PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Neutron Diffraction Studies of Proteins

G. A. Bentley and S. A. Mason

Phil. Trans. R. Soc. Lond. B 1980 **290**, 505-510
doi: 10.1098/rstb.1980.0111

Phil. Trans. R. Soc. Lond. B **290**, 505–510 (1980)
Printed in Great Britain

Neutron diffraction studies of proteins

By G. A. Bentley and S. A. Mason
Institut Max von Laue – Paul Langevin,
156X Centre de Tri, 38042 *Grenoble cedex*, *France*

Neutrons interact differently with protein crystals than do X-rays. Not only do hydrogen or deuterium atoms diffract neutrons relatively more strongly, but in addition protein crystals suffer no radiation damage in a neutron beam. These and other differences are being exploited for a few selected proteins, at three reactors.

At the Institut Laue–Langevin, Grenoble, the crystal structure of the triclinic form of hen egg-white lysozyme is being refined at high resolution (*d*-spacings down to 1.4 Å (0.14 nm)), from neutron diffraction measurements on a partly deuterated native crystal of volume 20 mm³. Results at the present stage of refinement are discussed. Prospects for neutron protein crystallography are examined in the light of progress in X-ray protein crystallography and neutron detector technology.

Introduction

The many successes of X-ray crystallographic studies of large molecules such as proteins, nucleic acids and viruses are too well known to need exposition here. It is sufficient to recall the pioneering work on myoglobin, haemoglobin, lysozyme or insulin, culminating in the 1960s and 1970s in high-resolution structural models of the molecules and their surroundings in the crystals. It seems particularly appropriate to refer specifically to a Discussion Meeting on the structure and function of lysozyme, organized on behalf of the Royal Society 14 years ago by M. F. Perutz, since lysozyme was the first enzyme for which the crystal structure was determined at high resolution (Phillips 1966; Blake *et al.* 1967*a*, *b*). Imoto *et al.* (1972) have reviewed in detail the structure and properties of lysozymes.

Further X-ray crystallographic work is in progress on the lysozyme structures. High-resolution refinements of the structures of tetragonal hen egg-white lysozyme and human lysozyme are near completion. A wealth of information should soon be available about the water structure around the lysozyme molecules, and about the flexibility of different parts of the protein molecules (Artymiuk *et al.* 1979). Of most direct relevance to the mechanism of action of lysozyme is the planned study of substrate complexes of a tortoise lysozyme (C. C. F. Blake, personal communication), since this lysozyme is apparently unique in crystallizing with the entire active cleft region exposed to solvent. Thus for the first time one should be able to study, by crystallogaphic methods, several steps in the pathway of reaction of lysozyme with a complete substrate analogue. Such an experiment would provide the best possible data for comparison with theoretical studies on the energetics of enzyme reactions.

The present paper considers progress in our study of native triclinic lysozyme by neutron diffraction, and the potential of this technique in protein crystallography. We have deliberately kept the discussion of our own work as brief as possible since the refinement is still in progress and has been considered elsewhere (Bentley *et al.* 1979).

Neutron diffraction and lysozyme

That protein crystals may usefully be studied by neutron diffraction was first shown for myoglobin (Schoenborn 1969; Norvell *et al.* 1975). Whereas X-ray crystallography is most convenient for definition of the protein non-hydrogen atoms, neutron diffraction can give much more information about the positions and orientations of water molecules and salt ions, as well as confirming the positions of hydrogen atoms and showing which hydrogens are exchangeable with deuterium.

Table 1. Some parameters of triclinic lysozyme

hen egg-white lysozyme, 129 residues
crystallized from 2% $NaNO_3$ 1% lysozyme, 0.025 M acetate buffer, pH 4.6, by seeding
triclinic, space-group P1, $Z = 1$
unit-cell volume 26500 Å^3, molec. mass 20200
one lysozyme molecule (molec. mass 14600), 300 H_2O (27% water)
crystal volume, up to 20 mm^3
partial deuteration at pD 4.6, by soaking

Table 2. Neutron measurements on lysozyme

instrument D8, 4-circle neutron diffractometer
single detector, high-pressure ^{3}He
wavelength 1.68 Å, flux 2×10^8 n cm^{-2} s^{-1} ($\lambda/2$ component about 4.8% at detector)
31-step omega scans, small detector aperture

deuterated crystal

$d > 1.7$ Å, 1–3 min per reflexion
1.7 Å $\geqslant d \geqslant$ 1.4 Å, 3–5 min per reflexion
14735 of 17768 unique reflexions have $I > 2\sigma(I)$
crystal volume 20 mm^3, $T = 23$ °C

native crystal (i.e. not deuterated)

$d > 1.8$ Å, 1–3 min per reflexion
crystal volume 20 mm^3, T = 5 °C

Table 3. Neutron and X-ray scattering factors (femtometres) (Bacon 1975)

element	neutron	X-ray (sin $\theta = 0$)
H	−3.7	2.8
D	6.7	2.8
C	6.7	16.9
N	9.4	19.7
O	5.8	22.5
Na	3.6	30.9
S	2.8	45.0

At the Institut Laue–Langevin reactor in Grenoble, we have made neutron diffraction measurements on triclinic lysozyme, for which we give some pertinent parameters in tables 1 and 2. Further details are available elsewhere (Bentley *et al.* 1979). It is worth noting that for the partly deuterated lysozyme crystals, only about 3½ months of neutron time were needed to record, with moderately good statistical precision, most of the reflexion data with *d*-spacings larger than 1.4 Å†.

Thus triclinic lysozyme is ideal for high-resolution neutron work, although many other factors combined to dictate the choice of this system (Bentley *et al.* 1979); with it we should be able to show convincingly how much new information can in practice come from neutron diffraction.

† 1 Å = 10^{-10} m = 10^{-1} nm.

For proteins, the fundamental advantage of neutrons over X-rays is that the neutron scattering length of hydrogen is about 60% of those of oxygen and carbon, and that of deuterium is about the same as that of carbon (see table 3). Thus hydrogen, and especially deuterium, are relatively much stronger scatterers of neutrons than of X-rays. For this reason, hydrogen and deuterium atoms are visible in properly phased neutron Fourier maps, provided their temperature factors are not too high.

Further, neutrons do not damage protein crystals, so that only one crystal is needed for the entire experiment. This, and the fact that measurements at low temperatures are routine in neutron diffraction, could be of great advantage for experiments such as that on tortoise lysozyme referred to earlier.

Neutron refinement of the triclinic lysozyme structure

Our neutron refinement based at present only on the data from deuterated crystals, started from a structural model determined from X-ray diffraction data by Jensen and coworkers (Kurachi *et al.* 1976; L. H. Jensen, personal communication). After a first attempt to introduce most of the hydrogen atoms at calculated positions, we tried the more cautious approach of starting from the non-hydrogen protein atoms as found in the X-ray work, and gradually including hydrogen or deuterium atoms, and also solvent atoms, in the model. The 1730 atoms placed so far include 165 solvent atoms and 623 hydrogen or deuterium atoms. The agreement factor after idealization of geometry is

$$R = \Sigma(|F_o| - |F_c|)/\Sigma\,|F_o| = 0.28$$

for data with *d*-spacings greater than 1.5 Å. Here $|F_o|$ and $|F_c|$ are the observed and calculated neutron structure amplitudes.

Our model is far from complete, and in particular we have made no deliberate attempt to fit complete D_2O molecules to the solvent density. Such fitting will certainly be possible for some of the solvent regions as found also for myoglobin (Schoenborn & Hanson 1979) and trypsin (Kossiakoff & Spencer 1979).

As has been pointed out by Finney (1979) in a comprehensive review of water structure and proteins, it is easy to over-interpret apparent electron density in the solvent regions around protein molecules in crystals. On the other hand, there is no fundamental reason why, in favourable cases such as triclinic lysozyme, an objective description in conventional crystallographic terms (positional and thermal parameters, and occupation factors to allow for disorder and in the neutron case for exchange with deuterium) should not be given of a large proportion of the solvent regions. Serious attempts are now being made to find criteria for the reality of these water molecules (see, for example, Watenpaugh *et al.* 1978) and neutron diffraction can make a unique contribution (Schoenborn & Hanson 1979).

For triclinic lysozyme we expect that direct comparisons of the nuclear scattering density in the deuterated crystal and in the non-deuterated crystal (for which the 1.8 Å resolution data are now being analysed), with the electron density from the X-ray work of Jensen and coworkers, will greatly help our interpretation of the solvent regions. Another approach to examining apparent water structure is to compare the relative positions of solvent atoms in different crystal forms of the same or similar proteins. A beginning has been made for lysozyme

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

in the tetragonal and triclinic crystals (Moult *et al.* 1976), but this approach needs to be followed up by comparing the many high-resolution maps of various similar lysozymes now available.

In view of the current interest in the meaning of the apparent temperature factors in lysozyme (Artymiuk *et al.* 1979), we note in passing that in the regions of our triclinic lysozyme model that have refined well there is reasonable qualitative agreement with the thermal parameters in tetragonal hen egg-white lysozyme and in human lysozyme (Artymiuk *et al.* 1979). Thus, for example residues 13–23, 47 and 70–75 have relatively high temperature factors. On average, the temperature factors in the triclinic crystals are somewhat lower (overall B about 9.5 Å^2) than those in the other lysozymes, and further the packing of molecules in the triclinic case is quite different.

Good correspondence of the temperature factor curves, if confirmed at a later stage in our refinement, would therefore add further weight to the hypothesis that molecular flexibility is relevant to the biological role of proteins.

The complementarity of X-ray and neutron diffraction

Many man-years of effort have gone into various approaches to data collection and crystallographic analysis and refinement in the neutron diffraction studies of myoglobin. The calculations were based on the original X-ray model established over 10 years ago when refinement techniques were relatively unsophisticated. It is certain that much tedious manual interpretation of Fourier maps could in the future be avoided by using the better quality X-ray models now available.

For the neutron work on ribonuclease A (Wlodawer 1979) it has been found necessary to record new X-ray data and further refine the X-ray structure to obtain an adequate starting model for the neutron refinement.

New work on trypsin now in progress at Brookhaven appears to show that with a highly refined X-ray model, even neutron data of modest resolution may help to answer some important biological questions. The analysis, based so far on neutron data to 2.2 Å resolution, has already led to determination of the correct orientations of the amide side chains that are well ordered, as well as showing unambiguously the position of an important proton in the active site (Kossiakoff & Spencer 1979; A. A. Kossiakoff, personal communication).

Taken together with our own experience in refining the lysozyme structure, all of the above experiments suggest that to expedite neutron refinements one should use to the maximum the X-ray diffraction data, which are both cheaper and inherently simpler to interpret since the hydrogen atoms are of minor importance in the refinement. A good starting model for the refinement is even more important in work with neutron diffraction since it is necessary to begin with well placed hydrogen atoms; the ratio of intensity observations to parameters refined is thus reduced by about a half. For example, considering the protein alone, this ratio is 4.5 for the X-ray refinement and 2.3 for the neutron refinement with data to 1.4 Å d-spacing.

Biological molecules at high resolution

In low to medium resolution studies of proteins (say $d > 2$ Å) it is quite normal to omit hydrogen atoms from the protein model in work with X-ray diffraction, but for neutrons this is unsatisfactory since hydrogen and deuterium atoms make up more than half of the unit cell

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

contents. At medium to high resolution ($d \approx 1.5$ Å), it is advisable to include protein hydrogen atoms in the X-ray model, at predicted positions; and it is imperative in the neutron model if the neutron intensity data are to be explained.

At very high resolution, $d < 1$ Å (i.e. $\lambda^{-1} \sin\theta > 0.5$ Å^{-1}, precise crystallographic studies of small molecules have shown that the X-ray model must be much more complex than the neutron model: in both cases, all identifiable nuclear or atomic sites (including hydrogen) are assigned the usual positional parameters, an occupation factor, and at least isotropic thermal parameters. But whereas the bonded electron density (that found in the so-called X–N difference maps) need not be described at all in the neutron model, it must be included in any model that is to predict accurately the X-ray intensities. Conversely, it is possible to determine bonding electron density only from the X-ray data; neutron data are then very useful to give unbiased estimates of the nuclear positions and thermal vibration parameters.

In view of the fact that many protein hydrogen atoms are already visible in high-resolution X-ray Fourier difference maps, e.g. insulin at 1.2 Å resolution (Sakabe *et al.* 1978), it is to be expected that in the next few years it will be possible to make joint refinements based on X-ray and neutron data, as is already done for smaller molecules (Duckworth *et al.* 1969).

Further, bonding electron density should be observable in highly ordered regions of small proteins. Direct observation of shifts of bonding electron density consequent on, e.g. oxidation or binding of different substrate molecules, would be the most impressive confirmation of hypotheses advanced to explain the biological interactions of proteins.

In parallel with such very time-consuming studies, it is important to study by neutron diffraction smaller molecules of biological importance, with their water of crystallization. Neutron studies of a vitamin B_{12} monocarboxylic acid (D. C. Hodgkin & B. T. M. Willis, personal communication), of vitamin B_{12} (J. L. Finney, personal communication) both at about 1 Å resolution, and of an α-cyclodextrin hydrate (Saenger 1979) are already well advanced. Certain β-cyclodextrin complexes, which may serve as models of enzyme action or of membrane diffusion transport are very interesting candidates for neutron diffraction, particularly as several structures have already been determined carefully at about −120 °C by X-ray diffraction (Stezowski *et al.* 1978). For completeness we mention that the most accurate structural information available on biological molecules comes from the many neutron studies of the individual amino acids, mostly at Brookhaven or Trombay. Neutron structures of three dipeptides and of the nucleotide uridine-5′-phosphate are also available (Sequeira 1979).

Position-sensitive detectors for neutrons

High-resolution data collection for proteins larger than lysozyme or cell dimensions larger than 30 Å would require a year or more of beam time with a single detector. In practice, such large amounts of beam time are not available in open competition. For this reason, neutron area detector systems suitable for protein diffraction spectra are being constructed.

Several such systems operate already. Data from CO sperm-whale myoglobin have been extended from 1.8 to 1.5 Å resolution (Schoenborn & Hanson 1979) by using a 20 cm × 20 cm area detector system with spatial resolution about 3 mm (Schoenborn *et al.* 1978). A segmented linear detector system (Cain *et al.* 1976) has given good-quality data to 2.2 Å resolution from a 1.6 mm^3 crystal – small by neutron standards – of an inhibited trypsin (Kossiakoff & Spencer 1979).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A single linear position-sensitive detector has been used at a medium-flux reactor (Wlodawer 1979) to record reflexions to 2 Å resolution for ribonuclease A. At the I.L.L., Grenoble, a curved area detector to intercept the diffraction pattern over a solid angle of $4 \times 64°$ is under construction. This detector will have 16 vertical wires and 512 horizontal wires to give a resolution of 0.25° horizontally and 0.125° vertically.

Experience with such detectors, and perhaps also with detectors for time-of-flight modified Laue diffractometers at neutron spallation sources, will revolutionize neutron crystallography. Indeed, a careful theoretical study has already shown that with optimization of experimental parameters, high-resolution neutron data on medium-sized proteins might be obtained in a matter of days (Jauch 1979).

The power of X-ray protein crystallography has increased dramatically in the last few years with the availability of synchrotron sources, area detectors and computer programs for least-squares refinement of very large molecules. Thus armed with very good X-ray starting models and high-resolution neutron data, protein crystallographers will be able to learn much about protein structure that could not be learned without neutrons.

We thank E. Duee and A. C. Nunes for their work on triclinic lysozyme and L. H. Jensen for providing the lysozyme X-ray coordinates.

References (Bentley & Mason)

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C. & Sternberg, M. J. E. 1979 *Nature, Lond.* **280**, 563–568.
Bacon, G. E. 1975 In *Neutron diffraction*, p. 39. Oxford University Press.
Bentley, G. A., Duee, E. D., Mason, S. A. & Nunes, A. C. 1979 *J. Chim. phys.* **76**, 817–821.
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. 1967*a* *Proc. R. Soc. Lond.* B **167**, 378–388.
Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. M. 1967*b* *Proc. R. Soc. Lond.* B **167**, 365–377.
Cain, J. E., Norvell, J. C. & Schoenborn, B. P. 1976 *Brookhaven Symp. Biol.* **27** (5), 111–143.
Duckworth, J. A. K., Willis, B. T. M. & Pawley, G. S. 1969 *Acta crystallogr.* A **25**, 482–484.
Finney, J. L. 1979 In *Water: a comprehensive treatise*, vol. 6 (ed. F. Franks), pp. 47–122. New York: Plenum Press.
Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. 1972 In *The enzymes*, vol. 7, pp. 665–868. New York: American Press.
Jauch, W. 1979 Ph.D. thesis, Technical University of Berlin.
Kossiakoff, A. A. & Spencer, S. A. 1979 In *Abstracts, Summer Meeting, Am. Crystallogr. Assoc., Boston University, Boston, Mass.*, p. 28.
Kurachi, K., Sieker, L. C. & Jensen, L. H. 1976 *J. molec. Biol.* **101**, 11–24.
Moult, J., Yonath, A., Traub, W., Smilansky, A., Podjarny, A., Rabinovich, D. & Saya, A. 1976 *J. molec. Biol.* **100**, 179–195.
Norvell, J. C., Nunes, A. C. & Schoenborn, B. P. 1975 *Science, N.Y.* **190**, 568–570.
Phillips, D. C. 1966 *Scient. Am.* **215** (5), 78–90.
Saenger, W. 1979 *Nature, Lond.* **279**, 343–344.
Sakabe, N., Sakabe, K. & Sasaki, K. 1978 In *Proc. Symp. on Proinsulin, Insulin and C-peptide*, Tökushima, July 1978.
Schoenborn, B. P. 1969 *Nature, Lond.* **224**, 143–146.
Schoenborn, B. P., Alberi, J., Saxena, A. M. & Fischer, J. 1978 *J. appl. Crystallogr.* **11**, 455–466.
Schoenborn, B. P. & Hanson, J. C. 1979 In *A. C. S. Symposium on Water Structure* (ed. J. A. Rupley & I. D. Kuntz). (In the press.)
Sequeira, A. 1979 In *Proc. International Symp. Biomolecular Structure, Conformation and Evolution*, Madras, January 1978 (ed. R. Srinivasan). Oxford: Pergamon.
Stezowski, J. J., Jogun, K. H., Eckle, E. & Bartels, K. 1978 *Nature, Lond.* **274**, 617–619.
Watenpaugh, K. D., Margulis, T. M., Sieker, L. C. & Jensen, L. H. 1978 *J. molec. Biol.* **122**, 175–190.
Wlodawer, A. 1979 *Acta crystallogr.* (Submitted.)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES