Introduction to X-ray Crystallography

Jesse Rinehart, PhD
Than you for contributions to this lecture:

Yong Xiong, PhD
Yale University Department of Molecular Biophysics & Biochemistry

Recommended Course @ Yale: MB&B 720a
Macromolecular Structure and Biophysical Analysis

Additional Resources:

CMCC Home Page: http://spdbv.vital-it.ch/TheMolecularLevel/CMCC/index.html


“Introduction to X-ray crystallography” http://vimeo.com/7643687

http://ucxray.berkeley.edu/~jamesh/movies/
movies demonstrating diffraction, resolution, data quality, and refinement.
There's a \textit{limit} to how small an object can be seen under a light microscope.

\textbf{The diffusion limit:} you can not image things that are much smaller than the wavelength of the light you are using.

- The wavelength for visible light is measured in hundreds of nanometers, while atoms are separated by distances of the order of 0.1nm, or 1Å.

\textbf{We need to use x-rays to resolve atomic features.}

Distances between atoms are small:
Lab x-ray sources use CuKα radiation. Wavelength = 1.54 Å.
Synchrotron radiation wavelengths in the range 0.5 Å - 2.5 Å.
The 2014 Nobel Prize in Chemistry: Eric Betzig, W.E. Moerner, and Stefan Hell
“The development of super-resolved fluorescence microscopy"
Experimental Determination of Atomic Resolution Structures

**X-ray**
- Direct detection of atom positions
- Crystals

**NMR**
- Indirect detection of H-H distances
- In solution

Other methods for determining protein structures:
- EM, Cryo-EM, ESR/Fluorescence
Determination of Protein Crystal Structure

1. Subcloning
2. Expression
3. Purification
4. Crystallization: Hanging Drop with Protein

Reservoir with Precipitant

X-ray → crystal → diffraction → detector → refinement modeling

http://www.noble.org/PlantBio/Wang/crystallography.html
Why Crystals?

X-rays are scattered by electrons, too weak to record scattering from a single molecule. Crystals are therefore used because they present many molecules (N) in exactly the same orientation. The scattering from each of the N molecules interferes constructively to give a measurable diffraction pattern (enhanced \( \sim N^2 \) fold).

Yong Xiong
Some Crystallization Methods:

Vapor diffusion
Hanging-drop

Sitting-drop

Batch:
micro batch under oil

Dialysis

Energy $\Delta G$

Time

Critical Nuclei
Specific aggregates
Non-specific aggregates
Protein in solution

Yong Xiong
Data Collection

4-Circle Goniometer (Eulerian or Kappa Geometry)
Synchrotron X-ray Sources

Lab x-ray sources @ 1.54 Å VS. Synchrotron @ 0.5 Å - 2.5 Å.
Image of diffraction
Most famous X-ray diffraction pattern
Most famous X-ray diffraction pattern
The information we get from a single diffraction experiment

(a) space group of the crystal
(b) unit cell dimensions

analyze the pattern of the reflections

Cubic
\[ a = b = c, \]
\[ \alpha = \beta = \gamma = 90^\circ \]

Hexagonal
\[ a = b \neq c, \]
\[ \alpha = \beta = 90^\circ, \gamma = 120^\circ \]

Trigonal
\[ a = b \neq c, \]
\[ \alpha = \beta = 90^\circ, \gamma = 120^\circ \]

Tetragonal
\[ a = b \neq c, \]
\[ \alpha = \beta = \gamma = 90^\circ \]

How to understand symmetry?
Crystal = lattice + unit cell content
(asymmetric units (asu) content)
The phase problem: $F(hkl)$ is a complex vector. Measured diffraction data give the amplitude $|F(hkl)|$. The phase information $\alpha_{hkl}$ is lost!

How important are amplitude and phase?

Fourier Duck and his Fourier transform
Phase is color coded

Fourier Cat and his Fourier transform
Phase is color coded

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)| e^{-2\pi i (hx + ky + lz) + i\alpha_{hkl}}$$

Duck phase and Cat amplitude

Cat phase and Duck amplitude

Yong Xiong
Electron density map

Building a structure model
The importance of resolution

Experimental electron density map created from multi-wavelength data collected at SSRL beam line 1-5 on a Gold derivative of tetanus C fragment.

Example of high quality Experimental data where very little refinement has been applied to fit a tyrosine into the density map.

http://www.ruppweb.org/Xray/101index.html
Crystal structure of the nucleosome core particle at 2.8 Å resolution

Karolin Luger, Armin W. Mäder, Robin K. Richmond, David F. Sargent & Timothy J. Richmond

Institut für Molekularbiologie und Biophysik ETHZ, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

The X-ray crystal structure of the nucleosome core particle of chromatin shows in atomic detail how the histone protein octamer is assembled and how 146 base pairs of DNA are organized into a superhelix around it. Both histone/histone and histone/DNA interactions depend on the histone fold domains and additional, well ordered structure elements extending from this motif. Histone amino-terminal tails pass over and between the gyres of the DNA superhelix to contact neighbouring particles. The lack of uniformity between multiple histone/DNA-binding sites causes the DNA to deviate from ideal superhelix geometry.
The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution

Nenad Ban,¹* Poul Nissen,¹* Jeffrey Hansen,¹ Peter B. Moore,¹,² Thomas A. Steitz¹,²,³†

Yale's Thomas Steitz shared 2009 Nobel Prize in Chemistry for this structure
Protein Structure Databases

Where does protein structural information reside?
- PDB: http://www.rcsb.org/pdb/
- FSSP: http://www.ebi.ac.uk/dali/fssp/
- SCOP: http://scop.mrc-lmb.cam.ac.uk/scop/
- CATH: http://www.biochem.ucl.ac.uk/bsm/cath_new/

# of structures
2011: 70,813
2014: 97,180
PDB Growth from 2011 to 2014: \( \Delta \) structures: 26,367

compared to \( \Delta \) protein interactions: 107,018

What species are the structures from?

- Homo sapiens (24792)
- Escherichia coli (4813)
- Mus musculus (4223)
- Saccharomyces cerevisiae (2295)
- Bos taurus (2270)
- Rattus norvegicus (2056)
- Escherichia coli K-12 (1848)
- Other (53253)

Which methods?

- X-RAY (16470)
- Solution NMR (3268)
- Electron Microscopy (57)
- Solution Scattering (25)
- Other (16)
- Solid-State NMR (9)
- Electron Crystallography (7)
- Neutron Diffraction (5)
- Hybrid (5)
- Fiber Diffraction (3)
Tools for Viewing Structures

- Jmol
- PyMOL
  - http://pymol.sourceforge.net
- Swiss PDB viewer
- Mage/KiNG
  - http://kinemage.biochem.duke.edu/software/king.php
- Rasmol
  - http://www.umass.edu/microbio/rasmol/