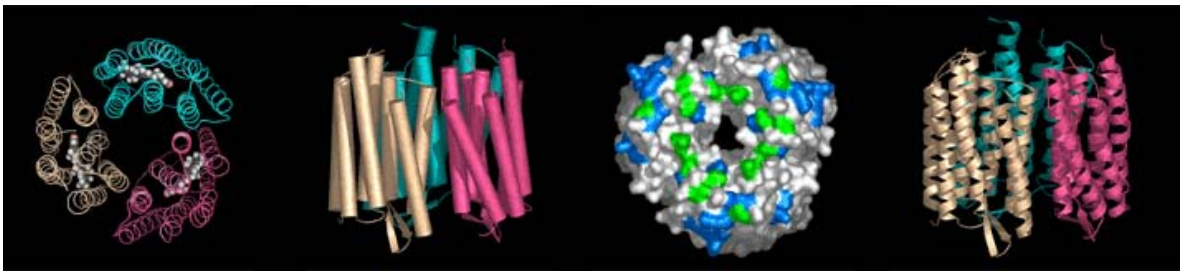
Experiment No. 2

Protein Structure Analysis

Introduction

PyMOL is a powerful molecular visualization package from DeLano Scientific, LLC. PyMOL is increasingly being used for very advanced molecular graphics applications in research laboratories and in the pharmaceutical industry. You will use PyMOL to view molecular structures and to make distance measurements.



Display a Molecule

PyMOL reads .pdb files, which come from the Protein Data Bank (PDB). These files are used to describe the 3D structure of any molecule, not just proteins but also nucleic acids (like RNA and DNA) and even small molecules.

Click on Start to go the Programs menu where you will find PyMOL. Start the program by double clicking on the icon. TWO windows will open up – one with a command line prompt and the other a viewer window.

PyMOL allows the user to easily manipulate the orientation of the molecule with respect to the viewer. The easiest way is to use the mouse to do this. Move the molecule around to see how the structure rotates in the window.

Distance Measurements (To find the distance use the measurement wizard)

1. Access the wizard by clicking on the Wizard menu and then choose the Measurement option.
2. You will be prompted in the upper left corner of the molecule-display region to select the first atom of interest. To select an atom, as shown in the Mouse region of the display, left click on the atom in the molecular display part of the window. When you have successfully selected an atom, a little pink square will appear on that atom.
3. Next choose the second atom. Left-click on another atom in the molecule to complete the measurement.
4. You will see a yellow dashed vector drawn between the two atoms you selected and a measurement in Angstroms drawn above it.

The view menu lists the various objects that are displayed at a given time. In our display at present there is only one molecule, so there is just one named object.

The view menu shows five popup menus to the right of each listed object: In order, let's call these the A, S, H, L, and C(rainbow) menus. If you left-click on these and drag down, you will be able to choose various options from the pop-up menu that comes up.

- The A (Action) menu allows the user to control the view of the molecule. The main option you might use is zoom, which resets the size of the molecule to fit in the display
- The S (Show) menu shows different modes of viewing a molecule. The lines and spheres options are the ones that we will use at first.
- The H (Hide) menu hides the different modes. So if you have both lines and spheres modes turned on (using the S menu), clicking H and then selecting one of the modes turns one of them off.
- The L (Labels) menu controls different ways of labeling the molecule.
- The C (Color/rainbow) menu chooses different styles of coloring the different elements in the molecule.

You will be given a protein file (PDB file) e.g.1CRN

1. Determine the number of each amino acid type present in the protein and record your results in a table. Calculate the percentage occurrence of each residue type in the protein.

   To select a specific residue name in the command line - write:

   - Select resn ALA (if you want to select ALA residues only)
   - You can then color them differently to help you count them more easily – use the C menu on the right side panel to color it.

2. Identify the helices and sheets present in the protein. Determine the percentage of residues present as helices and sheets and calculate the percentage present as random structure.
3. Calculate the distances from the N-terminus to the C-terminus of each helix and determine the pitch
4. Calculate the distances from the N-terminus to the C-terminus of each strand and determine the average distance between the C_α atoms for each strand
5. Construct a helical wheel of the helices present to determine the amphipathic nature of the helix if present.
6. From the number of aromatic residues present in the protein estimate the extinction coefficient of the protein.