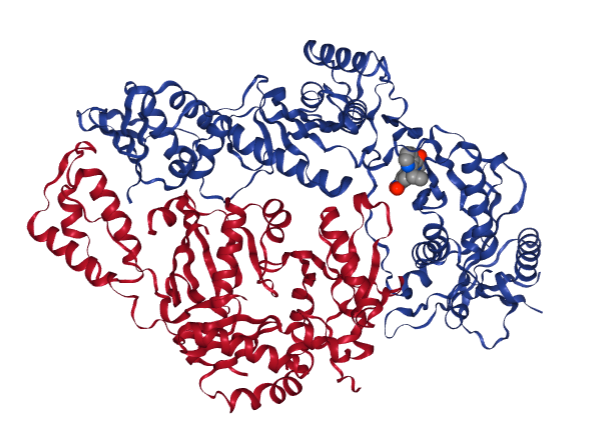
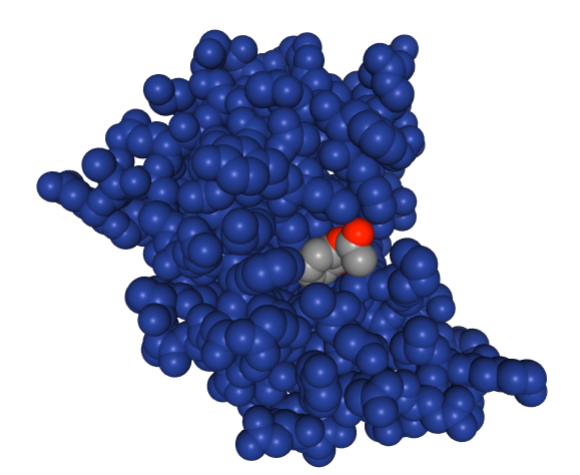
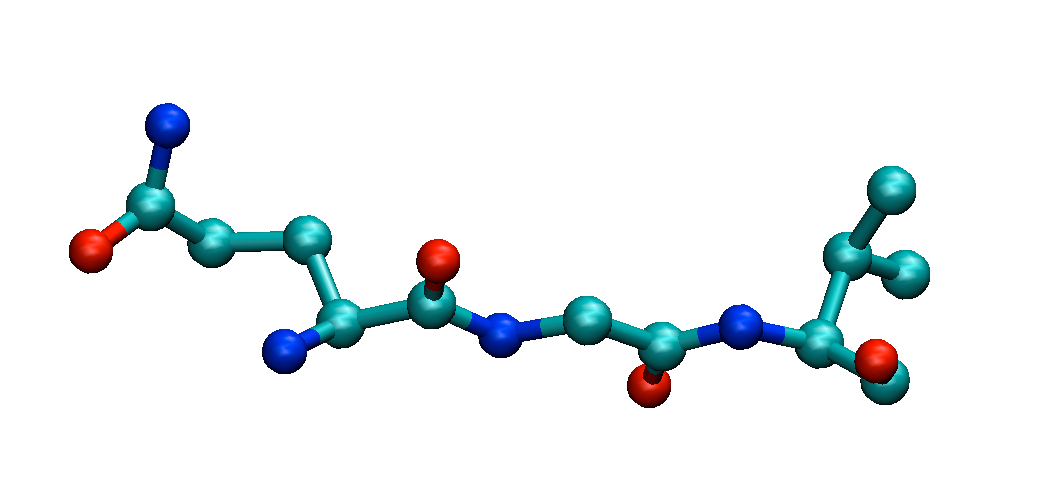
**Supplemental Materials**

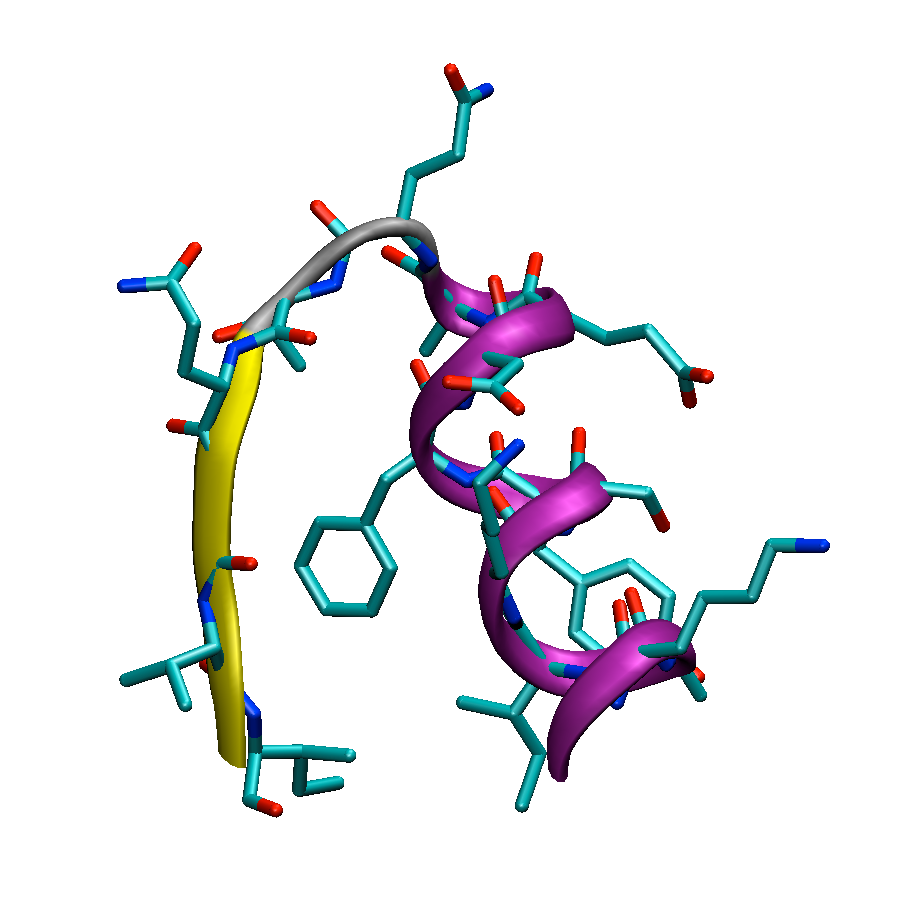
**I. Color figures**

**Figure 2A**

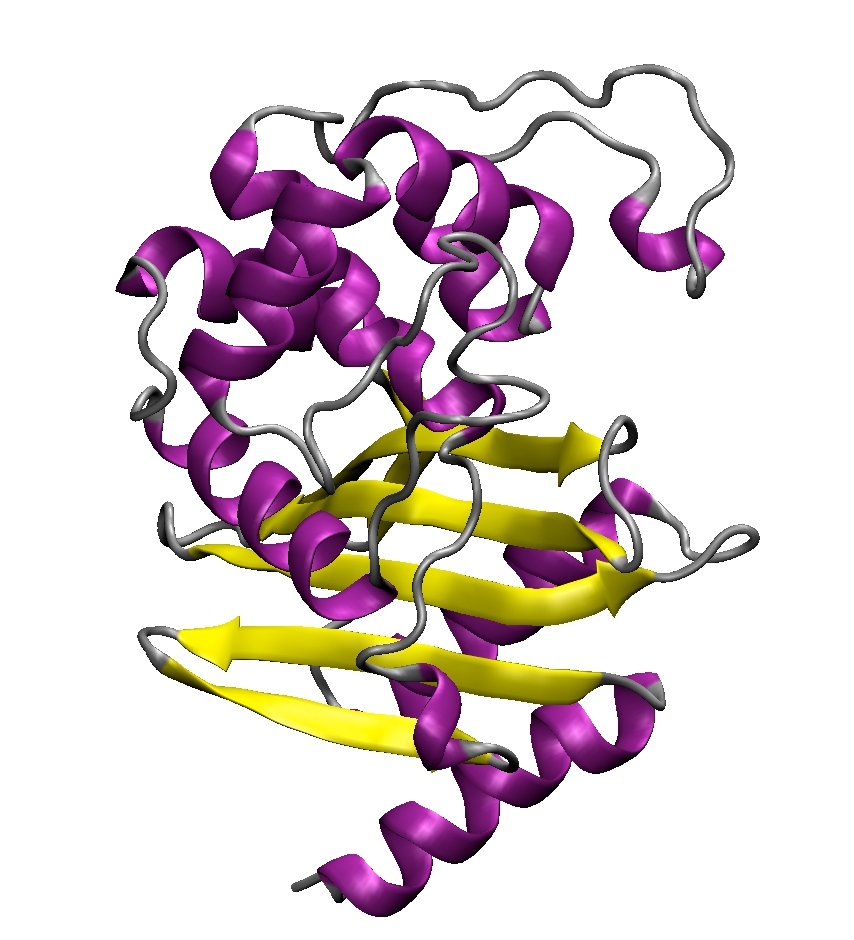
**Figure 2B**

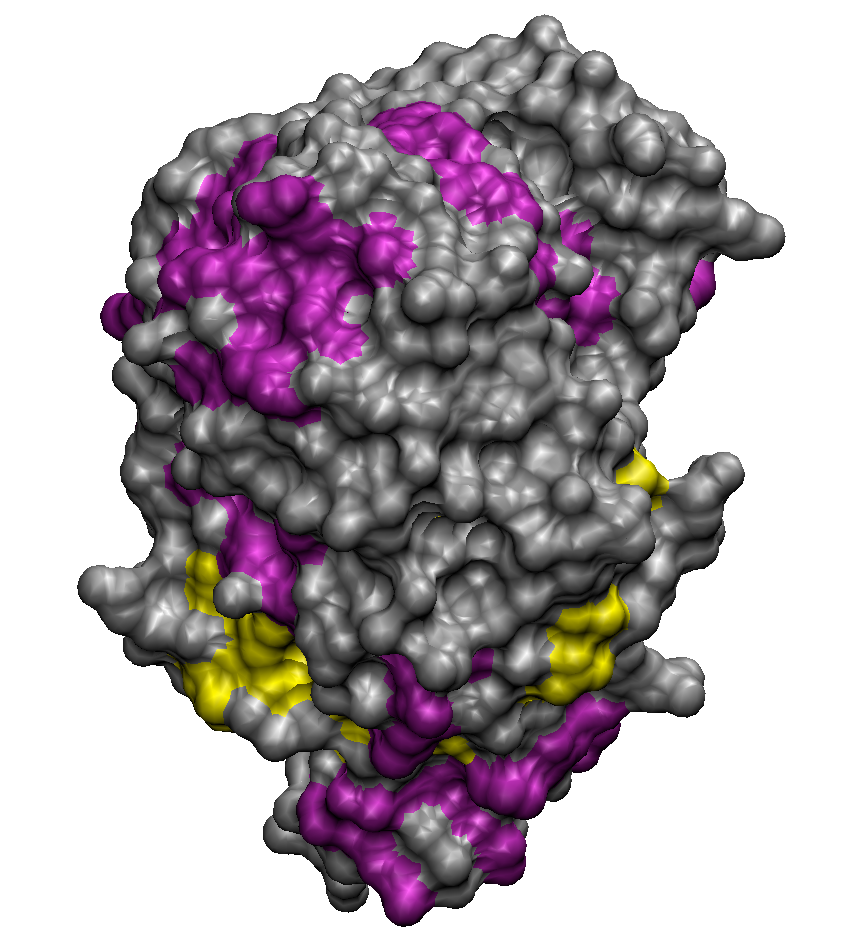


Figure 3A

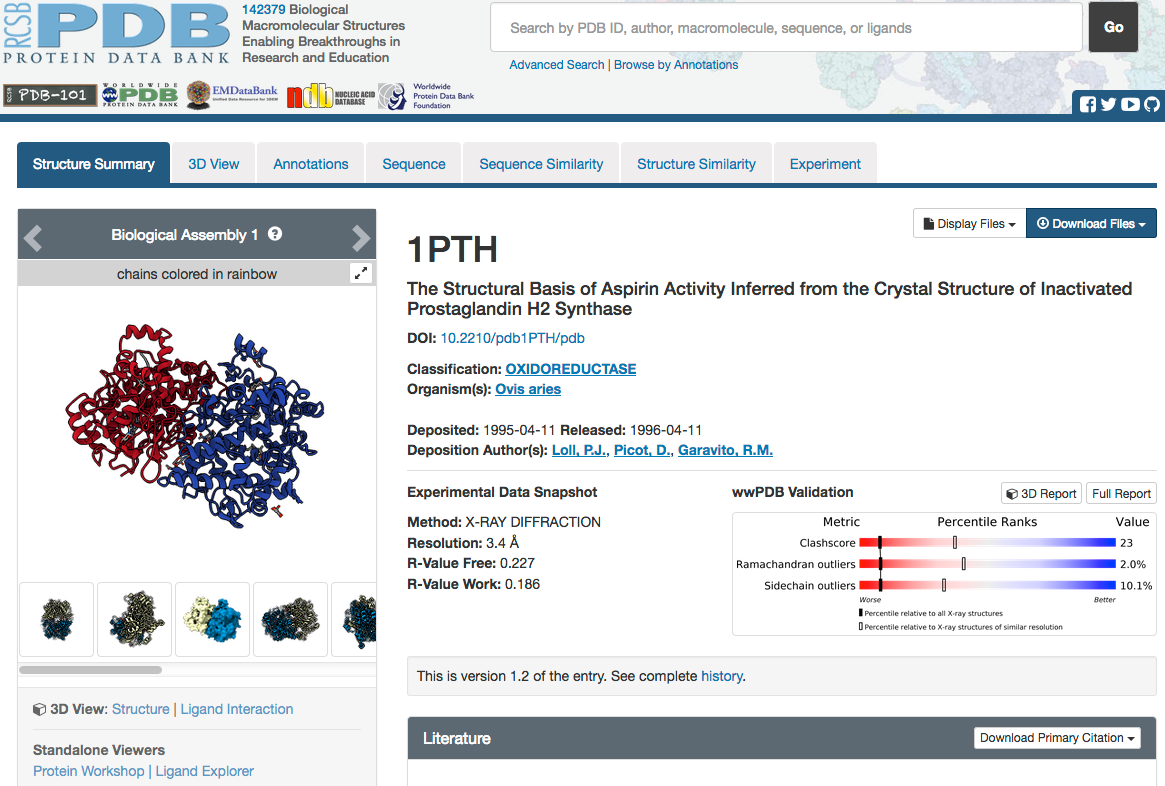
Figure 3B

**Figure 3C**

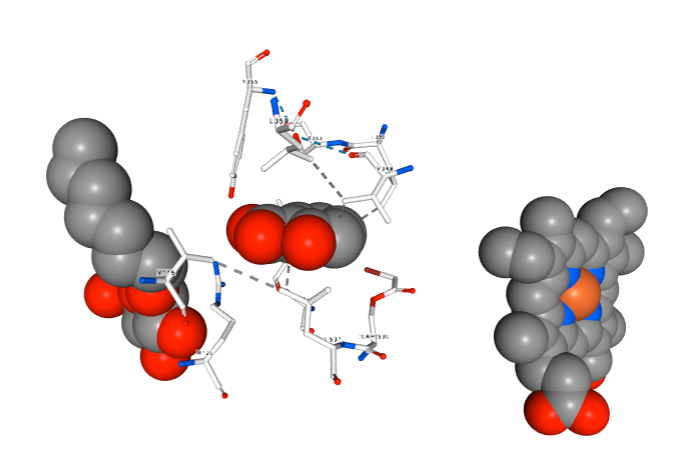


**Figure 3D**

**Figure 4**



**Figure 5**



**II. Comments and tips on programs and databases employed in lessons 1-4.**

Protein Data Bank. Students should be careful typing PDB ID’s. Misspelled PDB ID’s often correspond to a different protein structure than the one selected for analysis.

Once the desired protein structure record has been accessed (Figure 4), the structure file can be downloaded from the PDB using the Download tab (Figure 4 upper right) and selecting PDB format. The downloaded file is editable. If students would like to view and/or edit the file, we recommend WordPad (Windows) or TextEdit (MacIntosh). MS Word should be avoided when editing PDB-formatted files. Editing pdb files is necessary if the user decides to use only a fragment of the protein structure, e.g., (1) a single protein chain instead of all the chains present on the crystal asymmetric unit, (2) a part of the protein chain, or (3) protein with ligands or water molecules removed from the crystal structure. Note that, except for special cases, X-ray crystallographic structures do not contain hydrogen atoms. Thus, a water molecule is represented by a single atom, oxygen. Amino acids lack hydrogen atoms as well.

To view the structure students can choose to (in order of easiness)

1. use 3D View Structure (Figure 4 lower left) or Ligand Interactions;
2. use Protein Workshop and/or Ligand Explorer (Figure 4 lower left); note that Java updates may be required and installation permissions may need to be changed on students’ computers to run these applications;
3. use VMD (VMD needs to be downloaded, installed, and used in conjunction with the tutorial).

Ligand Structure Editing. Some of the programs (e.g., ProDrg, One-Click Docking, VMD) that we have used, are free but require registration, upon which the user receives either a temporary token or login credentials. Students can use ProDrg or One-Click Docking to draw their own ligand structures using very simple editors and then use those two-dimensional formulae to generate spatial coordinates for the ligands. New ligands can also be designed by first searching ZINC for known drugs, downloading their structure files and modifying those structures, for example by changing the functional groups present in the known drug molecules.

Using Avogadro and its “Builder” module may be more enjoyable for students as the program builds a spatial model of the molecule that can be rotated. It also allows for automatic addition of hydrogen atoms (at a specific pH for polar groups) and simple geometry optimization to ensure the built structure is chemically correct.

Spatial structure files for ligands can be generated in multiple different formats; the two most common are pdb and mol2. We recommend that students save their ligands’ coordinates in the mol2 format (the only format accepted by SwissDock). If the ligand has been saved or retrieved from a database as pdb, it can be easily converted to mol2 using VMD by opening a pdb file and saving coordinates as mol2. Avogadro offers multiple formats to save coordinates.

Docking Simulations. While in most cases the simulations are easy to set up and submit to a server, some problems may occur at times. For example, the server may report an error within the protein or ligand structure file. It is most likely to occur when working with SwissDock. We found One-Click Docking to be more robust in this respect. The reasons for errors in SwissDock are intrinsic problems with the selected crystal structure (e.g., missing atoms or residues, non-standard residues) or designed ligand (errors in bond topology or lack of parameters for highly unusual structures). It is not straightforward to suggest here a quick, certain, way to correct the files, so if the problem occurs, we recommend:

1. submitting the protein/ligand to One-Click Docking or Docking Server instead;
2. if multiple protein chains are present, submitting a different chain to SwissDock
3. looking for a target structure via SwissDock internal database and checking all available protein chains;
4. choosing a different structure for the same target (for example, another receptor structure, a homologous structure, structure with higher resolution);
5. inspecting the protein chain in VMD for “breaks” in the backbone and selecting only the contiguous segment of the protein as long as it does not disrupt the tertiary and/or domain structure;
6. in the worst case scenario, the instructor can assign any of the structures from Table 3, which we have tested and set up correctly in SwissDock.

Advanced or highly motivated students (working on the longer-term projects) can try using Chimera Dock Prep (<https://www.cgl.ucsf.edu/chimera/docs/ContributedSoftware/dockprep/dockprep.html#references>) or VMD to edit and prepare protein structure files for simulations in problematic cases or to learn more about the docking protocols in general.

**III. Example worksheets.**

Worksheet 2A: Protein structure and folding.

1. Protein sequence is also called the primary structure. Explain the word “sequence.”
2. How many different aminoacids form protein molecules? Can you name a few amino acids? Search the Internet and find structures, 1- and 3-letter codes for a few selected amino acids, for example alanine, glutamine, aspartic acid, and histidine. Draw the structures side-by-side: what is common for all these amino acids and what is different?
3. Explain the following terms:
   1. peptide bond,
   2. polypeptide chain,
   3. amino-acid residue,
   4. alpha carbon atom,
   5. N-terminus,
   6. C-terminus,
   7. backbone,
   8. side-chains.
4. Search the Web for the structural formulas of lysine, glycine, and valine – can you draw a chemical structural formula of KGV tripeptide?
5. In the figure below, can you circle the amino acids’ side chains? Where is the protein backbone?

 Fig. 5A Fig. 5B

1. Protein chain *folds* and forms *secondary structure* elements. These elements are \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Do you see either one in Fig. 5A?

1. Correctly folded protein chain (the protein’s *native state*) is compact and atoms in it are densely packed. The protein’s function depends on its unique 3D structure, which depends on the protein sequence.

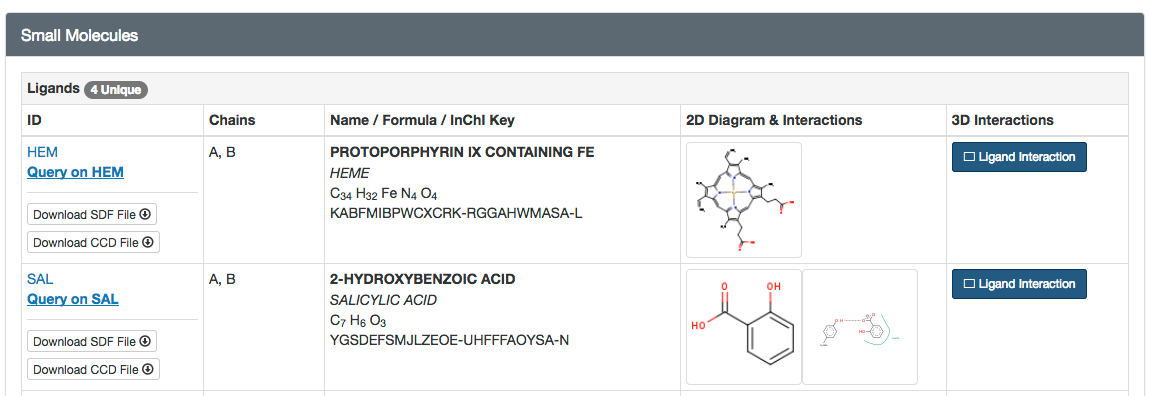
Try to find information on what different functions proteins perform in a cell and list a few examples. Try to find and compare the structures of your example proteins.

Now hypothesize why an amino acid chain must adopt a unique shape to be functional in a cell.

Worksheet 2B: Protein-ligand complexes and binding pockets

Using 1PTH structure in the Protein Data Bank as an example, complete the following:

* + 1. This structure is a *complex* of aspirin and protein, prostaglandin synthase. What does “complex” mean?
    2. Is this a structure of a human protein? If not, what organism does it come from?
    3. Why do you think we can glean information about human proteins from studies of their *homologs* in other species?
    4. Scroll down the PDB summary page for 1PTH till you see the section “Small Molecules.” See below for a snapshot of this section.



List all the ligands present in this structure.

* + 1. Focus your attention on ligands HEM and SAL.

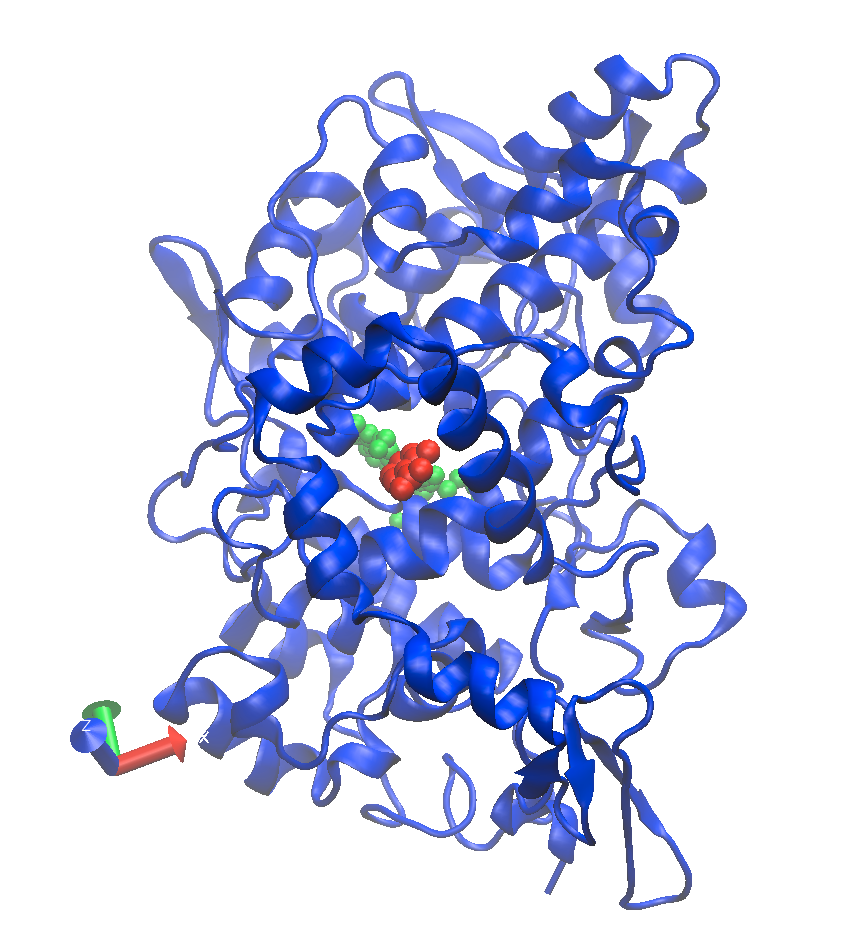
What do you think the function of HEM is? Is it a drug as well?

What is the function of SAL?

How does SAL inhibit the enzyme? Use the material in the PDB “Molecule of the Month” as well as enzyme database BRENDA (you can “google” “BRENDA cyclooxygenase”).

* + 1. Analyze the binding pocket via “2D Diagram and Interactions.” Draw or paste in the schematic of interactions between protein and the drug. List amino acids involved in the contacts. What types of interactions are “holding” the drug in the pocket?
    2. Now examine the 3D view analysis and complete the list of amino acids that form interactions with the ligand. Based on the 3D view try to propose how SAL can be modified chemically to increase the number of interactions with the pocket residues.
    3. View the entire structure either using “3DView Structure” or “Protein Workshop” or VMD. Rotate the structure to find where HEM and SAL are bound. SAL inhibits the enzyme by blocking the “channel” that leads a substrate to HEM – can you find this channel? Paste in the image of the protein and label the ligands.

If finding the channel leading to HEM is hard, try to orient the protein as shown in the figure below. SAL is shown in red. You can also see HEM in green and deeper in the background. Note that only one chain of 1PTH is displayed here.

****