Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/20619 holds various files of this Leiden University dissertation.

Author: Nederlof, Igor
Title: Protein three dimensional nano-crystallography by imaging and diffraction
Issue Date: 2013-03-14
Chapter 1

Introduction¹

1.1 Outline

The biological activity of most proteins is determined by their 3D structure. Determining the 3D structure of proteins can therefore be very useful to show how they function: protein structures can be used to interpret molecular processes of life. This can, not only give us a better understanding of life itself, but also can facilitate discovering new drugs or improving existing ones. For instance, a substantial number of molecular diseases are caused by protein structural alterations, which are genetically encoded. Drugs operate by binding to proteins, inducing changes to their functional structure and thereby affecting their biological activity. Hence the design and improvement of drugs is greatly facilitated by knowledge of the 3D structures of their macromolecular targets. In the light of these considerations, it is clear that elucidation of the 3D structure of proteins is of prime importance for understanding the underlying mechanisms of molecular diseases.

The most widely used techniques for the structure determination of proteins are X-ray crystallography, NMR and single particle cryo-EM. X-

¹ Part of the introduction is published as: Nederlof, I, van Genderen, E., Hoedemaeker, F., Abrahams, J.P., Georgieva, D., Protein Crystal Growth, Modern Aspects of Bulk Crystal and Thin Film Preparation: InTech.
ray crystallography is the most powerful in terms of resolution but requires well diffracting protein crystals. NMR yields structures of proteins in solution but is limited by the size of proteins that can be analysed. Larger protein structures and structures of protein complexes can be solved using cryo-EM but the resolution is not as high compared to the other methods.

Using crystals for structure determination has major advantages. As all the molecules are in the same orientation, their signal adds up and is easier to distinguish from the noise. Hence crystallographic methods have so far yielded the most detailed information of the structures of the molecules of life. I will first discuss how we can generate protein crystals, including pitfalls and bottlenecks of current procedures. Then, after a brief discussion of how crystals are currently used for structure determination, I will address new methods discussed in this thesis that aim at allowing structure determination even if crystals are sub-microscopic.

1.2 Protein 3D Micro- and Nano-Crystallogenesis

1.2.1 Relevance
It was initially believed that any soluble protein that could be purified would be relatively easy to crystallize. However, the results have indicated that solubility and purity of proteins, although being important factors, do not secure a yield of useful crystals. The crystallization behaviour of proteins turns out to be very complex.
In an effort to identify the naturally occurring protein folds, large structural genomics consortia were set up. The somewhat disappointing outcome of these efforts is that only about 3% of all proteins that were targeted by these consortia yielded a crystal structure, despite massive investments in high-throughput, automated protein production, purification and crystallization. It is clear that in order to improve the current situation, better strategies for protein crystallization are required, combined with techniques that allow the use of smaller nano-crystalline material.

1.2.2 Background and Bottlenecks

Crystallization of proteins involves the three classical steps of nucleation, growth and cessation of growth, even though the protein crystals contain on average 50% of disordered solvent (1). However, crystal growth of biological molecules differs substantially from small molecule crystallogenesis. The reason is the much larger number of parameters involved in protein crystallization, as well as the specific physico-chemical properties of the biological compounds. The main difference with small molecule crystal growth is the conformational flexibility and chemical versatility of macromolecules and their greater sensitivity to external factors. An overview of different parameters affecting the crystallization of bio-macromolecules is presented in table 1 (2).
<table>
<thead>
<tr>
<th>Intrinsic physicochemical properties</th>
<th>Biochemical and biophysical parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Supersaturation</td>
<td>• Sensitivity of conformation to physical parameters</td>
</tr>
<tr>
<td>• Temperature, pH</td>
<td>• Binding of ligands</td>
</tr>
<tr>
<td>• Ionic strength and purity of chemicals</td>
<td>• Specific additives</td>
</tr>
<tr>
<td>• Pressure, electric and magnetic fields</td>
<td>• Aging of samples</td>
</tr>
<tr>
<td>• Vibration and sound</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biological parameters</th>
<th>Purity of macromolecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Rarity of biological macromolecules</td>
<td>• Macromolecular contaminants</td>
</tr>
<tr>
<td>• Bacterial contaminants</td>
<td>• Sequence (micro) heterogeneity</td>
</tr>
<tr>
<td>• Biological sources of organisms and cells</td>
<td>• Conformational (micro) heterogeneity</td>
</tr>
<tr>
<td></td>
<td>• Batch effects</td>
</tr>
</tbody>
</table>

Table 1.1 Parameters affecting crystal growth.

Another important prerequisite for successful crystallization is the quality of the macromolecular samples. Bio-macromolecules are extracted from living cells or synthesized in vitro and they are frequently difficult to prepare at a high degree of purity and homogeneity. Besides traces of impurities, the different treatments proteins are subjected to may decrease their stability and activity through different kinds of alterations. As a general rule, purity and homogeneity are of prime importance. Accordingly, purification, stabilization, storage and handling of macromolecules are essential steps prior to crystallization.
Purity

The concept of purity has a special meaning in biological crystallogenesis. Molecular samples need to be not only chemically pure, but they must also be conformationally uniform (3). This concept is based on the fact that the best crystals are grown from solutions containing well-defined entities with identical physico-chemical properties. For X-ray crystallographic studies, the aim is to grow 'single crystals' diffracting to high resolution with a low mosaicity and prolonged stability in the X-ray beam. However, contaminants may compete for sites on the growing crystals and generate lattice errors leading to internal disorder, dislocations, poor diffraction or early cessation of growth (4). Because of the high molecular weight of molecules in a single crystal (up to millions of Daltons), and hence low molarity of their solutions, even relatively small amounts of contaminant may induce formation of non-specific aggregates, alter macromolecular solubility, or interfere with nucleation and crystal growth (5,6).
Figure 1.1 Solubility curve of a protein, where the phase state of the protein is plotted against the concentration of both protein and precipitant. At point [1], the protein may precipitate so fast that an amorphous precipitate or at best shower of micro-crystals is formed. At [2] the conditions are just right for the protein to form a stable crystal nucleus, which will start to grow – passing [3] – into a stable protein at equilibrium with the mother liquor [4]. At [5], the concentration of protein and precipitant are too low for crystal nucleation or growth, and the solution will remain clear. Note that the true solubility curve of any protein is highly multidimensional, with every parameter affecting protein solubility (cf. Table 1) representing a different independent axis.

Solubility, Supersaturation and Phase Transition

Biological macromolecules follow the same thermodynamic rules as inorganic or organic small molecules concerning supersaturation, nucleation and crystal growth. However, protein macromolecules are organized in tertiary and quaternary structures. The intra-molecular interactions responsible for their tertiary structure, the intermolecular interactions involved in the crystal contacts, and the interactions necessary to solubilise them in a solvent are similar.
To crystallize a biological macromolecule, its solution must have reached supersaturation, which is the driving force for crystal growth. The under- and supersaturated states are defined by the solubility of the macromolecules. When the concentrations of the crystallization agent and the macromolecules correspond to the solubility condition, the saturated macromolecule solution is in equilibrium with the crystallized macromolecules. Below the solubility curve (Figure 1.) the solution is under-saturated and the system is thermodynamically stable. In this case, phase transition (crystallization) will not occur. Above the solubility curve, the concentration of the biological macromolecules is higher than the concentration at equilibrium. A supersaturated macromolecular solution contains an excess of macromolecules that will fall out of solution as a solid phase, reducing the concentration of the dissolved species until the macromolecular concentration reaches the solubility value. The higher the supersaturation, the faster this solid phase appears. However, at very high supersaturation levels, precipitation, rather than crystallization occurs, as insoluble macromolecules rapidly separate from the solution in an amorphous state.
Figure 1.2 Glucose Isomerase crystallization condition yielding phase separation (far left) amorphous precipitation (near left) micro-crystals (near right) and macro-crystals (far right) bar on the top left represents 200 micrometer.

1.2.3 Growing and Handling Protein Micro- and Nano-Crystals
Crystallization of biological macromolecules is still one of the major bottlenecks in X-ray crystallography. There are many proteins of which the structures cannot be solved using current techniques due to their inability to grow well diffracting crystals. Protein crystals that are smaller than 1 micrometer are beyond the scope of standard X-ray techniques. Membrane proteins and large protein complexes often have stacking faults decreasing the size and diffraction potential of their crystals. Membrane proteins are located in the lipid layer that separates the cell from its environment and are essential in many signalling pathways, making them important drug targets. Structures of large (dynamic) protein-nucleic acid complexes can provide more information on how proteins and DNA interact with each other.

Many projects are abandoned when proteins yield crystals that cannot be optimized to diffraction grade, and thus structural information on these important drug targets is therefore severely lacking. Current trends in X-ray crystallography focus on decreasing the size limit of protein crystals, proving it is worthwhile to focus on this big supply of
poorly diffracting protein crystals. Micro-focused X-ray beams and improved quantum area detectors, like the Pilatus, have decreased the size limits of protein crystals (7,8). Also serial femtosecond crystallography using X-ray free-electron lasers will help in expanding the methods towards smaller crystals (9,10).

1.2.3.1 Nucleation

Nucleation is the rate-limiting step in protein crystallization (11,12). When a supersaturated protein solution is obtained, as described in 1.2.2, the growth of crystals is dependent on precipitant concentration, pH and temperature. However for crystal growth to be induced there needs to be a starting point, a crystal nucleus. These nuclei are formed in random events when a number of protein molecules come together in the same region in the crystallization drop and by chance coalesce into an ordered or semi-ordered cluster. If these nuclei are stable, more protein molecules can stick to them in the same orientation as the proteins of the nucleus and form a protein crystal. Increased nucleation usually increases the chances of nano- and micro-crystal growth at the expense of the growth of large single crystals, since multiple nucleation sites are competing for the same protein.

Heterogeneous nucleation, which is induced by introducing a properly chosen nucleant into a crystallization condition, may promote nucleation. The first report of a nucleant inducing crystallisation of macromolecules was the epitaxial growth of protein crystals on minerals (13). Other nucleants followed, like zeolites, silicates, charged surfaces, and other porous materials, that all have been tested for
multiple proteins (14,15). Previous results showed that horsehair and
dried seaweed showed increased hits when added to sparse-matrix
crystallization trials. The increase in crystallization was 35% when
horsehair was added to 10 test proteins (16). The underlying
mechanism is explained with epitaxial nucleation in the case of
minerals, electrostatic interactions if the nucleants contain charged
surfaces, nucleation through specific favourable protein-protein
interactions or physical entrapment in the caves of porous materials.
Introducing heterogeneous nucleants can increase nucleation sites
decreasing the size of the crystals obtained.

1.2.3.2 Automation

In recent years, setting up protein crystallisation trials and analysis of
the results have become largely automated. Lab automation includes
the use of dispensing robots, imaging robots, in situ crystal analysis as
well as automated diffraction analysis (17,18).

In general, as a beginning, one or more standard screens are used with
pre-defined crystallization conditions. These screens are designed on
the basis of statistical analysis of results obtained at structural genomics
initiatives. When initial hits are found with screens like these,
optimisation screens need to be performed to produce diffracting
crystals. In this stage, use of a Lab Information Management System
(LIMS), where experimental design can be coupled to experiment
preparation and analysis, greatly enhances the potential throughput in a
lab and thereby the success rate (19).
1.2.3.3 Detecting Crystals in Crystallization Drops

The increase in throughput obtained by using dispensing robots makes it impossible to routinely scan the results manually/interactively using a standard microscope. The dynamic nature of these experiments can cause the crystallographer to miss events, even crystals, especially micro- or nano-crystals that are often difficult to see by eye. Imaging robots vary from semi-automated microscopes with a camera and a moving stage, to fully automated incubators that are capable of following all lab experiments from start to finish without human intervention.

In recent years, three new techniques have been developed for detecting crystals and assess their quality while they are still in the crystallization plates. These techniques are in situ diffraction analysis, UV detection and second harmonic microscopy. Several years ago a device for X-ray diffraction analysis of crystals in the plate where they were grown was developed (20,21). Crystals are centred using visual alignments and X-ray diffraction patterns are obtained. When suitably diffracting crystals are found they still need to be harvested and frozen for complete diffraction analysis. UV detection is becoming more popular for detecting crystals. This technique makes use of the fluorescence in UV by proteins, mostly caused by tryptophan (22). Having visible light and UV cameras integrated in a single imaging device greatly enhances its usefulness for distinguishing protein from salt crystals. Second harmonic generation (SHG), more often referred to as “frequency doubling” (23) is an even more recent development. When an intense laser pulse travels through a crystal with a non centro-symmetric space group, light emerges with exactly half of the
wavelength of the incident beam. The explanation is that two photons of the incident beam merge, creating a single photon with twice the energy. If the incident beam is near-infrared, the emerging beam will be in the visible range. The technique is very sensitive, making it suitable for detecting nano- and micro-crystals that cannot be seen using conventional UV and light microscopy.

1.2.3.4 Harvesting and Mounting

X-ray diffraction data is collected at liquid nitrogen temperature (-196 °C) to reduce radiation damage to the crystal and to reduce noise caused by thermal motion. Crystals are transferred to a cryo-protectant solution before flash-freezing to prevent them from cracking. The cryo-protectant also minimizes ice-crystals from forming which interfere with data collection (24).

Protein crystals are traditionally harvested using a cryo-loop. A cryo-loop is a small pin with a nylon loop at the end to “scoop” a single protein crystal. This step is done by hand and can be quite tedious, especially for very small crystals. When a single crystal is captured using a cryo-loop, it can be transferred into an X-ray beam for data collection. Micro-crystals too small to be transferred with a cryo-loop can be transferred to a micro-mesh by transferring the whole or part of the droplet they grew in. They can then be examined using UV or SHG methods and data can be collected using a micro-focus X-ray beam. If data is collected on an X-ray Free Electron Laser source using femtosecond exposures, nano-crystals in suspension are “sprayed” through the beam using a liquid jet (25).
For examining nano-crystals and 2D crystals by cryo-electron microscopy, a drop of mother liquor containing the nano-crystals is pipetted onto an EM-grid. However since electron microscopes operate using vacuum, the drop needs to be vitrified. This is done by blotting away excess liquid and plunge-freezing the sample in liquid ethane cooled by liquid nitrogen. The thin aqueous layer is cooled so abruptly, there is no time for ice crystals to form and the water turns vitreous. The sample is now encased in a 10 – 100 nm layer of amorphous ice thin enough to let electrons through.

### 1.3 X-ray Micro- and Nano-Diffraction

#### 1.3.1 Background

X-ray diffraction is the most widely used technique for solving crystal structures. If a well-ordered protein crystal is obtained that is large enough, it can diffract incoming X-rays resulting in a diffraction pattern. From a set of these diffraction patterns an electron density map can be calculated into which the protein structure can be modelled.

**Bragg Model of Diffraction**

When a protein crystal diffracts an incoming X-ray beam, it is diffracted into well-defined angles. The further away from the centre the X-rays are scattered (or: the higher the scattering angle), the better the resolution of the data set. The maximum angle to which the crystal diffracts the incoming X-rays determines how fine the details are that can be distinguished in the protein structure. High resolution is thus
desirable. Knowing the wavelength $\lambda$ of the X-rays and the diffraction angle $2q$ of a reflection, its resolution $d$ can be calculated using the following equation:

$$d = \frac{1}{2} \left( \frac{\lambda}{\sin \theta} \right)$$

This is just a reformulation of the famous Bragg’s Law, formulated almost exactly 100 years ago (26):

$$\lambda = 2d \sin \theta$$

In this Bragg equation the variable $d$ is the distance between atomic layers in a crystal. In a diffraction experiment we measure the intensities of diffraction spots and the position of the reflections. From the position of a reflection we can determine its index $(h,k,l)$, which is associated with the direction of the parallel atomic planes from which the diffraction spot is reflected. By integrating the signals of each of the reflections, we can assign the appropriate intensities to them. This intensity is proportional to the square of the structure (factor) amplitude, $F_{hkl}$, and is a scalar value, a simple number. One dataset usually contains a few ten thousand of reflections collected for each crystal.
The structure factor $F_{(h,k,l)}$ for a reflection $h,k,l$ itself is a complex number and is derived as follows:

$$F_{(h,k,l)} = \sum_{j=1}^{\text{atoms}} f_{(j)} \exp \left[ 2\pi i \left( hx_{(j)} + ky_{(j)} + lz_{(j)} \right) \right]$$

This summation extends over all atoms $j$ of the repeating unit of the crystal ($x,y,$ and $z$ are the fractional coordinates of the atoms). The parameter $f_{(j)}$ is the scattering factor of atom $j$ and depends on the kind of atom and the diffraction angle of the corresponding reflection $(h,k,l)$.

For X-ray crystallography, at $h,k,l = 0$, $f$ equals the atom's number of electrons times the unit scattering of a single electron ($2.82 \times 10^{-15}$ m).

The $i$ in the equation is the unit complex number: the square root of $-1$. This equation shows that if the structure is known the structure factors can easily be calculated.

However since diffraction is measured in Fourier space and not in real space the Fourier summation is used:

$$\rho(x,y,z) = \frac{1}{V} \sum_{h,k,l} |F(h,k,l)| \exp \left[ -2\pi i (hx + ky +lz) + \alpha(h,k,l) \right]$$

In which $\rho$ is the electron density, $|F(h,k,l)|$ is the structure factor amplitude of reflection $(h,k,l)$, which includes a temperature factor correction; $\alpha(h,k,l)$ is the phase angle; as in the previous equation, $x,y,z$ are the fractional coordinates of the atoms in the unit cell. From the diffraction pattern $I(h,k,l)$ can be obtained after applying several correction for experiment-specific parameters. Because $I(h,k,l) = |F(h,k,l)|^2$ the amplitudes $|F(h,k,l)|$ can be calculated. Unfortunately, no
information is available on the phase angels $\alpha(h,k,l)$. Without this information the structure can't be solved. Four techniques are used for solving this so-called the ‘phase problem’. They are molecular replacement (MR), multiple wavelength anomalous diffraction method (MAD), multiple / single isomorphous replacement method (MIR, SIR) and finally direct methods, which require resolutions that are mostly not reached by protein crystals.

Molecular Replacement
This technique is used if a structure of a highly homologous protein is available, the phases calculated from the model can be combined with the experimental amplitudes, allowing the calculation of an initial electron density map. Through model building and refinement, this map is improved in an iterative way, in several rounds until the process converges and leads to a final model or structure.

Single or Multiple Isomorphous Replacement (SIR or MIR)
If heavy, electron dense atoms (lead, mercury, etc.) are soaked into the protein crystal without changing the crystal lattice, the positions of the heavy atoms can be calculated directly from difference Patterson maps. From the differences between the native and derivative data sets, the phases of the native data set can be inferred. These heavy atoms also cause anomalous dispersion of X-rays, causing differences between the Friedel pairs of reflections. This additional information is also used to estimate starting phases.
Multiple Anomalous Dispersion (MAD)
For MAD-phasing, a variant of the protein is produced that contains a seleno-methionine (methionine in which the sulphur atom is replaced by a selenium atom) instead of methionine. If X-Rays are used at a specific wavelength, these selenium atoms will produce a large enough anomalous signal to obtain phases. A “native” dataset collected at a different wavelength is also needed. The advantage of MAD over MIR (or SIR) is that the data sets can be recorded from the same crystal, thereby ensuring that the datasets are isomorphous (apart from radiation damage).

Model Building and Refinement
If an electron density map is obtained the protein model can be automatically built using ARP/wARP, buccaneer or other programs (27,28). If an initial model is obtained this can be refined with the original data using refinement software like Refmac or Phenix (29,30). This is done to better fit the model to the measured diffraction data. The model can be manually checked using coot (31), a graphical model building software tool. Coot can also be used for rebuilding and refining selected parts of the protein model.

1.3.2 Micro-Diffraction
Obtaining diffraction grade crystals is a real bottleneck in macromolecular crystallography, thus it has become a trend to improve current techniques to make structure solution of suboptimal crystals possible. In the last ten years micro-focus beams have become routinely used for data collection of micro-crystals, which are too small for
structure solution using traditional synchotron X-ray beams. Micro-focus beams were developed at the ESRF (32) and use apertures and mirrors to focus beams, thus decreasing their size and increasing their flux. Improvements not only increase maximum resolution, but also isotropy of diffraction data. Since diffraction data can be obtained from only a small part of the crystal, also twinned or mosaic crystals can now yield useful data. Radiation damage is reduced since less of the crystal is submerged to X-rays. Many synchrotrons now have micro-focused beams incorporated at their macromolecular beamlines (33).

1.3.3 Nano-Diffraction
A novel technique called serial femtosecond crystallography used X-ray Free Electron Lasers (XFELs) to collect diffraction data of protein crystals ranging from 200 nm to 2 μm. Protein crystals in suspension are sprayed through the XFEL beam and millions of randomly oriented diffraction patterns are collected (9). Exposures of 10 to 200 fs are used at power densities of $10^{16}$ W cm$^{-2}$, which vaporizes the protein crystals. Although this method is still in the development stage, it shows high potential for the future.

1.3.4 Detecting X-rays
In an X-ray experiment a crystal diffracts an incoming beam, which is scattered into reflections or spots. In the early days these spot patterns of diffracted X-rays were collected on photographic film. These have since been replaced by image plate scanners and CCD detectors. The latest developments have been pixel detector like the Pilatus (34),
which significantly increase the signal to noise ratio, thus decreasing exposure times and beam damage.

**1.4. Electron Diffraction and Imaging of Nano-Crystals**

**1.4.1 Relevance**

Since the development of the electron microscope by Max Knoll and Ernst Ruska in the early 1930's, the applications of transmission electron microscopy (TEM) have been extensive and impressive. Originally, material scientists used it for examining metals, alloys, ceramics, glasses, polymers, semiconductors and mixtures of these materials. With the development of cryo-electron microscopy in the 1970's (35-37) the applications have also shifted to biological samples. The advantages of using TEM for the structure determination of proteins and protein complexes over X-ray diffraction is that when employing single particle analysis, no crystals are needed. In electron crystallography thinner crystals can be used in the form of 2D crystals or 3D nano-crystals. Electrons are less damaging to proteins than X-rays by several orders of magnitude per diffracted quantum (Table 2. (38)), minimizing beam damage by decreasing energy deposited in the sample.
<table>
<thead>
<tr>
<th></th>
<th>Electrons (200 keV)</th>
<th>X-rays (1.5 Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio (inelastic/elastic)</td>
<td>~3</td>
<td>~10</td>
</tr>
<tr>
<td>scattering events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy deposited per</td>
<td>~20 eV</td>
<td>8 keV</td>
</tr>
<tr>
<td>inelastic event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy deposited relative</td>
<td>1</td>
<td>~10^3</td>
</tr>
<tr>
<td>to electrons per elastic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>event</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.2 *Electrons are less damaging to proteins than X-rays* (38).

1.4.2 Principles of Electron Diffraction and Imaging

Since the wavelength of electrons is significantly smaller than that of X-rays (or visible light), a much higher resolution can be obtained in theory. The higher the accelerating voltage of the electrons, the smaller their wavelength, which increases the maximum resolution that can be obtained. Unlike X-rays, electrons are charged, and therefore they can be focused by electromagnetic lenses. This means an electron microscope can be build in similar fashion to a light microscope, see figure 3.

1.4.2.1 Basic Principles of a Transmission Electron Microscope

We distinguish three basic systems in an electron microscope: the illumination system, the image formation system and the image detection system (for the latter two, see figure 3).
Figure 1.3 Formation of diffraction pattern (left) and image (right) in a TEM. (From Williams and Carter 2nd edition 2009(58)).

The illumination system consists of an electron source and a condenser system. The electron source is either an emission source (either tungsten or LaB₆ crystal) or field emission gun. The condenser consists of one or more condenser lenses and apertures and delivers a beam of electrons to the specimen. Depending on the optical settings, this beam
can be parallel, di- or convergent, can be precessing over the sample or scan the sample.

After the electrons have passed through the specimen they travel through the image formation system, which contains an objective lens an objective aperture, a selected area aperture, intermediate lenses and a projector lens.

The image is then formed on a fluorescent screen or a detector. In the past photographic film was used for detection but it has been replace by image plates and CCD cameras. Recently direct electron detectors like the Falcon or the K2 were developed (39,40). For measuring diffraction data, a quantum area detector like the Medipix is ideal, which significantly increases the signal to noise ratio, thereby decreasing exposure times and hence beam damage.

Since electrons are charged, they interact much stronger with matter compared to light or X-rays. This property means that electrons can be used to study much smaller objects. There are different ways electrons can interact with matter: they can scatter elastically, inelastically and dynamically.

Elastic scattering occurs when the electron conserves its kinetic energy but changes direction due to interaction with the electrons and nuclei of the specimen. If an electron loses kinetic energy, inelastic scattering has occurred. Inelastic scattering is the main scattering process of electrons in high voltage electron microscopy (see table 2) but for high-resolution imaging, the inelastic scattering events are mostly ignored. The reason is that elastic scattering, but not inelastic scattering, can be analysed
using Bragg's law, in a similar manner as was described for X-rays in 3.1.

Elastically scattering electrons can interact multiple times with atoms in the specimen, giving rise to dynamic scattering. These multiple scattered electrons can only be analysed using complex calculations and there is no straightforward relation between the structure of the specimen and the location where they up in a diffraction spot or image.

Spherical aberration (Cs) is the limiting factor for high resolution in modern day electron microscopy. However Electron microscopes can be equipped with a Cs-corrector system to correct for spherical aberration. A Cs-corrector consists of two hexapole and four additional lenses, which make it possible for high-resolution images to be obtained.

1.4.2.2 High Resolution Imaging of Biological Samples

Single particle analysis is the most widely used imaging technique for determining the structure of large protein complexes, viruses and other biological samples that are too large to be analysed by X-ray diffraction. Samples are negatively stained or vitrified and transferred into the electron microscope. Thousands of electron-images are taken at low dose to minimize beam damage. To increase the signal, multiple images are averaged and analysed using various image analysis techniques.

The first step in image processing is to group images of different orientation. This is done by classifying the images using multivariate statistical analysis. The images in each class are aligned and through averaging increase the signal to noise ratio. To reduce the influence of
high or low spatial frequency information a band pass filter is usually applied to images. These filters make use of the Fourier transform algorithms to weight or improve the results of alignment. To enhance the contrast in single particle analysis, EM images are taking 'under-focus'. This means images are blurred by a point spread function. This can be corrected for using the contrast transfer function, which is a function in reciprocal space and can allow higher resolution reconstructions. Since the images contain particles at random projections, their orientation angles need to be inferred in order to make a three dimensional reconstruction. To assign these 'Euler angles' common projection lines can be used. In the case of a sample containing symmetry like viruses or crystals this can enhance reconstructions by assigning angles in reciprocal space.

1.4.2.3 Electron Crystallography of Protein 2D Crystals

Electron crystallography can be used to determine the structure of two-dimensional (membrane) protein crystals. Resolutions better than 3 Ångstrom can be achieved (41,42). The method has similarities to single particle analysis but because of the internal order of the sample, the resolution can be enhanced though crystal unbending, internal symmetry and measuring diffraction patterns. First electron images are collected. These contain phase and amplitude information. Then electron diffraction patterns are obtained to further improve the resolution of the data. Since the crystals consist of one flat layer that is horizontally aligned they are often tilted during data collection to obtain three-dimensional datasets. 2D crystals of membrane proteins generally
need lipids and detergents to grow and are generally vitrified, negatively stained or embedded in sugar.

1.4.2.4 Electron Crystallography of 3D Crystals (Materials vs. Proteins)

In material sciences, electron diffraction is a well-developed tool for structure determination of inorganic crystals. Selected area electron diffraction (SAED) can be used to determine three-dimensional unit cells from different zone axes. To obtain more information about the symmetry of the crystal convergent beam electron diffraction (CBED) (43) can be used. The development of precession electron diffraction (PED) (44,45), automated diffraction tomography (ADT) (46,47) and rotation electron diffraction (RED) (48) have contributed to the toolbox of electron crystallographers, for inorganic or small organic molecules. However 3D nano-crystals of large macromolecules have so far resisted structure determination. The main reasons are the beam sensitivity, the large unit cell and the thickness of the crystals. The latter factor contributes to increased multiple scattering and non-linear effects in electron diffraction and imaging. However multi-slice least square methods (49) have tackled such problems in electron crystallography of small molecules. To increase resolution and refine structural data, electron diffraction data can be combined with electron microscopy images, which contain phase information (50-53).

4.3 Trends

Current trends in Cryo-EM are going towards sorting images to find more than one heterogeneous group of molecules (54). This means
dynamic processes of biological systems can be studied. Another
development is the use of phases from electron micrograms to solve
structures with X-ray diffraction where phases are missing (55,56).

In this thesis I describe my work on the analysis of 3D protein nano-
crystals using cryo-EM imaging and electron diffraction. I discuss the
use of heterogeneous nucleants to increase the success of crystallization
in automated trials. The structural studies of the enzyme Small Laccase
and the role of Tyr108 in its enzyme mechanism are also discussed.

References
Crystal Growth 76, 554-561.
Growth 152, 209-220.
Growth 168, 63-73.
Biol.
(8) Eikenberry, E. F., Bronnimann, C., Hulsen, G., Toyokawa, H., Horisberger, R., Schmitt, B.,
Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 501, 260-266.
Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L.,
Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmes, M., Caleman, C., Boutet, S., Bogan,
M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudak, B., Erk, B., Schmidt, C., Homke, A.,
Reich, C., Pietschner, D., Struder, L., Hauser, G., Gorke, H., Ulrich, J., Herrmann, S., Schaller, G.,
Schopper, F., Soltan, H., Kuhnle, K.-U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M.,
Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Haju, J., Timmeanu, N., Seibert, M. M.,
Andreasson, J., Rocker, A., Jonsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schroeter, C.-D,


Textbooks:


