

Ultra-high resolution protein crystallography

Ashley Deacon¹, Thomas Gleichmann¹, S. Harrop¹, John R. Helliwell^{1*} and A.J. Kalb (Gilboa)²

¹Department of Chemistry, University of Manchester, Manchester M13 9PL, England.

²The Weizmann Institute, Israel

* To whom correspondence should be addressed.

Many years ago the idea of collecting voluminous quantities of weak reflections to form a high resolution data set, within the constraints of radiation damage, was a particular challenge [1]. At CHESS the bringing together of a 24 pole wiggler with prodigious x-ray flux output at short x-ray wavelengths, along with very sensitive CCD detectors and freezing of crystals have provided a means to certainly match those best hopes. So much so that the data that can be realized are arguably best described as ultra-high resolution, at least, as evidenced in our studies on the plant protein concanavalin A of 25 kDa molecular weight. The intrinsic property of this protein is to bind sugar molecules. It is implicated in cell to cell recognition processes and is widely used as a laboratory diagnostic tool including biological staining.

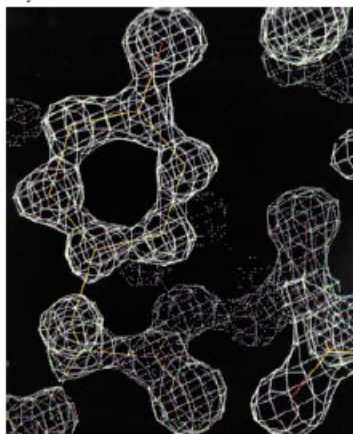
The study of concanavalin A at Cornell is particularly appropriate as it was here that Prof. J.B. Sumner originally isolated and crystallized this protein in 1919. One awe inspiring feature of J.B. Sumner was that as a young man he lost his arm in a shooting accident. Undaunted he pursued his scientific laboratory work. He won the Nobel Prize for

Chemistry in 1946, with J.H. Northrop, for demonstrating that enzymes were indeed proteins; Sumner had isolated and crystallized jack bean urease and Northrop chymotrypsin.

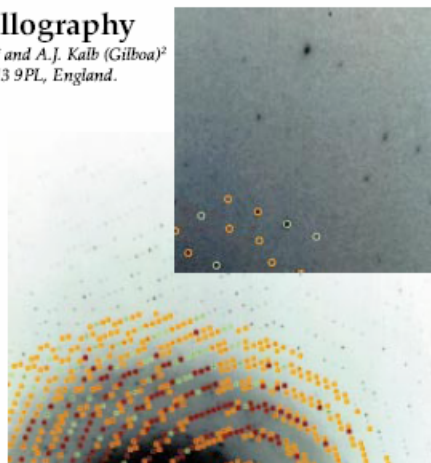
Our structural studies have been in overall collaboration with Dr. Joseph Yariv at the Weizmann Institute.

The objectives of our program of work have been to determine the structures of sugar bound forms of concanavalin A and to refine these and the saccharide free protein to as high a resolution as possible. In previous work, our refined structures were completed at 2 Å and 1.6 Å respectively [2, 3] using room temperature film and image plate synchrotron techniques respectively on SRS 7.2 [4] and DORIS XI1. The basic structure of the sugar free form of concanavalin A was solved some 20 years ago by two groups [5,6]. The structural changes associated with sugar binding, namely reorientation of amino acid side chains and the displacement of bound water molecules [2] can provide a model system for theoretical chemistry studies of molecular recognition.

With the further development of synchrotron facilities it is possible to consider the measurement of ideal or near-ideal diffraction data [7] from which as precise a structure as possible can be obtained. The definition of an ideal diffraction data set is one that is as free as possible of systematic and random errors as well as being complete to as high a resolution as possible. The use of a wavelength of 0.5 Å (or shorter) could eliminate absorption errors at better than the 1%



(Figure 2) A portion of a $(2F-F_2)$ electron density map of a tyrosine side chain. Resolution 1.1Å.



(Figure 1) (bottom) Diffraction pattern recorded from concanavalin A (CHESS A1, CCD, frozen crystals) extending to 0.90 Å resolution. Prediction to 1.6Å resolution (circles) superimposed. (top) Zoom in to the outermost parts of this pattern and increase of the gray scale. Prediction to 1.0Å resolution (circles) superimposed.

level (i.e. less than the random errors that can so far be realized). A finely focussed 0.5 Å beam is an objective for the future.

At CHESS we were able to use a 0.9 Å wavelength beam of very high intensity on the A1 station. The 50x50 mm² Princeton 1K CCD detector [8] was used. Concanavalin A crystals freeze readily. The mosaicity of the crystals was kept as small as 0.25° by using small crystals (0.1 x 0.2 x 0.5 mm³), since bigger crystals (0.2 x 0.4 x 1.0 mm³) showed mosaicity increases to 0.7°. Also, by keeping the CCD detector close to the crystal and using a 0.1mm collimator spot sizes were kept small thus matching the small pixel size of the CCD (50µm). A short sample to CCD detector distance (40mm) was a necessity anyway since it was observed that the diffraction extended to beyond 0.9 Å (Fig. 1), the geometric limit with the direct beam position off the bottom of the detector. The first and second solvent rings are discernible and then an 'ocean' of diffraction spots spreading towards the outermost part of the diffraction pattern. The dynamic range of the intensities is too large to display in figure 1a, so figure 1b shows a zoom into the outermost corner of the diffraction pattern (0.9 Å at the corner). The exposure

times per image were 20 seconds per 0.5° rotation. CHESS injection currents were 80 mA - 150 mA and are set to increase further. The readout time of the detector was 20 seconds per image, some 4 times quicker than the current image plate systems and quicker than any planned image plate device. The CCD readout can in any case be further reduced to ≤ 5 seconds. A voluminous quantity of data were collected then with little penalty on deadtime due to poor duty cycle.

Largely due to a limited number of scheduled slots and lack of a cold storage device, a total of 5 frozen concanavalin A crystals were used. Several slow pass runs, as described above, were used. Also one quick pass run was done involving 5 seconds exposure per image to record the diffraction pattern to 2.0 \AA , which consisted previously of overloaded spots. These crystals had similar unit cell dimensions to within 0.1 \AA . There is scope here then to improve the dynamic range of the detector. A total of 500 diffraction images have so far been processed using DENZO and merged with SCALEPACK (programs of Z. Otwinowski and W. Minor) to yield 89000 unique reflections with merging R factor on I of 6.6% and 63.6% complete to 0.98 \AA resolution. Owing to the geometry of the setup and the processing done so far, the data set has a very low completeness between 0.98 \AA and 1.07 \AA . It is 83% complete up to 1.07 \AA . An example of a $2F_o - F_o$ electron density map is shown (Fig. 2), with a tyrosine residue and several water molecules beautifully defined.

An intriguing prospect with data of this quality and at this resolution is to contemplate novel direct structure solution strategies based on metal atoms in a protein structure. The idea here being that, since the metal atoms can readily be located from wavelength optimized anomalous scattering, perhaps the partial model of the atoms [9] can yield starting phases. Concanavalin A contains a Mn and a Ca atom. They are some of the most ordered atoms in the structure so that the fractional contribution of their scattering will increase with Bragg angle. Overall, to 1.1 \AA resolution, the mean phase error between the

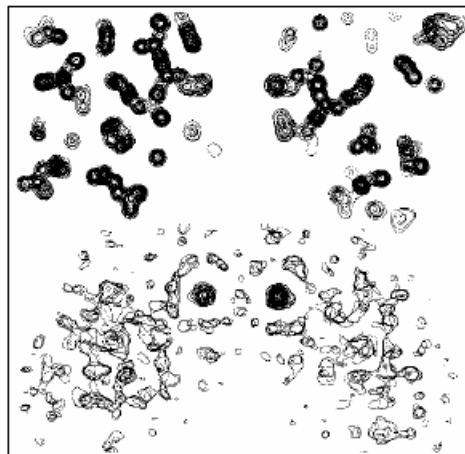
Mn+Ca (2-atom) phases) and the whole model (1839 atom phases, excluding water) is 83° . Random agreement would be 90° . An electron density map calculated to 1.1 \AA with calculated (perfect) phases and then with coefficients $(2F_o - F_o)_{\text{Mn+Ca}}$ is shown in figure 3. It clearly shows the outline of the protein subunit in the unit cell. The connectivity is poor. Computationally the map can be subject to a variety of image modification techniques now available, and will be tried to see how close the quality of the 'perfect' map can be approached.

Experimentally more can be done, as referred to above, to go to shorter wavelengths than 0.9 \AA and so realize ideal data as defined above. Most importantly at this stage though is to gather the data beyond the aperture of the CCD detector. A bigger aperture is readily achieved with an image plate. At CHESS a Fuji image plate scanner is available along with a set of $250\text{mm} \times 200\text{mm}$ Fuji image plates. It was possible to place a single image plate cassette some 52mm from the crystal. This distance is quite similar to that used for the CCD, although slightly more favorable in fall off of x-ray background by $(52/40)^2$. Diffraction spots were visible at 0.9 \AA resolution. However, the data processing results beyond 1 \AA were disappointing. Even at 1 \AA the merging R on I was 44% (versus 13%

for the CCD to 0.98 \AA). The poor performance of the image plate with weak spots has, of course, been documented [10]. One might say then that the image plate had the aperture but not the sensitivity and the CCD had the sensitivity but not the aperture! At CHESS there is a program (funded by the Keck Foundation) to provide bigger CCD devices and tiled array devices in the future. This is certainly worthwhile.

These diffraction resolution limits (0.90 \AA and beyond) open the prospect of not only better refined models (e.g. as done with rubredoxin, a 6 kDa molecular weight protein [11]) and new structure solving strategies as mentioned above, but also even valence electron density distribution analysis around the metal atoms. This had been pointed out many years ago for zinc insulin [12] but could be more general. Concanavalin A at 25kDa molecular weight sets a wide range of possibilities as there are obviously many proteins with molecular weights up to this level. Such a molecular weight ceiling applies too in NMR spectroscopy. This ceiling encompasses many proteins. In synchrotron x-ray crystallography there are still further developments in store, including bigger (or tiled) CCD detectors as well as higher currents and stronger beams still at CHESS.

Acknowledgments. None of this work would have been possible without the dedicated efforts of Prof. S. Ealick, the MacCHESS team, and all the CHESS staff.



(Figure 3) Electron density map of concanavalin A to 1.1 \AA with calculated (perfect) phases (top) and phased on two atoms (Mn+Ca) as partial model (bottom).

- [1] J.R. Hellwell (1979) Daresbury Study Weekend DLSCI R13, p1-6.
- [2] J.H. Naismith et al. (1984) Acta Cryst D60, 847-868.
- [3] G. Enmerich et al. (1994) Acta Cryst D60, 749-756.
- [4] J.R. Hellwell et al. (1982) J. Phys. E. 15 1363.
- [5] K.D. Hardman et al. (1982) J. Mol. Biol. 157, 69-96.
- [6] G.N. Reeko et al. (1975) J. Biol. Chem. 250, 1525-1547.
- [7] J.R. Hellwell et al. (1993) Acta Cryst D49, 120-128.
- [8] M. Tate et al. (1995) J. Appl. Cryst. 28, 196-206.
- [9] G. Sheldrick personal communication.
- [10] R. Lewis (1994) J. Synchrotron Radiation, 1, 43-53.
- [11] G. Sheldrick et al. (1993) Acta Cryst D49, 19-23.
- [12] N. Saksela et al. (1981) Structural Studies on Molecules of Biological Interest. Volume in honour of Dorothy Hodgkin.