united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
**abdus salam**
international centre for theoretical physics

SMR/1270-18

# SCHOOL ON SYNCHROTRON RADIATION

**6 November – 8 December 2000**

***Miramare - Trieste, Italy***

*Supported in part by the Italian Ministry of Foreign Affairs in connection with the SESEME project*

*Co-sponsors: Sincrotrone Trieste, Società Italiana di Luce di Sincrotrone (SILS) and the Arab Fund for Economic and Social Development*

## *SAXS*

## *Nanocrystalline materials and fractal systems*

A. Craievich
Universidade de Sao Paulo
Brazil

J. Appl. Cryst. (1988). 21, 781–785

# Small-Angle Scattering by Fractal Systems

By J. Teixeira

*Laboratoire Léon Brillouin,* CEN-Saclay, 91191 Gif-sur-Yvette CEDEX, France*

(*Received* 28 *October* 1987; *accepted* 4 *January* 1988)

## Abstract

Fractal structures are characterized by self similarity within some spatial range. The mass distribution in a fractal object varies with a power $D$ of the length $R$, smaller than the dimension $d$ of the space. When the range of physical interest falls below 1000 Å, scattering techniques are the most appropriate way to study fractal structures and determine their fractal dimension $D$. Small-angle neutron scattering (SANS) is particularly useful when advantage can be taken of isotopic substitution. It is easy to show that the scattering law for a fractal object is given by $S(Q) \sim Q^{-D}$, where $Q$ is the magnitude of the scattering vector. However, in practice some precautions must be taken because, near the limits of the fractal range, there are important deviations from this simple law. Some relations are derived which can be applied in relatively general situations, such as aggregation and relation. The effects of polydispersity, important, in particular, in situations described by percolation models, are also shown.

## Introduction

Since its introduction by Mandelbrot (1977), the concept of fractals has been used in many situations and systems, such as aggregates and gels. The fractal dimension very often summarizes and averages over very complex structural details. In physics this kind of description is useful if some physical property can be related to the fractal dimension. This is, for example, the case for aggregation, where different models, such as diffusion-limited aggregation (Witten & Sander, 1983), predict a well defined fractal dimension. However, the determination of the fractal dimension is not always easy, and some precautions must be taken when analysing experimental data.

## Theory

For the purpose of this paper, a fractal object can be described by two properties. The first one is self similarity, which means that a detail of the fractal object is structurally identical to the whole or, in other words, that the structure of the object is independent of the characteristic length scale of observation. It is clear that such a situation will break down at some scale in a real experiment. This means that, in the real world, an object can be considered as a fractal object and be described by a fractal dimension only within some spatial range, covering at least one decade. Near the borders of the fractal domain the ideal scattering law of the fractal object will no longer be obeyed and a specific treatment of the system cannot be avoided.

The second property which characterizes the fractal object is its fractal or Hausdorff dimension, $D$ (Mandelbrot, 1977), defined as the exponent of the linear dimension $R$ in the relation $M(R) \sim (R/r_0)^D$ where $M$ represents the mass and $r_0$ is the gauge of measurement. The exponent $D$ characterizes mass fractals. It is smaller than the dimension $d$ of the space where the fractal object is embedded.

The concept of fractal dimension can easily be generalized to fractal surfaces. The fractal surface dimension $D_s$ relates the measure of the surface and the linear length. Its limits are $d-1$ and $d$.

In computer simulations, fractal objects are generated and their dimensions are determined from images in real space or their projections in a two-dimensional space, *e.g.* on photographs. If the simulation is performed in three dimensions, the fractal dimension is conserved in the projection if it is smaller than 2 (Weitz & Oliveria, 1984). Real experiments have been analysed as well in this way. The statistics of such a technique is in general poor. Conversely, scattering techniques, because of averaging in reciprocal space, are the appropriate way to determine a structure. This statement applies of course to fractal structures and some review papers have been devoted recently to the subject (Teixeira, 1986; Martin & Hurd, 1987).

Because the fractal structure never extends below the atomic separation, evaluations of fractal dimensions by scattering techniques concern small scattering vectors $Q$, say below 0·5 Å$^{-1}$. In a small-$Q$ experiment the sample is analysed with a spatial resolution of the order of $2\pi/Q$. Typical values for $Q$ in a small-angle neutron scattering (SANS) experiment extend from $10^{-3}$ to 1 Å$^{-1}$. Consequently, the fractal structure is easily observed by SANS if self similarity holds in some spatial range below 1000 Å.

*Laboratoire commun CEA–CNRS.

0021-8898/88/060781-05$03.00

© 1988 International Union of Crystallography

If the fractal behaviour extends over the entire experimental range the scattering law is particularly easy to derive. Actually, from the mass distribution $M(R) \sim R^D$, one obtains the density distribution $\rho \sim R^{D-d}$ (which goes to zero at infinite distances) and, by Fourier transform, the scattering law $S(Q) \sim Q^{-D}$. This gives the well known straight line with slope $D$ in a logarithmic plot of $S(Q)$. Of course, this law applies only if one remains at values of $Q$ large enough to be independent of the non-zero density of the sample at large scales. This law also breaks down for large values of $Q$ when one analyses the sample over length scales small enough to be sensitive either to individual scatterers (the form factor) or to the surface (Porod region).

If the individual scatterers are relatively monodisperse, the scattered intensity $I(Q)$ can be decomposed as

$$I(Q) = \Phi P(Q) S(Q). \tag{1}$$

In this expression, $\Phi = N/V_0$, where $N$ is the number of scatterers and $V_0$ the volume of the sample. $\Phi$ is a number density expressed in $cm^{-3}$.

$P(Q)$ is a function of the form factor $F(Q)$ which will be defined below. It depends on the spatial distribution of the scattering lengths of the atoms constituting the sample. A precise determination of $P(Q)$ is not always possible, but simple expressions can be derived for some simple geometrical shapes. In principle, one must start from the coherent scattering length $b_i$ of each individual atom $i$ and evaluate, from their instantaneous positions $\mathbf{r}_i$, the density

$$\rho(\mathbf{r}) = \sum_i b_i \delta(\mathbf{r} - \mathbf{r}_i). \tag{2}$$

[Because the distribution $\delta(\mathbf{r} - \mathbf{r}_i)$ has the inverse dimensions of a volume, $\rho(r)$ is expressed in $cm^{-2}$.]

For practical purposes, in SANS the spatial resolution is low enough to allow some spatial averaging. In this context it is useful to introduce the notion of contrast in the following way. $\rho(\mathbf{r})$ is averaged over distances small compared with $1/Q$. The medium where the fractal object is immersed (often called the solvent) is in general well represented by a single number $\rho_0$ because its structure cannot be observed with the resolution of the SANS experiment. Very often the particle itself is supposed to have a uniform density. The contrast is given by $\rho(\mathbf{r}) - \rho_0$. One of the important advantages of SANS is that the contrast can very often be modified by isotopic substitution which has only minor effects on the structure. This is currently used with hydrogenated samples, owing to the large differences in the scattering lengths of hydrogen and deuterium.

The form factor $F(Q)$ is the Fourier transform of the contrast evaluated over the volume of the particle and $P(Q)$ is the square of the modulus of $F(Q)$ averaged over the ensemble of particles:

$$P(Q) = \langle |F(Q)|^2 \rangle = \left\langle \left| \int_{\substack{\text{volume of} \\ \text{the particle}}} [\rho(\mathbf{r}) - \rho_0] \exp(i\mathbf{Q}.\mathbf{r})\, d\mathbf{r} \right|^2 \right\rangle$$

[$P(Q)$ is expressed in $cm^2$].

For some simple shapes $P(Q)$ is easily determ (Guinier & Fournet, 1955). For example, for unif spheres of density $\rho$, volume $V$ and radius $r_0$,

$$P(Q) = V^2(\rho - \rho_0)^2 \times [3\sin(Qr_0) - Qr_0\cos(Qr_0)/(Qr_0)^3]^2.$$

At large values of $Q$ the scattered intensity beco independent of the shape of the particles an dominated by surface scattering. This is particul true if a small polydispersity is present, as is always case in real systems. In this domain, $P(Q)$ is giver the Porod law

$$P(Q) = 2\pi(\rho - \rho_0)^2 S/Q^4$$

where $S$ is the total surface. In this derivation, density profile is supposed to vary sharply o distances smaller than $1/Q$. If the density profile va less sharply, one obtains smoother $Q$ depender (Auvray, 1986).

The Porod law can be generalized (Bale & Schm 1984) to fractal surfaces of dimension $D_s$. Writing

$$S = (r/r_0)^{D_s}$$

for the measurement of the fractal surface v dimension $D_s$, one obtains

$$P(Q) = \pi(\rho - \rho_0)^2 S\Gamma(5 - D_s) \times \sin[\pi(D_s - 1)/2] Q^{-(6-D_s)},$$

which gives (5) when $D_s = 2$.

Because $6 - D_s$ falls between 3 and 4, the $Q$ pendence of the scattering law of a fractal surfac not easily distinguishable from that of a smo interface.

$S(Q)$ is an effective structure factor (Chen & Ben douch, 1985), given by

$$S(Q) = 1 + \frac{|\langle F(Q) \rangle|^2}{\langle |F(Q)|^2 \rangle} [S(Q) - 1].$$

For a centrosymmetric particle $S(Q) = S(Q)$. $S($ describes the spatial distribution of the individ scatterers. It is the Fourier transform of the p correlation function, $g(r)$. [$g(r)$ is dimensionle $\Phi g(r)$ represents the probability of finding a particle a distance $r$ from a particle situated at the origin.]

Explicitly, $S(Q)$ can be written as [see, for exam Squires (1978)]

$$S(Q) = 1 + \Phi \int [g(r) - 1] \exp(i\mathbf{Q}.\mathbf{r})\, d\mathbf{r}.$$

For isotropic systems

$$S(Q) = 1 + 4\pi\Phi \int_0^\infty [g(r) - 1] r^2 \sin(Qr)/Qr \, dr. \quad (10)$$

If the individual scatterers are sufficiently far apart (dilute system), $S(Q)$ reduces to unity and the total scattered intensity depends only on the form factor. Whenever possible, dilute systems are used to determine $P(Q)$.

Above some concentration, the individual scattering objects can no longer be seen as isolated. There will be at least a steric effect. In situations where there are, moreover, interparticle interactions, such as in systems containing charged particles, structural effects can appear even at extremely low concentrations, avoiding the direct determination of $P(Q)$. Methods have been derived to evaluate the structure factor in such situations (Guinier & Fournet, 1955).

The structure factor $S(Q)$ very often has a liquid-like shape with one or more oscillations, and goes necessarily to unity at large $Q$. It is unfortunately frequent practice, in the difficult interpretation of $S(Q)$, that a physical significance is attributed to the position of the first structural peak. Such a procedure is incorrect, because the structure factor is the Fourier transform of a pair correlation function. The first structural peak cannot in general be associated with some enlarged Bragg peak, as in a disordered solid material. At most, its position will give the order of magnitude of some characteristic distance in the sample. Because the measured $Q$ range in a typical SANS experiment does not allow the evaluation of the Fourier integral, the determination of $g(r)$ must start from a plausible model able to reconstitute the measured scattering function.

For a fractal object, the structure factor $S(Q)$ can be derived in the following way. From the meaning of $g(r)$, the total number of particles within a sphere of radius $r$ centred in a central particle is given (for $d = 3$) by

$$N(r) = \Phi \int_0^r g(r) 4\pi r^2 \, dr \quad (11)$$

or

$$dN(r) = \Phi g(r) 4\pi r^2 \, dr. \quad (12)$$

On the other hand, a fractal object is characterized by a spatial distribution of the individual scatterers given by

$$N(r) = (r/r_0)^D, \quad (13)$$

where $r_0$ is the gauge of the measurement, which has the magnitude of the characteristic dimension of each individual scatterer.

Differentiation of (13) and identification with (12) gives

$$\Phi g(r) = (D/4\pi r_0^D) r^{D-3}. \quad (14)$$

Because $D$ is smaller than 3, $g(r)$ goes to zero at large $r$, as we have seen above. This is clearly unphysical. At some large scale, the sample will show a macroscopic density. A good knowledge of the sample allows in general a reasonable assumption for the large-scale behaviour of $g(r)$. We will adopt here a procedure first introduced by Sinha, Freltoft & Kjems (1984), which proved to be very useful, even if not absolutely general. It involves an expression of the form $\exp(-r/\xi)$, where $\xi$ is a cut-off distance, to describe the behaviour of the pair correlation function at large distances. To derive the analytical form of $S(Q)$ within this assumption, one can use the general theory of liquids, where the uniform density is subtracted to avoid a divergence in the evaluation of $S(Q)$ (Chen & Teixeira, 1986; Teixeira, 1986). We then write

$$\Phi[g(r) - 1] = (D/4\pi r_0^D) r^{D-3} \exp(-r/\xi) \quad (15)$$

The introduction of the cut-off distance $\xi$ is done in analogy with the Ornstein–Zernike treatment of critical phenomena (Stanley, 1971), which is represented by (15) with $D = 2$.

The meaning of $\xi$ is only qualitative and has to be made precise in any particular situation. Generally speaking, it represents the characteristic distance above which the mass distribution in the sample is no longer described by the fractal law. In practice, it can represent the size of an aggregate or a correlation length in a disordered material.

Fourier transformation of (15) gives, in accordance with (10),

$$S(Q) = 1 + \frac{D}{r_0^D} \int_0^\infty r^{D-1} \exp(-r/\xi) \frac{\sin(Qr)}{Qr} \, dr$$

$$= 1 + \frac{1}{(Qr_0)^D} \frac{D\Gamma(D-1)}{[1 + 1/(Q^2\xi^2)]^{(D-1)/2}} \times \sin[(D-1)\tan^{-1}(Q\xi)] \quad (16)$$

where $\Gamma(x)$ is the gamma function of argument $x$.

This expression reduces of course to $S(Q) \sim Q^{-D}$ when $\xi^{-1} < Q < r_0^{-1}$. Expressions (1), (3) to (7) and (16) have been used with success to fit experimental data in many different situations. As an example, we give in Fig. 1 the scattered intensity from different samples of silica aerogel and their fit by the preceding equations (Courtens & Vacher, 1987).

The unity term in $S(Q)$ becomes dominant at large $Q$. Then, the scattered intensity $I(Q)$ is dominated by the form factor $P(Q)$ [see (1)]. By contrast, at small values of $Q$ the intensity $I(Q)$ is dominated by $S(Q)$ and for $Q \ll \xi^{-1}$ has the generalized Guinier-type behaviour

$$\lim_{Q \to 0} S(Q) = \Gamma(D+1)(\xi/r_0)^D \times \{1 - [D(D+1)/6] Q^2\xi^2\}. \quad (17)$$

A comparison of this expression with Guinier's law shows that $\xi$ is related to a generalized radius of gyration $R_g$ by

$$R_g^2 = D(D+1)\xi^2/2.$$

In a general way, $S(Q)$ can be considered as the form factor of an ensemble of fractal objects distributed at a larger scale in the space.

## Polydispersity

A particular situation is that of gelation processes described by percolation models (Stauffer, 1976; de Gennes, 1979), where a large polydispersity is present. During the gelation process of branched polymers, there is a very large distribution $n(s)$ of polymers of different sizes $s$. This distribution is characterized (Fig. 2) by two numbers: the average size, $N_w$, and the maximum cut-off size, $N_z$.

Within the percolation theory, $N_w$ and $N_z$, near the percolation threshold, are associated with the two universal exponents, $\gamma$ and $\sigma$.

$$N_w \sim \varepsilon^{-\gamma} \tag{18}$$

$$N_z \sim \varepsilon^{-1/\sigma} \tag{19}$$

where $\varepsilon$ measures the distance to percolation.

The distribution $n(s)$ has been derived by Stauffer (1985) and depends on the exponent $\tau$:

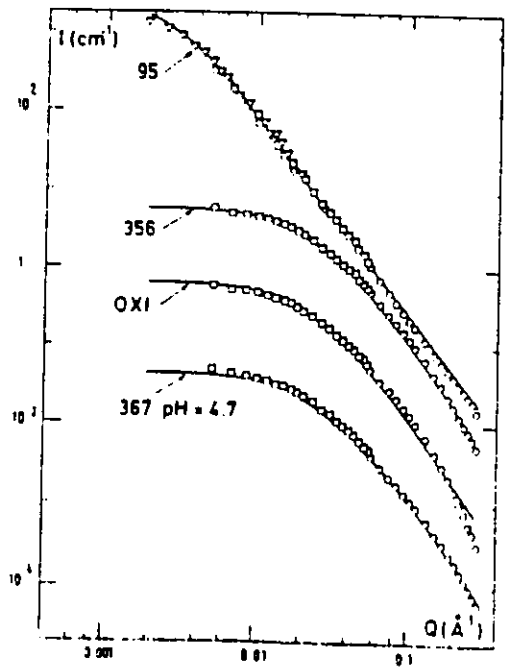


Fig. 1. Scattered intensity from several samples of silica aerogels. Numbers labelling the curves give the densities in g cm$^{-3}$ and OXI means an oxidized sample with density equal to 1.56 g cm$^{-3}$. The sample labelled pH = 4.7 was prepared with acid catalysis. All the others are neutrally reacted samples. Solid lines represent fits with equations (1), (5) and (16). Each intensity is divided by ten compared with the previous one, starting from the top. (From Courtens & Vacher, 1987.)

$$n(s) = s^{-\tau} f_1(\varepsilon s^{\sigma}) = s^{-\tau} f_1(s/N_z), \tag{20}$$

where $f_1(x)$ is an exponentially decaying function which accounts for the cut off of the distribution at large $s$.

Near the percolation threshold, each polymer has a fractal structure with dimension $D$. The radius of gyration of each polymer can be associated with the cut-off parameter $\xi$. Consequently, one can write

$$s = \xi^D \tag{21}$$

where $\xi$ is now the radius of gyration of the polymer of size $s$.

The pair correlation function of a single polymer containing $s$ monomers (the individual scatterers) has the form of (15), where $D$ is the fractal dimension of the polymer. In a more general way, one writes instead of (15),

$$\Phi[g(r)-1] = (D/4\pi r_0^D) r^{D-3} f_2(-r/\xi)$$
$$= (D/4\pi r_0^D) r^{D-3} f_2(-r/s^{1/D}), \tag{22}$$

where we use (21). $f_2(-r/\xi)$ represents a decreasing function of $r$.

The scattering function of such a polymer is obtained, following the same steps as for the derivation of (16), as

$$S(s,Q) = 1 + \frac{D}{r_0^D}\int_0^{\infty} r^{D-3} f_2(-r/s^{1/D}) \frac{\sin(Qr)}{Qr}\,dr$$

$$= 1 + \frac{D}{(r_0 Q)^D} f_3(Q\xi)$$

$$= 1 + s f_4(Q^D s) \tag{23}$$

where $f_3(x)$ and $f_4(x)$ can be determined if $f_2(-r/\xi)$ is known.

The total scattering function at infinite dilution is now obtained (Martin & Ackerson, 1985; Bouchaud, Delsanti, Adam, Daoud & Durand, 1986) by summing all the contributions $S(s,Q)$ with their relative

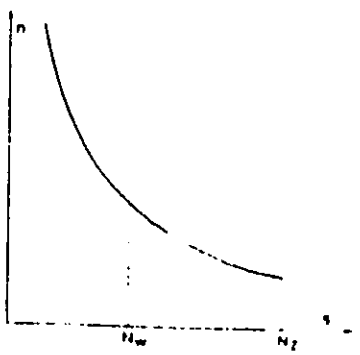


Fig. 2. Schematic representation of the number $n(s)$ of polymers with size $s$, near the percolation threshold. $N_w$ and $N_z$ are the average and the maximum size, respectively. This distribution is described by equation (20).

weights, $n(s)$, given by (20):

$$S(Q) - 1 = \int ds\, sn(s)[S(s, Q) - 1]$$
$$= \int ds\, s^2 s^{-\tau} f_1(s/N_z) f_4(Q^D s)$$
$$= Q^{-(3-\tau)D} f_5(QR_z) \quad (24)$$

where $R_z \sim N_z^{1/D}$. This result applies only for $\tau > 2$.

For large values of $Q$ ($Q \gg R_z^{-1}$), $f_5(QR_z) \sim 1$. Comparing (16) and (24), one notices that, in the fractal region, the scattered intensity is now represented in a logarithmic plot by a straight line with a slope $(3-\tau)D$, instead of $D$. The theory of percolation predicts $D = 2$ and $\tau = 2{\cdot}2$. In a recent experiment Bouchaud *et al.* (1986) have measured these exponents using fractionated samples containing only the polymers of large size. In the polydisperse sample, the dependence on $Q$ of the scattered intensity gives $(3-\tau)D = 1{\cdot}59$ and, in the fractionated sample, $D$ is evaluated directly giving $D = 2$ (Fig. 3), in excellent agreement with the theory.

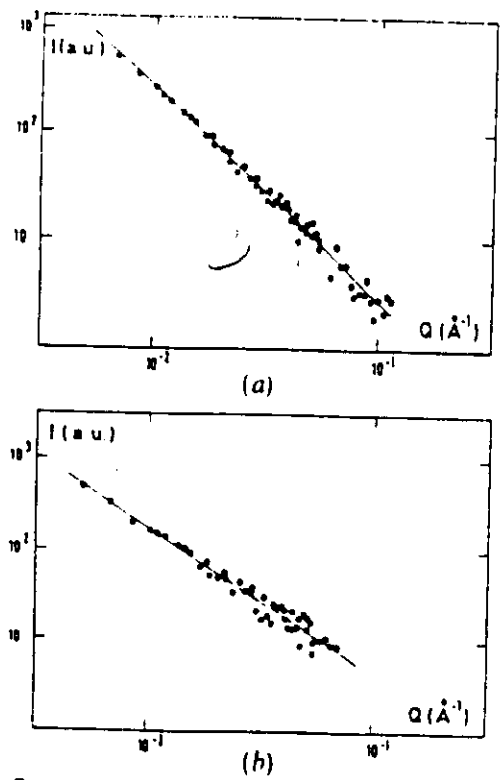


Fig. 3. (*a*) Scattered intensity by a fractionated sample containing only polymers with a large size. The slope gives $D = 1{\cdot}98$. (*b*) Scattered intensity by the dilute polydisperse sample. The slope gives $D(3-\tau) = 1{\cdot}59$. (From Bouchaud *et al.* 1986.)

These results show that fractal concepts are useful in interpreting complex situations in disordered materials and are the most convenient method for studying physical phenomena such as aggregation and gelation. Scattering techniques and SANS, in particular, appear to be the most appropriate technique for the determination of the fractal dimension. We have summarized some simple and relatively general expressions which can be derived to describe the intensity scattered by fractal objects.

## References


AUVRAY, L. (1986). *C. R. Acad. Sci.* **302**, 859–862.

BALE, H. D. & SCHMIDT, P. W. (1984). *Phys. Rev. Lett.* **53**, 596–599.

BOUCHAUD, E., DELSANTI, M., ADAM, M., DAOUD, M. & DURAND, D. (1986). *J. Phys. (Paris)*, **47**, 1273–1277.

CHEN, S.-H. & BENDEDOUCH, D. (1985). In *Enzyme Structure, Method in Enzymology*, edited by C. H. W. HIRS & S. N. TOMASHEFF. New York: Academic Press.

CHEN, S.-H. & TEIXEIRA, J. (1986). *Phys. Rev. Lett.* **57**, 2583–2586.

COURTENS, E. & VACHER, R. (1987). *Z. Phys. B*, **68**, 355–361.

GENNES, P. G. DE (1979). *Scaling Concepts in Polymer Physics*. Cornell Univ. Press.

GUINIER, A. & FOURNET, G. (1955). *Small-Angle Scattering of X-rays*. New York: Wiley.

MANDELBROT, B. B. (1977). *Fractals, Form and Dimension*. San Francisco: Freeman.

MARTIN, J. E. & ACKERSON, B. J. (1985). *Phys. Rev. A*, **31**, 1180–1182.

MARTIN, J. E. & HURD, A. J. (1987). *J. Appl. Cryst.* **20**, 61–78.

SINHA, S. K., FRELTOFT, T. & KJEMS, J. (1984). In *Kinetics of Aggregation and Gelation*, edited by F. FAMILY & D. P. LANDAU, pp. 87–90. Amsterdam: North Holland.

SQUIRES, G. L. (1978). *Introduction to the Theory of Thermal Neutron Scattering*. Cambridge Univ. Press.

STANLEY, H. E. (1971). *Introduction to Phase Transitions and Critical Phenomena*. Oxford: Clarendon Press.

STAUFFER, D. (1976). *J. Chem. Soc. Faraday Trans. 2*, **72**, 1354–1364.

STAUFFER, D. (1985). *Introduction to Percolation Theory*. London: Taylor & Francis.

TEIXEIRA, J. (1986). In *On Growth and Form*, edited by H. E. STANLEY & N. OSTROWSKY, pp. 145–162. Dordrecht: Nijhoff.

WEITZ, D. A. & OLIVERIA, M. (1984). *Phys. Rev. Lett.* **52**, 1433–1436.

WITTEN, T. A. & SANDER, L. M. (1983). *Phys. Rev. B*, **27**, 5686–5697.

*J. Appl. Cryst.* (1997). 30, 623–627

# Formation and Growth of Semiconductor PbTe Nanocrystals in a Borosilicate Glass Matrix

A. F. Craievich,[a,b]* O. L. Alves[c] and L. C. Barbosa[d]

[a]*National Synchrotron Light Laboratory/CNPq, Campinas SP, Brazil,* [b]*Institute of Physics/USP, São Paulo SP, Brazil,* [c]*Solid State Chemistry Laboratory, Institute of Chemistry/UNICAMP, Campinas SP, Brazil, and* [d]*Institute of Physics/UNICAMP, Campinas SP, Brazil. E-mail: aldo@lnls.br*

*(Received 23 July 1996; accepted 29 January 1997)*

## Abstract

Pb- and Te-doped borosilicate glasses are transformed by appropriate heat treatment into a composite material consisting of a vitreous matrix in which semiconductor PbTe nanocrystals are embedded. This composite exhibits interesting non-linear optical properties in the infrared region, in the range 10–20 000 Å. The shape and size distribution of the nanocrystals and the kinetics of their growth were studied by small-angle X-ray scattering (SAXS) during *in situ* isothermal treatment at 923 K. The experimental results indicate that nanocrystals are nearly spherical and have an average radius increasing from 16 to 33 Å after 2 h at 923 K, the relative size dispersion being time-invariant and approximately equal to 8%. This investigation demonstrates that the kinetics of nanocrystal growth are governed by the classic mechanism of atomic diffusion. The radius of nanocrystals, deduced by applying the simple Efros & Efros [*Sov. Phys. Semicond.* (1982), **16**, 772–775] model using the energy values corresponding to the exciton peaks of optical absorption spectra, does not agree with the average radius determined by SAXS.

## 1. Introduction

A number of doped glasses are transformed by appropriate heat treatment into composite materials consisting of a vitreous matrix in which semiconductor nanocrystals are embedded. When the semiconductor crystals are small, electronic quantum confinement effects occur. The optical properties of a number of silicate and borosilicate glasses containing CdS (Potter & Simmons, 1988), CdSe (Borrelli, Hall, Holland & Smith, 1987), CdTe (Medeiros Neto, Barbosa, Cesar, Alves & Galembeck, 1991) *etc.*, have been investigated by optical measurements in the visible wavelength range. Recently, a new material, a borosilicate glass–PbTe nanocrystal composite, which exhibits quantum confinement effects in the infrared region, was discovered (Reynoso *et al.*, 1995). The consequent non-linear optical properties, being in the infrared, make this material a good candidate for applications in telecommunication devices.

The structural characterization of nanocrystals embedded in glassy matrix is usually carried out by analyzing the experimental optical absorption spectrum of the material. This spectrum exhibits a more or less sharp absorption edge and, in some cases, one or more exciton peaks, their position in the spectrum being dependent on the average size, on the size distribution and on the shape of the nanocrystals.

The method most frequently used for determining the average radius of the nanocrystals from optical data applies the Efros & Efros (1982) equation. This equation is based on a theoretical model for optical absorption, which assumes spherical nanocrystals and simple parabolic energy bands for electrons and holes. It is worthwhile, however, to carry out independent experimental studies of the composites using a technique that provides more direct structural information. Because of the small size of the nanocrystals, the small-angle scattering (SAXS) technique is a good choice for obtaining relevant structural information.

Applications of glass–semiconductor nanocrystal composites in non-linear optical devices usually require the nanocrystals to be nearly isodiametric and have a narrow size distribution. This paper reports the first SAXS study of a borosilicate glass–PbTe nanocrystal composite.

## 2. Sample preparation and optical characterization

The nominal composition of the studied glass is $52SiO_2$–$8B_2O_3$–$20ZnO$–$20K_2O$. The raw material was doped with 2wt% PbO and Te, homogenized at 1623 K during 50 min and quenched to room temperature. By appropriate polishing, a transparent lame about 0.1 mm thick was obtained. It was then placed inside a cell at 923 K for *in situ* SAXS measurements.

Optical absorption spectra of an equivalent sample, also held at 923 K for different periods of time, were recorded and are plotted in Fig. 1. Taking into account that the energy gap of a bulk PbTe semiconductor is 0.337 eV, the wavelength corresponding to the absorption edge is 36 800 Å. The various spectra in Fig. 1 show that the absorption edges are located at wavelengths, $\lambda$, of less than 20 000 Å, clearly exhibiting the known blue-shift effect associated with the nanometric size of the crystals. Well defined exciton peaks are also apparent, their position shifting from $\lambda$ = 12 000 Å to $\lambda$ = 19 000 Å

© 1997 International Union of Crystallography
Printed in Great Britain – all rights reserved

for periods of time increasing from 10 to 60 min. This confirms that the crystals are nanometric and that they progressively grow. The average sizes of the nanocrystals in the several samples were inferred from the optical absorption spectra plotted in Fig. 1, using the Efros & Efros (1982) equation. The results are reported in §4.

## 3. Experimental SAXS set-up

The SAXS experiments were performed at the D24 workstation of the synchrotron radiation laboratory LURE, Orsay. In this workstation, the beam is horizontally focused and monochromated using a bent silicon crystal ($\lambda$ = 1.49 Å). Three sets of slits define a collimated beam with a point-like cross section at the detection plane. The sample was held at constant temperature ($T$ = 923 K) using a high-temperature cell, stable within 1 K, during the *in situ* SAXS study.

The quenched and initially homogeneous doped glass was placed inside the high-temperature cell and maintained at a constant temperature during SAXS measurements. The X-ray scattered intensity, $I(q)$, was measured as a function of the modulus of the scattering vector $q$, using a one-dimensional position-sensitive X-ray detector. An ionization chamber, placed downstream from the sample, was used to monitor the intensity decay of the transmitted beam and to determine the sample attenuation. Parasitic air and slit scattering were subtracted from the total intensity. The counting time for each SAXS spectrum was 300 s and the measurements were repeated at intervals of about 500 s.

We have estimated the effects on the small-angle scattering intensity from the small, but non-zero, cross section of the incident beam at the detection plane and from the finite resolution of the position-sensitive detector. We concluded that these effects on the experimental SAXS intensity are negligible.

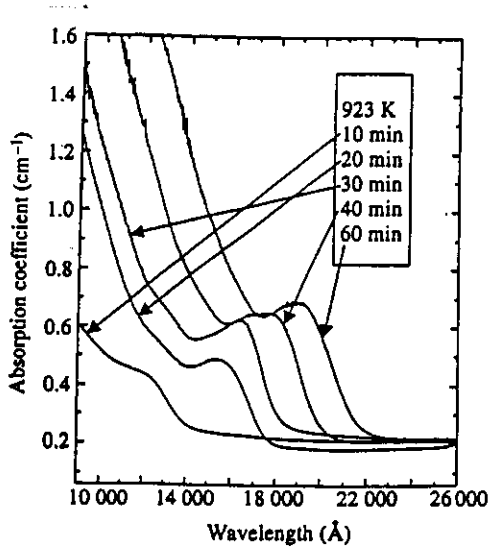


Fig. 1. Coefficient of optical absorption of the glass–PbTe composites for *ex situ* isothermal treatment during the indicated time.

## 4. Results and discussion

The formation of PbTe nanocrystals introduces heterogeneities in the electronic density of the initially quasi-homogeneous material. PbTe nanocrystals embedded in borosilicate glass are efficient X-ray scatterers because of their very high electronic density as compared with the density of the light matrix. If we assume a simple two-electronic density model for the nanocrystal–glass composite, with nanocrystals or 'particles' occupying a small fraction of the total volume, the Guinier equation applies to the SAXS intensity function in the small $q$-range limit:

$$I(q) = I(0)\exp[-(\langle R_g\rangle^2 q^2/3)] \qquad (1)$$

with

$$I(0) = (\Delta\rho)^2 N\langle v^2\rangle, \qquad (2)$$

$\langle R_g\rangle$ being the so called Guinier average of the radius of gyration of the particles, $\Delta\rho$ the difference in electron density between the particles and the matrix, $N$ the number of particles per unit volume and $\langle v^2\rangle$ the average of the particle square volume (Guinier & Fournet, 1955). In the case of a set of spherical particles with a radius distribution given by $g(R)$ and taking into account that in this case $R = (5/3)^{1/2}R_g$, the Guinier average $\langle R_g\rangle$ is given by

$$\langle R_g\rangle = (3/5)^{1/2}R_G = (3/5)^{1/2}\left[\int R^8 g(R)\mathrm{d}R / \int R^6 g(R)\mathrm{d}R\right]^{1/2} \qquad (3)$$

where $R_G$ is the Guinier average radius of the nanocrystals. Equation (1) is, in practice, obeyed within a large $q$ range when the particles are identical or have a rather narrow size distribution.

The experimental SAXS intensity functions are displayed as log $I$ *versus* $q^2$ (Guinier plot) in Fig. 2. The increase in SAXS intensity for samples held at 923 K during increasing periods of time is a consequence of the progressive formation and growth of PbTe nanocrystals, which are responsible for the variation in the optical absorption reported in §2.

We note that the Guinier plots of all curves in Fig. 2 show a rather wide linear range with increasing slope as the heat-treatment time increases. The increasing slope in the linear $q$ range indicates progressive nanocrystal growth, in agreement with the conclusion inferred from the optical absorption measurements. The Guinier average radius of the nanocrystals, $R_G$, was determined using (1) and (3) as a function of the isothermal treatment time. We also noted in Fig. 2 that the linear behavior of $I(q)$ is followed by a negative deviation as $q$ increases.

This occurs when the system is composed of quasi-isodiametric scattering particles with narrow size distribution.

The scattering curves in Fig. 2 exhibit an abrupt increase in intensity at very small $q$. This was also observed in the SAXS intensity function from the same sample before heating. The very large negative slope at very small $q$ indicates that this contribution to the experimental scattering intensity is a consequence of the existence of large-sized electronic density heterogeneities in the glass matrix. Since this contribution to SAXS intensity remains approximately constant for increasing periods of time, we concluded that the coarse heterogeneities do not evolve significantly. The same kind of effect has already been observed in a borosilicate glass doped with CdTe (Craievich, Alves & Barbosa, 1994). This contribution to SAXS intensity, being confined to very small $q$, is not considered in the further analysis associated with the formation and growth of PbTe nanocrystals.

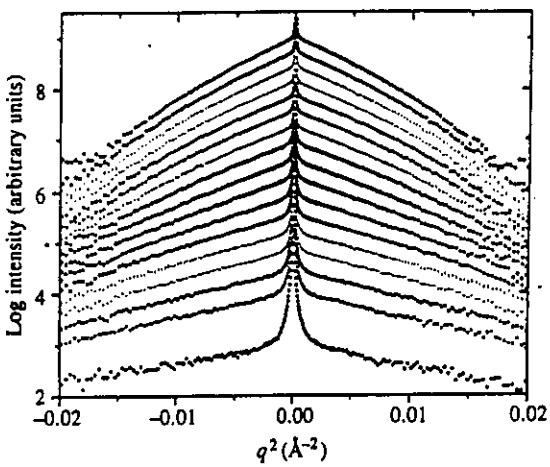


Fig. 2. Guinier plots of SAXS intensity from the glass–PbTe composite during *in situ* isothermal treatment. The weakest intensity curve corresponds to 10 min and the strongest to 2 h at 923 K.

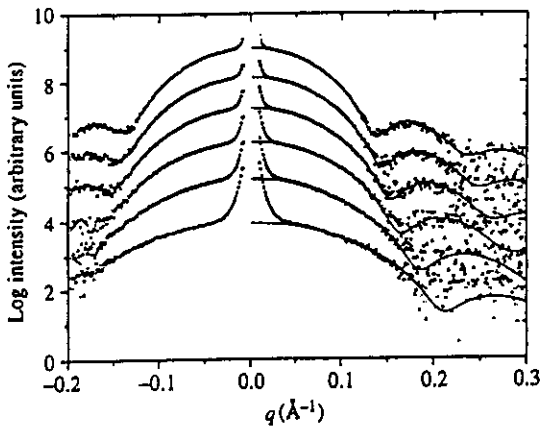


Fig. 3. Plots of the X-ray scattering intensity for different periods of time. The continuous lines are the theoretical functions for a set of spherical nanocrystals having a radius distribution with a relative standard deviation of 8% and the following average radii: 21.0 Å (19 min), 24.0 Å (38 min), 27.0 Å (56 min), 29.0 Å (78 min), 31.0 Å (98 min), 32.5 Å (119 min). The curve corresponding to the shortest time is on the bottom, the others are vertically displaced for clarity.

For long periods of time at 923 K a satellite peak is clearly apparent in the experimental SAXS intensity functions plotted in Fig. 3. This effect in scattering curves is, in general, produced by spherical or spheroidal particles, all of them having approximately equal radii. The conclusion about the spheroidal shape of the nanocrystals is equivalent to that previously established by transmission electron microscopy for glass–CdS composites (Liu & Risbud, 1990*b*).

We obtained more quantitative structural information on the studied materials by fitting the experimental SAXS intensity with a theoretical scattering function for a set of spherical nanocrystals. The scattering function, $I_s(q)$, for a sphere with a radius $R$ is given by (Guinier & Fournet, 1955):

$$I_s(q) = K(4\pi R^3/3)^2[\varphi(qR)]^2$$

where $K$ is a constant and

$$\varphi(qR) = 3\{[\sin(qR) - qR\cos(qR)]/(qR)^3\}.$$

Assuming a set of spherical nanocrystals with a Gaussian radius distribution embedded in a homogeneous matrix, the fitting function for the experimental SAXS intensity $I(q)$ is given by

$$I(q) = A\int g(R)R^6[\varphi(qR)]^2 dR + B \qquad (4)$$

where

$$g(R) = [(2\pi)^{1/2}\sigma]^{-1}\exp\{-[(R - R_A)^2/(2\sigma^2)]\}, \qquad (5)$$

$R_A$ being the nanocrystal average radius, $\sigma$ the standard deviation of the radius distribution, $A$ a scale constant and $B$ a constant that accounts for a constant contribution to the scattering intensity produced by short-range electronic density fluctuations in the matrix.

It was assumed in the fitting procedure using the function defined by (4) that the quotient $\sigma/R_A$ is a time constant equal to 8% during the whole isothermal treatment. The best fitting functions of several SAXS curves for different heat-treatment times are shown in Fig. 3. Good agreement between the theoretical and experimental SAXS intensities is apparent for all heat-treatment times in the full $q$ range, except the very small $q$ range, in which the contribution from the previously mentioned coarse heterogeneities is predominant.

We determined another Guinier average of the nanocrystal radii, $R_G'$, using (3) and (5) with the parameters $R_A$ and $\sigma$ of the fitting functions plotted in Fig. 3. The values of the average radii $R_G$ and $R_G'$ are expected to be the same provided that the basic assumptions of the model (dilute system, spherical

nanocrystals and Gaussian size distribution) actually correspond to the real system.

Another average radius, $R_o$, was determined from the optical spectra plotted in Fig. 1, using the Efros & Efros (1982) equation based on a simple model of spherical nanocrystals and parabolic energy bands for electrons and holes:

$$E - E_b = h^2/(8\mu R_o^2) \qquad (6)$$

where $E_b$ is the energy gap for the bulk semiconductor, $E$ the energy of the first exciton peak, $h$ the Planck constant and $\mu$ the exciton reduced mass determined by

$$1/\mu = 1/\mu_e + 1/\mu_h$$

with $\mu_e = \mu_h = 0.058m$ (Ishida, Matsuura, Mizuno & Fujiyasu, 1987), $m$ being the electron mass.

The Guinier average $R_G$ obtained from the slopes of the linear portions of the Guinier plots in Fig. 2, the average radius $R_A$ with an error bar $\pm\sigma$ corresponding to the fitting functions plotted in Fig. 3, the Guinier average of the radius $R_G'$ determined using the Gaussian size distribution and the above mentioned $R_A$ and $\sigma$ values, and the Efros & Efros (1982) average radius $R_o$, obtained from the spectroscopic data (Fig. 1), are plotted in Fig. 4 as functions of the heat-treatment time.

We observe in Fig. 4 that the $R_G$ and $R_G'$ values are larger than $R_A$. This is expected because the large crystals contribute much more to the Guinier average of nanocrystal radius, defined by (3), than small ones. However, in Fig. 4, a small but noticeable difference between $R_G$ and $R_G'$ is apparent. This indicates that the real structure of the borosilicate glass–PbTe nanocrystal composites deviates slightly from the assumed structural model consisting of a dilute system containing spherical shaped nanocrystals with Gaussian size distribution. We have estimated that the difference between $R_G$ and $R_G'$ is a consequence of slight deviations from the sphericity in nanocrystal shape, as was previously observed in electron microscopy studies of CdSe and CdS nanocrystals (Liu & Risbud, 1990*b*). We then concluded that the PbTe nanocrystals are approximately isodiametric or spheroidal.

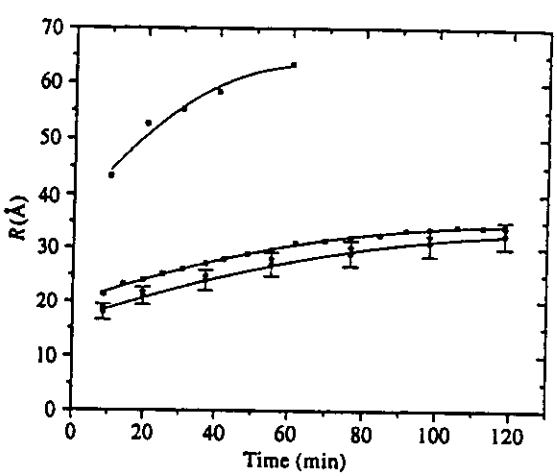


Fig. 4. The upper curve represents the Efros & Efros (1982) radius, $R_o$, as a function of time. The lower curves correspond to the time dependence of the average radius $R_A$ (circles) with $\pm\sigma$ bars, and the Guinier averages $R_G$ and $R_G'$ (squares and diamonds, respectively).

As can be seen in Fig. 4, the $R_o$ value calculated from the optical absorption spectra is about twice the average nanocrystal radius, $R_A$, determined by SAXS. This strong difference between $R_o$ and $R_A$ implies that the Efros & Efros (1982) equation [(6)] does not apply to PbTe nanocrystals. The disagreement cannot be due to the slight deviation of the nanocrystals from the spherical shape but can be explained, more probably, by the strong nonparabolicity of the energy bands of the studied semiconductor (Martinez, Schluter & Cohen, 1975).

In order to characterize the mechanism of nanocrystal formation and growth, we determined the time dependence of the integral of the SAXS intensity $I(q)$ in reciprocal space $Q(t)$, which is related to structural parameters by (Guinier & Fournet, 1955):

$$Q(t) = \int q^2 I(q, t)\mathrm{d}q = (1/4\pi)(\Delta\rho)^2 N(t)\langle v\rangle(t) \qquad (7)$$

where $\langle v\rangle$ is the average volume. $I(q,t)$ is the SAXS intensity after subtracting the constant intensity $B$ [(4)]. The integral $Q$ was determined by numerical integration from $q_{min} = 0.02$ Å$^{-1}$ to $q_{max} = 0.25$ Å$^{-1}$ and using Guinier and Porod laws (Guinier & Fournet, 1955), respectively, for $q < q_{min}$ and $q > q_{max}$. The fraction of $Q$ corresponding to the $q$ ranges below $q_{min}$ and above $q_{max}$ is small: it amounts to less than 3% of the total value.

In the case of spherical particles, (2) and (7) yield

$$I(0) \propto (\Delta\rho)^2 N\langle R^6\rangle \quad \text{and} \quad Q \propto (\Delta\rho)^2 N\langle R^3\rangle. \qquad (8)$$

The quotient $I(0)/(QR^3)$ is expected to be invariant when the growing spherical nanocrystals are identical. This approximately occurs for the quotient $I(0)/(QR_G^3)$ determined from our experimental results plotted in Fig. 5. This result provides complementary evidence of the narrowness of the nanocrystal size distribution in the studied material.

The monotonic increase in the experimental function $Q(t)$, defined in (7), for increasing periods of time (Fig. 5) indicates that the volume fraction occupied by the nanocrystals, $N\langle v\rangle$, progressively increases. This proves that the structure evolution of the nanocrystal formation is not governed by a pure coarsening or coalescence mechanism. The increase of the integral $Q(t)$ can be caused by a progressive increase in the number $N$ of nanocrystals.

The experimental SAXS results suggest that nanocrystals nucleate and grow as a consequence of the diffusion of isolated Pb and Te elements through the glass matrix. Under this assumption, the average crystal radius, $R_A$, is expected to have the simple time

dependence given by (Liu & Risbud, 1990*a*):

$$R^2 = Kt + R(0)^2, \quad (9)$$

where $K$ is a constant and $R(0)$ is the initial nanocrystal radius.

The experimental values of $R_G$ for increasing periods of time obey the potential law of (9), as can be seen in Fig. 6. This result, together with the time dependence of the integral $Q$ (Fig. 5), is consistent with the proposed mechanism of nanocrystal growth by pure atomic diffusion up to about 1 h of heat treatment. As the concentration of Pb and Te approaches the solubility limit, coarsening becomes the predominant mechanism responsible for nanocrystal growth and the time dependence of the radius is no longer given by (9) and becomes $R \propto t^{1/3}$ (Lifshitz & Slyozov, 1959). This explains the deviation, observed in Fig. 6, of the $R^2$ *versus* $t$ plot from the linear behavior predicted by (9).

These results show that nuclei having an appreciable critical size (with average radius ~16 Å) are formed initially. The growth of the nanocrystals is governed by the classical atomic diffusion mechanism in the first stages and by coarsening in advanced stages of isothermal heat treatment.

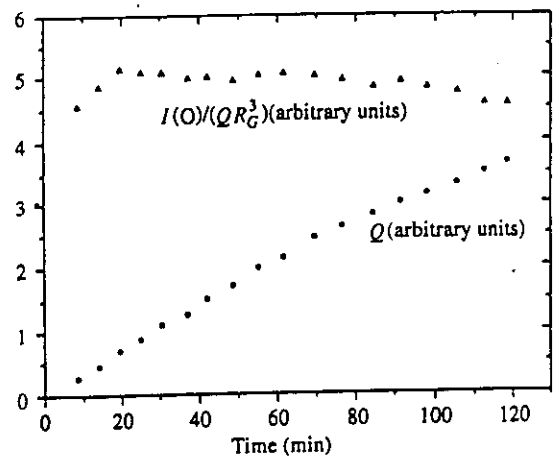


Fig. 5. Time dependence of the integral $Q(t) = \int q^2 I(q,t)dq$ and of the quotient $I(0)/(QR_G^3)$.

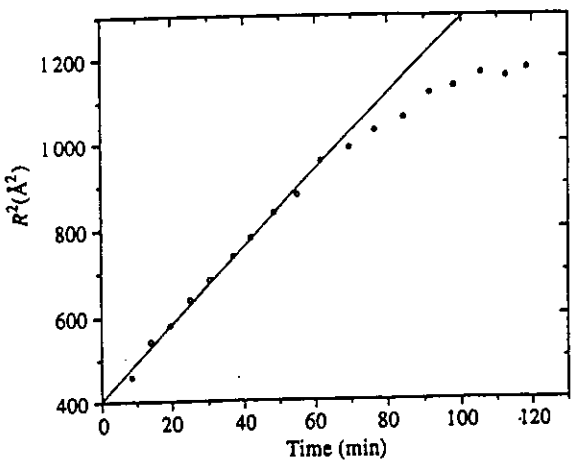


Fig. 6. Square of the Guinier average radius ($R_G^2$) as a function of time.

## 5. Conclusions

The conclusions regarding the structural evolution of Pb- and Te-doped borosilicate glass materials, quenched from 1623 K and isothermally treated at 923 K from 0 to 120 min, are:

(i) nearly spherical PbTe nanocrystals having an average radius of 16 Å are initially formed;

(ii) during the isothermal treatment the average nanocrystal radius grows to 33 Å;

(iii) the relative standard deviation of the size distribution function is about 8%;

(iv) the predominant mechanism for nanocrystal growth is classical atomic diffusion;

(v) the Efros & Efros (1982) equation does not yield a precise estimate of nanocrystal radii, probably due to the strong nonparabolicity of electron and hole energy bands.

The authors acknowledge C. L. Cesar for his useful comments on the determination of the average nanocrystal radius from the optical absorption spectra.

## References

Borrelli, N. F., Hall, D. W., Holland, H. J. & Smith, R. W. (1987). *J. Appl. Phys.* **61**, 5399–5409.

Craievich, A. F., Alves O. L. & Barbosa, L. C. (1994). *Rev. Sci. Instrum.* **66**, 1338–1341.

Efros, Al. L. & Efros, A. L. (1982). *Sov. Phys. Semicond.* 16, 772–775.

Guinier, A. & Fournet, G. (1955). *Small-Angle Scattering of X-rays*. New York: Wiley.

Ishida, A., Matsuura, S., Mizuno, M. & Fujiyasu, H. (1987). *Appl. Phys. Lett.* **51**, 478–480.

Lifshitz, I. M. & Slyozov, V. V. (1959). *Sov. Phys. JETP*, 35, 331–339.

Liu, B. C. & Risbud, S. H. (1990*a*). *J. Appl. Phys.* **68**, 28–32.

Liu, B. C. & Risbud, S. H. (1990*b*). *Philos. Mag. Lett.* 61, 327–332.

Martinez, G., Schluter, M. & Cohen, M. (1975). *Phys. Rev. B*, **11**, 651–658.

Medeiros Neto, J. A., Barbosa, L. C., Cesar, C. L., Alves, O. L. & Galembeck, F. (1991). *Appl. Phys. Lett.* **59**, 2715–2717.

Potter, B. C. & Simmons, J. H. (1988). *Phys. Rev. B*, **37**, 10838–10845.

Reynoso, V. C. S., de Paula, A. M., Cuevas, R. F., Medeiros Neto, J. A., Alves, O. L., Cesar, C. L. & Barbosa, L. C. (1995). *Electr. Lett.* **31**, 1013–1015.

*J. Appl. Cryst.* (1997). **30**, 750–754

# Thermostimulated Sol–Gel Transition in Suspensions of Sulfate–Zirconium Oxychloride

LEILA A. CHIAVACCI,[a] CELSO V. SANTILLI,[a]* SANDRA H. PULCINELLI[a] AND ALDO F. CRAIEVICH[b,c]

[a]*Institute of Chemistry, UNESP, Araraquara, SP, Brazil,* [b]*National Synchrotron Light Laboratory, CNPq, Campinas, SP, Brazil, and* [c]*Institute of Physics, USP, São Paulo, SP, Brazil. E-mail: santilli@iq.unesp.br*

(*Received 23 July 1996; accepted 29 January 1997*)

**Abstract**

The sol–gel transition of an aqueous solution of zirconium oxychloride modified by sulfuric acid (Zr:S 3:1) has been studied by small-angle X-ray scattering (SAXS) performed *in situ* during one cycle of heating and cooling between 298 and 358 K. The experimental SAXS curves exhibit three regions, at small, medium and high $q$ values, characteristic of the Guinier, 'fractal' and Porod regimes, respectively. The value of 5.5 Å for the radii of the primary particles, obtained from the cross over of the Porod and fractal regimes, is consistent with the size of the inner core of the polynuclear $Zr_{18}O_4(OH)_{38.8}(SO_4)_{12.6}.33H_2O$ molecule. These molecules aggregate as small clusters (31 Å) of fractal structure, with a dimensionality $D = 2.16$ in the sol. The value of the fractal dimensionality, which is characteristic of ideal branched polymers, decreases during heating, attaining a value of 2.0, characteristic of swollen polymers. Above the critical temperature (329 K), the average size of aggregates increases abruptly to 200 Å and the fractal dimensionality decreases to $D = 1.75$, this value being characteristic of a cluster–cluster aggregation process.

## 1. Introduction

Zirconia-based ceramic materials have been obtained through the sol–gel process, aiming at improving their thermomechanical, electrical, optical and catalytic properties (Bartlett, Woolfrey, Percy, Spiccia & West, 1994). Zirconia gels are prepared by hydrolysis and condensation of alkoxides such as zirconium *n*-propoxide or ethoxide. The zirconium alkoxide precursors are highly reactive and these reactions have to be controlled by modifying the precursors using complexing ligands (Chatry, Henry, In, Sanchez & Livage, 1994). There is some evidence (Henry, Jolivet & Livage, 1992) that mineral complexing ligands, like $SO_4^{2-}$, $PO_4^{3-}$ and $NO_3^-$, can also control the condensation reactions of aqueous zirconium solutions prepared from metal halides. While studying the possibility of controlling these reactions by adding a mineral complexing ligand, we discovered that an aqueous solution of zirconium oxychloride in the presence of sulfate exhibits a thermoreversible sol–gel transition.

Thermoreversible gels are two-phase systems which acquire elastic properties under temperature changes as a result of the formation of a connected network by physical linkages (Djabourov, 1987). Thermoreversible gelation is often observed in biological systems and mixtures of synthetic polymers, but this behaviour is not usual for inorganic gels. To our knowledge, only one article reports the formation of an inorganic thermoreversible gel, namely of aluminium and sodium polyphosphate, which was attributed to the formation of a network of polycations and polyanions (Lima & Galembeck, 1994).

In this paper we present the first SAXS study concerning the thermoreversible sol–gel transition of zirconium oxychloride modified by sulfuric acid. Structural parameters, such as the fractal dimensionality and the upper and lower cut-off lengths, are determined from the SAXS results.

## 2. The structure function for fractals

Many investigations have clearly established that the structures of a number of gels and aerogels can be described by the concept of fractals (Chaput, Boilot, Dauger, Devreux & De Geyer, 1990; Freltoft, Kjems & Sinha, 1986; Teixeira, 1988). The structure function and the SAXS intensity, $I(q)$, which is proportional to the structure function, exhibits the following remarkably simple behaviour:

$$I(q) = Kq^{-\alpha} \qquad (1)$$

where $q$ is the modulus of the scattering vector and $\alpha$ and $K$ are constants; the exponent $\alpha = D$, where $D$ is the fractal dimensionality of the mass fractal object. The infinite $q$ range for which (1) holds only applies to ideal fractals having extremely large size and very small basic units. In real systems, the potential dependence of $I(q)$ [(1)] is limited both at large and at small $q$ values because of the finite size or correlation length of the aggregates, $\xi$, and because of the non-zero size of the elementary particles, $a$, respectively. The effect of these cut-offs is the appearance of two crossovers in the scattering curve: the first at $q = 1/\xi$, between the Guinier and the fractal regimes, and the second at $q = 1/a$, between the fractal and the Porod regions.

© 1997 International Union of Crystallography
Printed in Great Britain – all rights reserved

The scattering intensity function $I(q)$ is given by

$$I(q) = \varphi P(q) S(q) \tag{2}$$

where $S(q)$ is the structure function, $P(q)$ is the form factor of the elementary units and $\varphi$ is the number of scattering units per unit volume.

Freltoft, Kjems & Sinha (1986) and Teixeira (1988) established the structure function for a fractal composed of elementary particles of size $a$ having a cut-off for the correlation represented by a function $C(r) \propto \exp(-r/\xi)$. The dependence of $S(q)$ on $D$, $\xi$ and $a$ is given by

$$\begin{aligned} S(q) = C\big(1 + \{[1/(qa)^D] \\ \times \{[D\Gamma(D-1)]/[1 + (1/q^2\xi^2)]^{(D-1)/2}\} \\ \times \sin[(D-1)\tan^{-1}(q\xi)]\}\big) \end{aligned} \tag{3}$$

where $\Gamma$ is the gamma function and $C$ is a constant. Note that this expression becomes $S(q) \simeq q^D$ when $\xi^{-1} \ll q \ll a^{-1}$. At small $q$, $S(q)$ exhibits Guinier-type behaviour (Chaput, Boilot, Dauger, Devreux & De Geyer, 1990). The form factor was approximated by a Gaussian function assuming spherical-shaped elementary particles:

$$P(q) \simeq \exp[-(1/5)q^2a^2]. \tag{4}$$

Equation (3) holds for $q < 1/a$. For $q \gg 1/a$ the SAXS intensity exhibits the classical Porod asymptotic behaviour.

## 3. Sample preparation

The colloidal suspensions were prepared by adding an aqueous solution of $ZrOCl_2$ (0.2 mol $l^{-1}$, room temperature) to a hot (353 K) aqueous solution of sulfuric acid (0.2 mol $l^{-1}$), drop by drop under magnetic stirring. The volume of the $ZrOCl_2$ solution was adjusted to yield suspensions containing the following $Zr^{4+}$:$HSO_4^-$ molar ratios: 1:1, 1:2, 1:3, 2:1, 3:1 and 4:1. Aliquots of 15 ml of the suspensions were put inside acetylcellulose membrane tubing with molecular weight cut-off smaller than 14 000, and then submitted to static dialysis against 200 ml of doubly distilled water during 24 h. Dialysis was carried out in order to decrease the ionic strength of the suspensions by extracting the excess of ions; consequently, peptization of the precipitate was favoured. Peptization occurs only for samples prepared with molar ratios of 3:1 and 2:1, leading to the formation of transparent and stable suspensions, which gelify as the pH increases from 1.6 to 3.0. Elemental analysis of these samples showed that the nominal molar ratio was preserved after dialysis. We have also noted that the isocoric gelation may be obtained by increasing the temperature of the sol prepared with a molar ratio of 3:1. Moreover, the liquefaction of the system is observed as the temperature decreases to values smaller than the critical temperature. This thermoreversible behaviour is observed if the system is kept at temperatures higher than the critical temperature for a short period of time (less than 1 h).

## 4. SAXS measurements

The SAXS study was performed using the D24 SAXS workstation at the synchrotron X-ray source DCI at LURE, France. The workstation is equipped with a bent silicon crystal monochromator which yields a monochromatic ($\lambda$ = 1.49 Å) and horizontally focused beam. A set of slits defines the beam vertically. A position-sensitive X-ray detector and a multichannel analyser were used to determine the SAXS intensity $I(q)$ as a function of the modulus of the scattering vector $q$. Because of the small size of the incident-beam cross section at the detection plane, no mathematical desmearing of the experimental function was needed. The scattered intensity was determined as the difference between the experimental curve recorded with the sample and the parasitic scattering curve obtained with the solvent (water). An ionization chamber, placed downstream from the sample, was used to monitor the intensity decay of the transmitted beam and to determine the sample attenuation. The natural decay intensity of the incident beam was monitored also by recording the electronic current in the synchrotron source.

*In situ* SAXS measurements were carried out during one cycle of heating and cooling (rate 1 K $min^{-1}$) between 303 and 358 K. Each spectrum was recorded during 150 s, which corresponds to a temperature rise of 1.5 K. The sample was put into a cell between two thin films of Kapton, separated by 0.8 mm, placed in a Metler FP82 'hot stage' to control the thermal cycle.

## 5. Results and discussion

The thermoreversibility of this system was evidenced by turbidimetric measurements carried out at different rates of heating and cooling, as shown in Fig. 1, in which gelation during heating is characterized by an abrupt increase in the turbidity at the sol–gel transition temperature. During cooling, the decrease in turbidity revealed the liquefaction, the temperature of which (gel–sol transition) is always lower than the gelation temperature, resulting in a hysteresis loop. This behaviour is associated with differences in the mobilities of the small and large particles. The small particles in the sol have high mobility, favouring aggregate growth. On the other hand, in the large aggregates in the gel, the interaction forces between the clusters make liquefaction (disaggregation) more difficult. It is noteworthy that fast rates of cooling prevent the system from returning to the sol state (Fig. 1*b*). The *in situ* SAXS measurements were carried out at 1 K $min^{-1}$. In order to obtain complete reversibility, a much slower cooling rate is necessary.

This slow type of cycle was not tried because of the limited time available for the experimental run.

Fig. 2 shows the log–log plot of the scattered intensity $I(q)$ as a function of the modulus of scattering vector $q$, measured at different temperatures during sample heating (Fig. 2*a*) and cooling (Fig. 2*b*). For the region of high $q$ ($q > 0.05$ Å$^{-1}$) the curves measured at different temperatures are approximately overlapping and two power-law regimes can be distinguished. The first one, at high $q$, presents a slope of $-4$, in agreement with the Porod law. This behaviour is confirmed by the plot presented in Fig. 3, which shows that the product $I(q)q^4$ is a constant in the high $q$ region, either for the sol or the gel, indicating that the primary particles have well defined surfaces. The Porod limit, which is proportional to the surface area of the interface, is nearly invariant during the sol–gel transition. This suggests that aggregation occurs by simple branching with local contacts between the primary particles and without the contribution of any density-increasing or phase-separation mechanisms, which would produce a noticeable variation in the interface area.

The second power-law regime at medium $q$ produces a linear behaviour in the double logarithmic plots (Fig. 2). The slope is equal to $-2.2$ at room temperature and decreases as the temperature is raised. Moreover, the $q$ range of the linear region increases toward small $q$ values for higher temperatures, indicating an increase in size of the fractal aggregates. A third regime is observed at small $q$ values where the scattered intensity is approximately independent of $q$, which is characteristic of noninteracting clusters. This feature is expected for fractal clusters with finite correlation range which behave as independent objects.

Fig. 4 shows the fitting for the data displayed in Fig. 2 with the theoretical function $I(q)$ given by (2) using $S(q)$ and $P(q)$ defined by (3) and (4), respectively. The parameter $a$ was determined from the crossover between the fractal and Porod regions. The other parameters of the theoretical function were determined by a nonlinear least-squares refinement procedure. The theoretical equation gives, as expected, a good agreement with the experimental observations only below $1/a$.

The $q$ value of the crossover of the two power regimes remains approximately constant during the sol–gel–sol transition and corresponds to $a = 5.5$ Å for the size of the primary particles. This value is in good agreement with the radius of the inner core of a molecule of $Zr_{18}O_4(OH)_{38.8}(SO_4)_{12.6}.33H_2O$ (Squattrito, Rudolf & Clearfield, 1987), which is formed under conditions of synthesis similar to those used in the present study. In this structure, the inner core of the molecule is formed by

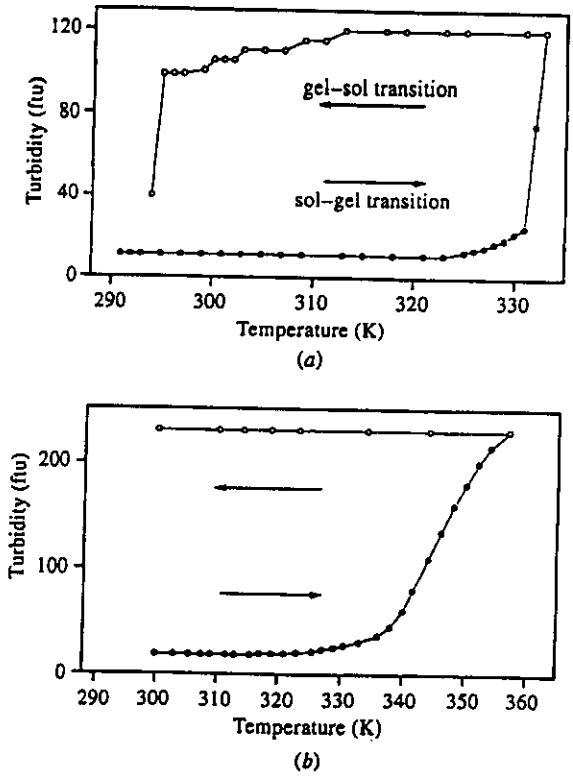


Fig. 1. Temperature evolution of the turbidity of the colloidal suspension of zirconium oxychloride modified by sulfuric acid (Zr:S 3:1 and [Zr] = 0.15 mol l$^{-1}$) during a cycle of heating and cooling at (*a*) 0.01 and (*b*) 1 K min$^{-1}$.

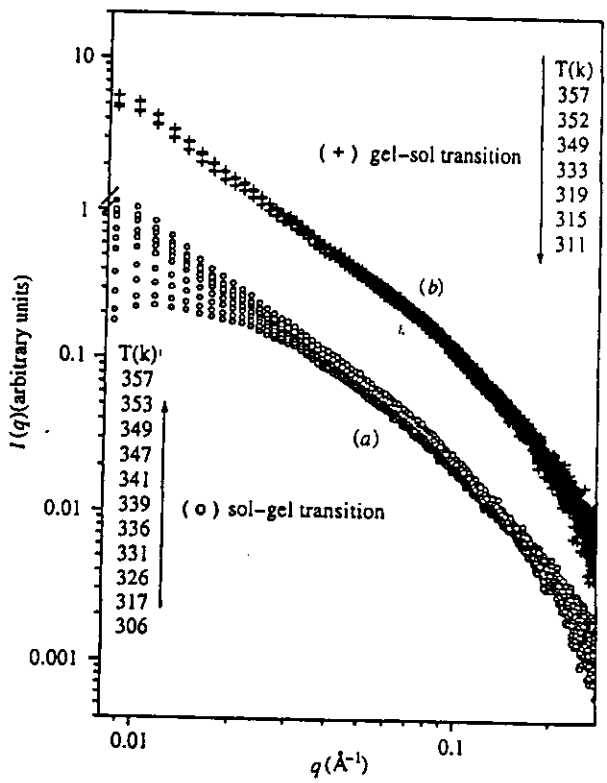


Fig. 2. Log–log plot of the experimental SAXS curves measured at different temperatures during (*a*) heating and (*b*) cooling. The set of SAXS curves corresponding to cooling are vertically displaced for clarity.

hydroxy and oxo groups bridging Zr atoms, whereas the sulfate ligands are located at the surface of the $Zr(OH)_xO_y$core. These molecules are held together by electrostatic interactions with the remaining water molecules.

Fig. 5 shows the temperature evolution of the fractal dimensionality $D$ (Fig. 5*a*) and the correlation length of the aggregates $\xi$ (Fig. 5*b*) during the sol–gel–sol transition. Note that, during cooling, the values of both $D$ and $\xi$ are not the same as during heating, indicating a strong hysteresis effect. This was expected from the turbidimetric measurements at the same heating rate.

Both sol and gel can be characterized by a fractal structure model, the dimensionality $D$ decreasing progressively from 2.2 to about 2.0 in the sol phase and remaining constant and equal to 1.75 in the gel phase.

The value of $D$ of 2.2 at room temperature is theoretically expected for ideal branched polymers, as is $D = 2.0$ for swollen branched polymers (Chaumont, Craievich & Zarzycki, 1992*a*,*b*, and references therein). This implies that in solution at room temperature, ideal branching occurs with monomers interlinked by a matrix formed by water, similar to the crystalline structure of $Zr_{18}O_4(OH)_{38.8}(SO_4)_{12.6}.33H_2O$. In solution, these interactions relax as the temperature increases due to the diminution of the dielectric constant of water. This leads to swelling of the cluster, which produces a decrease in the fractal dimensionality from 2.2 to 2.0 and an increase in the cluster correlation length from 31 to 46 Å as the temperature increases from 298 to 328 K. However, the change in $\xi$ seems to be small when compared to the large relative variation in the fractal dimensionality, suggesting that relaxation involves only small and local structural modifications.

At about 333 K, the size of aggregates, $\xi$, increases abruptly and the fractal dimensionality decreases from 2.0 to 1.75 and remains almost constant up to 343 K. The value of $D$ of 1.75 is close to the value expected for the mechanism of random-walk cluster–cluster aggregation (Chaumont, Craievich & Zarzycki, 1992*a*,*b*, and references therein). The enhanced mobility of the clusters and the change of the electrostatic interactions as the temperature increases may play a role in the modification of the aggregation mechanism above the gelation temperature.

The thermoreversible behaviour observed for the colloidal suspensions of zirconium oxychloride modified by sulfuric acid is qualitatively similar to that observed in some polymers in solution with lower critical solution temperature (LCST), in which phase separation and gelation occurs above a critical temperature. This behaviour is often assigned to a decrease in the

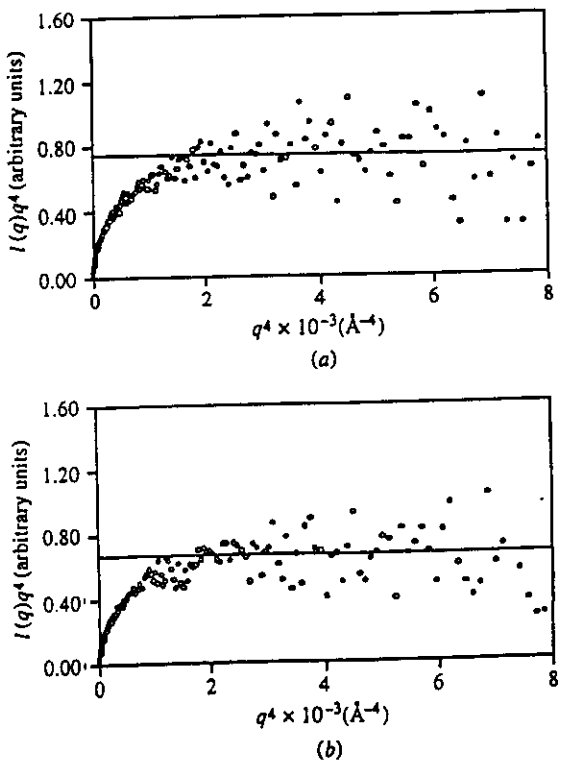


Fig. 3. Porod plots corresponding to (*a*) the gel phase ($T$ = 343 K) and (*b*) the sol phase ($T$ = 303 K). In spite of the large statistical errors in this $q$ range, the Porod asymptotic behaviour seems to be obeyed for both sol and gel.

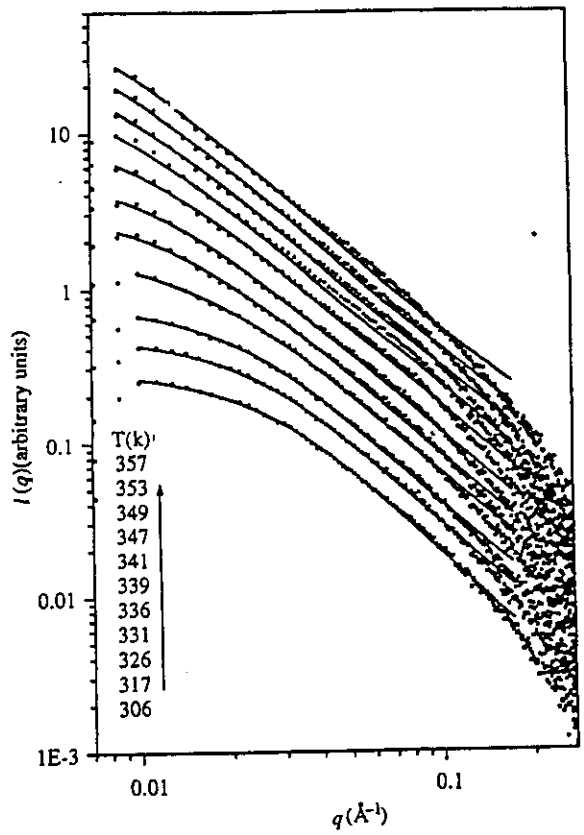


Fig. 4. Experimental points and fitting of the structure function $S(q)$ for the SAXS curves. The curves are vertically displaced for clarity.

effectiveness of the solvent, causing a decrease in the volume entropy due to the density-increase mechanism of the polymers (Sperling, 1992). Moreover, in the case of a transition from the solution to a phase-separated structure, a significant variation in the surface area is expected.

In contrast, our system shows swelling of the fractal cluster with increasing temperature, indicating that the effectiveness of the solvent increases. Moreover, the Porod limit does not indicate any variation in the surface area (Fig. 3) during the sol–gel transition. Therefore, it is difficult to explain our results from the thermodynamic concepts of LCST.

It seems more consistent to attribute the sol–gel–sol transition to a change in the affinity between the solvent and the clusters. In general, the enthalpy of hydration becomes less negative as the temperature increases. As a consequence, the solvation tends to decrease during heating, favouring the formation of more efficient electrostatic interactions between the $Zr_{18}O_4(OH)_{38.8}(SO_4)_{12.6}$ molecules by means of sulfate bridging.

In fact, we have observed by IR spectroscopy the presence of a bidentate $SO_4$ ligand in the irreversible gel, obtained after ageing at 353 K for 3 h (Chiavacci, 1996). Therefore, the thermoreversibility seems to be a consequence of an intermediate step involving the formation of a sulfate complex, in which the short-lived nodes of the network are due to the electrostatic interaction between the clusters.

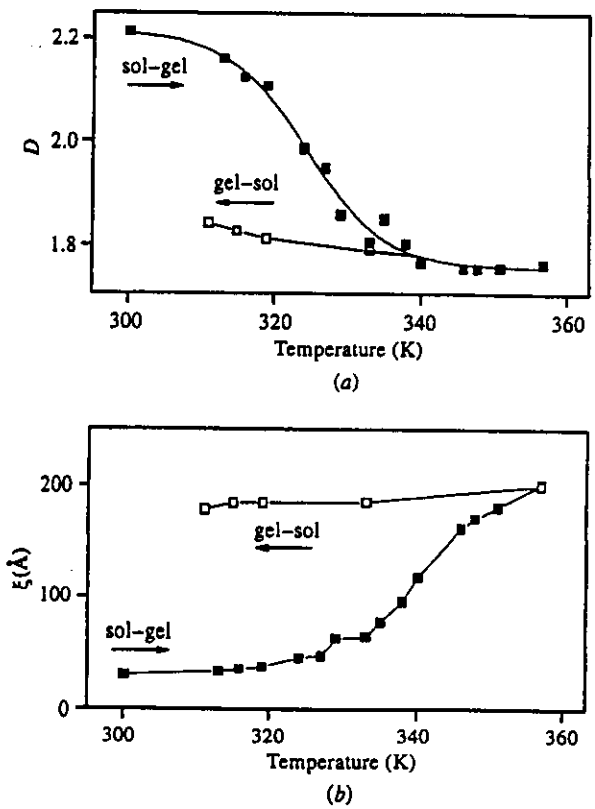


Fig. 5. Temperature dependence of (*a*) the fractal dimensionality $D$ and (*b*) the size of aggregates $\xi$, obtained from the fittings using (2).

## 6. Conclusion

An aqueous solution of zirconium oxychloride modified by sulfuric acid exhibits a thermoreversible sol–gel transition, as evidenced by turbidimetric measurements. This SAXS study, during a cycle of heating and cooling, demonstrates that both the sol and the gel are formed by clusters or aggregates having a fractal structure. The size of the elementary structural unit is similar to that of the inner core of the $Zr_{18}O_4(OH)_{38.8}(SO_4)_{12.6}.33H_2O$ molecule. The size of the basic building block does not vary along the whole heating–cooling cycle.

As temperature increases, the initial clusters of perfect branched polymers (with $D = 2.2$) in the sol phase progressively swell (with $D$ decreasing to 2.0).

Gelation is produced by the enhanced electrostatic interaction between the clusters resulting in a network with a fractal dimensionality $D$ of 1.75, which is consistent with the value expected for the mechanism of cluster–cluster aggregation.

This work was supported by CNPq, CAPES/COFECUB and FAPESP (Brazil).

### References

Bartlett, J. R.,Woolfrey, J. L., Percy, M., Spiccia, L. & West, B. O. (1994). *J. Sol–Gel Sci. Technol.* **2**, 215–220.
Chaput, F., Boilot, J. P., Dauger, A., Devreux, F. & De Geyer, A. (1990). *J. Non-Cryst. Solids*, **116**, 133–139.
Chatry, M., Henry, M., In, M., Sanchez, C. & Livage, J. (1994). *J. Sol–Gel Sci. Technol.* **1**, 233–240.
Chaumont, D., Craievich, A. & Zarzyćki, J. (1992*a*). *J. Non-Cryst. Solids*, **147/148**, 41–46.
Chaumont, D., Craievich, A. & Zarzycki, J. (1992*b*). *J. Non-Cryst. Solids*, **147/148**, 127–134.
Chiavacci, L. A. (1996). Masters thesis, UNESP, Brazil.
Djabourov, M. (1987). *Rev. Gen. Therm. Fr.* **306/307**, 369–373.
Freltoft, T., Kjems, J. K. & Sinha, S. K. (1986). *Phys. Rev. B*, **33**, 269–275.
Henry, M., Jolivet, J. P. & Livage, J. (1992). *Structure and Bonding*, Vol. 77, edited by R. Reisfeld & C. K. Jorgensen, pp. 153–203. Berlin: Springer-Verlag.
Lima, E. C. O. & Galembeck, F. (1994). *J. Collect. Interface Sci.* **166**, 309–315.
Sperling, L. H. (1992). *Introduction to Physical Polymer Science*, pp. 127–130. New York: Wiley Interscience.
Squattrito, P. J., Rudolf, P. R. & Clearfield, A. (1987). *Inorg. Chem.* **26**, 4240–4244.
Teixeira, J. (1988). *J. Appl. Cryst.* **21**, 781–785.

*J. Appl. Cryst.* (1995). **28**, 768–773

# *CRYSOL* – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates

BY D. SVERGUN,* C. BARBERATO AND M. H. J. KOCH

*European Molecular Biology Laboratory, Hamburg Outstation, Notkestrasse 85, D-22603 Hamburg, Germany*

(*Received 7 March 1995; accepted 24 May 1995*)

## Abstract

A program for evaluating the solution scattering from macromolecules with known atomic structure is presented. The program uses multipole expansion for fast calculation of the spherically averaged scattering pattern and takes into account the hydration shell. Given the atomic coordinates (*e.g.* from the Brookhaven Protein Data Bank) it can either predict the solution scattering curve or fit the experimental scattering curve using only two free parameters, the average displaced solvent volume per atomic group and the contrast of the hydration layer. The program runs on IBM PCs and on the major UNIX platforms.

## Introduction

Comparisons between experimental X-ray solution scattering [small-angle X-ray scattering (SAXS)] curves and those evaluated from crystallographic structures have been widely used to validate theoretical models, to verify the structural similarity between proteins and nucleic acids in crystals and in solution and to predict quaternary structures (see *e.g.* Langridge *et al.*, 1960; Ninio, Luzzati & Yaniv, 1972; Fedorov, Ptitsyn & Voronin, 1972; Fedorov & Denesyuk, 1978; Müller, 1983; Pavlov, Sinev, Timchenko & Ptitsyn, 1986; Grossmann *et al.*, 1993). Moreover, for multisubunit macromolecules, accurate evaluation of the solution scattering amplitudes from the atomic coordinates of separate domains allows one to determine their relative positions from the SAXS data (Svergun, 1991, 1994).

The main problem in evaluating the solution scattering from atomic coordinates is to adequately take into account the solvent scattering. Several methods have been developed that basically differ in the representation of the particle volume inaccessible to the solvent. In the effective-atomic-scattering-factor method (*e.g.* Langridge *et al.*, 1960; Fraser, MacRae & Suzuki, 1978; Lattman, 1989), the excluded volume is built by dummy solvent atoms placed at the positions of the atoms in the macromolecule. This approach is well justified at resolutions down to 1–2 nm [i.e. in the range of momentum transfer $0 \lesssim s \lesssim 3$ nm$^{-1}$, $s = (4\pi/\lambda) \sin \theta$, $2\theta$ is the scattering angle, $\lambda$ the wavelength]. At higher resolution, the inhomogeneous filling of the excluded volume may introduce systematic deviations. The cube method (Fedorov, Ptitsyn & Voronin, 1972; Ninio, Luzzati & Yaniv, 1972) and its modifications (Müller, 1983; Pavlov & Fedorov, 1983) homogeneously fills this volume with cubic elements and thus provides better accuracy at higher scattering vectors ($s \gtrsim 3$ nm$^{-1}$).

The above methods do not take into account the hydration shell surrounding macromolecules in solution. Omission of this shell can lead to significant systematic errors even near the origin of the scattering curves and therefore to misinterpretation of the results. Attempts to include the hydration shell have been made, *e.g.*, by Hubbard, Hodgson & Doniach (1988) and Grossmann *et al.* (1993), but no general-purpose program has been developed. *CRYSOL*, the program described below, evaluates the SAXS profiles from crystallographic structures taking into account the scattering from the hydration shell.

## Theory

Macromolecules in solution can be schematically represented as illustrated in Fig. 1. The particle with scattering density $\rho_a(\mathbf{r})$ is surrounded by a solvent with an average scattering density $\rho_0$. The hydration shell is approximated by a border layer of effective thickness $\Delta$ and density $\rho_b$ that may differ from $\rho_0$. The SAXS intensity from such particles in dilute solution is proportional to the averaged scattering of a single particle:

$$I(s) = \langle |A_a(\mathbf{s}) - \rho_0 A_c(\mathbf{s}) + \delta\rho A_b(\mathbf{s})|^2 \rangle_\Omega, \qquad (1)$$

where $A_a(\mathbf{s})$ is the scattering amplitude from the particle *in vacuo*, $A_c(\mathbf{s})$ and $A_b(\mathbf{s})$ are, respectively, the scattering amplitudes from the excluded volume and the border layer, both with unitary density, $\delta\rho = \rho_b - \rho_0$, and $\langle\ \rangle_\Omega$ stands for the average over all particle orientations [$\Omega$ is the solid angle in reciprocal space, $\mathbf{s} = (s, \Omega)$].

* On leave from the Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia.

© 1995 International Union of Crystallography
Printed in Great Britain – all rights reserved

Spherical averaging in (1) is greatly facilitated by use of the multipole expansion (Stuhrmann, 1970*a*; Lattman, 1989). For the atomic coordinates $\mathbf{r}_j = (r_j, \omega_j) = (r_j, \theta_j, \varphi_j)$ and the corresponding atomic form factors $f_j(s)$, the scattering amplitude *in vacuo* of a particle consisting of $N$ atoms is

$$A_a(\mathbf{s}) = \sum_{j=1}^{N} f_j(s) \exp(i\mathbf{s}\mathbf{r}_j). \quad (2)$$

Substituting the relation (Edmonds, 1957)

$$\exp(i\mathbf{s}\mathbf{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(sr) Y^*_{lm}(\omega) Y_{lm}(\Omega), \quad (3)$$

where the $j_l(sr)$ are the spherical Bessel functions and the $Y_{lm}(\Omega)$ are the spherical harmonics, into (2), one can write

$$A_a(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}(s) Y_{lm}(\Omega), \quad (4)$$

where $A_{lm}(s)$ are the partial amplitudes:

$$A_{lm}(s) = 4\pi i^l \sum_{j=1}^{N} f_j(s)\, j_l(sr_j) Y^*_{lm}(\omega_j). \quad (5)$$

With the excluded volume represented as a superposition of dummy atoms with form factors $g_j(s)$ centered at the same coordinates $\mathbf{r}_j$, the amplitude $A_c(\mathbf{s})$ is expressed in the form of (4) with the partial amplitudes

$$C_{lm}(s) = 4\pi i^l \sum_{j=1}^{N} g_j(s)\, j_l(sr_j) Y^*_{lm}(\omega_j). \quad (6)$$

Following Stuhrmann (1970*b*), the border layer can be described by a two-dimensional angular function $F(\omega)$ (Fig. 1) as

$$\rho_b(\mathbf{r}) = \begin{cases} 1 & F(\omega) \le r \le F(\omega) + \Delta \\ 0 & 0 < r < F(\omega) \text{ or } r > F(\omega) + \Delta. \end{cases} \quad (7)$$

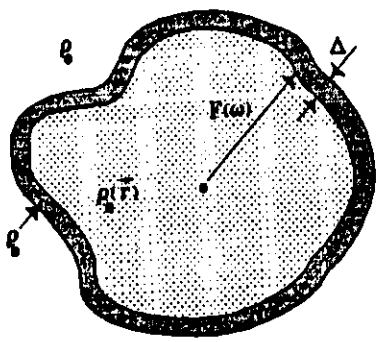


Fig. 1. Schematic representation of a macromolecule in solution. For explanations see text.

As the partial amplitudes are the Hankel transforms of the real-space radial functions:

$$B_{lm}(s) = i^l (2/\pi)^{1/2} \int_0^{\infty} \rho_{lm}(r) j_l(sr) r^2\, dr, \quad (8)$$

where

$$\rho_{lm}(r) = \int_{\omega} \rho_b(\mathbf{r}) Y^*_{lm}(\omega)\, d\omega, \quad (9)$$

it is readily shown that

$$B_{lm}(s) = i^l (2/\pi)^{1/2} \int_{\omega} Y^*_{lm}(\omega)\, d\omega \int_{F(\omega)}^{F(\omega)+\Delta} j_l(sr) r^2\, dr. \quad (10)$$

By (4)–(6) and (10), the three amplitudes entering into (1) are represented *via* their multipole components. Owing to the orthogonal properties of the spherical harmonics, all cross terms cancel out in the average over $\Omega$, leading to a simple expression for the SAXS intensity:

$$I(s) = \sum_{l=0}^{L} \sum_{m=-l}^{l} |A_{lm}(s) - \rho_0 C_{lm}(s) + \delta\rho B_{lm}(s)|^2, \quad (11)$$

where the truncation value $L$ defines the resolution of the representation of the particle.

## Evaluation of the partial amplitudes

The form factors $f_j(s)$ and $g_j(s)$ necessary for the evaluation of the partial amplitudes $A_{lm}(s)$ and $C_{lm}(s)$ are specified as follows. Since usually only the coordinates of the non-H atoms are available in the Protein Data Bank (Bernstein *et al.*, 1977), the summation in (2) runs only over these atoms, but those having covalently bound hydrogens are replaced by the corresponding atomic groups according to Table 1. A similar approach was used earlier by Pavlov & Fedorov (1983) and Lattman (1989). The spherically averaged form factors $f_j(s)$ of each atomic group are tabulated on the interval $0 \le s \le 10$ nm$^{-1}$ using a five-Gaussian approximation of the form factors of the individual atoms (*International Tables for X-ray Crystallography*, 1974) and the interatomic distances.

The form factors of the dummy atoms entering (6) are expressed as (Fraser, MacRae & Suzuki, 1978)

$$g_j(s) = G(s) V_j \exp(-\pi s^2 V_j^{2/3}), \quad (12)$$

where $V_j = (4\pi/3) r_{wj}^3$ is the solvent volume displaced by the $j$th atom, or atomic group, represented by the Gaussian sphere of radius $r_{wj}$ (see Table 1), and $G(s)$ is an overall expansion factor

$$G(s) = (r_0/r_m)^3 \exp[-(4\pi/3)^{3/2} \pi s^2 (r_0^2 - r_m^2)]. \quad (13)$$

Here, $r_m = N^{-1} \sum_{j=1}^{N} r_{wj}$ is the actual average radius of the atomic group (for proteins, $r_m$ is normally around

Table 1. *Parameters of the atomic groups and heteroatoms*

| Atom or atomic group | Displaced solvent volume ($nm^3$) | Radius (nm) |
|---|---|---|
| H* | 0.00515 | 0.107 |
| C* | 0.01644 | 0.158 |
| CH† | 0.02159 | 0.173 |
| $CH_2$† | 0.02674 | 0.185 |
| $CH_3$† | 0.03189 | 0.197 |
| N* | 0.00249 | 0.084 |
| NH† | 0.00764 | 0.122 |
| $NH_2$† | 0.01279 | 0.145 |
| $NH_3$ | 0.01794 | 0.162 |
| O* | 0.00913 | 0.130 |
| OH† | 0.01428 | 0.150 |
| S‡ | 0.01986 | 0.168 |
| SH† | 0.02510 | 0.181 |
| Mg‡ | 0.01716 | 0.160 |
| P‡ | 0.00573 | 0.111 |
| Ca‡ | 0.03189 | 0.197 |
| Mn‡ | 0.00920 | 0.130 |
| Fe‡ | 0.00799 | 0.124 |
| Cu‡ | 0.00878 | 0.128 |
| Zn‡ | 0.00985 | 0.133 |

* Observed displaced volumes according to Fraser, MacRae & Suzuki (1978).

† Evaluated by adding the displaced volume of corresponding hydrogens.

‡ Radii taken from *International Tables for X-ray Crystallography* (1968).

0.162 nm), and $r_0$, the effective atomic radius, is a variable parameter that can be used to change the displaced volume per atomic group and thus to adjust the total excluded volume. Note that the expansion factor (13) does not depend on the atomic positions and can be taken out of the summation in (6).

To evaluate the envelope function $F(\omega)$, the particle is moved so that its geometrical center coincides with the origin. A quasiuniform grid of $N_g$ angular directions is evaluated using Fibonacci numbers as described by Svergun (1994). Each non-H atom in the macromolecule updates the envelope function in the direction $\omega_i$ if the minimum distance between the atom and this direction is less than the sum of the atomic radius $r_{gj}$ and the radius of the water molecule, $r_w = 0.15$ nm. The updated value is

$$F(\omega_i) = \max\{F'(\omega_i), (r_{ji} + 0.5r_{gj})\}, \qquad (14)$$

where $F'(\omega_i)$ is the current value of the envelope function and $r_{ji}$ is the projection of $\mathbf{r}_j$ onto the direction $\omega_i$ (Fig. 2).

After all atoms have been sorted, $F(\omega)$ contains the distances between the origin and the particle surface for each $\omega_i$. The amplitudes $C_{lm}(s)$ are then evaluated by numerical integration of (10). The integral over $r$ does not depend on the envelope function and can be tabulated in advance.

The density of the bound solvent can differ noticeably from that of the bulk solvent within only a few ångströms distance from the surface (see *e.g.* Cheng & Schoenborn, 1990; Badger, 1993), *i.e.* the condition $F(\omega) \gg \Delta$ is usually fulfilled. This means that the contribution from the border layer depends mostly on the product $\delta\rho\Delta$ and without loss of generality one of the two parameters can be fixed. In *CRYSOL*, the effective width of the border layer is taken to be 0.3 nm to simulate the most ordered first hydration layer. The SAXS intensity (11) depends on two parameters: (i) the average displaced volume per atomic group and (ii) the contrast of the border layer. The former parameter is expressed *via* the effective atomic radius, which should not differ much from $r_m$ (we found $0.96r_m \le r_0 \le 1.04r_m$). The upper limit of the latter is $(\delta\rho)_{max} \simeq 70$ e $nm^{-3}$ and corresponds to the difference in the electron density between free and bound water molecules (Perkins, 1986).

## Program implementation

The above algorithms are implemented in the interactive Fortran77 program *CRYSOL*, which performs the following steps:

(1) The atomic coordinates are read from the data file in the Brookhaven Protein Data Bank (PDB) format into a temporary heap storage by blocks of a thousand to determine the geometrical center of the macromolecule and the origin is shifted to this point. The use of the heap storage allows one to avoid limitations on the number of atoms in the input file.

(2) A quasiuniform grid of angular directions $\omega_i$, $i = 1 \ldots N_g$ is evaluated ($N_g \le 4185$), the form factors and the integrals of Bessel functions (10) are tabulated in the range of momentum transfer and at the resolution level ($L \le 15$) specified by the user.

(3) The atomic types and coordinates are read from the PDB file into a temporary heap storage once again. For each atom or heteroatom, the atomic group and the displaced volume are identified according to Table 1 and the contributions to the partial amplitudes $A_{lm}(s)$, $C_{lm}(s)$ and the envelope function $F(\omega)$ are evaluated. O atoms belonging to water molecules are ignored.

(4) The amplitudes $B_{lm}(s)$ are evaluated from the function $F(\omega)$.

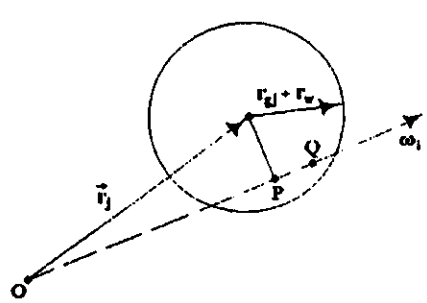


Fig. 2. Evaluation of the envelope function. The current atomic coordinate relative to the origin $O$ is $\mathbf{r}_j$, the current direction $\omega_i$, $|OP| = r_{ji}$, $|PQ| = 0.5r_{gj}$, $F(\omega_i) = |OQ|$. For details see text.

(5) The SAXS intensity, $I(s, r_0, \delta\rho)$ is calculated using (11) for a value of $\rho_0 = 334$ e nm$^{-3}$ corresponding to the bulk water with the default parameters, $r_0 = r_m$ and $\delta\rho = 30$ e nm$^{-3}$.

(6) If the experimental curve $I_e(s)$ is given, the parameters are adjusted so as to fit the experimental data. A plain grid search is made for $0.96\, r_m \leq r_0 \leq 1.04\, r_m$ and $0 \leq \delta\rho \leq 60$ e nm$^{-3}$ to minimize the functional

$$\chi^2(r_0, \delta\rho) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left[ \frac{I_e(s_i) - cI(s_i, r_0, \delta\rho)}{\sigma(s_i)} \right]^2, \quad (15)$$

where $N_p$ is the number of experimental points, the $\sigma(s_i)$ are the experimental errors and

$$c = \left[ \sum_{i=1}^{N_p} \frac{I_e(s_i) I(s_i, r_0, \delta\rho)}{\sigma(s_i)^2} \right] \left[ \sum_{i=1}^{N_p} \frac{I(s_i, r_0, \delta\rho)^2}{\sigma(s_i)^2} \right]^{-1} \quad (16)$$

is the scale factor. The fit is presented on a graphic display and the parameters can also be changed manually.

(7) The results (integral parameters of the particle, its envelope function, partial amplitudes, intensities and the fit to the experimental data) are stored in the corresponding ASCII and binary files. The data can be retrieved for further calculations with other parameters and/or experimental data sets.

The program is compiled on IBM PC computers using the Microsoft Fortran PowerStation 1.0 with the Phar Lap MS-DOS extender and requires DOS version 3.3 or later, 2 Mbytes of free memory (conventional + extended) and EGA/VGA/SVGA video display. Versions for the major UNIX platforms (Sun, Silicon Graphics, DEC Alpha), which use the public domain graphical package *GNUPLOT*, are also available.

## Examples of applications

To illustrate the use of *CRYSOL*, we present the results obtained for hen egg white lysozyme (molecular weight 14 KDa), which has already been used for illustrative purposes by various authors (Pickover & Engelman, 1982; Pavlov & Fedorov, 1983; Lattman, 1989). The X-ray scattering curve from a lysozyme solution (Fig. 3) was recorded using standard procedures on the X33 camera of the EMBL in HASYLAB at the Deutsches Electronen Synchrotron (DESY) in Hamburg. The coordinates of the lysozyme (Diamond, 1974) were taken from the PDB file pdb6lyz.ent. Fig. 4 displays the scattering curves $I_a(s) = \langle A_a(\mathbf{s}) \rangle^2$, $I_c(s) = \langle \rho_0 A_c(\mathbf{s}) \rangle^2$ and $I_b(s) = \langle \delta\rho A_b(\mathbf{s}) \rangle^2$ evaluated with $L = 12$ and $N_g = 2585$. In Fig. 3, the best fit to the experimental data ($\chi = 0.477$) obtained at $r_0 = 0.161$ nm (total excluded volume 17.4 nm$^3$) and $\delta\rho = 25$ e nm$^{-3}$ corresponding to a hydration of 0.4 g g$^{-1}$ (gram of $H_2O$ per gram of protein) is presented. The experimental radius of gyration is $R_g = 1.52$ (2) nm; the theoretical value is $R_{gt} = 1.48$ nm. Note that when the hydration shell (its radius of gyration is 1.88 nm) is not taken into account by fixing of $\delta\rho = 0$, the fit to the experimental data is poorer ($\chi = 0.765$, $R_{gt} = 1.43$ nm).

The results of *CRYSOL* were compared to those of the program of Pavlov & Fedorov (1983), which uses the modified cube method and the numerical average in reciprocal space. The excluded volume 16.8 nm$^3$

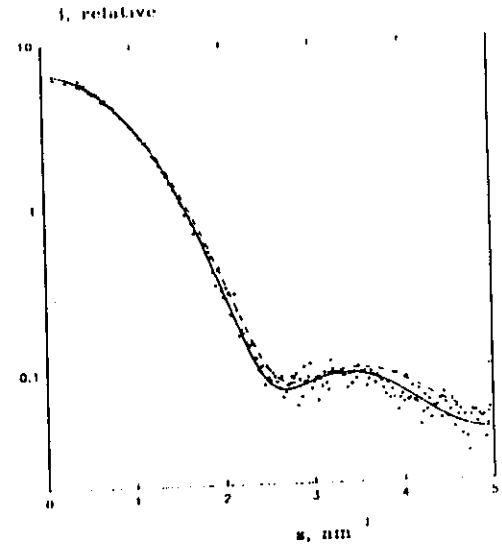


Fig. 3. Experimental solution scattering from lysozyme (dots) and the best fit by *CRYSOL* (solid line). The dashed curve indicates the best fit at $\delta\rho = 0$.

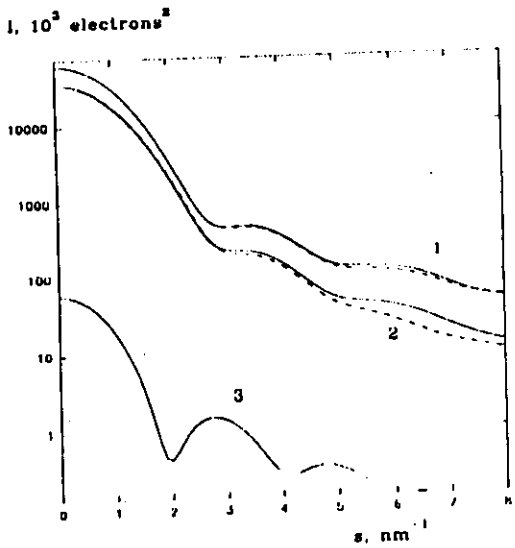


Fig. 4. Scattering from the atomic structure *in vacuo* $I_a(s)$ (1), shape scattering $I_c(s)$ (2) and scattering from the border layer $I_b(s)$ for lysozyme (solid lines). Dashed lines: curves evaluated by Pavlov's program with the shape scattering scaled to the same excluded volume.

corresponds well to the value obtained by *CRYSOL* and the curves $I_a(s)$ and $I_c(s)$ are in good agreement up to $s \simeq 4$ nm$^{-1}$. For higher angles, there are deviations in the shape scattering from *CRYSOL* owing to the inhomogeneously filled excluded volume. The scattering curve calculated by Pavlov's program, which does not take the hydration shell into account, has noticeable systematic deviations at small angles ($\chi = 0.687$, $R_{gl} = 1.47$ nm). The deviations can be reduced by artificial changing of the solvent density to $\rho_0 = 310$ e nm$^3$, which gives $\chi = 0.581$ and $R_{gl} = 1.45$ nm. The total CPU time required by *CRYSOL* on an IBM AT/486 DX50 was 310 s. In comparison, Pavlov's program, for which the user has to run three separate executable modules, requires a total of 660 s of CPU time.

We have also attempted to make a comparison with the program of Lattman (1989) which uses the effective atomic factors method and the multipole expansion. This comparison failed, apparently owing to software limitations encountered in running this program with our experimental data.

Fig. 5 illustrates the use of CRYSOL for the *E. coli* aspartate transcarbamylase (ATCase), which is a dodecamer with a molecular weight of 303 KDa. The coordinates of the *T* form of the ATCase (Kantorowitz & Lipscomb, 1988) were generated using the appropriate symmetry operations from the PDB file pdb6at1.ent. The experimental curve recorded at the small-angle scattering installation of the synchrotron-radiation laboratory LURE in Orsay, France (Herve *et al.*, 1985) yields $R_g = 4.68$ (3) nm. *CRYSOL* provides the best fit ($\chi = 1.16$, $R_{gl} = 4.64$ nm) at $r_0 = 0.168$ nm and $\delta\rho = 58$ e nm$^{-3}$ corresponding to a hydration of 0.21 g g$^{-1}$. The best fit achieved without the hydration shell at $r_0 = 0.167$ is poor ($\chi = 4.89$, $R_{gl} = 4.43$ nm) and displays a significant shift of the subsidiary maxima (such a shift has already been reported by Altman, Ladner & Lipscomb, 1982).

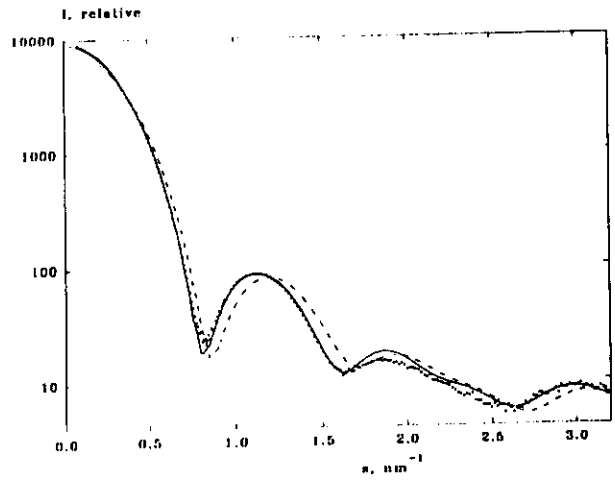


Fig. 5. Experimental solution scattering from ATCase and the fits with and without solvation shell. Notation is as in Fig. 3.

*CRYSOL* has been successfully used on a number of protein structures in ongoing projects at the EMBL Outstation (*e.g.* hexokinase, ribonucleotide reductase proteins *R*1 and *R*2 *etc.*). A beta-release of the program was also tested at the Stanford Synchrotron Radiation Laboratory (Stanford University, USA).

## Concluding remarks

The importance of the contribution of the hydration shell to the scattering has been discussed by various authors (Hubbard, Hodgson & Doniach, 1988; Schoenborn, 1988; Badger, 1993; Grossmann *et al.*, 1993). The structures available from the Protein Data Bank usually contain less than 50% of the bound waters and these can hardly be used to represent the hydration shell in solution. In fact, attempts to include these waters in the calculations on lysozyme with Pavlov's program degraded the results.

The border layer introduced in *CRYSOL* is, of course, a simplified model of the hydration shell. For macromolecules with a complicated shape, the envelope function may not be single valued and the use of $F(\omega)$ would fill the inner cavities. Although $F(\omega)$ is thus not suitable to evaluate the shape scattering itself, one is still well justified to use it for the description of the outer hydration shell. By use of the shell of a constant density and the fixed thickness of 0.3 nm, the primary solvation layer is approximated. The primary layer is known to contain the most ordered waters (see, *e.g.*, Thanki, Thornton & Goodfellow, 1988; Cheng & Schoenborn, 1990; Badger, 1993) and thus dominates the scattering from the solvation shell. For all proteins studied up to now, we found that the contribution from the border layer significantly improved the fit to the experimental data (the hydration ratio was normally 0.2–0.3 g g$^{-1}$; the value 0.4 g g$^{-1}$ for lysozyme was exceptionally high).

*CRYSOL* has been proven to adequately evaluate the SAXS profiles up to $s \lesssim 4$ nm$^{-1}$ (*i.e.* a resolution of about 1.5 nm), where the deviations due to the inhomogeneously filled excluded volume and the finite number of multipoles are negligible. At higher angles, the cube methods are expected to be more accurate.

The executable code of the program for IBM PCs and UNIX workstations and a user manual are available from the authors (e-mail svergun@embl-Hamburg.de).

The authors thank Drs M. Pavlov and E. Lattman for providing their programs, Dr A. Semenyuk for his help at the early stage of this project and Dr P. Vachette for providing the data on ATCase. This work was supported by the NATO Linkage Grant LG 921231, INTAS grant 93-645 and the CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) fellowship of C. Barberato.

## References


ALTMAN, R. B., LANDER, J. E. & LIPSCOMB, W. N. (1982). *Biochem. Biophys. Res. Commun.* **108**, 592–595.

BADGER, J. (1993). *Biophys. J.* **65**, 1656–1659.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

CHENG, X. & SCHOENBORN, B. P. (1990). *Acta Cryst.* B**46**, 195–208.

DIAMOND, R. (1974). *J. Mol. Biol.* **82**, 371–391.

EDMONDS, A. R. (1957). *Angular Momentum in Quantum Mechanics*. Princeton Univ. Press.

FEDOROV, B. A. & DENESYUK, A. I. (1978). *J. Appl. Cryst.* **11**, 473–477.

FEDOROV, B. A., PTITSYN, O. B. & VORONIN, L. A. (1972). *FEBS Lett.* **28**, 188–190.

FRASER, R. D. B., MACRAE, T. P. & SUZUKI, E. (1978). *J. Appl. Cryst.* **11**, 693–694.

GROSSMANN, G., ABRAHAM, Z. H. L., ADMAN, E. T., NEU, M., EADY, R. R., SMITH, B. E. & HASNAIN, S. S. (1993). *Biochemistry*, **32**, 7360–7366.

HERVÉ, G., MOODY, M. F., TAUC, P., VACHETTE, P. & JONES, P. T. (1985). *J. Mol. Biol.* **185**, 189–199.

HUBBARD, S. R., HODGSON, K. O. & DONIACH, S. (1988). *J. Biol. Chem.* **263**, 4151–4158.

*International Tables for X-ray Crystallography* (1968). Vol. III. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

*International Tables for X-ray Crystallography* (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

KANTOROWITZ, E. R. & LIPSCOMB, W. N. (1988). *Science*, **241**, 669–674.

LANGRIDGE, R., MARVIN, D. A., SEEDS, W. E., WILSON, H. R., HOOPER, C. W., WILKINS, M. H. F. & HAMILTON, L. D. (1960). *J. Mol. Biol.* **2**, 38–64.

LATTMAN, E. E. (1989). *Proteins*, **5**, 149–155.

MÜLLER, J. J. (1983). *J. Appl. Cryst.* **16**, 74–82.

NINIO, J., LUZZATI, V. & YANIV, M. (1972). *J. Mol. Biol.* **71**, 217–229.

PAVLOV, M. YU. & FEDOROV, B. A. (1983). *Biopolymers*, **22**, 1507–1522.

PAVLOV, M. YU., SINEV, M. A., TIMCHENKO, A. A. & PTITSYN, O. B. (1986). *Biopolymers*, **25**, 1385–1397.

PERKINS, S. J. (1986). *Eur. J. Biochem.* **157**, 169–180.

PICKOVER, C. A. & ENGELMAN, D. M. (1982). *Biopolymers*, **21**, 817–831.

SCHOENBORN, B. P. (1988). *J. Mol. Biol.* **201**, 741–749.

STUHRMANN, H. B. (1970*a*). *Acta Cryst.* A**26**, 297–306.

STUHRMANN, H. B. (1970*b*). *Z. Phys. Chem. Frankfurt*, **72**, 177–184, 185–198.

SVERGUN, D. I. (1991). *J. Appl. Cryst.* **24**, 485–492.

SVERGUN, D. I. (1994). *Acta Cryst.* A**50**, 391–402.

THANKI, N., THORNTON, J. M. & GOODFELLOW, J. M. (1988). *J. Mol. Biol.* **202**, 637–657.

Biophysical Journal Volume 76 June 1999 2879–2886



# Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing

D. I. Svergun

European Molecular Biology Laboratory, Hamburg, Germany and Institute of Crystallography, Russian Academy of Sciences, Moscow, Russia

ABSTRACT A method is proposed to restore ab initio low resolution shape and internal structure of chaotically oriented particles (e.g., biological macromolecules in solution) from isotropic scattering. A multiphase model of a particle built from densely packed dummy atoms is characterized by a configuration vector assigning the atom to a specific phase or to the solvent. Simulated annealing is employed to find a configuration that fits the data while minimizing the interfacial area. Application of the method is illustrated by the restoration of a ribosome-like model structure and more realistically by the determination of the shape of several proteins from experimental x-ray scattering data.

## INTRODUCTION

The fundamental aim of structural studies in molecular biology is to establish a relationship between the structure (or, more precisely, structural changes) and function of biological macromolecules. Over the past years, a tremendous amount of structural information has been obtained using macromolecular crystallography and nuclear magnetic resonance (NMR). These high-resolution methods apply only in rather specific conditions: it is often difficult to grow crystals of high molecular weight (MW) assemblies that are suitable for diffraction, and the application of NMR is fundamentally limited to small (MW $<$ 30 kd) proteins. As most cellular functions are performed by macromolecular complexes, the structure of which depends on their environment, alternative ways of obtaining information on structures and the factors governing their often subtle changes must be explored.

X-ray and neutron small angle scattering (SAS) in solution can yield low-resolution information only (from ~1–100 nm) but are applicable in a broad range of conditions and particle sizes (Feigin and Svergun, 1987). SAS permits analysis of biological macromolecules and their complexes in nearly physiological environments and direct study of structural responses to changes in external conditions.

Scattering intensity from a dilute monodisperse solution of macromolecules (e.g., of purified proteins) is proportional to the spherically averaged single-particle scattering $I(s) = \langle A^2(s)\rangle_\Omega$, where $s = (s, \Omega)$ is the scattering vector, $s = (4\pi/\lambda)\sin\theta$, $\lambda$ the wavelength, and $2\theta$ the scattering angle. The sampling theorem (Shannon and Weaver, 1949; Moore, 1980; Taupin and Luzzati, 1982) estimates the number of degrees of freedom associated with $I(s)$ on an interval $s_{min} < s < s_{max}$ as $N_s = D_{max}(s_{max} - s_{min})/\pi$, where $D_{max}$ is the maximum particle diameter. As the SAS curves decay rapidly with $s$ they are reliably registered only at low resolution and, in practice, $N_s$ does not exceed 10–15. Based on this estimate, SAS is commonly considered to be not only a low-resolution but also a low-information technique.

Additional information about the particle structure is provided by contrast variation (Stuhrmann and Kirste, 1965). The contrast of a particle or its component with a scattering density distribution $\rho(r)$ in a solvent of density $\rho_s$ is the average effective density $\Delta\rho = \langle\rho(r)\rangle - \rho_s$. For single-component macromolecules (e.g., proteins), measurements at different $\rho_s$ allow extraction of the scattering due to the particle shape. For particles consisting of distinct components with different scattering length densities (e.g., lipoprotein or nucleoprotein complexes), contributions from the components can be extracted, allowing analysis of their individual structures and mutual positions. Neutron contrast variation studies employing isotopic H/D exchange are especially effective due to a remarkable difference in the scattering length of H and D atoms (Koch and Stuhrmann, 1979; Capel et al., 1987).

Only a few particle parameters (radius of gyration $R_g$, volume, $D_{max}$) are directly evaluated from the SAS data. A common way of further analysis by trial-and-error modeling requires *a priori* information and can by no means guarantee uniqueness. The degree of uncertainty is reduced when the structure of individual domains is available; this also permits construction of biologically meaningful models (Krueger et al., 1997; Ashton et al., 1997; Svergun et al., 1998a). An ab initio approach for restoration of low-resolution envelopes (Stuhrmann, 1970a; Svergun and Stuhrmann, 1991; Svergun et al., 1996) has been applied to shape determination of proteins (Svergun et al., 1997a) and contrast variation analysis of ribosomes (Svergun, 1994; Svergun et al., 1997b). An interesting procedure for ab initio shape determination has recently been developed by Chacón et al. (1998) using a genetic algorithm to produce models described by densely packed beads. The present paper in-

*Received for publication 28 January 1999 and in final form 30 March 1999.*

Address reprint requests to D. I. Svergun, European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany. Tel.: 49-40-89902-125; Fax: 49-40-89902-149; E-mail: svergun@embl-hamburg.de.

© 1999 by the Biophysical Society

0006-3495/99/06/2879/08 $2.00

troduces a general method for ab initio low-resolution shape and internal structure retrieval and presents its application to a model system and to real objects.

## MATERIALS AND METHODS

### Dummy atom model

First, a general model of a $K$-phase particle ($K \geq 1$) is constructed and its scattering is calculated. A volume is defined which encloses the particle (e.g., a sphere of sufficiently large radius $R$) and filled with $N$ dummy atoms (e.g., closely packed spheres of radius $r_0 \ll R$; see example of such packing in Fig. 1, *middle row*). Each dummy atom is assigned an index $X_j$ indicating the phase to which it belongs ($X_j$ ranges from 0 (solvent) to $K$). Given the fixed atomic positions, the shape and structure of the dummy atom model (DAM) are completely described by a phase assignment (configuration) vector $X$ with $N = (R/r_0)^3$ components.

Assuming that the dummy atoms of the $k$th phase have contrast $\Delta\rho_k$, the scattering intensity from the DAM is

$$I(s) = \left\langle \left[ \sum_{k=1}^{K} \Delta\rho_k A_k(s) \right] \right\rangle_{\Omega} \qquad (1)$$

where $A_k(s)$ is the scattering amplitude from the volume occupied by the $k$th phase. Representing the amplitudes using the spherical harmonics $Y_{lm}(\Omega)$

$$A_k(s) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}^{(k)}(s) Y_{lm}(\Omega) \qquad (2)$$

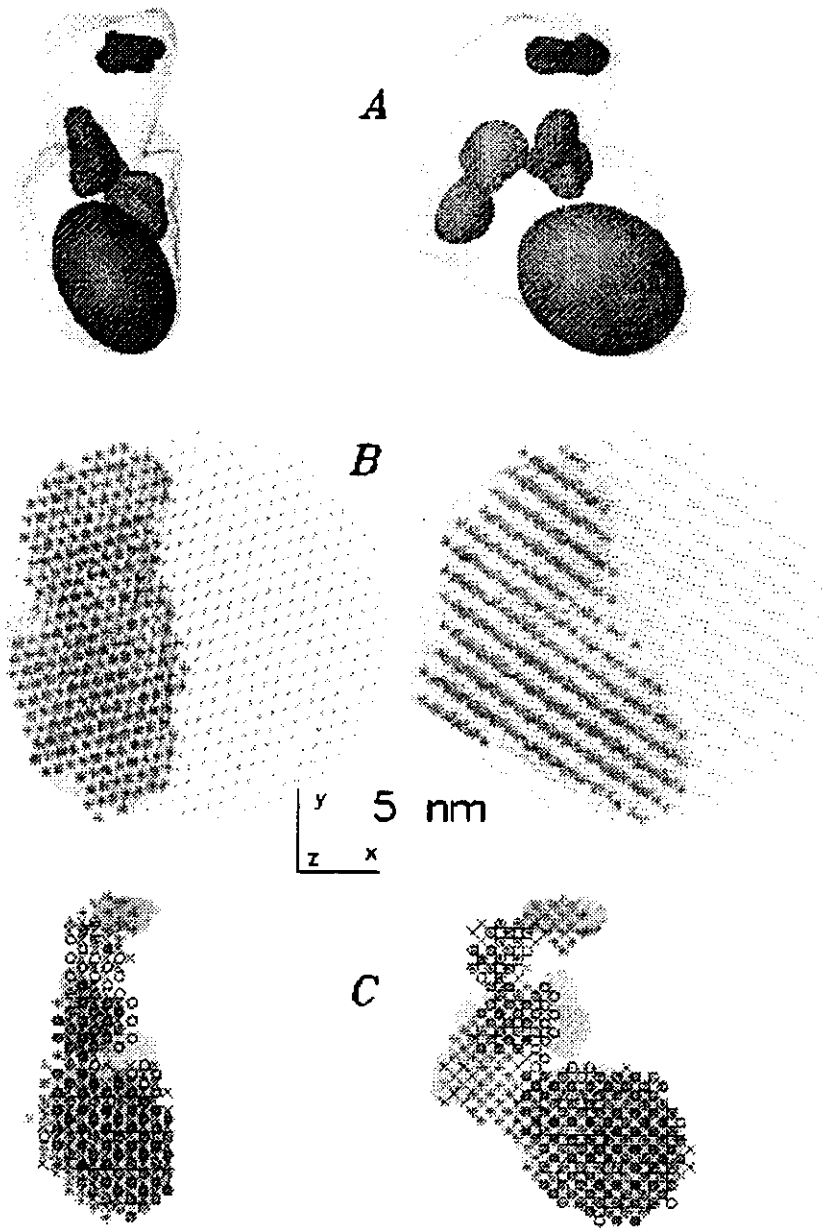


FIGURE 1 Restoration of a two-phase particle. (*A*) Model structure: outer semitransparent envelope and solids represent phase 1. (*B*) Shape determination: green dots represent the positions of dummy atoms in the search volume, red circles (superimposed on the outer envelope) show the configuration obtained by fitting scattering curve 1 in Fig. 2. (*C*) Internal structure retrieval: final configurations of phase 1 for three independent runs (circles, crosses, and asterisks) are superimposed to the semitransparent solids of phase 1. The right orientation is rotated counterclockwise around Y by 90°. Figs. 1 and 3 were displayed using the program ASSA (Kozin et al., 1997).

one obtains (Stuhrmann, 1970b; Svergun, 1994)

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left\{ \sum_{k=1}^{K} [\Delta\rho_k A_{lm}^{(k)}(s)]^2 + 2 \sum_{n>k} \Delta\rho_k A_{lm}^{(k)}(s) \Delta\rho_n [A_{lm}^{(n)}(s)]^* \right\} \quad (3)$$

The partial amplitudes from the volume occupied by the $k$th phase in a DAM are

$$A_{lm}^{(k)}(s) = i^l \sqrt{2/\pi} f(s) \sum_{j=1}^{N_k} j_l(sr_j) Y_{lm}^*(\omega_j) \quad (4)$$

where the sum runs over the dummy atoms of the $k$th phase, $r_j$, $\omega_j$ are their polar coordinates, $j_l(x)$ the spherical Bessel function, and $f(s)$ the scattering from a single atom (form factor). Eqs. 3 and 4 allow computation of the scattering curves from a multiphase DAM for an arbitrary configuration $X$ and arbitrary contrasts $\Delta\rho_k$.

## Looseness criterion

Given a set of $M \geq 1$ contrast variation curves $I_{exp}^{(i)}(s)$, $i = 1, \ldots, M$, it is natural to search for a configuration $X$ minimizing the discrepancy

$$\chi^2 = \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{N(i)} [(I_{exp}^{(i)}(s_j) - I^{(i)}(s_j))/\sigma(s_j)]^2 \quad (5)$$

where $N(i)$ is the number of points in the $i$th curve and $\sigma(s)$ denotes the experimental errors.

For an adequate description of a structure the number of dummy atoms must, however, be large ($N \approx 10^3$). Even if the data are neatly fitted, uniqueness of such a model cannot be meaningfully discussed.

Let us require the model to have low resolution with respect to $r_0$. Qualitatively this means that the volumes occupied by the phases are not expected to contain only a single dummy atom or a few atoms, nor can the interfacial area be too detailed. For a quantitative estimate, a list of contacts (i.e., atoms at an offset $<2r_0$) is defined for each dummy atom. The number of contacts for hexagonal packing is $N_c = 12$ (or less for the atoms close to the DAM border). An individual connectivity of a nonsolvent atom is characterized by counting among its contacts the number of atoms $N_c$ belonging to the same phase. An exponential form $C(N_c) = 1 - P(N_c) = 1 - [\exp(-0.5N_c) - \exp(-0.5N_c)]$ can be taken to emphasize loosely connected dummy atoms. This function slowly decays from $C(12) = 1$ (ideal connectivity) to $C(6) = 0.943$ (half the contacts may indicate, e.g., the particle border), followed by a steep decrease for looser atoms toward $C(0) = 0.002$ (for a dummy isolated atom, which should never appear). The compactness of a given configuration $X$ can be computed as an average connectivity of all nonsolvent atoms $\langle C(N_c)\rangle$. In the following, a configuration will be characterized by the average looseness $P(X) = 1 - \langle C(N_c)\rangle$. This value depends mostly on the connectivity of the individual atoms, but also on the anisometry of the particle represented by the nonsolvent atoms. For example, at $K = 1$ and $N \approx 2000$, P $\approx$ 0.007 for a solid sphere, 0.012 for a prolate ellipsoid of rotation with an axial ratio 1:10. Filling the two volumes randomly with phase 0 (solvent) and phase 1 (particle) atoms yields $P \approx 0.1$ in both cases.

## Minimization procedure

The task of retrieving a low-resolution model from the scattering data can be formulated as follows: given a DAM, find a configuration $X$ minimizing a goal function $f(X) = \chi^2 + \alpha P(X)$, where $\alpha > 0$ is the weight of the looseness penalty. As usual when using penalties, the weight has to be selected in such a way that the second term yields a significant (say ~10–50%) contribution to the function at the end of the minimization. Because $\chi^2$ is expected to be around 1 for a correct solution and $P(X)$ is of order of $10^{-2}$ for compact bodies, $\alpha \approx 10^1$ is a reasonable choice.

Given the large number of variables and the combinatorial nature of the problem, simulated annealing (SA) (Kirkpatrick et al., 1983) seems to be an appropriate global minimization method. The main idea in this method is to perform random modifications of the system (i.e., of the vector $X$) while moving always to the configurations that decrease energy $f(X)$, but sometimes also to those that increase $f(X)$. The probability of accepting the latter moves decreases in the course of the minimization (the system is cooled). At the beginning, the temperature is high and the changes almost random, whereas at the end a configuration with nearly minimum energy is reached. The algorithm was implemented in its faster simulated quenching (Press et al., 1992; Ingber, 1993) version:

1. Start from a random configuration $X_0$ at a high temperature $T_0$ [e.g., $T_0 = f(X_0)$].

2. Select an atom at random, randomly change its phase (configuration $X'$), and compute $\Delta = f(X') - f(X)$.

3. If $\Delta < 0$, move to $X'$; if $\Delta > 0$, do this with a probability $\exp(-\Delta/T)$. Repeat Step 2 from $X'$ (if accepted) or from $X$.

4. Hold $T$ constant for $100N$ reconfigurations or $10N$ successful reconfigurations, whichever comes first, then cool the system ($T' = 0.9T$). Continue cooling until no improvement in $f(X)$ is observed.

Only one dummy atom is changed per move so that only a single summand in Eq. 4 must be updated to calculate the partial amplitudes. As the latter is the most time-consuming operation, this accelerates the evaluation of $f(X)$ about $N$ times. This acceleration makes it possible to use the SA, which is very robust (Ingber, 1993) but would otherwise be prohibitively slow, as millions of function evaluations are required for a typical refinement.

## Scattering experiments and data treatment

The synchrotron radiation x-ray scattering data from enolpyryvul transferase, elongation factor Tu, thioredoxin reductase, and reverse transcriptase were collected following standard procedures using the X33 camera (Koch and Bordas, 1983; Boulin et al., 1986, 1988) of the European Molecular Biology Laboratory at Deutsches Elektronen Synchrotron (Hamburg) and multiwire proportional chambers with delay line readout (Gabriel and Dauvergne, 1982). Details of the experimental procedures are given elsewhere (Schönbrunn et al., 1998; Bilgin et al., 1998; Svergun et al., 1997a, 1998b). The data processing (normalization, buffer subtraction, etc.) involved statistical error propagation using the program SAPOKO (Svergun and Koch, unpublished data). The maximum diameters were estimated from the experimental data using the orthogonal expansion program ORTOGNOM (Svergun, 1993).

# RESULTS

## Model example: two-phase particle

The method was first tested on simulated data from a model two-phase object in Fig. 1. The outer envelope was taken from the electron microscopic model of the 30S *Escherichia coli* ribosomal subunit (Frank et al., 1995). Phase 1 is represented by four bodies inside the envelope (a triaxial ellipsoid and several ribosomal proteins, see Table 1), phase 2 by the remaining volume. The curves in Fig. 2 were calculated in a typical experimental interval $0.06 < s < 1.5$ nm$^{-1}$ (a resolution of $2\pi/s_{max} = 4.1$ nm). The contrasts of the two phases were taken to correspond to those of protein and RNA, respectively, in a neutron experiment. Curve 1

**TABLE 1 Envelopes comprising the two-phase model**

| Model body. Protein Data Bank entry | Volume ($nm^3$) | $R_g$ (nm) |
|---|---|---|
| 1. Outer envelope | 1232 | 6.29 |
| 2. Ellipsoid 3.1 × 4.4 × 5.0 nm | 286 | 3.28 |
| 3. Dimer of protein S8. 1sei (Davies et al., 1996) | 43 | 2.59 |
| 4. Protein S5. 1pkp (Ramakrishnan and White. 1992) | 20 | 1.53 |
| 5. Protein S15. lab3 (Berglund et al., 1997) | 15 | 1.45 |

The envelopes of the ribosomal proteins were computed from their atomic coordinates in the Protein Data Bank (PDB. Bernstein et al., 1977) using the CRYSOL program (Svergun et al., 1995).

corresponds to infinite contrast ($\Delta\rho_1 = \Delta\rho_2$, deuterated particle in $H_2O$), curves 2–5 to a protonated particle in solvents with $D_2O$ concentrations of 0%, 40% (protein matched out), 70% (RNA matched out), and 100%. Only three of these five curves are independent, the redundancy being required, as in real experiments, to account for random errors simulated here (3% relative noise was added to the intensities).

Model calculations below were performed with the correctly scaled data sets (absolute scale) and with those multiplied by arbitrary factors (relative scale, by fitting only geometry of the curves) led to similar final models. Series 2–3 over spherical harmonics were truncated at $l = 14$ and the atomic scattering $f(s) = 1$ was taken (it can be shown that a constant form factor, not that of a sphere with radius

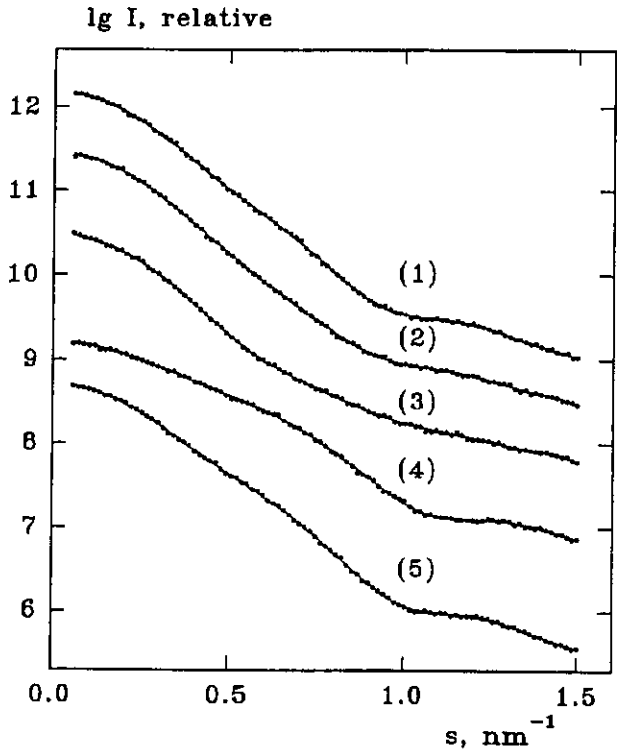


FIGURE 2 Simulated scattering from the two-phase model in Fig. 1 *A* (*dots*) and the fits (*solid lines*; results of different restorations are indistinguishable). For the contrasts of curves 1–5, see text.

$r_0$, ensures adequate computation of the partial amplitudes). The simulated and experimental curves were always neatly fitted ($\chi^2 \approx 1$) and the final looseness was around $P(X) \approx 0.02$.

At infinite contrast, the object is a single phase particle and ab initio shape determination can be done against curve 1 ($K = M = 1$). A sphere of radius $R \approx D_{max}/2 = 11$ nm was filled by dummy atoms with $r_0 = 0.8$ nm ($N = 1925$). Annealing yields stable results for different starting points and the restored configurations (a typical one is presented in Fig. 1 *B*) match the theoretical envelope well. The shape is, of course, recovered in an arbitrary orientation and handedness, the enantiomorph yielding the same scattering curve.

The envelope of the DAM in Fig. 1 *C* was computed, radially expanded by 0.5 nm to enclose 1.25 times the volume of the model particle and filled with $N = 2098$ dummy atoms at $r_0 = 0.5$ nm. The two-phase refinement performed against all five scattering curves yields nearly perfect restoration of the overall shape. The reconstruction of the inner structure is illustrated in Fig. 1 *C* displaying the atoms assigned to phase 1 for three independent runs. The shape and location of the largest ellipsoidal particle are well recovered, whereas the uncertainty in the representation of the smaller bodies is relatively large. This is not surprising given that these smaller bodies occupy only a few percent of the model volume and their radii of gyration are smaller than the resolution of the data (Table 1). It is rather surprising that the method is sensitive to their presence: the solutions for all runs (more than a dozen) displayed atoms of phase 1 in the volume around the correct positions of the small particles. As can be seen from Fig. 1 *C*, averaging the results of independent runs provides a way to further refine the solution and to estimate its uncertainty.

## Practical example: shape determination of proteins

Is it possible to use the method if no contrast variation data are available, e.g., for shape determination of proteins from x-ray scattering? For proteins with MW > 30 kd, the shape scattering dominates the inner part of the x-ray curve. Scattering from the internal structure is nearly a constant that can be subtracted from the data to ensure that the intensity decays as $s^{-4}$ at higher angles, following Porod's (1982) law. Fig. 3 *A-D* and Table 2 illustrate ab initio shape restoration from the experimental data of several proteins with known atomic resolution crystal structures. The synchrotron radiation scattering curves (Fig. 4) were recorded as part of ongoing projects at the European Molecular Biology Laboratory, Hamburg Outstation (see Materials and Methods). The data on a relative scale were used and the diameters of the search spheres $D_{max}$ were determined from the individual experimental curves. The value of $r_0$ was selected to have $N \approx 1500$ atoms, and the results obtained were stable to the starting configuration. Comparison with the appropriately rotated atomic models indicates that the low-resolution structure is well restored.

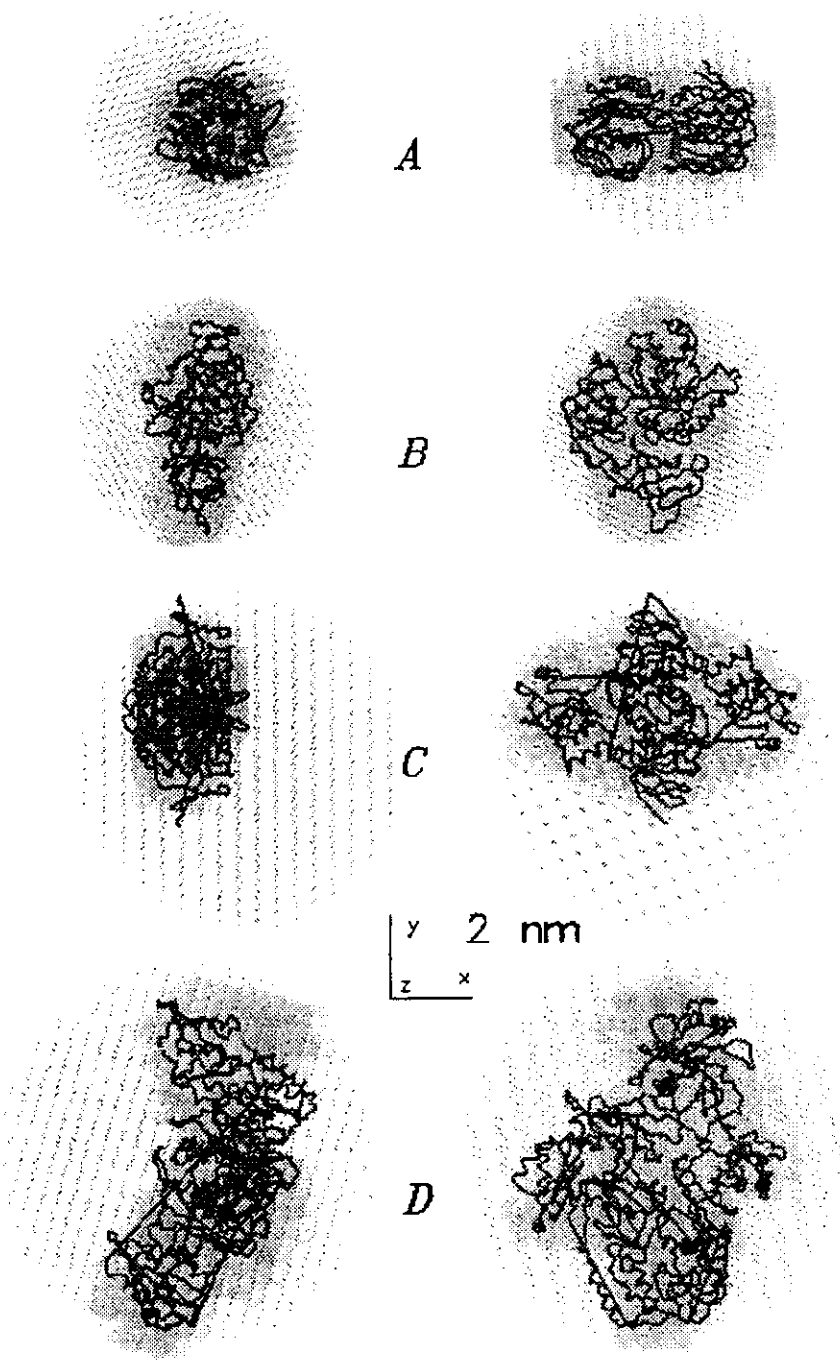


FIGURE 3 Shape determination of proteins in solution. Notations (*A*) through (*D*) correspond to Table 2. *Dots*: search volumes, semitransparent spheres of radius $r_0$ restored configurations superimposed to the $C_\alpha$ chains of the atomic models (*lines*). The right orientation is rotated as in Fig. 1.

The volumes occupied by the final DAMs are in all cases larger than the dry volumes of the proteins computed from their MWs. This apparent swelling is due to the higher density of the bound water in the hydration shell (Ashton et al., 1997; Svergun et al., 1998b).

## DISCUSSION

How can the predictions of the sampling theorem be reconciled with the restoration of the models described by $N \gg N_s$ atoms? First, $N_s$ alone does not define the degrees of freedom for a data set. Redundancy of the experimental data measured with an angular step much smaller than the width of the Shannon's channel ($\Delta s = \pi/D_{max}$) increases the information content; this is successfully used for superresolution in optical image reconstruction (Frieden, 1971). The effective number of degrees of freedom was shown to range from zero at the signal-to-noise ratio of 1 to $15N_s$ at signal-to-noise ratio of $10^3$ (Frieden, 1971). This should not be taken as a proof that one is entitled to build models described by $15N_s$ independent parameters, but rather as an indication that the number of degrees of freedom strongly

**TABLE 2 Proteins recovered and their parameters**

| Protein, Protein Data Bank entry | MW kd | $D_{max}$ nm | $s_{max}$ nm$^{-1}$ | $N_s$ | $r_0$ nm | $N$ | $N_1$* |
|---|---|---|---|---|---|---|---|
| A. Enolpyruvyl transferase 1uae (Skarzynski et al., 1996) | 45 | 7.5 | 2.2 | 5.3 | 0.30 | 1442 | 524 |
| B. Elongation factor Tu 1eft (Kjeldgaard et al., 1993) | 45 | 8.0 | 1.9 | 4.8 | 0.30 | 1763 | 545 |
| C. Thioredoxin reductase 1tde (Waksman et al., 1994) | 68 | 11.0 | 2.0 | 7.0 | 0.45 | 1341 | 267 |
| D. Reverse transcriptase 3hvt (Wang et al., 1994) | 105 | 12.5 | 2.0 | 8.0 | 0.50 | 1483 | 313 |

*$N_1$ denotes the number of dummy atoms of phase 1 in the final model.

depends on data accuracy. For SAS, it was demonstrated by Svergun et al. (1996) that a unique determination of particle envelope is also achieved with a number of model parameters up to $1.5N_s$. Second, the number of independent parameters in a DAM is much lower than $N$ due to the looseness penalty. At later annealing stages the program searches for a compact solution with the smallest interfacial area, whereas the fit acts as a constraint (the penalty, rather than $\chi^2$, is decreased). The more information provided by the data, the more stringent is the constraint, i.e., the more detail should be kept by the DAM. Among the proteins presented, the most detail is obtained for that with the largest MW and the largest $N_s$ value (Fig. 3 *D* and Table 2).

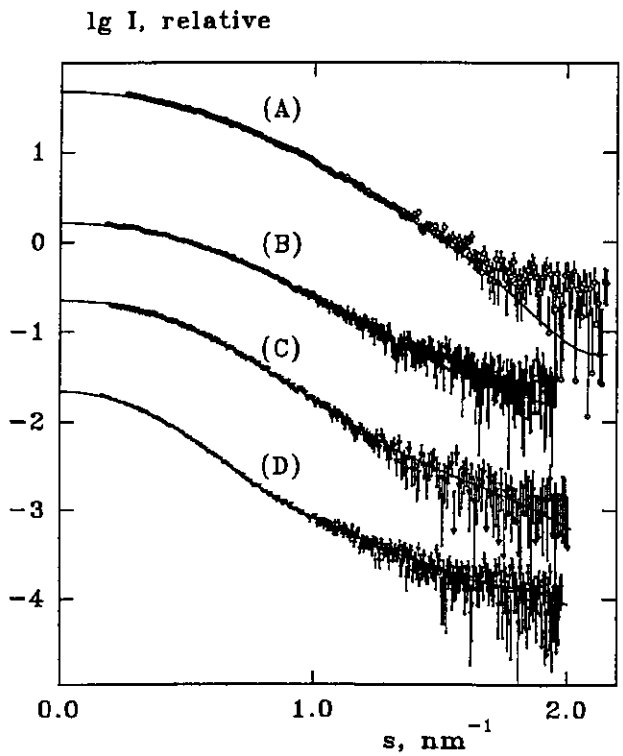


FIGURE 4 Experimental curves from the proteins in Fig. 3 (symbols with error bars; notations (*A*) through (*D*) as in Table 2) and scattering from the restored models (*solid lines*). Appropriate constants were subtracted from the experimental data before fitting; see text.

For single-phase particles ($K = 1$), the shape representation using DAM is equivalent to that employed in the bead modeling of Chacón et al. (1998). The ab initio shape determination from a single scattering curve ($K = M = 1$) is the least favorable case from the informational point of view, as the cross-terms are missing in Eq. 3. Svergun et al. (1996) demonstrated that bodies sharing similar gross features but differing in finer details may produce nearly identical scattering curves in a given interval. A unique solution can then be obtained only by restricting the resolution of the model. In the method of Svergun et al. (1996, 1997a) this was done by representing the particle envelope with limited number of spherical harmonics. Chacón et al. (1998) did not use an explicit compactness criterion; instead, the genetic algorithm procedure started from a relatively large bead radius $r_0$ and several cycles with decreasing $r_0$ were performed. Although the effective resolution of the model was lowered by the reduction of the search volume after each cycle, noticeable portions of loosely connected beads could be seen in the final models.

In the model calculations performed, and also for the examples in Figs. 3 and 4, the SA procedure yielded very similar compact solutions for different starting approximations (again, up to an arbitrary rotation, shift, and handedness). The weight of the looseness penalty may be changed by a factor of up to five without distorting the low-resolution features of the solution, and comparison of several independent runs can be used to estimate the uncertainty. One should stress, however, that the ab initio shape determination must be used with caution, especially if the scattering from the internal inhomogeneities is not negligible. In particular, it would not be justified to expect a detailed shape restoration when using x-ray scattering curves from low MW proteins presented in Fig. 3. Further analysis of the uniqueness of the shape restoration using the SA procedure, including the influence of the systematic errors and comparison with other methods, are in progress. Test calculations made on several other proteins with known atomic structure yielded good ab initio restorations of their low-resolution structure similar to those presented in Fig. 3. It is thus tempting to say that the looseness penalty forces the method to select the level of detail required for uniqueness.

On a 180-MHz SGI workstation with an R10,000 processor, single-phase DAM refinement against one curve takes ~5–6 h of CPU time. For the two-phase system, typical times were longer (40–50 h). These times correspond to the annealing conditions listed above; practice will show to what extent the number of function evaluations can be reduced without affecting the convergence. In particular, it was found that reconfigurations of $50-70N$ are sufficient to equilibrate the system at each temperature, which halves the CPU power required. For a single-phase DAM, significant acceleration can be achieved by reducing the search volume at a later annealing stage, when the particle shape is already well defined. Clearly the method could gain considerably from parallel implementation. Global minimization techniques that are claimed to be faster, e.g., taboo search (Glover, 1989), will also be tested.

Further applications of the method include, first, the analysis of the internal structure of multi-component macromolecular complexes, which in many cases is facilitated by using electron microscopic models of the overall shape. In particular, in studies on ribosome, where single crystals have long been available, little information has been reported so far about the mutual distribution of ribosomal components despite remarkable recent progress in x-ray crystallography and cryo-electron microscopy (Ban et al., 1998). The main reason for this is the small contrast between ribosomal proteins and RNA in these studies, and the most detailed results are still those obtained by neutron scattering using triangulation of individual proteins in the ribosomal subunits (Capel et al., 1987; May et al., 1992). The method presented is being used to construct the map of the protein-RNA distribution in the *E. coli* ribosome based on the earlier neutron scattering data from selectively deuterated particles (Svergun et al., 1997b). Second, ab initio retrieval of the quaternary structure of macromolecules in terms of low-resolution particle shape could, albeit with some caveats, also be done without contrast variation, using x-ray scattering data only. The executable codes of the shape determination program for IBM-PC and major UNIX platforms are available from the author upon request.

The author is indebted to M. H. J. Koch for helpful discussions and encouraging criticism and to M. B. Kozin for assistance with computer graphics. I thank E. Schönbrunn, N. Bilgin, S. Kuprin, and L. Goobar-Larsson for providing the experimental scattering data. The work was supported by European Union grant BIO4-CT97–2143.

## REFERENCES

Ashton, A. W., M. K. Boehm, J. R.Gallimore, M. B. Pepys, and S. J. Perkins. 1997. Pentameric and decameric structures in solution of serum amyloid P component by X-ray and neutron scattering and molecular modelling analyses. *J. Mol. Biol.* 272:408–422.

Ban, N., B. Freeborn, P. Nissen, P. Penczek, R. A. Grassucci, R. A. Sweet, J. Frank, P. B. Moore, and T. A. Steitz. 1998. A 9 Å resolution X-ray crystallographic map of the large ribosomal subunit. *Cell.* 93: 1105–1115.

Berglund, H., A. Rak, A. Serganov, M. Garber, and T. Härd. 1997. Solution structure of the ribosomal RNA binding protein S15 from Th. *thermophilus*. *Nat. Struct. Biol.* 4:20–21.

Bilgin, N., M. Ehrenberg, C. Ebel, G. Zaccai, Z. Sayers, M. H. J. Koch, D. I. Svergun, C. Barberato, V. Volkov, P. Nissen, and J. Nyborg. 1998. Solution structure of the ternary complex between aminoacyl-tRNA, elongation factor Tu, and guanosine triphosphate. *Biochemistry.* 37: 8163–8172.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Boulin, C., R. Kempf, M. H. J. Koch, and S. M. McLaughlin. 1986. Data appraisal, evaluation and display for synchrotron radiation experiments: hardware and software. *Nucl. Instrum. Meth.* A249:399–407.

Boulin, C. J., R. Kempf, A. Gabriel, and M. H. J. Koch. 1988. Data acquisition systems for linear and area X-ray detectors using delay line readout. *Nucl. Instrum. Meth.* A269:312–320.

Capel, M. S., D. M. Engelman, B. R. Freeborn, M. Kjeldgaard, J. A. Langer, V. Ramakrishnan, D. G. Schindler, D. K. Schneider, B. P. Schoenborn, I.-Y. Sillers, S. Yabuki, and P. B. Moore. 1987. A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science.* 238:1403–1406.

Chacón, P., F. Morán, J. F. Díaz, E. Pantos, and J. M. Andreu. 1998. Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biophys. J.* 74:2760–2775.

Davies, C., V. Ramakrishnan, and S. W. White. 1996. The crystal structure of ribosomal protein S8: implications for modelling the central domain of 16A rRNA. *Structure* 4:1093–1104.

Feigin, L. A., and D. I. Svergun. 1987. Structure Analysis by Small-Angle X-Ray and Neutron Scattering. Plenum Press, New York.

Frank, J., J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R. K. Lata, and R. K. Agrawal. 1995. A model of protein synthesis based on a new cryo-electron microscopy reconstruction of the *E. coli* ribosome. *Nature.* 376:441–444.

Frieden, B. R. 1971. Evaluation, design and extrapolation methods for optical signals, based on the use of the prolate functions. *In* Progress in Optics, vol. 9. E. Wolf, editor. North-Holland, Amsterdam. 312–407.

Gabriel, A., and F. Dauvergne. 1982. The localization method used at EMBL. *Nucl. Instrum. Meth.* 201:223–224.

Glover, F. 1989. Tabu Search–Part I. *ORSA J. Comput.* 1:190–206.

Ingber, L. 1993. Simulated annealing: practice versus theory. *Math. Computer Modelling.* 18:29–57.

Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecci. 1983. Optimization by simulated annealing. *Science.* 220:671–680.

Kjeldgaard, M., P. Nissen, S. Thirup, and J. Nyborg. 1993. The crystal structure of elongation factor EF-Tu from Thermus aquaticus in the GTP conformation *Structure.* 1:35–50.

Koch, M. H. J., and H. B. Stuhrmann. 1979. Neutron scattering studies of ribosomes. *Meth. Enzymol.* 59:670–706.

Kozin, M. B, V. V. Volkov, and D. I. Svergun. 1997. ASSA: a program for three-dimensional rendering in solution scattering from biopolymers. *J. Appl. Crystallogr.* 30:811–815.

Krueger, J. K., G. A. Olah, S. E. Rokop, G. Zhi, J. T. Stull, and J. Trewhella. 1997. Structures of calmodulin and a functional myosin light chain kinase in the activated complex: a neutron scattering study. *Biochemistry.* 36:6017–6023.

May, R. P., V. Nowotny, P. Nowotny, H. Voss, and K. H. Nierhaus. 1992. Inter-protein distances within the large subunit from *Escherichia coli* ribosomes. *EMBO J.* 11:373–378.

Moore, P. B. 1980. Small-angle scattering. Information content and error analysis. *J. Appl. Cryst* 13:168–175.

Porod, G. 1982. General theory. *In* Small-Angle X-Ray Scattering. O. Glatter and O. Kratky, editors. Academic Press, London. 17–51.

Press, W. H., S. A. Teukolsky, W. T. Wetterling, and B. P. Flannery. 1992. Numerical Recipes. University Press, Cambridge.

Ramakrishnan, V., and S. W. White. 1992. The structure of ribosomal protein S5 reveals 2 sites of interaction with 16S RNA. *Nature.* 358: 768–771.

Schönbrunn, E., D. I. Svergun, N. Amrhein, and M. H. J. Koch. 1998. Studies on the conformational changes in the bacterial cell wall biosynthetic enzyme UDP-N-acetylglucosamine enolpyruvyltransferase (MurA). *Eur. J. Biochem.* 253:406–412.

Shannon, C. E., and W. Weaver. 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL.

Skarzynski, T., A. Mistry, A. Wonacott, S.E. Hutchinson, V. A. Kelly, and K. Duncan. 1996. Structure of UDP-n-Acetylglucosamine enolpyruvyl transferase, an enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-n-Acetylglucosamine and the drug fosfomycin. *Structure.* 4:1465–1474.

Stuhrmann, H. B. 1970a. Ein neues Verfahren zur Bestimmung der Oberflächenform und der inneren Struktur von gelösten globulären Proteinen aus Röntgenkleinwinkelmessungen. *Zeitschr. Physik. Chem. Neue Folge.* 72:177–198.

Stuhrmann, H. B. 1970b. Interpretation of small-angle scattering of dilute solutions and gases. A representation of the structures related to a one-particle scattering functions. *Acta Crystallogr.* A26:297–306.

Stuhrmann, H. B., and Kirste, R. G. 1965. Elimination der intrapartikulären Untergrundstreuung bei der Röntgenkleinwinkelstreuung am kompakten Teilchen (Proteinen). *Zeitschr. Physik. Chem. Neue Folge* 46:247–250.

Svergun, D. I. 1993. A direct indirect method of small-angle scattering data treatment. *J. Appl. Crystallogr.* 26:258–267.

Svergun, D. I. 1994. Solution scattering from biopolymers: advanced contrast variation data analysis. *Acta Crystallogr.* A50:391–402.

Svergun, D. I., and H. B. Stuhrmann. 1991. New developments in direct shape determination from small-angle scattering. 1. Theory and model calculations. *Acta Crystallogr.* A47:736–744.

Svergun, D. I., C. Barberato, and M. H. J. Koch. 1995. CRYSOL: a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773.

Svergun, D. I., V. V. Volkov, M. B. Kozin, and H. B. Stuhrmann. 1996. New developments in direct shape determination from small-angle scattering. 2. Uniqueness. *Acta Crystallogr.* A52:419–426.

Svergun, D. I., V. V. Volkov, M. B. Kozin, H. B. Stuhrmann, C. Barberato, and M. H. J. Koch. 1997a. Shape determination from solution scattering of biopolymers. *J. Appl. Crystallogr.* 30:798–802.

Svergun, D. I., N. Burkhardt, J. Skov Pedersen, M. H. J. Koch, V. V. Volkov, M. B. Kozin, W. Meerwink, H. B. Stuhrmann, G. Diedrich, and K. H. Nierhaus. 1997b. Solution scattering structural analysis of the 70S *Escherichia coli* ribosome by contrast variation. *J. Mol. Biol.* 271: 588–601, 602–618.

Svergun, D. I., I. Aldag, T. Sieck, K.-H. Altendorf, M. H. J. Koch, D. J. Kane, M. B. Kozin, and G. Grueber. 1998a. A model of the quaternary structure of the *Escherichia coli* F1 ATPase from X-ray solution scattering and evidence for structural changes in the δ subunit during ATP hydrolysis. *Biophys. J.* 75:2212–2219.

Svergun, D. I., S. Richards, M. H. J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai. 1998b. Protein hydration in solution: experimental observation by X-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA.* 95: 2267–2272.

Taupin, D., and V. Luzzati. 1982. Informational content and retrieval in solution scattering studies. I. Degrees of freedom and data reduction. *J. Appl. Crystallogr.* 15:289–300.

Waksman, G., T. S. R. Krishna, C. H. Williams, Jr., and J. Kuriyan. 1994. Crystal structure of *Escherichia coli* thioredoxin reductase refined at 2 angstrom resolution: implications for a large conformational change during catalysis. *J. Mol. Biol.* 236:800–816.

Wang, J., S. J. Smerdon, J. Jaeger, L. A. Kohlstaedt, J. Friedman, P. A. Rice, and T. A. Steitz. 1994. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. *Proc. Natl. Acad. Sci. USA.* 91:7242–7245.