

In-silico Prediction of Surface Residue Clusters for Enzyme-Substrate Specificity of highly homologous enzymes

Gong-Xin Yu¹, Byung-Hoon Park¹, Praveen Chandramohan, Al Geist, and Nagiza F. Samatova²

¹ Both authors have contributed equally to this work

² Corresponding author: samatovan@ornl.gov

Introduction

One of the most remarkable properties of enzyme substrate binding is high specificity. Several possible mechanisms have been suggested to illustrate such exquisite substrate specificity, e.g. substrate-binding in the catalytic centers of enzymes, loop-based hinge-motion, and cofactor binding and intra- or inter-molecule (domain-domain) interactions. We believe that there must be sets of *specificity-determining residues* (i.e. clusters of amino acids whose structural, dynamic and physico-chemical properties directly or indirectly affect interaction and transformation) that enable different enzymes to recognize their unique substrates.

Objective

This study aims to develop a surface patch ranking (SPR) to identify co-evolved surface residue clusters, called *Specificity-Determining Residue (SDR)* clusters, for substrate specificity determination among homologous enzymes.

Method

Surface patch ranking (SPR) for finding SDR clusters

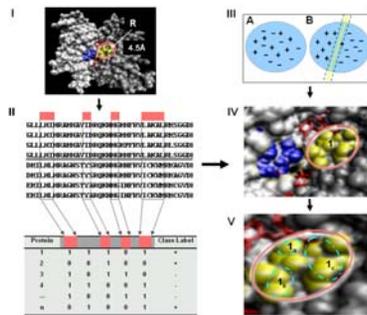


Figure 1. Basic steps of SPR method. I. Construct surface patches for each surface residue. II. Build surface patch-specific matrix. III. Assess each surface patch by leave-one-out error. IV. Identify putative functional surface patches (IV.1). V. Locate all minimal groups of residues as SDR clusters within the putative functional surface patches that are best able to discriminate functional sub-types and are statistically significant (V1a, b, c).

Residue categorization: For RuBisCo CO₂/O₂ specificity, we further categorize how an SDR residue within SDR clusters changes among cyano-bacteria, green plants and marine non-green algae. The aim is to study when the residues mutate during the RuBisCo evolution and how they correspond to the changes of specificity.

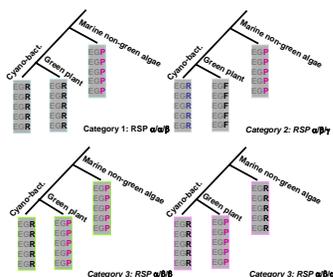


Figure 2. Residue Substitution Patterns (RSP) of residues within SDR clusters. Four different RSPs can be found (RSP $\alpha/\alpha/\alpha$, $\alpha/\alpha/\beta$, $\alpha/\beta/\alpha$, and $\alpha/\beta/\beta$) based on the Substitution Patterns among marine cyano-bacterial, green plant and marine non-green algal RuBisCo.

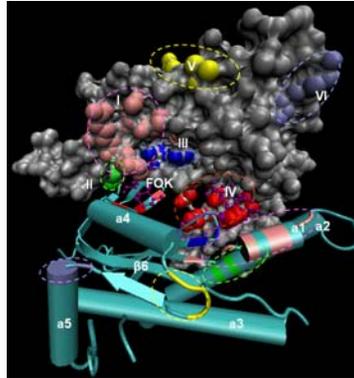


Figure 3. Mapping of SDR clusters on 1ab8, an adenyl cyclase from *Rattus norvegicus*. Six residue groups are mapped and each corresponds to surface regions of different functions, I: interface regions with forskolin, magnesium, adenine and ribose; II: either residues of the VC1:IIVC2 interface or interface with an activator Gsa protein or both; III: the residues that bind with forskolin and Magnesium and IV: ligand binding (Ribose, pyrophosphate) and substrate specificity.

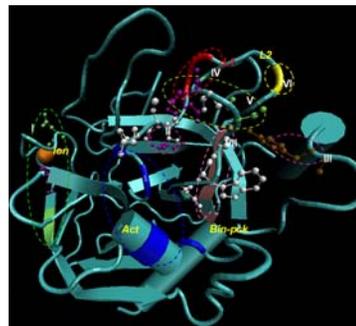


Figure 4. Mapping of SDR clusters on 5tpa, a Beta trypsin from *Bos taurus*. The residues in single-residue SDR clusters are colored in purple (I, II, IV), and residues in multi-residue SDR clusters are colored in green (I, V), ochre (III), red (IV), yellow (VI), and pink (VII), respectively. The catalytic residues are colored in blue, ion in orange and water in pink

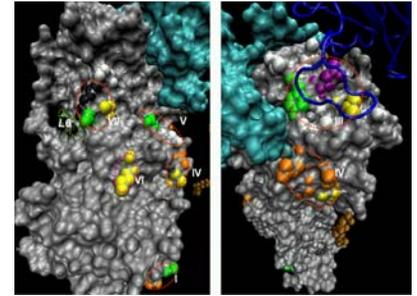


Figure 5. Mapping of the CO₂/O₂ specificity-determining residues on 1RCO a large subunit of RuBisCo protein from *Spinacia oleracea*. A total of seven residue clusters are mapped on the large subunit and residues within the clusters are painted in different colors according to their residue substitution pattern (RSP): orange for RSP $\alpha/\beta/\alpha$; yellow for non-consensus RSP $\alpha/\beta/\alpha$; green for RSP $\alpha/\beta/\gamma$; black for RSP $\alpha/\alpha/\beta$; white for non-consensus RSP $\alpha/\alpha/\beta$; purple for RSP $\alpha/\beta/\beta$.

Results

1. The SDR clusters include both highly conserved residues and non-conserved yet complementary residues.
2. Some of the identified SDR clusters, primarily the mono-residue ones, represent residues that are directly involved in enzyme-substrate interactions. Others, mostly the multi-residue ones, represent residues vital for domain-domain and regulator-enzyme interactions.
3. Specificity-determining residues within the SDR clusters of RuBisCos are further categorized. Each category associates uniquely with given RuBisCos groups, their evolutionary history and substrate specificity levels.

Conclusions

1. The SPR algorithm has the ability to identify co-evolved surface residue clusters for substrate specificity determinations.
2. The RSP-based RuBisCo analysis enable us to associate CO₂/O₂ specificity-determining residues within the identified SDR clusters with the RuBisCo evolutionary history and specificity levels, thus, providing additional information for the selection of residues and SDR clusters for a successful site-directed mutagenesis.

Acknowledgements

This work was funded in part or in full by the US Department of Energy's Genomes to Life program under project, "Carbon Sequestration in *Synechococcus* Sp.": From Molecular Machines to Hierarchical Modeling. The work of N.F.S. was sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory. This research used resources of the Center for Computational Sciences and the eXtreme TORC cluster at Oak Ridge National Laboratory, contractor of the U.S. Government under Contract No. DE-AC05-00OR22725. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.