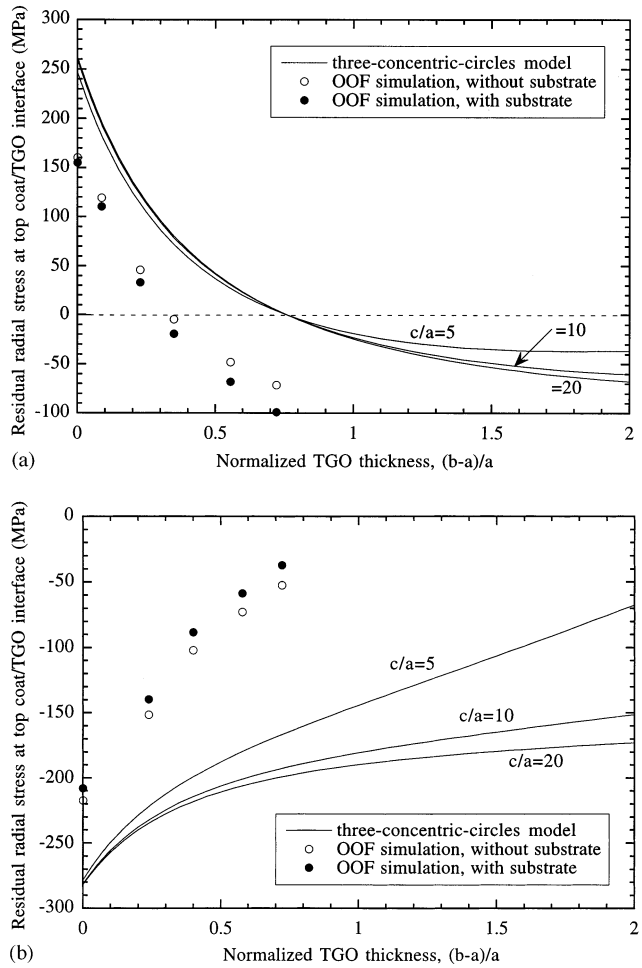


Fig. 13. Comparison between simulated results and analytical results for the residual stress normal to the interface within the top coat and at the tip of (a) a convex asperity, and (b) a concave asperity.

3.2.3. Comparison with simulated results

Despite the different geometries adopted in numerical simulations (Fig. 5) and analytical modeling (Fig. 10) for the TBC system, the predicted residual stresses normal to the interface in the top coat and at the asperity tip as functions of the normalized TGO thickness, $(b-a)/a$, are shown in Fig. 13(a–b), respectively, for convex and concave asperities. For numerical simulations, results for the coating system without substrate constraint and with substrate constraint are both shown. For the three-concentric-circles analytical

model, results for different c/a ratios are shown. It can be seen that the variations of these residual stresses with the TGO thickness are qualitatively the same between numerical simulations and analytical modeling.

4. Conclusions

The effects of (1) the curvature and the height of the interface asperity, (2) the thickness of the thermally grown oxide (TGO), and (3) interaction between asperities on residual thermal stresses in a plasma-sprayed TBC system were numerically simulated. The following qualitative trends were concluded from the present numerical simulations.

1. In the absence of TGO, the residual stress normal to the interface, σ_y, in the tip region of convex asperity is tensile within the top coat which increases with both the curvature and the height of the asperity (Figs. 2 and 3).

2. The residual stress, σ_y, can be decomposed into two components: (1) $\sigma_{y(m)}$, the one due to the thermal–mechanical mismatch among the top coat, the TGO, and the bond coat, and (2) $\sigma_{y(c)}$, the one due to the substrate constraint imposed on the coating system (Fig. 5). Whereas $\sigma_{y(m)}$ dominates, the contribution of $\sigma_{y(c)}$ to the resultant σ_y increases with the TGO thickness.

3. The residual stress, σ_y, at the asperity tip within the top coat is defined as σ_{yo}. For a convex asperity, σ_{yo} is tensile for thin TGOs but becomes compressive for thick TGOs (Fig. 6). The tensile stress state, σ_y, moves from the tip to the side of the asperity as the TGO grows (Fig. 4). Also, the stress state, σ_y, far above the asperity persists to be tensile even when σ_{yo} becomes compressive for thick TGOs (Fig. 4).

4. For a concave asperity, σ_{yo} in the ceramic top coat is compressive for thin TGOs and becomes less compressive for thick TGOs (Figs. 7 and 8). However, σ_y can becomes tensile above the asperity tip and within the top coat for thick TGOs (see σ_y in Fig. 7c).

5. In the presence of multiple asperities, σ_{yo} is reduced due to the interaction between asperities (Figs. 3 and 9).

The physical meaning of the trend of the variations of residual stresses, σ_y, with the TGO thickness was qualitatively interpreted using an analytical model of three concentric circles. Finally, it is noted that the present study aims at the qualitative trend of effects of interface asperities and TGO thickness on residual stresses, and the elastic solution is used. Including non linearity and creep in the finite element modeling, residual stresses in the TBC have been analyzed elsewhere [14].

Acknowledgements

The authors thank Drs P.F. Becher, P.F Tortorelli, M.J. Lance, T.J. Chuang, and E.J. Garboczi for reviewing the manuscript. CHH acknowledges jointly support of the U.S. Department of Energy, Division of Materials Sciences, Office of Basic Energy Sciences, and Assistant Secretary for Energy, Efficiency and Renewable Energy, Office of Industrial Technologies, Advanced Turbine Systems Program, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp. ERF acknowledges support of the Alexander von Humboldt Foundation.

References

[1] R.A. Miller, Surf. Coat. Technol. 30 (1) (1987) 1–11.
[2] W.Y. Lee, D.P. Stinton, C.C. Berndt, F. Erdogan, Y.D. Lee, Z. Mutasim, J. Am. Ceram. Soc. 79 (12) (1996) 3003–3012.
[3] J.A. Haynes, E.D. Rigney, M.K. Ferber, W.D. Porter, Surf. Coat. Technol. 86–87 (1996) 102–108.
[4] C.H. Hsueh, J.A. Haynes, M.J. Lance, P.F. Becher, M.K. Ferber, E.R. Fuller Jr, S.A. Langer, W.C. Carter, W.R. Cannon, J. Amer. Ceram. Soc. 82 (4) (1999) 1073–1075.
[5] D.J. Wortman, B.A. Nagaraj, E.C. Duderstadt, Mater. Sci. Eng. A121 (1989) 433–440.
[6] J.T. DeMasi-Marcin, K.D. Sheffler, S. Bose, J. Eng. Gas Turbines Power 112 (1990) 521–526.
[7] C.H. Hsueh, A.G. Evans, J. Appl. Phys. 54 (11) (1983) 6672–6686.
[8] A.G. Evans, G.B. Crumley, R.E. Demaray, Oxid. Metals 20 (5/6) (1983) 193–216.
[9] G.C. Chang, W. Phucharoen, R.A. Miller, Surf. Coat. Technol. 30 (1987) 13–28.
[10] H. Gao, Int. J. Solids Structures 28 (6) (1991) 703–725.
[11] V.K. Tolpygo, H.J. Grabke, Oxid. Metals 41 (5/6) (1994) 343–364.
[12] J.K. Wright, R.L. Williamson, R.M. Cannon, Mater. Sci. Eng. A230 (1997) 202–212.
[13] M.Y. He, A.G. Evans, J.W. Hutchinson, Mater. Sci. Eng. A245 (1998) 168–181.
[14] A.M. Freborg, B.L. Ferguson, W.J. Brindley, G.J. Petrus, Mater. Sci. Eng. A245 (1998) 182–190.
[15] J. Chao, J.L. González-Carrasco, Mater. Sci. Eng. A230 (1997) 39–48.
[16] J. Cheng, E.H. Jordan, B. Barber, M. Gell, Acta Mater 46 (16) (1998) 5839–5850.
[17] X.Y. Gong, D.R. Clarke, Oxid. Metals 50 (5/6) (1998) 355–377.
[18] B.A. Pint, I.G. Wright, W.Y. Lee, Y. Zhang, K. Prü(ner, K.B. Alexander, Mater. Sci. Eng. A245 (1998) 201–211.
[19] S.A. Langer, W.C. Carter, E.R. Fuller, Object Oriented Finite Element Analysis for Materials Science: Version 1.0, National Institute of Standards and Technology, Nov. 1998.
[20] C.H. Hsueh, P.F. Becher, E.R. Fuller Jr, S.A. Langer, W.C. Carter, Materials Science Forum, Vols. 308–311 (1999) 442–449.
[21] R.V. Hillery, B.H. Pilsner, R.L. McKnight, T.S. Cook, M.S. Hartle, Thermal barrier coating life prediction model development, NASA Contractor Report 180807 (1988).
[22] R.J. Christensen, V.K. Tolpygo, D.R. Clarke, Acta Mater. 45 (4) (1997) 1761–1766.
[23] C.H. Hsueh, E.R. Fuller, Jr., Scripta Materialia, in press.

Analyses of edge effects on residual stresses in film strip/substrate systems

Chun-Hway Hsueh[a)]
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

(Received 17 April 2000; accepted for publication 9 June 2000)

The residual stress distribution in a thin-film strip overlaid on a substrate is influenced by the edges of the strip. An analytical model is developed to derive a closed-form solution for the stress distribution along the film width. Because the film is much thinner than the substrate, the stress variation through the film thickness is ignored; however, the stress variation through the substrate thickness is considered in the analysis. Compared to the existing analytical models, the present model is more rigorous and the analytical results agree better with both finite element results and experimental measurements. © *2000 American Institute of Physics.* [S0021-8979(00)02218-0]

I. INTRODUCTION

Thin-film strips on a substrate have extensive applications in microelectronic devices.[1–5] The film/substrate system is subjected to residual stresses, which result from the film deposition process, the lattice mismatch, and the thermal expansion mismatch. These stresses provide a driving force to form defects which, in turn, degrade the device performance. If the film has an infinite lateral dimension, the stress in the film would be uniform when the substrate is much thicker than the film and bending of the system is ignored. However, a finite film strip has free surfaces at the edges which, in turn, modify the stress distribution in the film/substrate system. There have been many studies to analyze the film/substrate stresses arising from the edges of an otherwise uniformly stressed film.[6,7] Both analytical modeling[6,8–12] and the finite element analysis (FEA)[7,13–15] have been performed to predict the film-edge effects. However, by comparing the analytical solutions with finite elements results and experimental measurements, it is concluded that predictions for the film-edge effects from the existing analytical models are not satisfactory.[7,15] Also, numerical procedures are required in some of the existing analytical models in order to solve some complex equations.

The purpose of this study is to develop a better analytical model with a closed-form analytical solution for the film-edge effects. Unlike the FEA which requires a finite element package and extensive numerical computations, the closed-form analytical solution is ready to be used to examine the roles of the essential material parameters in the system. First, the existing analyses are reviewed. Then, an analytical model is developed to derive a closed-form solution to predict the stress distribution in the film strip/substrate system. Finally, comparison is made between the present analytical results, existing analytical/numerical results,[7–11,15] and experimental measurements.[16–19]

[a)]Electronic mail: hsuehc@ornl.gov

II. SUMMARY OF EXISTING ANALYSES

Both analytical modeling[6,8–12] and the FEA[7,13–15] have been performed to analyze the film-edge effects, and they are summarized as follows.

A. Analytical modeling

Analytical solutions for the stress distribution in a film strip similar to that shown in Fig. 1 have been previously derived. A film strip with a width $2l$ and a thickness t is overlaid on a substrate with a thickness s. The film is subjected to a mismatch strain, $\Delta\varepsilon$. The Cartesian coordinates, x, y, and z, are shown in Fig. 1, such that the film is bonded to the substrate at $x=0$, and the edges of the film are located at $y=\pm l$. The dimension of the film/substrate system along the z axis is assumed to be infinitely long. Among the existing analytical models, the most cited ones are developed by Hu[8,9] and Suhir.[10,11] However, in all analytical models, the variation of the stress through the film thickness is neglected because the film is much thinner than the substrate. Other approximations are also used to simplify the problem in order to obtain analytical solutions.

In Hu's analysis,[8,9] an infinitely thick substrate (i.e., $s\rightarrow\infty$) is considered and the solution is derived for a single-edged film having a semi-infinite width. A concentrated force model was derived first, such that a concentrated body force (per unit length in the z direction), F, is induced at the film edge and acts tangentially on the substrate.[8] In this case, the stress distribution in the substrate can be obtained readily from Timoshenko and Goodier's solutions,[20] in which the stress distribution in a half space subjected to a concentrated force tangential to the surface is analyzed. The solution obtained from the concentrated force model lacks self-consistency because the film stress becomes distributed as it relaxes with the substrate deformation. Consequently, the concentrated force model[8] was modified to a distributed force model,[9] in which the change in film stress is balanced by a distributed tangential force on the substrate. However, for the distributed force model, the formulation for the stress distribution in the system is an integral equation and a numerical method is required to obtain the solutions.[9]

 © 2000 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

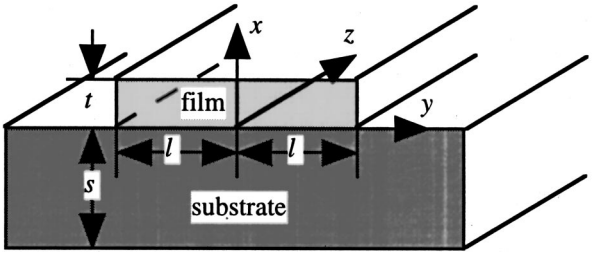


FIG. 1. Schematic showing a thin-film strip deposited on a substrate.

When the film becomes a strip, Hu's solutions for a single-edged film and the principle of superposition are used to obtain the solutions for a two-edged film.[2,12] It has been noted that the stress distribution in the film obtained by Hu for a single-edged film satisfies the free surface condition at the edge. However, the stress distribution in the film obtained by using Hu's solutions and the principle of superposition for a two-edged film does not satisfy the free surface condition at the edges.[7] A simple analytical solution for the stress distribution in the substrate can be obtained when Timoshenko and Goodier's solutions (i.e., the concentrated force model)[20] and the principle of superposition are used, such that the stress component in the y direction in the substrate, σ_y^s, for the geometry described by Fig. 1 is given by[2,7]

$$\sigma_y^s = \frac{-2F}{\pi}\left\{\frac{(y+l)^3}{[(y+l)^2+x^2]^2} - \frac{(y-l)^3}{[(y-l)^2+x^2]^2}\right\}. \tag{1}$$

Suhir's analysis is a modification of Timoshenko's model[21] on analyzing the residual stresses in a bi-metal thermostat, in which the width is infinite (i.e., $l \to \infty$). By considering a finite film width, $2l$, Timoshenko's model is modified by Suhir. Because of the mismatch, $\Delta\varepsilon$, between the film, the substrate, and the edge effect, a nonuniform force along the y axis (per unit length in the z direction) $P(y)$ is induced in the film and a force of the same magnitude but of the opposite direction is induced in the substrate. These forces can be regarded as concentrated forces at the half thicknesses of the film (i.e., at $x=t/2$) and the substrate (i.e., at $x=-s/2$), respectively. They can also be treated as forces distributed uniformly through the thicknesses of the film and the substrate. A shear force $q(y)$ is induced at the interface such that $q(y)$ and $P(y)$ are related by an equilibrium equation. Bending of the system is considered. Then, the displacements of the film and the substrate at the interface are formulated, respectively, based on the mismatch strain, the two force distributions, and the curvature of the system. It is noted that a Fourier series is adopted in formulating the displacement due to the interfacial shear force, and t/l is assumed to be small (e.g., $t/l<0.2$) in order to simplify the equations.[10] By satisfying the displacement compatibility condition at the interface and the boundary conditions that $P(l)=0$ and $q(0)=0$, the solution for $P(y)$ [and hence $q(y)$] can be obtained, such that[11]

$$P(y) = \frac{-tE_f\Delta\varepsilon}{1-\nu_f}\left[1-\frac{\cosh(\kappa y)}{\cosh(\kappa l)}\right]. \tag{2}$$

The parameter, κ, is defined by

$$\kappa = \left[\frac{\dfrac{1-\nu_f}{tE_f}+\dfrac{1-\nu_s}{sE_s}+\dfrac{(t+s)^2}{\dfrac{t^3E_f}{1-\nu_f}+\dfrac{s^3E_s}{1-\nu_s}}}{\dfrac{2t(1+\nu_f)}{3E_f}+\dfrac{2s(1+\nu_s)}{3E_s}}\right]^{1/2}, \tag{3}$$

where E and ν are Young's modulus and Poisson's ratio, and the subscripts, f and s, denote the film and the substrate, respectively. Because the film is relatively thin, the bending moment in the film (proportional to t^3) is small, and the stress in the film, $\sigma(y)$, is due to the force $P(y)$ only, such that

$$\sigma(y) = \frac{-E_f\Delta\varepsilon}{1-\nu_f}\left[1-\frac{\cosh(\kappa y)}{\cosh(\kappa l)}\right]. \tag{4}$$

Unlike the stress in the film, the stress in the substrate is due to both the force and the bending moment. While the stress component due to the force, P, is uniform through the substrate thickness, the stress component due to the bending moment varies through the substrate thickness. However, when the thickness of the substrate is infinite, the bending moment becomes zero, and the stress in the substrate becomes zero although the force is nonzero.

B. Finite element analyses

The stress distribution in the geometry depicted by Fig. 1 has been analyzed using the FEA.[7,14,15] The tractions on all free surfaces are zero except at the bottom of the substrate (i.e., at $x=-s$) which is assumed to be rigid (i.e., zero displacement). An initially uniform mismatch strain, $\Delta\varepsilon$, is assigned to the film. Then, this mismatch strain is allowed to relax, and a numerical code is used to find the minimum energy configuration for the whole system. It is noted that singularities exist at the points where the film edges meet the substrate, and a refined finite element mesh is used in the region close to the edges.

III. THE PRESENT ANALYTICAL MODEL

The present analytical model preserves the essence of Suhir's analysis but adopts different analytical procedures to derive the stress distribution in the film/substrate system. Also, since the substrate is much thicker than the film, the variation of the stress through the thickness of the substrate is included in the analysis; however, bending of the system is ignored. Incidentally, the present model can also be regarded as a shear lag model in a planar geometry. The shear lag model was first developed by Cox to analyze the stress transfer problem for a cylindrical geometry (i.e., a fiber embedded in a coaxial cylindrical matrix).[22] It has then been modified and improved by many others.[23–27] The analytical procedures in deriving the present model are as follows.

A. Condition for a bonded interface

The film and the substrate are bonded at the interface, and the film is subjected to a mismatch strain, $\Delta\varepsilon$. Without the constraint in the x direction (see Fig. 1), the stress normal to the interface is zero, and the displacement compatibility condition at the interface requires

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

$$\frac{1}{E_f}(\sigma_y^f-\nu_f\sigma_z^f)+\Delta\varepsilon=\frac{1}{E_s}(\sigma_y^s-\nu_s\sigma_z^s)\quad(\text{at } x=0), \tag{5a}$$

$$\frac{1}{E_f}(\sigma_z^f-\nu_f\sigma_y^f)+\Delta\varepsilon=\frac{1}{E_s}(\sigma_z^s-\nu_s\sigma_y^s)\quad(\text{at } x=0), \tag{5b}$$

where σ_y and σ_z are the stresses in the y and the z directions, and the superscripts, f and s, denote the film and the substrate, respectively. When the substrate is much thicker than the film, bending of the system can be ignored. Furthermore, when the system has an infinite length in the z direction, σ_z^f is uniform in the film and σ_z^s is zero in the substrate. With $\sigma_z^s=0$, the combination of Eqs. (5a) and (5b) yields

$$\sigma_y^s=\frac{(1-\nu_f{}^2)E_s}{(1-\nu_f\nu_s)E_f}\sigma_y^f+\frac{(1+\nu_f)E_s}{1-\nu_f\nu_s}\Delta\varepsilon\quad(\text{at } x=0). \tag{6}$$

If the system is also infinitely long in the y direction, σ_y^f is uniform in the film, σ_y^s is zero in the substrate, and the solution of σ_y^f from Eq. (6) is

$$\sigma_y^f=\sigma_{\text{fb}}=\frac{-E_f}{1-\nu_f}\Delta\varepsilon\quad(\text{for } 0\leqslant x\leqslant t). \tag{7}$$

Hence, for the special case that the substrate is much thicker than the film and the system is infinitely long in both the y and the z directions, the film is subjected to a biaxial stress ($\sigma_y^f=\sigma_z^f$) described by Eq. (7) and the substrate is stress-free. For convenience, this biaxial stress in the film is redefined as σ_{fb}. However, when the film has a finite width in the y direction (e.g., a strip), the displacement compatibility condition at the interface is described by Eq. (6).

B. Stress gradient in the film

When the system has an infinite length in the z direction, there is no stress variation in the z direction. The equilibrium equation between the normal stress, σ_y, and the shear stress, $\tau_{xy}(=\tau)$, is

$$\frac{\partial\sigma_y}{\partial y}+\frac{\partial\tau}{\partial x}=0. \tag{8}$$

Integration of Eq. (8) over the thickness of the film and then division by its thickness yields

$$\frac{\partial}{\partial y}\left(\frac{1}{t}\int_0^t\sigma_y dx\right)+\frac{1}{t}\int_0^t\frac{\partial\tau}{\partial x}dx=0. \tag{9}$$

Because the film is relatively thin, the variation of σ_y through the thickness can be ignored. Letting σ_f represent the average value of σ_y over the thickness of the film and using the free surface condition that $\tau=0$ at $x=t$, Eq. (9) becomes

$$\frac{d\sigma_f}{dy}=\frac{\tau_0}{t}, \tag{10}$$

where τ_0 is the shear stress at the interface (i.e., $\tau=\tau_0$ at $x=0$). The stress gradient in the film is described by Eq. (10).

C. Stress distribution in the substrate

To strictly satisfy the equilibrium equation, the stress–displacement relation, and the essential boundary conditions, solutions for the stress distribution require extensive numerical analysis. However, simple analytical solutions are attainable when some of the above equations are approximately satisfied and the rest of the equations are exactly satisfied. Because the main interest in the system is the variation of σ_y in the y direction, the variation of σ_y in the x direction is ignored in the equilibrium equation such that the gradient of σ_y in the y direction in the substrate can be approximated by a function of y; i.e.,

$$\frac{\partial\sigma_y}{\partial y}=f(y)\quad(\text{for } -s\leqslant x\leqslant 0). \tag{11}$$

The physical meaning of this approximation [i.e., Eq. (11)] in the equilibrium equation [Eq. (8)] is that the equilibrium equation is satisfied in an average sense (with respect to the x direction). A similar approach has been adopted in the shear lag model in analyzing the stress transfer problem in a cylindrical geometry. The approximate x dependence of σ_y can be obtained by solving the approximate equilibrium equation and is shown as follows.

Continuity of the shear stress at the interface (i.e., $\tau=\tau_0$ at $x=0$) is required, and solution of Eqs. (8) and (11) subjected to the continuity condition gives

$$\tau=\left(1+\frac{x}{s}\right)\tau_0-\frac{x}{s}\tau_s\quad(\text{for } -s\leqslant x\leqslant 0), \tag{12}$$

where τ_s is the shear stress at $x=-s$.

With $\sigma_x=0$ and the displacement in the x direction being negligible, τ can be related to the displacement in the y direction, w, by

$$\tau=\frac{E_s}{2(1+\nu_s)}\frac{dw}{dx}\quad(\text{for } -s\leqslant x\leqslant 0). \tag{13}$$

Combination of Eqs. (12) and (13) and integration yields

$$\tau_0=\frac{E_s(w_0-w_s)}{s(1+\nu_s)}-\tau_s, \tag{14a}$$

$$\tau=\left(1+\frac{x}{s}\right)\frac{E_s(w_0-w_s)}{s(1+\nu_s)}-\left(1+\frac{2x}{s}\right)\tau_s\quad(\text{for } -s\leqslant x\leqslant 0), \tag{14b}$$

where w_0 and w_s are the displacements in the y direction at $x=0$ and $-s$, respectively. When the substrate is fixed at $x=-s$, $w_s=0$ and the displacement, w, can be derived from Eqs. (13) and (14b) such that

$$w=\left(1+\frac{x}{s}\right)^2 w_0-\frac{2(1+\nu_s)}{E_s}\left(x+\frac{x^2}{s}\right)\tau_s\quad(\text{for } -s\leqslant x\leqslant 0). \tag{15}$$

Using the conditions that $\sigma_x=\sigma_z=0$ in the substrate, differentiation of Eq. (15) with respect to y and then multiplication by E_s gives

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

$$\sigma_y=\left(1+\frac{x}{s}\right)^2\sigma_0-2(1+\nu_s)\left(x+\frac{x^2}{s}\right)\frac{d\tau_s}{dy}$$
$$(\text{for } -s\leqslant x\leqslant 0), \tag{16}$$

where σ_0 is the substrate stress in the y direction at $x=0$. The approximate x dependence of σ_y in the substrate is described by Eq. (16).

The mechanical equilibrium condition for the stress distribution in the system requires that

$$t\sigma_f+\int_{-s}^{0}\sigma_y dx=0. \tag{17}$$

Solution of Eqs. (16) and (17) gives

$$\frac{d\tau_s}{dy}=\frac{-3t}{s^2(1+\nu_s)}\left(\sigma_f+\frac{s}{3t}\sigma_0\right). \tag{18}$$

The aforementioned equation is required in deriving the governing differential equation for σ_f.

D. Stress distribution in the film

Combination of Eqs. (10), (14a), and (18) and differentiation with respect to y yields

$$\frac{d^2\sigma_f}{dy^2}=\frac{2}{st(1+\nu_s)}\left(\frac{3t}{2s}\sigma_f+\sigma_0\right). \tag{19}$$

The solution of the stress distribution σ_f is contingent upon the determination of σ_0. For a bonded interface, σ_0 is related to σ_f by Eq. (6), where σ_y^s at $x=0$ becomes σ_0 and σ_y^f at $x=0$ becomes σ_f when the variation of σ_y^f through the film thickness is ignored. Hence, the differential equation governing the stress distribution in the film can be obtained by substituting Eq. (6) into Eq. (19), such that

$$\frac{d^2\sigma_f}{dy^2}=\frac{2}{st(1+\nu_s)}\left[\left(\frac{3t}{2s}+\frac{(1-\nu_f^2)E_s}{(1-\nu_f\nu_s)E_f}\right)\sigma_f-\frac{(1-\nu_f^2)E_s}{(1-\nu_f\nu_s)E_f}\sigma_{\text{fb}}\right], \tag{20}$$

where σ_{fb} is defined by Eq. (7).

The solution of σ_f from Eq. (20) is subject to the boundary conditions that the film is stress-free at the edge and the system is symmetric with respect to $y=0$; i.e.,

$$\sigma_f=0 \quad (\text{at } y=l), \tag{21a}$$

$$\tau_0=0 \quad (\text{at } y=0). \tag{21b}$$

With the above two boundary conditions, the solution of σ_f is

$$\sigma_f(y)=\frac{\sigma_{\text{fb}}}{1+\dfrac{3t(1-\nu_f\nu_s)E_f}{2s(1-\nu_f^2)E_s}}\left[1-\frac{\cosh(\alpha y)}{\cosh(\alpha l)}\right], \tag{22}$$

where

$$\alpha=\left[\frac{2}{st(1+\nu_s)}\left(\frac{3t}{2s}+\frac{(1-\nu_f^2)E_s}{(1-\nu_f\nu_s)E_f}\right)\right]^{1/2}. \tag{23}$$

It is noted that the present solution for the film stress distribution [Eq. (22)] has a similar formulation as that in Suhir's model [Eq. (4)].

The maximum stress within the film occurs at the center of the film (i.e., at $y=0$), such that

$$\sigma_f(0)=\frac{\sigma_{\text{fb}}}{1+\dfrac{3t}{2s}\dfrac{(1-\nu_f\nu_s)E_f}{(1-\nu_f^2)E_s}}\left[1-\frac{1}{\cosh(\alpha l)}\right]. \tag{24}$$

When the film strip is sufficiently wide such that $\cosh(\alpha l)\rightarrow\infty$, this maximum stress has a constant value: i.e.,

$$\sigma_f(0)=\frac{\sigma_{\text{fb}}}{1+\dfrac{3t}{2s}\dfrac{(1-\nu_f\nu_s)E_f}{(1-\nu_f^2)E_s}} \quad (\text{for } l\gg 0). \tag{25}$$

Furthermore, when the substrate is much thicker than the film, this constant maximum stress becomes

$$\sigma_f(0)=\sigma_{\text{fb}} \quad (\text{for } l\gg 0 \text{ and } s\gg t). \tag{26}$$

IV. RESULTS

The FEA has been performed to predict the stress distribution in GeSi film strips on silicon substrates.[7,14,15] The material properties of this system are: $E_f=147.2$ GPa, $E_s=163.1$ GPa. $v_f=0.2765$, and $v_s=0.28$. The thickness and the width of the substrate are 50 and 100 units respectively, and different thicknesses and widths are used for the film in the FEA. The width of the substrate is not involved in all the analytical solutions. Unless noted otherwise, the above material properties and a thickness of 50 units for the substrate are used to calculate the results from the analytical models. First, comparison of the stress distribution along the film width is made between the present analytical results, Suhir's analytical results, and finite element results. Second, the stress at the center of the film (i.e., at $y=0$) is calculated and compared with the results obtained from other models and experimental measurements. Third, comparison of the stress distribution in the substrate is made between the results obtained from the present analytical model, Hu's model, and the FEA. Finally, the effects of the Young's modulus ratio of substrate to film and the aspect ratio of the film on the stress distribution in the film are shown. Because the geometry of the system is symmetric with respect to $y=0$, only the solutions for $y\geqslant 0$ are shown in the following.

A. Stress distribution along the film width

Based on the finite element results, the stress is not uniform throughout the thickness of the film. The film stress is largest at the interface and decreases with the increasing distance from the substrate. However, the stress at the half thickness of the film (i.e., at $x=t/2$) is a good approximation of the average film stress.[28] Since the analytical solutions give only the average stress throughout the film thickness, the stresses at $x=t/2$ from the finite element results are chosen to compare with the analytical results. Using the film thickness $t=1$ unit and the half width $l=20$ units, the present analytical results, Suhir's analytical results, and the finite elements results are shown in Fig. 2. Compared to

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

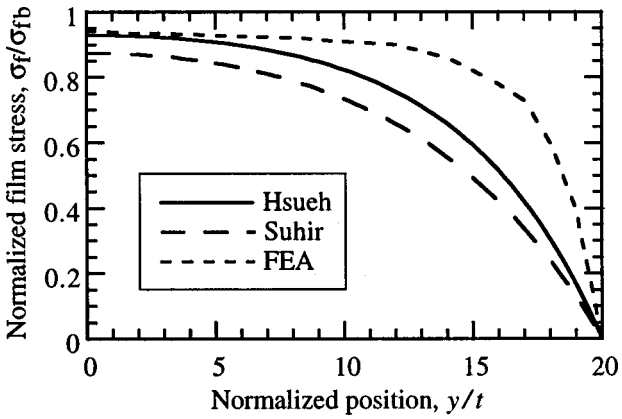


FIG. 2. Comparison of variations of normalized film stress, σ_f/σ_{fb}, along the film width between the results obtained from the present model, Suhir's model, and the FEA for $s=50$ units, $t=1$ units, and $l=20$ units.

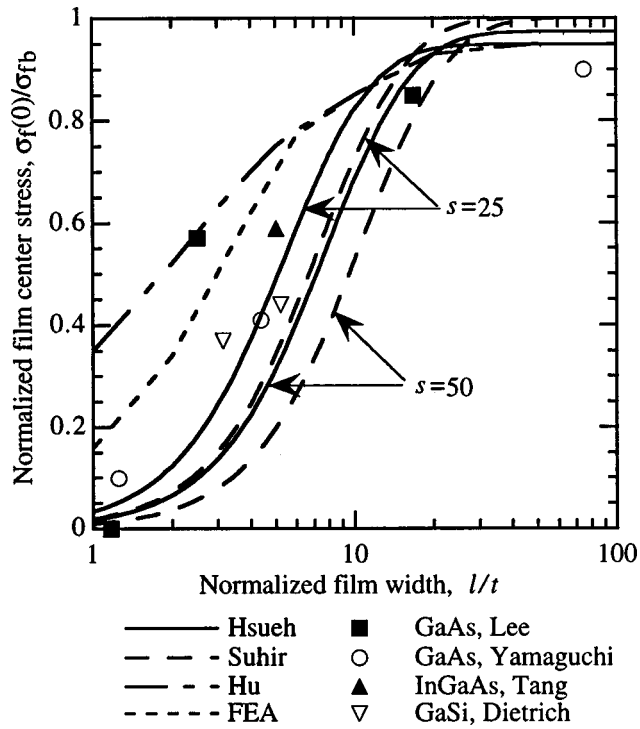


FIG. 3. Comparison of the normalized stress at the center of the film, $\sigma_f(0)/\sigma_{fb}$, as a function of l/t between the results obtained from the present model, Suhir's model, Hu's model, the FEA, and experimental measurements.

Suhir's results, the present results are closer to the finite elements results. Suhir's model predicts a lower stress in the film. This discrepancy could result from the different boundary conditions adopted in the modeling. While the substrate is fixed at $x=-s$ in the present analytical model and the finite element model, the system is free to bend in Suhir's model which, in turn, results in stress relaxation. However, when the substrate is much thicker than the film (e.g., when $s/t=50$ in Fig. 2), the difference in the predicted results caused by the different boundary condition assigned at $x=-s$ is negligible. Hence, the lower stresses obtained in Suhir's model are believed to be due to the assumption of the force $P(y)$ being uniformly distributed through the substrate thickness.

B. Stress at the center of the film

Stress measurements have been made on several semiconductor films deposited on Si substrates. Using cathodoluminescence, the stresses in GaAs films were measured by Lee (with $t=3$ μm, $2l=7$, 15, and 100 μm)[16] and Yamaguchi (with $t=4$ μm, $2l=10$, 35, and 600 μm),[17] respectively. Using micro-Raman, the stress in a $In_{0.25}Ga_{0.75}As$ film (with $t=0.1$ μm, $2l=1$ μm) was measured by Tang,[18] and the stresses in $Ge_{0.14}Si_{0.86}$ films (with $t=96$ nm, $2l=0.6$ μm and 1 μm) were measured by Dietrich.[19] It is noted that a sharp film edge is difficult to obtain in film processing due to the formation of ''bird's beak,''[2,7] which distorts the film edge. This complicates the stress measurement at the film edges; nevertheless, the stresses have been measured at the film center (i.e., at $y=0$). However, because of the finite beam spot diameter (e.g., 1 μm for cathodoluminescence and 0.7 μm for micro-Raman), the measured stress is not just at the center but some average around the center.

The FEA[7] has been performed to predict the stress at the center of the film as a function of l/t to compare with the measurements. Not all the films involved in the stress measurements are GeSi. However, the elastic constants are neither sensitive to small changes in Ge concentration for GeSi films nor are they very different from GaAS and InGaAs films,[7] and the material properties of GeSi film are used to calculate the results from modeling. In Hu's model, the substrate thickness is assumed to be infinite, and the ratio, l/t, is sufficient to determine the solutions. However, the parameters, s, t, and l, are involved in both the present model and Suhir's model. Using $s=50$ units, $t=1$ unit, and different units for l, the results from the present model and Suhir's model are calculated for comparison. The normalized stress at the center of the film, $\sigma_f(0)/\sigma_{fb}$, as a function of l/t is shown in Fig. 3. Higher stresses are predicted based on Hu's model. Compared to Suhir's model, the present model predicts results closer to the finite element results. While the stresses predicted from the present model are lower than those predicted from FEA, the present predicted results qualitatively agree with the measurements.

It has been argued that for the geometry described by Fig. 1, the thin film influences the stress field in the substrate only within a certain boundary zone in the neighborhood of the film. Outside the boundary zone, the stress field in the substrate is not perturbed by the presence of the film.[29] Hence, an effective thickness should be defined, such that this effective thickness instead of the actual substrate thickness should be used in analytical modeling. However, a criterion for defining the effective substrate thickness has not been established. Based on the finite element results for the film thickness being one unit, the perturbation of the stress field in the substrate is negligible when the position in the substrate is >25 units away from the interface.[7] Using 25 units instead of 50 units as the substrate thickness, the results from the present model and Suhir's model are also shown in Fig. 3 and they agree better with both the finite element results and the measurements.

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

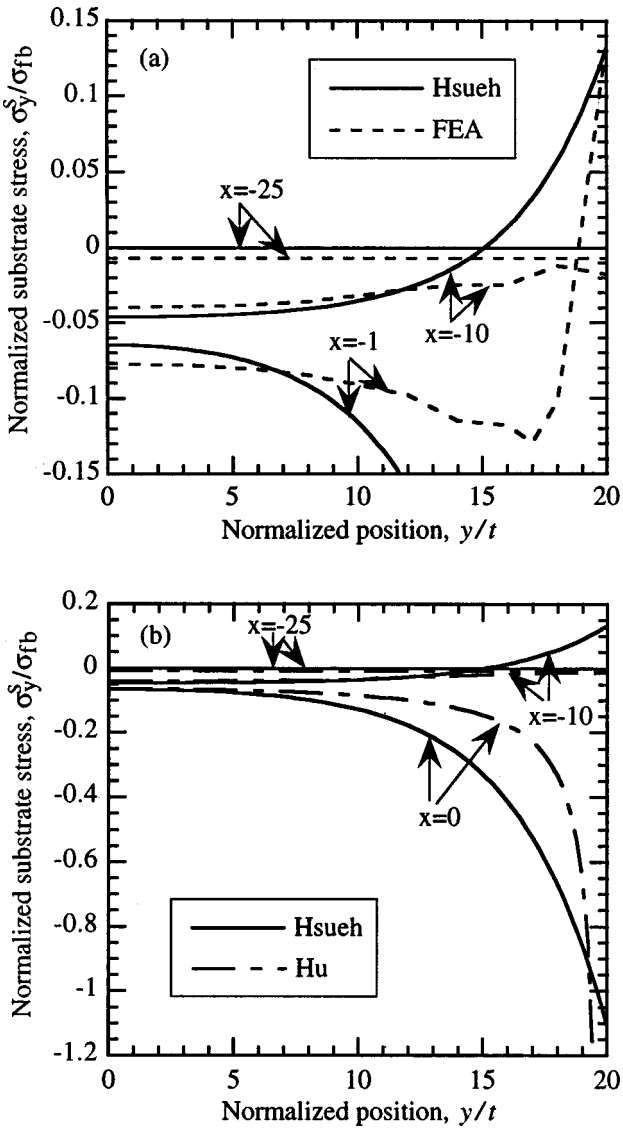


FIG. 4. Comparison of the variations of normalized substrate stress, σ_y^s/σ_{fb}, along the film width between the results obtained from (a) the present model and the FEA, and (b) the present model and Hu's model.

C. Stress distribution in the substrate

The stress distribution in the substrate has been analyzed using the FEA. Using one unit for the film thickness and 20 units for the half film width l, the stresses in the substrate have been calculated at positions 1, 10, and 25 units away from the interface, (i.e., at $x=-1$, -10, and -25) respectively, and the results are shown in Fig. 4(a). Choosing 25 units as the effective substrate thickness, the results predicted from the present model are also shown in Fig. 4(a). The stress distribution in the substrate can also be obtained from Hu's model [Eq. (1)], in which the concentrated edge force acting on the substrate, F, becomes $t\sigma_{fb}$. The comparison between the present results and Hu's results is shown in Fig. 4(b). For y positions away from the edge position (e.g., $y/t<10$ for $l/t=20$), the stresses in the substrate predicted from the present model, the FEA, and Hu's model are similar. However, very different stresses are predicted from the three models for y positions close to the edge, and these differences are discussed as follows. First, without the edge effect, mechanical equilibrium of the system requires that the stresses in the substrate and the film have the opposite sign, such that the normalized substrate stress, σ_y^s/σ_{fb}, is negative. The condition that the film has a free surface at the edge (i.e., $\sigma_f=0$ at $y=l$) is satisfied in both the present model and the finite element model. Hence, the integral of the substrate stresses at $y=l$ over the substrate thickness is zero which, in turn, results in both positive and negative signs for σ_y^s/σ_{fb} at $y=l$ depending on the x position [Fig. 4(a)]. However, a concentrated force, F, at the edge position acting on the surface of the substrate [i.e., at $(x,y)=(0,l)$] is considered in Hu's model, the integral of the substrate stresses at $y=l$ over the substrate thickness must balance F, and σ_y^s/σ_{fb} at $y=l$ is always negative through the substrate thickness [see Eq. (1) and Fig. 4(b)]. Second, Hu's results show that the normalized substrate stress has a negative singularity at $(x,y)=(0,l)$. In the present analytical model, the stress singularity cannot be handled exactly; however, the magnitude of the negative σ_y^s/σ_{fb} reaches a maximum at $(x,y)=(0,l)$. In the finite element results, the stress singularity is reflected by an abrupt change of the sign for σ_y^s/σ_{fb} at positions close to the interface [see $x=-1$ in Fig. 4(a)] and near the edge. Considering σ_y^s/σ_{fb} at $y=l$, while the finite element results show positive values at locations close to the interface and negative values at locations away from the interface, the present results show the opposite [Fig. 4(a)].

D. Effects of E_s/E_f and l/t on the stress distribution in the film

To examine the effects of Young's modulus ratio on the stress distribution in the film, the normalized stress in the film, σ_f/σ_{fb}, as a function of the normalized position, y/t, are shown in Fig. 5(a) for $s/t=25$, and $l/t=20$ at different values of E_s/E_f. The stress in the film, σ_f, is zero at the free edge and increases with the distance from the edge; however, the rate of increase of σ_f with the distance from the edge increases with increasing E_s/E_f.

To examine the effects of the film aspect ratio on the stress distribution in the film, the normalized stress in the film, σ_f/σ_{fb}, as a function of the normalized position, y/t, is shown in Fig. 5(b) for $s/t=25$ at different values of l/t. The stress at the center of the film reaches an equilibrium value defined by Eq. (25) if the film is sufficiently wide [e.g., $l/t \geqslant 30$ in Fig. 5(b)].

V. CONCLUSIONS

A closed-form analytical solution is derived in the present study to analyze the edge effects on the residual stress distribution in a thin-film strip on a substrate. In the existing analytical models, the stress variation through the film thickness is always ignored, and there are two major approaches; i.e., Hu's model and Suhir's model. The first approach starts with a single-edged uniformly stressed film on an infinitely thick substrate and considers a concentrated force exerted by the film edge on the substrate.[8] Because of the deformation of the substrate under the force, the stress in the film is modified. In this case, the stress distribution in the

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

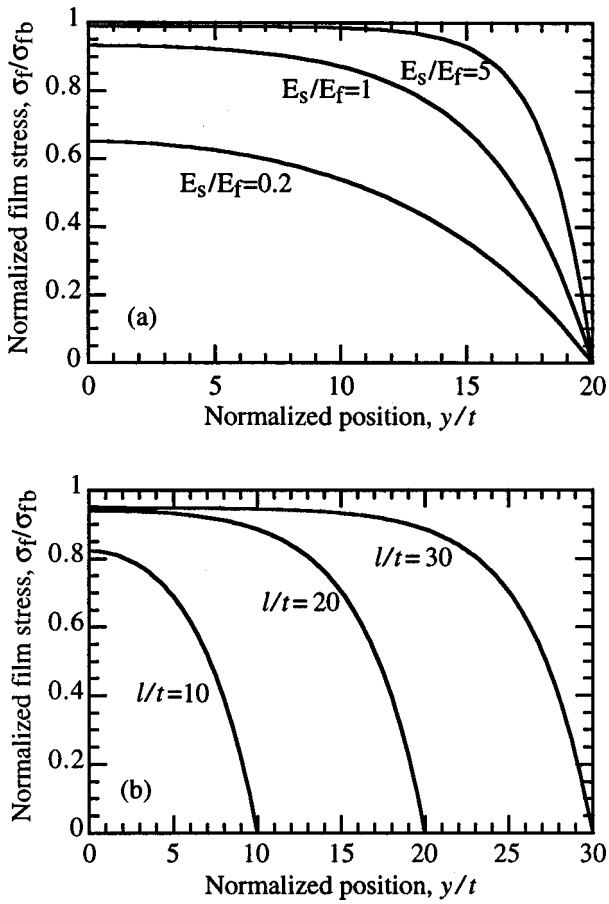


FIG. 5. Effects of (a) E_s/E_f, and (b) l/t on the stress distribution in the film.

substrate has been formulated by Timoshenko and Goodier.[20] The concentrated force model[8] lacks self-consistency due to the film stress relaxation and is, consequently, improved to a distributed force model.[9] In this case, the stress distribution in the system is an integral equation and a numerical method is required to obtain the solution. When the film is a strip with two edges, the principle of superposition is used to obtain the solutions from the solutions for the single-edged film. It has been noted that the solution obtained from the principle of superposition does not satisfy the free surface condition at the film edges. The second approach considers a distributed force $P(y)$ within the film and a distributed force of the opposite sign within the substrate. A distributed shear force $q(y)$ is introduced at the interface between the film and the substrate to equilibrate the nonuniform force distribution $P(y)$. However, the force $P(y)$ is assumed to be uniformly distributed through the thickness of the substrate.

The present analytical model keeps the essence of the second approach but adds the stress variation through the substrate thickness in the analysis. Compared to the existing analytical models, the present model is more rigorous and the analytical results agree better with the finite element results. Also, a qualitative agreement for the stress at the center of the film is obtained between the present analytical results and the experimental measurements. However, two issues remain unsolved. First, the stress singularity, which occurs at the points where the film edges meet the substrate, cannot be exactly handled in analytical models. Second, the thin film influences the stress field in the substrate only within a certain boundary zone in the neighborhood of the film. Hence, when the substrate is much thicker than the film, an effective substrate thickness instead of the actual substrate thickness should be used in analytical modeling. This effective substrate thickness can be extracted from the finite element results. Using the effective substrate thickness, the analytical results agree better with both the finite element results and the experimental measurements. However, an analytical model capable of predicting the effective substrate thickness has not been established.

ACKNOWLEDGMENTS

The author thanks Dr. P. F. Becher, Dr. G. E. Ice, and Dr. M. J. Lance for reviewing the manuscript. Research sponsored by U.S. Department of Energy, Division of Materials Sciences and Engineering, Office of Basic Energy Sciences, under Contract No. DE-AC05-00OR22725 with UT-Battelle.

[1] J. Vanhellemont, S. Amelinckx, and C. Claeys, J. Appl. Phys. **61**, 2170 (1987).
[2] I. De Wolf, J. Vanhellemont, A. Romano-Rodríguez, H. Norström, and H. E. Maes, J. Appl. Phys. **71**, 898 (1992).
[3] D. A. Faux, J. Appl. Phys. **75**, 186 (1994).
[4] I. S. Yeo, P. S. Ho, and S. G. H. Anderson, J. Appl. Phys. **78**, 945 (1995).
[5] Y. L. Shen, S. Suresh, and I. A. Blech, J. Appl. Phys. **80**, 1388 (1996).
[6] S. M. Hu, J. Appl. Phys. **70**, R53 (1991).
[7] S. C. Jain, A. H. Haker, A. Atkinson, and K. Pinardi, J. Appl. Phys. **78**, 1630 (1995).
[8] S. M. Hu, Appl. Phys. Lett. **32**, 5 (1978).
[9] S. M. Hu, J. Appl. Phys. **50**, 4661 (1979).
[10] E. Suhir, J. Appl. Mech. **53**, 657 (1986).
[11] E. Suhir, Mater. Res. Soc. Symp. Proc. **91**, 73 (1987).
[12] A. Fischer and A. Richter, J. Appl. Phys. **75**, 657 (1994).
[13] S. Sakai, K. Kawasaki, and N. Wada, Jpn. J. Appl. Phys., Part 2 **29**, L853 (1990).
[14] P. Van Mieghem, S. C. Jain, J. Nijs, and R. Van Overstraeten, J. Appl. Phys. **75**, 666 (1994).
[15] A. H. Harker, K. Pinardi, S. C. Jain, A. Atkinson, and R. Bullough, Philos. Mag. **A71**, 871 (1995).
[16] H. P. Lee, X. Liu, H. Lin, J. S. Smith, S. Wang, Y. H. Huang, P. Yu, and Y. Z. Huang, Appl. Phys. Lett. **53**, 2394 (1988).
[17] M. Yamaguchi, M. Tachikawa, M. Sugo, S. Kondo, and Y. Itoh, Appl. Phys. Lett. **56**, 27 (1990).
[18] W. C. Tang, H. J. Rosen, S. Guha, and A. Madhukar, Appl. Phys. Lett. **58**, 1644 (1991).
[19] B. Dietrich, E. Bugiel, J. Klatt, G. Lippert, T. Morgenstern, H. J. Osten, and P. Zaumseil, J. Appl. Phys. **74**, 3177 (1993).
[20] S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity* (McGraw-Hill, New York, 1951) p. 97.
[21] S. P. Timoshenko, J. Opt. Soc. Am. **11**, 233 (1925).
[22] H. L. Cox, Brit. J. Appl. Phys. **3**, 72 (1952).
[23] A. Takaku and R. G. C. Arridge, J. Phys. D: Appl. Phys. **6**, 2038 (1973).
[24] M. R. Piggott, Polym. Compos. **8**, 291 (1987).
[25] C. H. Hsueh, J. Mater. Sci. Lett. **7**, 497 (1988).
[26] Y. C. Gao, Y. W. Mai, and B. Cotterell, J. Appl. Math. Phys. (ZAMP) **39**, 550 (1988).
[27] C. H. Hsueh, Mater. Sci. Eng., A **123**, 1 (1990).
[28] A. I. Sauter and W. D. Nix, Mater. Res. Soc. Symp. Proc. **188**, 15 (1990).
[29] G. P. Cherepanov, J. Appl. Phys. **75**, 844 (1994).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/japo/japcpyrts.html.

ELSEVIER

Composites Science and Technology 60 (2000) 2671–2680

COMPOSITES SCIENCE AND TECHNOLOGY

www.elsevier.com/locate/compscitech

Young's modulus of unidirectional discontinuous-fibre composites☆

Chun-Hway Hsueh *

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Received 5 November 1999; received in revised form 7 March 2000; accepted 6 June 2000

Abstract

Young's modulus of a unidirectional discontinuous-fibre composite is derived in the present study by the use of a modified shear-lag model. To achieve this, correct determination of the stress transfer at the fibre ends and selection of the representative volume element in the analytical model are required. The present predicted stress distribution along the fibre length during uniaxial loading of the composite is in excellent agreement with that obtained from a finite-element analysis. Also, by a judicious choice of the representative volume element for the modified shear-lag model, the predicted Young's modulus of the composite is in excellent agreement with those predicted from the Halpin–Tsai semi-empirical equation, the Eshelby model, and numerical analyses. Comparison is also made with experimental measurements of Young's moduli of aligned particulate- or whisker-reinforced SiC/aluminum-alloy composites, and good agreement is obtained. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: A. Short-fibre composites; B. Modelling; C. Stress transfer; C. Elastic properties

1. Introduction

Substantial reinforcement can be achieved by incorporating aligned fibres in the matrix with the fibre axis parallel to the loading direction, and there has been a rapid growth in the use of short-fibre composites because of their great versatility, high performance, and ease of fabrication [1–7]. There has been extensive study of the mechanical properties of these composites, and a comprehensive review of predictions of the elastic constants of these composites has been performed recently by Tucker and Liang [8]. Among the elastic constants, the longitudinal Young's modulus is the most fundamental. Theoretical analyses of Young's modulus can be classified into four types: the Halpin–Tsai semi-empirical equation [9], the Eshelby model [10–13], the shear-lag model [13–17], and numerical methods [8,18]. Among the four types of analyses, the shear lag model is considered to be the least accurate one on account of its crude approximation in the analysis [13]. However, among the analytical models, the stress-transfer mechanism between the fibre and the matrix during loading of the composite can only be derived with the shear-lag model. Hence, a question is raised. Can the shear-lag model be improved to allow better prediction of the Young's modulus of the composite?

The purpose of the present study is to improve the prediction of Young's modulus of an aligned short-fibre composite by the shear-lag model and the stress-transfer mechanism. To achieve this, a newly developed modified shear-lag model [19,20] is adopted. First, both the Halpin–Tsai semi-empirical equation and the Eshelby model are summarized. Second, the essence of the modified shear-lag model is presented, and its solutions are then extended to derive the Young's modulus of the composite. Third, the predicted stress-transfer phenomena is compared to the results obtain from a finite-element analysis (FEA) [21]. Fourth, the predicted Young's modulus of the composite is compared to those predicted from the Halpin–Tsai equation and the Eshelby model. Comparison is also made with experimental measurements of the Young's moduli of particulate- or whisker-reinforced SiC/aluminum-alloy composites [22–24]. It is noted that the results derived from the shear-lag model are affected by the selection of the representative volume element (RVE) in modeling,

☆ The submitted manuscript has been authorized by a contractor for the US Government under contract No. DE-ACO596OR2246. According, the US Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

* Tel.: +1-865-576-6586; fax: +1-865-574-8445.

E-mail address: hsuehc@ornl.gov.

0266-3538/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0266-3538(00)00128-7

and a simple method is developed in the present study to determine the RVE. Finally, the analytical solutions for the Young's modulus of the composite are compared to those obtained from finite-difference [18] and finite-element [8] analyses.

2. Halpin–Tsai semi-empirical equation and Eshelby model

Both the Halpin–Tsai semi-empirical equation and the Eshelby model can predict the longitudinal Young's modulus of an aligned short-fibre composite quite well, and they are summarized as follows while a more thorough review can be found elsewhere [8,25].

2.1. Halpin–Tsai semi-empirical equation

The Halpin–Tsai semi-empirical equation has been adopted extensively because of its simple formulation. The longitudinal Young's modulus of an aligned short-fibre composite, E_o, can be described by [9]

$$\frac{E_o}{E_m} = \frac{1 + \xi\eta f}{1 - \eta f} \tag{1}$$

where f is the volume fraction of fibres, and η is given by

$$\eta = \frac{(E_f/E_m) - 1}{(E_f/E_m) + \xi} \tag{2}$$

The subscripts, f and m, denote the fibre and the matrix, respectively. The parameter, ξ, is obtained by curve-fitting of the predicted E_o with experimental results, and it represents a measure of the fibre geometry, the packing geometry, and the loading condition. Good agreement between E_o and experimental measurements has been obtained by letting ξ equal twice the aspect ratio of the fibre [9].

2.2. Eshelby model

The Eshelby model was originally derived for the case of loading of an infinite matrix containing an ellipsoidal inclusion [10]. It was then modified by Mori and Tanaka to include the effect of a finite volume fraction of inclusions [11], and has been applied to the case of uniaxial loading of an aligned short-fibre composite [12,26]. Whereas the stress in an ellipsoidal inclusion is uniform on account of the nature of its geometry, the stress in an embedded short fibre is non-uniform. Hence, the Eshelby model cannot be used to predict the stress-transfer phenomenon, which is a local property, in a short-fibre composite. However, in analyzing Young's modulus, which is a global property, of a short-fibre composite, the Eshelby model gives a good approach and it is summarized as follows.

The Cartesian coordinates, X_1, X_2, and X_3, are used, and X_3 is the axis of symmetry of the ellipsoidal inclusion. A stress, σ_o, is applied to the composite in the X_3 direction. In this case, the remote applied strains, ε^a_{ij}, become

$$\varepsilon^a_{11} = \varepsilon^a_{22} = \frac{-\nu_m \sigma_o}{E_m} \tag{3a}$$

$$\varepsilon^a_{33} = \frac{\sigma_o}{E_m} \tag{3b}$$

where ν is Poisson's ratio. In the absence of the inclusion, the deformation throughout the system is uniform; i.e., the applied strain ε^a_{ij}. However, because of the presence of the inclusion, the deformation in the inclusion is perturbed, and the strain due to this perturbation is defined as the constrained strain, ε^c_{ij}. To solve this problem, the true inclusion is replaced by an equivalent inclusion, which has the same elastic constants as the matrix and undergoes an equivalent transformation strain, ε^t_{ij}. However, this replacement is subjected to the requirement that the displacement and the stress within the true and the equivalent inclusions are identical, such that [10]

$$K_e(\varepsilon^c + \varepsilon^a) = K_m\left(\varepsilon^c + \varepsilon^a - \varepsilon^t\right) \tag{4a}$$

$$G_e\left('\varepsilon^c_{ij} +' \varepsilon^a_{ij}\right) = G_m\left('\varepsilon^c_{ij} +' \varepsilon^a_{ij} -' \varepsilon^t_{ij}\right) \tag{4b}$$

where K and G are the bulk modulus and the shear modulus, the subscript, e, denotes the ellipsoidal inclusion, and ε and $'\varepsilon_{ij}$ are the dilatational and the deviatoric strains, respectively.

When the volume fraction of inclusions is finite, the constrained strain in the inclusion consists of two components, such that [11]

$$\varepsilon^c_{ij} = \varepsilon^{c\,\prime}_{ij} + \varepsilon^{c\,\prime\prime}_{ij} \tag{5}$$

where $\varepsilon^{c\,\prime}_{ij}$ is the constrained strain component when a single inclusion is embedded in an infinite matrix, and it can be related to the transformation strain by [10]

$$\varepsilon^{c\,\prime}_{ij} = S_{ijkl}\varepsilon^t_{kl} \tag{6}$$

where S_{ijkl} is the Eshelby tensor and is a function of the aspect ratio of the inclusion and Poisson's ratio of the matrix. The second constrained strain component, $\varepsilon^{c\,\prime\prime}_{ij}$, is introduced to account for the effect of the presence of other inclusions on an inclusion, and can be related to the transformation strain by [11]

$$(1-f)\varepsilon_{ij}^{c''} + f\left(\varepsilon_{ij}^{c} - \varepsilon_{ij}^{t}\right) = 0 \tag{7}$$

The solutions of the strain components, ε_{ij}^{c} and ε_{ij}^{t}, can be obtained by solving Eqs. (3)–(7).

The stress in the ellipsoidal inclusion in the loading direction, σ_e, can be derived from the stress/strain relationship, such that [10]

$$\sigma_e = \frac{\nu_e E_e(\varepsilon^c + \varepsilon^a)}{(1+\nu_e)(1-2\nu_e)} + \frac{E_e\left(\varepsilon_{33}^{c} + \varepsilon_{33}^{a}\right)}{1+\nu_e} \tag{8}$$

The effective Young's modulus of the composite along the loading direction, E_o, can be obtained by dividing the applied stress with the average strain in the composite in the loading direction, such that [27]

$$E_o = \frac{\sigma_o}{\varepsilon_{33}^{a} + f\varepsilon_{33}^{t}} \tag{9}$$

It is noted that formulations of the solutions of ε_{ij}^{c} and ε_{ij}^{t} (and hence σ_e and E_o) are lengthy and have been obtained by Tandon and Weng [12]. Only the results are presented in the present study.

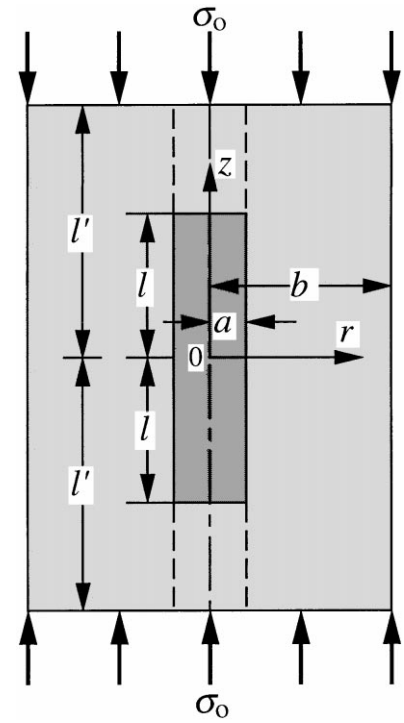


Fig. 1. A schematic showing the representative volume element in the shear-lag model for aligned short-fibre composites. The imaginary fibres are added to analyze the stress transfer at the bonded fibre ends.

3. The modified shear-lag model

The RVE adopted in the shear-lag model is shown in Fig. 1. A discontinuous fibre with a radius a and a length $2l$ is embedded at the centre of a coaxial cylindrical matrix with a radius b and a length $2l'$, such that a^2l/b^2l' corresponds to the volume fraction of fibres in the composite, f.[1] The cylindrical coordinates, r, θ, and z, are used. A tensile stress, σ_o, is applied to the RVE in the z direction. The fibre is bonded to the matrix at both the interface (i.e. at $r = a$) and the ends (i.e., at $z = \pm 1$). The stress is transferred from the matrix to the fibre through both the interface and the fibre ends. The geometry in Fig. 1 is symmetric, and only the region $z \geqslant 0$ needs to be considered.

In the classical shear-lag model [14], the fibre is assumed to sustain only the axial load, the matrix is assumed to transmit only the shear stress, and no load is transferred through the fibre ends (i.e. the fibre ends are debonded from the matrix). It is noted that, in the case of an embedded fibre with bonded fibre ends, the stress at the bonded ends has a finite value, and this stress is required as a boundary condition in order to derive the stress distribution in the fibre. However, this finite stress is not a predetermined value, which, in turn, results in difficulties in using the classical shear-lag model. Various assumptions of this finite stress have been made to solve the stress-transfer problem [16,17].

In the modified shear-lag model [19,20], the radial dependence of the axial stress in the matrix is included in the analysis, the axial-stress distribution in the RVE is in equilibrium with the applied load, the shear-stress distribution in the matrix satisfies the free surface condition at $r = b$, and the boundary condition at the bonded fibre ends is defined by the continuity condition at $z = \pm l$. In order to satisfy this continuity condition, a technique of adding imaginary fibres to the shear-lag model has been developed [19,20]. The matrix connecting the fibre ends to the surface of the RVE (see the area between the dash lines in Fig. 1) is treated as two imaginary fibres, which have the matrix properties. Hence, the stress-transfer problem depicted by Fig. 1 can be solved using two shear-lag models: one for $0 \leqslant z \leqslant l$ and one for $l \leqslant z \leqslant l'$. Although the stress at the fibre end is unknown, the two shear-lag models are jointed at $z = l$ with the continuity condition, which, in turn, determines the stress at the fibre end. Incidentally, a technique developed earlier by Fukuda and Chou [28] by introducing a partially axial load into the matrix in the classical shear-lag model is found to be similar to the "imaginary fibre" technique. Using the methodology described above, the stress transfer problem has been analyzed previously by also including Poisson's effect and residual thermal stresses [20]. The predicted stress

[1] In the classical shear-lag model, the matrix radius, b, is defined by the fibre packing arrangement [8], and the formulation of the fibre volume fraction is different from that defined in the present study.

distribution along the fibre length is in a good agreement with the fluorescence spectroscopy measurement performed on a soda-lime-silicate (SLS) glass matrix containing a single embedded PRD-166 alumina/zirconia fibre subjected to thermo-mechanical loading [20]. It is noted that whereas Poisson's effect is important when the interface at $r = a$ is debonded and frictional [29], it can be ignored when the interface is bonded, in analyzing the stress-transfer problem [30]. Also, the primary interest in the present study is Young's modulus of the composite. Hence, in order to simplify the equations, both Poisson's effect and residual thermal stress are removed from the analysis in Ref. [20], and the axial-stress distributions in the fibre and the imaginary fibre, σ and σ_f', become:

$$\sigma_f = \frac{b^2 E_f \sigma_o}{a^2 E_f + (b^2 - a^2) E_m} + A[\exp(\alpha z) + \exp(-\alpha z)] \quad (\text{for } 0 \leqslant z \leqslant l) \tag{10a}$$

$$\sigma_f{}' = \sigma_o + A'\{\exp(\alpha' z) - \exp[\alpha'(2l' - z)]\} \quad (\text{for } l \leqslant z \leqslant l') \tag{10b}$$

where

$$\alpha = \frac{1}{a} \times \left\{ \frac{a^2 + (b^2 - a^2)\dfrac{E_m}{E_f}}{b^2\left[\dfrac{b^2}{b^2 - a^2}\ln\left(\dfrac{b}{a}\right) - \dfrac{3b^2 - a^2}{4b^2}\right](1 + \nu_m)} \right\}^{1/2} \tag{11a}$$

$$\alpha' = \frac{1}{a}\left\{ \frac{1}{\left[\dfrac{b^2}{b^2 - a^2}\ln\left(\dfrac{b}{a}\right) - \dfrac{3b^2 - a^2}{4b^2}\right](1 + \nu_m)} \right\}^{1/2} \tag{11b}$$

$$A = \frac{(b^2 - a^2)(E_m - E_f)\sigma_o}{a^2 E_f + (b^2 - a^2) E_m} \times \left\{ \exp(\alpha l) + \exp(-\alpha l) - \frac{\alpha[\exp(\alpha l) - \exp(-\alpha l)] \times \{\exp(\alpha' l) - \exp[\alpha'(2l' - l)]\}}{\alpha'\{\exp(\alpha' l) + \exp[\alpha'(2l' - l)]\}} \right\}^{-1} \tag{12a}$$

$$A' = \frac{\alpha[\exp(\alpha l) - \exp(-\alpha l)]A}{\alpha'\{\exp(\alpha' l) + \exp[\alpha'(2l' - l)]\}} \tag{12b}$$

It is noted that although the modified shear-lag model is much more rigorous than the classical shear-lag model, it is not exact and approximations are adopted in deriving the axial stress in the matrix. Also, the axial displacement is non-uniform on the faces of the RVE. In this case, the effective Young's modulus of the RVE in the axial direction is dictated by the average axial displacement at $z = l'$ under the axial loading stress, σ_o. The average axial displacement at $z = l$ is

$$w = \frac{2}{b^2}\left(\int_0^a r \int_0^l \frac{\sigma_f}{E_f}\,dz\,dr + \int_a^b r \int_0^l \frac{\sigma_m}{E_m}\,dz\,dr\right) \tag{13}$$

where σ_m is the axial stress in the matrix. It is noted that whereas the radial dependence of σ_f is ignored in both the classical and the modified shear-lag models as a result of the relatively small fibre radius, the radial dependence of σ_m is considered in the modified shear-lag model [i.e. $\sigma_f(z)$ and $\sigma_m(r, z)$ are considered in the analysis]. However, the explicit expression of σ_m is not required in Eq. (13) in order to solve w. This is because the mechanical equilibrium condition requires

$$2\int_0^a r\sigma_f\,dr + 2\int_a^b r\sigma_m\,dr = b^2\sigma_o \tag{14}$$

The combination of Eqs. (13) and (14) yields

$$w = \frac{l\sigma_o}{E_m} + \frac{a^2}{b^2}\left(\frac{1}{E_f} - \frac{1}{E_m}\right)\int_0^l \sigma_f\,dz \tag{15}$$

Substitution of Eq. (10a) into Eq. (15) gives

$$w = \frac{l\sigma_o}{E_m} + \frac{a^2}{b^2}\left(\frac{1}{E_f} - \frac{1}{E_m}\right) \times \left\{\frac{lb^2 E_f \sigma_o}{a^2 E_f + (b^2 - a^2)E_m} + \frac{A}{\alpha}[\exp(\alpha l) - \exp(-\alpha l)]\right\} \tag{16}$$

Similarly, the average axial displacement at $z = l'$ relative to $z = l$ can be derived from

$$w' = \frac{2}{b^2}\left(\int_0^a r \int_l^{l'} \frac{\sigma_f'}{E_m}\,dz\,dr + \int_a^b r \int_l^{l'} \frac{\sigma_m}{E_m}\,dz\,dr\right) \tag{17}$$

such that

$$w' = \frac{(l' - l)\sigma_o}{E_m} \tag{18}$$

Hence, the average axial displacement at $z = l'$, $w_o(= w + w')$, is

$$w_o = \frac{l'\sigma_o}{E_m} + \frac{a^2}{b^2}\left(\frac{1}{E_f} - \frac{1}{E_m}\right) \times \left\{\frac{lb^2 E_f \sigma_o}{a^2 E_f + (b^2 - a^2)E_m} + \frac{A}{\alpha}[\exp(\alpha l) - \exp(-\alpha l)]\right\} \tag{19}$$

The effective Young's modulus of the RVE in the axial direction, E_o, can be related to the loading stress, the average axial displacement at $z = l'$, and the length of the RVE, such that

$$E_o = \frac{l'\sigma_o}{w_o} \tag{20}$$

A closed-form solution for E_o is hence obtained. However, it is noted that while the elastic constants other than E_o can also be calculated using the Halpin–Tsai equation and the Eshelby model, they cannot be obtained from the shear-lag model.

4. Results

The present predicted stress distribution along the fibre length during uniaxial loading of the composite is compared to that obtained from a FEA. Also, the present analytical solution of longitudinal Young's modulus of an aligned short-fibre composite is compared to the results obtained from the Halpin–Tsai semi-empirical equation, the Eshelby model, experimental measurements, the finite-difference analysis, and the FEA.

4.1. Comparison of stress transfer with finite-element analysis

The stress-transfer problem for the configuration depicted in Fig. 1 has been analyzed using a finite-element analysis, ABAQUS, in which an eight-node, biquadratic axisymmetry solid element is adopted [21]. The material properties used in the FEA are those pertinent to a PRD-166/SLS composite, such that $E_f = 380$ GPa, $E_m = 74$ GPa, $\nu_f = 0.27$, $\nu_m = 0.3$. The dimensions adopted in the FEA are: $b/a = 20$, $l/a = 30$, and $l'/a = 60$, and the solutions of the fibre-axial stress are taken from the fibre elements next to the interface. The results of fibre-axial stress along the loading direction normalized by the loading stress are shown in Fig. 2. The present predicted results of σ_f/σ_o and σ_f'/σ_o are also shown. It is noted that the fibre-axial stress defined in the modified shear-lag model is the average value with respect to the fibre cross-section [19,20] which is different from that defined in the FEA. However, as a result of the small fibre radius, the difference in defining the fibre axial stress between the modified shear-lag model and the FEA is expected to be small. Good agreement between the present analytical results and the FEA results is obtained in Fig. 2. It is also noted that the comparison is based on the solutions obtained from the RVE containing only one fibre. In the case of composites containing multiple fibres, the solutions also depend on the fibre packing arrangement [8].

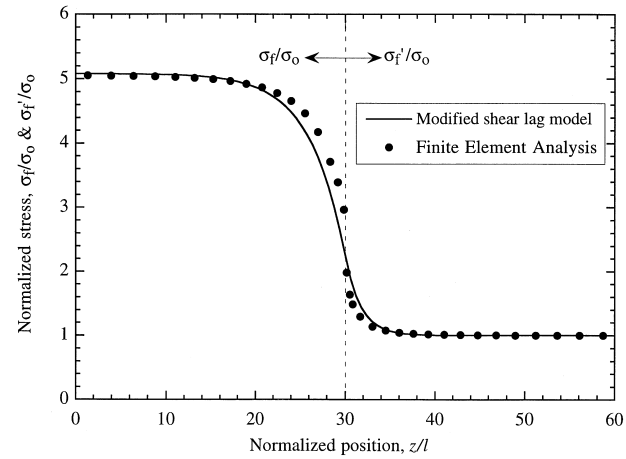


Fig. 2. Comparison of distributions of the normalized axial stress, σ_f/σ_o and $\sigma_{f'}/\sigma_o$, along the axial loading direction between the modified shear-lag model and the finite-element solutions.

4.2. Comparison with the Halpin–Tsai equation, the Eshelby model, and experiments

Aluminum alloys reinforced with SiC particles or whiskers have extensive applications in aircraft engines and aerospace structures. The composite is commonly made by extrusion, and the SiC reinforcement is generally oriented in the extrusion direction. Unless noted otherwise, the material properties of the SiC/Al composite, $E_f = 450$ GPa, $E_m = 72$ GPa, $\nu_f = 0.2$, and $\nu_m = 0.3$, are adopted in the present study to predict Young's modulus of the composite in the extrusion direction. It is noted that with the above material properties, only the aspect ratio and the volume fraction of the reinforcement are needed in using the Halpin–Tsai equation or the Eshelby model to predict the Young's modulus of the composite. However, the specific dimensions of the RVE (i.e. b and l' in Fig. 1) are also required in using the modified shear-lag model because different combinations of b and l' can give the same volume of the RVE. In the literature, two different geometries have been found in choosing a RVE to model a unidirectional discontinuous fibre composite: the RVE scaling with the fibre (i.e. $l'/b = l/a$) [31], and the RVE and the fibre having the same length (i.e. $l' = l$) [13,16,17,32]. It has been noted that $l' = l$ is a good assumption for the case of long embedded fibres and

would induce significant errors for the case of short embedded fibres [13]. The present calculated results using the above two geometries are presented as follows.

To summarize the above two geometries, the radius of the RVE, b, can be related to the fibre radius, a, by

$$b = af^{-p} \tag{21}$$

where $p = 1/3$ when $l'/b = l/a$, and $p = 1/2$ when $l' = l$. Using these two geometries, the predicted effective Young's modulus, E_o, as a function of the fibre-aspect ratio, l/a, is shown in Fig. 3 for $f = 0.2$. The results from the Halpin–Tsai equation and the Eshelby model are also shown. It can be see that the present results are in excellent agreement with the Halpin–Tsai equation for small fibre-aspect ratios ($l/a <\sim 2$) when $l'/b = l/a$ is used, and are in excellent agreement with both the Halpin–Tsai equation and the Eshelby model for large fibre-aspect ratios ($l/a >\sim 10$) when $l' = l$ is used. This trend agrees with the statement that $l' = l$ can only be used for the case of long embedded fibres. Based on the above trend (Fig. 3), a logical selection of the RVE would have p in Eq. (21) increase from 1/3 at small fibre aspect ratios and approach 1/2 as the fibre-aspect ratio increases. A simple equation describing p adopted in the present study is

$$p = \frac{1}{2} - \frac{1}{6}\exp\left(\frac{-l}{5a}\right) \tag{22}$$

The numerical factor in the exponential term in Eq. (22) (i.e. 5) can be varied to adjust how fast p approaches 1/2 as l/a increase. Using Eq. (22), the parameter, p, as a function of the fibre-aspect ratio, l/a, is shown in Fig. 4. Using Eqs. (21) and (22), l' can be obtained by satisfying $a^2 l/b^2 l' = f$, such that

$$l' = lf^{2p-1} \tag{23}$$

Using Eqs. (21)–(23) to define the RVE, the predicted results of E_o for $f = 0.2$ are shown in Fig. 5a, and excellent agreement with both the Halpin–Tsai equation and the Eshelby model is obtained. The results for $f = 0.1$ and 0.3 are shown in Fig. 5b and c, respectively, and excellent agreement is also obtained.

Using $l'/b = l/a$, $l' = l$, and Eq. (23) in the modified shear-lag model, the normalized axial stress distributions, σ_f/σ_o and σ'_f/σ_o, along the axial loading direction are shown in Fig. 6 for $l/a = 3$ and $f = 0.2$. The stress at the loading surface is σ_o (i.e. $\sigma'_f/\sigma_o = 1$ at $z = l'$). When $l' = l$, the stress at the fibre end is the loading stress, σ_o (Fig. 6). The normalized uniform stress in the ellipsoidal inclusion based on the Eshelby model, σ_e/σ_o, is also shown in Fig. 6. While the stress distribution along the fibre length (and hence the stress-transfer phenomena) can be obtained from the modified shear-lag model depending on the RVE selected, the uniform stress in the ellipsoidal inclusion is obtained from the Eshelby model. In the following comparison, Eqs. (21)–(23) are used to define the RVE for the modified shear-lag model.

Experimental measurements of Young's moduli of particulate- or whisker-reinforced SiC/aluminum alloy composites in the extrusion direction, E_o, have been performed. Depending on the composition and the treatment of the aluminum alloy, its Young's modulus can be varied. The material properties of SiC/2124 (T4) [22], SiC/6061 (T6) [23], and SiC/8090 (T6) [24] listed in Table 1 are used for theoretical calculations. The predicted Young's modulus, E_o, based on the modified shear-lag model, the Halpin–Tsai equation, and the Eshelby model are shown in Fig. 7a–c for SiC/2124, SiC/6061, and SiC/8090 composites, respectively. The experimental data [22-24] are also shown, and good

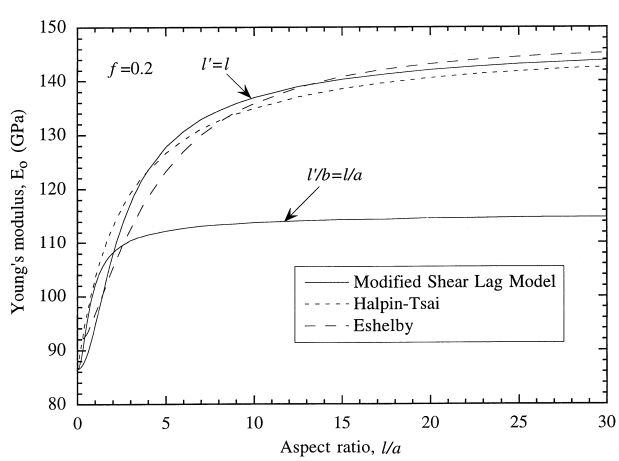


Fig. 3. Longitudinal Young's modulus of SiC/Al composite, E_o, as a function of fibre-aspect ratio, l/a, for $f = 0.2$. The results from the modified shear lag model (assuming $l'/b = l/a$ and $l' = l$), the Halpin–Tsai equation, and the Eshelby model are shown.

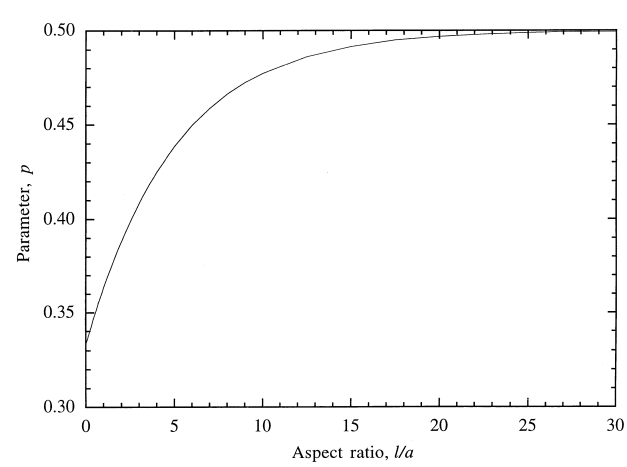


Fig. 4. The parameter, p, in determining the geometry of the representative volume element for aligned short-fibre composites as a function of fibre-aspect ratio, l/a.

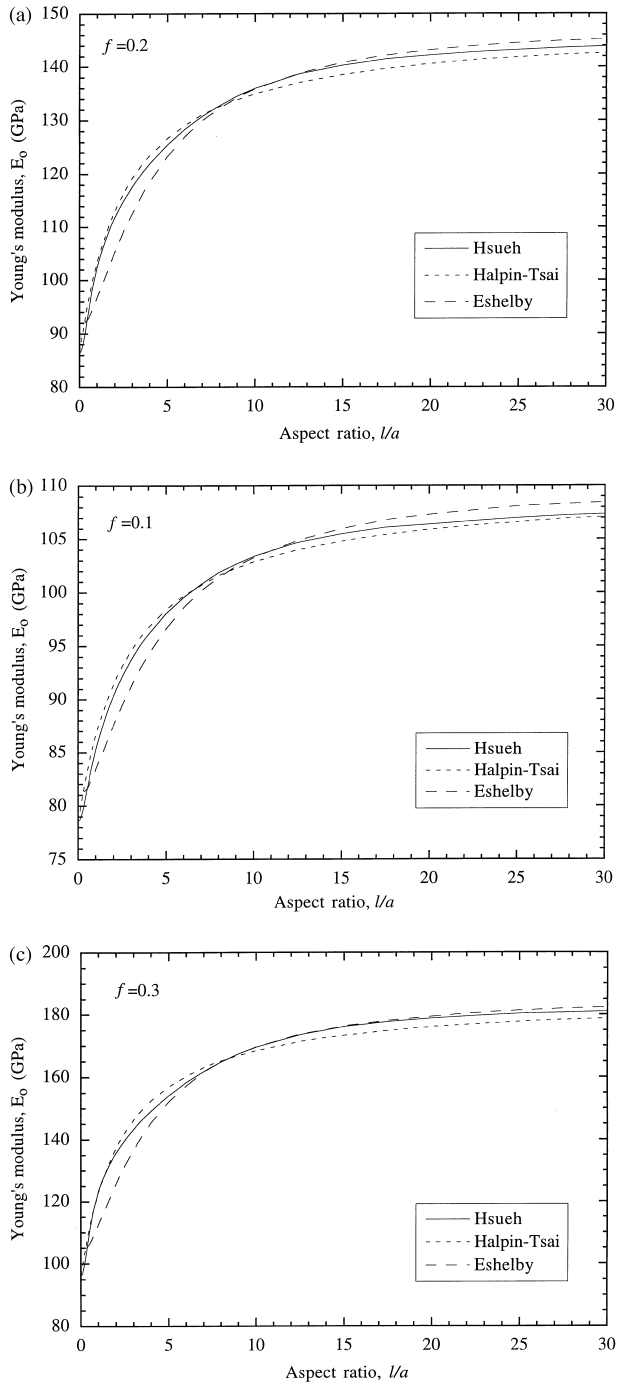


Fig. 5. Longitudinal Young's modulus of SiC/Al composite, E_o, as a function of fibre-aspect ratio, l/a, for (a) $f = 0.2$, (b) $f = 0.1$, and (c) $f = 0.3$. The results from the modified shear-lag model, the Halpin–Tsai equation, and the Eshelby model are shown.

agreement between the three predictions and measurements is obtained.

4.3. Comparison with numerical analyses

Young's modulus of an aligned short-fibre composite has been analyzed numerically using a finite-difference method [18]. However, in order to simplify the analysis, cubic lattices were adopted for both the fibre and the matrix in the finite-difference analysis. Hence, instead of the cylindrical geometry depicted by Fig. 1, the geometry considered in the finite-difference analysis is a rectangular parallelepiped. Also, a multiple-fibre composite was adopted in the finite difference analysis. The difference in geometry between the modified shear-lag model and the finite-difference model would result in differences in the magnitude of the calculated results. However, since Young's modulus of the composite, E_o, is a global property, the difference in its prediction as a result of different geometries is expected to be small. The mechanical properties used in the finite-difference analysis are: $E_f/E_m = 20$ and $\nu_f = \nu_m = 0.3$, and the results were calculated for $l/a = 2$ and 10, respectively, at different fibre-volume fractions. The finite-difference results of E_o/E_m as a function of the fibre-volume fraction, f, are shown in Fig. 8. The predictions based on the modified shear-lag model, the Halpin–Tsai equation, and the Eshelby model are also shown, and good agreement is obtained. It is noted that the results from the modified shear lag model almost overlap with those

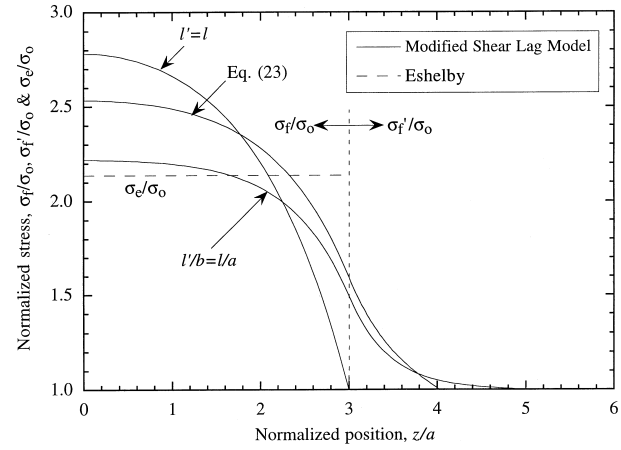


Fig. 6. The normalized axial stress distributions, σ_f/σ_o and $\sigma_{f'}/\sigma_o$, from the modified shear lag model using $l'/b = l/a$, $l' = l$, and Eq. (22), and the uniform normalized stress, σ_e/σ_o, in the ellipsoidal inclusion in the loading direction based on the Eshelby model. The results are shown for SiC/Al composites with $l/a = 3$ and $f = 0.2$.

Table 1
Properties of particulate- or whisker-reinforced SiC/aluminum alloy composites

Composite	Reinforcement	E_e or E_f (GPa)	E_m (GPa)	ν_e	ν_m	Aspect ratio, l/a
SiC/2124 (T4)[a]	Particulate	450	72	0.2	0.3	1.6
SiC/6061 (T6)[a]	Particulate	450	71.7	0.2	0.3	3
SiC/8090 (T6)	Whisker	480	78.2	0.2	0.3	4

[a] T4; room temperature aged; T6; heat treated.

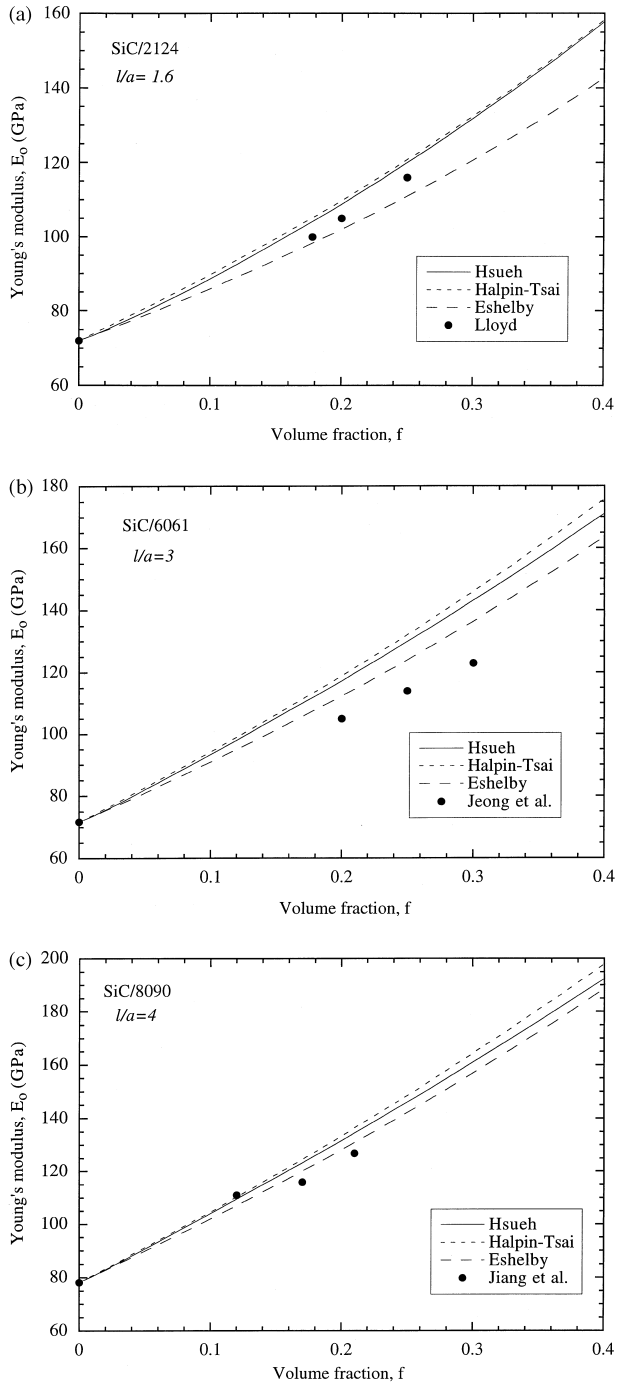


Fig. 7. Longitudinal Young's modulus, E_o, as a function of fibre-volume fraction, f, for (a) SiC/2124, (b) SiC/6061, and (c) SiC/8090 composites. The results from the modified shear lag model, the Halpin–Tsai equation, the Eshelby model, and experimental measurements are shown.

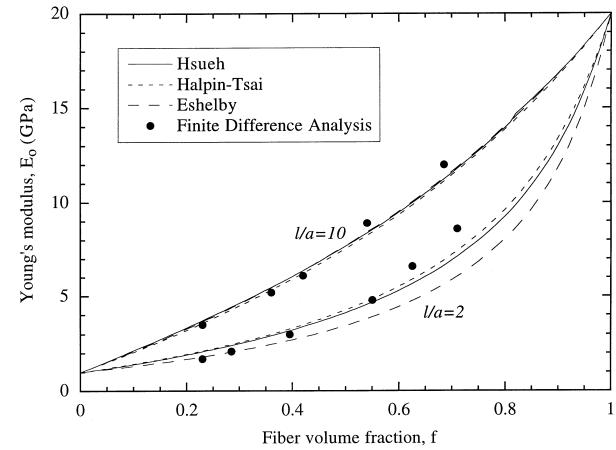


Fig. 8. Longitudinal Young's modulus of aligned short-fibre composites, E_o, as a function of fibre-volume fraction, f, for $l/a = 2$ and 10. The results from the modified shear-lag model, the Halpin–Tsai equation, the Eshelby model, and the finite-difference analysis are shown.

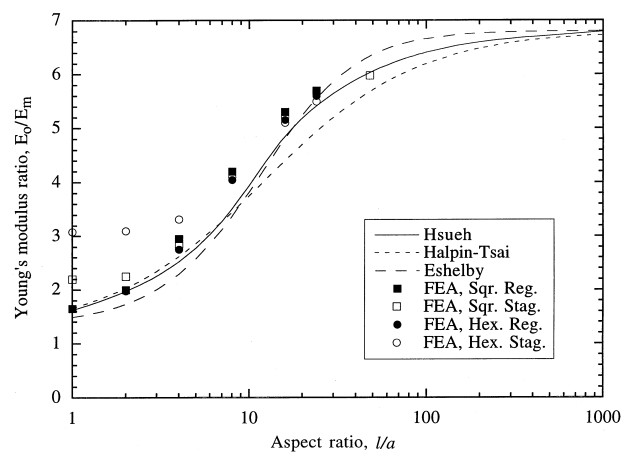


Fig. 9. Normalized longitudinal Young's modulus of aligned short fibre composites, E_o/E_m, as a function of fibre aspect ratio, l/a, for $f = 0.2$. The results from the modified shear-lag model, the Halpin–Tsai equation, the Eshelby model, and the finite-element analysis are shown.

from Eshelby model when $l/a = 10$, and they can hardly be distinguished in Fig. 8.

Young's modulus of an aligned short-fibre composite has also been analyzed using a FEA [8]. Two types of fibre packing arrangements and two types of periodic, three-dimensional arrays of fibres are considered in defining the RVE for the FEA which, in turn, result in four different geometries for the RVE. The two fibre packing arrangements are square and hexagonal, and the two fibre arrays are regular and stagger [8]. The material properties used in the FEA are those pertinent to fibre-reinforced engineering thermoplastics, such that $E_f/E_m = 30$, $\nu_f = 0.2$, $\nu_m = 0.38$. The volume fraction of fibres adopted is 0.2, and different fibre-aspect ratios are used to calculate the effective Young's modulus, E_o, of the RVE. The calculated results of E_o/E_m as a function of the fibre-aspect ratio, l/a, are shown in Fig. 9. The different fibre packing/arrays create scatter in the finite element results. The scatter is small for $l/a \geqslant 8$, but it is significant for $l/a \leqslant 4$. It is believed that the lower finite-element values are more representative of the actual packing and the actual stiffness of composites with very short fibres [8]. The predictions based on the modified shear-lag model, the Halpin–Tsai equation, and the

Eshelby model are also shown in Fig. 9, and good agreement is obtained between the predictions and the lower finite-element values.

5. Conclusions

The Young's modulus of an aligned short-fibre composite is of interest. The theoretical prediction of this value from a shear-lag model is known to be less accurate than those predicted from the Halpin–Tsai semi-empirical equation and the Eshelby model. This is a result of the crude approximation adopted in the shear-lag model. It is noted that the shear-lag model was originally developed to address the stress-transfer problem between the fibre and the matrix. However, whereas the Halpin–Tsai equation and the Eshelby model offer a good prediction of the Young's modulus, they cannot be used to analyze the stress transfer problem for an aligned short-fibre composite. The present study aims at improving the shear-lag model to give better predictions of both Young's modulus and the stress-transfer phenomena in the composite. To achieve this, a rigorous modified shear-lag model is required. This modified shear-lag model has been developed [19,20], and it is then used in the present study to derive the Young's modulus of the composite. However, an inherent problem in the shear-lag model remains in the modified shear lag model; i.e. a representative volume element (RVE) is required in modelling. To resolve this problem, a simple criterion for selecting the RVE is established in the present study [see Eqs. (21)–(23)]. The present predicted stress distribution along the fibre length during uniaxial loading of the composite (Fig. 1) is compared to that obtained from a finite-element analysis, and excellent agreement is obtained (Fig. 2). Also, by a judicious choice of the representative volume element for the modified shear-lag model, the predicted Young's modulus of the composite is in excellent agreement with those predicted from the Halpin–Tsai semi-empirical equation, the Eshelby model, and the numerical analyses (Figs. 5, 8, and 9). Comparison is also made with experimental measurements of Young's moduli of particulate- or whisker-reinforced SiC/aluminum alloy composites, and good agreement is obtained (Fig. 7a–c).

Acknowledgements

The author is grateful to Professor N.A. Yu for making the finite element data in Ms. Y. Zhang's Master Thesis available, and thanks Drs. P.F. Becher, K. Breder, and M.J. Lance for reviewing the manuscript. Research sponsored by US Department of Energy, Division of Materials Sciences, Office of Basic Energy Sciences, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

References

[1] Chou TW, Kelly A. Fibre composites. Mat Sci Eng 1976;25:35–40.
[2] Kelly A. Composites in context. Compos Sci Technol 1985;23:171–99.
[3] Mai YW. In: Shah SP, editor. Applications of fracture mechanics to cementitious composites. Netherlands: Martinus Nijhoff Publ., 1985. p. 399–429.
[4] Arsenault RJ. The strengthening of aluminum alloy by fiber and platelet silicon carbide. Mater Sci Eng 6061;1984:64171-181.
[5] McDanels DL. Analysis of stress-strain, fracture, and ductility behavior of aluminum matrix composites containing discontinuous silicon carbide reinforcement. Metall Trans 1985;16A:1105–15.
[6] Becher PF, Hsueh CH, Angelini P, Tiegs TN. Toughening behavior in whisker-reinforced ceramic composites. J Am Ceram Soc 1988;71:1050–61.
[7] Styles CM, Pitcher PD. Forging behavior and properties of metal matrix composites based on mechanically alloyed Al–Mg–Li alloy. Mater Sci Technol 1998;14:913–9.
[8] Tucker CL, Liang E. Stiffness predictions for unidirectional short-fiber composites: review and evaluation. Compos Sci Technol 1999;59:655–71.
[9] Halpin JC. Primer on composite materials: analysis. Lancaster (PA): Technomic, 1984. p. 130
[10] Eshelby JD. The determination of the elastic field of an ellipsoidal inclusion, and related problems. Proc Roy Soc London, A241:376–96
[11] Mori T, Tanaka K. Average stress in matrix and average elastic energy of materials with misfitting inclusions. Acta Metall 1973;21:571–4.
[12] Tandon GP, Weng GJ. The effect of aspect ratio of inclusions on the elastic properties of unidirectionally aligned composites. Polymer Compos 1984;4:327–33.
[13] Taya M, Arsenault RJ. A comparison between a shear lag type model and an Eshelby type model in predicting the mechanical properties of a short fibre composite. Scripta Metall 1987;21(3):349–54.
[14] Cox HL. The elasticity and strength of paper and other fibrous materials. Brit J Appl Phy 1952;3:72–9.
[15] Piggott MR. Load bearing fiber composites. Elmsford (NY): Pergamon Press, 1980. p. 62
[16] Clyne TW. A simple development of the shear lag theory appropriate for composites with a relatively small modulus mismatch. Mater Sci Eng, A122:183–92
[17] Starink MJ, Syngellakis S. Shear lag models for discontinuous composites: fibre end stresses and weak interface layers. Mater Sci Eng, A270:270–7
[18] Termonia Y. Computer model for the elastic properties of short fibre and particulate filled polymers. J Mater Sci 1987;22:1733–6.
[19] Hsueh CH. A modified analysis for stress transfer in fiber-reinforced composites with bonded fibre ends. J Mater Sci 1995;30:219–24.
[20] Hsueh CH, Young RJ, Yang X, Becher PF. Stress transfer in a model composite containing a single embedded fiber. Acta Metall 1997;45(4):1469–76.
[21] Zhang Y. Modeling of stress transfer in fiber-reinforced ceramic matrix composites. Master thesis, Department of Mechanical and Aerospace Engineering and Engineering Science, The University of Tennessee, Knoxville, 1998.

[22] Lloyd DJ. Particle reinforced aluminum and magnesium matrix composites. Inter Mater Rev 1994;39(1):1–23.
[23] Jeong H, Hsu DK, Shannon RE, Liaw PK. Characterization of anisotropic elastic constant of silicon-carbide particulate reinforced aluminum metal matrix composites: part I. Experiment. Metall Mater Trans 1994;25A:799–809.
[24] Jiang Z, Lian J, Yang D, Dong S. An analytical study of the influence of thermal residual stresses on the elastic and yield behaviors of short fiber-reinforced metal matrix composites. Mater Sci Eng 1998;A248:256–75.
[25] Mlekusch B. Thermoelastic properties of short-fibre-reinforced thermoplastics. Compos Sci Technol 1999;59:911–23.
[26] Hsueh CH. Effects of aspect ratios of ellipsoidal inclusions on elastic stress transfer of ceramic composites. J Am Ceram Soc 1989;72:344–7.
[27] Withers PJ, Stobbs WM, Pedersen OB. The application of the Eshelby method of internal stress determination to short fiber metal matrix composites. Acta Metall 1989;37(11):3061–84.
[28] Fukuda H, Chou TW. An advanced shear-lag model applicable to discontinuous fiber composites. J Comp Mater 1981;15:79–91.
[29] Hsueh CH. Some considerations of evaluation of interfacial frictional stress from the indentation technique for fibre-reinforced ceramic composites. J Mater Sci Lett 1989;8:739–42.
[30] Hsueh CH. Some applications of fiber pullout analyses. J Mater Sci Lett 1991;10:29–32.
[31] Levy A, Papazian JM. Tensile properties of short fibre-reinforced SiC/Al composites: part II. Finite-element analysis. Metall Trans 1990;21A:411–20.
[32] Kelly A, Street KN. Creep of discontinuous fibre composites, II. Theory for the steady-state. Proc R Soc Lond 1972;A328:283–93.

Influence of residual stress on thermal expansion behavior

X.-L. Wang,[a] C. M. Hoffmann, C. H. Hsueh, G. Sarma, C. R. Hubbard, and J. R. Keiser
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6064

(Received 17 August 1999; accepted for publication 20 September 1999)

We demonstrate that the thermal expansion behavior of a material can be substantially modified by the presence of residual stresses. In the case of a composite tube made of two layers of dissimilar steels, *in situ* neutron diffraction measurements revealed a significant difference in the coefficients of thermal expansion along the radial and tangential directions. It is shown that the observed difference in thermal expansion is due to the change of residual stresses with temperature. © *1999 American Institute of Physics.* [S0003-6951(99)01647-2]

Thermal expansion is a fundamental property of a material, arising from the anharmonic vibration of the crystal lattice. For a single crystal under no external influence, its coefficient of thermal expansion (CTE) is compliant with the crystal symmetry.[1] For instance, the CTE of a cubic crystal is isotropic along different crystallographic directions. While the thermal expansion behaviors of simple materials are relatively well understood, there have been continuing studies of the thermal expansion behaviors of complex materials systems, such as ceramic–ceramic and ceramic–metal composites.[2–4] Depending on the microstructure, the overall thermal expansion of a composite material may differ from that given by constitutive models.[2]

In practical applications, thermal expansion is an important consideration, not only from the thermophysical property point of view but also from the standpoint of the mechanical behavior of the material. This is because in a fabricated component, differences in thermal expansion between constituent materials lead to thermal residual stresses, which affect the integrity and lifetime of the component.

The design of replacement materials for composite tubing used in Kraft black liquor recovery boilers[5,6] is a case in point. A composite tube is made of a carbon steel core and a corrosion-resistant clad layer. Because of the thermal expansion mismatch between the carbon steel and clad layers, thermal residual stresses are generated as the boiler goes through operating cycles. Experimental and finite-element studies have shown that the signs and magnitudes of these residual stresses are dictated by the difference in CTE (between the two layers of dissimilar materials) and the yield strength of the clad layer.[6] If the clad layer is in tension, as is the case for composite tubing currently in use, the boiler is vulnerable to stress-corrosion cracking, which could lead to grave economical consequences and causes safety concerns as well.

Accurate modeling of thermal residual stresses depends on detailed knowledge of the thermal expansion behaviors of the constituent materials. We show here that under the influence of residual stresses, the thermal expansion behavior of a material may be substantially different from that in the unconstrained condition. As will be shown below, for samples containing type-I or macroresidual stresses, the effective coefficient of thermal expansion as measured in a diffraction experiment may be *macroscopically* anisotropic.

The specimen is a 2.5-in.-o.d. composite tube. The core part of the tubing is SA 210 carbon steel, with the composition 0.18C–0.28Si–0.65Mn–0.006P–0.15S wt %.[7] The clad layer is Sanicro 28, a nickel-based austenitic alloy steel, with the composition 0.10C–26.7Cr–30.4Ni–3.33Mo–36.2Fe–0.01Nb–1.06Cu wt %.[8] A cross-sectional view of the specimen is shown in Fig. 1. The geometric dimensions are $a = 24.8$ mm, $b = 30.0$ mm, $c = 31.8$ mm, and $l = 300$ mm. Table I lists the coefficients of thermal expansion and elastic properties[7,8] for each material comprising the tube specimen.

The neutron diffraction experiment was conducted at the High Flux Isotope Reactor of Oak Ridge National Laboratory, using the HB-2 spectrometer modified for residual stress mapping.[9] The specimen was mounted upright and slits of $1 \times 30\ \mathrm{mm}^2$ and $1 \times 50\ \mathrm{mm}^2$ were inserted before and after the specimen, which together defined a sampling volume of approximately 30 mm.[3] A heating tape wrapped around the tube specimen provided controlled heating up to 250 °C. Temperatures within the tube were monitored with eight thermocouples, four on the inside and four on the outside surfaces, respectively. A rather uniform temperature field was achieved with the present experimental setup; the maximum difference in temperatures recorded by the thermocouples was less than 2 °C. Detailed strain measurements were made in the middle of the carbon steel, using the bcc (2 1 1) peak. At each location, diffraction peaks were recorded for two specimen orientations, with the scattering vector par-

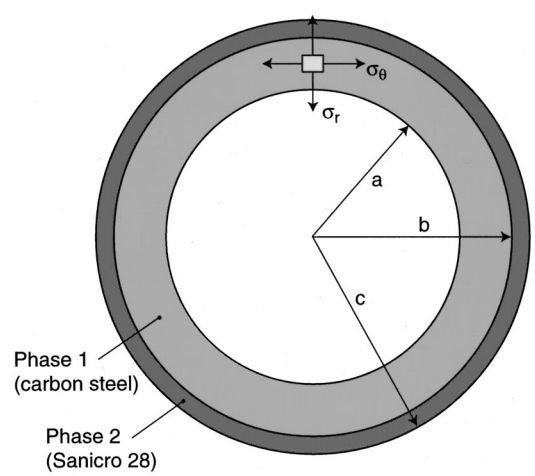


FIG. 1. Schematic of a cross-sectional view of the tube specimen used in neutron diffraction measurements.

[a]Electronic mail: wangxl@ornl.gov

 © 1999 American Institute of Physics

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

TABLE I. Coefficients of thermal expansion (α), Young's modulus (E), and Poisson's ratio (ν) for the carbon steel core[a] and Sanicro 28 clad layer.[b]

	α(30–250 °C) (10^{-6} °C^{-1})	E(30 °C) (GPa)	ν(30 °C)
Carbon steel core	13.3	207	0.29
Sanicro 28 clad layer	16.0	195	0.29

[a]Reference 7.
[b]Reference 8.

allel to the radial and tangential directions, respectively. The fitted peak position was then used to calculate the effective thermal expansion for each orientation, using the following equation:

$$\epsilon = \frac{d(T) - d(T_0)}{d(T_0)} = \frac{\sin\theta(T_0)}{\sin\theta(T)} - 1, \tag{1}$$

where d is the lattice spacing corresponding to the measured peak, 2θ is the diffraction angle, and T_0 is room temperature (~30 °C).

Figure 2 shows the experimentally determined thermal expansion for two specimen orientations, with the scattering vector parallel to radial and tangential directions, respectively. As can be seen, the effective thermal expansion varies linearly with temperature. No hysteresis was observed for cycling between room temperature and 250 °C, indicating that the thermally induced deformation is fully elastic within the temperature range of study (see Ref. 6 and also below). At a given temperature, the effective thermal expansion in the tangential direction ϵ_θ is always greater than that in the radial direction ϵ_r. This shows that the effective thermal expansion, as determined in the present neutron diffraction experiment, is macroscopically anisotropic. The effective CTE were determined by fitting the data to straight lines, and the results are $\alpha_\theta^* = 13.7\times10^{-6}\,\mathrm{C}^{-1}$ and $\alpha_r^* = 12.8\times10^{-6}\,\mathrm{C}^{-1}$. These values can be compared with the published CTE for carbon steel, which is $13.3\times10^{-6}\,\mathrm{C}^{-1}$.[7]

The difference in α_θ^* and α_r^* can be qualitatively understood by considering that the measured lattice spacing contains a temperature-dependent elastic component, i.e.,

$$d(T) = d_0(T)(1+\epsilon(T)), \tag{2}$$

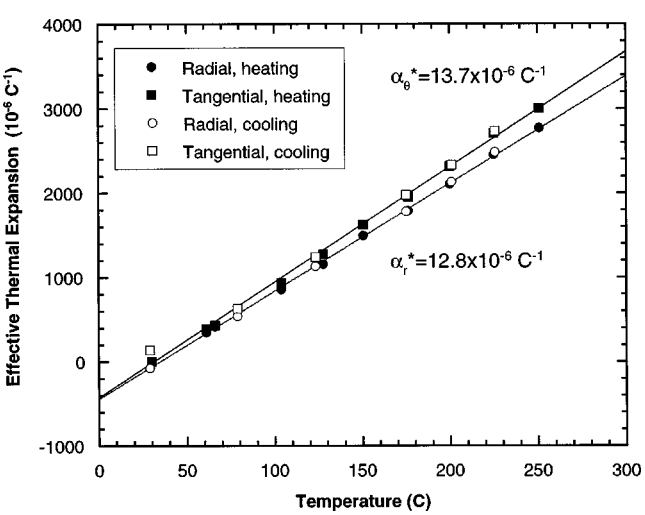


FIG. 2. Effective thermal expansion in carbon steel as a function of temperature. The solid lines are linear fits to the experimental data.

where d_0 denotes the lattice spacing under the unconstrained (stress-free) condition and $\epsilon(T)$ the residual strain at a given temperature. To facilitate the discussions, we rewrite $\epsilon(T)$ as

$$\epsilon(T) = \epsilon_0 + \epsilon'(T), \tag{3}$$

where ϵ_0 is the residual strain at room temperature, and $\epsilon'(T)$ is the thermally induced elastic strain such that $\epsilon' = 0$, when $T = T_0$. By differentiating Eq. (2), the effective CTE is obtained,

$$\alpha^* = \frac{1}{d}\frac{\partial d(T)}{\partial T} = \alpha + \frac{\partial \epsilon'}{\partial T}, \tag{4}$$

where $\alpha = (1/d_0)\partial d_0(T)/\partial T$ is the CTE under the unconstrained condition. Equation (4) shows that the effective CTE as measured in a diffraction experiment is modified by the thermal stress induced during heating. Moreover, the effective CTE is *macroscopically* anisotropic (i.e., dependent on specimen directions), unless the thermally induced residual stress is hydrostatic. For the composite tubing used in this study, $\epsilon'_\theta(T)$ in the carbon steel layer is tensile upon heating, due to the larger CTE of the clad layer, and increases with temperature. As a consequence, $\alpha_\theta^* > \alpha$. Because of the small wall thickness, the radial stress is small throughout the tube[10] and, therefore, the elastic strain in the radial direction is mostly due to the Poisson's effect. This leads to the observed $\alpha_r^* < \alpha$.

To further understand the difference between α_θ^* and α_r^*, we have carried out an analytical calculation of the thermal residual stresses in composite tubes due to thermal expansion mismatch. The analytical model is illustrated in Fig. 1, in which a cylindrical shell of phase 1 with an inner radius a and an outer radius b is surrounded by a concentric cylindrical shell of phase 2 with an outer radius c. Stresses are analyzed for locations away from the ends of the cylindrical shells and the free-end effects are not considered. The stresses can be determined by the procedure of first allowing the two phases to exhibit unconstrained thermal strains during the temperature change ΔT. Then, the radial traction σ_b is placed at the interfaces $r = b$, and axial tractions $\sigma_{1,z}$ and $\sigma_{2,z}$ are placed on phases 1 and 2, respectively, to restore the displacement continuity between the two phases. The axial tractions are subjected to the requirement that the resultant axial force be zero since no external force is applied on the system. Details of the analysis are reported elsewhere,[11] but the solution for stresses in phase 1 is given below:

$$\sigma_{1,r} = \frac{b^2(r^2-a^2)}{r^2(b^2-a^2)}\sigma_b, \tag{5a}$$

$$\sigma_{1,\theta} = \frac{b^2(r^2+a^2)}{r^2(b^2-a^2)}\sigma_b, \tag{5b}$$

$$\sigma_{1,z} = \frac{P_1 - P_3}{P_4 - P_2}\sigma_b, \tag{5c}$$

where σ_b is given by

$$\sigma_b = \frac{(P_4 - P_2)E_1(\alpha_2 - \alpha_1)\Delta T}{P_1P_4 - P_2P_3}. \tag{6}$$

In Eq. (6), $\Delta T = T - T_0$, and E and α are the Young's modulus and unconstrained CTE. The subscripts in E and α denote

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.

phases 1 and 2, respectively. P_1, P_2, P_3, and P_4, are constants related to a, b, c, and E, ν, α of both phases, where ν is the Poisson's ratio of the material. Following Hooke's law, the elastic strains in phase 1 were obtained,

$$\epsilon'_{1,j}=\frac{(1+\nu_1)\sigma_{1,j}-\nu_1(\sigma_{1,r}+\sigma_{1,\theta}+\sigma_{1,z})}{E_1}, \tag{7}$$

where $j=r$, θ, and z.

From Eqs. (5)–(7), it can be seen that the thermally induced elastic strains are proportional to σ_b, which varies linearly with T. This explains why the effective thermal expansion shown in Fig. 2 remains linear with temperature. Substituting Eq. (7) into Eq. (4) and using the parameters in Table I, the contribution by thermal stress towards the effective CTE was determined. For middepth in the carbon steel, where the neutron diffraction measurements were made, these contributions are -0.5×10^{-6} C^{-1}, 0.6×10^{-6} C^{-1}, and 0.7×10^{-6} C^{-1}, respectively, along radial, tangential, and axial directions. Adding to the published value for α_1 (see Table I), we obtain the effective CTE in the carbon steel: $\alpha'_{1,r}=12.8\times10^{-6}$ C^{-1}, $\alpha'_{1,\theta}=13.9\times10^{-6}$ C^{-1}, and $\alpha'_{1,z}=14.0\times10^{-6}$ C^{-1}. The calculated values of $\alpha^*_{1,r}$ and $\alpha^*_{1,\theta}$ are in excellent agreement with the experimental data. With a similar analysis, it can be shown that the effective CTE in the clad layer is also anisotropic, but with an inverse anisotropy in the sense that $\alpha^*_{2,z}\simeq\alpha^*_{2,\theta}<\alpha^*_{2,r}$. This prediction was confirmed by subsequent neutron diffraction measurements in the clad layer of a similarly prepared composite tube.[11]

The example discussed in this letter illustrates the influence of residual stress on the thermal expansion behavior of a material. This phenomenon is ubiquitous in all materials whose residual stresses vary with temperature. It is generally accepted that residual stresses can be classified into three categories, depending on their length scale relative to the microstructure. Macro-, or type-I, stresses are uniform across several grains. Micro-, or type-II, stresses are uniform within a grain, but vary from grain to grain. Type-III stresses are referred to those that fluctuate within a grain. The example of composite tubing is specific to macrostresses, whose signs and magnitudes usually vary with specimen directions. Thus, in general, the effective CTE in specimens containing macroresidual stresses is expected to be macroscopically anisotropic. Note that this anisotropy is characteristically different from the intrinsic anisotropy along various crystallographic directions. The latter is a fundamental property of a material and is governed by the symmetry of the crystal lattice. The macroscopic anisotropy discussed here is due to external influence and, therefore, has no crystal symmetry requirement.

Microresidual stresses also affect the thermal expansion behavior. The effect of microresidual stresses is perhaps more readily realized in a two-phase composite material, where microresidual stresses are generated due to the thermal expansion mismatch between the two phases. Analysis of the residual stress and thermal expansion data obtained for Al_2O_3–ZrO_2 ceramic composites[12] indicates that the effective CTE in each phase is modified towards a reduced mismatch in CTE between the two phases. In addition, because residual stresses in a composite material depend in part on the microstructure, the overall thermal expansion of the composite may also differ from that given by constitutive models, as illustrated by Hsueh, Becher, and Sun.[2]

The interaction of residual stress and thermal expansion in a material has significant implications in practical applications. From a materials design point of view, the reliability of finite-element models that simulate the thermal–mechanical behavior of a material depends on the accuracy of the CTE used in the models. As noted earlier, the experimentally determined CTE may be substantially different from that in the unconstrained condition. A small deviation in the CTE of one phase could lead to a large error in the calculation of the CTE mismatch, which usually determines the thermal stress in the materials system. The complexity resulting from interaction between residual stress and thermal expansion also brings about, potentially, a new method in residual stress determination. As can be seen from Eqs. (3) and (7), by comparing the experimentally determined CTE for various specimen orientations with that in the unconstrained condition, it should be possible to derive information about the residual stresses in the material. This method can become particularly useful when a stress-free specimen required in conventional neutron diffraction measurements cannot be made. Of course, to apply this method, the unconstrained CTE must be known in advance.

In summary, the influence of residual stress on thermal expansion behavior was demonstrated with a composite tube made of two layers of dissimilar steels. The presence of a temperature-dependent residual stress causes the effective coefficients of thermal expansion to be significantly different along different specimen directions. Particular care must be taken when interpreting the experimental results from thermal expansion measurements.

This work was sponsored by the Advanced Industrial Materials Program, U.S. Department of Energy. Oak Ridge National Laboratory is managed by Lockheed Martin Energy Research Corporation for the U.S. Department of Energy under Contract No. DE-AC05-96OR22464.

[1] J. F. Nye, *Physical Properties of Crystals* (Oxford University Press, London, U.K., 1985), pp. 93–109.
[2] C. H. Hsueh, P. F. Becher, and E. Y. Sun (unpublished).
[3] Y.-L. Shen, Mater. Sci. Eng., A **252**, 269 (1998).
[4] M. Hoffman, S. Skirl, W. Pompe, and J. Rodel, Acta Mater. **47**, 565 (1999).
[5] J. R. Keiser, B. Taljat, X.-L. Wang, R. W. Swindeman, P. J. Masiasz, L. E. Meyers, R. L. Thomas, S. T. Elliot, D. L. Singbeil, and R. Prescott, in *Proceedings of the 9th International Symposium on Corrosion in the Pulp and Paper Industry* (Pulp and Paper Research Institute of Canada, Vancouver, Canada, 1998), pp. 213–220.
[6] B. Taljat, T. Zacharia, X.-L. Wang, J. R. Keiser, and R. W. Swindeman, in *Proceedings of the 9th International Symposium on Corrosion in the Pulp and Paper Industry* (Pulp and Paper Research Institute of Canada, Vancouver, Canada, 1998), pp. 193–197.
[7] *Materials Handbook, Desk Edition* (ASM International, Metals Park, OH, 1992).
[8] Data sheet for Sandvik Sanicro 28 austenitic stainless alloy, S-1885-ENG (March 1997 edition), Sandvik Steel AB, Sweden.
[9] X.-L. Wang, C. R. Hubbard, S. Spooner, S. A. David, B. H. Rabin, and R. L. Williamson, Mater. Sci. Eng., A **211**, 45 (1996).
[10] X.-L. Wang, E. A. Payzant, B. Taljat, C. R. Hubbard, J. R. Keiser, and M. J. Jirinec, Mater. Sci. Eng., A **232**, 31 (1997).
[11] X.-L. Wang (unpublished).
[12] K. B. Alexander, P. F. Becher, X.-L. Wang, and C. H. Hsueh, J. Am. Ceram. Soc. **78**, 291 (1995).

Downloaded 09 Feb 2001 to 128.219.47.178. Redistribution subject to AIP copyright, see http://ojps.aip.org/aplo/aplcpyrts.html.