

The PDB and Molecular Visualization

Biochemistry Boot Camp 2017
Session #7
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
 - 119,000 structures total
 - 107,000 are crystal structures

Building an X-Ray Structure

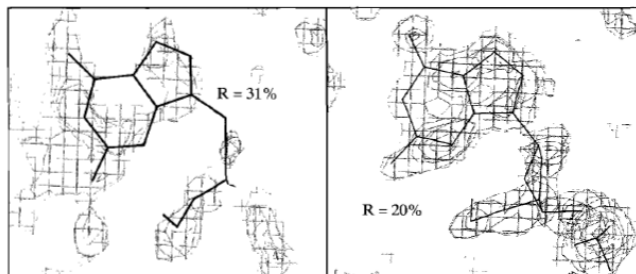


Figure 6.31 Effect of refinement on structure. The guanine nucleotide of a DNA fragment is shown with its electron density map prior to refinement and after refinement. Prior to refinement, the R factor is 31%. The structure is refined against the data to an R factor of 20%, which is one criterion of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

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

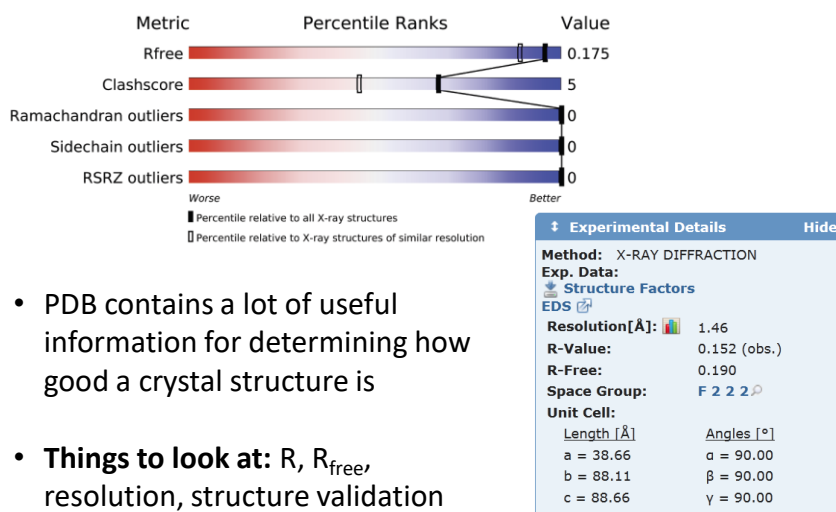
4

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

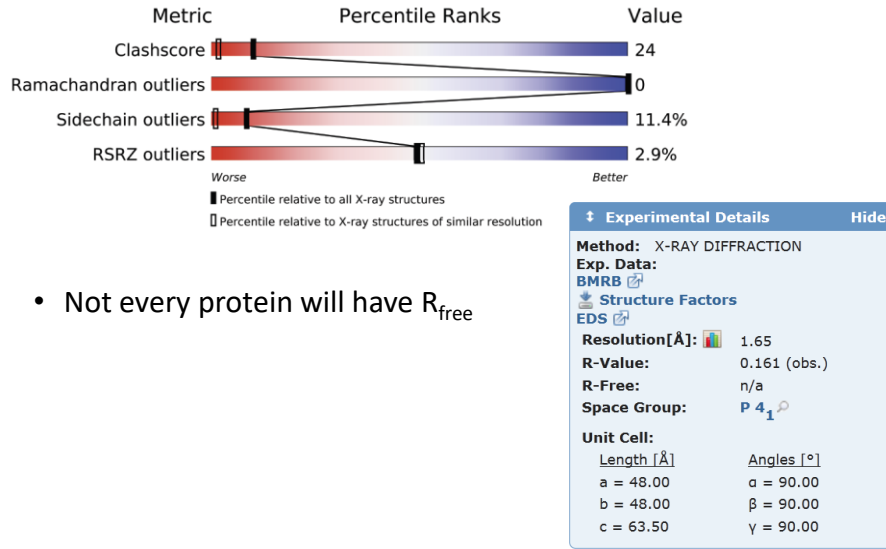
Assessing a Crystal Structure: 3TJW



- 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

6

Assessing a Crystal Structure: 1SNC



PDB Files: A Second Look

Obtaining PyMOL

- Be wary of the most recent educational release. It's buggy, and critical features (like ray tracing) are disabled.
- Older educational builds can be obtained by registering at:
<http://www.pymol.org/educational/>
- I recommend version 1.3r1, although some features are still buggy (e.g. torsion angle calculation)

9

Obtaining PyMOL

- PyMOL Website: <http://pymol.org/ep>
Username: jun2017
Password: histidine
- For Boot Camp, you can download PyMOL versions here (version 1.3r1, no password required):
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 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`

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 xjf <source file name>
```
- This will extract the source code to a directory with a similar name as the tar.bz2 archive (e.g. `pymol-v1.8.6.0.tar.bz2` extracts to `pymol-v1.8.6.0`)

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

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

16