

PyMOL tutorial



PyMOL 2.3 used in this tutorial

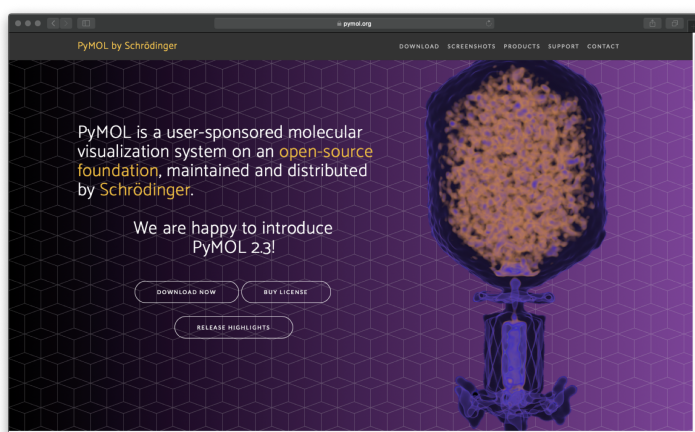
Date: April 15, 2020

Kuen-Phon Wu

Institute of Biological Chemistry, Academia Sinica

<https://kpwulab.wordpress.com>

Copyright reserved.



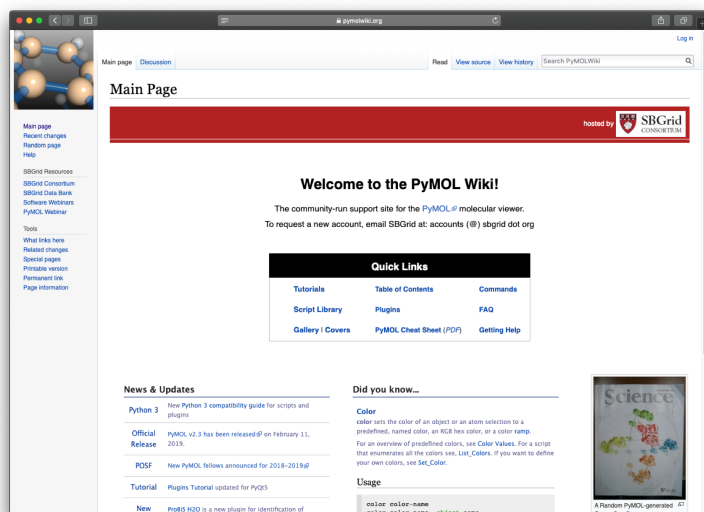
<https://pymol.org/2/>

About PyMOL:

- Originally developed by Dr. Warren DeLano
- Now owned by Schrödinger, Inc.
- Current version (2020 April): 2.3.5
- Support OS: Windows, MacOS and Linux (and iPad)
- Open sourced at GitHub.
- Commercial PyMOL: Incentive PyMOL.
- Education license: no cost for full-time students. (go to this link to apply <https://pymol.org/edu/?q=educational>)

Online resources for PyMOL users

PyMOL Wiki https://pymolwiki.org/index.php/Main_Page



PyMOL Webinar (1 hour)

https://www.youtube.com/watch?v=C_lewbvUyGk

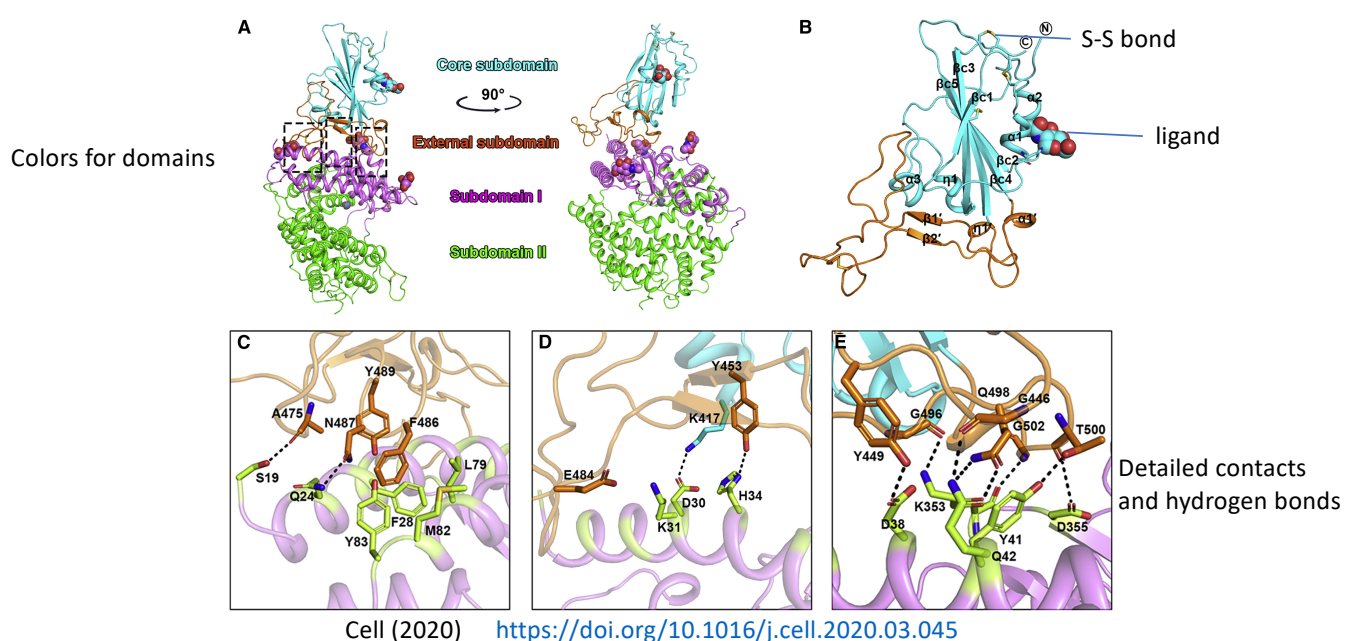
PyMOL: Active Sites in Minutes (Using only Sequence Info!)

<https://www.youtube.com/watch?v=mBIMi82JRfI>

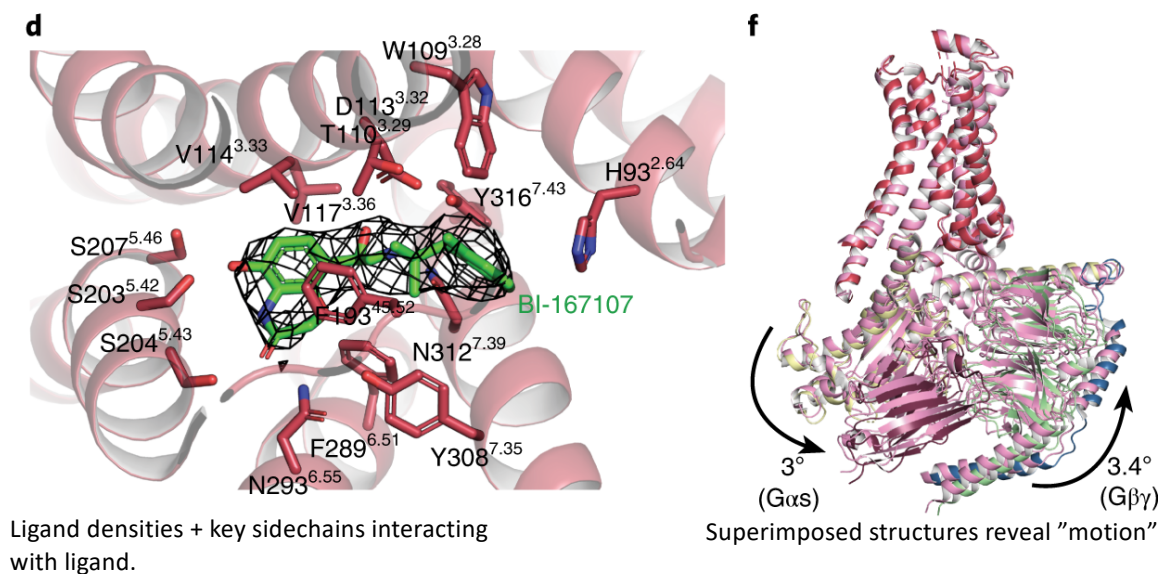
PyMol Lab Assignment

<https://www.youtube.com/watch?v=tkxNuTADMWg>

Examples of PyMOL-made figures in published papers

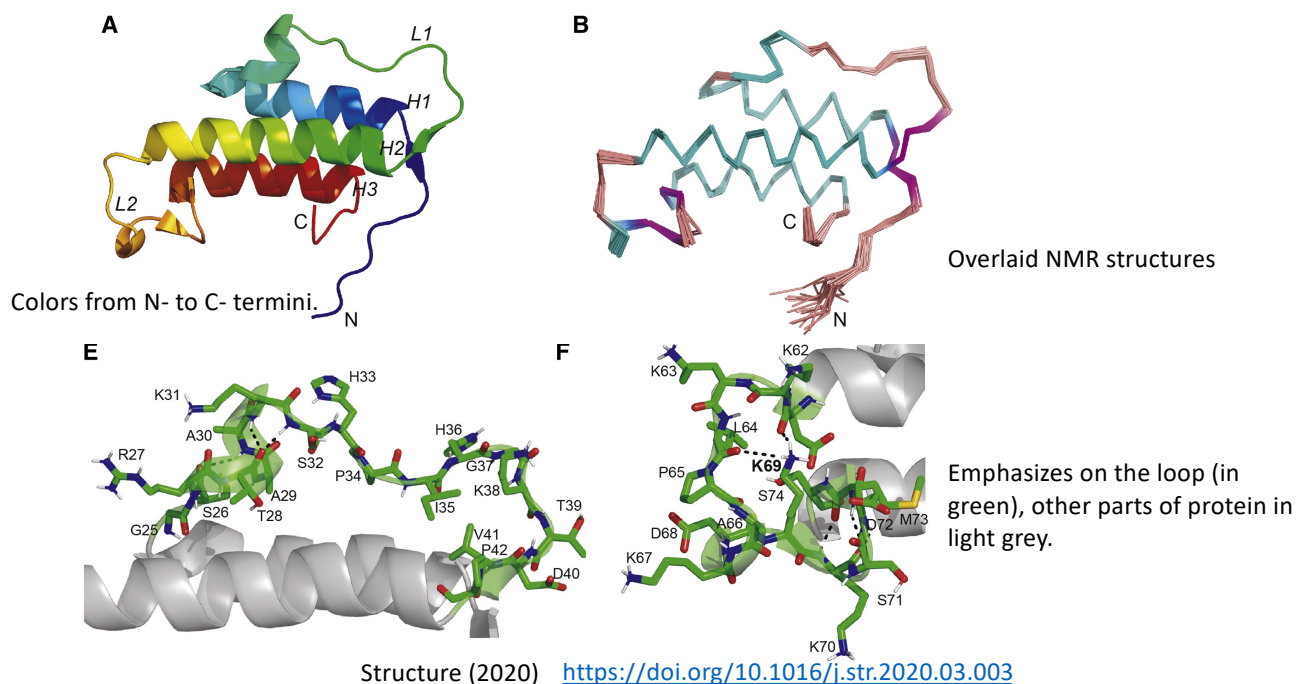


Examples of PyMOL-made figures in published papers



NSMB (2019) GPCR <https://www.nature.com/articles/s41594-019-0330-y>

Examples of PyMOL-made figures in published papers



Protein Data Bank stores all publicly deposited structures of biomolecules

The screenshot shows the RCSB PDB website for entry 1UBQ. The interface includes a top navigation bar with links like 'Deposit', 'Search', 'Visualize', 'Analyze', 'Download', 'Learn', and 'More'. Below this, there are tabs for 'Structure Summary', '3D View', 'Annotations', and 'Experiment'. The 'Structure Summary' tab is active, displaying a 3D ribbon diagram of the protein structure. To the right of the diagram, key information is provided: '1UBQ' (PDB accession ID), 'STRUCTURE OF UBIQUITIN REFINED AT 1.8 ANGSTROMS RESOLUTION', DOI: 10.2210/pdb1UBQ/pdb, BMRB: 5387, Classification: CHROMOSOMAL PROTEIN, Organism(s): Homo sapiens, Deposited: 1987-01-02, Released: 1987-04-16, and Deposition Author(s): Vijay-Kumar, S., Bugg, C.E., Cook, W.J. Below this, there is a section for 'Experimental Data Snapshot' showing Method: X-RAY DIFFRACTION, Resolution: 1.80 Å, and R-Value Observed: 0.176. A 'wwPDB Validation' section provides a detailed breakdown of the structure's quality metrics, including Clashscore, Ramachandran outliers, Sidechain outliers, and RSQR outliers. The 'Literature' section lists the primary citation: 'Structure of ubiquitin refined at 1.8 Å resolution. Vijay-Kumar, S., Bugg, C.E., Cook, W.J. (1987) J Mol Biol 194: 531-544'. The 'Macromolecule Content' section shows the total structure weight and atom counts. The '3D View' section offers options for viewing the structure, including standalone viewers and protein workshop tools. The 'Annotations' section provides information on global symmetry and stoichiometry. The 'Experiment' section offers options for downloading files and viewing the structure in 3D. The 'Structure Summary' section provides a detailed overview of the structure, including the title, authors, and depositions. The '3D View' section offers options for viewing the structure, including standalone viewers and protein workshop tools. The 'Annotations' section provides information on global symmetry and stoichiometry. The 'Experiment' section offers options for downloading files and viewing the structure in 3D.

More details

Download file and more information

PDB accession ID

Structure title, authors and depositions

Raw data and model quality

Associated literature(s)

Preview

Online 3D viewer

Assembly info

Molecular info: MW, # of atoms, # of aa..

<https://www.rcsb.org/structure/1UBQ>

Content of PDB file

```

HEADER      CHROMOSOMAL PROTEIN                      02-JAN-87   1UBQ
TITLE       STRUCTURE OF UBIQUITIN REFINED AT 1.8 ANGSTROMS RESOLUTION
COMPND      MOL ID: 1;
COMPND      2 MOLECULE: UBIQUITIN;
COMPND      3 CHAIN: A;
COMPND      4 ENGINEERED: YES
SOURCE      MOL ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;
SOURCE      3 ORGANISM_COMMON: HUMAN;
SOURCE      4 ORGANISM_TAXID: 9606
KEYWDS      CHROMOSOMAL PROTEIN
EXPDTA      X-RAY DIFFRACTION
AUTHOR      S.VIJAY-KUMAR,C.E.BUGG,W.J.COOK
REVDAT      5 09-MAR-11 1UBQ 1 REMARK
REVDAT      4 24-FEB-09 1UBQ 1 VERSN
REVDAT      3 01-APR-03 1UBQ 1 JRNL
REVDAT      2 16-JUL-87 1UBQ 1 JRNL REMARK
REVDAT      1 16-APR-87 1UBQ 0
JRNL        AUTH S.VIJAY-KUMAR,C.E.BUGG,W.J.COOK
JRNL        TITL STRUCTURE OF UBIQUITIN REFINED AT 1.8 A RESOLUTION.
JRNL        REF J.MOL.BIOL. V. 194 531 1987
JRNL        REFN ISSN 0022-2836
JRNL        PMID 3041007
JRNL        DOI 10.1016/0022-2836(87)90679-6
REMARK      1
REMARK      1 REFERENCE 1
REMARK      1 AUTH S.VIJAY-KUMAR,C.E.BUGG,K.D.WILKINSON,R.D.VIERSTRA,
REMARK      1 AUTH 2 P.M.HATFIELD,W.J.COOK
REMARK      1 TITL COMPARISON OF THE THREE-DIMENSIONAL STRUCTURES OF HUMAN,
REMARK      1 TITL 2 YEAST, AND OAT UBIQUITIN
REMARK      1 REF J.BIOL.CHEM. V. 262 6396 1987
REMARK      1 REFN ISSN 0021-9258
.....
REMARK 525 THE SOLVENT MOLECULES HAVE CHAIN IDENTIFIERS THAT
REMARK 525 INDICATE THE POLYMER CHAIN WITH WHICH THEY ARE MOST
REMARK 525 CLOSELY ASSOCIATED. THE REMARK LISTS ALL THE SOLVENT
REMARK 525 MOLECULES WHICH ARE MORE THAN 5A AWAY FROM THE
REMARK 525 NEAREST POLYMER CHAIN (M = MODEL NUMBER;
REMARK 525 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE
REMARK 525 NUMBER; I=INSERTION CODE):

```

```

REMARK 525
REMARK 525 M RES CSSEQI
REMARK 525 HOH A 93 DISTANCE = 5.65 ANGSTROMS
REMARK 525 HOH A 114 DISTANCE = 5.45 ANGSTROMS
REMARK 525 HOH A 126 DISTANCE = 5.71 ANGSTROMS
DBREF 1UBQ A 1 76 UNP P62988 UBIQ_HUMAN 1 76
SEQRES 1 A 76 MET GLN ILE PHE VAL LYS THR LEU THR GLY LYS THR ILE
SEQRES 2 A 76 THR LEU GLU VAL GLU PRO SER ASP THR ILE GLU ASN VAL
SEQRES 3 A 76 LYS ALA LYS ILE GLN ASP LYS GLU GLY ILE PRO PRO ASP
SEQRES 4 A 76 GLN GLN ARG LEU ILE PHE ALA GLY LYS GLN LEU GLU ASP
SEQRES 5 A 76 GLY ARG THR LEU SER ASP TYR ASN ILE GLN LYS GLU SER
SEQRES 6 A 76 THR LEU HIS LEU VAL LEU ARG LEU ARG GLY GLY
FORMUL 2 HOH *58(H2 O)
HELIX 1 H1 ILE A 23 GLU A 34 1
HELIX 2 H2 LEU A 56 TYR A 59 5
SHEET 1 BET 5 GLY A 10 VAL A 17 0
SHEET 2 BET 5 MET A 1 THR A 7 -1
SHEET 3 BET 5 GLU A 64 ARG A 72 1
SHEET 4 BET 5 GLN A 40 PHE A 45 -1
SHEET 5 BET 5 LYS A 48 LEU A 50 -1
CRYST1 50.840 42.770 28.950 90.00 90.00 P 21 21 21 4
ORIGX1 1.000000 0.000000 0.000000 0.00000
ORIGX2 0.000000 1.000000 0.000000 0.00000
ORIGX3 0.000000 0.000000 1.000000 0.00000
SCALE1 0.019670 0.000000 0.000000 0.00000
SCALE2 0.000000 0.023381 0.000000 0.00000
SCALE3 0.000000 0.000000 0.034542 0.00000
ATOM 1 N MET A 1 27.340 24.430 2.614 1.00 9.67 N
ATOM 2 CA MET A 1 26.266 25.413 2.842 1.00 10.38 C
ATOM 3 C MET A 1 26.913 26.639 3.531 1.00 9.62 C
ATOM 4 O MET A 1 27.886 26.463 4.263 1.00 9.62 O
ATOM 5 CB MET A 1 25.112 24.880 3.649 1.00 13.77 C
ATOM 6 CG MET A 1 25.353 24.860 5.134 1.00 16.29 C
ATOM 7 SD MET A 1 23.930 23.959 5.904 1.00 17.17 S
ATOM 8 CE MET A 1 24.447 23.984 7.620 1.00 16.11 C
ATOM 9 N GLN A 2 26.335 27.770 3.258 1.00 9.27 N
ATOM 10 CA GLN A 2 26.850 29.021 3.898 1.00 9.07 C
ATOM 11 C GLN A 2 26.100 29.253 5.202 1.00 8.72 C
ATOM 12 O GLN A 2 24.865 29.024 5.330 1.00 8.22 O
ATOM 13 CB GLN A 2 26.733 30.148 2.905 1.00 14.46 C
ATOM 14 CG GLN A 2 26.882 31.546 3.409 1.00 17.01 C
ATOM 15 CD GLN A 2 26.786 32.562 2.270 1.00 20.10 C

```

Record type	Atom info		Amino acid (ligand) name	Chain ID	Residue ID	Atom coordinates			Occupancies		Element name
	number	type				X	Y	Z		Temperature factor (b factor)	
ATOM	1	N	MET A	1	1	27.340	24.430	2.614	1.00	9.67	N
ATOM	2	CA	MET A	1	1	26.266	25.413	2.842	1.00	10.38	C
ATOM	3	C	MET A	1	1	26.913	26.639	3.531	1.00	9.62	C
ATOM	4	O	MET A	1	1	27.886	26.463	4.263	1.00	9.62	O
ATOM	5	CB	MET A	1	1	25.112	24.880	3.649	1.00	13.77	C
ATOM	6	CG	MET A	1	1	25.353	24.860	5.134	1.00	16.29	C
ATOM	7	SD	MET A	1	1	23.930	23.959	5.904	1.00	17.17	S
ATOM	8	CE	MET A	1	1	24.447	23.984	7.620	1.00	16.11	C
ATOM	9	N	GLN A	2	2	26.335	27.770	3.258	1.00	9.27	N
ATOM	10	CA	GLN A	2	2	26.850	29.021	3.898	1.00	9.07	C
ATOM	11	C	GLN A	2	2	26.100	29.253	5.202	1.00	8.72	C
ATOM	12	O	GLN A	2	2	24.865	29.024	5.330	1.00	8.22	O
ATOM	13	CB	GLN A	2	2	26.733	30.148	2.905	1.00	14.46	C
ATOM	14	CG	GLN A	2	2	26.882	31.546	3.409	1.00	17.01	C
ATOM	15	CD	GLN A	2	2	26.786	32.562	2.270	1.00	20.10	C
.....											
Heavy atoms	HETATM	648	O	HOH A	121	39.646	23.959	9.699	0.41	18.25	O
	HETATM	649	O	HOH A	122	34.405	45.181	13.420	0.87	26.13	O
	HETATM	650	O	HOH A	123	26.517	24.300	27.592	0.41	21.02	O

PyMOL graphic user interface

Control panel

Inputs for commands

List of objects

Display area

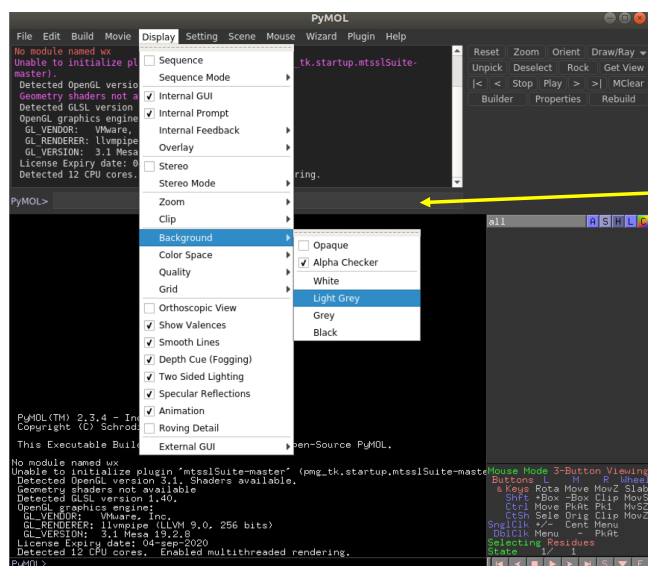
Mouse control

Inputs for commands

Movie control

S = sequence mode/view
F = full screen mode/view

Change your background in PyMOL



Change your default background color from black to white or other color:

1. Using “display” toolbar to change background color
2. Using command line by typing: `bg_color blue` (any pre-defined colors, see next page).

Available pre-defined color codes in PyMOL

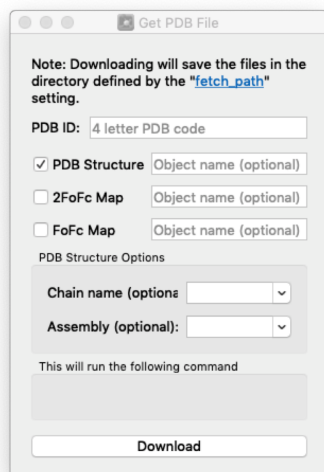
actinium	darksalmon	iodine	palecyan	sodium
aluminum	dash	iridium	palegreen	splitpea
americium	deepblue	iron	paleyellow	strontium
antimony	deepolive	krypton	palladium	sulfur
aquamarine	deeppurple	lanthanum	phosphorus	tantalum
argon	deepsalmon	lawrencium	pink	teal
arsenic	deepsalmon	lead	platinum	technetium
astatine	deepteal	lightblue	plutonium	tellurium
atomic	default	lightmagenta	polonium	terbium
auto	density	lightorange	potassium	thallium
barium	deuterium	lightpink	praseodymium	thorium
berkelium	dirtyviolet	lightteal	promethium	thulium
beryllium	dubnium	lime	protactinium	tin
bismuth	dysprosium	limegreen	pseudoatom	titanium
black	einsteinium	limon	purple	tungsten
blue	erbium	lithium	purpleblue	tv_blue
bluewhite	europium	lonepair	radium	tv_green
bohrium	fermium	lutetium	radon	tv_orange
boron	firebrick	magenta	raspberry	tv_red
brightorange	fluorine	magnesium	red	tv_yellow
bromine	forest	manganese	rhodium	uranium
brown	francium	marine	rhodium	vanadium
cadmium	gadolinium	meitnerium	rubidium	violet
calcium	gallium	mendelevium	ruby	violetpurple
californium	germanium	mercury	ruthenium	warmpink
carbon	gold	molybdenum	rutherfordium	wheat
cerium	gray	neodymium	salmon	white
cesium	green	neon	samarium	xenon
chartreuse	greencyan	neptunium	sand	yellow
chlorine	gray	nickel	scandium	yelloworange
chocolate	hafnium	niobium	seaborgium	ytterbium
chromium	hassium	nitrogen	selenium	yttrium
cobalt	helium	nobelium	silicon	zinc
copper	holmium	olive	silver	zirconium
curium	hotpink	orange	skyblue	
current	hydrogen	osmium	slate	
cyan	indium	oxygen	smudge	

Loading files

First, you need to know your PDB ID of your protein or have your own PDB files prepared.

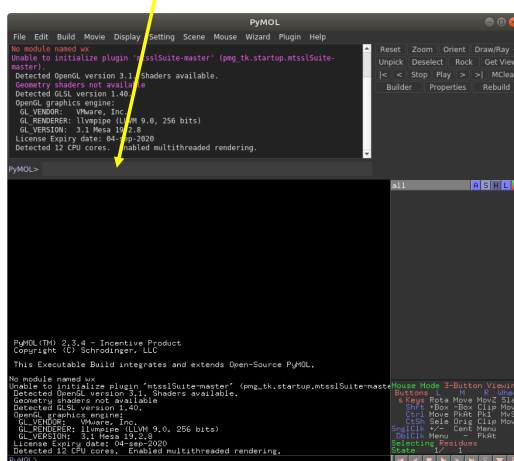
(1) From toolbar:

File → Get PDB, then type the PDB ID



(2) At command line, type following:

fetch 1ubq (1ubq is the PDB ID)

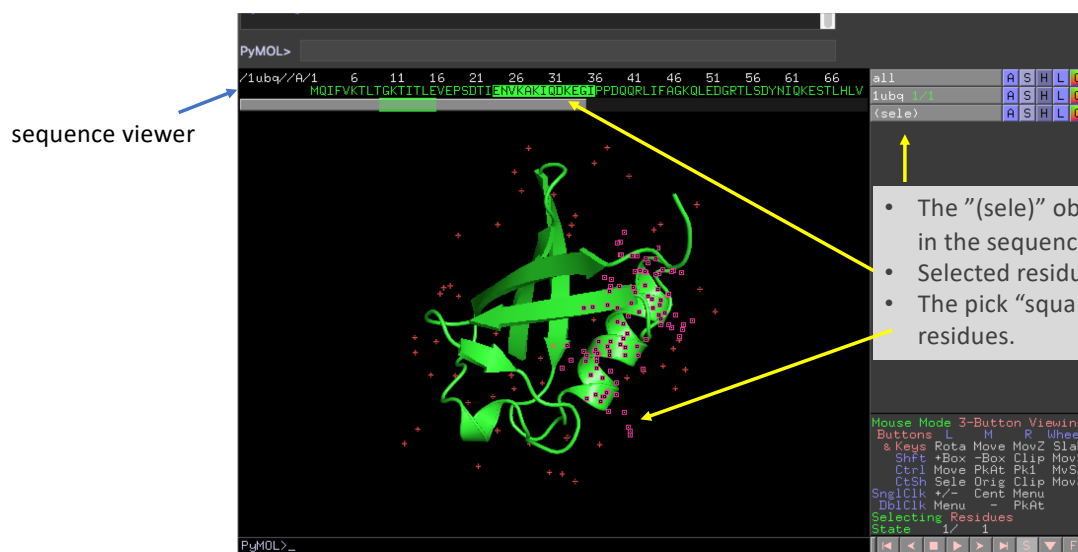


(3) Load PDB file:

File → Open, find your PDB files in local drive

Sequence viewer

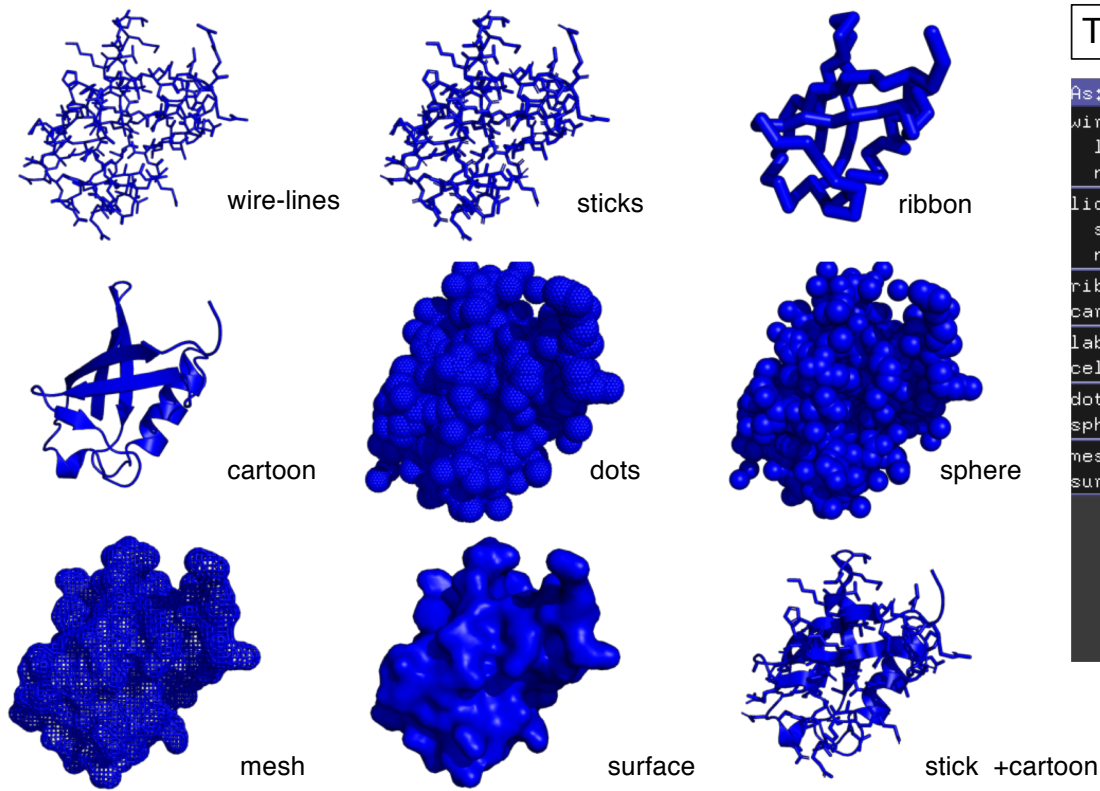
First, you need to know your PDB ID of your protein or have your own PDB files prepared.



sequence viewer

- The "(sele)" object is what I just selected in the sequence viewer panel.
- Selected residues are highlighted.
- The pick "squares" indicate selected residues.

Click "S" to turn on sequence viewer



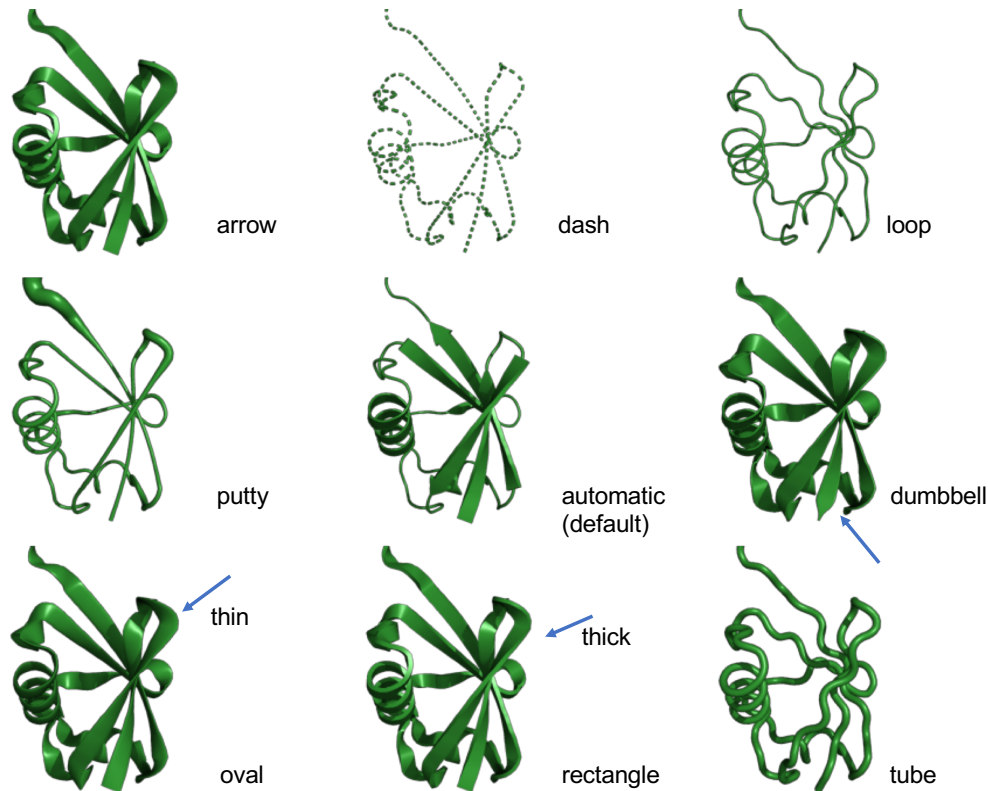
The “Show”

As:	Show:
wire	as
lines	wire
nonbonded	lines
licorice	nonbonded
sticks	licorice
nb_spheres	sticks
ribbon	nb_spheres
cartoon	ribbon
label	cartoon
cell	label
dots	cell
spheres	dots
mesh	spheres
surface	mesh
	surface
	organic
	main chain
	side chain
	disulfides
	valence

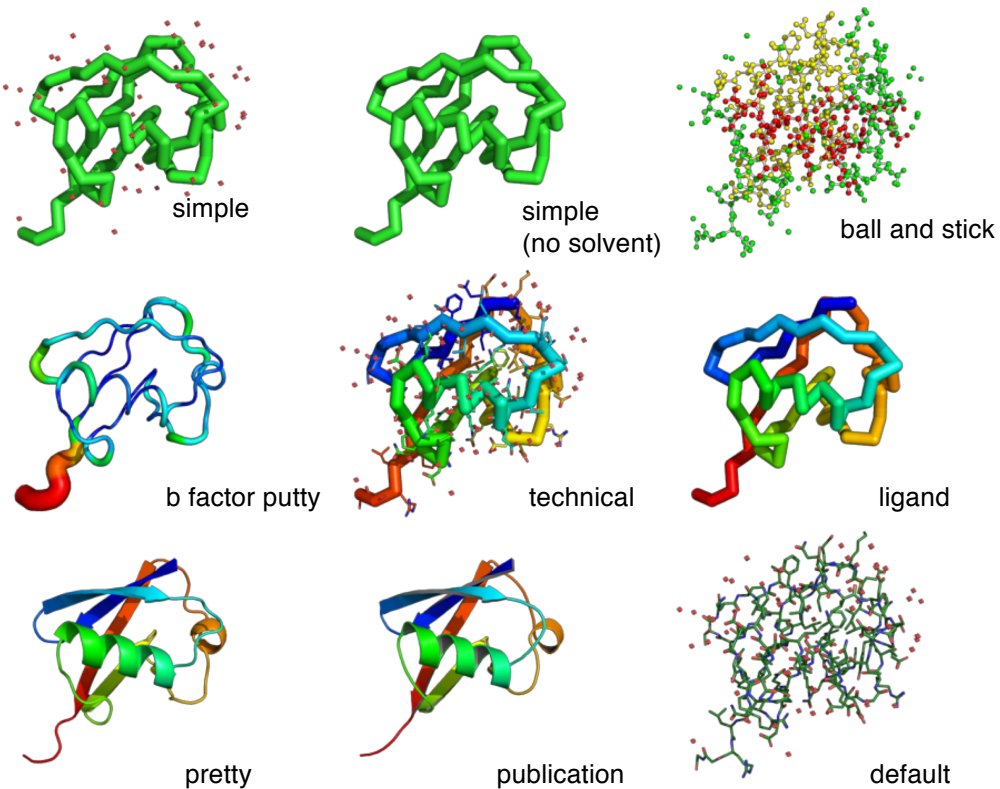
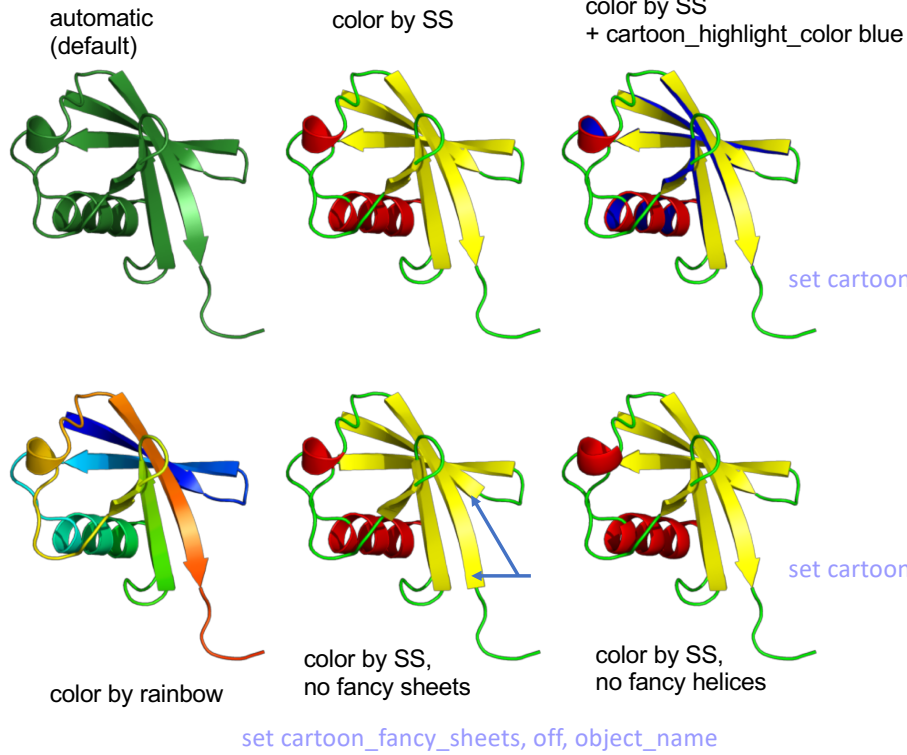
Cartoon options:

- arrow
- dash
- loop
- putty
- skip
- automatic
- dumbbell
- oval
- rectangle
- tube

Type:
cartoon option, object_name



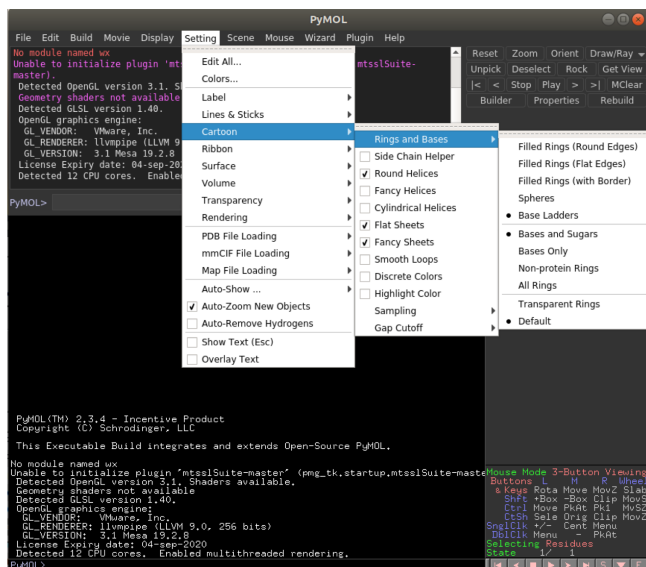
More on “cartoon”



The “Action --> Preset”

```
Preset:
classified
simple
simple (no solvent)
ball and stick
b factor putty
technical
ligands
ligand sites
pretty
pretty (with solvent)
publication
publication (with solvent)
protein interface
default
```

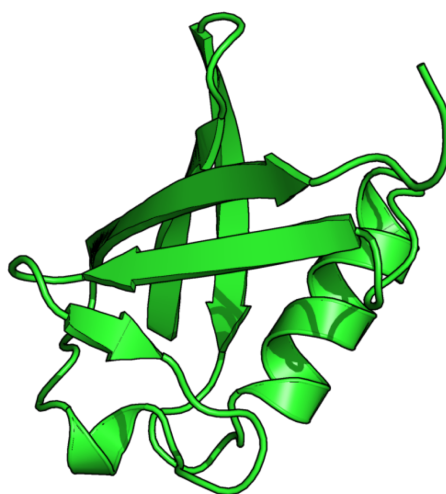
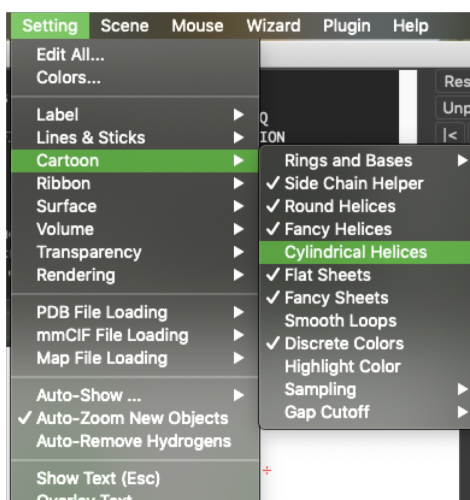
The “cartoon” mode



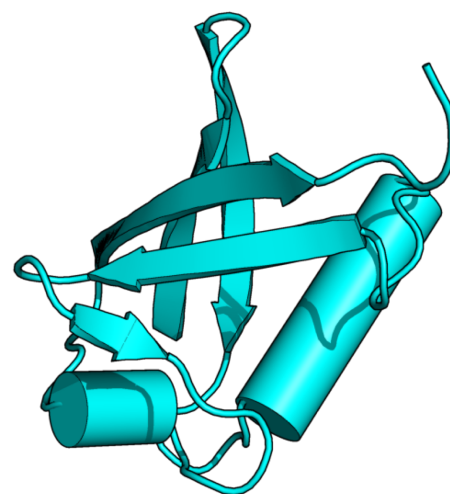
Cartoon settings:

- Useful to turn on highlight colors, discrete color.
- For a large/complex molecule, present helices in “cylindrical view” is less distracting
- DNA or compound rings can be “filled” with colors.

The “cylindrical helices” option in cartoon



cartoon automatic

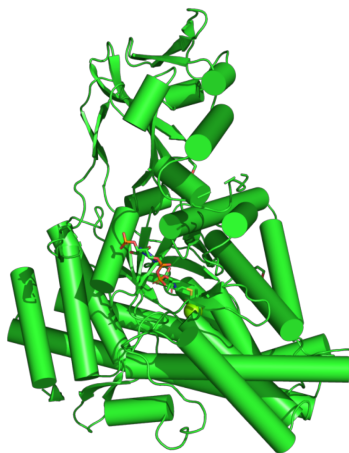


cartoon automatic
+ set cartoon_cylindrical_helices, on, object

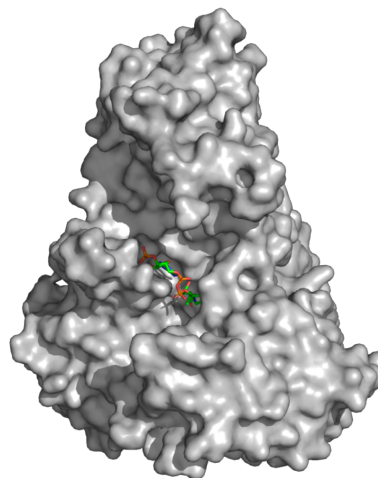
To present a ligand or binding site, there are options...



cartoon automatic
+ ligand in sticks



cartoon automatic
+ ligand in sticks +
set cartoon_cylindrical_helices, on, object



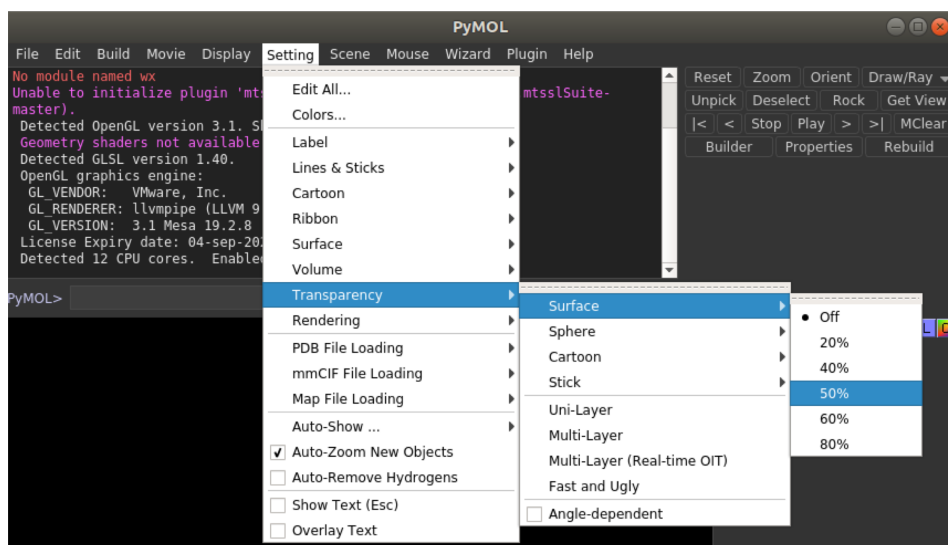
Surface view of protein
+ ligand in sticks

Very busy in the binding site

Reduced distractions

More focused on ligand

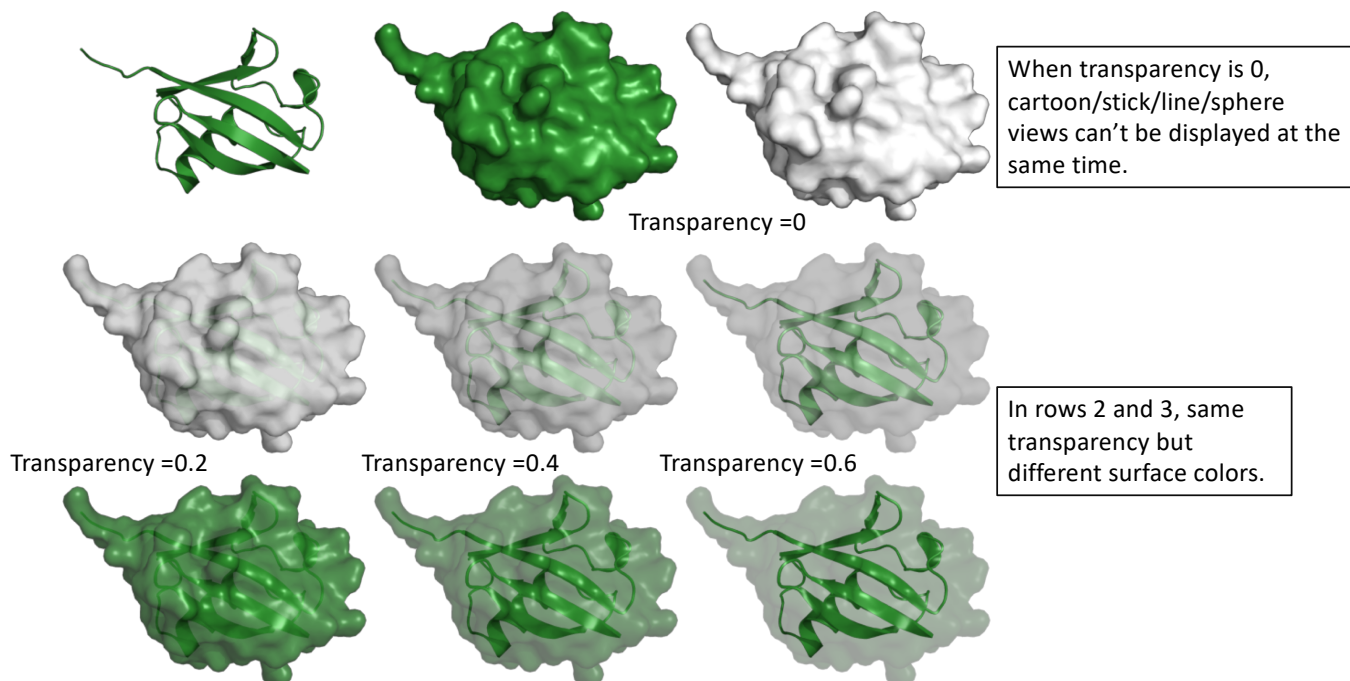
The “transparency” settings



Transparency settings

- Multiple options, partial transparent surface is commonly used.
- See examples in next page.

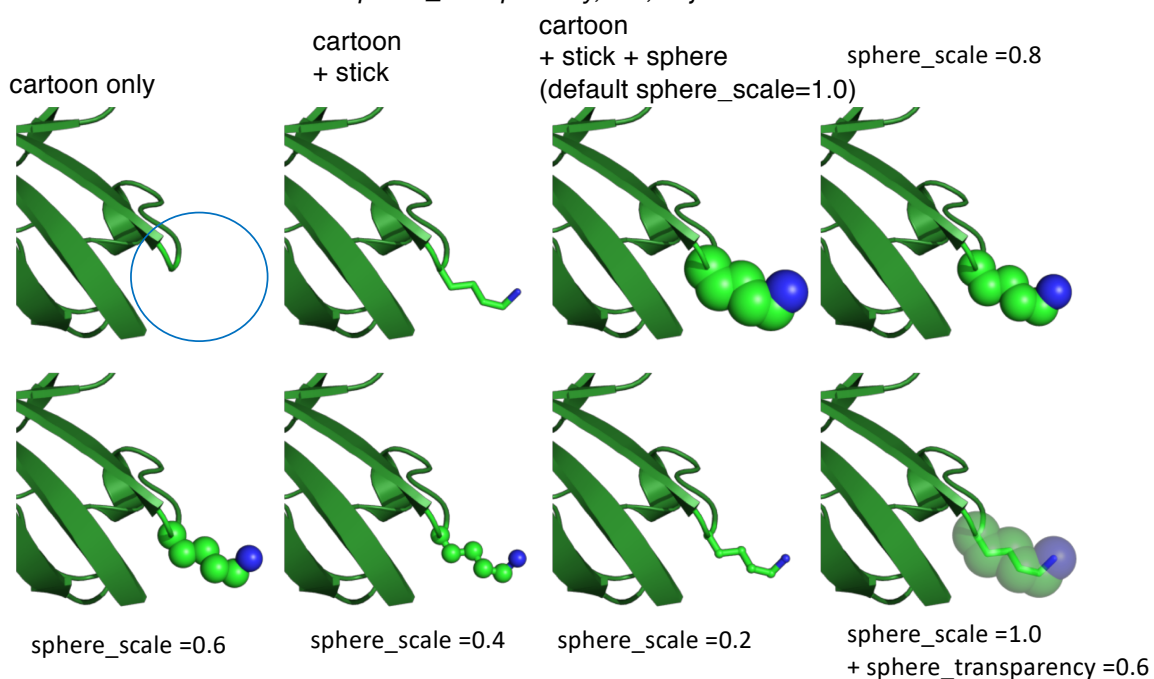
White surface with 0.2-0.4 (20-40%) transparency works well with cartoon view



The “sphere” mode

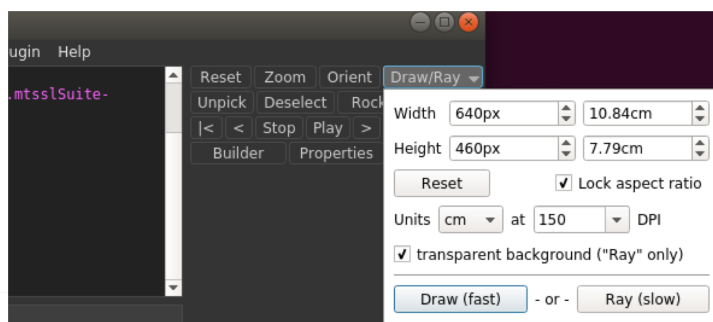
commands:

- `set sphere_scale, 0.2, object`
- `set sphere_transparency, 0.4, object`



Rendering your molecules

- Draw or Ray
- Draw is quick, with background color
- Ray takes time (depends on cases), but rendering light, shadows and can output a “transparent” image.
- Choose your image size (default or larger/smaller)
- Choose your image resolution (300 dpi for publication quality).



draw



ray

The “ray_trace” options give you “outlines” of the molecules

When using “Ray” to generate light and shadows, default trace mode is 0. There are “trace” options.
The commands include:

set ray_trace_mode, [0, 1, 2, 3] (choose a number)
set ray_trace_color, color_name

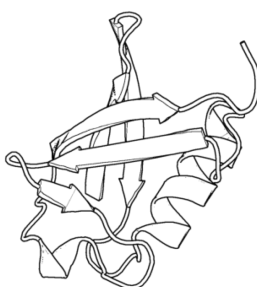
Here are the example.



ray_trace_mode
=0



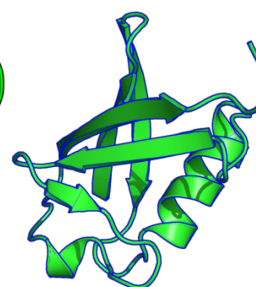
ray_trace_mode
=1



ray_trace_mode
=2

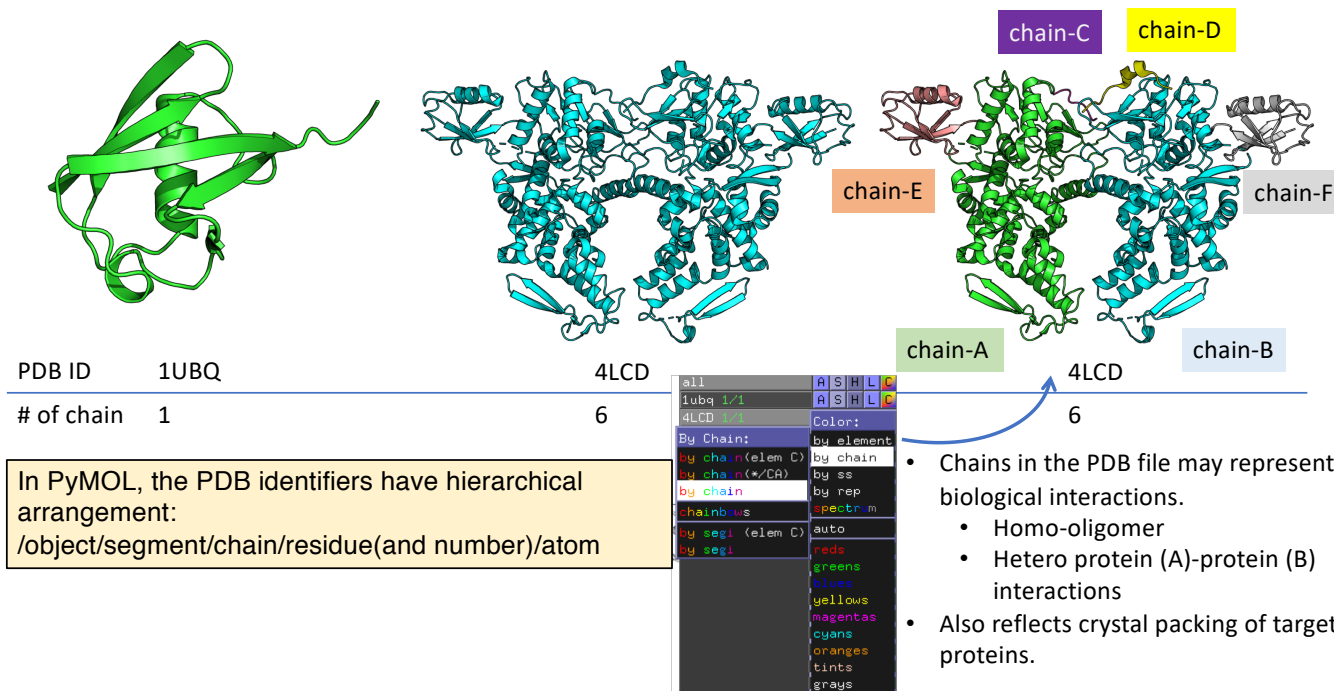


ray_trace_mode
=3

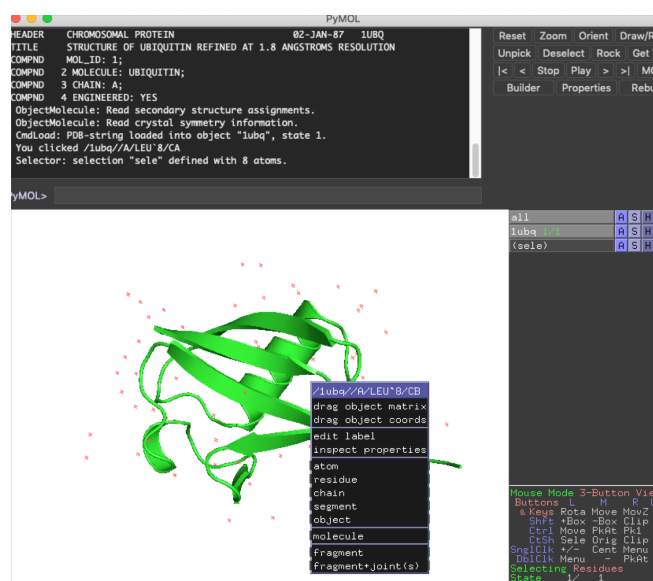


ray_trace_mode
=1
ray_trace_color
=blue

Chains and colors



PDB identifiers



hierarchical PDB identifiers:

/object/segment/chain/residue(and number)/atom

/1ubq//A/LEU`8/CB

At here:

Object name is 1ubq

segment is empty as this is a simple protein crystal data, no ligands included.

Chain ID is A

Residue I picked is Leu`8

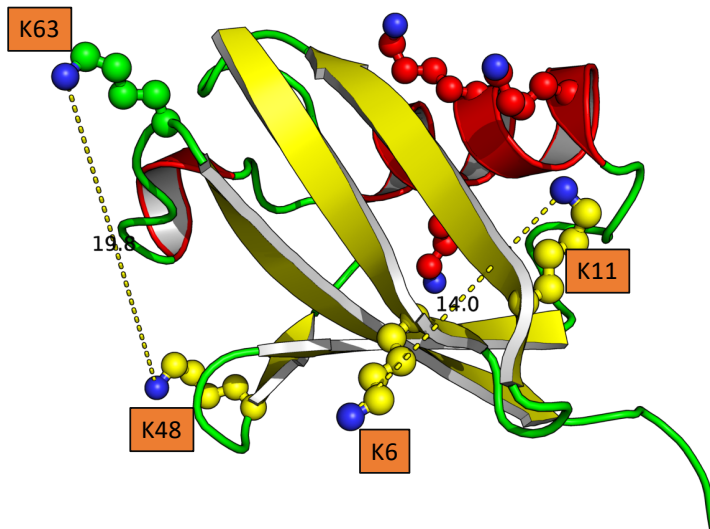
Atom I picked is CB (Cβ)

drag object matrix
drag object coords
edit label
inspect properties
atom
residue
chain
segment
object
molecule
fragment
fragment+joint(s)

Step-by-step to illustrate your target protein

Using ubiquitin (PDB 1UBQ) as an example

Steps highlighted in light blue are “commands”, others can be done by clicking the function “S/H/C” buttons”



1. Load 1ubq
2. Hide waters
3. `set cartoon_highlight_color, grey80`
4. select lys_all, resn lys
5. Show "lys_all" sticks
6. Show "lys_all" spheres
7. Color "lys_all" by element
8. `set sphere_scale, 0.4`
9. Measure distance:
 - command: `dist` (or `distance`)
 1. `dist K6-K11, ///6/NZ, ///11/NZ`
 2. `dist K48-K63, ///48/NZ, ///63/NZ`

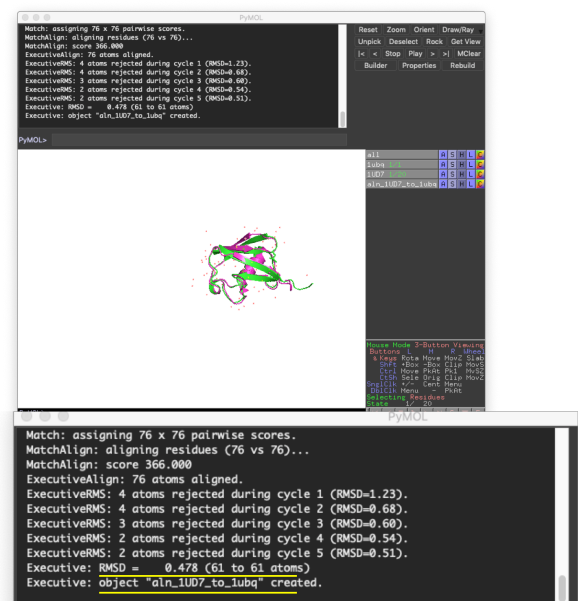
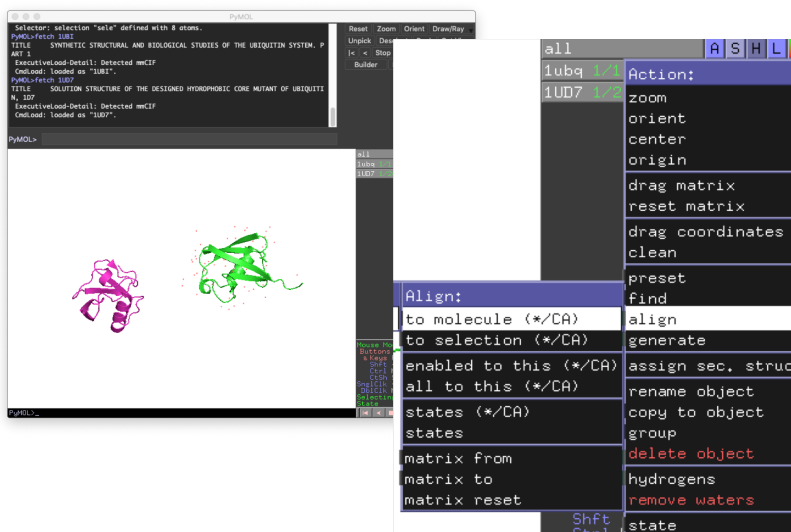
hierarchical PDB identifiers:

/object/segment/chain/residue(and number)/atom

Align your interested proteins

Examples: many ubiquitin structures are available at RCSB PDB.

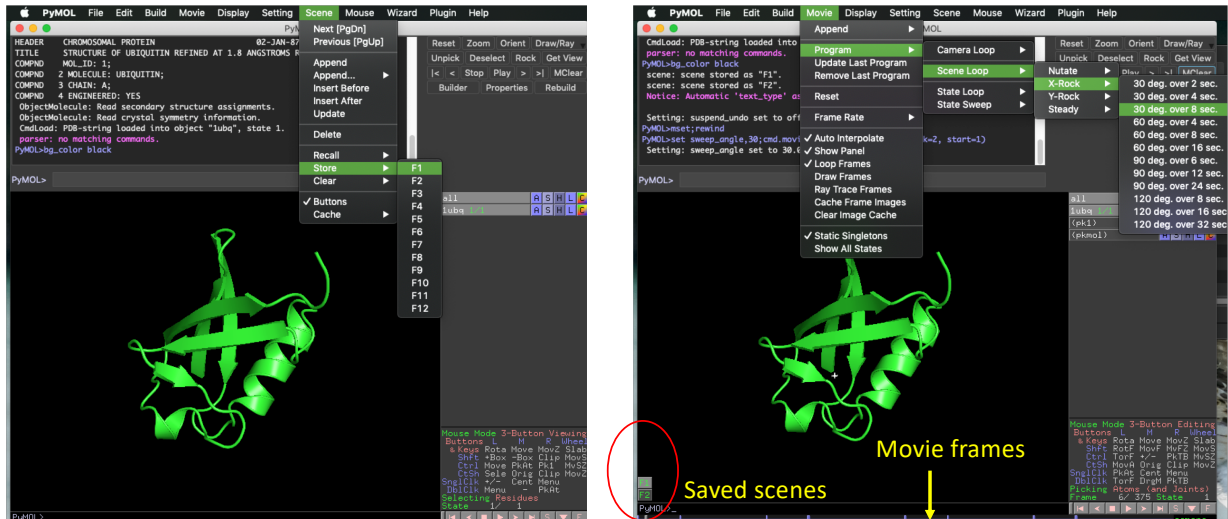
When loading **1UBQ** (crystal structure) and **1UD7** (NMR structure), the two molecules are not in the same 3D space (not even close to each other).



Make your own movies by PyMOL

Tips:

1. Save "scenes"
2. Set up movie loop times
3. Export as Movie (From File -> Export Movie as)



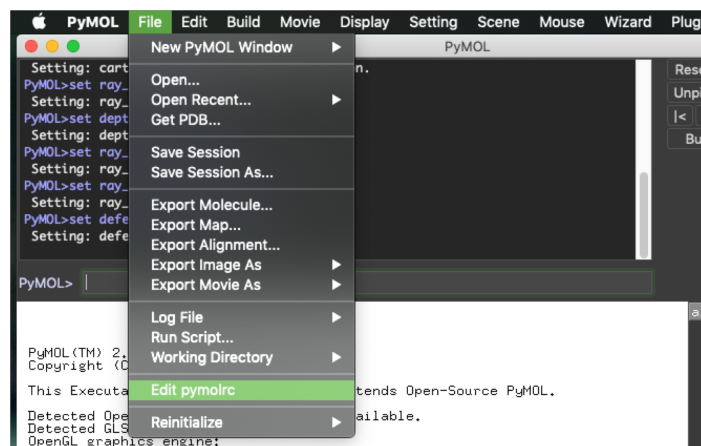
Pymolrc: a pre-defined environmental settings for your PyMOL

<https://pymolwiki.org/index.php/Pymolrc>

```
# save fetched PDB files here
set fetch_path,
/Users/user/Documents/PDB

# set background color white
bg white

set ribbon_width, 8
set antialias, 2
set cartoon_fancy_helices, 1
set cartoon_discrete_colors, 1
set cartoon_side_chain_helper, 1
set ray_trace_mode, 1
set depth_cue, 0
set ray_trace_fog, 0
set ray_opaque_background, 0
set defer_builds_mode, 3
```



Save session files: save pymol works as xxx.pse format, can be reopened at saved states. Also good to transport to another computer for other PyMOL users.

