Introduction to Protein Structure through Homology Modeling Using Bioinformatics

Woon-Kai Low, PhD
Associate Professor
St. John’s University, College of Pharmacy & Health Sciences
Queens, NY, 11439
loww@stjohns.edu
Molecular Toxicogenomics

- 3 credits (one semester)
- Associated lab (1 credit, 3 hours/week)
- Students:
  - 3rd year toxicology majors (BS)
  - 2nd year Clinical Laboratory Sciences (BS)
Molecular Toxicogenomics (3 credits, one semester)

- Introduction to the molecular and genetic basis for human diversity
- Impact on toxicological and pharmacological response
- Scientific basis of personalized (precision medicine)
Lab #2: Protein Structure and Homology

- Introduction (and review) of amino acids, and the four levels of protein structure
- Introduction to the concept of homology
- Introduction to bioinformatics

Image Source: Getty Images
Resources (Part II):

RCSB PDB protein data bank
http://www.rcsb.org/pdb/home/home.do

Universal Protein Resource (UniProt)
http://www.uniprot.org/

Clustal Omega (sequence comparison software)
http://www.ebi.ac.uk/Tools/msa/clustalo/

NCBI National Center for Biotechnology Information

BLAST (Basic local alignment search tool)

© SOT 2018
Student Activities

I. PDB entry for structure of arsenic methyltransferase (*Cyanidioschyzon* sp.)

II. UniProKB entries for arsenic methyltransferase

III. Sequence alignment of arsenic methyltransferase from different species using Clustal Omega

IV. NCBI BLAST using arsenic methyltransferase

- Students maintain record of activities in lab notebook
I. Protein Data Bank

- Search in PDB “arsenic methyltransferase”

http://www.rcsb.org/pdb/home/home.do
I. Protein Data Bank

- Results – no structure for human protein (2 hits associated with Homo sapiens are related proteins)

Search Parameter:

Text Search for: arsenic methyltransferase

Refinements

<table>
<thead>
<tr>
<th>ORGANISM</th>
<th>UNIPROT MOLECULE NAME</th>
<th>TAXONOMY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanidioschyzon sp. 5508</td>
<td>Arsenic methyltransferase (8)</td>
<td>Eukaryota only (11)</td>
</tr>
<tr>
<td>Homo sapiens (2)</td>
<td>Histidine protein methyltransferase (1)</td>
<td></td>
</tr>
<tr>
<td>Chlamydomonas reinhardtii</td>
<td>Arsenite methyltransferase (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histone-lysine N-methyltransferase (1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Refine Query</td>
<td></td>
</tr>
</tbody>
</table>

http://www.rcsb.org/pdb/home/home.do
3 search hits from *Cyanidioschyzon* sp. (algae)

### 4FR0
ArsM arsenic(III) S-adenosylmethionine methyltransferase with SAM

*Ajees, A. A., Marapakaia, K., Packianathan, C., Sankaran, B., Rosen, B. P.*

(2012) *Biochemistry* 51:5476-5485

- **Macromolecule:** Arsenic methyltransferase (protein)
- **Unique Ligands:** SAM

### 4FS8
The structure of an As(III) S-adenosylmethionine methyltransferase: insights into the mechanism of arsenic biotransformation

*Ajees, A. A., Marapakaia, K., Packianathan, C., Sankaran, B., Rosen, B. P.*

(2012) *Biochemistry* 51:5476-5485

- **Macromolecule:** Arsenic methyltransferase (protein)
- **Unique Ligands:** CA

### 4FSD
ArsM arsenic(III) S-adenosylmethionine methyltransferase with As(III)

*Ajees, A. A., Marapakaia, K., Packianathan, C., Sankaran, B., Rosen, B. P.*

(2012) *Biochemistry* 51:5476-5485

- **Macromolecule:** Arsenic methyltransferase (protein)
- **Unique Ligands:** ARS, CA, CL

http://www.rcsb.org/pdb/results/results.do?tabtoshow=Current&qrid=37EECA42

Explore entry 4FSD (with arsenic ligand)
• 3D view – identify residues involved in binding

[Image of a 3D molecular structure with labeled residues]

http://www.rcsb.org/pdb/ngl/ngl.do?pdbid=4FSD&bionumber=1
Sequence view

Chain A: Arsenic methyltransferase

Chain Info
Polymer: 1
Length: 333 residues
Chain Type: polypeptide(L)
Reference: UniProtKB (COV/19)
Up-to-date UniProt IDs are provided by the SIFTS project

Display Parameters
Currently displayed
SEGRES sequences
Display external SEQUEST (sequence)

Sequence Chain View

http://www.rcsb.org/pdb/explore/remediatedSequence.do?structureId=4FSD
II. Universal Protein Knowledgebase

- Explore entries for arsenic methyltransferase from human, rat and mouse.
- Collect sequences for each in FASTA format.

http://www.uniprot.org/uniprot/?query=AS3MT&sort=score
III. Clustal Sequence Alignment

- Using AS3MT sequences from human, rat, mouse and Cyanidioschyzon sp.
Note conservation surrounding binding pocket residues

Residues involved in function (co-ordination of ligands, arsenic)
IV. BLAST

[Image of BLAST webpage]

Explore the BLAST results

Post-Lab Activities

- Final lab report – questions
- Bioinformatics – databases and their utility (PDB vs. UniProtKB)
- Homology – evolutionary conservation of amino acid sequence and protein function
- Clustal alignment results vs. BLAST search results