A 5 Macromolecular Crystallography

T. E. Schrader

Jülich Centre for Neutron Science

Forschungszentrum Jülich GmbH

Contents

1   Introduction

Ideally, one knows each atom position in a cell or even in the whole body of an organism at any given time. With this amount of data one could reconstruct each metabolic pathway. One could understand for example the synthesis of fatty acids or the immune response to an HI-Virus. Unfortunately, it is not easy to retrieve each atom position in such a complex environment as a bacterial cell. One problem is the quantum mechanical uncertainty relation. This is a principal problem which cannot be overcome by any structure resolving technique. Here, only quantum mechanical calculations in combination with structural techniques can help. But mostly quantum mechanics is not the limit. A whole organism or even a single living cell is already too complex in order to obtain structural information on all molecules involved in metabolic processes. Here, one usually constructs model situations "in vitro" meaning: one just dissolves the molecules under investigation in a buffer solution mimicing pH and ionic strength of the cell environment. Small molecules (of less than 1000 g/mol in weight) usually have only few degrees of strutural freedom, so one is more interested in the "macromolcules" which often are polymers like starch (made out of suger units) or DNA/RNA consisting of Nucleotides often matched in base pairs and proteins which are polymers of 20 different aminoacids. Among those macromolecules used by living cells in their metabolism there are two classes which have a strong structure-function relationship and are therefore subjected to many structural studies: DNA or RNA and proteins. But it is already difficult to retrieve high (=atomic) resolution structural information from single molecules in solution. The NMR technique is one example but is often limited in the size of the macromolecule (see lecture A7). When it comes to scattering techniques, only very bright x-ray sources in the future (e. g. XFEL) might enable one to gain enough information out of many single protein molecules in solution, averaging many exposures and orientations[1].

Crystallography provides here the most elegant solution. In a crystal, the macromolecules are all oriented in a very regular way. This is why their scattering intensities add coherently and the contrast is increased. Both DNA/RNA and proteins form crystals. But by far the biggest number of structures recorded on biological macromolecules is on protein crystals using x-rays as a probe.

Apart from water, proteins are the most abundant molecules in living cells. Proteins fulfil numerous functions in the cell, for example enzymatic catalysis which enhances the speed with which molecules like fatty acids are synthesized. Proteins take care for transport and storage of important molecules like oxygen or they are involved in immunology, just to name a few of their functions [2]. In order to perform these functions proteins adopt a unique three dimensional structure with a carefully controlled mixture of flexibility and stiffness. For an understanding of their function knowledge of this three dimensional structure is a prerequisite. Ideally one would like to produce a movie where one can follow the functioning protein in action in slow motion with atomic resolution. In practice, there are techniques available which have a sufficient time resolution (in the fs-regime) but do only provide very limited structural information like time resolved infrared spectroscopy. However, there are some promising attempts to do time resolved x-ray scattering to have both strutural resolution at the atomic level and time resolution [3]. But this technique of time resolved x-ray scattering is often limited to certain systems which undergo a photo-initiated process. On the other hand there are methods which provide full atomic resolution but with essentially no time resolution.

With those methods one often stops the functioning process of a protein under investigation in an intermediate state by trapping methods using inhibitor molecules which stop the catalytic process of the protein leaving it trapped it in a certain intermediate state. This lecture focusses on this type of methods since most of the structural information deposited in the protein data bank has been obtained with this "slow" x-ray crystallography technique. Due to its similarity to x-ray crystallography we will also discuss neutron protein crystallography which uses neutrons instead of x-ray photons as a probe.

The protein data bank (www.rcsb.org) is a well known source of structural information on proteins. Data from many different experimental techniques are entered in a standardized format, a .pdb file which essentially contains not only three dimensional coordinates of all observed atoms in a protein but also information on their mean square displacements. The number of stored entries exceeds 120534 as of 31st of October 2017.
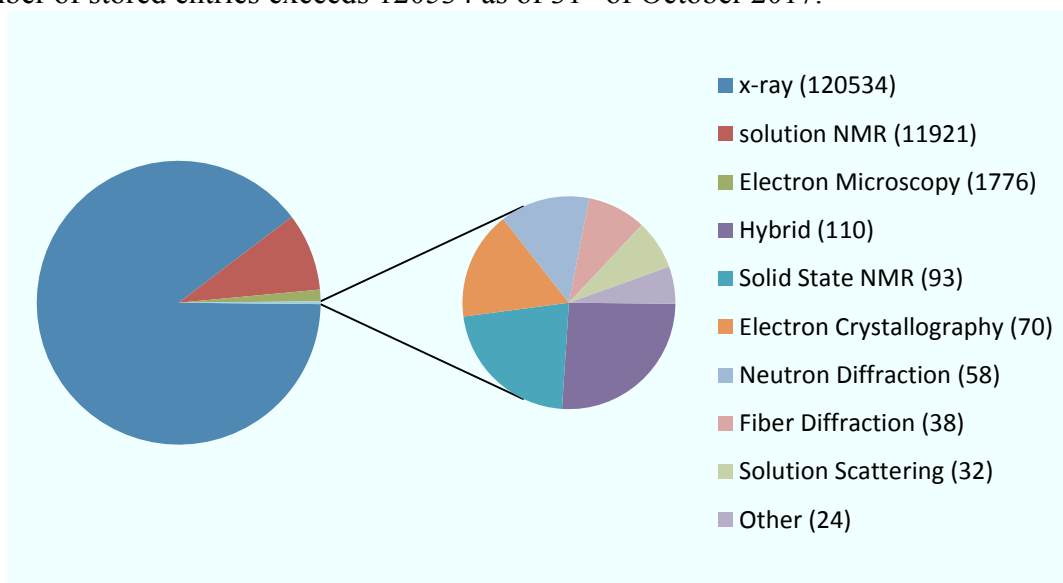

Fig. 1:   *Entries of structual data with respect to the method used as found in the protein data base (www.rcsb.org) up to the 31st of October 2017. The total number of entries for each method is given in brackets. The circle on the left totals to 134656 entries of which 425 entries are taken out to form the second circle on the right. This lecture discussed the method of x-ray (dark blue section, left circle) and neutron (light blue section, right cercle) crystallography with 120534 and 58 entries respectively.*

Among them x-ray crystallography has contributed more than 89.5 %. The next in line method with 11921 entries in the data bank is solution NMR spectroscopy. Electron microscopy as a method was used in more than 1776 entries and is the fastest expanding technique at the moment. A whole lecture (A2) in this spring school is dedicated to this technique. The remaining methods count mostly less than 100 entries each including neutron protein crystallography. Since the latter technique is represented by an instrument in the Jülich Centre for Neutron Science (JCNS) and because of its similarity to x-ray crystallography it will be given some space in this lecture. Finally two example case studies are discussed where both techniques give complimentary information.

2   The Physics of Macromolecular Crystallography

The following chapter will show how most of this structural information has been obtained. For both techniques x-ray and neutron protein crystallography a single crystal of the protein of interest is required since the scattering of one protein molecule is very weak. So, in general a crystal has to be grown, especially large ones in case of neutron crystallography since the neutron luminosity of modern sources is much smaller than the x-ray flux reached by synchrotron sources. To grow sufficiently large crystals is a big challenge in the case of many proteins, especially membrane proteins. Here, one has to adjust a large parameter set of protein concentration, pH condition, salt concentration, percipitant concentration and type just to name a few.
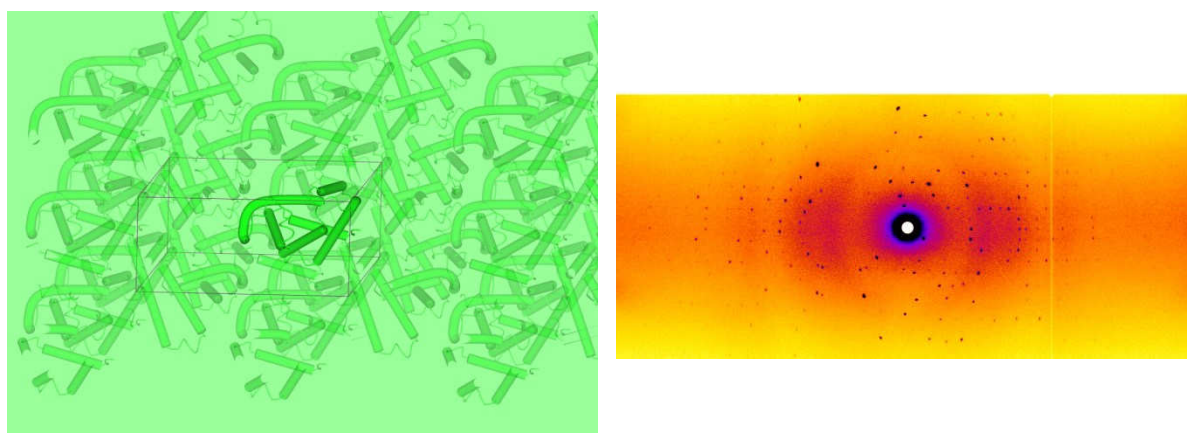

Fig. 2:   *Real space arrangement of myoglobin molecules in a crystal of space group P2₁ (on the left) versus diffraction pattern (right) of a myoglobin crystal.*

The crystal then serves as a noiseless amplier of the diffraction signal. But the arrangement of proteins in a crystals brings in another advantage, since the orientational averaging can be avoided, which is always present in the solution phase. Fig. 2: lef hand side shows the regular arrangement of myoglobin molecules in a crystal lattice. The unit cell of the monoclinic lattice (space group $P2_1$) is indicated by black lines. It bears two myoglobin molecules in one unit cell. The picture on the right shows a diffraction pattern recorded with the instrument BioDiff on a myoglobin crystal. The crystal is rotated by ca. 0.5° while recording one diffraction pattern. In order to map the reciprocal space completely one has to put the crystal in many different orientations into the beam and record diffraction patterns as mentioned above. Fortunately, crystal symmetry helps that some orientations are equivalent to each other and need not be recorded.

2.1   An X-ray macromolecular crystallography beamline

Synchrotron beamlines provide extremely high photon flux for x-ray crystallography. Due to the high demand from the structural biology community, often more than one macromolecular crystallography beamline is operated at a synchrotron. Those beamlines are optimized for special wavelengths and focal diameters. Here as an example the beamline BL14.2 is used to elucidate a typical x-ray protein crystallography experiment.

Fig. 3: lef hand side shows a schematic view of the beam path of beamline BL14.2. The beam paths to the other beamlines BL14.1 and BL14.3 have been omitted for clarity. A supraconducting magnet-structure interacts with the electron orbit of the storage ring and creates the so called synchrotron radiation. This radiation is used as a white light X-ray source for the three beamlines. A double crystal monochromator is used to select a very narrow energy band (2 eV at 9 keV) from the broad spectrum of the X-ray source. The mechanics of the double crystal monochromator keeps the out-going beam path constant when changing the wavelength. Focussing mirrors and collimators in the beam path ensure an efficient photon transport from the source to the sample and a small beam size at the sample position of 150 μm x 100 μm (hxv, FWHM). The Rayonix MX-225 detector has a pixel size of 37 μm. Without on chip binning one frame amounts to 6144 x 6144 pixels. The exposure time per frame is typically between 3 to 10 seconds.
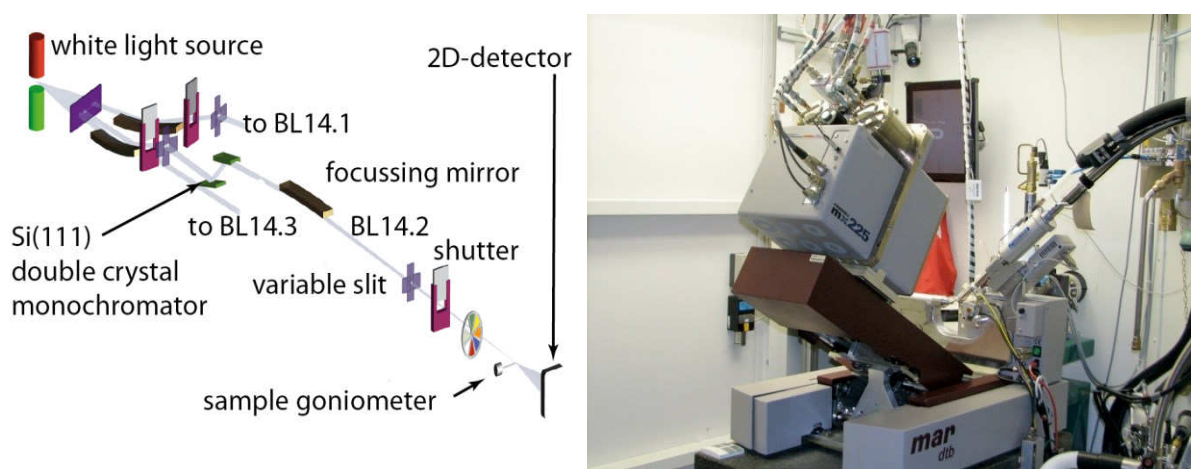

Fig. 3: *Schematic view of the beamline 14.2 of BESSY beginning with the magnet structure which is used as a white light x-ray photon source. On the right a picture taken from the experimental hutch of beamline 14.1 is shown. The beam enters from the right and the sample goniometer is mounted horizontally from the left. A cryostream sample environment to stabilize the sample temperature is discernible pointing towards the sample from the top right corner. Pictures kindly provided by Dr. Uwe Müller.*

Typically, the sample crystals are kept at liquid nitrogen temperatures to avoid radiation damages. To record a full data set takes about 10-30 minutes. The largest diagonal of a typical protein crystal ranges between 10 to 500 μm

2.2   A neutron macromolecular crystallography instrument

Since x-rays are scattered from the electrons in the crystal and neutrons from the nuclei, hydrogen atoms are hardly seen in x-ray crystallography experiments. Only at very high resolutions of 1 Å or less there is a chance to observe hydrogen atom positions. This resolution is often not within reach because of the crystal quality. Here neutron protein crystallography must be employed to retrieve the hydrogen atom positions. Moreover, neutron scattering can distinguish between different isotopes, especially between hydrogen and

deuterium. Whereas from x-ray crystallography the electron density in the unit cell of the crystal can be calculated, neutron protein crystallography yields the nuclear scattering length density, which is a signed quantity. In fact, the coherent scattering length of hydrogen is negative and the one from deuterium is positive (cf. Fig. 4: left).

Nucleus	atomic number	scattering length [10^{-12} cm]
^{1}H	1	-0.378
^{2}H	1	0.667
^{12}C	6	0.665
^{15}N	7	0.921
^{16}O	8	0.581

Fig. 4: *The table on the left lists scattering lengths of selected atoms of biological relevance. On the right there is a comparison of x-ray scattering cross section with scattering lengths from neutron scattering. The circles are scaled to match at the carbon atom.*

A major drawback of the method neutron protein crystallography is the required crystal size. Due to the much smaller neutron flux as compared to x-ray flux the crystals required for a neutron crystallography study must be much larger as compared to x-ray crystallography. Here, often crystal diagonals of 1 mm and more have to be reached.

As an example of a neutron diffraction instrument optimised for protein crystallography the instrument BioDiff at the FRM II shall be introduced. It is a collaboration between the Forschungszentrum Jülich (FZJ) and the Forschungs-Neutronenquelle Heinz Maier-Leibnitz (FRM II).
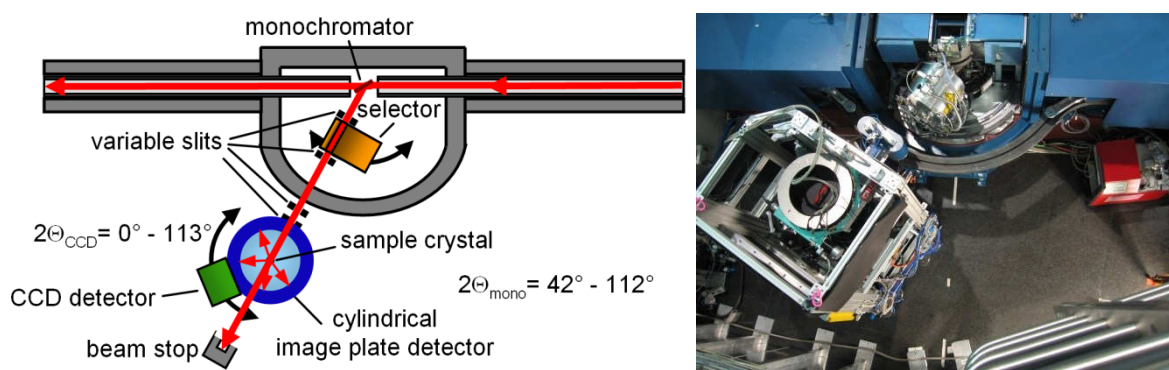
Fig. 5: *Schematic view of the BioDiff instrument (left) and a picture taken from a similar view point with the biological shielding removed (right).*

Figure 10 shows a schematic view of the instrument from the top and a corresponding picture when the biological shielding has been removed. The neutron beam from the cold source of

the FRM II reactor enters from the right. By Bragg reflection from a pyrolytic graphite crystal (002-reflex) neutrons are taken out from the white neutron spectrum of the neutron guide NL1 and pass a first boron carbide adjustable slit and a velocity selector. The velocity selector acts as a λ/2 filter. Together with the pyrolytic graphite crystal it forms a monochromator with a Δλ/λ of ca. 2.5 %. Behind the velocity selector the beam passes a second variable slit and the main instrument shutter, named γ-shutter. Additionally, a boron carbide neutron shutter is placed directly after the monochromator crystal for a faster shutter operation. Before entering the detector drum of the image plate detector through a Zirconium window a collimator made out of two manually exchangeable boron carbide apertures with fixed diameters between 3 mm and 5 mm shape the beam to fit to the sample size. At present the sample is usually contained in a glass tube (either a thin walled capillary or a NMR-tube for larger crystals). It is fixed to a standard goniometer which is mounted upside-down from the sample stage on top of the instrument. After passing the sample the main neutron beam exits the detector drum through a second Zirconium window and hits finally the beam stop which consists of a cavity of 4 mm thick boron carbide plates surrounded by a 13 cm thick wall of lead shielding bricks. The cylindrical image plate detector is covering roughly half of the total 4π solid angle. It is 200 mm in diameter and 450 mm in height. It can be read out with three different resolutions of 125 µm, 250 µm and 500 µm. As an alternative, one can lower the image plate detector and swing in a neutron sensitive scintillator which is imaged onto a CCD-chip. This CCD-camera set up serves as a second detector. In particular it is used for a fast alignment of the sample crystal with respect to the neutron beam.

With the image plate detector the diffraction pattern shown in Fig. 2: right hand side has been recorded. In fact, a complete crystallographic data set on a myoglobin crystal has been recorded allowing for the calculation of a nuclear scattering length density map. The exposure time was 17 minutes per frame and the crystal was rotated by 0.5° during exposure. 331 frames were recorded before the crystal was manually rotated by ca. 90° in the capillary and another set of 243 frames were recorded. Altogether ca. 9 days of beam time were necessary to record the complete data set. The achieved resolution with sufficient completeness was 1.7 Å. The required time to record this data set was much longer as the 30 minutes from x-ray diffraction.

2.3   Some general aspects of diffraction by a protein crystal

Having recorded a complete data set on a crystal some data treatment is necessary in order to calculate meaningful atom positions. Here only a brief introduction can be given more details can be found in text books [4,5].
Assuming a number of n atoms per unit cell the structure factor of a single unit cell can be written as:

$$F(\mathbf{S}) = \sum_{j=1}^{n} f_j \exp(2\pi i \mathbf{r}_j \mathbf{S}) \qquad\qquad (1)$$

Here $\mathbf{r}_j$ denote the atom position of atom j and $\mathbf{S}$ is the scattering vector perpendicular to the plane which reflects the incident beam. In the prvious lecture A4 the scattering vector q was defined. It relates to $\mathbf{S}$ with the following relation: $q = 2\pi\mathbf{S}$. In crystallography it is just more convenient to use $\mathbf{S}$ instead of q. f_j can be seen here either as the scattering length of atom j in the neutron scattering case or the atomic scattering factor in case of x-ray diffraction. One can generalize this approach by switching form the summation to an integration to yield:

$$F(\mathbf{S}) = \int_{unitcell} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r}\mathbf{S}) d^3\mathbf{r} \tag{2}$$

where $\rho(\mathbf{r})$ is the electron density distribution or the scattering length density respectively. Since a crystal consists of AxBxC unit cells, the structure factor of the crystal can be composed as

$$F_{cryst.}(\mathbf{S}) = F(\mathbf{S}) \sum_{u=0}^{A}\exp(2\pi i u\mathbf{a}\mathbf{S}) \sum_{v=0}^{B}\exp(2\pi i v\mathbf{b}\mathbf{S}) \sum_{w=0}^{C}\exp(2\pi i w\mathbf{c}\mathbf{S}) \tag{3}$$

The vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ denote basis vectors of the unit cell. For an increasing number of unit cells the sums can be represented by delta functions leading to the Laue conditions for the structure factor being non-zero:

$$\mathbf{a}\mathbf{S} = h, \mathbf{b}\mathbf{S} = k, \mathbf{c}\mathbf{S} = l \tag{4}$$

This means that one only gets constructive interference, when the scattering vector is perpendicular to planes in the crystal which can be denoted by the index vector $\mathbf{h} = hkl$. For this reason the diffraction pattern of a single crystals shows distinct peaks, the so called Bragg peaks. The Bragg law can be easily derived from equation 4. Figure 11 shows the Ewald sphere construction. It is a tool to construct the direction of the diffracted beam. The Ewald sphere has its origin at the position of the crystal. Its radius is the reciprocal wavelength used in the scattering experiment. The origin of the reciprocal space lattice is placed at the intersection of the sphere with the incident beam direction. Whenever the orientation of the reciprocal space is such that another point of the reciprocal space lies on the Ewald sphere a diffracted beam results in the direction of line running from the centre of the Ewald sphere through that point.
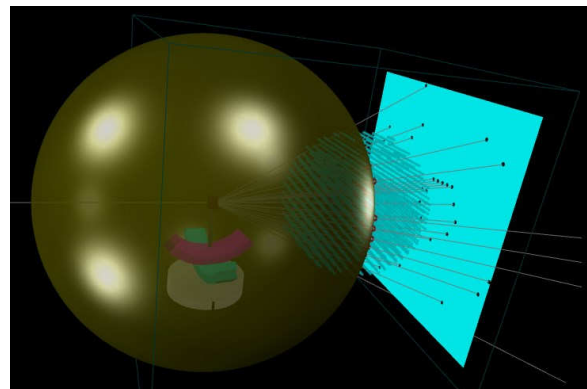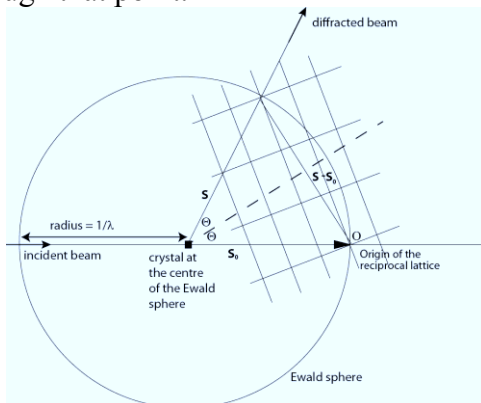

Fig. 6: *Ewald sphere: On the left the schematic shows how to construct the Ewald sphere. On the right an Ewald sphere (golden colour) construction is shown in three dimensions. The blue square represents a flat two dimensional detector.*

When the crystal is rotated the reciprocal space rotates with it resulting in other lattice points to cause diffracted beams. In practice the incident beam is not strictly monochromatic but has a wavelength distribution which causes the Ewald sphere to be elongated to form a spherical shell of a certain thickness. This increases the number of diffracted beams observed. The beam divergence adds also to its thickness.

So, the positions of the diffracted beams on the detector only depend on the reciprocal lattice. The structure of the protein inside the unit cell is encoded in the amplitude and phase of the structure factor.

To obtain the electron density or the nuclear scattering length density one has to perform the inverse Fourier transformation:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{h}} F(\mathbf{h}) \exp(-2\pi i \mathbf{rh}) \tag{5}$$

Here V is the volume of the unit cell. Unfortunately only the modulus squared of the structure factor is measured as intensity on the detector. The phase information is lost which is known as the phase problem of crystallography.

There are several solutions to the phase problem which are only applicable for the x-ray diffraction case:

- isomorphous replacement: Several crystals of the same crystal structure have to be available for this method. First a crystallographic data set is recorded on an untreated crystal. Then crystals are soaked in at least two different heavy atom salt solutions. In the best case, the different heavy atom ions occupy different regular positions in the unit cell. From these (at least) two crystallographic data sets recorded on the heavy atom treated crystals phase information can be retrieved which is then used to determine the phases of the data set of the untreated crystal.

- anomalous dispersion: Often it is possible to replace one distinct methionine amino acid with an artificial selenomethionine one. The selenium atom has a suitable absorption edge on which anomalous scattering can be performed by tuning the wavelength of the beamline to the anomalous regime. Crystallographic data sets are then recorded at different wavelengths from with the phase information can be calculated. In some cases this approach can also be adopted for sulfur atoms present in naturally occuring cysteine residues.

- molecular replacement: From the primary structure one can search the protein data base (pdb) for proteins with a similar amino acid sequence. If one finds enough fragments which seem to be sufficiently homologous to the unknown structure one can use those fragments for the calculation of initial phases. In further refinement steps these phases can be improved further. Since the number of unique structures entered in the protein data base is growing this method is increasingly favoured over other methods.

The phase problem of the neutron data sets is solved by using the x-ray structure and the molecular replacement technique.

2.4   Model building and refinement

With the data treatment one has now arrived at a contour map $\rho(\mathbf{r})$ be it either a nuclear scattering length density or an electron density. Now the information on the primary structure of the protein is used and either manually or employing software first the backbone is coarsely fitted into the contour map. Then from this model new amplitudes and phases of the structure factor are calculated using eq. 1. The modulus squared of the structure factor is again compared with the data. One could now think of a least square based fitting procedure to find the optimum arrangement of the protein atoms in the unit cell. In practice however maximum likelihood and simulated annealing molecular dynamics simulations are used because those are superior to the least square approach in terms of overcoming local minima. In these molecular dynamics simulations a lot of stereochemical information is used as restraints for example known bond lengths of CC single bonds or bond angles. The agreement between model and observed contour map is often measured by calculating a so called R-factor:

$$R = \frac{\sum_{\mathbf{h}} \left\| F_{obs}(\mathbf{h}) \right| - \left| F_{calc}(\mathbf{h}) \right\|}{\sum_{\mathbf{h}} \left| F_{obs}(\mathbf{h}) \right|} \tag{6}$$

The index "obs" denotes the observed structure factors and the index "calc" the calculated structure factors from the model. The value of the R-factor lies in the limits between 0 and 1. A good agreement between model and measured data leads to an R-factor of about 0.2. R-factors of 0.5 and above are indicative for a random agreement between model and data.
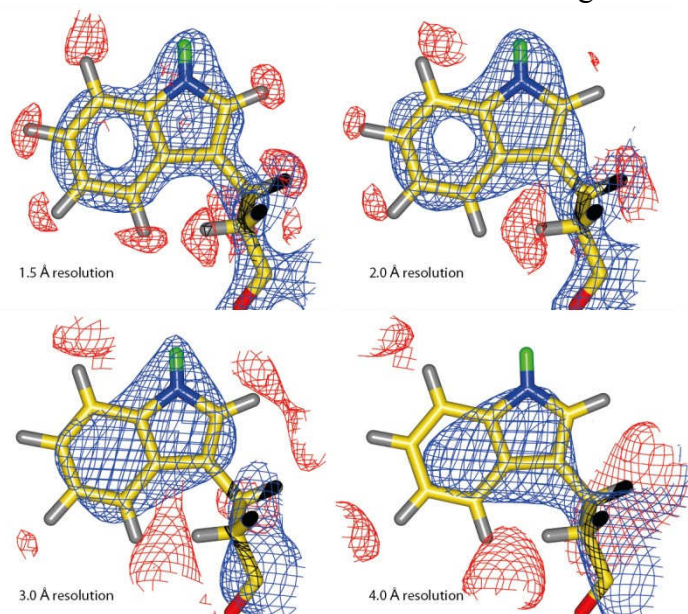

Fig. 7: *The side chain of the amino acid tryptophan no. 7 of myoglobin measured with neutron diffraction at different resolutions. The contour level of the shown nuclear map is +1.5σ (blue) and -1.75σ (red). Exchanged (liable) hydrogen atoms (green) and C- (yellow), N- (blue) and O-atoms (red) appear on a positive contour level. Only not liable hydrogen atoms are seen on the negative contour level.*

But even a good R-factor does not guarantee that the model fits to the data. In fact, it is possible in special cases to obtain a reasonably low R-factor when using the amino acid chain in the wrong direction as a model [6]. Here, Brünger et al. [7] have suggested to divide the measured Bragg reflections into two subset one working set denoted by "A" and one test set denoted by "T". With the working set the fitting procedure is performed, whereas the test set only serves to control the model quality by calculating the R_{free} factor.

$$R_{free} = \frac{\sum\limits_{\mathbf{h} \in T} \left\| F_{obs}(\mathbf{h}) \right| - \left| F_{calc}(\mathbf{h}) \right\|}{\sum\limits_{\mathbf{h} \in T} \left| F_{obs}(\mathbf{h}) \right|} \qquad (7)$$

The test set usually consists of 5 to 10 % of all structure factors, uniformly distributed over the resolution range.

The R-factor and R_{free} factor should not differ too much from each other. In general, it is good practice to always look at the resulting model and its fit to the calculated map after each refinement step. Ramachandran plots can also be used to judge whether the amino acid backbone adopts a reasonable fold. With decreasing resolution (cf. Fig. 12) of the data it becomes more and more difficult to find the right orientations of side chains or even errors in the registry of the protein backbone can occur, whereby for example one amino acid is left out.

3   A first case study: Water network around myoglobin

This example is chosen since it nicely shows the interplay between x-ray and neutron crystallography. Myoglobin has been used quite frequently as an example throughout this lecture. Its function is to take over the oxygen molecules from the blood heme proteins in the red blood cells and to transport and store it within the muscle cells. Therefore its binding affinity to oxygene must be stronger than that of the hemoglobin. In order to perform the transport tasks myoglobin has to be highly soluble and movable within the context of a muscle cell. Let alone therefore it is interesting to study the water network surrounding of myoglobin. Since it is a joint neutron and x-ray diffraction study the crystal under investigation has been soaked in D_2O in order to exchange all liable hydrogen atoms with deuterium atoms.

Fig. 8: shows a contact region between two myoglobin molecules. In the centre a sulfate ion is observed. Here, in the nuclear map the central sulfur atom is hardly seen because of its small scattering length. The oxygen atoms of the sulfate ion and of the water molecules are readily observed by both techniques. The deuterium atoms of the water molecules are discernible only in the nuclear map. When the water molecule is fixed by hydrogen bonds, all three atoms can be observed. These water molecules exhibit a triangular shape in the nuclear map. In case it is free to rotate around the OH-axis the nuclear scattering length of the rotating deuterium atom is distributed over a large volume and is therefore averaged out. Those water molecules

are denoted as "short ellipsoidal" by Chatake et al. [8]. The long ellipsoidal water molecules are fixed at both deuterium atoms but only the oxygen can rotate freely around the DD-axis. In the fourth case only the oxygen atom is observed. In this case the orientation of the water molecule is not fixed, only the oxygen atom is held in place.
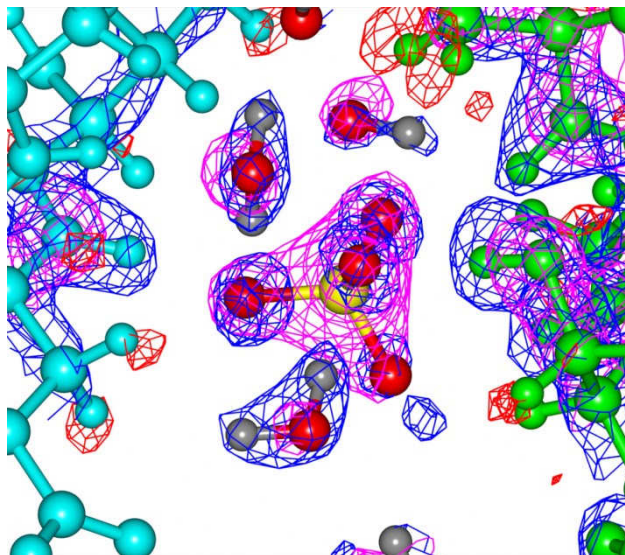

Fig. 8: *Water network in the contact region between two myoglobin molecules in the crystal. In grey colour on the left amino acids 51 to 52 of one myoglobin molecule are shown. On the right amino acids 58 to 60 (from top to bottom) are depicted in green. In the centre of picture a sulfate ion SO_4^- is seen with the sulfur atom shown in yellow, the oxygen atoms shown in red. The deuterium atoms of the water molecules are coloured grey. The x-ray map is shown in magenta at a contour level of +2.7s. The nuclear map is shown at a contour level of -1.75s in red and at +2.3s in blue. Data taken from ref. [9]. (The picture is similar to the one shown in [8] or [10] since it is based on the same data.)*

This classification of water molecules helps to judge the flexibility of the water network around proteins. It can also be used to classify observed water molecules and their hydrogen bonding pattern in molecular crystals in general.

4 A second case study: Fighting Meticillin resistant bacteria

This example also shows the nice interplay between x-ray and neutron crystallography. It is chosen because of the relevance it has to our everyday life. The protein involved here is the β-lactamase. It is produced by bacteria and partly secreted to their sourrounding environment in order to split a certain bond in the four-membered ring (=β-lactam motif) of a series of very common antibiotics. This is one of the mechanisms which render such bacteria resistant against this type of antibiotics. The antibiotic drug called meticillin also bears such a four membered ring. Meticillin resistant *Staphylococcus aureus* is assumed to possess this

mechanism of resistance against β-lactam antibiotics. This is why it would be interesting to investigate the catalytic cylce of this protein in order to block it. This would result in the bacteria being affected by the antibiotica again.

The experiment was designed in a clever way. The catalytic cycle of the β-lactamase consists of an acylation step where a covalent adduct between protein and substrate (here the antibiotics molecule) is formed (Fig. 9: top part). Here, the CO double bond is split into a single bond and a tetrahedral intermediate between substrate and protein is produced. The second part is a deacylation step where the four-membered ring (the so-called β-lactam motif) is split and the product is released. The product cannot act as an antibiotic any more due to its different chemical struture.
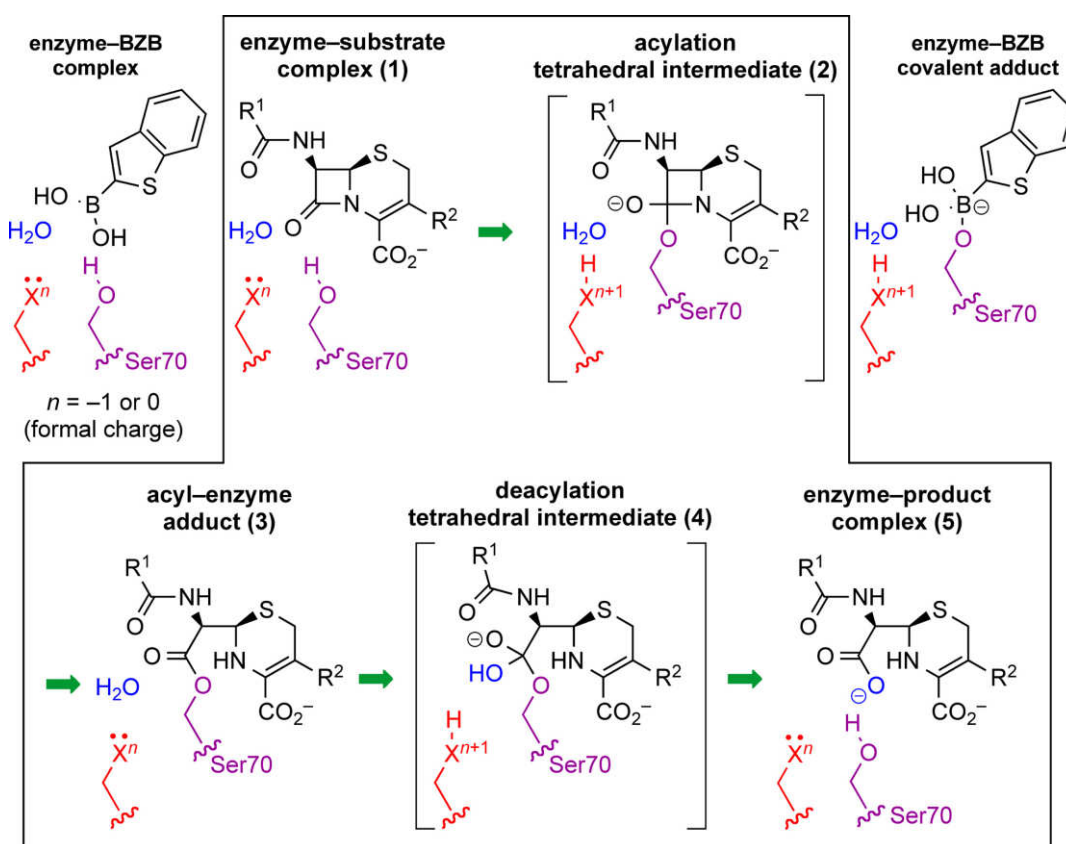

Fig. 9: *Catalytic cycle of a β -lactamase enzyme. (The Figure is taken from ref. [11]. This research was originally published in Journal of Biological Chemistry, Stephen J. Tomanicek et al. "Neutron and X-ray Crystal Structures of a Perdeuterated Enzyme Inhibitor Complex Reveal the Catalytic Proton Network of the Toho-1 β-Lactamase for the Acylation Reaction." © the American Society for Biochemistry and Molecular Biology.)*

Due to its relevance to clinics the enzyme family of β-lactamases have been subjected to countless studies mostly by x-ray crystallography but the main question yet not addressed was the nature of the active base which takes the excess proton in the acylation step. It was not

clear which part, which side chain of the protein takes over this role. But this was the key information for improving drugs to block this enzyme. Since a base is best detected when it takes a proton in the acylation step, the problem consisted of stopping the catalytic cycle in the acylation step and hunting for a proton which was not there in the ground state of the protein when no substrate was bound.
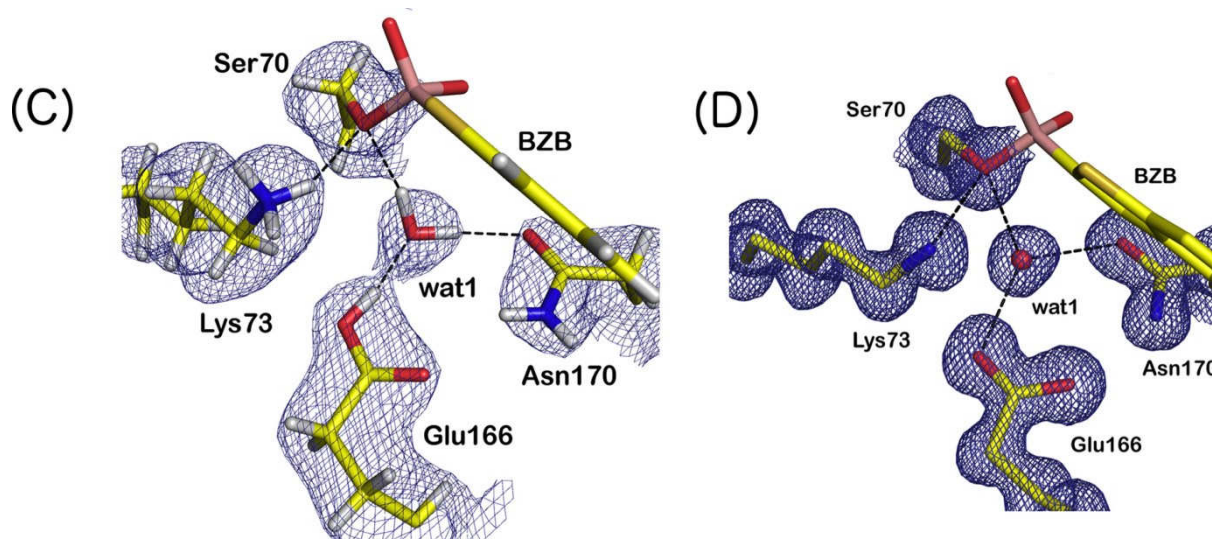

Fig. 10: *X-ray (right) and neutron (left) data on the BZB aduct in the catalytic cycle of a β-lactamase enzyme. (The Figure is taken from ref. [11]. This research was originally published in Journal of Biological Chemistry, Stephen J. Tomanicek et al. "Neutron and X-ray Crystal Structures of a Perdeuterated Enzyme Inhibitor Complex Reveal the Catalytic Proton Network of the Toho-1 β-Lactamase for the Acylation Reaction." © the American Society for Biochemistry and Molecular Biology.)*

Here, a different ligand was used as the natural antibiotics: benzothiophene-2-boronic Acid (BZB). This lignand was known to stop the protein in the acylation step. In addition to that, this ligand was also known to mimic this covalent intermediate between substrate and protein. In addition to that, neutron protein crystallography was used to detect the proton unambiguously. For the latter technique the protein was expressed in deuterated media such that a fully deuterated protein resulted. The BZB was also synthesized using deuterium instead of hydrogen atoms. In fact, even a special Boron isotope was used which shows less neutron absorption than the natural abundant boron which is known as a good neutron absorber. Fig. 10: left side shows the x-ray structure and Fig. 10: right side shows the neutron structure. The proton is clearly seen in the neutron structure and therefore the side chain Glutamate 166 could be identified as the base in the acylation step. With this knowledge better antibiotic drugs can be designed which bind this sidechain firmly blocking the protein's catalytic cycle in its first step.

Acknowledgement

References

[1] F. Maia, T. Ekeberg, N. Timneanu, D. van der Spoel, and J. Hajdu, Physical Review E **80**, 031905 (2009).

[2] L. Stryer, *Biochemie* (Spektrum Akad. Verlag, Heidelberg Berlin New York, 1991), 4. edn.

[3] P. Fromme and J. C. H. Spence, Curr. Opin. Struct. Biol. **21**, 509 (2011).

[4] J. Drenth, *Principles of Protein X-Ray Crystallography* (Springer Science+Business Media, LLC, New York, 2007).

[5] N. Niimura and A. Podjarny, *Neutron Protein Crystallography - Hydrogen, Protons, and Hydration in Bio-macromolecules* (Oxford University Press, Oxford, New York, 2011), IUCr Monographs on Crystallography, 25.

[6] G. J. Kleywegt and T. A. Jones, Structure **3**, 535 (1995).

[7] A. T. Brunger, Nature **355**, 472 (1992).

[8] T. Chatake, A. Ostermann, K. Kurihara, F. G. Parak, and N. Niimura, Proteins-Structure Function and Genetics **50**, 516 (2003).

[9] A. Ostermann, I. Tanaka, N. Engler, N. Niimura, and F. G. Parak, Biophys. Chem. **95**, 183 (2002).

[10] T. Chatake, A. Ostermann, K. Kurihara, F. G. Parak, N. Mizuno, G. Voordouw, Y. Higuchi, I. Tanaka, and N. Niimura, Journal of Synchrotron Radiation **11**, 72 (2004).

[11] S. J. Tomanicek, R. F. Standaert, K. L. Weiss, A. Ostermann, T. E. Schrader, J. D. Ng, and L. Coates, J. Biol. Chem. **288**, 4715 (2013).

Recommended Textbooks

on X-ray crystallography:

J. Drenth, *Principles of Protein X-Ray Crystallography* (Springer Science+Business Media, LLC, New York, 2007)

on neutron protein crystallography:

N. Niimura, and A. Podjarny, *Neutron Protein Crystallography - Hydrogen, Protons, and Hydration in Bio-macromolecules* (Oxford University Press, Oxford, New York, 2011)