Guide: Intro to Navigating PyMol

Cleaning up your PDB and finding mutation sites

Getting started in PyMol: download, open, cleanup your PDB file **1**
Opening and Orienting to PyMol Features 2
Removing sections of your protein in PyMol 3

Use PyMol to visualize mutation points from literature **4**
Identify mutations you are interested in studying **4**
Defining your mutation goal 5
Visualizing the mutation sites 5

Note: unformatted blue text lines that start with verbs are action items (steps) for you to follow throughout this guide!

Getting started in PyMol: download, open, cleanup your PDB file

Follow the steps shown in this video to **download the free education version**

Opening and Orienting to PyMol Features

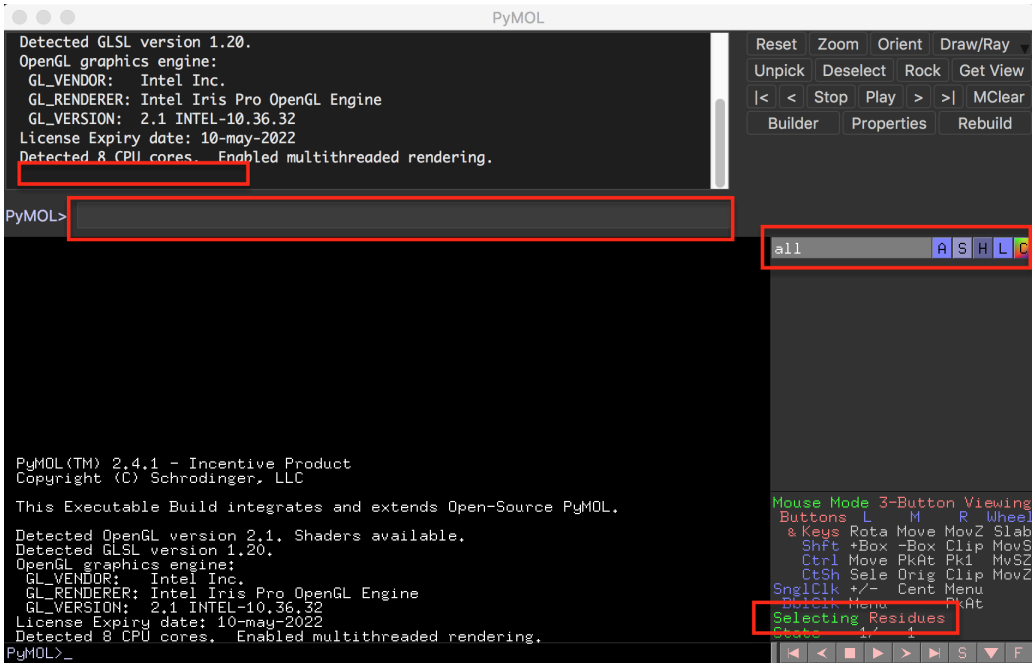


Figure 1. The "most" useful components of the GUI are shown in red boxes. From top down: (1) the code log and "prompter/hints", (2) the command line (also at the very, very bottom), (3) right panel actions/show/hide/color, (4) selection definer aka action panel---note, here ensure you are "viewing" and not "editing", which can mess up your file.

Open your minimized file using the menu dropdown at the top of your screen

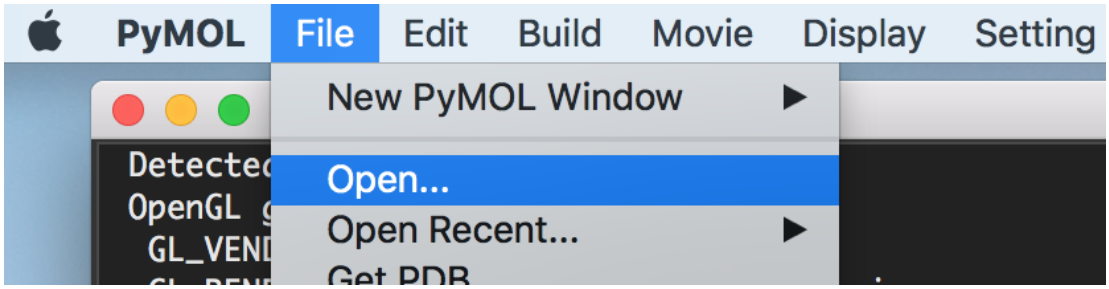


Figure 2. Opening the file

Removing sections of your protein in PyMol

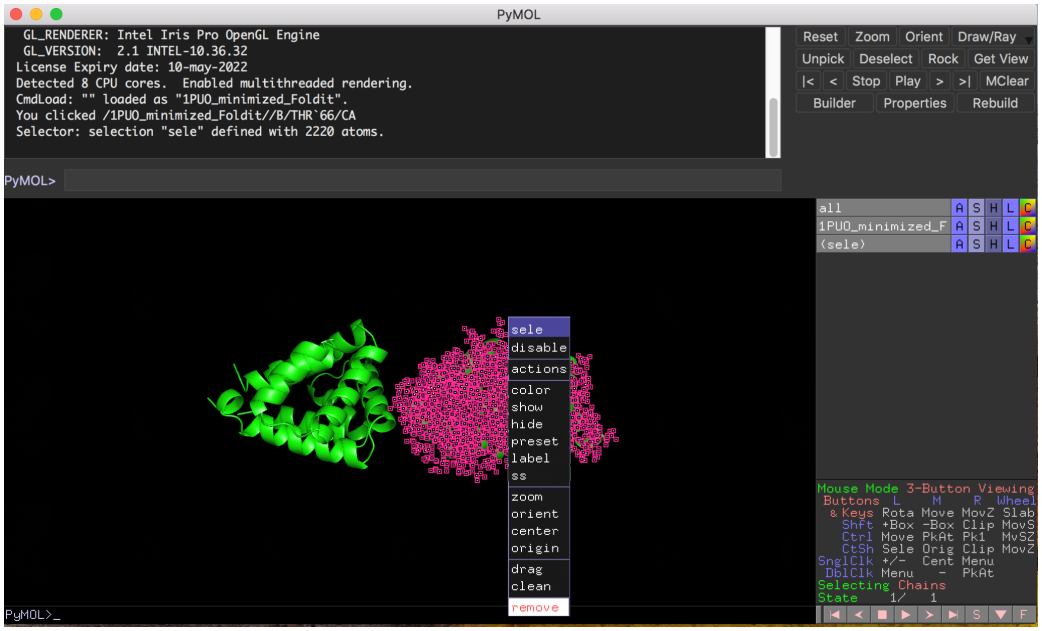


Figure 3. One of the duplicate monomers included in the original PSB needs to be removed. Here we've changed the selection option to "chains". Then click on the section to be removed, it will highlight as shown above. Right click on it and it will bring up the option dropdown menu (seen in the middle of the screenshot), lowest on the list is "remove", click that and it will go away.

Export the PDB **IF** you made a change in PyMol like removing a large section of the protein (this is optional, if you didn't make a significant change, you can use your original, PDB file in the next step).

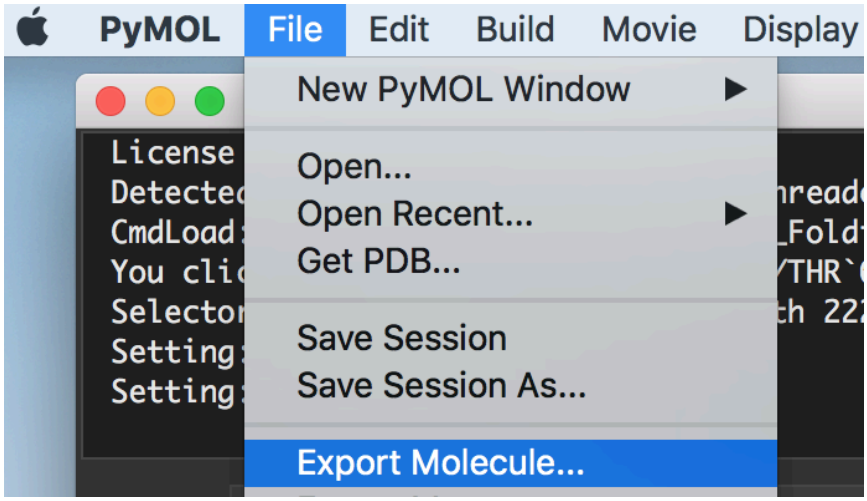


Figure 4. Navigate to export PDB.

Make sure you select the appropriate PDB to export, and that it's set to export PBD.

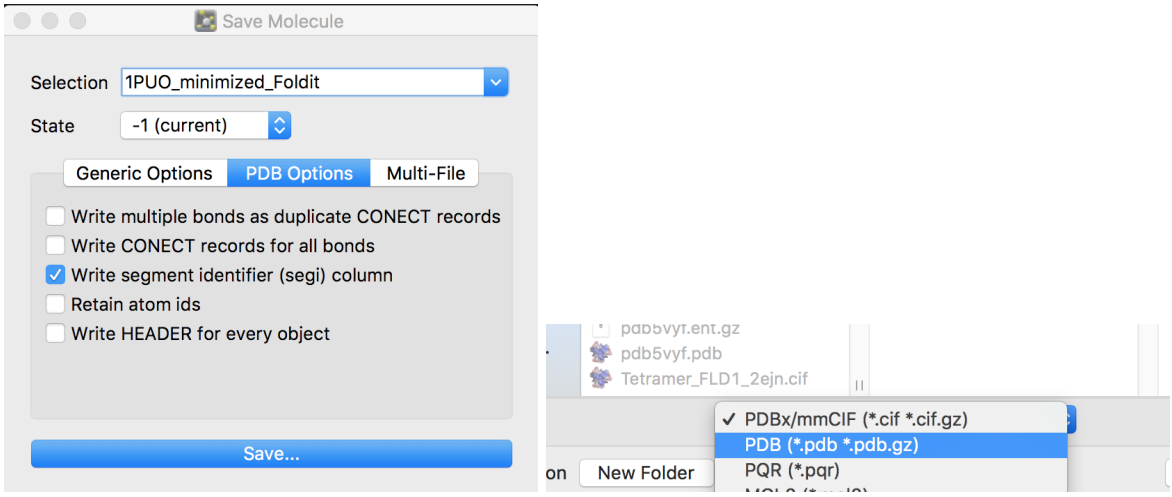


Figure 5. Navigate to export properly as a PDB

Use PyMol to visualize mutation points from literature

The goal of this step is to select the appropriate mutation points you've identified in your protein and make them "sticks" in your design

Identify mutations you are interested in studying

Returning to my cat example, the PDB shows mutations (Figure 1) but the paper linked in the PDB includes a report of the likely place to target and lists exposed residues (Figure 2). https://www.rcsb.org/sequence/1PUO

Three important IgE epitopes have been defined on the Fel d 1 allergen using 14-residue overlapping peptides spanning both chains, two in chain 1 (residues 25–38 and 46–59) and one in chain 2 (residues 15–28) (55). Epitopes 25–38 (corresponding to 117–130 in the construct), 46–59 (138–151 in the construct), and 15–28 (15–28 in the construct) are located on the neighboring helices H6–H7, H7–H8, and H1–H2, respectively, and connecting loops (Fig. 4). Interestingly, epitope 117–130 was recognized by 46% of the analyzed patient sera (55). The crystal structure of Fel d 1 reveals that the side chains of residues Gln-119, Leu-123, Pro-124, Glu-128, Ala-139, Glu-143, Glu-144, and Glu-147 in chain 1 and residues Phe-15, Asn-19, Glu-22, Leu-23, and Leu-27 in chain 2 are exposed to the solvent and are thereby available for IgE binding (Fig. 4).

Figure 6. The "Gold" found here. These are the spots I'll be mutating.

Defining your mutation goal

Articulate your mutation goal: (example): I want to mutate these residues (Figure 1) such that they are very different from the WT to inhibit normal IgE binding, however, I want to select mutants that don't ruin the stability of the structure of the Fel d 1 because I want the mutation to be safe for the cat; this is an important protein in many homeostatic pathways.

Produce a list that looks like this, that will help you navigate alternating between Foldit and PyMol and will serve as a way to organize your thoughts. The residue numbering between Foldit and PyMol is shifted. You will need to figure out the "conversion coefficient" for your particular system; however beyond this, if there are missing pieces of the structure, the numbering will be totally unpredictable and you'll have to "dance" back and forth to locate the residues you're interested in. Once you've got that sorted out, it's easy to program your excel/google sheet to help you find where the sites match for each mutation you want to study.

Visualizing the mutation sites

Bring up the sequence panel

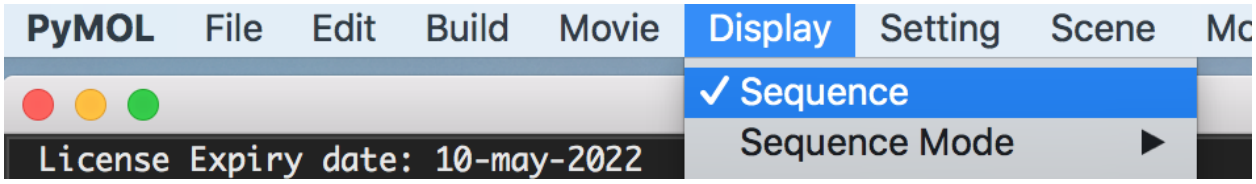


Figure 5. Navigation to bring up sequence panel

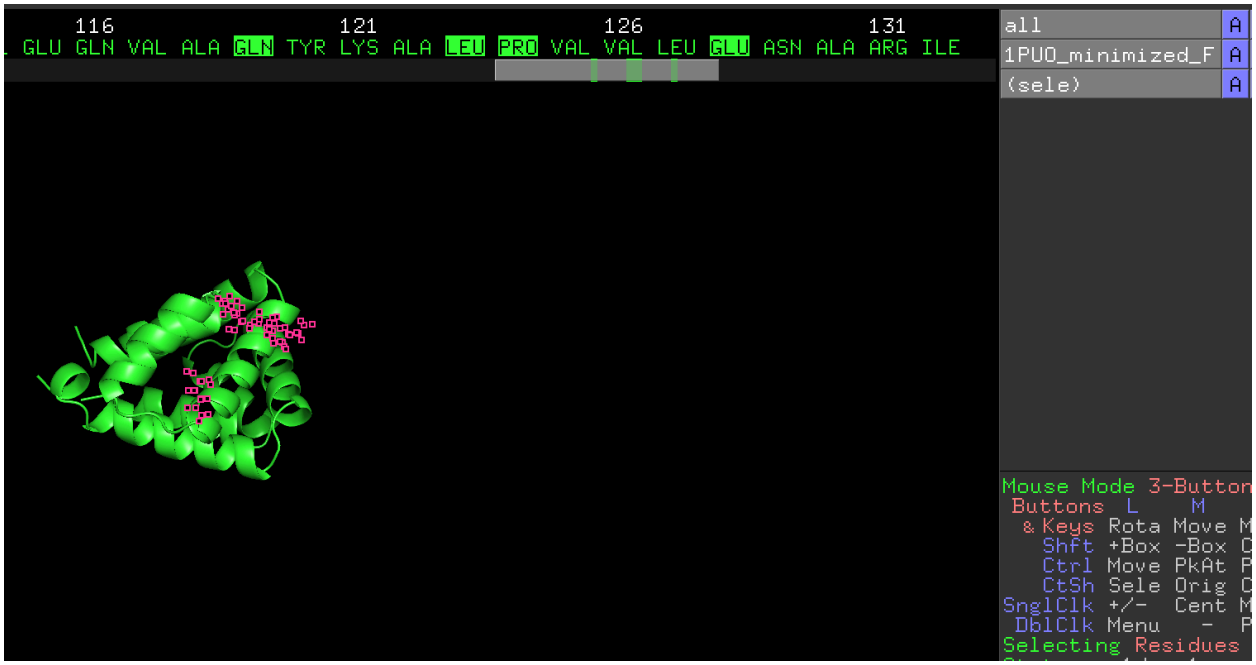


Figure 7. Here the residues of interest are highlighted. To do this you will need to ensure that "Selecting" is set to "Residues" (see bottom right), then just click on the residues of interest in the sequence panel.

Then using the “S” show features of the “(sele)” in the right panel to bring up the “sticks” option will show the residues

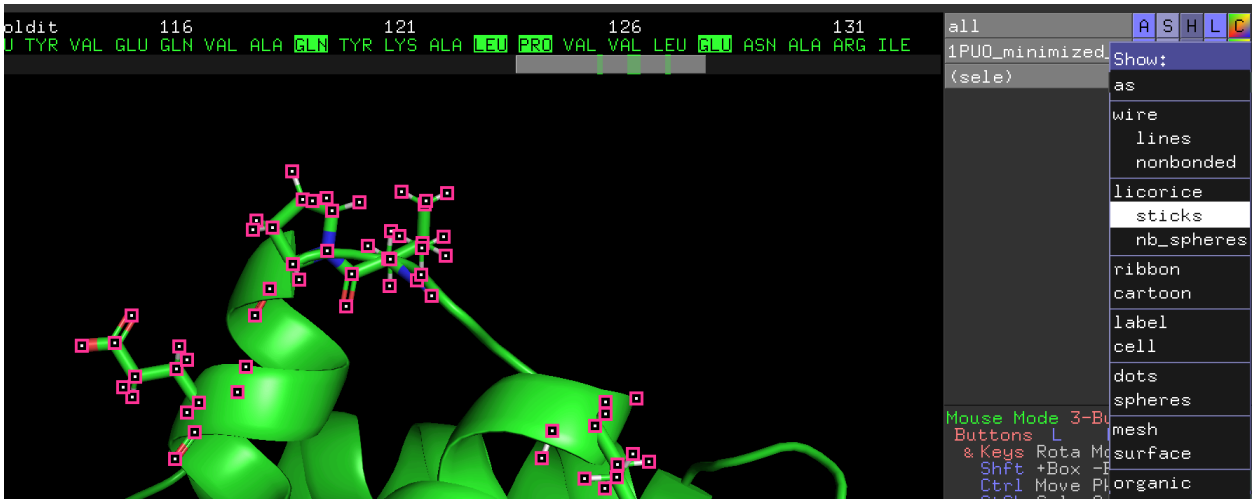


Figure 8. Sticks are shown