Pramāṇa – J. Phys., Vol. 34, No. 4, April 1990, pp. 303–312. © Printed in India.

Computation of small angle neutron scattering functions for molecules of arbitrary shapes

F U AHMED[1], P MARTEL[2] and A M KHAN[3]

[1]Institute of Nuclear Science and Technology, Atomic Energy Research Establishment, P.O. Box 3787, Ramna, Dhaka 1000, Bangladesh
[2]Chalk River Nuclear Laboratories, Chalk River, Ontario KOJ 1JO, Canada
[3]Physics Department, University of Dhaka, Dhaka 1000 Bangladesh and Bangladesh Centre for Advanced Studies

MS received 27 November 1989

Abstract. A computer program has been developed to analyze small angle neutron scattering (SANS) data by using the Debye method of spherical modification proposed by Glatter. In the calculational procedure the model shape is emulated with a large number of overlapping small spheres which fill the volume of the model shape. A technique is described for fitting experimental data to a resolution-broadened model scattering function. At each stage of the iterative procedure the radius of gyration is computed. The program is able to calculate the scattering function of the mixture of two different molecules. This facility even allows one to calculate the scattering function of the mixture of monomer and dimer of a particular molecule in aqueous solution. In case a portion of the molecule has a different weight from the rest, the program has a variation to calculate the scattering function of that model as well.

Keywords. Computation; small angle neutron scattering; arbitrary molecules.

PACS No. 61·20

1. Introduction

The small angle neutron scattering (SANS) technique is nowadays considered a very useful technique in understanding and determining the shape and dimension of complex molecules. Although the method is somewhat analogous to X-ray scattering, there are certain definite advantages this method has over the X-ray techniques e.g. the lower absorption cross section of neutrons than that of X-rays. Moreover thermal neutron scattering amplitudes do not decrease with scattering angles unlike in X-rays. Since the scattering around the centre is particularly independent of the short range order of the atoms, SANS has turned out to be an effective means for studying the dimensions and form of colloidal particles.

The general approach in the study and interpretation of small angle scattering (SANS) data of condensed matter involves the comparison of observed SANS data with some model function (Sjoberg 1978). This comparison is usually done in two ways, either with smeared scattering curves (Sjoberg 1978; Martel *et al* 1980) or with desmeared scattering curves (Glatter 1977, 1980). In either case one has to know the corresponding model functions. While the model functions for a sphere can be calculated analytically, scattering amplitudes of simple triaxial bodies e.g. ellipsoids, prisms and cylinders can be calculated for any specified orientation. However, the spatial average for orientation has to be performed numerically which is time

consuming. Such semianalytical calculations were done by Porod (1948), Mittelbach and Porod (1961) and Mittelbach (1962). Unfortunately, there are a large number of complex biological molecules which cannot be approximated by simple triaxial models.

Glatter (1977) in his paper had presented a powerful technique for analyzing small angle scattering data from monodisperse suspension of molecules of arbitrary shape. In his method, any arbitrary shape is filled with large number of small spheres and the scattering functions are calculated by means of the Debye equation (Debye 1915). However, Glatter had given no prescription for fitting the model calculations to the experimental data by convoluting the theoretical values with the resolution functions. We have therefore combined his calculational procedure with an iterative least squares procedure which determines various parameters by fitting different models to variance-weighted data. The outline of the method and the user guide were briefly reported elsewhere (Ahmed and Martel 1986).

2. Methodology

Small angle scattering is broadened by the finite resolution of the apparatus, because the perfectly collimated X-ray or neutron beams are impossible to obtain in practice. There are two philosophies for analyzing small angle scattering data. According to the Austrian School, as typified by the papers of Glatter (1977, 1980), the experimental data should be deconvoluted to obtain data which is directly comparable with theoretical scattering functions. Sjoberg (1978) and others (Martel *et al* 1980) suggest that the theoretical scattering function should be convoluted with the known experimental resolution for comparison with the data.

The scattering function of a sphere can be calculated analytically and this function is independent of orientation. As mentioned earlier the scattering functions of simple bodies such as ellipsoids, cylinders etc. can be calculated analytically but the orientational average (Sjoberg 1978; Glatter 1980) is necessary for their random orientations. In the present method of calculation the scattering function depends upon the distances between the centres of the small spheres. The following conditions of selecting the sizes of the small sphere have been imposed by Glatter (1980)

$$Q_{\max} < 2\pi/d$$

where d is the distance between two neighbouring spheres. If the spheres are small enough, little error will be incurred because the total volume of the spheres is close to the volume of the true molecular shape.

Sometimes an arbitrary shape may be represented by two or three simple subunits (figure 1). In our calculational method the overlapping of subunits is allowed to approximate the arbitrary molecules whereas in Glatter's calculation such a provision is absent and thus the present method is an improvement over the other. This facility allows us to approximate the arbitrary bodies with small number of subunits. When an arbitrary molecule can be approximated by a combination of three of the simple subunits, sphere, ellipsoids or cylinders, the adjustable parameters are the coordinates of the centres and dimensions of these subunits. As the dimensions of the subunits are adjusted in the filling procedure the program automatically adjusts the appropriate number of small spheres.

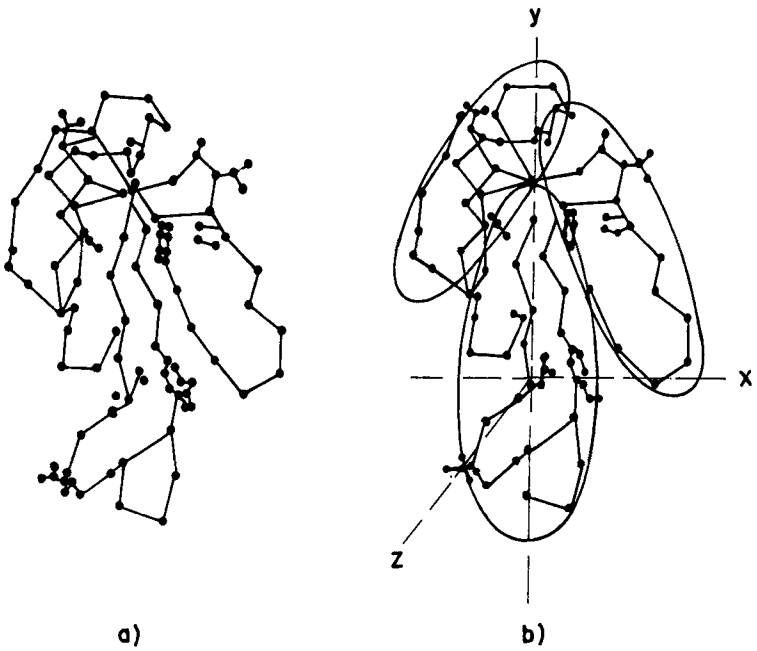

Figure 1. (**a**) Structure of neurotoxin obtained by X-ray diffraction (Wilkinshaw *et al* 1980). (**b**) Our triple ellipsoid approximation.

The scattering function for an arbitrary body having different densities at different portions can be calculated using our program. As the small angle scattering is not sensitive to the small scale structure of a molecule the approximation can be made of homogeneous density within an overall molecular shape. For most large biological molecules this approximation is surprisingly good. For example, in transferring (Martel *et al* 1980) it was possible to determine its molecular weight to $\pm 3\%$.

The program identified as COBRA calculates the convoluted scattering intensity, $J(Q)$, compares it with the experimental data and gives a criterion for the quality of the fit. The language used is Fortran 77. The program currently runs on IBM 4341 at Atomic Energy Research Establishment (AERE), Savar, Dhaka as well as at the University of Dhaka and on CDC Cyber 175 at Chalk River Nuclear Labs. (CRNL), Canada.

3. Mathematical formalism

The intensity of coherently scattered neutrons, $I(Q)$, at wave-vector transfer Q, from an orientational average of an aggregate of n overlapping spheres can be computed by the Debye formula (Porod 1948)

$$I(Q) = \sum_{i=1}^{n} w_i^2 S_i^2(Q) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_i w_j S_i(Q) S_j(Q) \sin(l_{ij}Q)/(l_{ij}Q) \tag{1}$$

where $Q = 4\pi \sin\theta/\lambda$, λ is the wavelength of neutrons, 2θ is the scattering angle and l_{ij} is the distance between the centres of the ith and jth spheres. The structure factor of each small sphere of radius R is

$$S(Q) = 3[\sin(QR) - (RQ)\cos(RQ)]/(RQ)^3. \tag{2}$$

Assuming simple cubic packing with packing fraction $\pi/6$, the radius of the small spheres is given by

$$R = D/2 \cdot (6/\pi)^{1/3} \tag{3}$$

where D is the distance between the centres of two neighbouring small spheres which overlap. Excluding boundary effects, the amount of overlap governed by (3) is just sufficient to ensure that the volume of all the small spheres equals the volume of the molecule. The scattering amplitude of sphere is

$$w = \rho v = \rho(4\pi/3)R^3 \tag{4}$$

where ρ is the neutron scattering length density of the sphere.

To compare with experiment, the intensity of (1) must be convoluted with the experimental resolution function (Mittelbach 1962). This leads to

$$J(Q) = K \iint \{I[(Q - Y)^2 + Z^2]^{1/2}\} W_h(Y) W_v(Z)\, dY\, dZ, \tag{5}$$

where W_h and W_v are normalized Gaussian weighting functions that specify the horizontal and vertical resolution respectively and k is a scale factor. The width of W_h was determined experimentally by scanning the detector in 2θ across the direct beam and fitting the results to a Gaussian. The width of the W_v function was calculated from known collimations. The Gaussian approximations for W_h and W_v allow integration by means of a Gauss-Hermite numerical algorithm.

The experimental intensity $J(Q)$ is given by

$$\begin{aligned} J(Q) = &\{[S(Q) - C(Q))/t_S] - [(E(Q) - C(Q))/t_E]\} \\ &- \delta\{[(R(Q) - C(Q))/t_R] - [(E(Q) - C(Q))/t_E]\} \end{aligned} \tag{6}$$

where S, R and E are the observed intensities (normalized to constant monitor) of the sample plus cell, solvent plus cell and empty cell respectively and t_S, t_R and t_E are the corresponding measured transmissions. The observed intensity with the incident beam blocked by cadmium is designated by C and δ is the volume fraction of solvent in solution.

4. Results and discussion

The scattering intensity calculated for different models, and its comparison with analytical calculations have been shown in figures 2, 3, 4 and 5. Figure 6 shows the comparison of the calculated intensities for different models with experimental data. In the first four figures the neutron wavelength was 4 Å. In figure 6 the wavelength was 2·5 Å. A comparison of figures 4 and 5 provides some evidence for the necessity of utilizing as many spheres as possible. Whereas the results for a solid cylinder (figure 4) are satisfactory up to $\simeq 3{\cdot}6°$, a discrepancy at scattering angle of about 1·4° appears in the calculations illustrated in figure 5. The latter result arises because an insufficient number of spheres span the wall thickness of this hollow cylinder. In practice the number of subunits whose parameters are varied is limited. Typically it

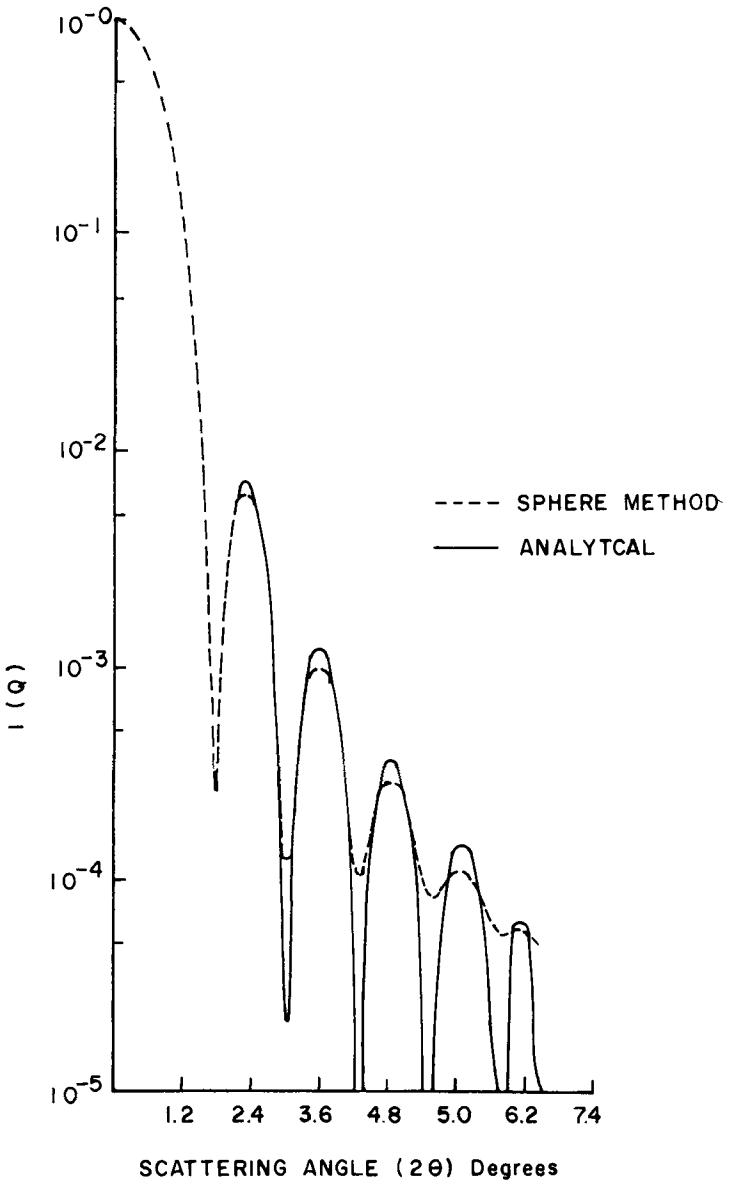

Figure 2. Scattering function from a sphere, $R = 100$ Å, $Rg = 77{\cdot}45$ Å, $R_1 = 13{\cdot}57$ Å, number of spheres = 402.

is possible to vary the size and position of one ellipsoid in a three-ellipsoid model. The program also allows calculations for mixtures of two different molecular species, a possibility where partial dimerization occurs.

Every successful run of the program produces plots of molecular profiles in the X–Y, Y-Z and Z-X planes. In the fitting procedure the quality of fit parameter is defined as

$$\chi^2 = \sum_{i=1}^{N} M_i[J(Q_i)_{\text{calc}} - J(Q_i)_{\text{expt}}]^2/(N-P) \tag{7}$$

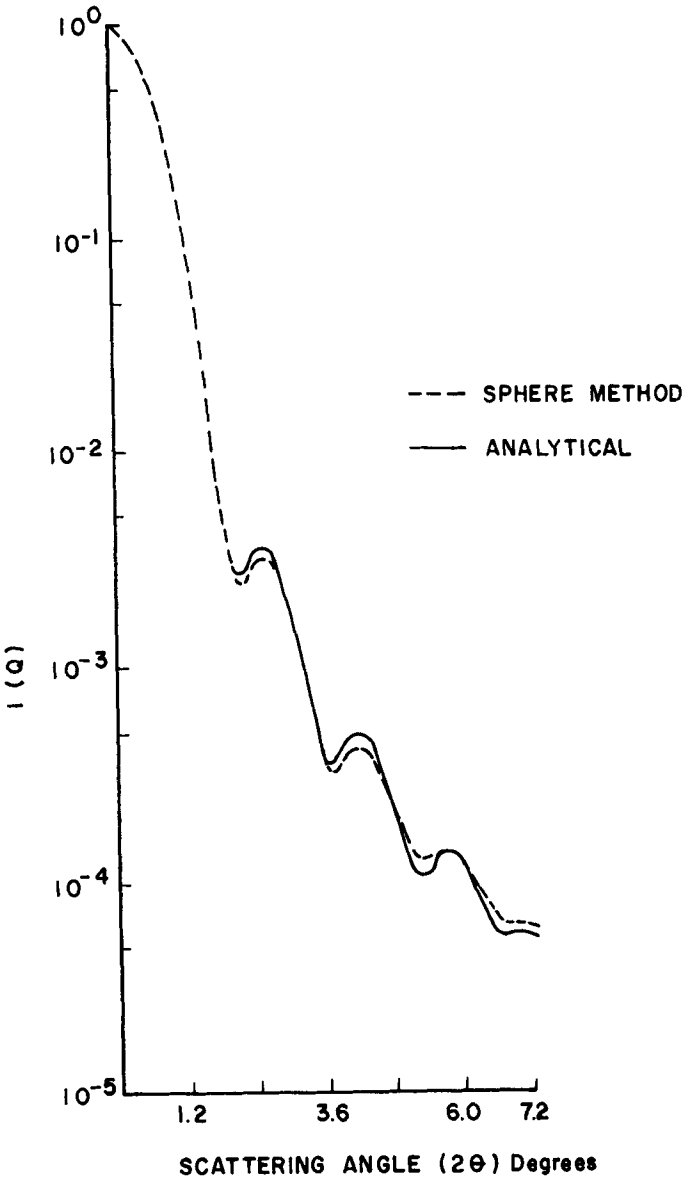

Figure 3. Scattering function from an ellipsoid, $A = 80$, $B = 80$, $C = 150$ Å, $Rg = 83{\cdot}87$ Å, $R_1 = 13{\cdot}39$ Å, number of spheres = 403.

where $J(Q_i)_{calc}$ and $J(Q_i)_{expt}$ are the calculated and experimental value respectively at the point i, N is the number of data points and P is the number of parameters which are varied. The weighting factor is $M_i = (1/J(Q_i)_{expt})^2$. The program varies the model parameters to minimize χ^2.

The running time in summing the Debye formula (eq. 1) goes up very fast as the number of small spheres, n, increases. To provide an idea of running times on an IBM 4341 computer, a fit was made of the scattering from neurotoxin molecule measured at a concentration of 8 g/l. The scan data comprised 301 points. The model

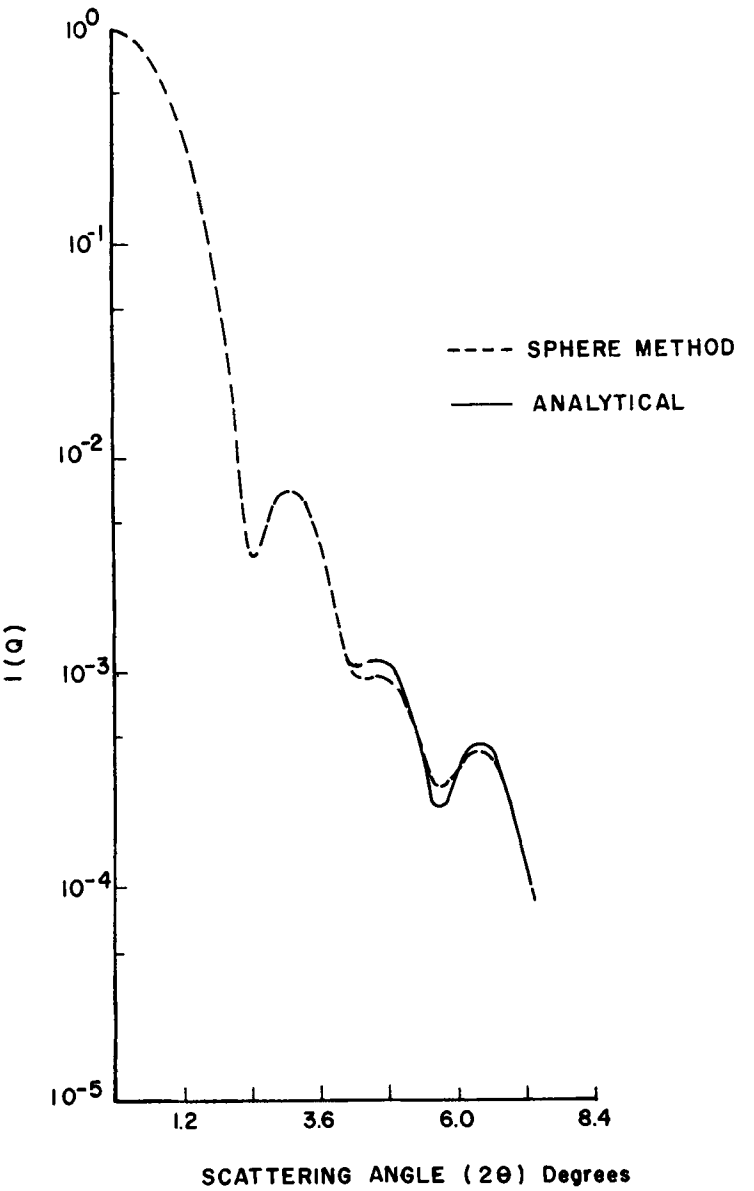

Figure 4. Scattering function from a cylinder, $R = 70$ Å, $H = 100$ Å, $Rg = 56{\cdot}93$ Å, $R_1 = 9{\cdot}15$ Å, number of spheres = 483.

chosen was an arbitrary shape as in figure 1. Time required was 360 s for convergence with $n = 20$, without convolution and with variation of four model-parameters. Time required was 12 s only without variation of model parameters. For $n = 40$, without convolution and with variation of four model parameters the time was 905 s, and without convolution and without variation of model parameters the time was only $\simeq 30$ s. When the analytic function was convoluted with the experimental resolution, the model volume increased slightly but the parameters remained the same within their combined errors. The time was 25 times greater for calculations with convolution

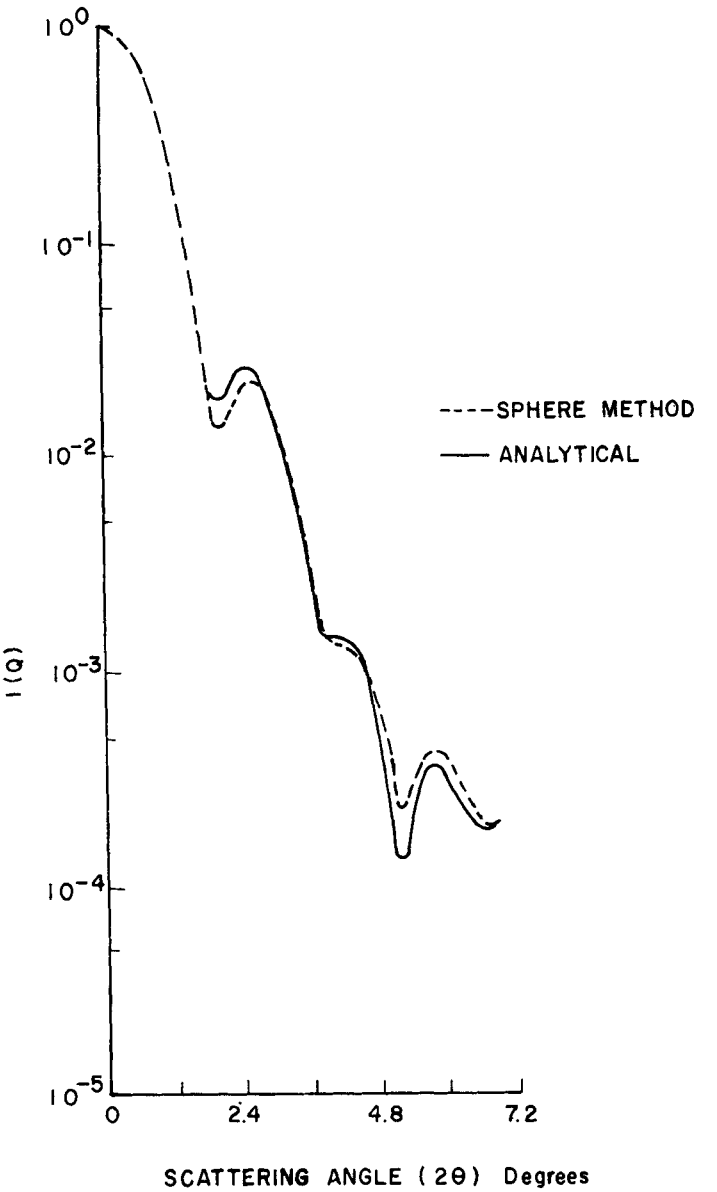

Figure 5. Scattering function from a hollow cylinder of outer radius = 80 Å and inner radius = 30 Å, $Rg = 66{\cdot}52$ Å, $R_1 = 9{\cdot}16$ Å, $H = 100$ Å, number of spheres = 537.

and with variation of four parameters for $n = 20$ and 36 times greater when $n = 40$. These results suggest that the preliminary calculations should be maintained without convolution until approximate values for model parameters are obtained.

The most reliable and commonly measured variable in small angle scattering experiments is the radius of gyration Rg. The radius of gyration is calculated from the expression

$$Rg^2 = \sum_{i=1}^{n} w_i I_{0i}^2 \Big/ \sum_{i=1}^{n} w_i \tag{8}$$

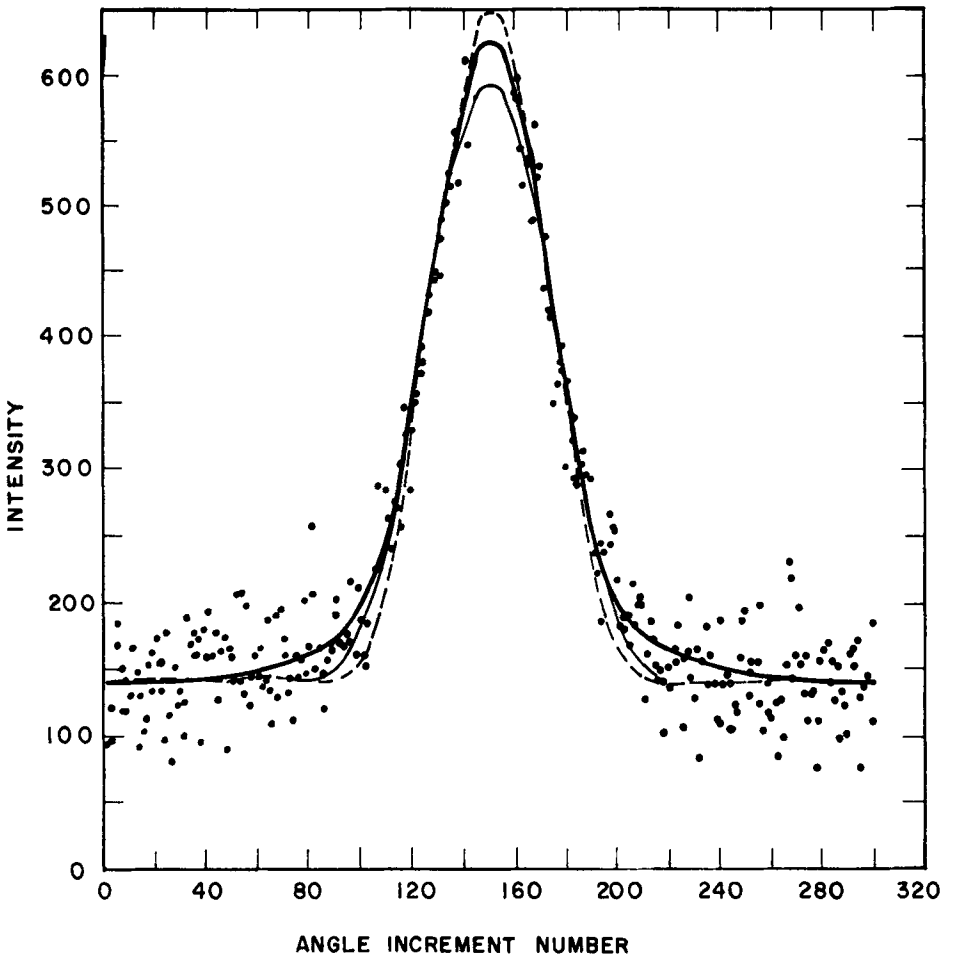

Figure 6. Scattering function from neurotoxin. The solid thick, solid thin and dashed lines represent the arbitrary, ellipsoid and sphere models respectively.

where

$$l_{0i}^2=(X_0-X_i)^2+(Y_0-Y_i)^2+(Z_0-Z_i)^2.$$

The X-coordinate, X_0, of the centre of scattering amplitude density is given by

$$X_0=\sum_{i=1}^{n} w_iX_i \Big/ \sum_{i=1}^{n} w_i. \tag{10}$$

Analogous expressions are used for Y_0 and Z_0. The coordinates of the centres of the small spheres are X_i, Y_i and Z_i.

The radii of gyration for different shapes calculated by using the present method are compared with analytical and experimental values (table 1).

Any useful model must yield a value of Rg in agreement with the measured value as determined by the Guinier approximation to the low-Q data. A valuable feature of our program is the calculation of Rg at every iteration; this is especially desirable when complicated molecular shapes are being generated. When the calculated Rg

Table 1. The radii of gyration for different shapes calculated by using the present method and its comparison with analytical and experimental values.

Model	Model parameters Å	Radius of gyration: Analytical Å	Radius of gyration: Present method Å
Sphere	$R = 100$	77·46	77·45
Ellipsoid	$A = 80$ $B = 80$ $C = 150$	84·02	83·87
Cylinder (height 2H)	$R = 70$ $H = 100$	57·30	56·93
Hollow cylinder	$R_1 = 80$ $R_2 = 30$ $H = 100$	66·95	66·52
Triple ellipsoid	Arbitrary model for cobra neurotoxin	Experimental $12{\cdot}2 \pm 0{\cdot}2$	12·24

differs significantly from the measured value the model is usually rejected. Results obtained by us shown in table 1 give sample indications as to the usefulness of the present method. With such advanced computational methods small angle neutron scattering techniques are gradually turning out to be a very important tool for the determination of the behavior and structure of biologically important compounds in nature.

References

Ahmed F U and Martel P 1986 Analysis of Neutron Small Angle Scattering Data from Molecules of Arbitrary Shape. (Internal Report), AECL-9185, Chalk River Nuclear Labs, Canada

Debye P 1915 *Ann. Phys.* **46** 809

Glatter O 1977 *J. Appl. Crystallogr.* **52** 243

Glatter O 1980 *Acta Phys. Austriaca* **52** 415

Mittelbach P and Porod G 1961 *Acta Phys. Austriaca.* **14** 185

Mittelbach P 1962 *Acta Phys. Austriaca.* **19** 53

Martel P, Kim S M and Powell B M 1980 *Biophys. J.* **31** 371

Porod G 1948 *Acta. Phys. Austriaca.* **2** 255

Sjoberg B 1978 *J. Appl. Crystallogr.* **11** 73

Wilkinshaw M D, Saenger W and Maelicke A 1980 *Proc. Natl. Acad. Sci. U.S.A.* **77** 2400