# The PDB and Molecular Visualization

Biochemistry Boot Camp 2018
Session #8
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

- Most PDB files are solved using X-ray crystallography
  - 141,000 structures total
  - 126,000 are crystal structures

2

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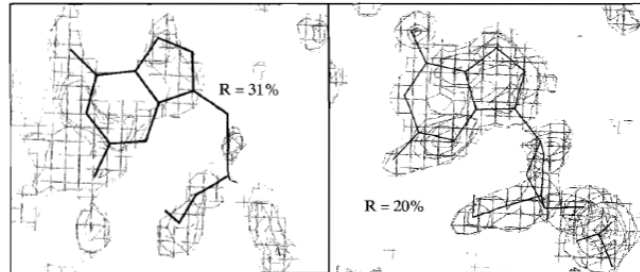
Figure 6.31 Effect of refinement on structure. The guanine nucleotide of a DNA fragment is shown with its electron density map prior to refinement and after refinement. Prior to refinement, the *R* factor is 31%. The structure is refined against the data to an *R* factor of 20%, which is one criterion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

3

Guidelines for X-Ray Quality

- **R-factor:** Less than 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.
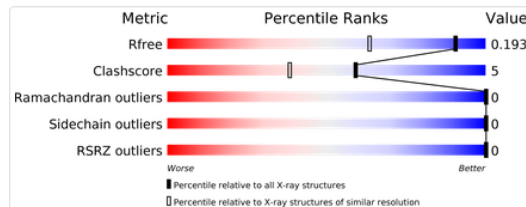- **Water**: 2-5 molecules / kDa

4

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

5

Assessing a Crystal Structure: 3TJW

Structure Validation

View Full Validation Report or Ramachandran Plots

Metric	Percentile Ranks	Value
Rfree		0.193
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		0

Worse — Better
■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

- 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
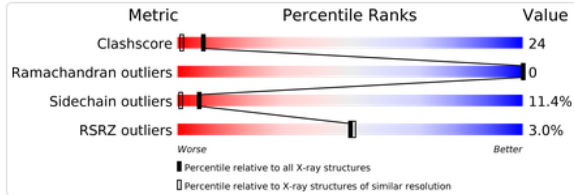R-Value Work: 0.150
Space Group: F 2 2 2

Unit Cell:

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

6

3

Assessing a Crystal Structure: 1SNC

Structure Validation

View Full Validation Report or Ramachandran Plots

Metric	Percentile Ranks	Value
Clashscore		24
Ramachandran outliers		0
Sidechain outliers		11.4%
RSRZ outliers		3.0%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

- Not every protein will have R_{free}

Experimental Data & Validation

Experimental Data

Method: X-RAY DIFFRACTION
Resolution: 1.65 Å
Space Group: P 4₁

Unit Cell:

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

7

PDB Files: A Closer Look

8

Obtaining PyMOL

- Recent versions of educational PyMOL (EduPyMOL) are great, although time-limited.

- Older educational builds can be obtained by registering at:
  http://www.pymol.org/educational/

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

- **Linux:** Open source (full version) can be obtained through most distributions (e.g, apt, yum)

9

Obtaining PyMOL

- PyMOL Website: http://pymol.org/ep
  Username: jun2018
  Password: caffeine

- Older versions for Windows XP, etc. (version 1.3r1) are available at:
  Windows: http://goo.gl/0mRH2O
  Mac: http://goo.gl/U68Hoc
  Linux: http://goo.gl/HMZvPU

10

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 <u>not</u> 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
  ```

11

Obtaining PyMOL on CentOS

- On Redhat/Centos (harder, and you still need admin access):

  ```
  1. sudo yum update
  2. sudo yum install epel-release
  3. sudo yum install subversion gcc gcc-c++
     kernel-devel python-devel tkinter python-pmw
     glew-devel freeglut-devel libpng-devel
     freetype-devel libxml2-devel apbs
  ```

- The commands above install the necessary libraries for compiling PyMOL
- The final command should be typed entirely on one line.

12

Obtaining PyMOL: CentOS (cont.)

- Obtain the latest tar.bz2 file for PyMOL from
  http://sourceforge.net/projects/pymol/?source=typ_redirect

- Extract the source code using the following command:
  `tar xfj <source file name>`

- This will extract the source code to a directory with a similar
  name as the tar.bz2 archive (e.g. `pymol-v2.1.0.tar.bz2`
  extracts to `pymol-v2.1.0`)

13

Obtaining PyMOL: CentOS (cont.)

- Finally, compile and install python. Change to the python
  source directory and type (will take some time):

  ```
  1. python setup.py build
  2. sudo python setup.py install
  ```

- You can then run PyMOL from the setup directory (you can
  also create a link):

  ```
  1. cd setup       (still from within the PyMOL directory)
  2. ./pymol
  ```

14

Alternatives to PyMOL

- UCSF Chimera (https://www.cgl.ucsf.edu/chimera/)
  - Absolutely free, although tricky to learn. Given the draconian practices of Schrodinger, this may be the future (for Boot Camp).

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

- MolScript (http://www.avatar.se/molscript/)
  - Old, and difficult to master (editing text files is required), but its images are still as good as any other software.

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

15

Quick Intro to PyMOL

16

PyMOL Tools: Protein Alignment

- Download two files from the web page: 1F8A and 1PIN
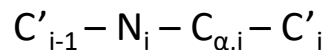 ```
  fetch 1f8a
  fetch 1pin
  ```

- Open both PDB files in PyMOL

- Align both structures using the following command (uses similar residues as cues):
  ```
  align 1f8a and resi 80-160, 1pin and resi 80-160
  ```

17

## **PyMOL Tools:** Measurement

- Measure distances and dihedral angles using PyMOL (Wizard → Measurement)

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

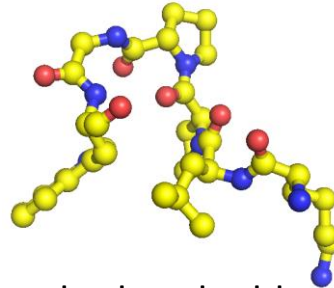$$C'_{i-1} - N_i - C_{\alpha,i} - C'_i$$

18

9

# **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

19

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)

20

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
     ```
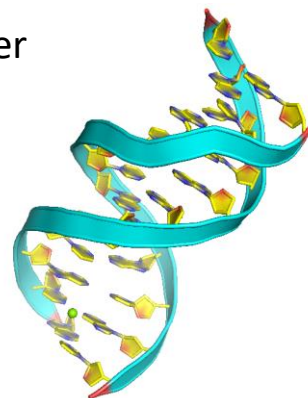
21

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://pymolwiki.org/index.php/Examples_of_nucleic_acid_cartoons

22

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); for best results, maximize slab so no fading
  - Display → Background: Can set color

- Save image: Use File → Export Image As… (or click Draw/Ray again)

23

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

24