ADVANTAGES OF NEUTRON SCATTERING FOR BIOLOGICAL STRUCTURE ANALYSIS

Benno P. Schoenborn

Biology Department

Brookhaven National Laboratory

Upton, New York 11973

Work performed under the auspices of the U.S. Energy Research and Development Administration.
INTRODUCTION

Neutron scattering experiments for the analysis of biological structures is a useful addition to the numerous physical techniques used to unravel the complexities of living things since they allow the determination of structural features not revealed by x-ray methods (1,2,3). The theories of neutron (3) and x-ray (4) scattering are rather similar; the main difference arising from the fact that x-rays are scattered by the electron cloud while neutrons are scattered by the atomic nucleus. For x-rays, the scattering magnitude is proportional to the electron density and the scattering factor, which relates the magnitude of the scattered wave to the incident wave amplitude for a given atom, increases with atomic number. For neutrons the magnitude of the scattering factor depends on the physical size of the nucleus and on the permissible nuclear energy changes caused by the incident neutron while it is momentarily bound to the target nuclei. These energy changes cause resonance effects which vary the scattering factors seemingly haphazardly from atom to atom and from isotope to isotope even within the same element (3,5). This resonance scattering effect is seen for the hydrogen isotopes with scattering factor values of $b_H = -0.38 \cdot 10^{-12}$ cm and $b_D = 0.65 \cdot 10^{-12}$ cm. The negative scattering amplitude of hydrogen and other elements like titanium and manganese is due to an additional resonance phase change of $\pi$ that is not observed for most other elements. Generally, however, neutron scattering factors are all of the same order of magnitude and do not show the large variation observed for x-rays (Table 1).

This narrow range of neutron scattering factors prevents the use of the "heavy atom" phasing method, but fortunately there are a number of isotopes
that exhibit strong anomalous scattering behavior which can be used to determine phases (6). The magnitudes of the scattering factors of hydrogen and deuterium and the large difference between them are of particular importance since they are responsible for the large scattering contrasts between the individual components of complex biological molecules that make neutron scattering experiments so powerful. Due to the negative scattering factor, hydrogen will appear as negative densities on a Fourier map with a magnitude sufficient for identification. Apart from this useful aspect, hydrogen atoms unfortunately also exhibit large incoherent scattering that is due to the unpaired spin of the hydrogen nucleus. This incoherent radiation does not contain any structural information but adds considerably to the background. To reduce this background, it is often desirable to reduce the number of hydrogen atoms present by replacing them with deuterium. Apart from the differences in scattering factors, neutrons exhibit a number of other physical properties important to the molecular biologist. The electrically neutral neutron does not produce free radicals as x-rays do and therefore does not cause significant radiation damage, thus enabling the collection of all diffraction data from one sample. Unfortunately though, neutron beams have relatively low flux and even the most intense neutron beams are about $10^5$ less intense than common x-ray sources. This handicap can, however, be partly overcome by the use of larger specimens and longer exposures. Larger samples cause fewer problems since absorption factors for neutrons are considerably smaller than those found for x-rays (Table 1).
PROTEIN CRYSTALLOGRAPHY

The main advantage of using neutrons instead of x-rays in the analysis of proteins is the ability to determine hydrogen atom position. Hydrogen accounts for about 1/3 of all atoms of a protein. They are intimately involved in enzymatic reactions and play a large role in maintaining the 3D structure of the protein (8). Exchangeable hydrogen positions can be determined by crystallizing the protein in D$_2$O. Exchangeable hydrogens will thus be replaced with deuterium and then indicated by positive instead of negative features on the Fourier map. Particular molecular groups can be titrated by changing the pH, thus establishing local charge and reactivity parameters. The orientation of groups such as histidine can be established since the scattering factor of nitrogen ($9.4 \cdot 10^{-13}$ cm) is significantly larger than that for either O or C; amide groups are therefore also distinguishable from carboxyl groups.

Neutron diffraction analysis is particularly useful in the determination of the location of water of hydration. For proteins crystallized in D$_2$O, scattering by the water molecule is directly comparable with the scattering from the protein itself, providing more clearly defined Fourier peaks than observable in x-ray maps. Experiments in H$_2$O and in D$_2$O further enhance contrast in the map, enabling detailed study of the water of hydration as well as of exchangeable hydrogen positions. A difference map calculated from data with a protein in H$_2$O and then in D$_2$O results in localizing hydrogen (deuterium) atoms only. This approach is particularly suitable in distinguishing water from other small molecules since changes in Fourier density features can be assessed in terms of atomic compositions.
Since no crystal damage is expected as observed in myoglobin studies, neutron crystallography dispenses with cumbersome crystal to crystal data scaling and should yield information on individual crystal differences as well as the type and location of radiation damage caused by x-rays. Phasing of neutron protein data is best done by using the x-ray structure (9). Phases are calculated from the x-ray derived structure without postulating H atom positions, e.g. phases are calculated from C, N and O atoms only since this reduces any bias in possible H atom location. The resultant Fourier map will depict H and D atom locations at approximately 1/2 weight. It can be easily demonstrated by model calculations that neutron phases calculated from a protein with and without H atoms exhibit an average deviation of only ~3° (9). If no x-ray data are available, phases can be determined by using anomalous scattering from isotopes such as samarium, gadolinium or cadmium. Even heavy atom type phasing is possible if molecules rich in hydrogen/deuterium are used (neopentane, etc.). Refinement techniques using real space or least squares techniques interspersed with new phase calculations based on all located atoms do improve the map and enable the location of all atoms. Present experience is, however, rather limited but the author is confident that significant improvements in the understanding of protein structures and function will result from such neutron studies.

SCATTERING FROM ORIENTED SYSTEMS

Problems in this category deal with the structural motive of larger assemblies such as membranes, muscle, collagen, etc. In these cases, where the location of subfragments within a matrix is wanted, it is helpful to consider the average scattering densities of the major components (Table 2). The data in
Table 2 lists such group scattering factors. These scattering densities vary by nearly three orders of magnitude compared to a factor of two for the equivalent x-ray scattering densities. This illustrates why neutrons are ideally suited for such structural investigations. It also should be noted that deuteration of any component drastically alters the scattering density while relative contrast variations can be achieved by altering the aqueous environment of these systems by mixing $H_2O$ with $D_2O$.

Significant advances in the elucidation of membrane structures have been made by x-ray diffraction (3,11). However, the phasing of the x-ray diffraction data has often been based on controversial methods (13-15).

The large difference between the neutron scattering of hydrogen and deuterium can be used like the heavy atom of protein crystallography and provides a convenient solution to the phase problem. Since all biological membranes contain water, a simple isomorphous replacement of $H_2O$ with $D_2O$ is sufficient to determine at least a low resolution structure. The exact positions of the water molecules within the membrane will determine the resolution achievable by this simple method (12). The necessary diffraction data is collected for the membrane soaked in $H_2O$ and in $D_2O$ Ringer's solution. The intensities are corrected for absorption and extinction with account taken of the quite different absorption corrections for $H_2O$ and for $D_2O$ due to the large incoherent scattering of hydrogen. The Lorentz function corrected structure factors are then used in a difference Patterson function with $(\Delta F)^2 = (|F|_{H_2O} - |F|_{D_2O})^2$ as coefficients. The difference Patterson function thus contains the information regarding distribution of water.
The above definition of \( (F)^2 \) uses the absolute values of the observed structure factors ignoring possible changes in phase of any diffraction orders. A gradual \( \text{H}_2\text{O}/\text{D}_2\text{O} \) exchange permits, however, direct observation of phase changes of particular reflections. Such a phase change observed for \( h=2 \) in a dipalmitoyl lecithin (DPL) cholesterol sample is shown in Fig. 1.

The presence of exchangeable hydrogen on lipid head groups and in proteins poses a problem by introducing multiple minor sites which add to the difficulty of interpreting the Patterson function, particularly for high resolution structures. To determine phases for a high resolution structural analysis it is therefore best to use a specially deuterated constituent especially if data are obtained for both the H and D versions. Determination of phases by the heavy atom technique is straightforward and follows the methods developed for protein crystallography (16). In the isomorphous replacement method the calculated structure factor \( (f_c) \) is obtained for the "heavy group" from the positions given by the Patterson map interpretation. The "heavy group" scattering factor in this case is the difference between the H and D versions \( (b_D - b_H) \). Since the scaling of \( f_c \) relative to the observed structure factor is often difficult, changes in magnitude and sign for \( F_D \) and \( F_H \) are best detected by following the gradual change in the scattering magnitude of the heavy group, which is easily brought about by mixing the H and D versions of the heavy group (it is particularly easy in the case of \( \text{H}_2\text{O}/\text{D}_2\text{O} \) exchange).

Specific deuteration and relative contrast variations by \( \text{H}_2\text{O}/\text{D}_2\text{O} \) exchange are not only useful for phase determination but can easily be used to locate particular constituents in known structures. This approach should prove valuable in binding site studies of inhibitors, activators and drugs.
SOLUTION SCATTERING

X-ray scattering (17-19) from proteins in solution has long been used to determine general shape parameters but has otherwise been of little help. The large difference in scattering densities for groups in which hydrogen is replaced by deuterium and the ease of adjusting scattering contrast by H₂O/D₂O mixing has opened up unique applications in the analysis of protein shapes (20-23) and quaternary structures of complexes (24-26). The low angle scattering profile \( F(s) \) of macromolecules in solution is determined by the difference in scattering density \( \Delta \rho \) between the solvent and the solute and the Fourier transform of its shape functions \( f(r) \) with \( F(s) = \Delta \rho \int f(r) \exp(2\pi i \mathbf{r} \cdot \mathbf{s}) dV \). As a first approximation Guinier (17) has shown that at very small angles the scattering intensity is a linear function of the scattering angle squared, yielding directly the radius of gyration; measuring the distribution of scattering density.

In particles such as viruses, ribosomes, vesicles, etc., which are composed of different macromolecules (proteins, nucleic acids, lipids, etc.), the determination of the radii of gyration \( R_g \) of individual components will allow conclusions on their relative distribution (20). The parallel axis theorem of classical mechanics shows that the total radius of gyration \( R_T \) of a two particle system is given by \( R_T^2 = \rho_1 R_1^2 + \rho_2 R_2^2 + \rho_1 \rho_2 \Delta^2 \) where \( \rho \) = scattering density and \( \Delta \) = separation between the two mass centers of radii \( R_1 \) and \( R_2 \). The radii of gyration of individual components can be determined by contrast variation. In a two particle system the solvent scattering density is adjusted by varying the H₂O/D₂O concentration to match the average scattering density of one of the
macromolecules in order to determine the $R_g$ of the second particle. The procedure is then reversed to measure $R_g$ for the first particle. To increase scattering contrast and therefore increase the signal to background ratio specific deuteration might be necessary (27).

At larger angles the scattering profile is determined by the general shape and density fluctuations within the sample (20,21). An interesting case arises when an object possesses two centers with scattering densities different from its surrounding (24). The scattering function will then contain an interference term equivalent to the scattering form from diatomic molecules as developed by Debye (28) and described by James (5). 

$$(F(s))^2 = (\rho)^2 \left( \frac{\sin 2\pi sd}{2\pi sd} \right)$$

where $d$ is the separation of the two scattering centers. This simple form applies ideally only to spherical scatterers and departures from the ideal will modify the shape somewhat.

Nevertheless this diatomic type scattering theory can be applied to the elucidation of the quaternary structure of protein complexes (24-26). In this case the solvent is density matched to the complex resulting in a featureless scattering profile. Two components are then replaced by their deuterated counterparts, imposing a $\sin x/x$ type scattering profile on the "background" scattering of the density matched complex in solution. The experiment is then repeated with two other constituents as the focal scatterers eventually yielding by triangulation a location map of the different components of the complex. Further analysis of the scattering transform by spherical harmonics, Patterson technique or simple model calculations will provide information on the shape of the individual particles.
CONCLUSION

While analysis of biological structures by neutron scattering experiments has just begun, the papers in this symposium demonstrate the successful use of neutron scattering techniques to solve many different structural problems. The rapid advancement of this field has been due in part to the ingenious use of old techniques and in part to the development of new experimental techniques, particularly those using two-dimensional counters. Most present experiments are however still hindered somewhat by the relatively low flux of present day neutron sources. While little direct increase in reactor flux can be expected, flux can often be improved by using radiation at the optimal wavelength (λ) for the resolution required using either hot or cold moderations to maximize flux. Improvements in beam geometry using efficient monochromators with a proper Δλ and the use of focussing devices will also often improve the effective flux. Significant gain can still be achieved by the development of still larger high efficiency and high resolution two-dimensional counters. The final degree of success will however always depend on the skill of producing good biological samples specially where selective deuteration is required.
REFERENCES


TABLE I. Cross Sections And Scattering Factors For Elements Occuring In Biological Systems

<table>
<thead>
<tr>
<th>Element</th>
<th>Total</th>
<th>Coherent</th>
<th>Neutron mass absorption coefficient ($\mu/\rho$, cm$^2$/g)</th>
<th>Scattering length $10^{-12}$ cm$^{-1}$</th>
<th>Neutrons (sin $\theta$ = 0)</th>
<th>X rays (sin $\theta$ = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>81.5</td>
<td>1.8</td>
<td>.11</td>
<td>.38</td>
<td>.28</td>
<td>.28</td>
</tr>
<tr>
<td>D</td>
<td>7.6</td>
<td>5.4</td>
<td>.0001</td>
<td>.65</td>
<td>.28</td>
<td>.28</td>
</tr>
<tr>
<td>C</td>
<td>5.5</td>
<td>5.5</td>
<td>.0002</td>
<td>.66</td>
<td>1.7</td>
<td>.7</td>
</tr>
<tr>
<td>N</td>
<td>11.4</td>
<td>11.0</td>
<td>.048</td>
<td>.94</td>
<td>2.0</td>
<td>.7</td>
</tr>
<tr>
<td>O</td>
<td>4.2</td>
<td>4.2</td>
<td>.0</td>
<td>.58</td>
<td>2.3</td>
<td>.7</td>
</tr>
<tr>
<td>Na</td>
<td>3.4</td>
<td>3.6</td>
<td>.008</td>
<td>.36</td>
<td>3.14</td>
<td>.7</td>
</tr>
<tr>
<td>Mg</td>
<td>3.7</td>
<td>3.6</td>
<td>.001</td>
<td>.53</td>
<td>3.4</td>
<td>.7</td>
</tr>
<tr>
<td>P</td>
<td>3.6</td>
<td>3.5</td>
<td>.002</td>
<td>.51</td>
<td>4.2</td>
<td>.7</td>
</tr>
<tr>
<td>S</td>
<td>1.2</td>
<td>1.2</td>
<td>.006</td>
<td>.28</td>
<td>4.5</td>
<td>.7</td>
</tr>
<tr>
<td>Cl</td>
<td>15.2</td>
<td>12.3</td>
<td>.3</td>
<td>.96</td>
<td>4.8</td>
<td>.7</td>
</tr>
<tr>
<td>Ca</td>
<td>3.2</td>
<td>3.0</td>
<td>.004</td>
<td>.46</td>
<td>5.6</td>
<td>.7</td>
</tr>
<tr>
<td>Mn</td>
<td>2.0</td>
<td>1.6</td>
<td>.083</td>
<td>-.36</td>
<td>7.0</td>
<td>.7</td>
</tr>
<tr>
<td>Fe</td>
<td>11.8</td>
<td>11.4</td>
<td>.015</td>
<td>.95</td>
<td>7.3</td>
<td>.7</td>
</tr>
<tr>
<td>Ni</td>
<td>18.0</td>
<td>13.4</td>
<td>.03</td>
<td>1.0</td>
<td>7.9</td>
<td>.7</td>
</tr>
<tr>
<td>Zn</td>
<td>4.2</td>
<td>4.2</td>
<td>.006</td>
<td>.56</td>
<td>8.3</td>
<td>.7</td>
</tr>
<tr>
<td>Ge</td>
<td>9.0</td>
<td>8.8</td>
<td>.01</td>
<td>.84</td>
<td>9.0</td>
<td>.7</td>
</tr>
</tbody>
</table>

TABLE II. The average neutron scattering length per unit volume of some common biological constituents given in their hydrogenated and deuterated form. The exact volumetric scattering densities depend on the actual atomic compositions and densities.

<table>
<thead>
<tr>
<th>Mater</th>
<th>Hydrocarbon</th>
<th>Lipid polar head group</th>
<th>Proteins</th>
<th>RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scattering length in $10^{-14}$ cm per $\text{A}^3$</td>
<td>hydrogen form</td>
<td>deuterium form</td>
<td>hydrogen form</td>
<td>deuterium form</td>
</tr>
<tr>
<td>Water</td>
<td>-0.6</td>
<td>6.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrocarbon</td>
<td>-0.3</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipid polar head group</td>
<td>1.7</td>
<td>2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins</td>
<td>4.0</td>
<td>9.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA</td>
<td>4.2</td>
<td>7.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

FIG. 1. Structure factor amplitudes for multilayers of dipalmitoyl lecithin (DPL) cholesterol at 95% humidity. Structure factors were measured as a function of $\text{H}_2\text{O}/\text{D}_2\text{O}$ concentration for samples with cholesterol with a hydrogenated and a deuterated hydrocarbon chain (d). The reflection orders are indicated. The data has only been corrected for absorption and sample geometry. The data between the hydrogen vs the deuterated sample has been scaled to equal shapes since the structure factor changes due to the $\text{H}_2\text{O}/\text{D}_2\text{O}$ exchange have to be equal for both isomorphous samples.