

## A New Method for the Determination of Biological Quarternary Structure by Neutron Scattering

(pair separations/triangulation/deuterated form/ribosomes/macromolecular aggregates)

DONALD M. ENGELMAN AND PETER B. MOORE

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520

Communicated by Frederic M. Richards, May 19, 1972

**ABSTRACT** Determination of quaternary structure by neutron scattering is proposed. The method gives the identity and relative spatial position of each component of a complex, provided that the complex can be obtained in deuterated form and can be reconstituted from its separated parts. If two components are rich in hydrogen and if the rest of the complex is deuterated, the contrast in scattering power permits the measurement of the separation of the pair from an interference cross term. The structure is obtained by triangulation from a set of measured pair separations.

Biochemists are becoming increasingly adept at reconstituting macromolecular aggregates from their components, as is shown by recent successes with ribosomes from *Escherichia coli* (1-3). Many of these aggregates are so large and complex that little progress in determination of their three-dimensional structures can be expected from traditional techniques such as x-ray diffraction and electron microscopy. We would like to point out the possibility of combining the techniques of neutron diffraction and reconstitution to obtain structural information about macromolecular aggregates.

A most important feature of neutron diffraction in biological systems is that hydrogen is easily distinguishable from all other normally encountered atomic species and from deuterium (see Table 1). (For an introduction to neutron scattering and previous applications to biological systems, see refs. 4-6.) Exploiting this property, Schoenborn (7) has been able to locate the hydrogen atoms in crystalline myoglobin, and Caspar and Kirschner have used the exchange of D<sub>2</sub>O for H<sub>2</sub>O to define the location of water in the myelin sheath (8).

What we are proposing is a logical extension of this prior work. The basic experiment consists of introduction of two H-rich components by reconstitution into an otherwise heavily deuterated structure. The thermal neutron-scattering profile of the object is measured in solution in D<sub>2</sub>O. The neutrons scattered from the two H-rich components interfere to give a sinusoidal ripple whose frequency depends on the separation of the centers of mass of the components. Because of the large difference in thermal neutron-scattering length between hydrogen and deuterium (see Table 1), the ripple is strong enough to be detected in the presence of the other scattering contributions, and the separation of the components is found. By systematic substitution, all the intercomponent distances can be found and a three-dimensional map can be obtained in which each component is explicitly identified. The map is indeterminate only with respect to hand. The advantages of the technique are (a) it can be done in solution without crystals, and (b) it positively identifies and locates each component in the structure. The technique is applicable

to any structure that can be obtained in fully deuterated form and that can be reconstituted from its component parts.

To illustrate the technique, let us consider how a 30S ribosomal subunit would scatter if it were fully deuterated with the exception of two proteins. At low resolution, the proteins and the subunit itself can be approximated as closed volumes within which the scattering length per unit volume ( $\rho$ ) is constant and equal to the average value for the structure in question. The theory of the scattering of such systems is similar to that of polyatomic gases, which is treated in standard texts on x-ray diffraction (9, 10). The scattering profile of a solution of such objects,  $I_t(\mathbf{R})$ , has six contributions, the spherically averaged Fourier transforms of the ribosome and of the two substituents,  $I_r(\mathbf{R})$ ,  $I_1(\mathbf{R})$ , and  $I_2(\mathbf{R})$ , and three cross terms representing the interference between the three regions in question, taken pairwise and spherically averaged,  $C_{r1}$ ,  $C_{r2}$ , and  $C_{12}$ :

$$I_t(\mathbf{R}) = I_r(\mathbf{R}) + I_1(\mathbf{R}) + I_2(\mathbf{R}) + C_{r1}(\mathbf{R}) + C_{r2}(\mathbf{R}) + C_{12}(\mathbf{R})$$

If the scattering of the fully deuterated ribosome,  $I_r(\mathbf{R})$ , is measured as well as the scattering of the two possible singly substituted subunits,  $I_{r1}(\mathbf{R})$  and  $I_{r2}(\mathbf{R})$ , then,

$$I_{r1}(\mathbf{R}) = I_r(\mathbf{R}) + I_1(\mathbf{R}) + C_{r1}(\mathbf{R})$$

$$I_{r2}(\mathbf{R}) = I_r(\mathbf{R}) + I_2(\mathbf{R}) + C_{r2}(\mathbf{R})$$

$C_{12}$  can be isolated from the doubly substituted profile:

$$C_{12}(\mathbf{R}) = I_t(\mathbf{R}) + I_r(\mathbf{R}) - I_{r1}(\mathbf{R}) - I_{r2}(\mathbf{R}).$$

When the two substituents can be approximated as spheres,

$$C_{12}(\mathbf{R}) = 2(\rho_{\text{protein}} - \rho_{\text{ribosome}})^2 S_1(\mathbf{R}) S_2(\mathbf{R}) \sin 2\pi d_{12} \mathbf{R} / 2\pi d_{12} \mathbf{R}$$

TABLE 1. Neutron scattering lengths of predominant nuclei

| Nucleus | $b \times 10^{13}$ cm |
|---------|-----------------------|
| H       | -3.74                 |
| D       | 6.67                  |
| C       | 6.65                  |
| N       | 9.40                  |
| O       | 5.80                  |
| P       | 5.10                  |

Coherent scattering lengths are shown as compiled by Schull, G. G. (recent data prepared at Massachusetts Institute of Technology, Cambridge, Mass., 1971). The phase shift on scattering from H is 360° rather than 180°, and its scattering length is shown as negative.

where  $S_1$  and  $S_2$  are the transforms of the two spheres and  $d_{12}$  is the distance between them. This function has a value of zero every  $n/2d_{12}$ ,  $n = \text{integer}$ , in reciprocal space, from which the intersubunit spacing can be obtained.

It can be shown that this analysis of the low-angle scattering pattern is independent of assumptions about the shape and density distribution of the main ribosome. The substituents should be fairly approximated as spherically symmetric density distributions at the resolution in question. If the substituents are merely centrosymmetric,  $C_{12}$  no longer has the form given above. However, it can be shown numerically that its fundamentally sinusoidal nature is quite insensitive to deviations from spherical symmetry.

From the intercomponent distances, the map of the substituent centers is obtained by triangulation. An important feature of the strategy is that, particularly for structures with many components, there is substantial redundancy in the information. For the 21 proteins of the 30S ribosomal subunit there are 210 pairs, yet the structure can be uniquely described by 74 pairs. This redundancy permits refinement of the structure and also allows the substitution of alternate measurements if a given pair proves experimentally difficult.

We have done a number of computations based on this theory to predict the form of the experimental data. The

TABLE 2. Average scattering length per unit volume for ribosomal components

| a. Components relative to vacuum           |                              |                                  |
|--|------------------------------|----------------------------------|
| Component                                  | $\rho \times 10^{14}$        | $\rho \times 10^{14}$            |
|  | cm/ $\text{\AA}^3$<br>H-rich | cm/ $\text{\AA}^3$<br>Deuterated |
| Ribosomal protein                          | 3.19                         | 7.95                             |
| Ribosomal RNA                              | 4.50                         | 7.36                             |
| 30S Ribosomal subunit                      | 3.90                         | 7.64                             |
| Water                                      | —                            | 6.38                             |
| b. Components relative to D <sub>2</sub> O |                              |                                  |
| Component                                  | $\rho \times 10^{14}$        | $\rho \times 10^{14}$            |
|  | cm/ $\text{\AA}^3$<br>H-rich | cm/ $\text{\AA}^3$<br>Deuterated |
| Ribosomal protein                          | -3.19                        | 1.57                             |
| Ribosomal RNA                              | -1.88                        | 0.98                             |
| 30S Ribosomal subunit                      | -2.48                        | 1.26                             |

The average scattering length per unit volume was calculated

as  $\rho = \frac{N}{M\bar{V}} \sum a_i b_i$ , where  $N$  is the Avogadro number,  $M$  the total molecular weight,  $\bar{V}$  the total partial specific volume,  $b_i$  the scattering length of the  $i$ th nucleus, and  $a_i$  the number of such nuclei. Values for  $b_i$  shown in Table 1 were used. For the ribosomal protein, the aminoacid composition reported by Spahr (11) and the value of  $\bar{V} = 0.74$  that he calculated were used. The base composition A, 25.2%; C, 21.6%; G, 31.5%; U, 21.7% (12) was used for the 16S RNA, and its partial specific volume was taken as  $\bar{V} = 0.55$ . The molecular weight of the 30S subunit is  $0.9 \times 10^6$  (13) and of the 16S RNA is  $0.55 \times 10^6$  (14), so the RNA is 54% of the ribosomal volume, and the average scattering power per unit volume for the ribosome is  $\rho_{\text{ribosome}} = 0.54 \rho_{\text{RNA}} + 0.46 \rho_{\text{protein}}$ . In the case of H-rich components, all H not bonded to C is counted as exchangeable and taken as D. In (a), scattering lengths are given relative to vacuum; in (b) they are shown relative to D<sub>2</sub>O.

average scattering lengths per unit volume used in these calculations are given in Table 2. As the values given in Table 2b show, there is a large difference in scattering length per unit volume between hydrogenated and deuterated components when these are measured relative to D<sub>2</sub>O. This difference leads to the contrast in neutron scattering that makes our technique possible. To generate a comparable level of contrast for the purposes of doing an analogous x-ray experiment, heavy atoms would have to be introduced at a proportion of about one uranium atom for every three amino acids.

In Fig. 1, it was assumed that all three objects were spheres. The main object is a sphere of the same volume as a 30S ribosome, and the two substituents are spheres equivalent to proteins of 25,000 molecular weight, separated from each other by 75  $\text{\AA}$  in the main ribosome. The whole system is in D<sub>2</sub>O. Fig. 1a shows the expected scattering profiles when

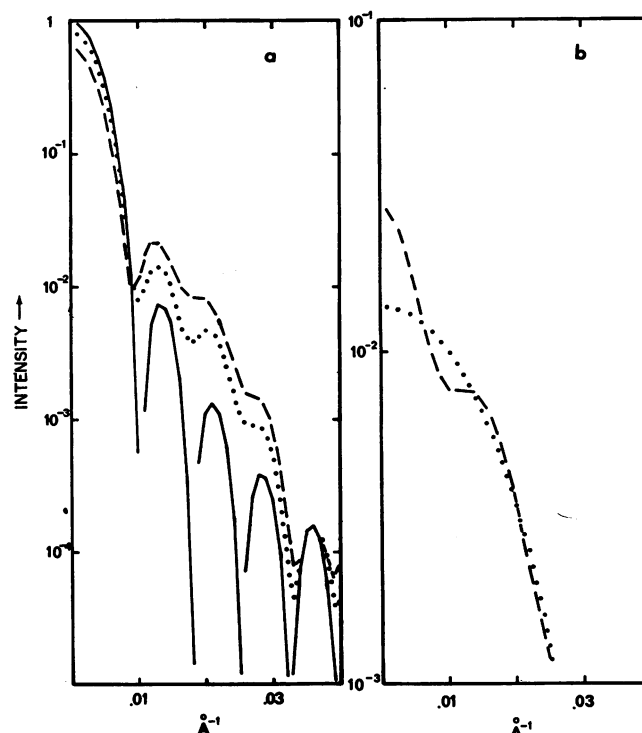


FIG. 1. Computed scattering profiles for 30S ribosomal subunits. The 30S ribosome was approximated as the sphere (radius = 69  $\text{\AA}$ ) whose volume equals that of the "best ellipsoid" found by Hill, Thompson, and Andregg (15) from x-ray scattering data. The substituents were taken as spheres (radius = 20  $\text{\AA}$ ) whose volume is that of an anhydrous protein of molecular weight 25,000. In both calculations it was assumed that the two substituents were located diametrically across from each other, 37.5  $\text{\AA}$  from the center of the 30S sphere. The scattering profiles are calculated for a solution of similar particles in D<sub>2</sub>O. The scattering lengths per unit volume used are given in Table 2b. In the case of the 69- $\text{\AA}$  sphere, the value was adjusted to allow for hydration. Intensities are given in arbitrary units, which are consistently maintained in all figures. (a) Fully deuterated 30S subunit: solid line, the subunit with no substitutions; dotted line, the subunit with one hydrogenated protein; broken line, the subunit with both hydrogenated proteins in place. (b) The 30S subunit whose scattering length per unit volume is adjusted to equal that of D<sub>2</sub>O: broken line, the subunit with both hydrogenated proteins in place; dotted line, the sum of the distributions observed when both substituents are included separately. Note that the scales of intensity differ in Figs. 1a and b.

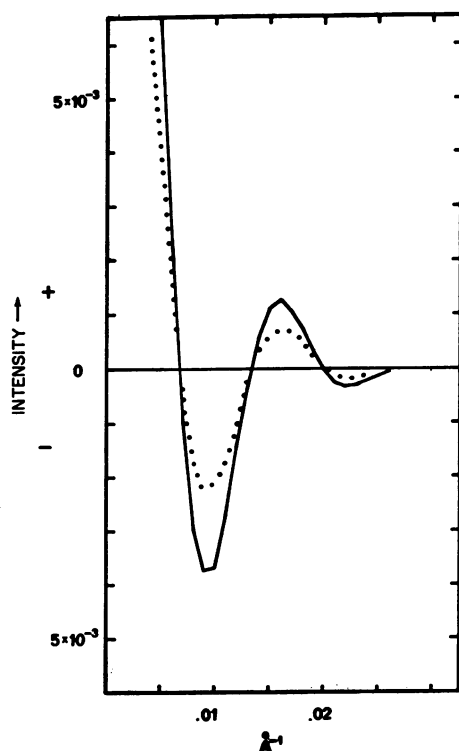


FIG. 2. The cross term,  $C_{12}$ .  $C_{12}$  was calculated from the model computations shown in Fig. 1. The intensity scale used here is linear, not exponential, but the values assigned are consistent with those used in Fig. 1. *Solid line*,  $C_{12}$  from Fig. 1a, no density matching; *dotted line*,  $C_{12}$  from Fig. 1b, density matching of the subunit with the medium.

the ribosome is fully deuterated. Fig. 1b shows what happens when the densities of the main ribosome and the medium are matched. The simplification obtained when densities are matched is obvious; the peaking of the interference ripple can be seen in the total scattering curve. Fig. 2 shows  $C_{12}$  for both cases.  $C_{12}$  has zeros every  $1/150 \text{ \AA}^{-1}$ , as it must. Over much of its range, the magnitude of  $C_{12}$  is greater than 10% of the largest value needed to determine it.

We have measured scattering curves for hydrogenated and deuterated ribosomal subunits with slit-collimated  $4.12\text{-\AA}$  neutrons (beam divergence =  $0.15^\circ$ ). From these preliminary curves, we are able to assess our calculated values for the scattering lengths per unit volume,  $\rho$ ; we find that the experimental ratio of scattering lengths between deuterated and hydrogenated particles is 80% of that expected. Using the x-ray scattering curves of Hill *et al.* (15), we obtain an approximate value for the forward scattering by extrapolation from our data, and we can calculate the expected scattered intensities. A 25-mg sample of deuterated 30S ribosomes (100 mg/ml in  $D_2O$ ) irradiated with  $10^6$  neutrons/min should have a forward scatter of about 6000 counts per hour (cphr) per  $4 \times 10^{-5}$  steradians of detector. Two hydrogenated substituents each of 25,000 molecular weight would contribute about 200 cphr to the forward scatter. The incoherent background, which is flux dependent, would be about 6 cphr; background from other sources should be limited to less than 6 cphr. (Incoherent background can be estimated from incoherent cross sections, atomic abundances in the sample,

and the incident flux, ref. 5.) If density matching were done, the substituent scatter would drop from 200 to 125 cphr. Using 10 detectors in parallel, 1 hr of operation per location measured would give 1250 counts of data in the forward direction against 60 of background. In the model situation used to compute the curves in Figs. 1 and 2, these figures lead to an overall expected standard deviation for  $C_{12}$  in the vicinity of its first zero of about  $\pm 40$  counts ( $\pm 0.7 \times 10^{-3}$  intensity units in Fig. 2), which corresponds to an uncertainty in  $d_{12}$  of  $\pm 4\%$ . Without density matching, the standard deviation of  $C_{12}$  would increase by 50%, but since the magnitude of  $C_{12}$  would also increase, the error in  $d_{12}$  would remain the same. Thus,  $d_{12}$  appears to be measurable in this case, but if the substituents were much smaller than 25,000 daltons, the experiment would be difficult.

In these calculations, we have assumed a conservative value for the incident neutron flux. Using focussing geometry and a liquid-hydrogen moderator, a gain in flux of up to 50 times may be possible at the Brookhaven high-flux beam facility. Higher fluxes would allow one to reduce counting times, work with smaller substituents, smaller samples, or at higher angular resolution (or some combination of these). However, higher fluxes would not allow much reduction of sample concentration, since the ratio of signal to incoherent background would drop proportionately.

It should be emphasized, however, that the ribosome is an extremely difficult object to examine in this way because of the low ratio of its component molecular weights to its total molecular weight. Many other biologically interesting structures (e.g., enzyme complexes) should require far less material and lower neutron fluxes. These objects are well within reach.

Using the high flux beam reactor at the Brookhaven National Laboratory, we have initiated experiments in collaboration with Dr. B. P. Schoenborn to develop and apply the necessary low-angle scattering techniques.

We thank Dr. B. P. Schoenborn for his help with the initial scattering experiments and Mrs. B. Rennie and Mr. M. Kottler for technical assistance. This work was supported by grants from the National Institutes of Health.

1. Traub, P. & Nomura, M. (1968) *Proc. Nat. Acad. Sci. USA* **59**, 777-784.
2. Nomura, M. & Erdmann, V. A. (1970) *Nature* **228**, 744-748.
3. Maruta, H., Tsuchiya, T. & Mizuno, D. (1971) *J. Mol. Biol.* **61**, 123-134.
4. Bacon, G. E. (1955) *Neutron Diffraction* (Clarendon Press, Oxford).
5. Egelstaff, P. E. (ed.) (1965) *Thermal Neutron Scattering* (Academic Press, New York).
6. Schoenborn, B. P., Nunes, A. C. & Nathans, R. (1970) *Ber. Bunsenges. Phys. Chem.* **74**, 1202-1207.
7. Schoenborn, B. P. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36**, 569-575.
8. Kirschner, D. A. & Caspar, D. L. D., *Proc. N.Y. Acad. Sci.*, in press.
9. James, R. W. (1967) *The Optical Principles of the Diffraction of X-Rays* (G. Bell & Sons, London).
10. Guinier, A. (1963) *X-Ray Diffraction* (Freeman and Co., San Francisco).
11. Spahr, P. F. (1962) *J. Mol. Biol.* **4**, 395-406.
12. Spahr, P. F. & Tissieres, A. (1959) *J. Mol. Biol.* **1**, 237-259.
13. Hill, W. E., Rosetti, G. P. & Van Holde, K. E. (1969) *J. Mol. Biol.* **44**, 263-277.
14. Kurland, C. G. (1960) *J. Mol. Biol.* **2**, 83-91.
15. Hill, W. E., Thompson, J. D. & Andregg, J. W. (1969) *J. Mol. Biol.* **44**, 89-102.