BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Ambiguity assessment of small-angle scattering curves from monodisperse systems

Maxim V. Petoukhov and Dmitri I. Svergun*

European Molecular Biology Laboratory, Hamburg Unit, EMBL c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany. *Correspondence e-mail: svergun@embl-hamburg.de

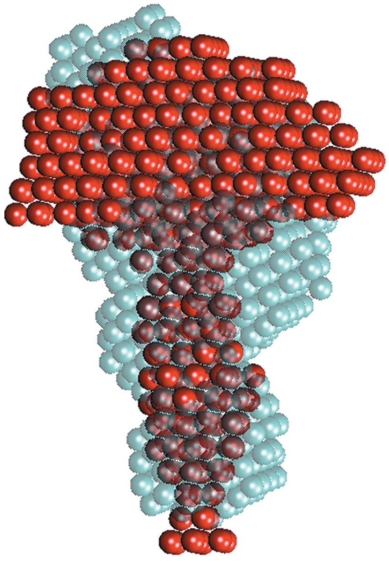
Received 27 November 2014
Accepted 6 February 2015

Edited by Z. Dauter, Argonne National Laboratory, USA

Keywords: solution scattering; uniqueness; ambiguity; structural modelling; shape determination; *AMBIMETER*.

Supporting information: this article has supporting information at journals.iucr.org/d

A novel approach is presented for an *a priori* assessment of the ambiguity associated with spherically averaged single-particle scattering. The approach is of broad interest to the structural biology community, allowing the rapid and model-independent assessment of the inherent non-uniqueness of three-dimensional shape reconstruction from scattering experiments on solutions of biological macromolecules. One-dimensional scattering curves recorded from monodisperse systems are nowadays routinely utilized to generate low-resolution particle shapes, but the potential ambiguity of such reconstructions remains a major issue. At present, the (non)uniqueness can only be assessed by *a posteriori* comparison and averaging of repetitive Monte Carlo-based shape-determination runs. The new *a priori* ambiguity measure is based on the number of distinct shape categories compatible with a given data set. For this purpose, a comprehensive library of over 14 000 shape topologies has been generated containing up to seven beads closely packed on a hexagonal grid. The computed scattering curves rescaled to keep only the shape topology rather than the overall size information provide a 'scattering map' of this set of shapes. For a given scattering data set, one rapidly obtains the number of neighbours in the map and the associated shape topologies such that in addition to providing a quantitative ambiguity measure the algorithm may also serve as an alternative shape-analysis tool. The approach has been validated in model calculations on geometrical bodies and its usefulness is further demonstrated on a number of experimental X-ray scattering data sets from proteins in solution. A quantitative ambiguity score (*a-score*) is introduced to provide immediate and convenient guidance to the user on the uniqueness of the *ab initio* shape reconstruction from the given data set.



1. Introduction

Small-angle scattering (SAS) of X-rays (SAXS) and neutrons (SANS) is a powerful method allowing one to study the low-resolution structure of diverse noncrystalline systems. In structural biology the technique is widely applied to probe the quaternary structure, folding state and flexibility of dissolved particles and complexes and to rapidly analyze structural changes in response to variations in external conditions (Jacques & Trewhella, 2010; Svergun *et al.*, 2013). In an SAS experiment, the intensity of X-rays or neutrons scattered at very small angles (from a few angular minutes to a few degrees) is recorded *versus* momentum transfer $s = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle and λ is the radiation wavelength. Subsequent analysis of $I(s)$ provides low-resolution (about $\sim$1 nm) information about the object. Low-resolution shapes of macromolecules are now routinely reconstructed *ab initio* within minutes thanks to recent progress in data-analysis methods (Chacón *et al.*, 1998; Franke & Svergun, 2009; Takahashi *et al.*, 2003).

Owing to chaotic positions and orientations of macromolecules, the scattering intensity $I(s)$ from a dilute monodisperse solution is proportional to the scattering by a single particle spherically averaged over all orientations (Svergun *et al.*, 2013). The ambiguity of the three-dimensional shape reconstruction from a one-dimensional scattering profile can be considered to be the major limitation of SAS as a structural method. The ambiguity issue has been tackled for a specific case of rigid-body modelling of two subunits with orientations known from NMR residual dipolar couplings (Gabel, 2012). The reliability of data interpretation in terms of structural models depends on the availability of complementary information about the object. Several distinct shapes may yield the same scattering curve and the results of *ab initio* modelling are potentially non-unique. To overcome this problem, multiple runs of Monte Carlo-based reconstruction algorithms are typically performed and the individual solutions are compared with each other and averaged and/or clustered (Petoukhov *et al.*, 2012; Volkov & Svergun, 2003). However, at present there are no objective means to assess the potential ambiguity of the shape reconstruction; given a scattering curve, it is not known *a priori* how extensive the modelling attempts to obtain the true positive in a set of individual solutions should be. Here, an approach is presented to evaluate the potential ambiguity inherent to a given scattering profile. For this, we generated a comprehensive set of shape topologies represented by a limited number of interconnected beads on a regular grid and computed scattering curves from all of these models. The curves were appropriately scaled to highlight the topological shape-related information in the data. The inherent ambiguity of any (experimental or theoretical) scattering pattern can, upon appropriate scaling, be quantitatively estimated through the number of objects with different topologies yielding the computed scattering indistinguishable from the given data.

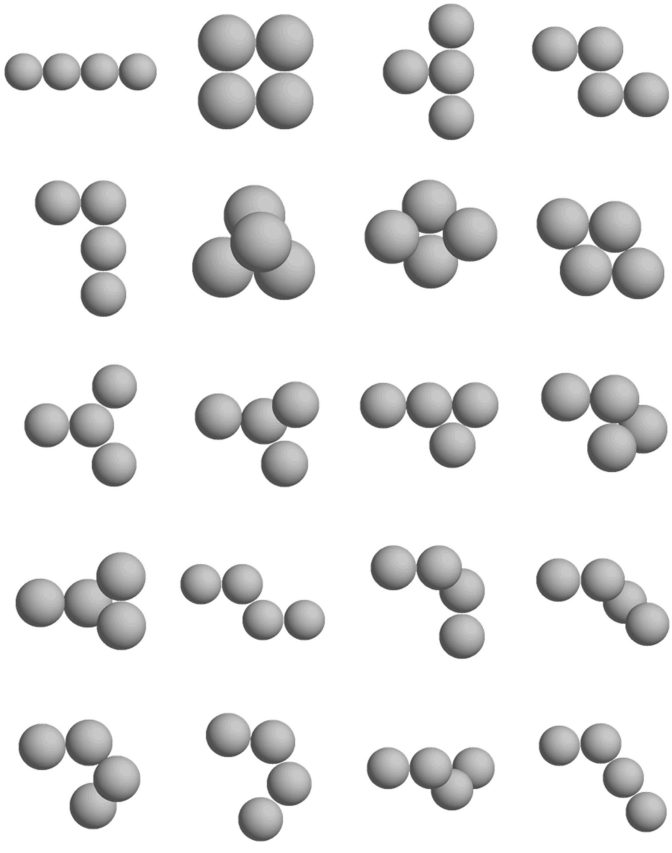


Figure 1
A set of possible topologies (20 skeletons) of four-bead assemblies on a dense hexagonal grid. The sphere radii are scaled such that all models have the same R_g.

2. Methods

2.1. Deconvolution of shape and size information

Each scattering profile contains information about the overall size and the shape topology of the macromolecule, and the first step in the assessment of the ambiguity lies in decoupling these two types of information. According to the Guinier approximation (Guinier, 1939), at very small angles the intensity is represented as $I(s) = I_0\exp[-(sR_g)^2/3]$, where the forward scattering I_0 is proportional to the molecular weight and the radius of gyration R_g gives a measure of the particle size. Representing the scattering profiles using dimensionless coordinates I/I_0 *versus* sR_g allows one to eliminate the mass and size information and to obtain a curve characterized solely by the shape topology. A similar scaling is used in a dimensionless Kratky plot for characterization of the protein folding state (Durand *et al.*, 2010). In the following, the analysis will be restricted to homogeneous approximation, *i.e.* to objects with a uniform density. The main information about the overall shape of such objects is concentrated in the low-resolution data, and the angular range up to an sR_g of about 5–10 typically employed in shape reconstruction will be considered.

2.2. Generation of a comprehensive set of shape topologies

To make the search of the shape topologies consistent with the given renormalized scattering profile possible, the shape topology should be described by just a few parameters. For this purpose, a shape skeleton on a dense hexagonal grid appears to be a suitable representation whereby the number of parameters is equal to the number of grid points in the skeleton. A collection of up to seven beads allows one to represent reasonably complex shape topologies with an anisometry ratio of up to 7:1 (which is sufficient for most biological macromolecules at low resolution). Such a representation is suitable for the fitting of the renormalized scattering curves over a restricted s range of about $3 < sR_g < 7$.

An in-house script was designed to generate an exhaustive library of assemblies containing up to seven interconnected beads on a dense hexagonal grid. Starting from a single bead, the library was compiled by adding neighbouring grid points to the existing members such that only interconnected assemblies were created. Redundant assemblies (for example rotated copies or enantiomorphs of the existing assemblies) were discarded. This yielded one two-bead assembly, four three-bead assemblies, 20 four-bead assemblies (Fig. 1), 133 five-bead assemblies, 1214 six-bead assemblies and 12 739

seven-bead assemblies, resulting in a total of 14 112 unique skeletons.

The scattering intensities from the members of the shape library were calculated using the Debye formula,

$$I(s) = f^2(s)\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\sin(sr_{ij})}{sr_{ij}}, \qquad (1)$$

where $f(s)$ is the form factor of a sphere with a radius equal to half of the grid size, r_{ij} is the distance between the two beads of the skeleton and n is the number of beads. The map of the appropriately renormalized intensities of the set is displayed in Fig. 2. The map area has been split into 60 (horizontally) × 50 (vertically) bins and the density of the curves in each bin was computed. The colouring in Fig. 2 corresponds to the logarithm of the number of curves in each bin from the lowest (blue) to the highest (red). As expected, all renormalized curves match each other up to $sR_g = 1$, according to Guinier's law. The most populated area of the map corresponds to the scattering of a flat disc (red dots); the lower bound is given by a solid sphere (cyan dots) and the upper bound is given by an infinitely thin rod (green dots).

2.3. Search of the consistent shape categories

A program, *AMBIMETER*, has been developed to assess the ambiguity of SAS data using the above library of normalized scattering curves. In this approach, the experimental scattering profile is renormalized and scanned against all intensities from the library of shape topologies to find the subset of skeletons consistent with the given data. A dissimilarity measure is introduced,

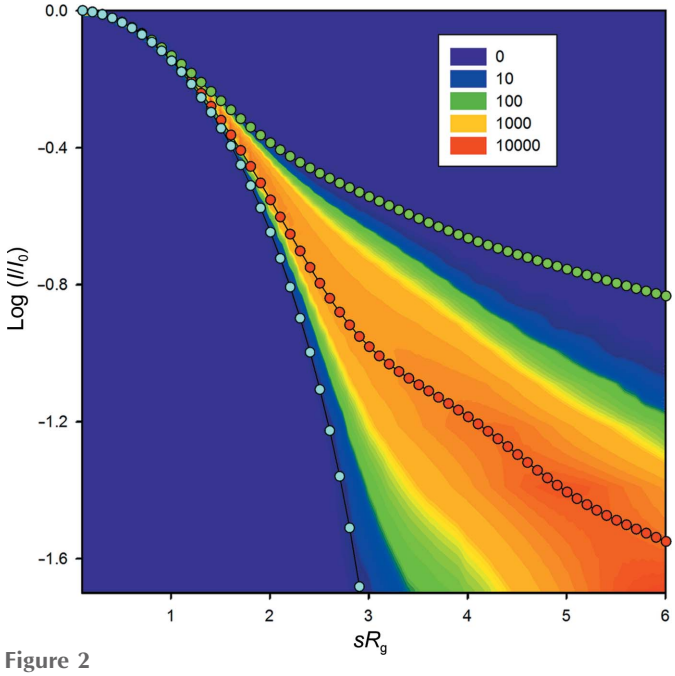


Figure 2
A density map of normalized scattering profiles computed from all unique shape topologies represented by up to seven densely packed beads (see the main text for further explanations).

$$d^2 = \frac{1}{N}\sum_{i=1}^{N}\{(s_iR_g)^2[I_{exp}(s_iR_g)/I_0 - I_{set}(s_iR_g)]\}^2, \qquad (2)$$

representing the discrepancy between the renormalized Kratky plots (here, N is the number of angular points). A threshold of $d^2 < 0.0016$ is applied to identify the shapes that are considered to be in agreement with the data. This threshold is chosen such that each curve from the library has at least one structural neighbour in the default range $sR_g \leq 4$ (restricting the comparisons to this range allows one to avoid the influence of the form factor of the bead of the skeletons). On output, the method selects the skeletons consistent with the given data set (appropriately scaled to the real size of the object) and their fits to the experimental data. For visualization of the results, space-filling models in which the spheres at the grid points are represented by collections of beads of smaller size are computed and written in Protein Data Bank (Berman *et al.*, 2000) format. The number of consistent shape topologies m provides an estimate of the inherent ambiguity of the data set and the variability of the possible solutions, and it is convenient to introduce an ambiguity measure, *a-score* = log(m). A higher number of different skeletons compatible with the given SAS data, yielding a higher *a-score*, points to a higher probability of finding a false positive during the shape reconstruction and, accordingly, more reconstructions should be needed in order to obtain the true positive in the set of solutions.

2.4. Data processing, shape determination, comparison and averaging

Theoretical scattering intensities of the model bodies as well as the experimental profiles from the test proteins have been processed using the indirect transform program *GNOM* (Svergun, 1992), which provides the maximum size D_{max}, the distance distribution function $p(r)$ and the regularized scattering intensity. Shape reconstructions have been performed using the bead-modelling program *DAMMIF* (Franke & Svergun, 2009). Starting from a random bead model, *DAMMIF* employs simulated annealing (SA) to build a compact interconnected assembly fitting the experimental data. Both *DAMMIF* and *AMBIMETER* read in the data processed by *GNOM* as input to obtain the scattering intensity, R_g and D_{max} values. Ten independent *DAMMIF* runs were performed for each shape. The results were compared by *SUPCOMB* (Kozin & Svergun, 2001), which aligns two arbitrary low- or high-resolution models represented by ensembles of points by minimizing a dissimilarity measure called the normalized spatial discrepancy (NSD). As demonstrated by Volkov & Svergun (2003), disagreement between individual *ab initio* bead reconstructions measured by the NSD provides an indication of the stability of the shape determination. Typically, average NSD values below 0.8 point to stable reconstructions, while values exceeding unity are obtained when the shape determination is ambiguous. Note that this criterion applies to the low-resolution models represented by collections of hundreds/thousands of densely packed beads (as in *DAMMIF*). For reconstructions with fewer numbers of

Table 1
Modelling with geometrical bodies.

Body	No. of compatible shape topologies from *AMBIMETER*, *m*	*a-score*	Averaged NSD of *DAMMIF* reconstructions	NSD of the averaged *DAMMIF* shape to the target shape	d^2 yielded by the averaged *DAMMIF* model *versus* the target shape
(*a*) Cube: a = 3 nm	2	0.3	0.52	0.55	0.0002
(*b*) Cylinder: R = 2, h = 4 nm	2	0.3	0.50	0.52	0.0001
(*c*) Prism: a = 2, b = 3, c = 4 nm	6	0.8	0.57	0.52	0.0002
(*d*) Cylinder: R = 2, h = 10 nm	19	1.3	0.57	0.47	0.0003
(*e*) Ring: R = 4, r = 2, h = 4 nm	23	1.4	0.64	0.71	0.0001
(*f*) Prism: a = 1, b = 2, c = 4 nm	120	2.1	0.64	0.58	0.0004
(*g*) Disc: R = 2, h = 1 nm	416	2.6	0.79	0.85	0.0007
(*h*) Disc: R = 5, h = 1 nm	927	3.0	1.38	1.66	0.0110

(more sparse) elements [for example *ab initio* dummy-residue reconstructions by *GASBOR* (Svergun *et al.*, 2001) and also *AMBIMETER* skeletons], models with the same low-resolution shapes may have an NSD exceeding unity. The program *DAMAVER* (Volkov & Svergun, 2003) has been applied to average the individual models (from *DAMMIF* and from *AMBIMETER*) aligned by *SUPCOMB*. In the averaging procedure here, a collection of aligned models is remapped onto a densely packed grid. For each knot in this grid, an occupancy factor is assigned equal to the number of beads belonging to any of the superimposed models in the vicinity of the knot, and the knots with nonzero occupancy form a total spread region. The averaged model corresponding to an interconnected ensemble of the most populated points is evaluated by filtering the map to yield the volume equal to the average excluded volume of all reconstructions. A more detailed analysis is performed by the program *DAMCLUST* (Petoukhov *et al.*, 2012), which is able to identify different types of solutions, classify them into individual clusters and perform averaging within these clusters (this approach is useful, for example, when performing *ab initio* reconstructions with symmetry constraints).

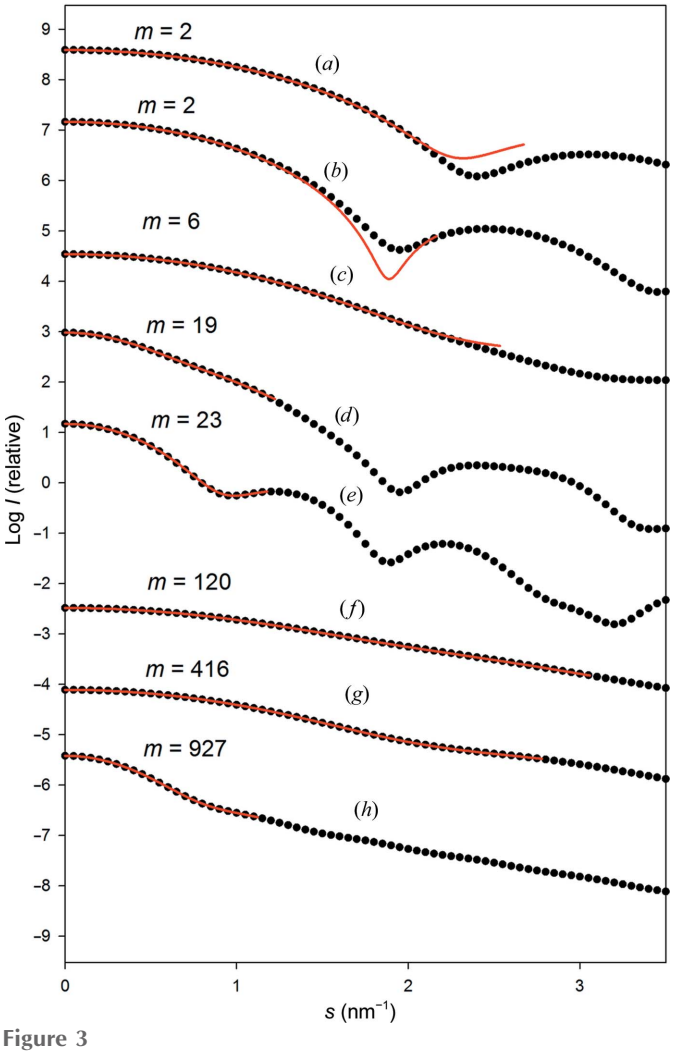


Figure 3
Scattering profiles from geometrical bodies (*a*)–(*h*) as described in the text and Table 1. Simulated data are shown as black dots and fits from the best skeleton selected by *AMBIMETER* in the range up to sR_g = 4 are shown as red solid lines.

3. Results

3.1. Model calculations

The algorithm has been tested on a set of geometrical bodies with various shapes and anisometries. The following geometrical shapes were considered: (*a*) a cube with an edge of 3 nm, (*b*) a cylinder with radius R = 2 nm and height h = 4 nm, (*c*) a rectangular prism with edges of a = 2, b = 3, c = 4 nm, (*d*) a rod with radius R = 2 nm and height h = 10 nm, (*e*) a ring with outer radius R = 4 nm, inner radius r = 2 nm and height h = 4 nm, (*f*) a rectangular prism with a = 1, b = 2, c = 4 nm, (*g*) a disc with R = 2 nm and h = 1 nm and (*h*) a flat disc with R = 5 nm and h = 1 nm. Theoretical scattering patterns (Fig. 3) of the model shapes were evaluated up to s = 3.5 nm^{-1} and processed by *GNOM* (Svergun, 1992), and *AMBIMETER* was applied to identify the topologies consistent with the given geometrical shape (Table 1). Not surprisingly, the number of consistent topologies *m* depends on the anisometry of the shape: the flatter the particle, the larger the number of structural neighbours. To correlate this number with the results of *ab initio* reconstruction, low-resolution shapes of these geometrical bodies were restored by *DAMMIF* (Franke & Svergun, 2009). The results of individual reconstructions were superposed and averaged using *SUPCOMB* (Kozin & Svergun, 2001) and *DAMAVER* (Volkov & Svergun, 2003), respectively. Both the averaged

NSD within the ten reconstructed models and the NSD between the averaged model and the target shape correlate well with the number of consistent shape topologies (Table 1). This result supports an intuitive expectation that the higher the number of structural neighbours, the more ambiguous is the reconstruction, leading to a higher probability of obtaining a false positive in the *ab initio* modelling.

It should also be mentioned that if the number of consistent shape topologies m is relatively small, *AMBIMETER* can be considered as a tool for rapid shape determination. For this purpose, each sphere of the skeleton is filled with (several dozens of) beads of smaller size and these models are averaged by *DAMAVER* following the conventional procedure. As an example, the shape generated in this way for the prism (*c*) with six consistent topologies yields an NSD of 0.76 to the original geometrical body. The opportunity for rapid shape determination makes *AMBIMETER* a reasonable alternative to Monte Carlo-based modelling.

3.2. Effect of experimental errors

The above calculations addressed the principal question of the ambiguity of the SAS data *per se* by characterizing the non-uniqueness of the reconstruction based on noise-free data. For real data, accounting for experimental errors is important and we also made simulations to assess their influence. *AMBIMETER*, similar to *DAMMIF*, utilizes a smooth regularized scattering intensity computed from the raw data by the indirect Fourier transform (*GNOM*). This processing makes the method robust with respect to the statistical errors in the data, which makes the method more stable but also makes it difficult to analyze the influence of the errors. Test calculations were therefore first performed to estimate the effect that noisy data without regularization would have caused. For this purpose, the scattering profile of the prism [case (*c*)] was randomized with various noise levels and the dissimilarity measure d^2 was computed against the ideal intensity curve in the default range $sR_g \leq 4$. For the experimental uncertainties of 5% relative error the obtained d^2 was well below the threshold of 0.0016, and only when the noise level reached about 10% was the threshold slightly exceeded. On the other hand, even at a noise level of 20% only seven models (the number obtained by *AMBIMETER* is six) yielded a χ^2 value below 1.5 (Supplementary Fig. S1). For the yet more oblate shape of disc (*g*), the corridor with noise of 3% accommodates 40 skeletons which yield $\chi^2 \leq 1.5$ to the target curve randomized with this noise level. For noise levels

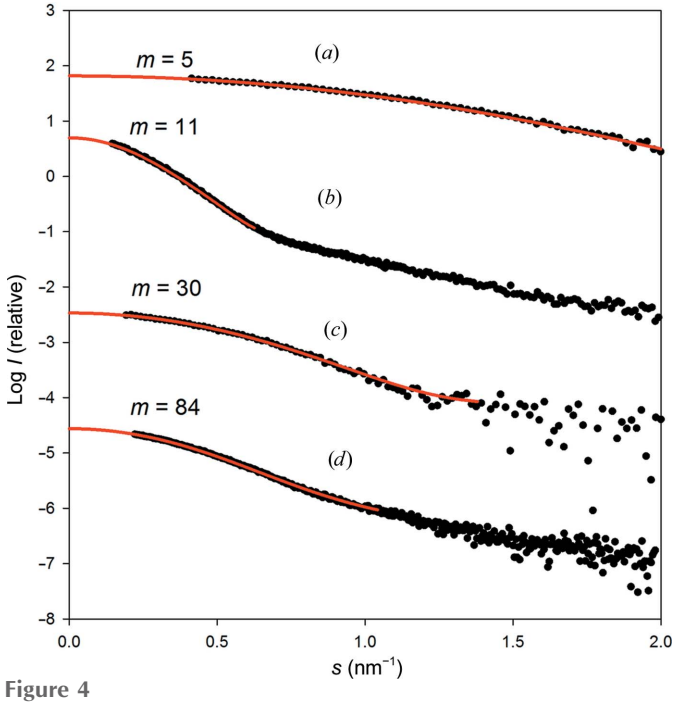


Figure 4
Scattering profiles from four test proteins: lysozyme (*a*), V_1-ATPase (*b*), BSA (*c*) and reverse transcriptase (*d*). Experimental data are shown as black dots and fits from the best compatible skeleton (selected in the range up to $sR_g = 4$) are shown as red solid lines. The number of structural neighbours are specified for each curve.

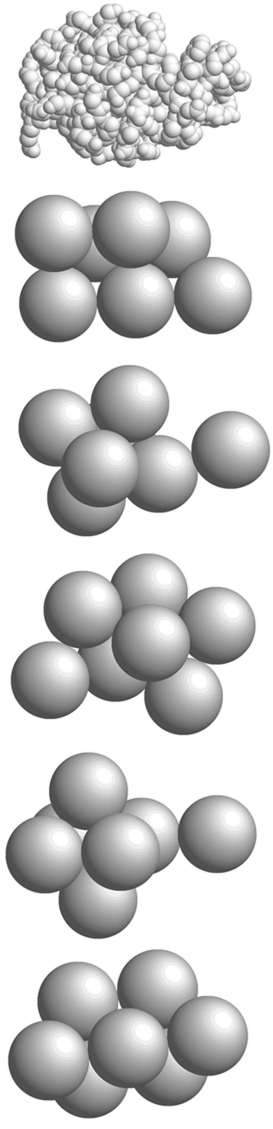


Figure 5
Five skeletons compatible with the lysozyme scattering profile. The crystal structure of lysozyme is displayed in space-filling mode for comparison.

of 5 and 10% the corresponding numbers were 164 and 694, respectively. The latter number is of the order of the *AMBIMETER* result (416 consistent topologies, *a-score* = 2.62) which suggests that the d^2 threshold of 0.0016 corresponds to about a 10% noise level. Finally, the randomized curves for cases (*c*) and (*g*) with noise levels of 10% were processed by *GNOM* (in sR_g ranges from 0.3 to 4.0) and passed to *AMBIMETER*. For the low-ambiguity prism (*c*), exactly the same six skeletons were obtained as for the ideal case. For the ambiguous disc (*g*), 401 compatible shape topologies were found, *i.e.* also very close to the *AMBIMETER* result for the ideal curve (416). Taken together, these results suggest that *AMBIMETER* is stable with respect to statistical errors up to relative uncertainties of about 10% (in a typical SAXS experiment, the data are typically much more accurate, especially in the used range of $sR_g \leq 4$; see, for example, the next section). Of course, in the presence of systematic errors (for example caused by sample polydispersity or buffer-subtraction issues), *AMBIMETER* would assess a lower limit of possible ambiguity in the reconstruction.

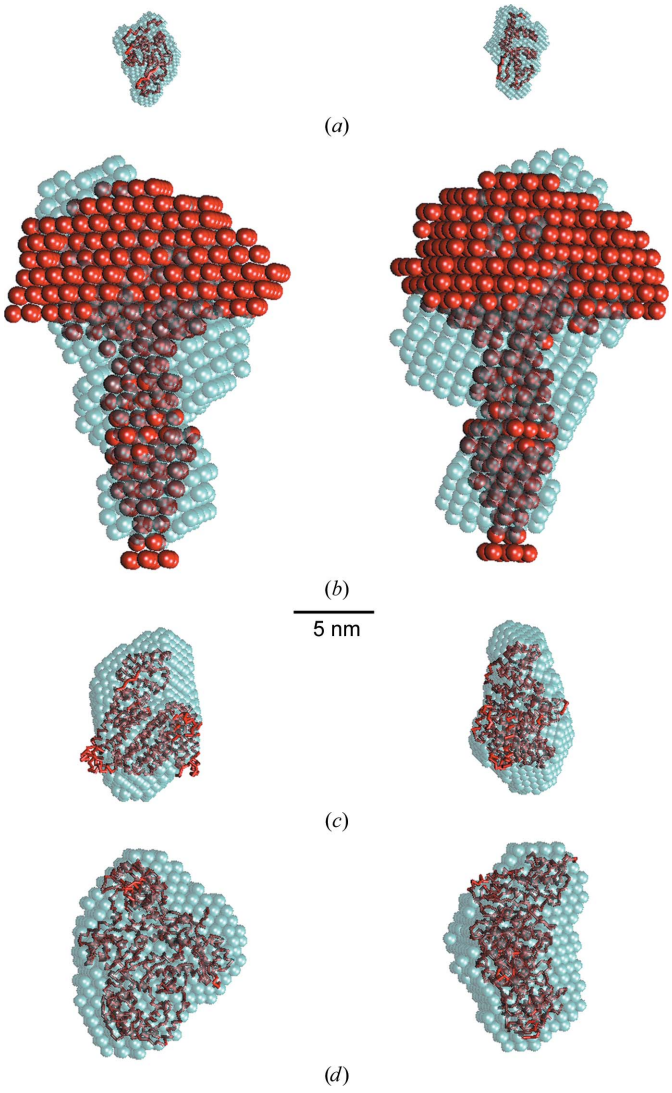


Figure 6
Averaged models built from the skeletons (semitransparent cyan beads) compared with the available structures (in red). (*a*) Lysozyme; (*b*) V_1-ATPase; (*c*) BSA; (*d*) reverse transcriptase. For lysozyme, BSA and reverse transcriptase the atomic models are shown as backbones; for V_1-ATPase the *ab initio* bead model published in Koch *et al.* (2003) is displayed in space-filling mode. The right view is rotated by 90° about the vertical axis.

3.3. Tests on experimental profiles from proteins

After testing on the above set of geometrical bodies, the algorithm was applied to experimental profiles from proteins with known atomic structures. Synchrotron-radiation X-ray scattering data from lysozyme, bovine serum albumin (BSA), V_1-ATPase, reverse transcriptase and immunoglobulin M (IgM) were collected following standard procedures on the X33 beamline (Blanchet *et al.*, 2012) at the European Molecular Biology Laboratory, DESY, Hamburg. Details of the experimental procedures are given elsewhere (Grüber *et al.*, 2000; Koch *et al.*, 2003; Svergun *et al.*, 2001; Volkov *et al.*, 2003). Data processing (averaging, buffer subtraction *etc.*) was performed using *PRIMUS* (Konarev *et al.*, 2003) and the resulting scattering profiles are shown in Fig. 4.

For lysozyme, with a globular shape, *AMBIMETER* found $m = 5$ consistent shape skeletons (Fig. 5). They all resembled the actual lysozyme shape and after averaging with *DAMAVER* produced a model with an NSD of 0.79 to the atomic model of lysozyme (Diamond, 1974; Fig. 6*a*), whereby the mean value of the NSD between the models was 1.08. V_1-ATPase (Fig. 6*b*) with an elongated shape (Grüber *et al.*, 2000) has $m = 11$ shape categories compatible with its scattering profile, with a mean value of the NSD in *DAMAVER* of 1.07. Interestingly, no oblate shapes were obtained by *AMBIMETER*, in contrast to Koch *et al.* (2003), where it was demonstrated that the use of symmetry constraints might yield wrong anisometry (a flat particle was generated in the *ab initio* reconstruction with the threefold axis). BSA, the shape of which can be approximated by a flat triangular prism, yields $m = 30$ consistent shape categories. The mean value of the NSD between the individual models was 1.06 and the averaged models yielded an NSD of 1.15 to the crystallographic model of BSA (Majorek *et al.*, 2012; Fig. 6*c*). The scattering curve of the complex boat-like shaped reverse transcriptase (Wang *et al.*, 1994; Fig. 6*d*) resulted in $m = 84$ skeletons generated by *AMBIMETER* with a mean value of the NSD of 1.22.

The case of IgM, a high-molecular-mass assembly with a rather flat shape, clearly stands out in these examples given that the shape reconstruction was found to be notoriously ambiguous (Volkov *et al.*, 2003). Indeed, given the experimental data in Fig. 7 *AMBIMETER* found a very large number $m = 677$ of consistent shape skeletons. Obviously, such a high ambiguity measure suggests that unconstrained *ab initio* modelling would probably not be able to depict the 'correct' solution. Indeed, 100 independent bead-modelling runs of *DAMMIF* in *P*1 resulted in four different clusters according to *DAMCLUST*, all of which neatly fitted the data (Fig. 7), and none of these clusters (Fig. 8) resembled the expected flat pentameric shape of the assembly (Volkov *et al.*, 2003). The

requirement for fivefold symmetry consistent with the pentameric stoichiometry of IgM coupled with the condition of shape oblateness yielded consistent results in individual bead-modelling runs. The averaged shape (Fig. 8) generated by *DAMAVER* from these constrained reconstructions agrees well with the previously suggested theoretical model of IgM (Perkins *et al.*, 1991). This example shows that for flat shapes that yield maximum ambiguity and which are the most difficult to restore, the addition of symmetry constraints may allow one to overcome the hurdle leading to meaningful shape reconstruction.

4. Conclusions

A novel concept presented here addresses perhaps the most challenging problem of modern small-angle scattering of biological macromolecules: the ambiguity of shape reconstruction from a given scattering profile. We generated a library of scattering patterns from shape skeletons comprehensively describing a manifold of low-resolution particle-shape topologies. Any experimental scattering profile can be readily positioned in the library on a normalized scale I/I_0 *versus* sR_g, eliminating the size information and keeping the topology information only. The number of skeletons consistent with the given SAS data provides a measure of ambiguity associated with the data and thus of potential non-uniqueness of the *ab initio* modelling. As an additional bonus, a simple averaging of the consistent skeletons constitutes a simple procedure which is sufficient to rapidly assess the overall appearance of the particle [of course, for more detailed analysis, the existing Monte Carlo-based approaches (see, for example, Franke & Svergun, 2009; Svergun, 1999) should still be used].

The results of *AMBIMETER* analysis of the theoretically generated and experimental data indicate, not unexpectedly, that the particles with isometric shapes generally represent the

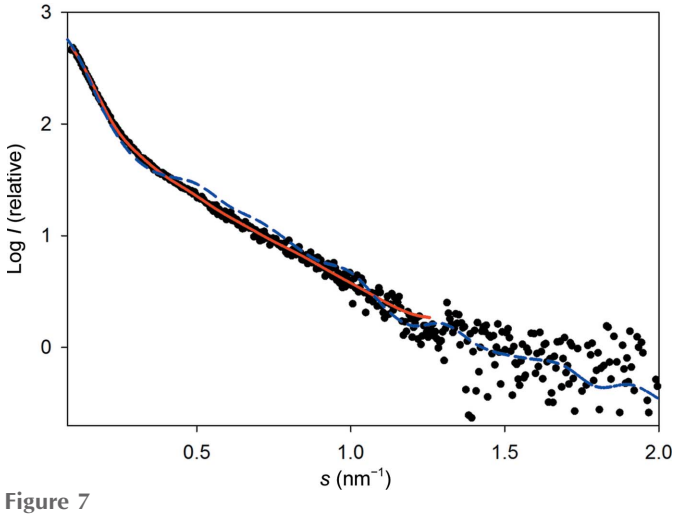


Figure 7
Scattering profiles from IgM. Experimental data are shown as black dots and fits from the typical shape reconstructed with symmetry and anisometry restrictions and from the theoretical model (Perkins *et al.*, 1991) are shown as red solid and blue dashed lines, respectively.

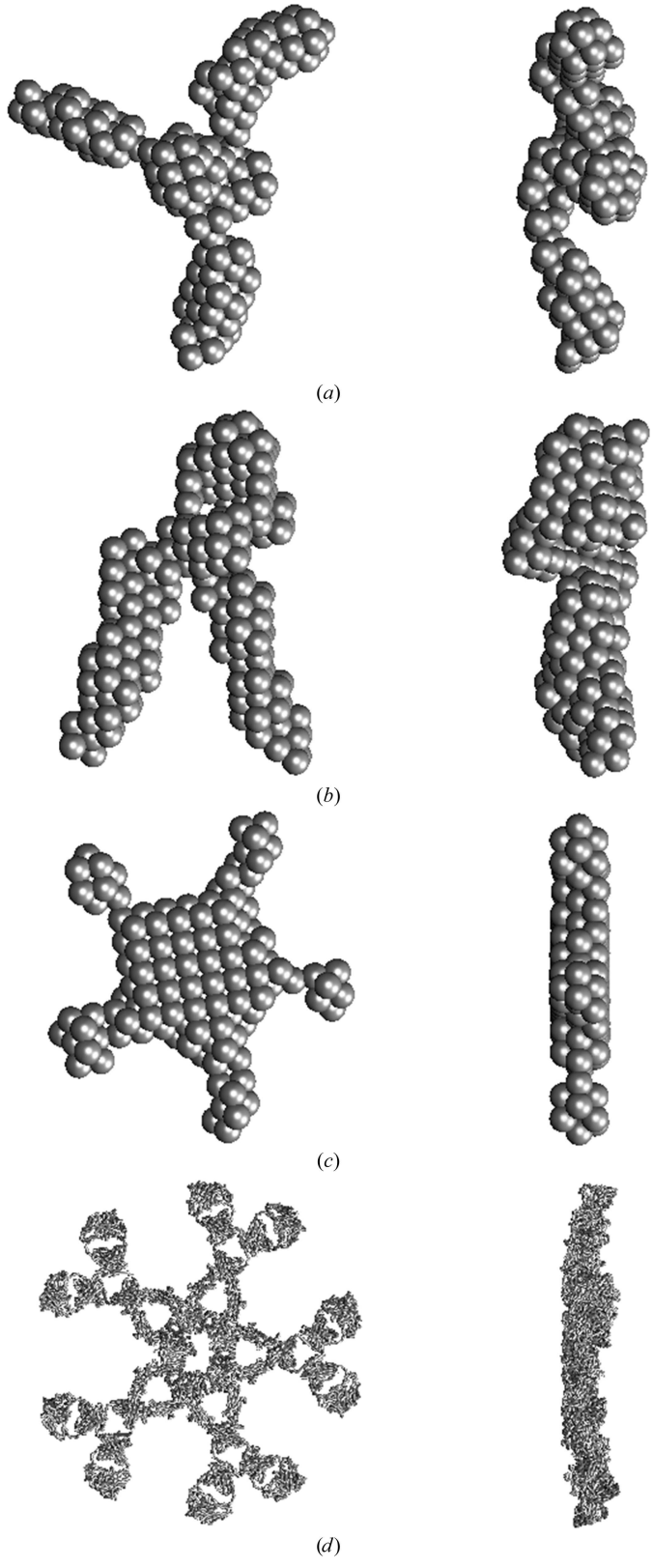


Figure 8
Structural models of IgM. (*a*, *b*) Representatives of the most distinct clusters reconstructed in *P*1. (*c*) Averaged model built with fivefold symmetry and the requirement for oblateness. (*d*) The theoretical model suggested by Perkins *et al.* (1991). Bead models are displayed in space-filling mode and the theoretical model is displayed as the backbone. The right view is rotated by 90° about the vertical axis.

easiest objects for reconstruction. Interestingly, the very fact of pronounced anisometry does not prevent unambiguous shape determination, and elongated particles are expected to be reconstructed well. Oblate particles are, as already empirically found in previous publications, the most difficult objects and their unique shape analysis is hardly possible without additional information. The use of *AMBIMETER* allows one to immediately obtain a 'validation stamp' for a given data set based on the number of compatible skeletons. Following our results, an *a-score* below 1.5 practically guarantees a unique *ab initio* shape determination, whereas when the *a-score* is in the range 1.5–2.5 care should be taken, perhaps involving cluster analysis, and for *a-score*s exceeding 2.5 unambiguous reconstruction without restrictions (for example, on symmetry and/or anisometry) is highly unlikely.

AMBIMETER is freely available in the v.2.6 release of *ATSAS* (http://www.embl-hamburg.de/biosaxs/software.html) and can be used either interactively or in data-analysis pipelines as an indicator of the potential ambiguity of shape determination or even as a rapid shape-determination tool.

References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Blanchet, C. E., Zozulya, A. V., Kikhney, A. G., Franke, D., Konarev, P. V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., Svergun, D. I. & Roessle, M. (2012). *J. Appl. Cryst.* **45**, 489–495.

Chacón, P., Morán, F., Díaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.

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

Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P. & Fieschi, F. (2010). *J. Struct. Biol.* **169**, 45–53.

Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.

Gabel, F. (2012). *Eur. Biophys. J.* **41**, 1–11.

Grüber, G., Svergun, D. I., Godovac-Zimmermann, J., Harvey, W. R., Wieczorek, H. & Koch, M. H. J. (2000). *J. Biol. Chem.* **275**, 30082–30087.

Guinier, A. (1939). *Ann. Phys. (Paris)*, **12**, 161–237.

Jacques, D. A. & Trewhella, J. (2010). *Protein Sci.* **19**, 642–657.

Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). *Q. Rev. Biophys.* **36**, 147–227.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.

Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.

Majorek, K. A., Porebski, P. J., Dayal, A., Zimmerman, M. D., Jablonska, K., Stewart, A. J., Chruszcz, M. & Minor, W. (2012). *Mol. Immunol.* **52**, 174–182.

Perkins, S. J., Nealis, A. S., Sutton, B. J. & Feinstein, A. (1991). *J. Mol. Biol.* **221**, 1345–1366.

Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.

Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.

Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.

Svergun, D. I., Koch, M. H. J., Timmins, P. A. & May, R. P. (2013). *Small Angle X-ray and Neutron Scattering from Solutions of Biological Macromolecules*. Oxford University Press.

Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.

Takahashi, Y., Nishikawa, Y. & Fujisawa, T. (2003). *J. Appl. Cryst.* **36**, 549–552.

Volkov, V. V., Lapuk, V. A., Kayushina, R. L., Shtykova, E. V., Varlamova, E. Y., Malfois, M. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 503–508.

Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.

Wang, J., Smerdon, S. J., Jäger, J., Kohlstaedt, L. A., Rice, P. A., Friedman, J. M. & Steitz, T. A. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 7242–7246.