

## SCE3153, 03.04.09

### X-RAY CRYSTALLOGRAPHY

Protein structure determination by X-ray crystallography requires successful production of sufficient amounts of pure protein and subsequent crystallisation. The central experiment is diffraction of incident X-rays by the lattice planes of the protein crystal. The necessary condition for diffraction is constructive interference of individual diffracted X-ray beams as formulated in Bragg's law. In contrast to optical microscopy, diffracted X-rays cannot be re-combined by an "X-ray lens". Mathematical procedures in form of Fourier transforms are necessary to perform recombination of diffracted X-ray waves in order to obtain the "micrograph" image. Additionally, the phase information of diffracted X-ray waves is lost during the experiment. This is known as the phase problem and requires a first good guess of phases by either one of three methods (protein model, multiple isomorphous replacement, multiple anomalous diffraction).

#### PROTEIN PURIFICATION

- Nowadays, mostly recombinant proteins are used: overexpression in (bacterial) cells
- Purification of protein material is done in most cases by liquid chromatography

#### PROTEIN CRYSTALLISATION

- Hanging drop vapour diffusion method
- Unit cell is the building block for a crystal; repetition of the unit cell into all three dimensions generates the crystal lattice
- Solvent content in protein crystals is rather high and ranges from 40% - 60%
- There are seven crystal systems (triclinic, monoclinic, orthorhombic, trigonal, tetragonal, hexagonal, cubic) with different metric requirements; by combination of all possible symmetry elements, this yields 32 crystal classes (point groups)
- Including translational components, this yields 14 Bravais lattices by combining the crystal systems with either primitive (P), body-centred (I) or face-centred (A, B, C or F)
- From combination of point groups with translational components one obtains 230 space groups

#### DATA COLLECTION

- Crystal is mounted in a loop and exposed to X-ray beam, while being frozen (100 K) in N<sub>2</sub> stream ("cryo technique")

#### X-ray production

##### *rotating anode generator*

in-house

monochromatic (Cu K $\alpha$ :  $\lambda = 1.5 \text{ \AA}$ )

##### *synchrotron*

Australian Synchrotron (Melbourne), ESRF (Grenoble), etc.

polychromatic (tuneable,  $\lambda = 4.0 - 0.5 \text{ \AA}$ )

#### Diffraction

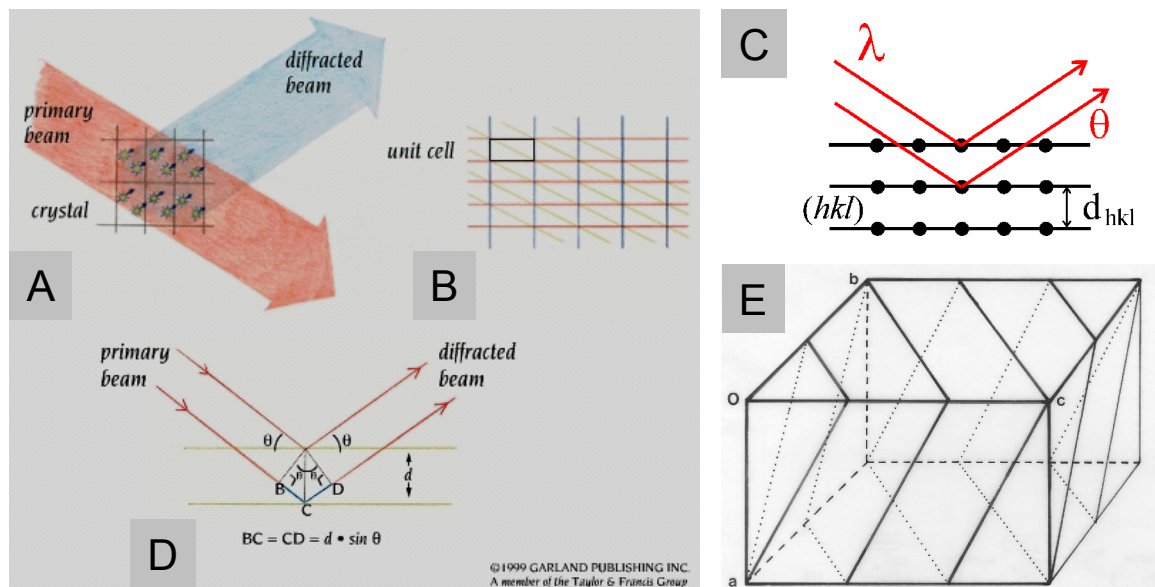
When the wavelength of incident light is comparable to the size of objects in its way, a pattern of destructive and constructive interference is observed: diffraction. In contrast to light in the visible spectrum, X-rays cannot be focussed from scattered radiation, since no material with a refractive index  $n > 1$  for X-rays has been found, so far.

Thus, X-ray diffraction patterns from crystals are acquired by a detector system and the re-combination of diffracted rays has to be carried out mathematically.

#### Theory of diffraction

**Miller indices** ( $hkl$ ) of lattice planes are the number of divisions the set of planes cuts into **a** ( $h$ ), **b** ( $k$ ), and **c** ( $l$ ). Each plane causes a spot of certain intensity in the diffraction pattern.

**Bragg's law:**  $n * \lambda = 2 * d * \sin \theta$ ;  $n$ : order of reflection (i.e. an integer);  $\lambda$ : wavelength of X-ray beam;  $d$ : distance of crystal lattice plane for a given set of  $h, k, l$ ;  $\theta$ : angle of reflection.



**Figure 1:** **A** The crystallographic unit cells are hit by the primary X-ray beam and the ordered array of molecules gives rise to a diffracted beam. **B/C** Due to arrangement in a lattice, molecules can be grouped into different planes that cut divisions into the cell axes. The diffraction of the primary X-ray beam into a certain direction (at angle  $\theta$ ) happens by these sets of lattice planes. **D** X-ray diffraction is only observed upon constructive interference of incident X-ray beams. **E** The set of (123) lattice planes; Miller indices:  $h=1$ ,  $k=2$ ,  $l=3$ .

## STRUCTURE SOLUTION (PHASING)

### Molecular replacement

If two proteins have a similar sequence (30% identity) or share a similar domain, the known structure (search model) can be used to model the unit cell packing of the unknown structure. The absolute prerequisite for successful molecular replacement is sufficient geometrical similarity between the “search model” and the structure to be determined.

### Multiple isomorphous replacement (MIR)

Heavy atom addition can change the phases and diffraction intensities. Heavy atom derivatives are prepared from native crystals by soaking in solutions of heavy atom salts. After collecting native and derivative X-ray data sets the heavy atom positions have to be solved. This information is used to estimate phases of the native protein. In iterative refinement, phases are improved.

### Multiple anomalous dispersion (MAD)

At certain resonance frequencies some heavier atoms (e.g. Fe, Se, Hg, Pt, I) absorb radiation and thus cause differences in diffraction intensities. Data sets measured at different wavelengths from *one crystal containing an anomalous scatterer* can be collected and the differences in intensities used to calculate the positions of the anomalous atoms. Subsequently, phases can be calculated for the native protein. This method relies on incorporation of an anomalous scatterer and usage of synchrotron radiation.

## STRUCTURE DETERMINATION (MODEL BUILDING AND REFINEMENT)

### Electron density

With help of the mathematical procedure of Fourier transform, the re-combination of diffracted X-rays is possible. It is in this stage, where the phases are needed.

The diffraction pattern acquired by the detector during the X-ray diffraction experiment is analysed with respect to position and intensity of diffraction spots. The spot position is correlated with Miller indices  $hkl$ , and thus the various crystal lattice planes. The intensity  $I$  and phase  $\alpha$  of all spots ( $hkl$ ) are needed to calculate an electron density map. The electron density  $\rho(x,y,z)$  is the Fourier transform of the structure factors  $F_{hkl}$ . Each structure factor  $F_{hkl}$  describes a specific reflection in the diffraction pattern and is the result of the X-ray beam diffracted by the arrangement of atoms in the crystal (i.e. it represents a wave!).

## Model building

- Computer programs are used to build an initial model using molecular graphics
- Problems in chain tracing are caused by discontinuous density
- Higher resolution data (2.0 to 1.5 Å) allow most side chains to be distinguished unambiguously

## ASSESSMENT OF PROTEIN STRUCTURES

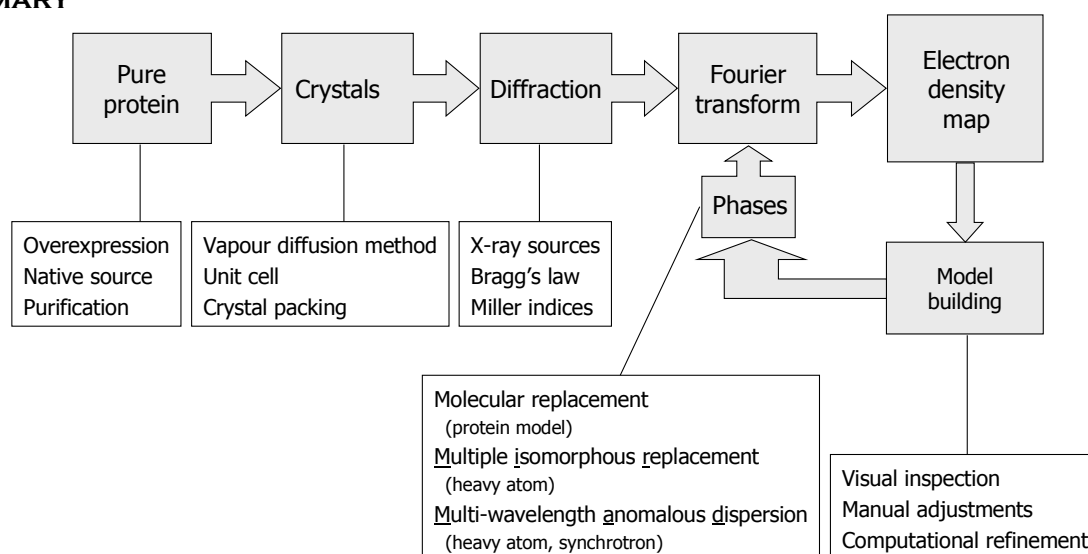
### Data quality

- Resolution should be better than 3 Å
- Completeness of data > 85%
- R-merge < 10% (how well multiply measured reflections agree)

### Model quality

- R-factor expresses the goodness of fit between the calculated and observed data. For a random (wrong) structure  $R = 0.59$ . A well determined structure has an R-factor between 0.15 and 0.20.  
$$R = \frac{\sum ||F_{hkl}^{obs}| - |F_{hkl}^{calc}||}{\sum |F_{hkl}^{obs}|}$$
- Deviation from standard bond lengths and angles should be about 0.02 Å and 2°
- $\Phi$  and  $\Psi$  angles should fall in the allowed regions of the Ramachandran plot
- A reasonable number of water molecules should be included in the structure (about one water per amino acid residue)
- Temperature factors (B-factors) measure the amount of motion in the atoms. They should normally be less than about 30 Å<sup>2</sup> (equivalent to an rms vibration of 0.6 Å)

## SUMMARY



**Figure 2:** Crystal structure solution by X-ray diffraction encompasses several steps with potential bottle necks. The electron density map in above scheme is the starting point for model building and subsequent iterative refinement.

## LITERATURE & REFERENCES

### Textbooks:

- Blow, D. (2002) *Outline of crystallography for biologists*, Oxford University Press.
- Creighton, T.E. (1993) *Proteins Structures and Molecular Properties*, 2<sup>nd</sup> ed., Freeman; Chapter 6.1: Three dimensional structures by X-ray diffraction.
- Glusker, J.P., Lewis, M., Rossi, M. (1994) *Crystal structure analysis for chemists and Biologists*, VCH Publishers.
- Massa, W. (2004) *Crystal structure determination*, 2<sup>nd</sup> ed., Springer
- Rhodes, G. (2000) *Crystallography made crystal clear – A guide for users of macromolecular models*, 2<sup>nd</sup> ed., Academic Press, Inc.

**Web tutorial:**

<http://www.ruppweb.org/Xray/101index.html>

<http://www.doe-mbi.ucla.edu/~sawaya/tutorials/tutorials.html>

<http://www.iucr.org/cww-top/edu.index.html>

*X-ray crystallography tutorial*

*Crystallographic tutorials*

*Cryst. Teaching Resources by IUCr*

**Databases and Tools:**

The Protein Data Bank (PDB)

Jmol

RasMol

Protein Explorer

<http://www.rcsb.org/>

<http://jmol.sourceforge.net/>

<http://www.umass.edu/microbio/rasmol/>

<http://molvis.sdsc.edu/protexpl/frntdoor.htm>

**IMPRESSUM**

Andreas Hofmann

Structural Chemistry Program, Eskitis Institute for Cell & Molecular Therapies, Griffith University

Email: [a.hofmann@griffith.edu.au](mailto:a.hofmann@griffith.edu.au), Web: <http://www.structuralchemistry.org/teaching/>