Exercise #1: Protein Data Bank Structural Analysis Exercise

Learning Objective: To be able to locate and analyze atomic coordinates deposited in the Protein Data Bank (PDB) resulting from the publication of a structure determination of a protein-inhibitor complex.

Approach and Analysis Queries:

1. Search for and obtain an electronic copy of the following publication:

***In situ* extension as an approach for identifying novel alpha-amylase inhibitors** by S. Numao, I. Damager, C. Li, T.M. Wrodnigg, A. Begum, C.M. Overall, G.D. Brayer, and S.G. Withers; J. Biol. Chem. 279, 48282-48291 (2004).

Hint: Articles of this type are free and easily accessed from the UBC library website or the PuBMed website, if using a computer at UBC.

2. Of the three protein-inhibitor complex structures described in this publication, determine what the Protein Data Bank (PDB) code is for the MeG2-GHIL/ human pancreatic alpha-amylase complex.

Hint: Start with an author search of the Protein Data Bank website.

Answer: The PDB code for this structure is: 1U33

3. Download the atomic coordinate file for the MeG2-GHIL/ human alpha-amylase complex from the Protein Data Bank website. By looking at this file in any convenient text reader program (such as MS Word), determine the following:

i) Total number of atoms in this structure determination? 4265
ii) Total number of protein atoms present (i.e. in amino acids)? 3945
iii) Number of water molecules in this structure? 268
iv) What is the N-terminal residue? A modified glutamine, pyroglutamic acid, labelled in the coordinate file as PCA. It turns out this is a common modification when glutamine is found at the N-terminus of a polypeptide chain.
v) What is the C-terminal residue? Leucine
vi) How many atoms are in the bound inhibitor? 36
vii) Why are there no hydrogen atoms in this coordinate file? Protein crystallography rarely has the resolution necessary to define hydrogen atom positions. However, knowing the positions of the larger atoms means that the hydrogen atom positions could be generated, if necessary, from well documented geometric bonding criteria.

4. Load the downloaded PDB coordinate file into the molecular graphics software you are using (e.g. PyMOL, RasMol, etc.) and determine the following structural information. **Hint:** See the graphics commands and menu items you have previously used as a guide to help in visualizing this enzyme-inhibitor complex structure.

A. This enzyme binds a calcium ion for stability (listed as residue 498; see PDB file).

i) What specific interactions bind this ion to the protein?

His 201 O - 2.4 Angstroms
Arg 158 O - 2.4 Angstroms
Asn 100 OD1 - 2.4 Angstroms
Asp 167 OD1/ OD2 - 2.5/ 2.5 Angstroms

ii) What are the interaction distances? (See above)
iii) Are there other non-protein interactions to the bound calcium ion?

Yes - it is not uncommon to see water molecules form interactions to bound ions.

Water #539 - 2.5 Angstroms
Water #622 - 2.5 Angstroms

Hint: If using PyMOL you might want to start by using the following commands to focus in on the relevant area of the enzyme-inhibitor complex:

select near 498, resi 498 around 4.0
select 498, resi 498
(Note: a "sphere" representation of atoms can often be very useful.)

B. A non-protein group (listed as residue 497) is found **covalently** bound to this form of human pancreatic alpha-amylase.

i) What is this non-protein group?

This form of the enzyme is glycosylated with a bound N-acetyl-D-glucosamine (often designated as a NAG group).

ii) To what is this non-protein group bound?

A covalent bond is made from Asn 461 ND2 to the C1 carbon of the NAG (bond distance - 1.43 Angstroms).

iii) Why does this non-protein group make so few interactions with other protein groups?

Its surface location, small size and relatively rigid structure, makes it difficult to reach other protential interacting groups on the protein.

C. A chloride ion (consult the downloaded PDB file for the residue number) is also bound in a separate binding pocket near the active site of human pancreatic alpha-amylase and it turns out that this ion is required for efficient enzymatic catalysis.

i) What protein interactions hold this chloride ion in position?

A search of the coordinate file shows that this chloride ion is labelled residue 499.
The primary interactions holding this ion in place are:

N298 ND2 - 3.4 Angstroms
R337 NH1 - 3.3 Angstroms
R337 NH2 - 3.1 Angstroms
R195 NH2 - 3.4 Angstroms
R195 NE - 3.2 Angstroms

ii) What other non-protein group is found in this binding site?

A water molecule (#508) interacts with the chloride ion (distance = 3.3 Angstroms).

D. The chloride ion discussed above, is bound towards one end of a pronounced structural feature, which forms most of 1 of the 3 structural domains of this enzyme. **HINT:** For the following questions represent the polypeptide chain with a "cartoon" or "ribbon" like mode and the chloride ion in a "sphere" mode.

i) What is this pronounced structural feature?

An 8-stranded parallel beta-barrel.

ii) How would you classify the domain that this structural feature is part of? Subgroup?

This domain would be designated a 'Parallel alpha/beta domain' and the subgroup would be 'Singly wound parallel beta-barrels'.

E. If you specifically color the bound inhibitor that is part of this structural analysis (numbered as residue 500) and use the "sphere" like mode, you will see how tightly it binds to the enzyme surface in the active site region. This modified carbohydrate has an N-O group at one end and a terminal methyl group at the other.

i) Of the three catalytic residues of human pancreatic alpha-amylase (D197, E233 and D300), which one is closest to the terminal N-O group? What is the shortest interaction distance?

The inhibitor N-O group is closest to E233 OE1 - 2.5 Angstroms.

ii) What is the shortest hydrophobic protein contact to the terminal methyl group?

The three shortest contacts to this methyl group are:

A106 CB - 5.5 Angstroms
G164 CA - 5.5 Angstroms
T163 C - 4.9 Angstroms

Having successfully completed this exercise, you now have the tools to look at any of the other structures in the Protein Data Bank.

LOOK UP AN ENZYME OF PARTICULAR INTEREST TO YOU AND EXPLORE!!

Protein Structure Rocks!!!