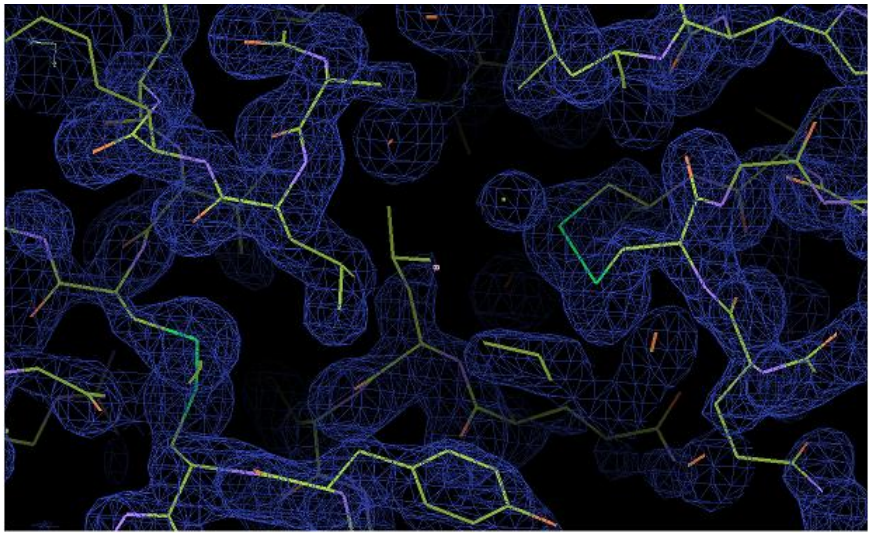
DESY Summer Student Programme 2009 Report

Carolin Bornemann

RWTH Aachen University

9th September 2009



Protein Crystallography at EMBL, Hamburg

Supervisor: Dr. Thomas Schneider,
Dr. Michele Cianci, Dr. Gleb Bourenkov

Contents

1 Summary

I worked in a project group for protein crystallography at EMBL Hamburg.
During the first term I was engaged in growing crystals, more precisely bovine insulin crystals. This protein is easy to crystallize and the properties of the crystal, for example cell parameters, mosaicity, resolution, freezing, are well reproducible. In addition insulin crystals are highly symmetric and therefore gladly used for all general test purposes.

Meanwhile the crystals were growing I worked through five EMBL tutorials "'for learning and teaching crystallography"' to introduce myself to the programs needed for data analysis. Thereby I evaluate collected data from several protein crystals like tetragonal thaumatin, monoclinic and tetragonal lysozyme and cubic insulin, too.

After that I might measure my own grown crystals and did the analysis for them to solve the structure of cubic insulin.

2 Growing Crystals

Proteins are the basic modules of cells built on amino acids. They do not only add structure to the cells, but also are molecular 'machines', transporting substances, pumping ions, catalysing chemical reactions and detecting molecular cues.
The spatial structure determines the functioning of the protein.

To study proteins it is easy to examine them in the crystalline state which is not their natural state. But the majority of crystal structures seems to be identical to the solution structure. Therefore proteins retain their function in the crystal and that is of capital importance because it is what you want to explore.
Protein crystallization is the most important part of crystallography. You need a well-ordered protein mono−crystal without any inclusions and large enough to diffract the x-ray beam with sufficient intensity.
Crystallization procedures are well known but it is impossible to make any predictions so the whole process is still empiric and by being successful you are lucky.

I used vapor diffusion in order to grow bovine insulin crystals. In this crystallization technique the protein and the precipitant solution equilibrate in a closed box which contains an aqueos reservoir whose precipitant solution promotes producing crystals. The further technique I used is called hanging−drop method.

Figure 1: Hanging−drop method [7] and a culture plate [7]

For the protein crystallization I used a well tissue culture plate in which 24 crystallization trials can be carried out. Each well is filled with a different mixture of Na_2HPO_4, Na_3PO_4, Na_4EDTA, Na_2EDTA and water, the recipitant and the buffer solution. So every composition holds a special pH−value and Na_2HPO_4/Na_3PO_4 concentration.
Then I put 2μl insulin solution with 2μl solution from one of the wells onto a microscope cover slide, inverted and closed the well with the cover slide (hanging−drop method) and repeated that for all wells. The trays were kept at room temperature (about 20°C) in the cold room.
It takes about 1−2 days until the crystals are grown. You can keep them 3 months at room temperature without decreasing quality.

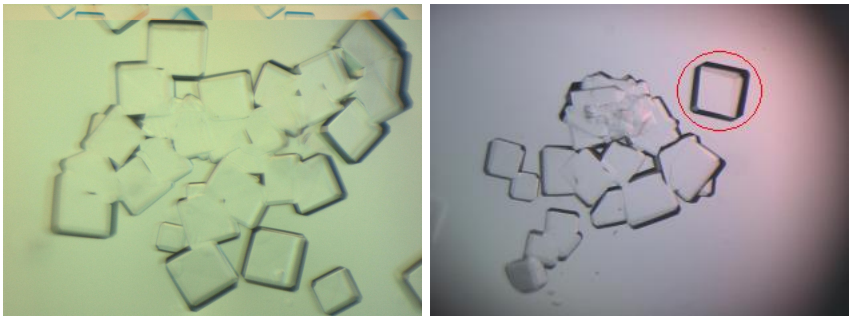


Figure 2: cubic insulin crystals

For data collection you need a single crystal therefore the one on the right side would be easily qualified.

3 Preparing The Experiment

Before starting the data collection you have to mount the crystal and even to check if the crystal diffracts well.

3.1 Mounting The Crystal

Using a microscope I selected nicely grown crystals and tried to take them out of the drop by catching them with a fiber loop. Before mounting the crystal on the sample holder I put them for about 10 seconds into a drop of insulin cyroprotectant. This is necessary because the crystal is measured in a cold nitrogen stream in order to reduce molecular motion and to improve diffraction. So you have to prevent the formation of ice crystals which would damage the protein crystal and even decrease the quality of diffration.

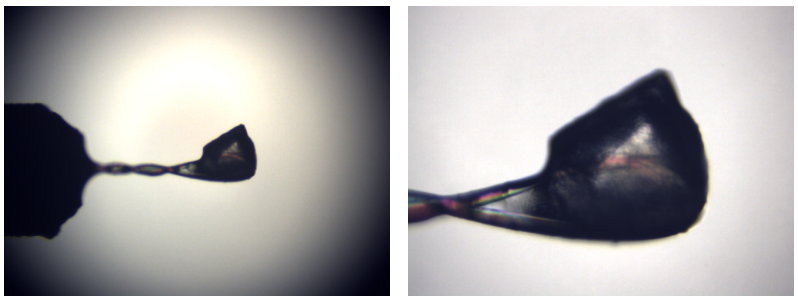


Figure 3: Crystal is mounted on a loop

Proteins in the crystal adhere to each other primarily by hydrogen bonds so they are much more fragile than inorganic crystals. Gentle pressure is enough to destroy them. Therefore handling and mounting a crystal for data collection is very challenging.

3.2 Judging Crystal Quality

For an accurate analysis you need the crystal to give sharp diffraction patterns. Therefore at first you have to check the crystal by eye for well mounting and inclusions or broken edges or even visible ice formations.
Then you carry out quickly a first measurement to examine the diffraction pattern and the shape of the reflections. Sometimes it happened that the mounted crystal looked even nice but did not diffract any more. This means that the crystal was dead, the crystal structure is destroyed.

4 Data Collection

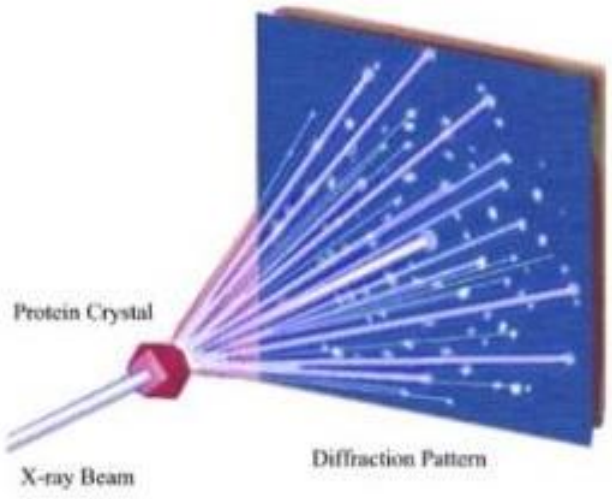


Figure 4: Principle of data collection [7]

The diffraction data were collected at the wiggler beamline BW7A for molecular crystallography.
The synchrotron source is a 1 Tesla multi−pole wiggler. BW7A holds two monochromators for a wide range of applications using different bandpasses with an energy range from 7−18keV, respectively 1,8−0,7 angstrom.
The station is optimized for SAD or MAD data collection and uses a MARCCD detector (165mm).

The data collection parameters are given below:

wavelength	0,874 A
detector distance	139,9 mm
oscillation range	1,0°
exposure time/image	1 sec

X-rays of wavelength λ are diffracted by electrons in parallel planes (hkl) in the crystal by holding Bragg's law. Here ϑ is the wave angle and n is an integer.

$$2d_{hkl}sin\vartheta = n\lambda$$

A diffraction pattern looks like this:

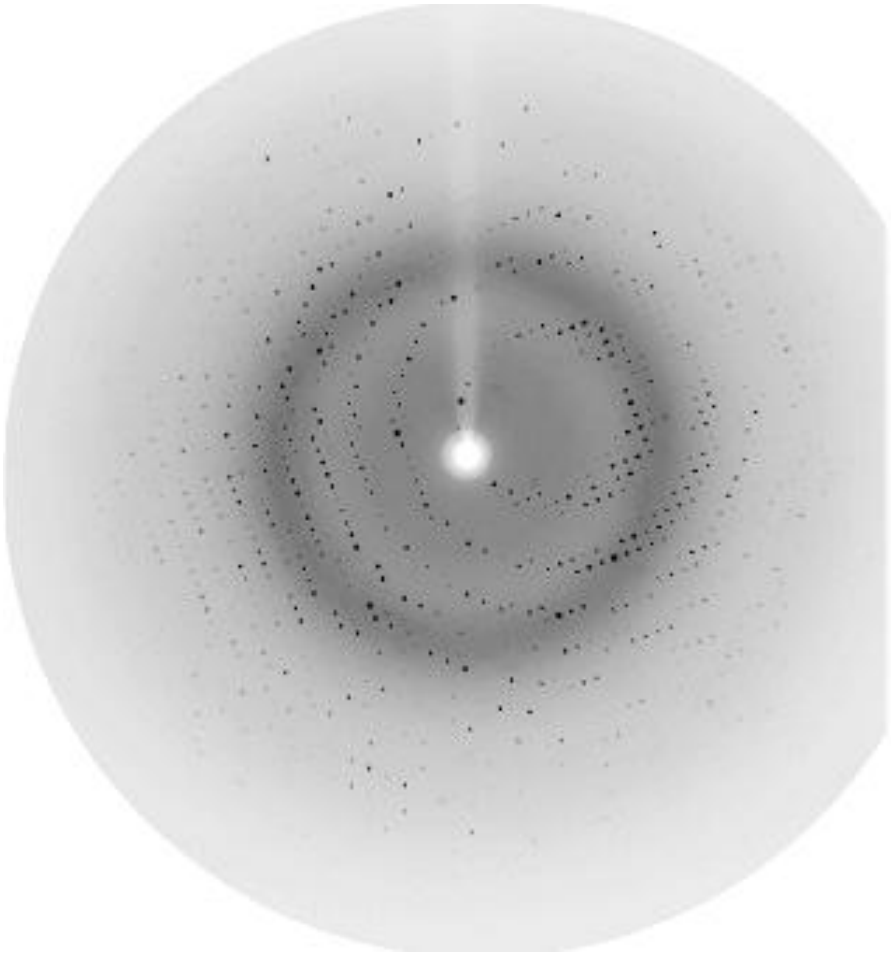


Figure 5: Diffraction pattern of insulin crystal

The points are the reflections needed for further analysis. The rings in the background are the so called ice rings which appear due to ice formation on the sample holder.

5 Analysis

5.1 Integration

The spacing of reflections in reciprocal space are inversely proportional to the unit cell dimensions. So defining real lattice spacings from reciprocal lattice spacings can just be solved by using geometric methods.

For the integration I use the XDS program (Kabsch, 2009).
At first it has to determine the origin (000) which lies in the direct path of the beam. So it just calculates the beam center position.
Then the reflections are searched and their positions are measured. After that the programm estimates the unit cell dimensions and the space group and attaches to each reflection an index hkl corresponding to the position in the reciprocal space.

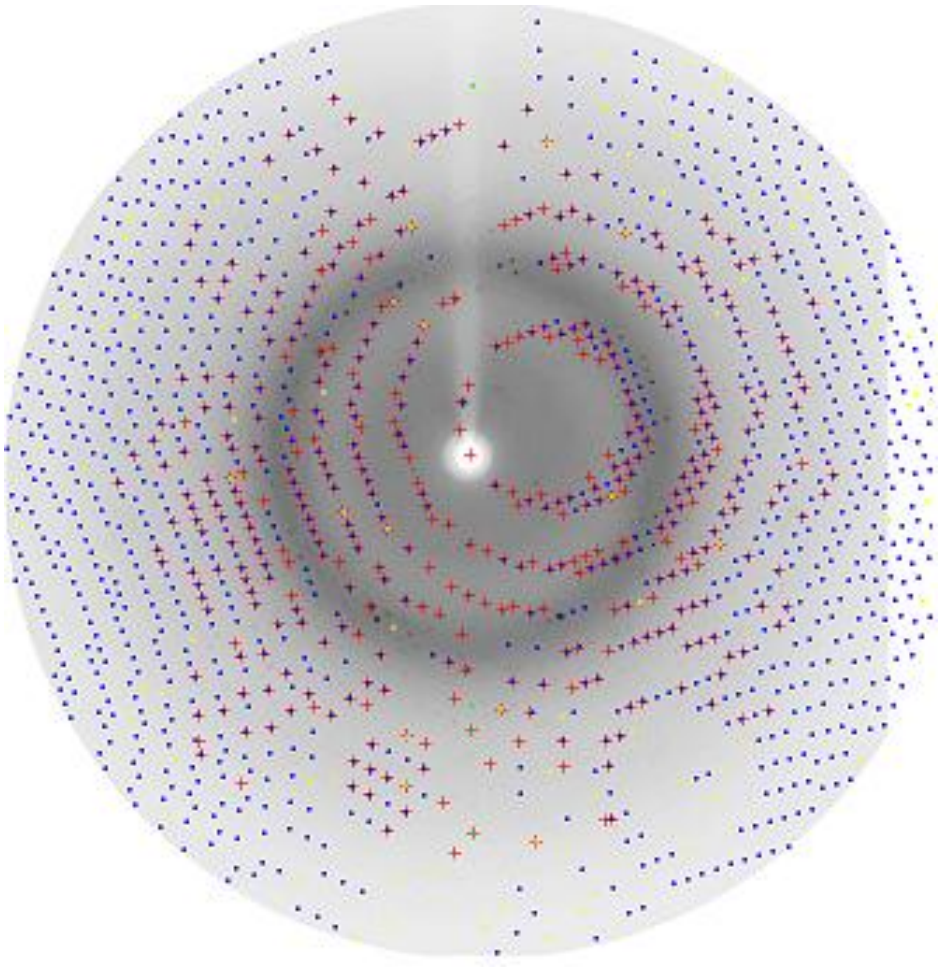


Figure 6: Diffraction pattern of crystalline insulin

The red crosses denote the visible reflections. The blue ones are a prediction due to the estimated unit cell. And as you can see it works fine.

Reflections near the origin are created by sets of widely spaced planes and carry

information about larger characteristics of the molecules in the unit cell. The reflections far from the origin are created by closely spaced planes and carry information about details of structure. During the subsequent integration the programm collects three dimensional profiles of all reflections and calculates their intensities.

The determined unit cell parameters are given below:

	a[A]	b[A]	c[A]	α[°]	β[°]	γ[°]
XDS	79.48	79.48	79.48	90	90	90

Cubic insulin has got the space group $I2_13$ or space group number 199.

5.2 Scaling

The collected images have to be on a common scale. Some reflections occur on several frames with different intensities but they should have the same because they are created by the same parallel planes (hkl) each time.
Corrections factors are calculated and applied to compensate radiation damage and absorption effects.
For this part of the analysis I use the scala program (Evans, 1996) which then scales together multiple observations of reflections and determines an average intensity for each reflection.
In this way identical reflections are attached to identical intensities. This process is called scaling.

5.3 Phasing

The relationship between an object and its diffraction pattern is described by Fourier transform. The transform allows us to convert a Fourier sum description of the reflections to a Fourier sum description of the electron density. A reflection is described by a structure factor equation containing one term for each atom in the unit cell.
In turn a Fourier sum in which each term is a structure factor describes the electron density. So we use the Fourier transform to convert the structure factor F_{hkl} to the desired electron density ρ(x,y,z).

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl}\ e^{-2\pi i(hx+ky+lz)}, \tag{1}$$

where V is the volume of the unit cell.
Looking at this equation you can see that the three frequencies h,k,l are the indices of reflection so determining the structure factor F_{hkl} still remain.
A diffracted ray is described by a structure factor which includes the following parameters: amplitude, wavelength and phase. The reflection intensity I_{hkl} is proportional to the squared amplitude $|A_{hkl}|^2$ so we just have to measure the intensity of the reflections. The frequency of the stucture factor can be calculated from $1/d_{hkl}$, making the wavelength the same as the spacing of the planes producing the reflection.
The only parameter which is not directly deducible from a single measurement is the phase.

For obtaining the phase I use the SHELX programs which estimates the phase due to anomalous dispersion.
Anomalous dispersion means that the intensity of symmetry related reflections are not equal. Friedel′s law $I_{hkl} = I_{\overline{hkl}}$ is not valid.
It is known that insulin contains 6 sulphur atoms arranged in 3 disulphide bridges. So you can determine the structure with sulphur−SAD, sulphur single−wavelength anomalous diffraction.
The program SHELXC (Sheldrick et al., 2001; Sheldrick 2008) prepares the data for macromolecular phasing with the programs SHELXD (Schneider and Sheldrick, 2002) and SHELXE (Sheldrick, 2002). SHELXD locates the heavy atoms, in this case the sulphur atoms, and SHELXE does the phasing and density modification by using SAD methods.

5.4 Displaying The Electron Density Map

To visualize the estimated electron density you can use the programm coot (Emsley and Cowtan, 2004). Coot displays a three dimensional diagram of the spatial structure which can be rotated and by zooming in you can detect heavy sulfur atoms and functional groups.
The electron density map of bovine insulin is shown below.
Of course atoms are definded by the electron density and therefore the spatial distribution of atoms in the unit cell can be visualized. This is shown due to the used colours in the electron density map. Yellow indicates carbon atoms, the blue ones are disulphide bridges, oxygen is displayed in red and the purple ones represent nitrogen atoms. The light hydrogen atoms can not be seen here but by knowing the structural formula of all amino acids you can imagine where they have to be.
With biological knowledge you can now identify the specific amino acids and see the spatial structure of the protein.

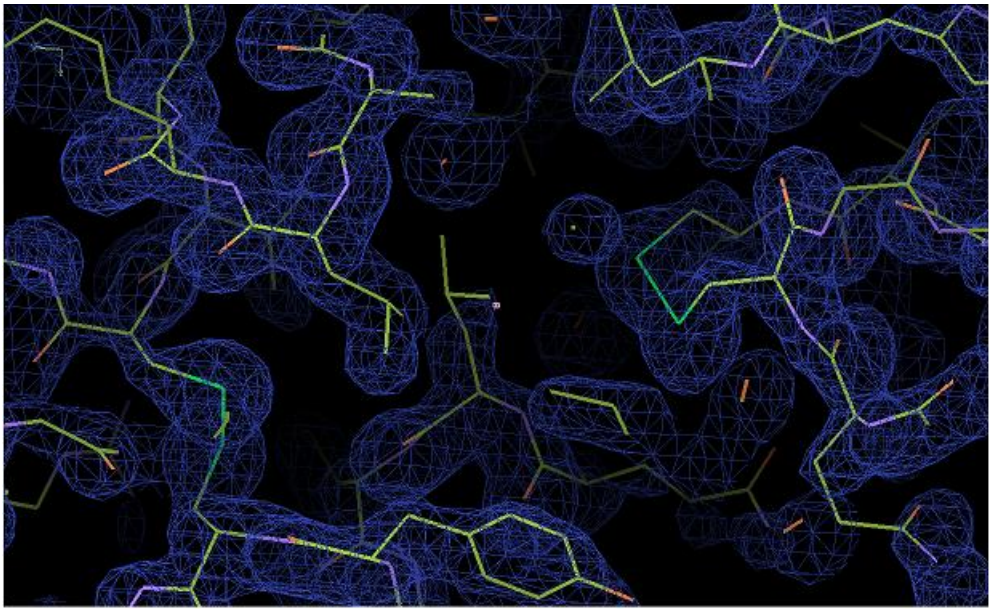


Figure 7: Electron density maps of cubic insulin

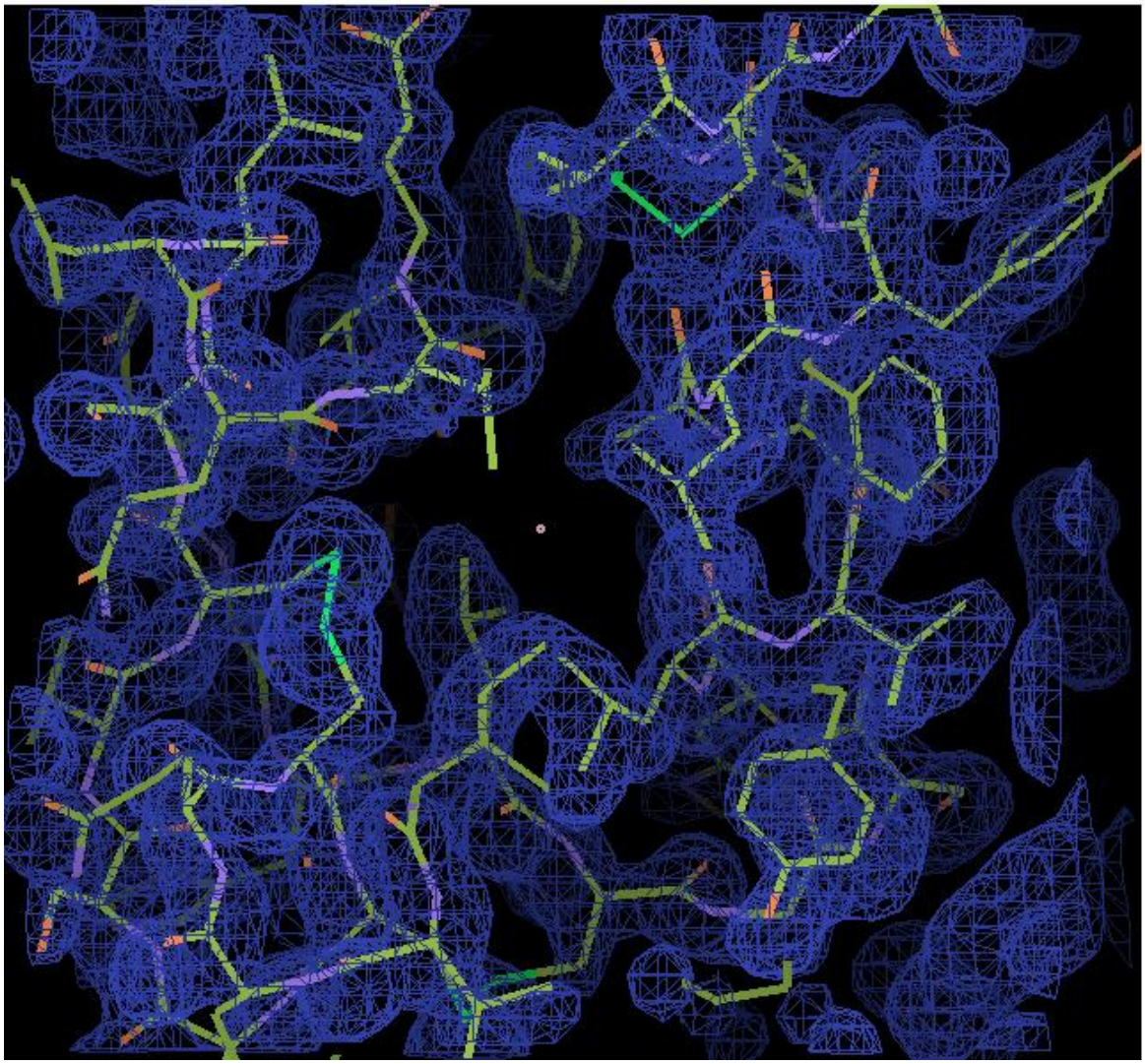


Figure 8: Electron density maps of cubic insulin

References

[1] G. Rhodes, Crystallography Made Crystal Clear, Third Edition, 2006

[2] http: //www.embl-hamburg.de/xraytutorial/experiment1.html, 08.09.2009

[3] http: //strucbio.biologie.uni−konstanz.de/xdswiki/index.php/XDS, 08.09.2009

[4] http: //www.ccp4.ac.uk/, 08.09.2009

[5] http: //strucbio.biologie.uni−konstanz.de/ccp4wiki/index.php/SHELXC/D/E, 08.09.2009

[6] http: //www.rcsb.org/pdb/home/home.do, 08.09.2009

[7] http: //userpage.chemie.fu−berlin.de/saenger/crystal.htm, 08.09.2009

[8] http: //proteincrystallography.org, 08.09.2009