

ribosomal 30 S subparticle in its lateral view. The amount of the particles represented in Fig. 11b is about 20% of the total amount of particles. This type of particle has a smaller head than the ribosomal 30 S subparticles. About 60% of particles of ribosomal 16 S RNA-protein S4 complex have a less clear subdivision into the head and body. Most of these particles are V-shaped; one shoulder is thicker than the other and often has at its end a bulge resembling the head (Fig. 11a). These structures can be considered as slightly unfolded structures of the type represented in Fig. 11b.

The electron microscopy study of the ribosomal 16 S RNA-protein S4 complex shows its structural similarity to the ribosomal 30 S subparticle. It can be concluded that the ribosomal 16 S RNA in the complex with protein S4 is able to form the principal elements of the three-dimensional structure of the ribosomal 30 S subparticle.

### Acknowledgments

The authors express their thanks to Professor A. S. Spirin for interest in the work.

## [49] On the Feasibility and Interpretation of Intersubunit Distance Measurements Using Neutron Scattering

By PETER B. MOORE and DONALD M. ENGELMAN

Several years ago we suggested that neutron solution scattering could be used to measure distances between subunits within macromolecular aggregates such as the bacterial ribosome.<sup>1</sup> The feasibility of such measurements has now been established.<sup>2-6</sup> In this chapter, we provide information that can be used by a biochemist or molecular biologist to decide whether the neutron method might be appropriate for studying his prob-

<sup>1</sup> D. M. Engelman and P. B. Moore, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1997 (1972).

<sup>2</sup> D. M. Engelman, P. B. Moore, and B. P. Schoenborn, *Brookhaven Symp. Biol.* **27**, IV-20 (1975).

<sup>3</sup> P. B. Moore, J. A. Langer, B. P. Schoenborn, and D. M. Engelman, *J. Mol. Biol.* **112**, 199 (1977).

<sup>4</sup> J. A. Langer, D. M. Engelman, and P. B. Moore, *J. Mol. Biol.* **119**, 463 (1978).

<sup>5</sup> D. M. Engelman, P. B. Moore, and B. P. Schoenborn, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3888 (1975).

<sup>6</sup> P. Stöckel, R. May, I. Strell, W. Hoppe, Z. Cejka, H. Heumann, W. Zillig, H. L. Crespi, J. J. Katz, and K. Ibel. Presented at the 4th International Conference on Small Angle Scattering, Gatlinberg, Tennessee, October, 1977.

TABLE I  
NEUTRON SCATTERING LENGTHS OF PREDOMINANT BIOLOGICAL NUCLEI<sup>a</sup>

Nucleus	Scattering lengths <i>b</i> ( $\times 10^{-12}$ cm)
H	-0.374
D	0.667
C	0.665
N	0.94
O	0.580
P	0.51
S	0.28

<sup>a</sup> The scattering lengths are taken from G. E. Bacon, "Neutron Scattering." Oxford Univ. Press (Clarendon), London and New York, 1975.

lem. The basic experimental design is described, and an expression is presented, based on our experience with the ribosome system, that can be used to test the parameters of a proposed experiment to decide on its feasibility. Additionally, the basic elements of the interpretation of experimental data are described to inform the reader about the kinds of information obtainable from measurements of this kind.

Neutrons differ in an important way from other kinds of radiation employed in biological investigations. Light, electrons, and X rays interact almost exclusively with the electrons in a molecule. Neutrons are relatively unaffected by the electrons, but interact with the atomic nuclei. It follows that neutrons can be scattered differently by atoms that represent different isotopes of the same element.<sup>7</sup> Distance-finding measurements rely upon the large difference in neutron scattering properties between hydrogen and deuterium. This difference is described by a quantity called the scattering length, which describes the relative phase of a scattered neutron and is related to the probability that a scattering event will take place. The scattering lengths for the abundant biological nuclei are shown in Table I, and it is apparent that H and D differ both in magnitude and sign. The sign difference means that an H atom will scatter 180° out of phase with D (or the other atoms in Table I) and will thus appear as a negative point in the structure as described by the scattered neutrons. Since the overall scattering properties of a macromolecule are the sum of the contributions of the individual atoms of which it is made, it follows that a deuterium-substituted macromolecule

<sup>7</sup> G. E. Bacon, "Neutron Diffraction." Oxford Univ. Press (Clarendon), London and New York, 1975.

is readily distinguished from its protonated cousin in a small-angle neutron-scattering experiment.<sup>1</sup>

Let us imagine that two deuterated subunits are placed in an otherwise protonated macromolecular aggregate, and that the neutron scattering of the product is measured in solution. The aggregate containing deuterated subunits will have a scattering profile different from that of the same aggregate fully protonated. [The scattering profile is simply the number of neutrons scattered per unit angle, where the scattering angle ( $2\theta$ ) is measured from the direction of the beam passing through the specimen.] Included in the difference will be an interference ripple resulting from the fact that the two deuterated subunits are scattering simultaneously while maintained in a fixed spatial relationship in the aggregate. This ripple is detectable even when the aggregates in the specimen have random orientations, as they do in solution. Its spatial frequency is inversely proportional to the distance between the two deuterated subunits, just as in the analogous two-slit diffraction experiment familiar from elementary physics (Young's experiment). The object of the experiment is to obtain a measurement of the interference cross term. From the measured interference curve, the separation of the centers of a pair of deuterated subunits can be obtained. A series of such distance measurements for pairs of subunits can be combined by triangulation to give a three-dimensional description of the relative positions of all the subunits in the assembly.<sup>1,8</sup> Such measurements could also be used to locate macromolecular ligands, study conformational changes, etc.

### Basic Experimental Design

The manner in which distance-finding experiments are done with neutrons owes much to theoretical considerations that have surrounded attempts to do analogous experiments using X-ray scattering.<sup>8,9</sup> In the case of X rays, a structure is labeled with heavy atoms rather than deuterium, but the principle is the same in either case. In order to obtain the rippling interference cross term describing the relationship of a pair of deuterated subunits, four samples are prepared (see Fig. 1): (1) a sample having the two subunits of interest deuterated; (2) and (3) the two possible samples having only one subunit deuterated; and (4) the macromolecular aggregate having no deuterated subunits. These four samples are combined pairwise in equal amounts as shown in Fig. 1. The scattering profiles of the two mixed samples are measured, and subtrac-

<sup>8</sup> W. Hoppe, *Isr. J. Chem.* **10**, 321 (1972).

<sup>9</sup> W. Hoppe, *J. Mol. Biol.* **78**, 581 (1973).

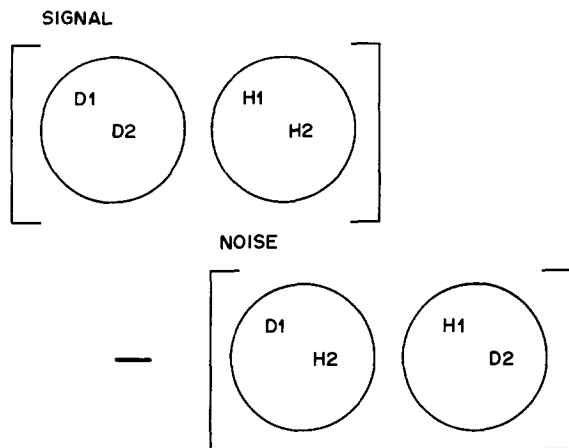


FIG. 1. Design of the basic experiment to obtain the interference function for a protein pair. The desired curve is obtained by measuring the difference between two scattering curves [W. Hoppe, *J. Mol. Biol.* **78**, 581 (1973)]. The first curve (the "signal") is collected on an equimolar mixture of ribosomes containing both of the proteins of a given pair in deuterated form with ribosomes containing the same proteins in the normal hydrogenated form. The second scattering curve ("noise") is collected using a mixture containing equal amounts of ribosomes containing one deuterated protein and ribosomes containing the other deuterated protein. Since both samples have the same amount of deuterated proteins, hydrogenated proteins and RNA, the only difference between them is in the arrangement of the scattering elements. The difference curve,  $I_x(s)$ , arises from the fact that the pair of deuterated proteins is held in a fixed spatial relationship in the "signal" sample.

tion of one from the other yields the desired interference cross term, provided that equal amounts of material were present in both samples.<sup>9</sup> In the case where the amounts of material are unequal, it is possible to correct for differences in sample concentration by including data from a scattering curve measured on a sample of buffer.

For anyone interested in applying these methods to solve biochemical questions, there are three key issues that must be confronted. (1) Is the system in question suitable for such studies? (2) How much effort will it take to get the work done? (3) What is the nature of the information that will be obtained? We will attempt to provide answers to these questions below. We will not provide details about data reduction or the theory of these measurements in this essay. Those who find the answers to the three questions mentioned above sufficiently to their liking, may find the more detailed aspects of the problem discussed in detail elsewhere.<sup>3,4,9-11</sup>

<sup>10</sup> P. B. Moore, J. A. Langer, and D. M. Engelman, *J. Appl. Crystallogr.*, in press (1978).

<sup>11</sup> P. B. Moore and D. M. Engelman, *J. Mol. Biol.* **112**, 228 (1977).

### Test of the Feasibility of an Experimental Measurement

In deciding whether to apply the neutron-scattering method to a given problem, the first issue is whether the measurement could be made, presuming that the biochemical problem of specimen preparation can be solved. In this section we will give a set of criteria intended to provide a test of the feasibility of a given experiment. The process of evaluating the experiment has, in fact, a number of aspects, which include the optimization of the experimental geometry. These more technical aspects are discussed in one of the accompanying articles.

For the purposes of this discussion we must make a number of assumptions. A critical aspect is that intensities should be measured to a sufficiently small angle of scattering so that

$$s = (2 \sin \theta)/\lambda \doteq 1/4d_{12}$$

where  $d_{12}$  is the largest expected value for the separation of subunit centers,  $2\theta$  is the scattering angle,  $\lambda$  is the wavelength, and  $s$  is the reciprocal space coordinate. The interference cross term,  $I_x(s)$ , has its largest value at zero angle. The indicated limit will give values that correspond to  $I_x(1/4d_{12}) \doteq I_x(0)/2$  or perhaps somewhat less. The measurement of the higher  $I_x$  intensities that are present at small angles is crucial to the success of the experiment. Presuming that the measurement of these angles is possible (see this volume [51] for details), and also presuming the use of a two-dimensional position-sensitive detector system, the following criterion can be applied:

$$M_1 X_1 M_2 X_2 E F V C t = K$$

where

$M_i$  = molecular weight of subunit  $i$

$X_i$  = fractional deuteration of the subunit in nonexchangeable locations (1.0 = all  $D$ )

$E$  = detector efficiency (1.0 = fully efficient)

$F$  = neutron flux through specimen in neutrons per second

$V$  = irradiated volume of sample ( $\text{cm}^3$ )

$C$  = Concentration of deuterated subunits in the sample ( $\text{mol cm}^{-3}$ )

$t$  = time required for the measurement (hours)

$K$  = a constant (see below)

Some of the assumptions which are made are:

1. The cell and sample attenuate beam no more than 30%.
2. The deuterated subunits are present at full occupancy in the samples and do not exchange between different aggregates.
3. The buffer scattering density is matched to that of the object.

4. The geometry is optimized for the possible range of distances to be measured.

From our experience with measurements of proteins in the small ribosomal subunit of *Escherichia coli* we can derive a value for the constant  $K$ . The following parameters hold for the experiments we have run at the Brookhaven National Laboratory facility:

$$M_1 \text{ and } M_2 \doteq 17,000 \text{ daltons}$$

$$X_1 = X_2 = 0.85$$

$$E \cdot F = 2 \times 10^5 \text{ neutrons per second through sample}$$

$$V = 6.3 \times 10^{-2} \text{ cm}^3$$

$$C = 3.6 \times 10^{-7} \text{ mol/cm}^3 \text{ of deuterated subunits}$$

$$t = 75 \text{ hr}$$

which gives

$$M_1 X_1 M_2 X_2 E F V C t \doteq 10^8 \text{ sec}^{-1} \text{ mol hr}$$

This expression can be used to evaluate (approximately) a proposed experiment. The most useful way to apply it is to derive an estimate of measurement time, given attainable values of the other parameters. If the answer is that the measurement will require more than  $\sim 200$  hr, it may not be within practical reach. However, several of the elements of the experiment should be examined before a final judgment is made. For example, the measurements can be made on highly concentrated samples, so the possibility of increasing the concentration by precipitation and/or centrifugation should be considered. In the case of the ribosome experiments, ribosome concentrations of  $\sim 25\%$  are routinely used.

Another aspect that might be adjusted to make an otherwise marginal experiment practicable is the collimation of the neutron beam used. The degree of collimation required is directly related to the range of distances one wishes to explore. The shorter the distances, the more relaxed collimation can be, and the more relaxed the collimation, the higher the flux available from a neutron source, everything else being equal. Adjustment of collimation to suit one's experiment might permit an increase in the flux through the sample or in the size of sample used.

The statistical quality of the data we employ, which is reflected in the value given for  $K$ , is, perhaps, higher than might be required for experiments with other complexes. We have noted that the important features of our measured curves can often be obtained with one-half to one-third of the data collection time that we customarily use. In addition, measurements of distances that are large in comparison with the diameters of the subunits give larger excursions of the interference curve from zero and can be made at reduced statistical precision. Compromises made on these aspects of the experiment could reduce its requirements for sample quantity and/or data collection time.

Since nucleic acids have fewer nonexchangeable proton sites on a mass basis than do proteins, the alterations in scattering resulting from deuteration are smaller. Thus,  $K$  must be multiplied by an additional factor of 2 for each of the subunits that consists of nucleic acid.

Thus the expression given above should be regarded as the starting point for deciding whether a given experiment is possible. A positive answer arrived at by its use should be regarded as a strong indication that the experiment in question is possible. A negative answer should not be taken as final until alternative ways of setting up the experiment have been carefully considered. Presuming that the feasibility of the measurement is established, one must next explore the feasibility of the biochemical preparation.

### Biochemical Requirements

It is obvious that there are two fundamental requirements that must be met by an aggregate to permit its study by this method. First, it must be possible to produce the material in massively deuterated form. This means that it must be derivable from an organism that tolerates growth in media containing  $D_2O$  at high concentrations and/or heavily deuterated nutrients. This requirement probably limits the experiment to macromolecular aggregates from bacteria and algae, although suitable material from higher organisms might be obtained using tissue culture cells. [In studies with higher organisms it has often been found that growth ceases in many instances at approximately 30%  $D_2O$ .] Second, it must be possible to reassemble the aggregate from its separated components so that structures having specifically deuterated subunits can be prepared at will.

In the preceding section we outlined the parameters that must be considered in deciding whether a measurement is feasible. These included the molecular weights of the subunits, the degree of deuteration, and the molar concentration of deuterated subunits in the sample. The last of these will involve the overall size of the object, since high molar concentrations will be more easily obtained if the overall size of the aggregate is small. Thus, measurements are facilitated in the case where the subunits are large and form an appreciable fraction of the total mass of the aggregate.

The importance of the level of deuteration obtained from the growth conditions is that it governs the magnitude of the scattering change produced by the labeling, and hence the strength of the observed signal. The number of deuterium atoms in each subunit is represented by the molecular weight and deuteration numbers ( $X_i$ ) in the feasibility equation given above. If both the subunits are deuterated to the same extent, as will usually be the case, the time required for the measurement will vary

inversely as the square of the deuteration. It is therefore imperative that a high level of deuteration be obtained in most cases.

One other biochemical parameter than must be considered is sample stability. Clearly the sample must be capable of remaining in the state one wishes to study for periods of time substantially longer than the duration of a measurement. In the ribosome case, we think in terms of a 2-week period between the time of final preparation of samples for measurement and completion of the measurement process. It is an advantage that radiation damage to the specimen is negligible in these studies; neutron radiation is relatively harmless to macromolecules.

### Data Interpretation

It is easy to show that the physical entity obtained from a solution scattering measurement is a distribution of lengths.<sup>12</sup> In this case it is the distribution of lengths of the vectors joining all deuterium atoms in one subunit to all the deuteriums in the other subunit. This distribution,  $p(r)$ , where  $r$  is the length, is related to the measured interference curve,  $I(s)$ , where  $s = 2 \sin \theta/\lambda$ , in the following manner:

$$p(r) = \int_0^{\infty} sI(s) \sin(2\pi rs) ds$$

where constants of integration are omitted. Thus once  $I(s)$  has been measured,  $p(r)$  can be calculated. A number of methods have been suggested for carrying out this calculation with experimental data.<sup>5,13</sup> An example of an interference curve and a corresponding  $p(r)$  is shown in Fig. 2.

What does  $p(r)$  tell us? Because deuteriums will be well distributed in any macromolecule,  $p(r)$  will be good approximation to the  $p(r)$  one would get if all atoms in the subunit contributed to the interference curve. The smallest  $r$  for which  $p(r)$  has significant values gives an estimate of the shortest distance between the two subunits, and the largest  $r$  an estimate of longest distance. The average distance,  $\bar{r}$ ,

$$\bar{r} = \sum_i p(r_i) r_i \Delta r$$

is very nearly equal to the distance separating the centers of mass of the two proteins.  $[\bar{r}]$  is usually a few angstroms larger than the centers of

<sup>12</sup> A. Guinier and G. Fournet, "Small Angle Scattering of X-Rays." Wiley, New York, 1955.

<sup>13</sup> O. Glatter, *Acta Phys. Austr.* **47**, 83 (1977).

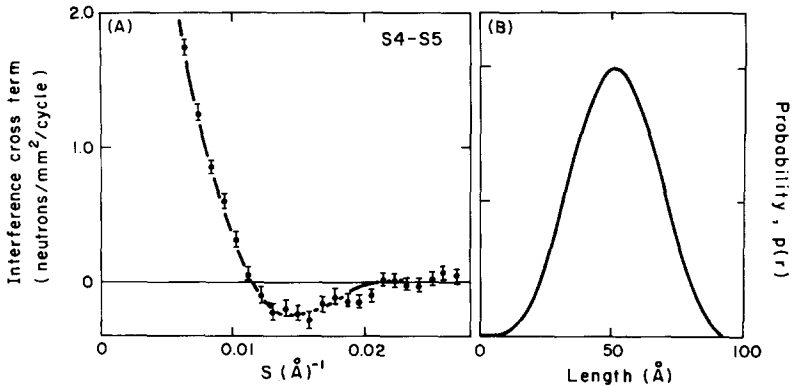


FIG. 2. The interference curve and distribution function for a pair of ribosomal proteins. As an example, we present the interference curve obtained for proteins S4 and S5 in the ribosome (A). Note that the curve is strongly damped and has no values significantly different from zero after its first negative excursion. The first point at which this curve crosses zero can be used to obtain an estimate of the center-to-center separation, since at this point  $s \approx 1/2d_{12}$ . The second curve (B) shows the Fourier inversion of the interference function to obtain the probability distribution,  $p(r)$ , of vector lengths relating the two deuterated subunits. The most probable length is somewhat larger than the center-to-center distance, but model calculations have shown that this difference is small.

mass distance in a structure like the ribosome.<sup>3]</sup> This value can be used in triangulation to obtain the three-dimensional organization of the aggregate. A crude, but simple, alternative way of extracting the separation of centers is simply to use the first point at which the interference curve crosses zero, where the equivalent Bragg spacing is equal to twice the separation.<sup>1</sup>

Another quantity of interest is  $\overline{r^2}$ , the second moment of  $p(r)$ ,

$$\overline{r^2} = \sum_i r_i^2 p(r_i) \Delta r$$

It turns out that  $r^2$  is related to the radius of gyration of the two proteins,  $R_1$  and  $R_2$ , and their center of mass separation,  $d_{12}$ .<sup>10</sup>

$$\overline{r^2} = (R_1^2 + R_2^2 + d_{12}^2)$$

What this means is that if three measurements are done relating three subunits in an aggregate, three such relationships will be found, which can be solved for the radii of gyration of the individual subunits *in situ*, in the aggregate, since the first moments give estimates of the  $d_{ij}$ 's.

The final aspect of the  $p(r)$  curves to consider is their breadth. The full width at half height of these distributions depends critically on both

the relative orientations and axial ratios of the protsins giving rise to them. The latter can be estimated from the second moments (which lead to radii of gyration) and the molecular volume of the subunits, which is estimated from their molecular weights. The breadths then can be used to determine the relative orientation of the axes of the subunits within the structure.<sup>11</sup>

A complete investigation of an aggregate like the ribosome should therefore yield (1) a complete set of distances relating the centers of the subunits in the structure, which gives their relative locations; (2) the radius of gyration of each subunit as it exists in the aggregate; and (3) the relative orientations of the subunits if they are appreciably extended in shape. (1) and (2) are obtained rigorously within the limits set by the precision of the experimental data; (3) will be more inferential, since the interpretation of distribution breadths requires the use of simplifying assumptions to interpret the radii of gyration in terms of subunit shapes. The usual assumption made in this regard is that the subunits can be represented as ellipsoids. The requirements for precision in the measurements are substantially higher for (2) and (3) than for (1), owing to the manner in which counting errors propagate.

We have reviewed straightforward criteria that can be used to test whether the application of the neutron method is likely to succeed and have indicated the kind of information that might follow from a set of experiments. Anyone familiar with biochemical procedures will realize that the difficulty of the biochemical aspects of such experiments is formidable. It follows that work of this kind should not be undertaken unless it is clear that the information obtained will be valuable in the context of a specific problem. It is a somewhat perverse fact that the amount of information obtainable about a structure by this means increases strongly with the number of different subunits it contains, so that the most difficult systems to work with biochemically are eventually the most rewarding.

### Acknowledgments

We wish to acknowledge the help of those who have participated in the work leading to the neutron scattering methods we have described: Drs. B. P. Shoenborn, J. A. Langer, and D. Schindler. This work has been supported by the NSF (BMS 75-03809) and NIH (AI-09167).