Abstract
In 2000, the Protein Structure Initiative (PSI) was started as a way to determine three-dimensional structures of proteins within a given family. Once solved, structures are deposited into the Protein Data Bank (PDB) and termed Structural Genomics (SG) proteins. Today, there are over 13,000 SG proteins in the PDB, but most lack a reliable functional assignment. Most of the SG proteins with assigned functions have only putative function based on global structural similarity with proteins of known function. One major challenge with putative functional assignments is that they are often incorrect. The Crotonase Superfamily (CS) contains five functionally different subfamilies and at least 70 SG proteins. Our approach is based on local structure matching at the computationally predicted active site. First, Partial Order Optimum Likelihood (POOL) and Structurally Aligned Local Sites of Activity (SALSA) are used to predict the local active site residues and compare them to the known members in the superfamily. We demonstrate in this analysis that the majority of the putative annotations in this superfamily are likely incorrect. Next, biochemical assays are being used to test these predictions. Preliminary results of biochemical assays show that one SG protein, classified as a probable enoyl-CoA hydratase (ECH), possesses hydratase activity as predicted by our methods. The main goals of this project are to classify SG proteins successfully based on their local structure at the predicted active sites and to provide a conceptual framework for the functional classification of the remaining SG proteins in the PDB.

Background
The Crotonase Superfamily
- Enzymes catalyze variety of metabolic reactions
- Share ability to stabilize the enolate anion internally
- Founding member: enoyl-CoA hydratase from Rattus norvegicus (PDB ID 1DUB)
- Conserved quaternary structure: dimer of trimers (each monomer shown in a different color)
- Monomer structure: three α/β repeating units fold so the pairs of β-sheets are almost perpendicular to each other
- Functional subgroups include: hydratase, isomerase, hydratase, and dehalogenase activities.
- Potential applications: agriculture, environmental remediation, medicine, and green chemistry

Results: SG protein “probable enoyl-CoA hydratase” from T. thermophilus
After computational analysis, SG protein tested for hydrolase-like activity:

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>$k_{cat}$ (mM)</th>
<th>$V_{max}$ (mM/sec)</th>
<th>$k_{cat}/V_{max}$ (sec^-1)</th>
<th>cat. eff. (mM sec^-1)</th>
<th>std. dev. (cat. eff.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2J5S</td>
<td>2.0</td>
<td>0.000078</td>
<td>0.053</td>
<td>0.026</td>
<td>0.0052</td>
</tr>
<tr>
<td>1W2B (SG)</td>
<td>7.1</td>
<td>0.000644</td>
<td>0.47</td>
<td>0.086</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Our method correctly predicted the function of this SG protein, showing that the originally reported annotation is incorrect.

References:

Source: Structural Genomics Proteins in the Crotonase Superfamily

Caitlyn L. Mills, Mary Jo Ondrechen, Penny J. Beuning
Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA 02115

Funding & Acknowledgements:
Northeastern University
Graduate Category: Physical and Life Sciences
Degree Level: Ph.D.
Abstract ID #627

Pool Server: www.pool.neu.edu