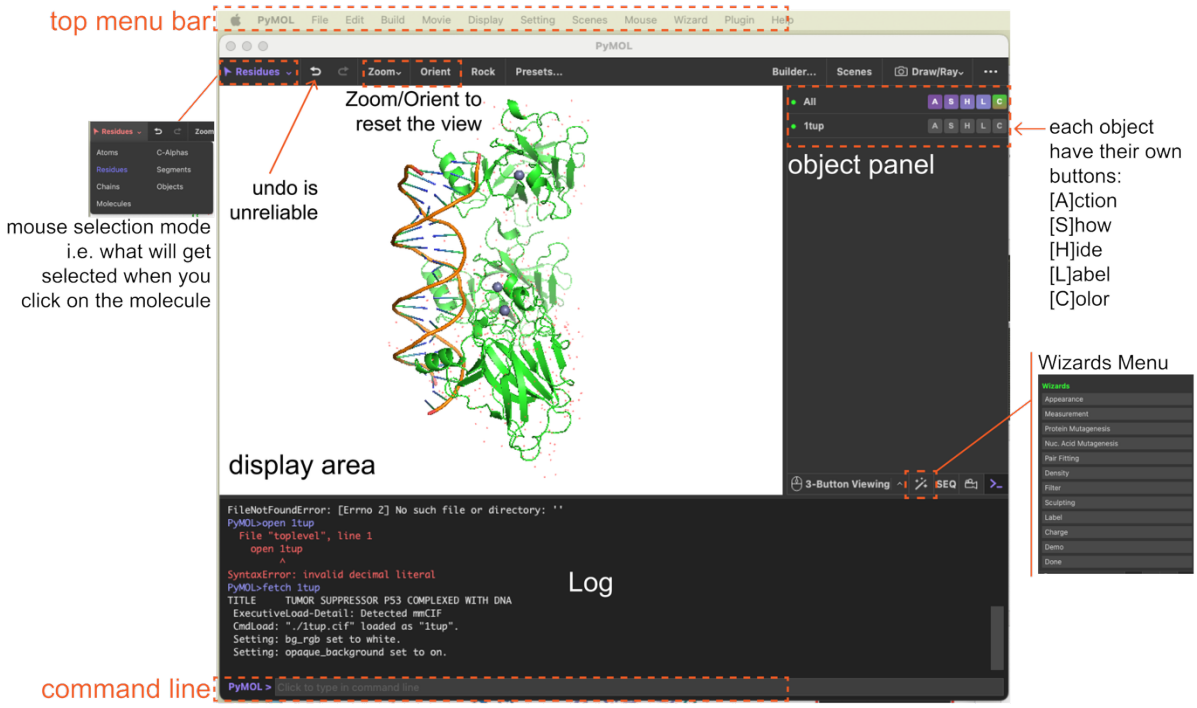
Introduction to PyMOL 2025

Work on p53/DNA (PDB ID 1TUP):

Background: p53 protein is a transcription factor that acts as a tumor suppressor. More than half of all cancers have mutations in the gene coding for p53, making it the most frequently mutated gene in cancer. p53 is a 53 kDa protein consisting of an N-terminal transactivation domain, a central DNA binding domain and a C-terminal tetramerization domain. In 1994, the group of Nikola Pavletich obtained the crystal structure of the central DNA binding domain of p53 bound to DNA. This structure has been deposited into the Protein Data Bank under the ID 1TUP.

PyMOL window overview:



1. Load the PDB file

On https://www.rcsb.org/, search for “1tup” and download the pdb file 1tup.pdb.
Open file in PyMOL using File > Open.
You can also fetch the file using File > Get PDB > 1tup, or in command line: `fetch 1tup`.

At any point, you can reset everything to this step by clicking: File > Reinitialize > Original settings
or enter the command: `reinitialize settings.`

2. Mouse basics (three button mouse):

Familiarize yourself with mouse controls:

Rotations: By clicking and dragging the molecule with the left button.

Translations: By clicking and dragging with the middle button.
You can center on an atom by clicking on it using the middle mouse button.

Zoom: By clicking and dragging with the right button, you can zoom in and out.

Clipping panels: By turning the wheel of the mouse, you can expand and shrink the visible slab (the portion of the Z-axis that you can see). By using control + turning wheel you can move the slab relative to camera.

Note: You are moving your point of view, not the molecule itself.

To reset the view to the overall view of your molecule, you can click the buttons on the top right: **Zoom** (center and scale your molecule to your screen) or **Reset** (same as zoom + reset viewing direction).

Exercise 1: Adjusting view and slab

- Reset the view.
- Locate DNA in the complex, adjust view and slab so you can see single base pair and move the slab so you can see single base-pair as you move the slab along Z-axis.
- How many protein chains are bound to this dsDNA?
- Do all of the protein chains bind DNA similarly?

3. Changing representations and colors

Reset the view.

Change the background color: You can change the background to white by clicking Display > Background > White, or enter the command: `set bg_rgb, white`.

Display molecule in cartoon/sticks/surface: In the Object Panel menu, use the [S]/Show button to change the representation of the molecule. Try sticks, cartoon, and surface. You can remove a representation by clicking on the [H]/Hide button and selecting the representation you want to hide.
Alternatively, you can use the **[S]/Show > as** to replace one kind of representation with another one completely.
In the command line, you can also do this by entering in:

```
show cartoon
show sticks
show surface
```

```
hide surface
hide sticks
hide cartoon
```

Change the color of the molecule: In the Object Panel menu, use the [C]/Color button to change the colors. Try [C] > By chain > By chain. Also try [C] > By SS > (pick a color palette).
Using the command line:
`util.color_chains()`
for fun: `util.chainbow`

Special representations are accessible using the [A]/Action menu and selecting Preset. Try different presets. Use [A] > preset > classified to turn back to the original view.

Note: Using presets will change multiple settings. It may undo work you did previously. The Undo function in PyMOL is unreliable if you need to go back.

4. Making selections and creating objects

Selections: Click on any bond or atom. In the Object Panel menu, you can see that <sele> has appeared. The entire residue to which the atom belongs has been selected in pink in the Display Area. If you click on the same residue, it will be deselected. If you click on another residue, it will be added to the selection.

Changing selection mode: By default, clicking on an atom selects the entire residue or molecule to which the atom belongs. You can change this behavior by clicking: Mouse > Selection Mode > Atoms/Residues/Chains etc.

If you want to reset the selection, click on <sele> in the Object panel menu. You can then make another selection.
Selections and objects can be deleted by clicking [A] > delete selection/object.

Exercise 2: Changing representations

- Reinitialize by clicking: File > Reinitialize > Original settings or enter the command: `reinitialize settings`
- Hide waters:
  [H] > waters
- Select individual protein chains and change their colors:
  Change the mouse selection mode to chains: Mouse > Selection Mode > Chains.
  Click on a chain. In the Object Panel, click on [C] button of <sele>, choose a color.
- Set the DNA representation to licorice sticks and change color by element:
  Make sure your mouse selection mode is still chains.
  Click the 2 DNA strands. <sele> > [S] > as > licorice – sticks.
  [C] > by element > choose a color palette.

Working with the sequence: You can also select or unselect residues from the sequence viewer.
Click on the menu bar: Display > Sequence to turn the sequence viewer on. This will display the sequence of the protein at the top of the Display Area. You can center residues by middle-clicking on a residue in the sequence. By right-clicking on a residue in the sequence you open a popup menu that allows you to change representation, colors, etc.

You can change the sequence viewer settings by clicking Display > Sequence Mode. For example, to change sequence names to 3-letter: Display > Sequence Mode > Residue Names.

Selection and Objects

- rename selection: <sele> [A] > rename selection
- copy selection to new object: <sele> [A] > copy to object
- select all residues within 4Å of selection: <sele> [A] > modify > around > residues within 4 A
  (this selects the atoms; you can show sidechains by going to [S] > sidechain > stick)

Exercise 3: Working with sequences

The Pavletich lab determined certain residues most frequently mutated in cancer are at or near the protein-DNA binding interface: R175, G245, R248, R249, R273, R282. Select and show the side chains for these residues and look at them to confirm they are at or near the DNA-binding interface. Note: there might be multiple protein molecules bound to this dsDNA.

Either use the GUI:

- click the residues in the sequence viewer
- change representation: [S] > sidechain > sticks

Or by using the command line:

```
select resi 175 or resi 245 or resi 248 or resi 249 or resi 282 or resi 273
```

more details on how to select residues through the command line:
https://pymolwiki.org/index.php/Selection_Algebra
change representation of selection:

```
show sticks, sel and sidechain
```

5. Making measurements

You can measure distances, angles, and dihedrals.

Exercise 4: Measure the coordinating distances of the Zn atom

- Select a Zn atom, expand the selection to select coordinating residues
  <sele> [A] > modify > expand > **by 4 A, residues**
- change the representation of the side chains to show sticks
  <sele> [S] > sidechain > sticks
- you can also change the color of this selection:
  <sele> [C] > by element > choose your color palette
- In the menu bar, Wizard > Measurement to turn on the measurement option
  Click two atoms to measure the distance between them.
  To exit the measurement mode, click “Done”.

6. Working with electrostatics

You can display surface representation and visualize electrostatic potential.

Exercise 5: Showing electrostatics

- Create new individual objects that contain the protein chains A, B and C individually
  (can click in GUI or use the command `split_chains 1tup`)
- Generate electrostatic potential surfaces quickly:
  [A] > generate > vacuum electrostatics > protein contact potential

- Do you see any electronegative or electropositive patches? Where are they located?

Note: The better way to generate electrostatics uses Adaptive Poisson-Boltzmann Solver (APBS). In papers you will often see figures with a scale from -5 to 5 that uses this:
Plugin > APBS electrostatics > choose your selection > Run (takes some time, wait till "Finished with Success" popup
If this errors out for you, try going into Prepare Molecule > Options > Method: change to "use formal_charge and vdw"

7. Doing alignments

You can use PyMOL to perform alignments to superimpose and compare protein structures.

Exercise 6: Align the three protein chains

- you should have the 3 protein chains in separate objects
- align chains A and C to chain B
  [A] > align > to molecule > chain B
- Are the proteins adopting different conformations?

This align function performs a sequence-based alignment, then a structural alignment. For more alignment options, go to menu bar Plugin > Alignment. for more information on comparing alignment methods such as super and cealign, see https://pymolwiki.org/index.php/Align

8. In silico mutagenesis

PyMOL can be used to perform in silico mutagenesis (without modelling) residues.

Exercise 7: perform an in silico alanine mutation

- Select R175 in chain A, color its sidechain.
- expand the selection to residues within 5 Å and show sidechains
- Go to menu bar Wizard > Mutagenesis
- In the panel on the right, click on 'No Mutation', select ALA.
- Click on the R175 that you want to mutate
- Then click Apply
- What could be an effect of this mutation?

9. Displaying symmetry mates

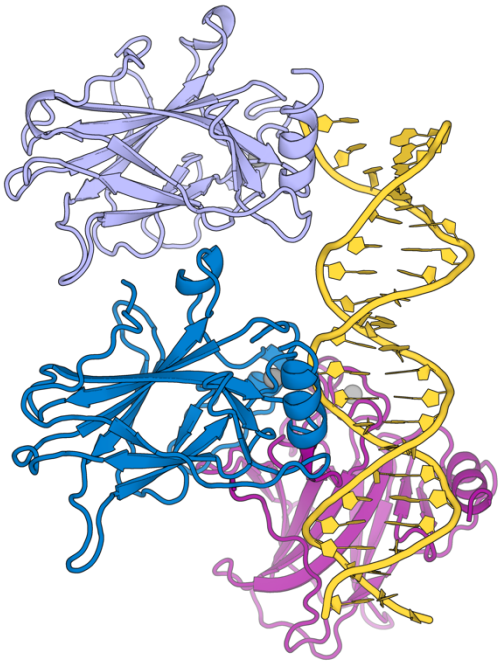
The structure of the p53/DNA complex has been solved by X-ray crystallography, which means that proteins are part of a crystal lattice. PyMOL can be used to visualize other molecules in crystal lattice.

To generate symmetry mates in PyMOL:

- On the object containing the entire original 1tup: [A] > generate > symmetry mates > within 20 A
- How are the DNA molecules arranged? Zoom into the interfaces of the DNA molecules. What facilitated the crystal lattice formation?

10. Make a figure

Generally, when you have a crystal structure, you want to publish it. In PyMOL, you can make figures through the Draw/Ray button. Draw will essentially make an image of your current display. Ray will make a figure with extra settings applied such as outlines, or transparent background.



Ha An's settings for the figure on the right, play around to see what the settings do:

```
set cartoon_ring_mode, 1
set cartoon_ladder_mode,1
set cartoon_ring_finder, 1

set ray_shadows, off
set ray_trace_mode, 1
```

For papers, I "Ray" at 300 dpi, transparent background, 20 cm width. As always, you can look at the PyMOL wiki for explanations and examples.

Useful resources:

Name	Link	Purpose
UniProt	https://www.uniprot.org/	Protein information database
Protein Data Bank (PDB)	https://www.rcsb.org/	Protein structure repository
PyMOL	https://pymol.org/2/	Download PyMOL
PyMOL wiki	www.pymolwiki.org	Information on PyMOL such as examples of settings
ConSurf	http://consurf.tau.ac.il	Evolutionary conservation of amino/nucleic acid sequences
PISA	www.ebi.ac.uk/pdbe/pisa/	Protein interface and assemblies
Dali	http://ekhidna2.biocenter.helsinki.fi/dali/	Search PDB for similar structures
EMDB	https://www.ebi.ac.uk/emdb/	Repository for EM/ET data
AlphaFold database	https://alphafold.ebi.ac.uk/	Database of AlphaFold predictions
ChimeraX	https://www.cgl.ucsf.edu/chimerax/	Another visualization software, works well with EM maps