X-Ray Crystallography

BECAUSE

The underlying principle of function is structure.
X-ray crystallography is the study of crystal structures through X-ray diffraction techniques. When an X-ray beam bombards a crystalline lattice in a given orientation, the beam is scattered in a definite manner characterized by the atomic structure of the lattice. This phenomenon, known as X-ray diffraction, occurs when the wavelength of X-rays and the interatomic distances in the lattice have the same order of magnitude.
• In 1912, the German scientist Max von Laue predicted that crystals exhibit diffraction qualities.

• Concurrently, W. Friedrich and P. Knipping created the first photographic diffraction patterns.

• A year later Lawrence Bragg successfully analyzed the crystalline structures of potassium chloride and sodium chloride using X-ray crystallography, and developed a rudimentary treatment for X-ray/crystal interaction (Bragg's Law).

• Bragg's research provided a method to determine a number of simple crystal structures for the next 50 years. In the 1960s, the capabilities of X-ray crystallography were greatly improved by the incorporation of computer technology.
Modern X-ray crystallography provides the most powerful and accurate method for determining single-crystal structures.

Structures containing 100-200 atoms now can be analyzed on the order of 1-2 days, whereas before the 1960s a 20-atom structure required 1-2 years for analysis.

Through X-ray crystallography the chemical structure of thousands of organic, inorganic, organometallic, and biological compounds are determined every year.
X-ray crystallography is an experimental technique that exploits the fact that X-rays are diffracted by crystals. It is not an imaging technique. X-rays have the proper wavelength (in the Ångström range, $\sim 10^{-8}$ cm) to be scattered by the electron cloud of an atom of comparable size. Based on the diffraction pattern obtained from X-ray scattering off the periodic assembly of molecules or atoms in the crystal, the electron density can be reconstructed. Additional phase information must be extracted either from the diffraction data or from supplementing diffraction experiments to complete the reconstruction (the phase problem in crystallography). A model is then progressively built into the experimental electron density, refined against the data and the result is a quite accurate molecular structure.
The knowledge of accurate molecular structures is a prerequisite for rational drug design and for structure based functional studies to aid the development of effective therapeutic agents and drugs.

Crystallography can reliably provide the answer to many structure related questions, from global folds to atomic details of bonding. In contrast to NMR (which is a spectroscopic method), no size limitation exists for the molecule or complex to be studied.

The price for the high accuracy of crystallographic structures is that a good crystal must be found, and that limited information about the molecule's dynamic behaviour in solution is available from one single diffraction experiment. In the core regions of the molecules, X-ray and NMR structures agree very well, and enzymes maintain their activity even in crystals, which often requires the design of non-reactive substrates to study enzyme mechanisms.
X-ray Crystallography is probably the most important technique to determine the structure of any molecule (from small molecules to large assemblies of biological macromolecules) at atomic resolution. Atomic resolution means that the position of every atom in the molecule is determined within a relatively small range of uncertainty.

This module will outline the technique and describe the different steps as well as highlight its strengths and discuss its weaknesses. The goal of this module is not to help the experienced crystallographer in certain phases of a structure determination but rather to explain to the novice or interested beginner what the method is all about.

This module should also help non-crystallographers to understand better what kind of information can be retrieved from a coordinate file of a structure, and what kind of hypotheses can be based upon it.
The results of a structure determination have implications in a wide variety of fields related to the function(s) of the studied molecule. That means that many non-crystallographers will make use of structures and they especially should have at least some idea of how to use them.

First I would like to point out three databases. Crystallographers like to put their results into databases, which are available to the public. For small molecules, the database of choice is the Cambridge Structural Database, CSD and for biological macromolecules there is the Protein Data Bank, PDB and the Nucleic Acid Database, NDB. The CSD contains about 190,307 structures of 170,409 small molecule compounds as of October 1998, the PDB contains 8,949 structures as of December 30, 1998 and the NDB contains about 520 Nucleic Acid structures as of December 1998. All the structures deposited in these databases are available to the public.

Refer to the pages of the three databases for more information.
What do Crystallographers do?

A schematic outline is shown in the flow chart on the right:

- The first step is to grow crystals of the molecule of interest. How that works and why crystals are necessary and why crystals can yield some information about the native structure of a molecule will be discussed later.
- The next step is the characterisation of the crystals in terms of their properties, which is mainly symmetry.
- Then, X-rays are used on the crystal to obtain diffraction images. These contain the information about the structure. Unfortunately the structure cannot be calculated directly from the diffraction images, because the phases of the diffracted waves (what that is will be discussed later) are lost during data collection.
What do Crystallographers do?

In an extra step these **phases** have to be determined.

Now, a so-called **electron density map** can be calculated which is nothing but an image of the structure. This **map needs to be interpreted** in terms of where the atoms and the amino acids sit. Then we have the first crude structure of the molecule is obtained, which can then be **refined** and **checked** thoroughly for possible errors.

Once that is done, we have the structure is determined and the data is publishable.

Although the process may sound easy and straight forward, it may easily take up to a couple of years, depending how difficult the single steps prove to be.
Why Crystals?

• Crystals are necessary because it is neither possible to see nor handle single molecules.
• Therefore a lot of molecules are placed into a single crystal.
• A defining feature of a crystal is that it is a three-dimensional periodic array of molecules.
• This means that all of the molecules in the crystal occur in only one or in very few orientations. This leads to a tremendous amplification of the diffraction image. And not only that, the diffraction image is also much simplified.
Why X-rays?

• X-rays are electromagnetic waves much like visible light. Only their wavelength is much shorter.

• Since the aim is to acquire information about the position of atoms it is necessary to use radiation of a wavelength close to the distance of atoms in molecules.

• Only then will it be possible to distinguish one atom position from the neighbouring atom position.

• X-rays have a wavelengths around $10^{-10}$ m which is close to the distance between two bonded atoms in a molecule.
What is the Role of Electrons?

• X-rays interact with the electrons in the molecules not with the atomic nuclei.
• The result of that is, that the diffraction image contains information about where the electrons are in our structure.
• From the electron density, which is just a term to describe how many electrons sit at a given site, we can then infer where the atomic positions have to be, atomic positions are inferred from electronic positions.
To perform a X-ray diffraction experiment, an x-ray source is required.

In many cases a laboratory X-ray source such as a rotating anode generator producing a X-ray beam of a characteristic wavelength is used.

Intense, tunable X-ray radiation produced by a synchrotron provides additional advantages.

The primary X-ray beam is monochromated by either crystal monochromators, focusing mirrors, or complex multilayer optics.
After the beam passes through a helium flushed collimator it passes through the crystal mounted on a pin on a goniometer head.

The head is mounted to a goniostat which allows to position the crystal in different orientations in the beam.

The diffracted X-rays are recorded using imaging plates, multiwire detectors (now obsolete), CCD detectors (most common) and the new superfast pixel array detectors (PADs).
Figure 8-1 A contemporary laboratory X-ray diffractometer for macromolecular crystallography. A rotating anode X-ray source is closely coupled with integrated focusing optics delivering high photon flux at low operating power. In the center of the diffractometer is a full 4-circle ω-goniostat for orienting and rotating the crystal in multiple positions in the X-ray beam, thus enabling redundant data collection and in-house S-SAD phasing experiments. The CCD area detector is located to the right, and the diffractometer is also equipped with a cryocooler and a video microscope. The 2θ- and the ω-axis are collinear, with 2θ the detector offset angle. Image courtesy Matt Benning, Bruker AXS.

(C) Bernhard Rupp 2010.
Growing Protein Crystals

• We all are familiar with crystals from rock collections or small molecules, such as salt or sugar.
• They are usually associated with properties such as hardness, durability, and prettiness.
• Unfortunately, only the latter is true for protein crystals.
Proteins

- Proteins consist of long macromolecule chains made up from 20 different amino acids.
- The chains can be several hundred residues long and fold into a 3-dimensional structure.
- It is therefore quite understandable that protein molecules have irregular shapes and are not ideally suited to be stacked into a periodic lattice, i.e., a crystal.
- Protein crystals are thus very fragile, soft (think of a cube of jelly instead of a brick) and sensitive to all kind of environmental variations.
Proteins

• Protein crystals contain on average 50% solvent, mostly in large channels between the stacked molecules on the crystal.

• The interactions holding the molecules together are usually weak, hydrogen bonds, salt bridges, and hydrophobic interactions, compared to strong covalent or ionic interactions in mineral crystals.

• This explains the fragility of the crystals, but allows for the possibility of soaking metal solutions (important for phasing) or even large enzyme substrates or inhibitors, into the crystals.
The Experimental Set-up

• In order to obtain a crystal, the protein molecules must assemble into a periodic lattice.
• One starts with a solution of the protein with a fairly high concentration (2-50 mg/ml) and adds reagents that reduce the solubility close to spontaneous precipitation.
• By slow further concentration, and under conditions suitable for the formation of a few nucleation sites, small crystals may start to grow.
• Often very many conditions have to be tried to succeed.
• Crystal size should to be from a few hundred down to about 20 micron in each direction to be useful for diffraction experiments.
The Experimental Set-up

1. Mix cocktail and protein on glass slide.
2. Turn slide and seal well.
3. Observe for crystal formation.
4. Well with crystallization cocktail (precipants, additives, detergents, etc. – unlimited combinations possible).
5. Vapor diffuses into well, concentrations in drop increase.
6. Harvest and mount crystals.

(C) Bernhard Rupp 2010
The Experimental Set-up

• The most common setup to grow protein crystals is by the hanging drop technique:
• A few microliters of protein solution are mixed with an about equal amount of reservoir solution containing the precipitants.
• A drop of this mixture is put on a glass slide which covers the reservoir.
• As the protein/precipitant mixture in the drop is less concentrated than the reservoir solution (The protein solution is mixed with the reservoir solution about 1:1), water evaporates from the drop into the reservoir.
• As a result the concentration of both protein and precipitant in the drop slowly increases, and crystals may form. There is a variety of other techniques available such as sitting drops, dialysis buttons, and gel and microbatch techniques.
Cryocrystallography- Radiation Damage

- Just as any organic or living material, proteins are sensitive to X-ray radiation damage.
- The energy range of X-rays used for diffraction is in the 6 to 15 keV range, which is in fact severely ionizing radiation.
- The ionizing absorption events create radicals, which rapidly destroy any protein crystal, particularly at dose rates experienced at synchrotrons.
- An efficient way to suppress radiation damage by slowing down the kinetics of the radical reactions is cryogenic cooling.
- Rapidly quenching or flash-cooling crystals to liquid nitrogen temperatures, either in cold nitrogen gas streams or directly into liquid nitrogen, will strongly reduce radiation damage.
- To prevent the formation of crystalline ice during flash-cooling of the crystals, cryoprotectants, present in the mother liquor or added to the mother liquor, are necessary.
Cryomounting

- Cryoprotection is effectively accomplished during harvesting, when the crystals are scooped up from the drop in cryo-loops and briefly swept though a cryoprotectant before being dipped into liquid nitrogen.
- Common cryoprotectants are ethylene glycol (the anti-freeze in automobile radiators), glycerol, higher alcohols, polyethylene glycols (PEGs), or high concentration solutions of sucrose or salts.
- Once the protein crystals are flash-cooled and stored in pucks in a liquid nitrogen dewar (dry shipper), they can be safely sent to a synchrotron for data collection.
Benefits of Cryo-cooling

• Many factors contribute to improvements in data quality during cryo-protection:
  • Obvious benefits are reduced thermal vibrations, enhanced signal-to-noise ratio, reduced conformational disorder, and in many cases, a higher limiting resolution.
  • The most important effect is the suppression of radiation damage, permitting a complete data set to be collected from one single crystal.
  • This in turn eliminates errors from merging and scaling of data sub-sets from multiple crystals or non-isomorphism between MAD data sets.
  • In addition, crystal mounting is vastly simplified over conventional capillary techniques.
  • These improvements combined lead to enhanced contrast and sharper detail in electron density maps, facilitating model building and reducing the total time required for structure determination.
What does High Resolution Mean?

- The figure on the following slide shows what a certain resolution, given in Ångström (Å) means for the user of structural models derived from X-ray data.
- One always has to remember that the obtained model you see was built into an experimental electron density.
- The model may look as good at 3 Å as it does at 1.2 Å, but is it a correct and unique description of reality?
Figure 1-6 Data quality determines structural detail and accuracy. The qualitative relation between the extent of X-ray diffraction, the resulting amount of available diffraction data, and the quality and detail of the electron density reconstruction and protein structure model are evident from this figure: The crystals are labeled with the nominal resolution $d_{\text{min}}$ given in Å (Ångström) and determined by the highest diffraction angle (corresponding to the closest sampling distance in the crystal, thus termed $d_{\text{min}}$) at which X-ray reflections are observed. Above each crystal is a sketch of the corresponding diffraction pattern, which contains significantly more data at higher resolution, corresponding to a smaller distance between discernible objects of approximately $d_{\text{min}}$. As a consequence, both the reconstruction of the electron density (blue grid) and the resulting structure model (stick model) are much more detailed and accurate. The non-SI unit Å ($10^{-8}$ cm or 0.1 nm = $10^{-10}$ m) is frequently used in the crystallographic literature, simply because it is of the same order of magnitude as atomic radii (~0.77 Å for carbon) or bond lengths (~1.54 Å for the C–C single bond).
Resolution

• The pictures of the electron density at different data set resolution of the same region of a molecule leave no question that a model of a phenylalanine-containing ligand (the 6-ring structure) can be correctly placed into the 1.2 Å data. This still can be done with confidence in the 2 Å case, but at 3 Å we already observe a deviation of the centroid of the ring from the correct model. The bottom panel visualizes the relation between diffraction limit, amount of data and nominal resolution. The more and better data, the more accurate and detailed the final structure model will be.

• Most protein crystals diffract between 1.8 and 3 Å, a few to very high resolution (the term high resolution is used loosely in macromolecular crystallography, we apply it to data of 1.8-1.2 Å, below 'atomic' or 'ultra-high' resolution are commonly used). The most efficient way to increase resolution (short of trying to grow better crystals) is to cryo-cool the crystals to near liquid nitrogen temperature.