The PDB and Molecular Visualization

Biochemistry Boot Camp 2023
Session #9
Nick Fitzkee
nfitzkee@chemistry.msstate.edu

Properties of PDB Files

- Experimental methodology:
  - X-Ray: Typically more precise
  - NMR: Need lots of “restraints;” sometimes hard to assess quality
  - Cryo-Electron Microscopy: Large complexes

- Most PDB files are solved using X-ray crystallography
  - 191,000 structures total
  - 166,000 are crystal structures

Building an X-Ray Structure

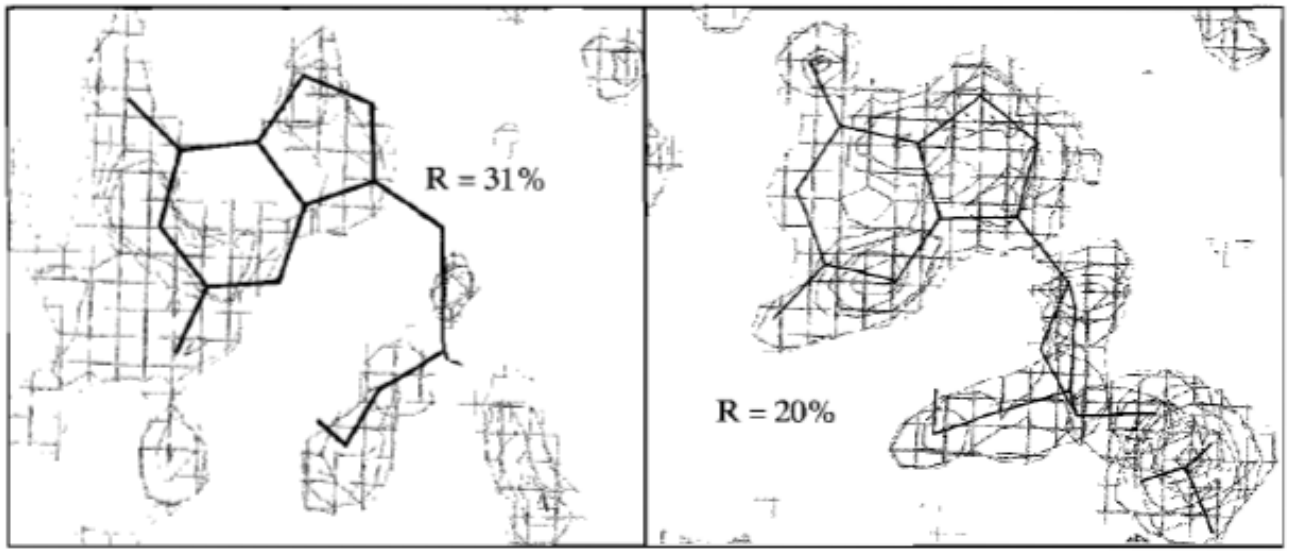


Figure 6.31 Effect of refinement on structure. The guanine nucleotide of a DNA fragment is shown with its electron density map prior to refinement and after refinement. Prior to refinement, the *R* factor is 31%. The structure is refined against the data to an *R* factor of 20%, which is one criterion of a good fit of the model to the data.

- **At first:** look for gross structural features (helix, backbone), then add side chains
- Molecular mechanics are used to help refine positions

Guidelines for X-Ray Quality

- **R-factor:** Less than 25% (or 0.25; ideally, less than 20%)
- **R-Free:** Bigger than R, but smaller than 25%
- **Resolution:** Less than 2.5 Å, but think about how much you need (1.5 Å usually very good)
  - At ~1 Å hydrogens become visible
- **Validation:** No clashes, good torsions, etc.

Guidelines for X-Ray Quality

- **Not too Much Bound Water**:
  - Higher resolution typically means reduced molecular motions, more bound water will be visible
  - However, water molecules can also be used to "sop up" electron density of unknown origin

- **Rules of Thumb (Approximate)**

1.5 Å or better	Worse than 2.0 Å
0.5-1.5 waters per AA	0.2-0.8 waters per AA
5-15 waters per kDa	2-8 waters per kDa

- **Example:** 7B4O (R_{free} = 0.19, 1.25 Å): 564 water molecules for 153 x 4 = 612 AA total (65 kDa)

Nittinger, *et al.* (2015) *J. Chem. Inf. Model.* 55: 771−783

Protein Data Bank Revisited

- http://rcsb.org/

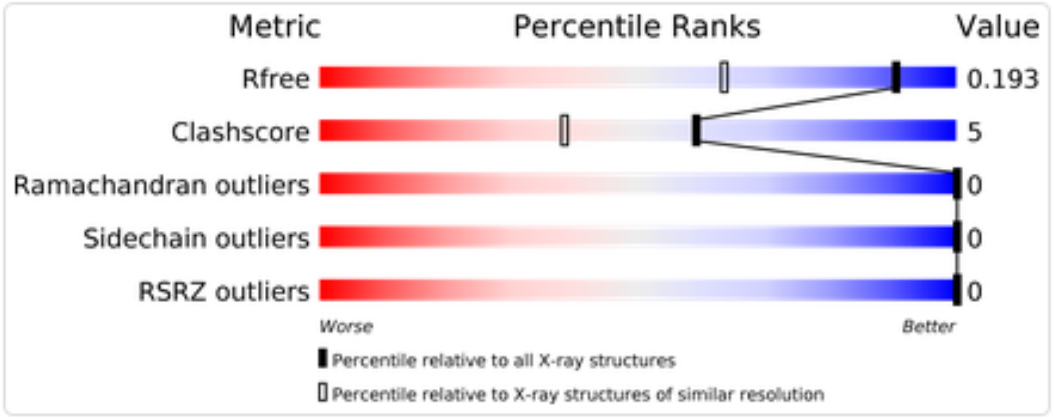
- **Input:** Protein name, PDB ID, authors, etc.

- **Output:** 3D coordinates of protein structures
  - Author information on methods
  - Cofactors and other information

Assessing a Crystal Structure: 3TJW

Structure Validation

View Full Validation Report or Ramachandran Plots



- PDB contains a lot of useful information for determining how good a crystal structure is
- **Things to look at:** R, R_{free}, resolution, structure validation

Experimental Data & Validation

Experimental Data

Method: X-RAY DIFFRACTION
Resolution: 1.46 Å
R-Value Free: 0.190
R-Value Work: 0.150
Space Group: F 2 2 2

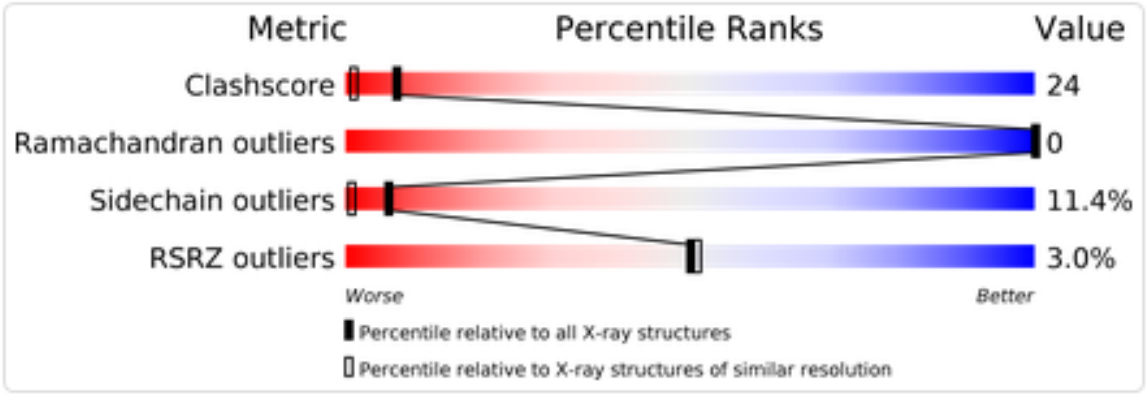
Unit Cell:

Length (Å)	Angle (°)
a = 38.656	α = 90.00
b = 88.112	β = 90.00
c = 88.663	γ = 90.00

Assessing a Crystal Structure: 1SNC

Structure Validation

View Full Validation Report or Ramachandran Plots



- Not every protein will have R_{free}

Experimental Data & Validation

Experimental Data

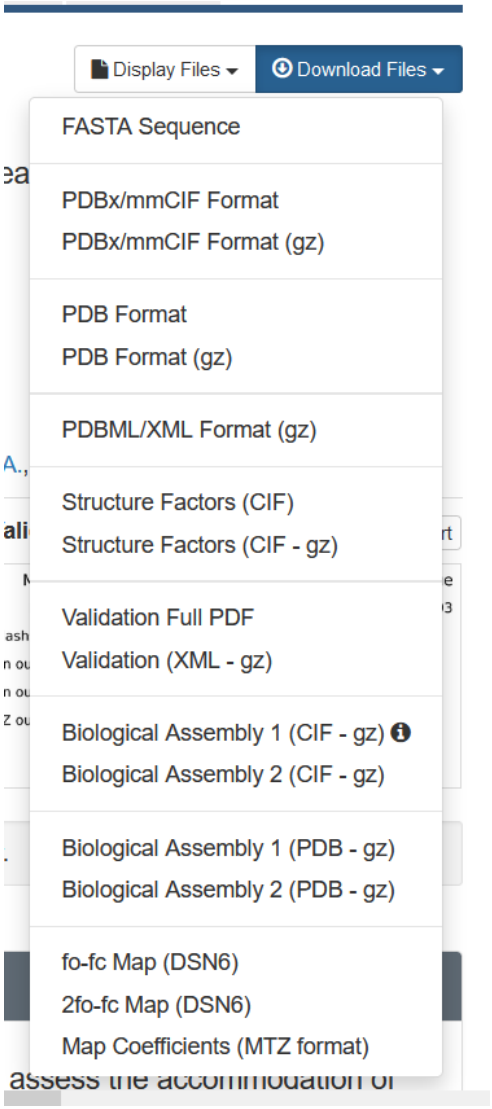
Method: X-RAY DIFFRACTION
Resolution: 1.65 Å
Space Group: P 4_1

Unit Cell:

Length (Å)	Angle (°)
a = 48.000	α = 90.00
b = 48.000	β = 90.00
c = 63.500	γ = 90.00

PDB Files: A Closer Look

- Standard PDB File (older)
- mmCIF File (newer, but sometimes not supported)
- Complete Validation Report



Obtaining PyMOL

(not for this class; see next slide)

- Obtain a license file by registering at:

  https://pymol.org/edu/

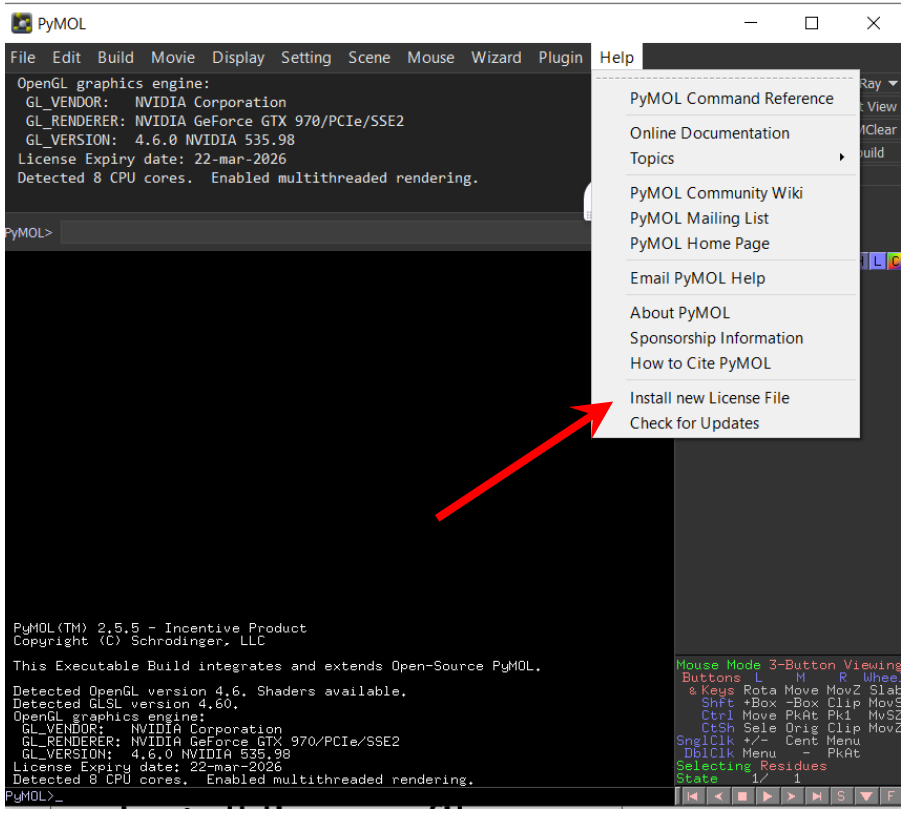
- Educational licenses must be renewed every year and have some limitations in functionality!

- **Mac or PC:** Fill in the form and follow the instructions for downloading; follow the standard procedure

- An open source can be obtained through most Linux distributions (e.g, apt, yum); this does not require a license!

Obtaining PyMOL

- Download PyMOL license file from Boot Camp Website
- Download PyMOL from link below: https://pymol.org/2/#download
- Install license file after PyMOL is installed



Obtaining PyMOL on Linux/Debian

- PyMOL is open source, and an alternative is downloading compiled versions on Linux

- Linux versions (i.e., versions compiled independently from Schrodinger) are not crippled, even the latest version. They do lack some Schrodinger-specific features.

- On Debian/Ubuntu (easy, but you need admin access):
  `apt-get install pymol`

- This is harder for CentOS/Red Hat – ask me if you are interested!

Alternatives to PyMOL

- UCSF Chimera (https://www.cgl.ucsf.edu/chimera/)
  - Absolutely free, although tricky to learn. A large and growing user base.

- VMD (https://www.ks.uiuc.edu/Research/vmd/)
  - Optimized for looking at MD simulations

- More information about history at http://www.umass.edu/microbio/rasmol/history.htm

Quick Intro to PyMOL

A Gallery of PyMOL Tricks
(we'll cover some of this)

- The following slides contain useful tricks I've picked up over the years; we won't have time to cover all of them

- Don't forget to visit the "actual" PyMOL Gallery on the PyMOL website:

  https://pymolwiki.org/index.php/Gallery

PyMOL Tools: Protein Alignment

- Open two PDB files within PyMOL: 1F8A and 1PIN (prolyl isomerase)

```
fetch 1f8a
fetch 1pin
```

- Align both structures using the following command (uses similar residues as cues):

```
align 1f8a and resi 80-160, 1pin and resi 80-160
```

PyMOL Tools: Complex Assemblies

- Before you fetch, type

```
set assembly, 1
fetch 3bw1
```

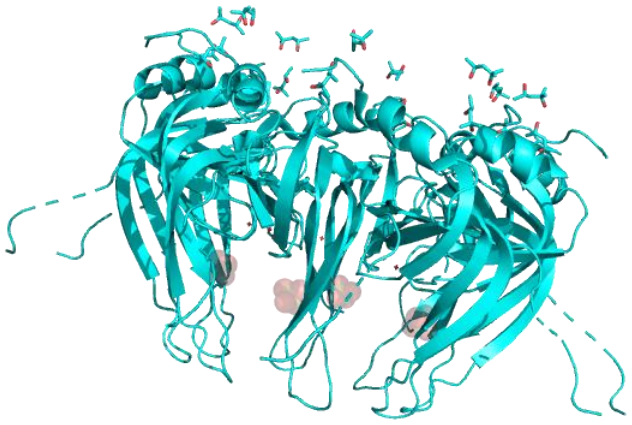
- This will grab the biological assembly if one is defined

- See more:

https://pymolwiki.org/index.php/Assembly

```
set assembly, ""
```

```
set assembly, 1
```



PyMOL Tools: Measurement

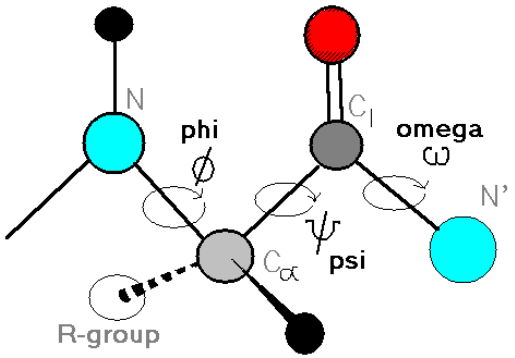
- Measure distances and dihedral angles using PyMOL (Wizard → Measurement, then click four atoms)

- Recall that phi (ϕ) is defined as the dihedral angle defined by:

$$C'_{i-1} - N_i - C_{\alpha,i} - C'_i$$

- Similarly, psi (ψ) is:

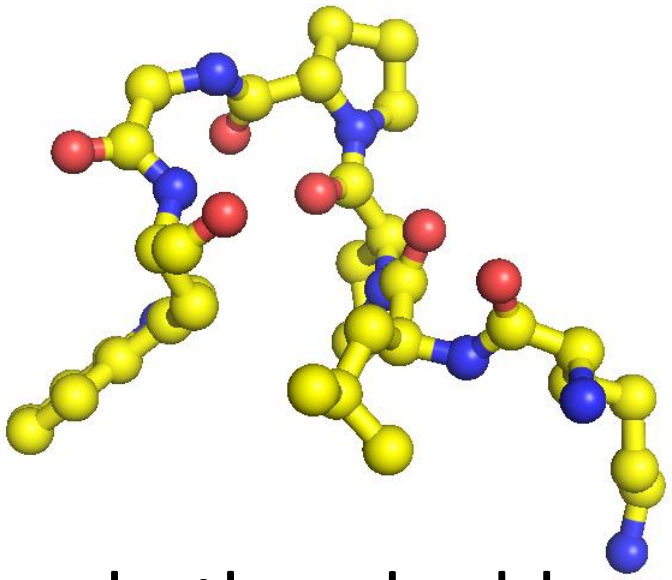
$$N_i - C_{\alpha,i} - C'_{i+1} - N_{i+1}$$



PyMOL Tools: Ball-and Stick

- Combining small spheres with stick model can produce a pleasing result

```
show sticks
show spheres
set sphere_scale, 0.3
set valence, 0
```



- The final command toggles whether double bonds are shown

PyMOL Tools: Transparent Surfaces

- Show surfaces using

  `show surface, selection`

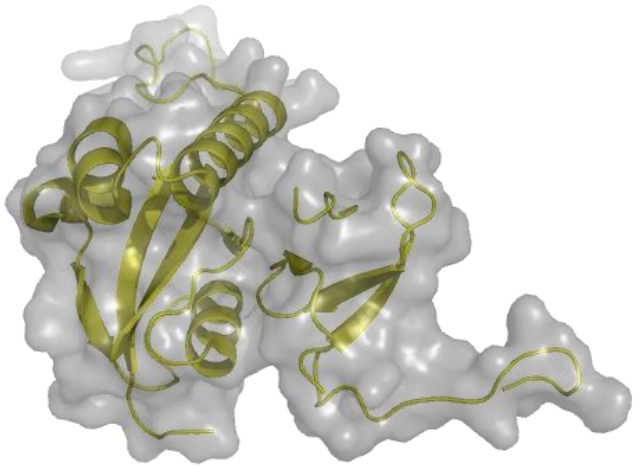
- Transparency can be applied to different renderings (e.g. cartoon vs. spheres)

  `set sphere_transparency, 0.65`

- Surface transparency is a global property, so a separate object must be created if mixed results are desired (see next slide)

PyMOL Tools: Transparent Surfaces

- **Example:** Cartoons under a semi-transparent surface
  1. First, create a duplicate object by clicking 1f8a (A) → Copy to Object → New
  2. Then show surface for the new object (obj01)
     ```
     show surface, obj01
     color grey, obj01
     ```
  3. Set transparency for the new object
     ```
     set transparency, 0.5, obj01
     ```
  4. Show cartoons for the original object
     ```
     hide everything, 1f8a
     show cartoon, 1f8a
     color yellow, 1f8a
     ```



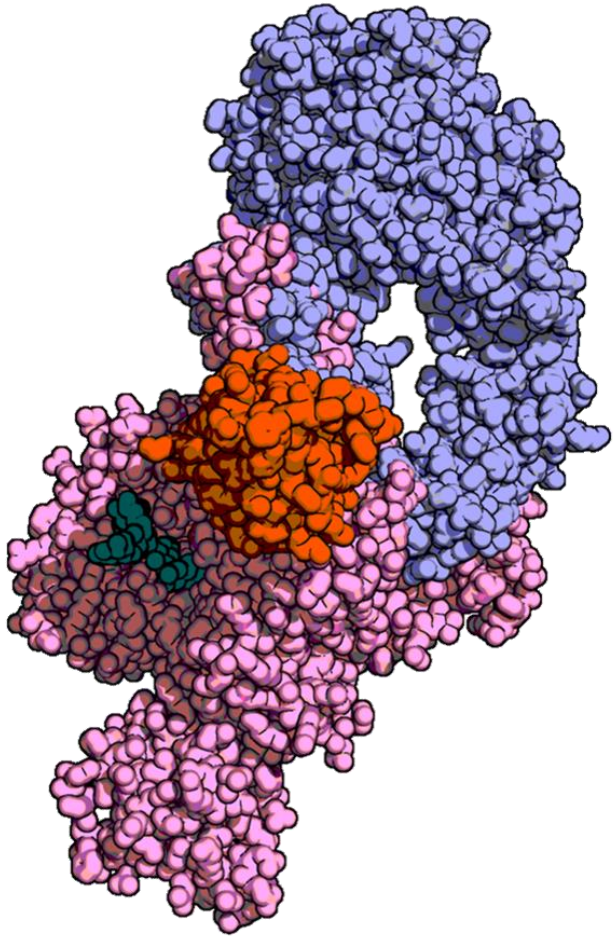
PyMOL Tools: David Goodsell-esque

1. Create your object and color domains accordingly
2. Show the molecule as “bulky sticks”

```
as sticks
set valence, 0
set stick_radius, 1.7
```

3. Ray trace options (“gain” makes black outlines, “mode” uses quantized color, “color” sets the outline color)

```
set specular, off
set depth_cue, 0
set ray_trace_gain, 0
set ray_trace_mode, 3
set ray_trace_color, black
```



PDB: 6RU5, Credit: @NikoMcCarty (Twitter)

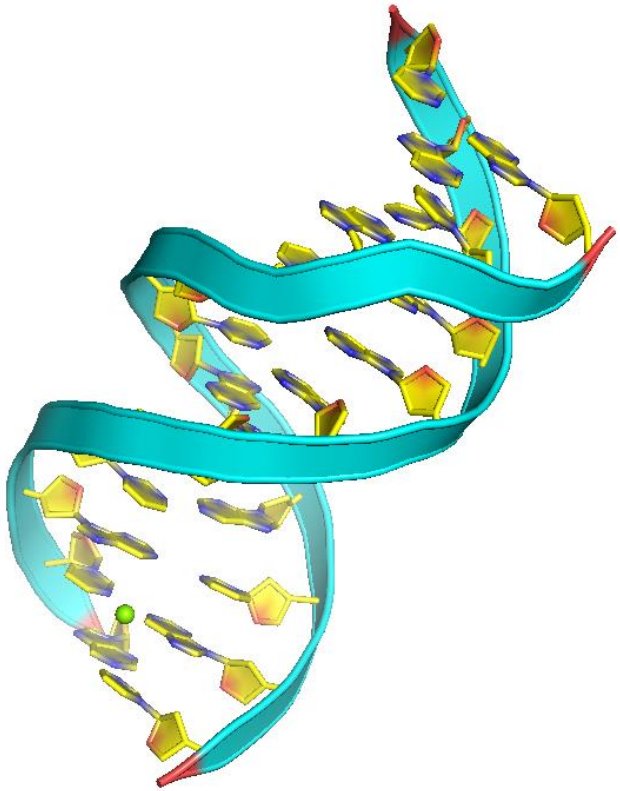
PyMOL Tools: Awesome DNA

- Grab the Dickerson dodecamer

```
fetch 4c64
```

- Reasonable first attempt

```
set cartoon_ring_mode, 3
cartoon dumbbell
set cartoon_dumbbell_radius, 0.25
color cyan, elem P
```



- My go-to pages for DNA:

https://kpwu.wordpress.com/2006/10/22/pymol-fancy-dna-helix-and-filled-rings/
https://kpwu.wordpress.com/2011/07/26/pymol-example-of-dna/
https://pymolwiki.org/index.php/Examples_of_nucleic_acid_cartoons

PyMOL Tools: Ray Tracing Tips

- **Ray Tracing:** Simulation of photon light paths, reflection, and scattering to give a photorealistic image

- New Draw/Ray Panel
  - Fine control over resolution/image size
  - Automatic removal of background (a bit buggy)
  - Display → Background: Can set color

- Always set slab to largest view area before ray tracing (mouse wheel up until fog is gone, or `set depth_cue, 0`)

- Save image: Use File → Export Image As… (or click Draw/Ray again)
  - **Always** save the file for publications/posters: The clipboard feature seems to reduce quality (especially in PowerPoint)

Summary

- PDB files are complex models, derived from experimental data
  - Need to assess these models, too
- Atom names and properties are stored in the PDB, can be used to investigate structures
- PyMOL can make interesting pictures, but it is also a very powerful analytical tool