The PDB and Molecular Visualization

Biochemistry Boot Camp
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

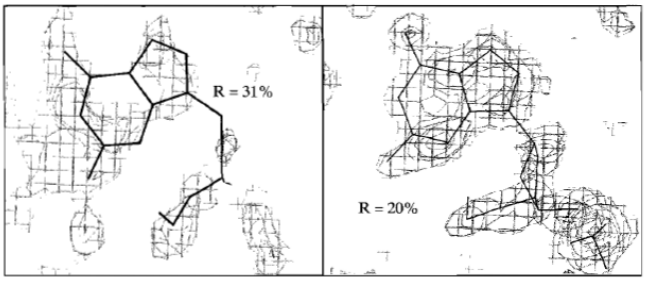


Figure 6.31 Effect of refinement on structure. The guanine nucleotide of a DNA fragment is shown with its electron density map prior to refinement and after refinement. Prior to refinement, the *R* factor is 31%. The structure is refined against the data to an *R* factor of 20%, which is one criterion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

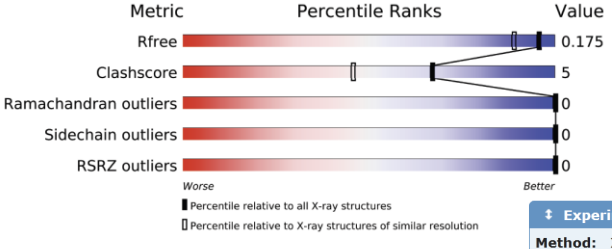
4

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

Experimental Details Hide

Method: X-RAY DIFFRACTION
Exp. Data:
Structure Factors
EDS

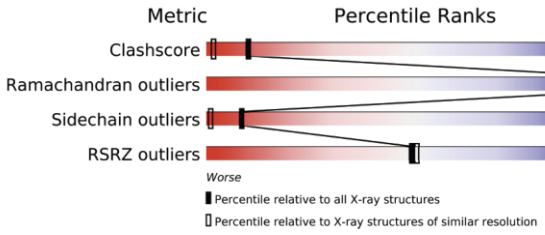
Resolution[Å]:	1.46
R-Value:	0.152 (obs.)
R-Free:	0.190
Space Group:	F 2 2 2

Unit Cell:

Length [Å]	Angles [°]
a = 38.66	α = 90.00
b = 88.11	β = 90.00
c = 88.66	γ = 90.00

6

Assessing a Crystal Structure: 1SNC



Experimental Details Hide

Method: X-RAY DIFFRACTION
Exp. Data:
BMRB
Structure Factors
EDS

Resolution[Å]:	1.65
R-Value:	0.161 (obs.)
R-Free:	n/a
Space Group:	$P\,4_1$

Unit Cell:

Length [Å]	Angles [°]
a = 48.00	α = 90.00
b = 48.00	β = 90.00
c = 63.50	γ = 90.00

7

- Not every protein will have R_{free}

PDB Files: A Second Look

8

Obtaining PyMOL

- Do NOT use the most recent educational release (1.7.4.4). 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

- For Boot Camp, you can download PyMOL versions here (version 1.3r1):
- Windows: http://goo.gl/0mRH2O
- Mac: http://goo.gl/U68Hoc
- Linux: http://goo.gl/HMZvPU
- PyMOL is open source, and an alternative is downloading compiled versions on Linux (e.g. "`apt-get install pymol`" on debian-based systems)
- Linux versions (i.e., versions compiled independently from Schrodinger) are not crippled, even the latest version. They do lack some Schrodinger-specific features.

10

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

11

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

12