

19.4. SMALL-ANGLE NEUTRON SCATTERING

macromolecule plus the surrounding perturbed volume. Both this equation and equation (19.4.2.1) implicitly assume that the medium surrounding each molecule is a vacuum, which is not true for the molecules in solution. The effect of solvent on low-angle scattering can be taken care of by subtracting the average scattering length of the solvent, ρ_0 , from the scattering-length densities within the molecule. Thus,

$$I(Q) \propto I_0 \int \int [\rho(\mathbf{r}_i) - \rho_0][\rho(\mathbf{r}_j) - \rho_0] [\sin(Qr_{ij})/(Qr_{ij})] dV_i dV_j. \quad (19.4.2.3)$$

The quantity $[\rho(\mathbf{r}_i) - \rho_0]$ is a *contrast*, and, as will be shown, contrast manipulation is a major component of SANS experiments.

Equation (19.4.2.3) can be evaluated a second way, because the $(\sin x)/x$ term in the integral depends only on the distances between volume elements, not on their locations in space. Thus, a function $p(r)$ can be defined as follows:

$$p(r) \equiv \int \int [\rho(\mathbf{r}_i) - \rho_0][\rho(\mathbf{r}_i + \mathbf{r}) - \rho_0] dV_i dV_r,$$

where the integral in \mathbf{r} runs over all \mathbf{r} such that $|\mathbf{r}| = r$, and the integral in \mathbf{r}_i runs over the entire molecular volume. Written in terms of $p(r)$, equation (19.4.2.3) becomes

$$I(Q) \propto I_0 \int p(r) [\sin(Qr)/(Qr)] dr, \quad (19.4.2.4)$$

where the integral runs from $r = 0$ to r_{\max} , the maximum atom-to-atom length within the molecule.

Note that if contrast was constant within a macromolecule, $p(r)$ would be proportional to the distribution of interatomic distances in the molecule, and for that reason $p(r)$ is often called the *length distribution*. Note also that $p(r)$ is simply the molecule's Patterson function, rotationally averaged about its origin. Note, finally, that $p(r)$ is the summation of a large, but finite, number of sharp, discrete interatomic distance peaks, each with its own weight. If the individual interatomic peaks in this 'length spectrum' could be assigned, *i.e.*, if the atoms responsible for each one could be identified, it would be possible to determine the three-dimensional structure of the molecule in question, save for uncertainty about its hand.

Since solution-scattering profiles can be computed by sine transformations of length distributions, it is reasonable to hope that a transformation might exist that enables one to compute length distributions once solution-scattering profiles have been measured. There is (Debye & Bueche, 1949; Debye & Pirenne, 1938):

$$p(r) \propto r \int QI(Q) \sin(Qr) dr. \quad (19.4.2.5)$$

Two practical issues must be addressed when carrying out the operation implied by equation (19.4.2.5) because the integral it contains runs from $Q = 0$ to ∞ . Firstly, scattering is never measured at $Q = 0$ due to interference with the direct beam. Secondly, the largest value of Q for which $I(Q)$ is measured is always less than ∞ . The absence of data at very small values of Q is easily addressed, because a soundly based method exists for extrapolating the low-angle data to $Q = 0$ (see below). The lack of data at high Q is harder to cope with, but it can be dealt with approximately using Porod's Law (Porod, 1951, 1952) and the impact of its absence on molecular parameters deduced from small-angle data is easy to estimate. In any case, it is important to realize that length distributions represent the sum total of the information that can be extracted from solution-scattering experiments.

The problem of extrapolating small-angle data to $Q = 0$ was solved by Guinier (1939). He demonstrated that, at very small angles,

$$I(Q) \propto I(0) \exp[-(QR_g)^2/3], \quad (19.4.2.6)$$

where R_g is the radius of gyration, and

$$R_g \equiv (\{\int [\rho(\mathbf{r}) - \rho_0] |\mathbf{r}|^2 dV\} / \{\int [\rho(\mathbf{r}) - \rho_0] dV\})^{1/2}. \quad (19.4.2.7)$$

The origin of the vector \mathbf{r} in this equation is the centre of gravity of the macromolecule's scattering-length density distribution, *i.e.*, it is the point where

$$0 = \{\int [\rho(\mathbf{r}) - \rho_0] \mathbf{r} dV\} / \{\int [\rho(\mathbf{r}) - \rho_0] dV\}.$$

It follows from equation (19.4.2.6) that if the lowest-angle data collected are plotted in the form $\ln[I(Q)]$ versus Q^2 , a straight line should result, the slope of which is $(R_g^2/3)$ and the intercept of which at $Q = 0$ is $I(0)$. Note that data have to be obtained at scattering angles well inside the region where $I(Q) \sim I(0)/2$ in order for this formula to hold; if the data are thus obtained, a radius of gyration estimate will emerge. The radius of gyration of an object is the root-mean-squared distance between its centre of gravity and the elements of which it is composed.

As might be expected, $I(0)$ and R_g can also be computed from $p(r)$. Consider the magnitude of $I(Q)$ at $Q = 0$. Since the $\sin x/x$ term in equation (19.4.2.4) is 1 at $Q = 0$,

$$I(0) \propto I_0 \int \int [\rho(\mathbf{r}_i) - \rho_0][\rho(\mathbf{r}_j) - \rho_0] dV_i dV_j = \int p(r) dr. \quad (19.4.2.8)$$

Thus, $I(0)$, the forward scatter, is proportional to the integral of the length distribution. It is easy to show that R_g equals $(M/2)^{1/2}$, where M is the second moment of $p(r)$ given by

$$M = [\int r^2 p(r) dr] / [\int p(r) dr]. \quad (19.4.2.9)$$

The average atom-to-atom distance in a molecule, r_{ave} , is easy to compute if $p(r)$ is known from

$$r_{\text{ave}} = [\int r p(r) dr] / [\int p(r) dr]. \quad (19.4.2.10)$$

The reason forward scatter, $I(0)$, is interesting is its dependence on molecular weight. As equation (19.4.2.8) suggests, the forward scatter measured for a sample is proportional to N times the square of the product of the average contrast between a molecule and its solvent and the molecular volume, where N is again the number of molecules contributing to the signal observed. Since average contrasts can be estimated from chemical compositions and partial specific volumes, $I(0)$ measurements can be used to estimate molecular weights. If the $I(0)$ values of solutions of a set of molecules of similar chemical composition are compared, it will be found that $I(0)$ divided by the weight concentration of each sample is proportional to molecular weight.

This procedure for estimating molecular weights can fail. Suppose $\int [\rho(\mathbf{r}_i) - \rho_0] dV = 0$, *i.e.*, the scattering-length density of the solvent is the same as the average scattering-length density of the macromolecule. Then $I(0)$ will be zero, the solution-scattering profile will lack a peak at small angles and no molecular-weight estimate will result. Under these conditions, the macromolecule is said to be 'contrast matched'. It is easy to contrast-match biological macromolecules in the context of SANS experiments, because all biological macromolecules that have not been labelled with ^2H have average scattering-length densities between those of H_2O and D_2O (see below).

19.4.3. Contrast variation

19.4.3.1. Variation of solvent density

The principle of contrast variation was studied in early work by Bragg & Perutz (1952), who observed that the magnitudes of low-order reflections in X-ray studies of protein crystals were reduced as the salt concentration in the solvent was raised. Following their concept, the effective scattering density of a dissolved particle is

$$\rho(\mathbf{r}) = \rho(\mathbf{r})_{\text{solute}} - \rho(\mathbf{r})_{\text{solvent}} = \rho(\mathbf{r})_{\text{solute}} - \rho_{\text{solvent}},$$

19. OTHER EXPERIMENTAL TECHNIQUES

where ρ_{solvent} is the average solvent scattering density. If we take a mean scattering density for the particle, a mean contrast, ρ , is defined:

$$\rho = \rho_{\text{solute}} - \rho_{\text{solvent}}. \quad (19.4.3.1)$$

If the solute and solvent have equal densities, they are ‘contrast matched’, and the scattering from the particle vanishes at zero angle. The particle will nonetheless scatter radiation at larger angles as a consequence of internal density fluctuations, which can be useful in obtaining structural information. In the case of neutron scattering, the solvent density is most often varied by mixing D₂O and H₂O to obtain different percentages of D₂O. This strategy gives a range of solvent densities that includes the densities of most biological molecules. However, biological molecules contain hydrogen atoms that exchange with solvent, so deuteration of the environment alters their scattering density to some extent (see below). In general, all hydrogen atoms not bonded to carbon are potentially exchangeable, but not all of these will actually exchange in a typical experiment.

To describe the variation of the radius of gyration with contrast, Stuhrmann derived the useful relationship (Stuhrmann, 1976; Stuhrmann *et al.*, 1976)

$$R_g^2 = R_c^2 + \alpha/\rho + \beta/\rho^2, \quad (19.4.3.2)$$

which separates the contributions of the internal structure of the particle (ρ_I) to its radius of gyration from the contributions of the shape (ρ_c). Scattering from the internal structure is independent of contrast; scattering from the shape is contrast dependent. The shape function is defined as having a value of one inside the particle and zero outside. The total scattering density is then

$$\rho(\mathbf{r}) = \rho\rho_c(\mathbf{r}) + \rho_I(\mathbf{r}). \quad (19.4.3.3)$$

The contrast-independent terms in the Stuhrmann equation are

$$\alpha = (1/V_c) \int \rho_I(\mathbf{r})r^2 d^3r, \quad (19.4.3.4)$$

$$\beta = (1/V_c^2) \int \int \rho_I(\mathbf{r})\rho_I(\mathbf{r}')\mathbf{r}\mathbf{r}' d^3r d^3r' \text{ and} \quad (19.4.3.5)$$

$$R_c^2 = (1/V_c) \int \rho_c(\mathbf{r})r^2 d^3r, \quad (19.4.3.6)$$

where R_c is the radius of gyration of the shape function and V_c is its volume. The sign and magnitude of α give information on the radial density distribution of scattering density in the particle: if the outer region is higher in density than the inner region, α is positive (as, for example, in lipoproteins); if the inner region is denser, α is negative. The β coefficient represents the displacement of the centre of mass as a function of the contrast and is always positive; in real cases, β is often negligible. The Stuhrmann equation leads to a useful way to represent graphically the radius of gyration data obtained from a series of contrasts: the observed R_g^2 is plotted *versus* $1/\rho$. If β is negligible, the plot is a straight line of slope α , intercepting the $1/\rho$ axis at R_c^2 . Thus, R_c is obtained by extrapolation to a point where $\rho = \infty$, and so is often termed the radius of gyration at infinite contrast. This quantity is a representation of the shape of the particle as if it had uniform internal scattering density. In a particle with two discrete regions of density, the radius of gyration for each region can be obtained from such a graph by evaluating R_g^2 where ρ is equal to the density of one region, so that R_g^2 of the non-contrast-matched region is determined. Such measurements can also be made by adjusting the solvent to match the scattering of one region to reveal the scattering of the other.

A parameter that is often useful is the contrast-match point for the particle, which reflects its overall composition including exchange.

$$\rho_M = \sum b_I/V + nd(b_D - b_H)/V, \quad (19.4.3.7)$$

where ρ_M , the match point, is the solvent scattering length density at which the contrast is zero and n is the number of exchanged hydrogens multiplied by d , the fractional deuteration of the water at the match point. Typically, the match point is obtained by measuring small-angle scattering at a series of D₂O:H₂O ratios, plotting each using a Guinier plot $\{\ln[I(Q) - I(0)] \text{ versus } Q^2\}$ to obtain a value for $I(0)$ by extrapolation, and then plotting $[I(0)/C]^{1/2}$ *versus* ρ_{solvent} , where C is the particle concentration. It is often convenient to represent ρ_{solvent} as per cent D₂O. The plot should be a straight line, passing through zero at the contrast-match point. As noted above, the vast majority of biological molecules have contrast-match points at densities between those of H₂O and D₂O. If the particles are compositionally heterogeneous, the observed plot will be a weighted sum of the curves for each of the compositions present and will deviate from a straight line at low contrasts. Thus, the contrast-matching experiment can provide information on both composition and homogeneity.

While contrast variation is most often based on variation of the deuteration level in water, it is also possible to create variation by adding molecules to the solvent. As an example, a study of hydration layers was conducted by adding solute molecules, such as glycerol, at high concentration; the solute molecules alter the solvent scattering length density but do not penetrate the hydration layer (Lehmann & Zaccai, 1984).

19.4.3.2. Variation of internal contrast

A second form of contrast variation can be achieved by replacing the hydrogen atoms in biological molecules with deuterium. If, in addition, it is possible to deuterate selected regions of a particle, internal contrasts can be modified to gain information about substructures. In some cases, the opportunity arises because significant biochemical differences are present, as between the RNA and protein portions of a ribosome, the DNA and protein of a nucleosome, or the lipid and protein of a lipoprotein particle. While the intrinsic contrast may be sufficient to provide key information [as, for example, the early finding that DNA is on the outside of nucleosomes (Bradbury *et al.*, 1976; Uberbacher *et al.*, 1982)], it can be accentuated by incorporation of deuterated biochemical precursors or reconstitution from separately labelled components.

Internal contrast can be created in single molecules by differential incorporation of biosynthetic precursors or by chemical synthesis. Differential incorporation was first used to test models of bacteriorhodopsin, using the incorporation of deuterated amino acids supplied to a culture of halophilic archae (Engelman & Zaccai, 1980), and internal labelling was used to document conformations of cholesteryl esters that had been chemically labelled in key positions (see below). Reconstitution from purified components has been used to place deuterated proteins in ribosomes and in other complexes, again with the aim of creating internal contrast to enhance the information obtained in a neutron experiment.

In general, the creation of internal contrast can be viewed as a strategy for enriching the low information content of a solution-scattering experiment by building additional information into the sample. By design, it is known what has been labelled, so the scattering given by contrasting elements provides information about the relationships of the labelled parts to each other or to the particle as a whole. A particularly informative (but difficult) strategy is to use internal contrast to measure distances between locations in a molecule or complex. Such measurements are discussed below.

19.4.3.3. Relationship of contrasting regions

Where a particle has two regions of different scattering density, the square of the total observed radius of gyration, R_g^2 , can be