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 $I(s) = \langle A^2(\boldsymbol{s})\rangle_\Omega$, where $\boldsymbol{s} = (s, \Omega)$ is the scattering vector, $s = (4\pi/\lambda)\sin\theta$, λ the wavelength, and 2θ 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 ρ_s is the average effective density $\Delta\rho = \langle\rho(\boldsymbol{r})\rangle - \rho_s$. For single-component macromolecules (e.g., proteins), measurements at different ρ_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.

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 $N \approx (R/r_0)^3$ components.

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

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

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

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

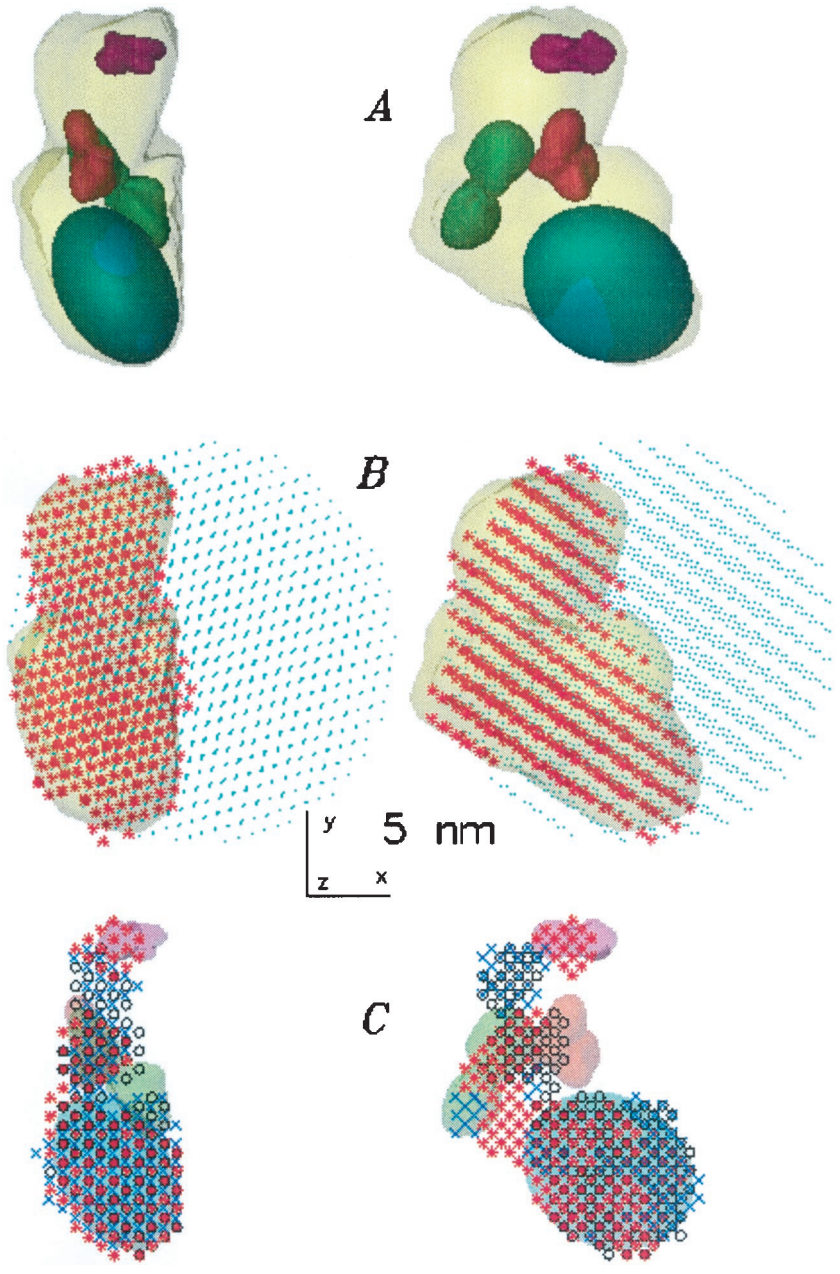

FIGURE 1 Restoration of a two-phase particle. (*A*) Model structure: outer semitransparent envelope and solids represent phase 1. (*B*) Shape determination: green dots represent the positions of dummy atoms in the search volume, red circles (superimposed on the outer envelope) show the configuration 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 kth 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 kth phase, r_j, ω_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 ith 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_e belonging to the same phase. An exponential form $C(N_e) = 1 - P(N_e) = 1 - [\exp(-0.5N_e) - \exp(-0.5N_c)]$ 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 $\langle C(N_e)\rangle$. In the following, a configuration will be characterized by the average looseness $P(X) = 1 - \langle C(N_e)\rangle$. 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 ≈ 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 χ^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 Syncrotron (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 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 $P(X) \approx 0.02$.

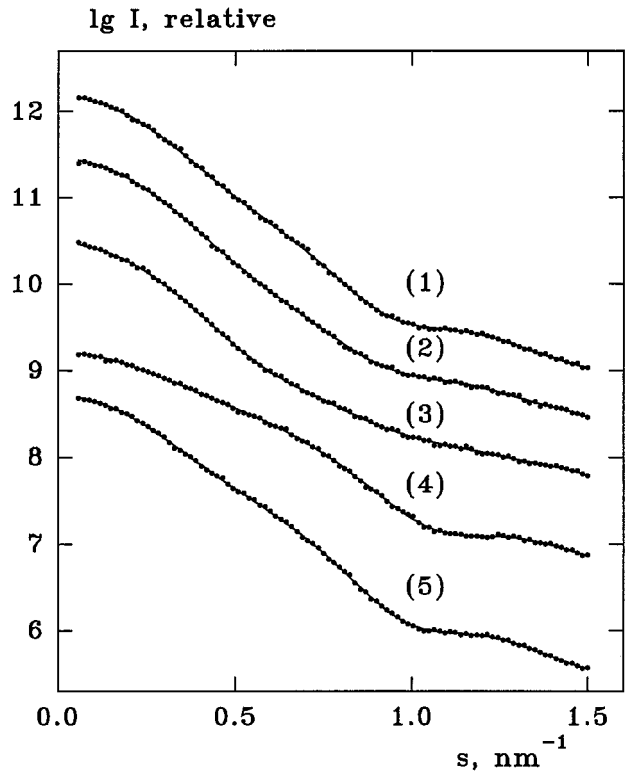

FIGURE 2 Simulated scattering from the two-phase model in Fig. 1 *A* (*dots*) and the fits (*solid lines*; results of different restorations are indistinguishable). For the contrasts of curves 1–5, see text.

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.

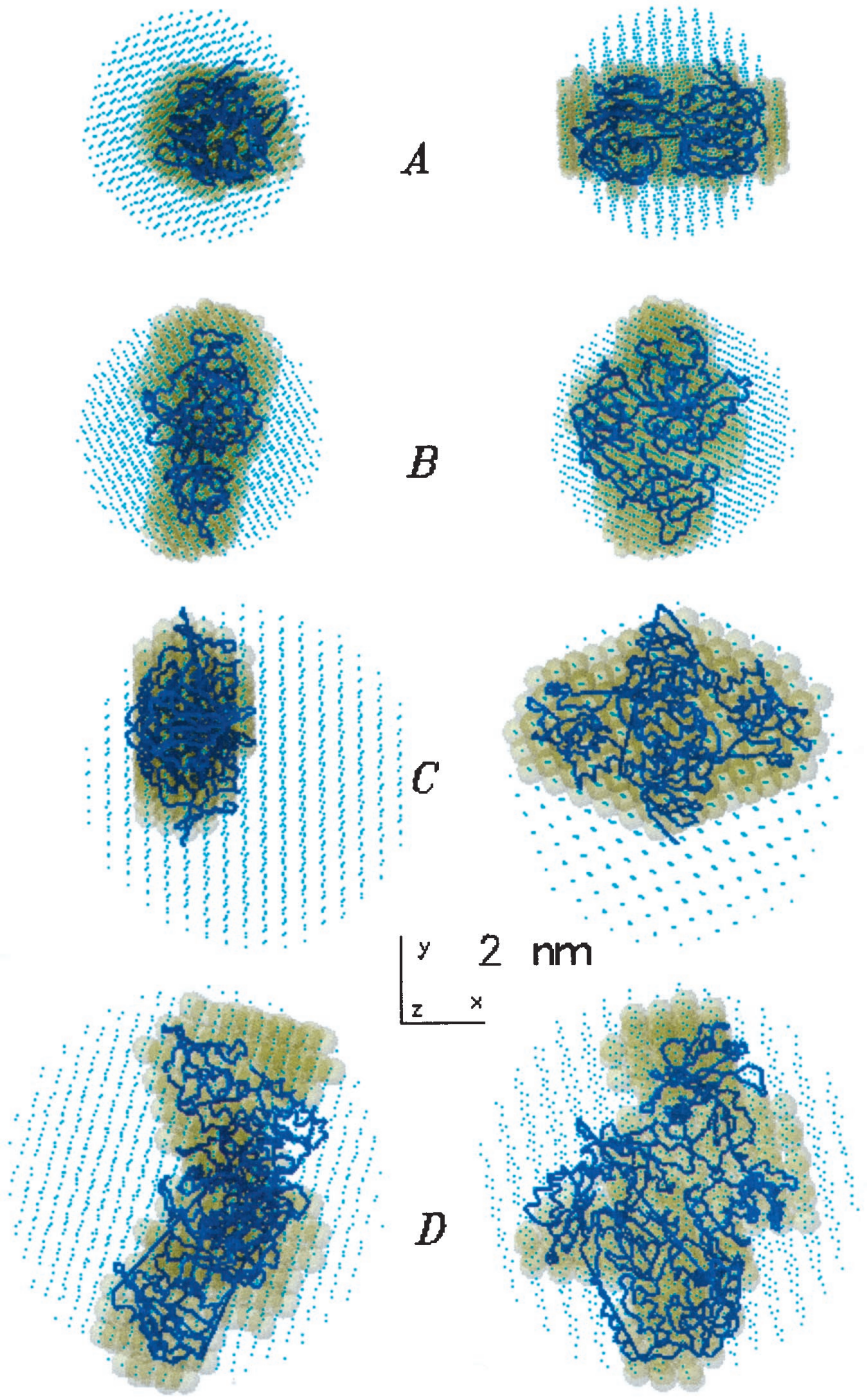

FIGURE 3 Shape determination of proteins in solution. Notations (*A*) through (*D*) correspond to Table 2. *Dots*: search volumes, semitransparent spheres of radius r_0 restored configurations superimposed to the C_α chains of the atomic models (*lines*). The right orientation is rotated as in Fig. 1.

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 χ^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).

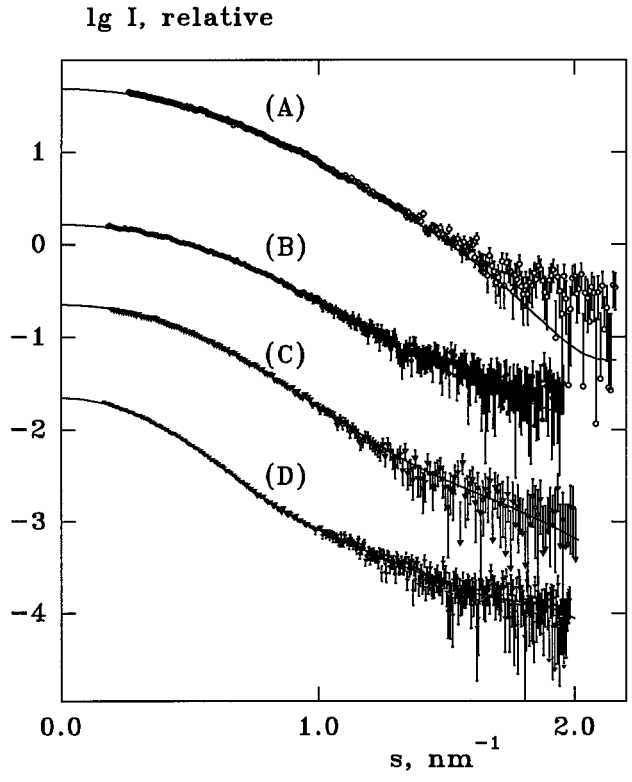

FIGURE 4 Experimental curves from the proteins in Fig. 3 (symbols with error bars; notations (*A*) through (*D*) as in Table 2) and scattering from the restored models (*solid lines*). Appropriate constants were subtracted from the experimental data before fitting; see text.

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 50–70*N* 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.