Originally prepared by Toni Lee

BMB/Bi/Ch 170 – Introduction to PyMOL

PyMOL is a powerful open-source molecular graphics and analysis program developed by Warren Delano. The program is available for various platforms (Windows, MacOS, Linux) and is extensible with the programming language Python. PyMOL allows a user to examine a molecular structure while creating high quality images and movies for publications and presentations. This tutorial will show the user how to install and navigate PyMOL on either Windows or MacOS.

Part I: Installation and Setup

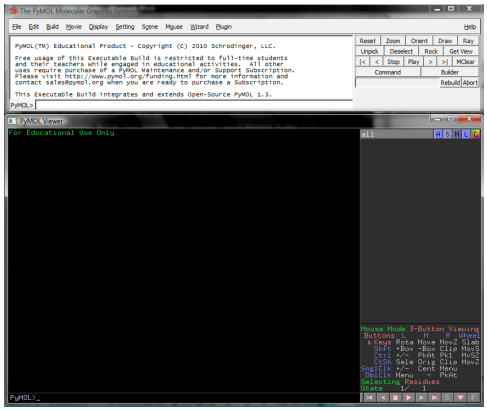
If PyMOL is not already present on your computer you'll need to download it from Caltech software download server: http://saf.bio.caltech.edu/cgi-bin/caltech/downloadcontrol.pl. Make sure your web server is connected from the campus wired and sign in Wi-Fi networks. (You are supposed to uninstall this version after the class ends.)

Alternatively, you can obtain it from http://pymol.org/edu/?q=educational by registering using a valid email address.

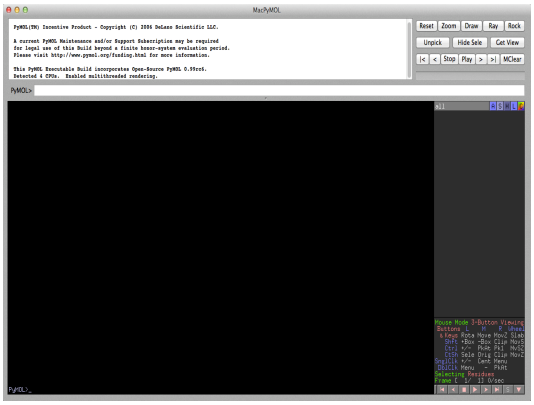
Download the version of PyMOL suitable for your operating system. You may have to decompress the file using WinZip (PC) or Stuffit Expander (Mac). If you have a recent operating system, extraction tools may be built into your file explorer. Installation should be straightforward on either a Windows or Mac.

Once installed, open the program. You should see two windows if you are on a Windows and a single window if you are using a Mac.

Windows Interface



Mac Interface

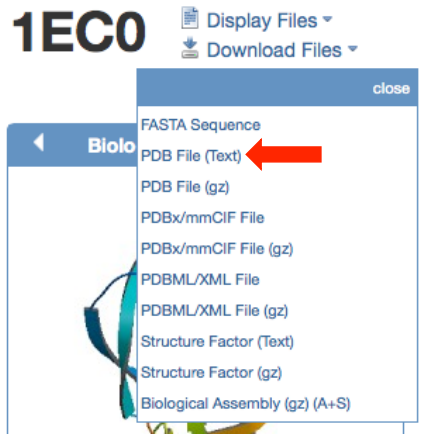


Part II: Opening a Molecular Structure

Structural data is contained in a worldwide database called the Protein Data Bank (PDB). The collection consists largely of structures solved by crystallographic- or NMR-based methods. For simplicity, we will focus primarily on the crystal structure of HIV-1 protease in complex with the inhibitor bea403 (PDB ID 1EC0).

1) First, navigate to the PDB: www.pdb.org
2) In the search box, type in the PDB ID code '1EC0.' You can also search by the protein name or any other identifier. Make sure you find the structure 1EC0 page, not the 1ECO page.

3) Expand "Display Files" on the right side of the page by clicking on it. Save the PDB text file (see red arrow) to a known destination.

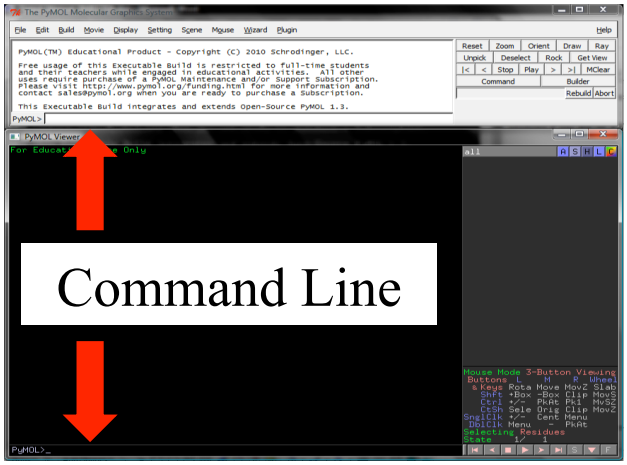


4) Open PyMOL and expand the File menu. Click on Open and locate the file 1EC0.pdb.
5) If you are connected to the internet and are using Windows, you can open files faster by opening PyMOL and expanding the Plugin menu (if you do not see a Plugin menu, do not worry about this feature). Open the PDB Loader Service and enter the ID code '1EC0.' The structure should pop onto the screen. You can also use the fetch command (i.e. 'fetch 1EC0').
6) Practice with other molecules until you feel comfortable opening them.

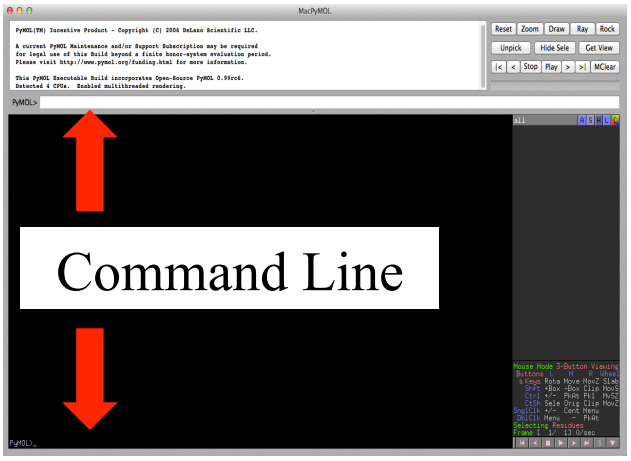
Part III: Viewing and Manipulating the Structure

The following information will demonstrate how to highlight certain features of the molecule. As with most aspects of PyMOL, there are usually many different ways to achieve the same effect. Using the command line can accomplish most tasks, but many common functions are built into the program as buttons toward the right side of the interface.

Windows:



Mac:



Viewing HIV protease (1EC0.pdb) with the mouse/touchpad:

Rotating the molecule around a fixed point: Left click and move the cursor. If you have a 1 or 2 button touchpad/mouse, navigate to the 'mouse' menu and select either 1- or 2-button viewing mode

depending on your hardware. Though commands are given for 2-button systems, it is highly recommended that you use a mouse with a scroll wheel (3-button viewing mode) to navigate a molecular structure without difficulty. **Note:** The commands for rotating, translating, etc. are listed in the box in the bottom right corner:

2-button viewing:

```
Mouse Mode 2-Button Viewing
 Buttons  L     M     R  Wheel
  & Keys Rota    -  MovZ    -
    Shft Pk1     -  Clip    -
    Ctrl Move    -  PkAt    -
    CtSh Sele    -  Cent    -
SnglClk PkAt     -  Menu
 DblClk Menu     -  Cent
Selecting Residues
Frame [  1/  1] 7/sec
```

3-button viewing:

```
Mouse Mode 3-Button Viewing
 Buttons  L     M     R  Wheel
  & Keys Rota Move MovZ Slab
    Shft +Box -Box Clip MovS
    Ctrl +/-  PkAt Pk1  MvSZ
    CtSh Sele Orig Clip MovZ
SnglClk +/-  Cent Menu
 DblClk Menu    -  PkAt
Selecting Residues
Frame [  1/  1] 0/sec
```

- **Zoom in or out**: Right click and move the cursor up or down.

- **Translating the molecule**: Hold the middle mouse button/scroll wheel and move the cursor. Again, for touch pads, go to the 'Mouse' menu and select the appropriate mode for your hardware (either 1- or 2-button viewing mode). For example, with 2-button viewing mode, left click and hold control while moving the cursor to translate the molecule.

- **Clipping the molecule**: The clipping function allows the user to slice the molecule at a plane parallel to the screen. This allows for better visualization of molecular interiors. To clip the molecule from the front and the back simultaneously, use the scroll wheel. Hold shift in 3-button viewing mode and scroll to perform the same operation on the touch pad. In 2-button viewing mode, hold shift + right mouse button and scroll. To use the command line to clip, use the following commands (comments appear after the # sign):

```
clip near, -10          #moves the clipping plane nearest to the user ten units away from
                         the user
clip far, 5             #moves the clipping plane furthest from the user five units towards
                         the user
```

To undo any clipping, go to the 'Display' menu, choose 'Clip' > Nothing

- **Showing and hiding objects**: There are many ways of displaying your molecule. Some of the most popular representations include cartoon, sticks, and ribbon. Cartoon highlights secondary structure, displaying β-sheets as arrows, α-helices as coils, and loops as thin tubes. The stick representation shows all bonds and atoms (excluding hydrogens) as thick sticks while the line representation shows bonds as thin lines. Ribbon structures trace the C_α backbone as a thin line. The structure is initially loaded as

lines. To change the representation, click the 'S' button next to 1EC0 and select the 'as' menu. Select 'lines,' 'ribbon,' 'sticks,' or 'cartoon.' Alternatively, type into the command line:

```
as cartoon, 1EC0          #shows 1EC0 as a cartoon, erases previous style
as sticks, 1EC0           #shows 1EC0 as sticks, erases previous style
```

The 'as' command will eliminate other representations before displaying your molecule in the new style. To display multiple representations, use the 'show' command:

```
show cartoon, 1EC0        #shows 1EC0 as a cartoon
show ribbon, 1EC0         #shows 1EC0 in ribbon form
```

Hiding objects is an extremely useful function. The 'H' menu can be used to hide specific representations without altering others. You can also select 'everything' to hide all parts of a certain object. To specifically hide unwanted styles from the command line use the 'hide' command:

```
hide cartoon, 1EC0        #hides cartoon 1EC0
```

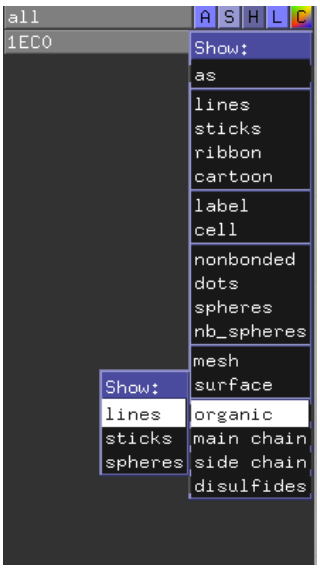
PDB files may not load with all portions of the molecule visible, especially in the case of ligands, metals, and glycosylations. To ensure you are seeing all atoms present, you can look at the PDB entry under 'Ligand Chemical Component' found under the summary tab:

↕ Ligand Chemical Component			Hide
Identifier	**Name**	**Formula**	**Interaction View**
BED	N,N-[2,5-O-DI-2-FLUORO-BENZYL-GLUCARYL]-DI- [1-AMINO-INDAN-2-OL]	$C_{38}H_{38}F_2N_2O_8$	Ligand Explorer

In the case of HIV protease there is 1 entry, the inhibitor N,N-[2,5-O-di-2-fluoro- benzyl-glucaryl]-di-[1-amino-indan-2-ol]. This molecule is known to PyMOL by the identifier 'BED.' To display the inhibitor, try typing these examples into the command line:

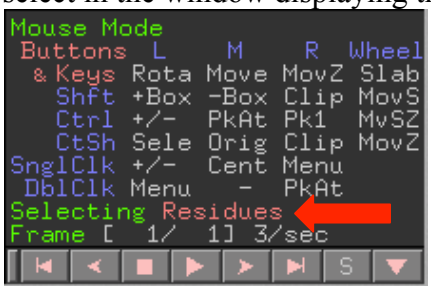
```
show sticks, resname BED        #shows the inhibitor in stick format
show spheres, resname BED       #shows the inhibitor as spheres
hide everything, resname BED    #hides all representations of the inhibitor
```

If you have small molecules lacking covalent bonds to your protein in your structure, you can easily display these molecules using the 'S' button. Click on the 'S' next to 1EC0 and select 'organic.' Choose lines, sticks, or spheres.



• **Showing Disulfides:** PyMOL has a special function for showing disulfide bonds. To do so, go to the 'S' menu then put the cursor on 'disulfides.' Display the disulfide in the representation of your choice. Practice with structure 1MF2 (anti HIV1 protease Fab complex). Download and open the structure. The molecule will load with the disulfides displayed. Hide the disulfides by choosing disulfides from the 'H' menu and attempt to display them again.

• **Selecting Residues/Atoms/Molecules, etc**.: You might want to look at or highlight a particular residue or region of the structure. To do that you can select the object of interest by clicking on what category of object you want to select (residues, chains, atoms, C-alphas, molecules, or objects) and then clicking on the residue/atom/molecule you would like to select. Reload 1EC0 and practice selecting objects. First click on the pink word on the bottom right of the window (shown below) to change your selection to what you want (residues, chains, atoms, C-alphas, molecules, or objects). Then click on the residue/atom/molecule you want to select in the window displaying the structure.



Your selection will be saved under the name '(sele)'. Selecting an object can also be accomplished in the command line:

```
select resi 1-10            #selects the first 10 residues
select chain A              #selects chain A
select resname BED          #selects inhibitor
```

• **Changing Colors**: The color of any selection (atom, molecule, chain, residue, etc.) can be changed. To change colors select what you would like to alter. Your selection is now named '(sele).' Expand the color menu next to (sele) on the right hand menu and choose a color. You can also color the chain in a rainbow from the N to C terminus by choosing 'chainbows' under the 'by chain' submenu.

To color by protein chain, navigate to the 'C' menu, choose 'by chain', then choose 'by chain' again. The molecule can also be colored from the command line:

```
color pink, resi 1:10          #colors the first 10 residues pink
color blue, resname BED        #colors the ligand blue
color gray60, chain A          #colors chain A gray60
color red, ss h                #colors helices red
color yellow, ss s             #colors β-sheets yellow
color green, ss l+             #colors loops green
```

Secondary structure can also be colored by using the 'by ss' function under the color menu.

Part IV: Advanced Operations

• **Distance/Angle/Dihedral Measurements:** To make a measurement go to the 'Wizard' menu and choose 'Measurements.' The type of measurement currently selected will appear in the top blue bar on the right side:

Using the cursor, click the atoms between which a measurement will appear. Click done on the right-hand sidebar to exit measurement mode. To delete a bad measurement, click 'Delete Last Object.' To delete any measurement, use the command line:

```
delete measure01               #deletes measure01
```

• **Labeling:** The 'L' button on the right-hand menu allows one to place labels on residues, atoms, chains, etc. The labels in PyMOL, however, are difficult to see and are conventionally not used in figures for publication. To create more clear labels, ray trace your image, save the image as a .png,

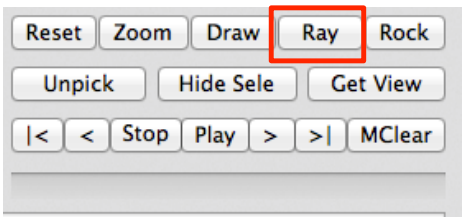
import the image into a picture editor (Powerpoint and Photoshop work fine) and add text labels.

• **Changing the Background Color:** For easier viewing, change the background color by going to the 'Display' menu and expanding the 'background' menu. The standard color for figure making is white. If you prefer to do this in the command line, type:

`bg white` #changes background color to white

`bg cyan` #changes background color to cyan (you can use any color supported by PyMOL)

• **Making Figures**: Making figures look good is an art. First, go to Display > Quality > Maximum Quality. Change the color space to publication grade (Display > Color Space > CMYK). Position your molecule for style and click the 'Ray' button on the GUI window:



To create images of greater size, use the command line.

`ray 500, 500` #creates an image 500x500 pixels

Large sizes may take a long time to generate. The program may look like it is crashing, but be patient. To save an image go to the 'File' menu and choose 'Save Image.' The image will save as a .png.

Part V: General Bits of Usefulness

• **Saving a Session**: To save your current session, go to File > Save Session. When you open the file all changes to your structure will reappear.

• **If you really screwed up:** There is no undo button in PyMOL, but you can reset the program.

`reinit` #resets the program

• **Creating Objects**: Sometimes you want to manipulate a certain part of a structure independently of all other parts. To do this, create that segment as a different object. The object will appear in the right-hand menu.

`create myAminoAcid, (sele)` #creates a new object called myAminoAcid that contains your selection

`create catalytic_residue, resi 4` #creates a new object called catalytic_residue that consists of residue 4

Now you can manipulate the new object:

`color pink, myAminoAcid` #colors myAminoAcid pink

`show surface,catalytic_residue` #shows catalytic_residue in surface mode

- **For More Information:** The PyMOL Wiki site has abundant information on PyMOL for the curious:
http://www.pymolwiki.org/index.php/Main_Page