**Protein Crystallography**

Part I — Technical Aspects of and Introduction to X-ray diffraction

Tim Grüne  
Dept. of Structural Chemistry  
Head of Dept.: Prof. G. Sheldrick  
University of Göttingen  
November/December 2005  
http://shelx.uni-ac.gwdg.de  
tg@shelx.uni-ac.gwdg.de

---

**Overview**

- Crystals and Protein Crystals
- Growing Protein Crystals
- X-ray diffraction
- Inside Crystals
- Bragg’s Law and Resolution
- From X-rays to Electron Density

---

**Crystals and Protein Crystals**

**Examples of Crystals**

- **Ice Crystals**
- **Salt Crystals**
- **Artificial Monocrystal for a kJ Laser**
- **Quartz Crystal**

Protein Crystals are generally more sensitive than minerals. They are usually grown from solution and must be kept in a humid environment throughout the experiment.

Protein Crystals grown in solution.

Crystals from inorganic and especially ionic compounds (like $NaCl$) tend to be rather stable. They can be kept at room temperature, for a long time and in a dry environment.
Protein Crystals

Proteins are generally more sensitive than minerals. They are usually grown from solution and must be kept in a humid environment throughout the experiment.

If proteins are not handled with care, they quickly degenerate/aggregate.

In order to grow crystals from a protein solution, the protein must be very pure (≥ 90–95%) and very concentrated (≈ 15 – 25 mg/ml, ranges go from 5 to > 100 mg/ml). The probably most common method is called the vapor diffusion method.

The reservoir contains a buffer, a precipitant (salts, PEG's with M_r = 400–20,000 Da, organic solvents) and additives (small molecules/salts that help packing, e.g. MgCl_2).

The drop is a (1:1)-mixture of protein and reservoir solution (mother liquor).

Volatile components are exchanged until equilibrium is reached. Since at the beginning the precipitant is diluted in the drop, the drop is being concentrated during equilibration.

Vapour Diffusion

**PRO:**
- Fast setup — robots can do up to 100 drops in 30 min
- Small amount (1 nl / drop with robot)

**CONTRA:**
- Sudden mixing may cause protein to aggregate

Batch

**PRO:**
- Fast setup
- Water permeable oil allows concentration above initial protein concentration

**CONTRA:**
- Slow setup; requires large amounts of protein

Other Techniques

**Batch method:** Protein and reservoir are mixed directly and covered with a layer of oil ⇒ no or little equilibration.

**Diffusion through Dialysis Membrane:** Protein solution and reservoir separated by a high molecular weight cut-off membrane ⇒ also non-volatile components are exchanged.

**Diffusion through Agarose:** Reservoir and Protein separated through a layer of agarose ⇒ concentration gradient
What are X-rays?

- (Visible) light is composed of electromagnetic waves: every colour has a wavelength/energy
- X-rays are the same, but with higher energy, i.e. shorter wavelength

The spectrum of electromagnetic waves

<table>
<thead>
<tr>
<th>Wavelength (Å)</th>
<th>Energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>~1.54</td>
</tr>
<tr>
<td>1</td>
<td>~1.25</td>
</tr>
<tr>
<td>10</td>
<td>~4.1</td>
</tr>
<tr>
<td>100</td>
<td>~850</td>
</tr>
<tr>
<td>1000</td>
<td>~85</td>
</tr>
<tr>
<td>10000</td>
<td>~8.5</td>
</tr>
</tbody>
</table>

X-rays are generated by accelerating or decelerating electrons. This happens either in rotating anodes or at synchrotrons.

X-ray Diffraction — Principle of an experiment

Electrons are accelerated and hit a (rotating) Cu-plate. This induces a transition of the inner shell electrons which in turn produces characteristic $CuK\alpha$ radiation (1.54 Å)

Bunches of electrons are circulating at high velocity and bent by strong magnets. This generates X-rays at range wavelengths, depending on the degree of the bent. Some set-ups allow to tune the wavelength (important for MAD-and SAD-phasing, see later in the lecture).
**X-rays: Diffraction pattern**

Light is scattered from an object in all directions. A lens (e.g., eye, a camera, a microscope, or telescope) collects some of the scattered light and focuses it on a screen (and eventually the retina). This creates an image of the object.

**Inside Crystals**

The Secret of Crystals — Regularity

A crystal consists of a unit cell which is "piled up" like bricks (but in three directions) to compose the whole crystal.

The unit cell is characterised by the three side lengths, \(a, b, c\) and angles \(\alpha, \beta, \gamma\).

Irrespective of the lengths and angles a unit cell can always be "piled" without gaps.

For X-rays, lenses do not exist, therefore one cannot create and X-ray image. With a "normal" object, no information would be collected on the screen.

What is special about crystals so that one can retrieve information through X-rays from them?

How does the crystal lattice help?

In order to understand why crystals produce more than just a blur under X-ray diffraction, one introduces the concept of crystal planes. The corner points of all unit cells create the crystal lattice. Three non-linear lattice points (corners of different unit cells) define a plane.

One set of planes is described by the number of section it divides each side of the unit cell into (in the figure, side \(a\) is divided into two parts by the green set, and \(b\) into five parts). These numbers are called the Miller-Indices \((hkl)\) (three in three dimensions).

I.e., the green set of planes has the Miller Indices \((250)\), the purple one the Indices \((110)\).
Bragg’s Law and Resolution

(Bragg) Reflection

X-rays are reflected at the lattice planes. Since a set of planes consists of a huge number of planes, constructive interference occurs at certain angles \( \theta \). At all other angles the waves extinct each other. Therefore one observes the spotty pattern.

The angle \( \theta \), the plane separation \( d \) and the wavelength of the incident X-rays are connected by Bragg’s law or the Bragg condition:

\[ \lambda = 2d \sin \theta \]

Remark: Bragg’s law can easily be derived from the above picture by using the fact that the path difference between rays reflected from two adjacent planes must be an integer multiple of the wavelength. This leads to the actual law \( n \lambda = 2d \sin \theta \), but higher order reflections (\( n > 1 \)) are generally too weak to be detected.

Indexing of Reflections

Given a setup of crystal orientation, detector, and beam, each set of planes leads to one reflection. Therefore the reflection that belongs to one plane is labelled with the same Miller Indices \((hkl)\). Assigning these three numbers to each recorded reflection is called indexing.

The distance \( d \) which can be calculated via \( \theta \) from Bragg’s law is called the resolution of that reflection. The resolution of the reflection with maximal \( \theta \) determines the resolution of the data set.

The resolution of data collected by X-ray diffraction is a measure for how much detail can be seen. It is related with the plane distance \( d \) via Bragg’s law by

\[ d = \frac{\lambda}{2 \sin \theta_{\text{max}}} \]

\( \theta_{\text{max}} \) is the maximum angle to which data (i.e. a reasonable number of reflections) could be collected.

The final goal of data collection is to calculate an electron density map for the crystallised molecule. The resolution corresponds quite well to the minimum distance between two atoms that can still be resolved in the electron density map.

Examples for the Resolution of Electron Density Maps

The images show three times the same region of a protein map at different resolutions.

N.B.: Crystallographers usually speak of “high resolution” when the number is small (e.g. 1.2Å) and of “low resolution” when the number is large (e.g. 3Å),

G. Sheldrick
How to calculate an Electron Density from Diffraction Data

An X-ray experiment measures intensities of many reflections. The intensity of a reflection $(hkl)$ is proportional to the square modulus of the structure factors:

$$I(hkl) \propto |F(hkl)|^2$$

Structure factors are complex quantities introduced in order to facilitate the calculation of the electron density, i.e.

$$F(hkl) = |F(hkl)| \cdot e^{i\phi(hkl)}$$

$|F(hkl)|$ is called the amplitude and $\phi(hkl)$ the phase of the reflection.

The electron density within the unit cell of the crystal is related to the structure factors via a Fourier transformation:

$$\rho(x, y, z) = \frac{1}{V_{\text{unit cell}}} \sum_{h,k,l=-\infty}^{h,k,l=\infty} |F(hkl)| \cdot e^{-2\pi i (hx + ky + lz)}$$

The Fourier transform can also be inverted:

$$F(h, k, l) = \int_{V_{\text{unit cell}}} d^3x \rho(x, y, z) e^{2\pi i (hx + ky + lz)}$$

Limits of Crystallographic Data Collection

If we knew all structure factors, we could calculate a perfect electron density and build a perfect model. Obviously it is not possible to collect data from an infinite number of reflections at infinite accuracy. This leads to truncation errors of the Fourier transformation, which basically means noise in the electron density map. Even with the best equipment there are physical limits why we cannot collect data to arbitrary resolution:

1. Bragg’s Law limits the resolution to $\lambda/2$, because $|\sin \theta| \leq 1$
2. Crystal imperfections, mostly mosaicity. Mosaicity describes that the unit cells do not pack without gaps. This leads to a broadening of the signal which at high resolution vanishes in the background.
3. Detectors also contribute to a broadening of the signal.
4. Every crystal has a limited size. This means that there is only a limited number of sets of planes that can cause reflections. While this is usually not a limiting factor, it becomes an issue for very small crystals ($< 50 \mu m$ side lengths).