FAST COMMUNICATIONS

Contributions intended for this section should be submitted to any of the Co-Editors of Acta Crystallographica *or* Journal of Applied Crystallography. *In the letter accompanying the submission authors should state why rapid publication is essential. The paper should not exceed two printed pages (about 2000 words or eight pages of double-spaced typescript including tables and figures) and figures should be clearly lettered. If the paper is available on 5.25" IBM PC compatible or 3.5" Apple /Macintosh diskettes it would be helpful if these could be sent with the manuscript together with details of the word-processing package used. Papers not judged suitable for this section will be considered for publication in the appropriate section of* Acta Crystallographica *or in* Journal of Applied Crystallography.

Acta Cryst. (1993). D**49**, 308–310

Protein single-crystal diffraction with 5 Å synchrotron X-rays at the sulfur *K*-absorption edge

BY M. S. LEHMANN [1], H.-H. MÜLLER [2] AND H. B. STUHRMANN [2]

[1] *Institut Laue–Langevin, Avenue des Martyrs, 156X, F-38042 Grenoble, France*

[2] *GKSS Forschungszentrum, D-2054 Geesthacht, Germany*

(*Received* 16 *October* 1992; *accepted* 13 *November* 1992)

Abstract. Sulfur atoms, an integral part of many proteins, are possible candidates for anomalous scattering in phase determination by multiple-wavelength methods. The main difficulty encountered is that a wavelength of about 5 Å is required to obtain a large anomalous signal from these atoms, leading to very large absorption effects. Initial experiments have been carried out using a synchrotron X-ray source, evacuated beam tubes, a diffractometer inside a vacuum chamber, a special sample holder and a suitable scattering geometry. The results are encouraging, showing that Bragg reflections can be measured, and that changes in their intensities around the absorption edge are observable.

Introduction. An important step in the determination of a new protein structure by crystallographic techniques is the location of some strongly scattering atom(s) (Green, Ingram & Perutz, 1954) which can then be used as a reference for phasing the diffraction data (Blow & Crick, 1959). These atoms are termed 'heavy atoms', and are mostly electron-rich atoms, *e.g.* metal atoms, which can be bound to the protein in its crystalline state.

It is also possible, however, to use the anomalous resonance scattering at an absorption edge for select atoms to perform the same task (Hendrickson, 1991), and in principle an ideal target would be sulfur, which is often both naturally present and well fixed inside the protein. The *K*-absorption edge for sulfur is found at $\lambda_K = 5.018$ Å, so major obstacles are the general availability of suitable X-ray sources and the high absorption of most materials to these X-rays, requiring the measurement to be performed in a vacuum or in a helium atmosphere. However, with present-day synchrotron sources and a specialized diffractometer these problems can be handled, and here we present some initial experiments investigating this approach.

General considerations. Both when heavy atoms are used and when the phases are obtained from multiple-wavelength observations, the approach is by difference techniques. Several sets of structure amplitudes, $F(hkl)$, are measured, and it is the strength of the difference signals, $\Delta F(hkl)$, that ensures the successful outcome.

For two sets of structure amplitudes the mean relative signal $\langle \Delta F(hkl) \rangle / \langle F(hkl) \rangle$ can be approximated (Crick & Magdoff, 1956; Hendrickson, Smith & Sheriff, 1985) by $1/2^{1/2}(\Sigma f_S^2/\Sigma f_P^2)^{1/2}$ where the f_S are the scattering factors for the atoms which contribute to the change, and f_P are the scattering factors for the normal atoms in the protein. If only one type of special atom is present, then this can be rewritten as $1/2^{1/2}(f_S/f_P)(N_S/N_P)^{1/2}$, where N_S is the number of special atoms, N_P is the total number of atoms in the protein and f_P is assumed to be the same for all atoms in the macromolecule.

If the anomalous signal from sulfur is used for crystallographic phasing, then the variation would be *ca* 8 electrons when moving across the *K*-absorption edge, thus $f_S \simeq 8$. Setting f_P to 6.7 electrons – the mean number of electrons for an atom in a protein – the expected average relative change in $F(hkl)$ then becomes $0.8(N_S/N_P)^{1/2}$.

In general we can count on finding 3–5% sulfur-containing residues (Dayhoff, Dayhoff & Hunt, 1976), and assuming seven atoms per amino-acid residue, of which 4% hold sulfur, the estimated average relative change will be 0.06. With precise data this would be a sufficient signal for structure solution, so a considerable number of proteins are thus potential candidates for the approach.

It is also worth noting that the relative signal does not depend on the size of the protein, although it is of course harder to measure precise data on a large system.

Soft X-ray diffraction. The measurements to verify these considerations were carried out on the A1 instrument (Stuhrmann, Goerigk & Munk, 1990) at HASYLAB, which uses X-rays from the DORIS storage ring. The whole beam path is in a vacuum, and there are no windows towards the storage ring. The sample is held on a

0907-4449/93/020308-03$06.00

Eulerian cradle, the position-sensitive detectors can cover the scattering angle from –49 to 117°, and for most of the experiments the reflections with d spacings in the range 4.2–7.2 Å were observed with one detector. To avoid the effects of vacuum on the protein crystal, this was held in a cylindrical cell (Müller, Lehmann & Stuhrmann, 1991) with mylar windows. The absorption coefficient for both the mylar and the protein crystal is around 500 cm^{-1}, so it is essential to use very thin foil. Likewise, for the protein crystal, transmission geometry would require very thin plates, ideally not much thicker than 20 μm, and this would be an obstacle to routine measurements.

However, to reach reasonable d spacings for structure analysis between *ca* 7 and 3 Å, the long wavelength imposes large scattering angles, so in this case reflection geometry becomes the obvious choice, leading to back-scattering as the limit of 2.5 Å is reached. This is a very favourable geometry in the case of strong absorption, and the intensity of the reflection is only restricted by the size of the crystal surface exposed to the beam. Most crystal shapes can be used in this geometry – except for thin plates where parts of the data would be recorded as transmission measurements.

Tetragonal crystals of hen egg-white lysozyme (Blake, Koenig, Mair, North, Phillips & Sarma, 1965) were used, and scans were made on 0.5 mm size crystals in reflection geometry. For simple intensity measurements the wavelength was chosen to be below the K-absorption edge, and two sets of data were recorded with the normal of the crystal surface making angles of 60 and 80° to the incoming beam direction, respectively. Fig. 1 shows one such scan. All 32 reflections predicted were observed, and the agreement with structure amplitudes recorded on a Fast diffractometer employing Cu radiation was $R_F = 0.16$. The counting statistics, expressed as $R_S = \Sigma\sigma(F)/\Sigma F$, was 0.014.

The detector was then moved to cover 2.9–3.7 Å in d, and a scan was made with the normal at 45° to the beam. This time 14 out of the 19 predicted reflections were found, R_F was 0.21 and R_S was 0.059. The degradation in quality is probably mainly due to radiation damage as this measurement was done 6 h after the first scans.

Altogether it shows that it is possible to observe intensities at this wavelength with a reasonable counting statistical error, but the agreement with traditional data is not satisfactory. Some of it may well be associated with large absorption effects, but the main source of error is at present undoubtedly the shadows from the window of the detector. This is a multiwire proportional counter where the thin mylar front window (towards the vacuum) is held with a steel grid. The wire thickness of this grid is 0.3 mm, and the spacing between wires of 1.3 mm is comparable to the Bragg spot size. While this is very suitable for small-angle scattering purposes, which until now has been the main use of the instrument, it is obviously not adapted for single-crystal work, and will be modified when quantitative phase measurements start.

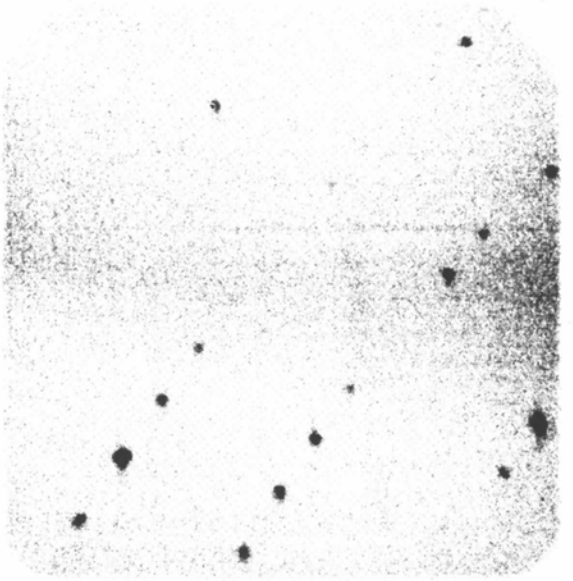

Fig. 1. Integrated scan for reflection geometry. The normal to the crystal face made an angle of 60° to the incoming beam and the detector covers d spacings in the range from 7.2 (top) to 4.2 Å (bottom). The crystal was at ambient temperature, the scan range was 2.4°, the individual step was 0.02° and the total measurement time was 1 h. The two rows of four Bragg reflections are 10,6,6 to 13,6,6 and 11,7,5 to 14,7,5. From the peak width and the distance between Bragg peaks, which in this case corresponds to the a axis of the crystal (79.2 Å), we note that the present instrument would be able to handle unit cells with axis lengths of more than 200 Å.

Measurements around the absorption edge. Absorption curves for both lysozyme and cystine were recorded across the K-absorption edge, and were used to select the wavelengths for a study of the anomalous-scattering effects. Measurements were then made for the two values indicated in Fig. 2. These were chosen to get near maximum change in f' and the same value for f'' in the two cases, thus leading to identical absorption behaviour for the two sets of measurements. The detector was set to cover d spacings from 4.2 to 7.2 Å, and the scan speed was typically 2° h^{-1} with steps of 0.01°, giving average counting statistical errors of less than 2%. Scans were performed twice, alternating between the two positions. This gave a measure of the reduction in intensity, which was 4% h^{-1}. Care was taken, therefore, that the identical reflections for the two wavelengths were all measured with the same time interval. In this way only one overall scale factor between the two sets was used.

The structure of tetragonal lysozyme is well known, and calculated structure factors were obtained using the coordinates of Kundrot & Richards (1987) as deposited with the Protein Data Bank (Bernstein, Koetzle, Williams, Meyer, Brice, Rodgers, Kennard, Shimanouchi & Tasumi, 1977; Abola, Bernstein, Bryant, Koetzle & Weng, 1987). Ratios between the observed structure factors above and below the absorption edge were then compared to the calculated value, and to show the influence of the anomalous scattering the calculation was performed for all possible combinations of the two wavelengths in the range around

the absorption edge. The best agreement should then occur for the set of wavelengths actually used in the experiment. This is depicted in Fig. 2, where the r.m.s. deviation between the observed and calculated ratios is given in a contour plot, where indeed the minimum is found for the set of wavelengths used in the experiment. The r.m.s. for the minimum is relatively high, and this is certainly mainly because of the detector-window effects discussed above.

Discussion and concluding remarks. Although it has only been possible to observe a few reflection ratios from a given crystal these preliminary results show that it is possible to see a signal, and that with higher flux and reduced radiation damage it should be possible to extract phase information. The flux at the specimen was found to be 10^9 photons s^{-1} mm^{-2}, but this can conceivably be increased by several orders of magnitude using insertion devices and sources with higher brilliance. The main obstacle is radiation damage, and here the only solution is to work with frozen crystals, a technique which in many recent cases has led to significant increase in the crystal lifetime. Following Henderson (1990), for 2.5 keV photons, the present flux and a penetration depth of 20 μm, the predicted limit would be one month at liquid-nitrogen temperatures before radiation damage occurs.

From the observed statistics the estimated time required to collect data for four wavelengths to better than 1% would be around 4 weeks, so only one or two crystals would be required to collect the full set of data. A similar argument would of course hold with a higher flux instrument. Therefore, we are presently setting up cryostat equipment.

Finally, the absorption is very high and in most cases no reliable correction can be made. This should not cause any problem, as the different sets of data are recorded for nearly identical geometries with the shift in Bragg angles being only a fraction of a degree.

At present there is only one famous case of a structure solution employing anomalous scattering of sulfur and Cu radiation (Hendrickson & Teeter, 1981), and because of the small signal of about 0.6 e it will probably not be generalized. If sulfur atoms – or phosphorus with λ_K = 5.784 Å – have to be used as phasing atoms, 5 Å X-rays are imposed. At first sight this appears a most inconvenient wavelength range, but our initial experiments, together with present-day developments of sources and cryogenic techniques, indicate that single-crystal diffraction with 5 Å X-rays from being a curiosity could become a useful tool for macromolecular structure solution.

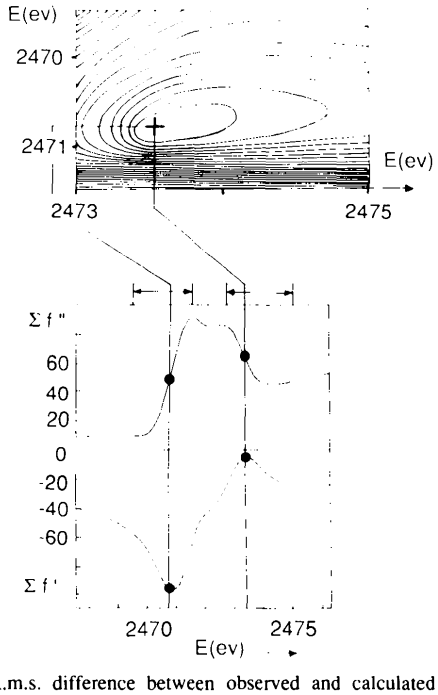

Fig. 2. R.m.s. difference between observed and calculated ratios of structure-factor amplitudes. In the bottom part f' and f'' for lysozyme are given based on eight Cys and two Met residues, where the anomalous values for the two residues have been observed independently (Stuhrmann, Goerigk & Munk, 1990) under identical instrumental conditions. In the contour diagram (contour intervals of 0.01) the r.m.s. difference for the ratios of the 19 reflections in this measurement is given, calculated for all wavelengths in the indicated regions. The vertical direction of the contour diagram corresponds to possible values below the absorption edge, while the horizontal direction is for values above the edge. The two wavelengths actually used are marked with large dots in the bottom part of the figure, and the connecting lines show that they do correspond to the minimum, which is 0.24.

We thank members of the staff of the EMBL and the MPI at HASYLAB for help during the experiments. Dr A. W. Thompson of the EMBL, Grenoble, is thanked for many enlightening discussions. We are also indebted to M. Hütch and O. Kühnholz for help and advice on data handling, and one of us (MSL) gratefully acknowledges hospitality and support during his stay at DESY, Hamburg.

References

ABOLA, E. E., BERNSTEIN, F. C., BRYANT, S. H., KOETZLE, T. F. & WENG, J. (1987). *Crystallographic Databases – Information Content, Software Systems, Scientific Applications*, edited by F. H. ALLEN, G. BERGERHOFF & R. SIEVERS, pp. 107–132. Bonn/Cambridge/Chester: IUCr.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BLAKE, C. C. F., KOENIG, D. F., MAIR, G. A., NORTH, A. C. T., PHILLIPS, D. C. & SARMA, V. R. (1965). *Nature (London)*, **206**, 757–761.

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.

CRICK, F. H. C. & MAGDOFF, B. S. (1956). *Acta Cryst.* **9**, 901–908.

DAYHOFF, M. O., DAYHOFF, R. E. & HUNT, L. T. (1976). *Atlas of Protein Sequence and Structure* 5, p. 301.

GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. R. Soc. London*, **225**, 287–307.

HENDERSON, R. (1990). *Proc. R. Soc. London Ser. B*, **241**, 6–8.

HENDRICKSON, W. A. (1991). *Science*, **254**, 51-58.

HENDRICKSON, W. A. & TEETER, M. M. (1981). *Nature (London)*, **290**, 107–108.

HENDRICKSON, W. A., SMITH, J. L. & SHERIFF, S. (1985). *Methods Enzymol.* **115**, 41–55.

KUNDROT, C. E. & RICHARDS, F. M. (1987). *J. Mol. Biol.* **193**, 157–170.

MÜLLER, H.-H., LEHMANN, M. S. & STUHRMANN, H. B. (1991). *HASYLAB Jahresber.* pp. 453–454 .

STUHRMANN, H. B., GOERIGK, G. & MUNK, B. (1990). *Handbook on Synchrotron Radiation* 4, edited by E. RUBINSTEIN, S. EBASHI & M. KOCH, pp. 555–580. Amsterdam: Elsevier.