

Topologically Informative Residues of Large Proteins Can Be Marked By Oxidation and Mapped by Mass Spectrometry

Carlee McClintock¹, Christine Shook², Vilmos Kertesz², Gary Van Berkel², Robert L. Hettich^{1,2}

¹UT-ORNL Genome Science and Technology Program, University of Tennessee, Knoxville, TN, ²Chemical Sciences Division, Oak Ridge National Lab, Oak Ridge, TN



OVERVIEW

- Protein structure is inextricably tied to biochemical function
- High resolution techniques such as XRC and NMR exclude many proteins due to size, conformational flexibility, quantity, and purity of protein required for structural analysis
- The "omics" age is demanding higher throughput methods for deriving structural information
- Covalent labeling of proteins under native conditions coupled with mass spectrometry can provide rapid and potentially high-throughput information on the solvent accessibility of reactive residues

INTRODUCTION

- Comparative protein modeling aims to predict structures based on homology of a protein sequence to one with a resolved structure that can be used as a template.
- De novo structural prediction evaluates the physical properties of the amino acid sequence to consider local interactions that may coalesce and give rise to secondary and tertiary structures by sampling lowest energy interactions with more distant regions.
- Essential to the success of computational structure prediction is the development of biophysical constraints for evaluating each model for an unknown structure based on empirical findings.
- Hydroxyl radicals place covalent labels on reactive, solvent accessible residues of proteins that can be pinpointed by MS to convey structural information.^{3,41}
- The goal of this project is to utilize the highly reactive hydroxyl radical to stably modify surface residues of model proteins on a "dynamic timescale" and use the information to map regions exposed by native structure.
 - Serum albumin proteins, the major protein constituent of blood plasma, are large in size and are physiologically important, as they bind and transport a wide range of ligands, such as fatty acids, bilirubin, and pharmaceutical compounds.
 - Although no high resolution structure exists for bovine serum albumin, its primary structure (amino acid sequence) is very similar to human serum albumin, for which there are many crystal structures.
 - Photochemical oxidation will be used to mark reactive sites on both BSA and HSA, and the data will be used to evaluate a published homology model for BSA based on the HSA structure.
 - Electrochemical oxidation has been explored as a method for online protein digestion, which has been shown to predominantly produce intact oxidized protein⁴² - a caveat that may prove useful for probing solvent accessibility of oxidizable residues.
 - Tagging of cysteine residues with a benzoquinone probe generated at a microarray electrode has been shown to probe their environment⁴³

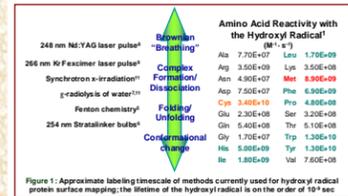


Figure 1: Approximate labeling timescale of methods currently used for hydroxyl radical protein surface mapping; the lifetime of the hydroxyl radical is on the order of 10⁻⁹ sec

EXPERIMENTAL

PHOTOCHEMICAL OXIDATION

- Model protein human serum albumin (HSA MW_{app} = 66.47 kDa) and bovine serum albumin (BSA MW_{app} = 66.43 kDa) were purchased from Sigma-Aldrich (St. Louis, MO)
- Proteins were dissolved in Tris buffer (50 mM Trizma-HCl, 10 mM CaCl₂, pH 7.6 @ 37°C) to 300 μM, then 5% to 15% hydrogen peroxide was added prior to 254 nm ultraviolet light exposure (Stratalinker, Stratagene, La Jolla, CA) to generate hydroxyl radicals
- Aliquots were taken prior to digestion and intact proteins were analyzed by direct infusion ESI-FT-ICR (Figure 2) to determine overall extent of oxidation
- Remaining sample volume was reduced and denatured in 6 M guanidine HCl / 20 mM DTT at 60°C for 1 hour, then diluted <1 M guanidine prior to addition of 1:100 trypsin
- Prior to MS analysis, samples were "cleaned up" by solid phase extraction with either C2 (intact protein) or C18 (peptide) Sep-Pak cartridges
- Tryptic peptides were analyzed by reversed-phase LC-MS/MS on Finnigan LCQ Deca XP (Figure 3) to pinpoint oxidized residues

- The DBDigger [12] algorithm searched for multiple oxidations in MS/MS spectra and matches were evaluated with a MASPIC [13] scorer to produce a DTASelect output list of peptide matches

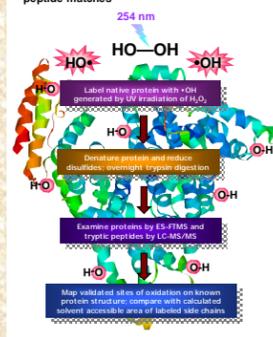


Figure 2: 9.4-Tesla FT-ICR MS (IonSpec, Lake Forest, CA) equipped with spray pump (Neward Research, Holliston, MA) delivery to the ESI source (Analytica, Branford, CT) for direct infusion of intact proteins



Figure 3: FemtoSwitch Ultimate HPLC system (Dionex, Sunnyvale, CA) coupled to a Finnigan LCQ Deca XP (Thermo, Waltham, MA) ion trap mass spectrometer

ELECTROCHEMICAL OXIDATION

- Model protein chicken egg lysozyme (MW_{app} = 14313 Da) was purchased from Sigma
- A potential of 1.0V or 1.2V was applied to a 0.8 nm porous graphite electrode (Figure 4) housed in an ESA 5021 electrochemical cell (ESA Biosciences Inc., Chelmsford, MA)
- 100 mM lysozyme solutions were passed through at a flow rate of 5 μL/min and collected for MS analysis during application of a potential of 1.0V or 1.2V
 - Non-native buffer: 47.5% ACN / 47.5% HPLC H₂O / 5% FA
 - Native buffer: PBS pH 7.5 (10 mM Na₂HPO₄, 150 mM NaCl)
- Intact proteins were analyzed by direct infusion ESI-FTICR-MS
- Tryptic peptides were analyzed by reversed-phase nanoLC-MS/MS
- Oxidation sites mapped to PDB 1AKI lysozyme crystal structure using RasMol [16]

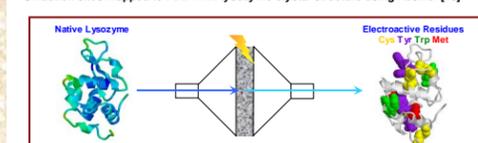


Figure 4: ESA 5021 electrochemical cell with 0.8 nm porous graphite electrode

PHOTOCHEMICAL OXIDATION RESULTS



Figure 5: Underline shows sequence coverage of functional HSA protein after light oxidation (5% peroxide 5 minutes UV exposure); oxidized residues highlighted by yellow ovals.

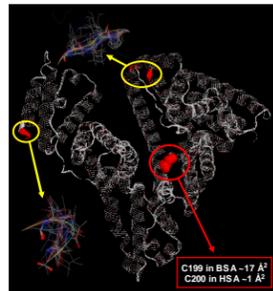


Figure 6: Homology model for BSA built with MOE software (Chemical Computing Group, Inc) using HSA template 1E7A shows high degree of structural similarity with disordered loop regions (circled above) showing the most variation. The SWISS-MODEL structure is shown above with the loop regions colored red to reference the MOE model insets, and the shifted cysteine is shown in space fill (discussed below).

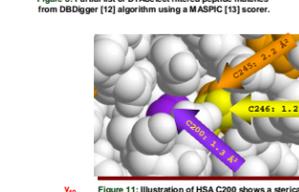


Figure 7: Partial list of DTASelect filtered peptide matches from DBDigger [12] algorithm using a MASPIC [13] scorer.

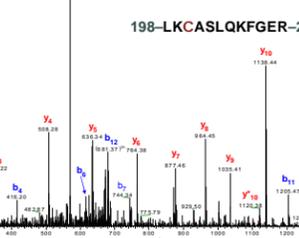


Figure 8: Tandem mass spectrum of HSA peptide showing no oxidation at C200.

HOMOLOGY MODELING

- 75% sequence identity between BSA and HSA
- Homology model generated by SWISS-MODEL pipeline based on HSA complexed with heme (PDB 1NSU)
- Dynamic regions obvious target for labeling, but differences in side chain accessibility may exist even in homologous regions
- Insertion of Y116 in HSA sequence relative to BSA sequence shifts alignment of homologous sequences - a common problem for comparative modeling
- If HSA crystal information were used directly to estimate solvent accessible area for the C200 side chain based on similarly positioned residue C199 in BSA, one would expect not to see oxidation at this site in the native structure
- However, the homology modeling method accounts for this positional shift and along with other considerations has produced a model that is supported by the covalent labeling technique detailed here
- Based on this homology model, differences in side chain orientation between HSA and BSA are predicted even for regions of strong sequence and structural similarity



Figure 9: Underline shows sequence coverage of functional BSA protein after light oxidation (5% peroxide 5 minutes UV exposure); yellow ovals indicate oxidized residues.



Figure 10: Partial list of DTASelect filtered peptide matches from DBDigger [12] algorithm using a MASPIC [13] scorer.

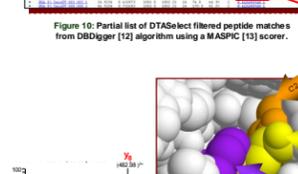


Figure 11: Illustration of HSA C200 showing a sterically hindered pocket relative to the region accessible in BSA (Figure 12).

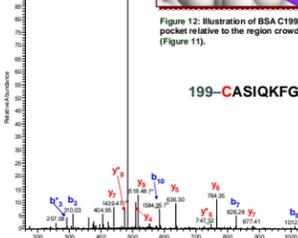


Figure 12: Tandem mass spectrum of BSA peptide showing triply oxidized C199.

ELECTROCHEMICAL OXIDATION

- Electrochemical oxidation has been explored for the online digestion of proteins prior to MS [14]
- Electroactive amino acids include cysteine (C), methionine (M), tyrosine (Y), and tryptophan (W)
- Application of 1-1.2V potential produced oxidations at C, M, Y, W
- Observed oxidation of phenylalanine (F) may suggest production of hydroxyl radicals in electrochemical cell

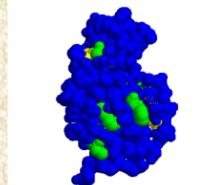


Figure 13: Lysozyme 1AKI shown in space fill model [16] without hydrogens; green residues are more heavily oxidized and yellow residues are slightly oxidized as noted in Figure 15.

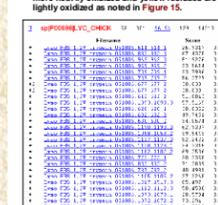


Figure 14: Mass spectrum of electrochemically oxidized intact lysozyme. Shift of +1 m/z for each charge state relative to native m/z peaks (listed below) suggesting the mono-oxidized species is most abundant. Inset shows deconvoluted spectrum of native protein mass (N) with successive oxidations as m(NO); note base peak at N-(NO).

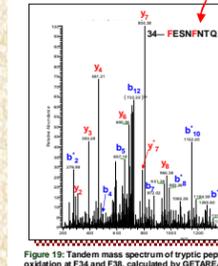


Figure 15: Tandem mass spectrum of partially tryptic peptide with oxidation at F34 and F38, calculated by GETAREA [2] to expose 51.86 Å² and 14.37 Å² of the sidechains, respectively.

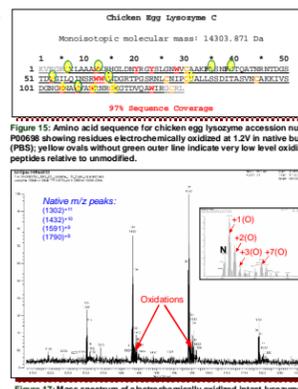


Figure 16: Tandem mass spectrum of partially tryptic peptide with oxidation at C6, calculated by GETAREA [2] to expose 59.55 Å² of the sidechain.

DISCUSSION

- Cysteines are consistent probes for solvent accessibility
 - Oxidation status of buried cysteines in native structures of model proteins can be used to gauge appropriate oxidation conditions for maintaining native structure in comparably-sized proteins
 - Incomplete disulfide oxidation in control implies some degree of remaining tertiary structure that protects buried cysteines/cysteines despite oxidation-induced unfolding
- Photochemical oxidation results are relatively consistent with the BSA homology model
 - Further studies will be conducted to optimize both native and control oxidation conditions for larger proteins
- Scoring decisions in manual validation of oxidation results
 - Filtering of peptide matches considers:
 - Score: Compares theoretical MS/MS spectrum to observed MS/MS spectrum for each peptide charge state (+1.24, +2.28, +3.46)
 - DeltaCN: Difference calculated for the best and second best peptide matches to evaluate likelihood of a confident match
 - Tandem mass spectra with multiple theoretical matches passing the filtering cutoffs were manually evaluated with values calculated for z and y fragment ions from each matched peptide
 - Search filters should be evaluated to guard against false positives that introduce inaccurate structural information
- A tandem mass spectrum may match multiple peptides
 - Interpretation of tandem mass spectra can be complicated by isoforms with overlapping retention times
 - Multiple potential oxidation sites near peptide termini often results in a fragment ion series insufficient to confidently identify the oxidized site
- Instrument capable of faster scanning and greater dynamic range will be used to improve gaps in sequence coverage
 - Faster full scans and MS/MS scans
 - Narrow mass window for trapping and CID

REFERENCES

- Buxton GV, Greenstock CL, Helman WP, Ross AB (1988) *Journal of Physical Chemistry Reference Data* 17: 513-886
- Fraczkiewicz R, Braun W (1998) *Journal of Computational Chemistry*, 19(3): 319-333
- Kislar JD, Maleknia SD, Sullivan M, Downard KM, Chance MR (2002) *International Journal of Radiation Biology*, 78(2): 101-114
- Maleknia SD, Kislar JD, Downard KM (2002) *Rapid Communications in Mass Spectrometry*, 16(1): 53-61
- Sharp JS, Becker JM, Hettich RL (2003) *Analytical Biochemistry*, 313: 216-225
- Sharp JS, Becker JM, Hettich RL (2004) *Analytical Chemistry*, 76: 672-683
- Sharp JS, Sullivan DM, Cavanagh J, Tomer KB (2006) *Biochemistry*, 45(20): 6260-6265
- Aye TT, Low TY, Sze SK (2005) *Analytical Chemistry*, 77: 5814-5822
- Hambly DM, Gross ML (2005) *Journal of the American Society for Mass Spectrometry*, 16(12): 2057-2063
- Xu G, Chance MR (2005) *Analytical Chemistry*, 77(14): 4549-4555
- Xu G, Chance MR (2005) *Analytical Chemistry*, 77(18): 2437-2440
- Tabb DL, Narasimhan C, Strader MB, Hettich RL (2005) *Analytical Chemistry*, 77: 2464-2474
- Narasimhan C, Tabb DL, VerBerkmoes NC, Thompson MR, Hettich RL, Uberbacher EC (2005) *Analytical Chemistry*, 77: 7581-7593
- Perreniwé HP, Bruins AP (2004) *Journal of the American Society for Mass Spectrometry*, 15: 1707-1716
- Dayon L, Roussel C, Girault HH (2006) *Journal of Proteome Research*, 5(4): 793-800
- Sayle RA, Milner-White EJ (1995) *Trends Biochem. Sci.*, 20(9):374-376

ACKNOWLEDGMENTS

This research was sponsored by the National Institutes of Health under 1R01-GM070754. Oak Ridge National Laboratory, managed by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. Carlee McClintock gratefully acknowledges support from the UT-ORNL Genome Science and Technology Program