Structural characterization of proteins and complexes using small-angle X-ray solution scattering

Haydyn D.T. Mertens, Dmitri I. Svergun *

European Molecular Biology Laboratory-Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany

ARTICLE INFO

ABSTRACT

Small-angle scattering of X-rays (SAXS) is an established method for the low-resolution structural characterization of biological macromolecules in solution. The technique provides three-dimensional low-resolution structures, using *ab initio* and rigid body modeling, and allow one to assess the oligomeric state of proteins and protein complexes. In addition, SAXS is a powerful tool for structure validation and the quantitative analysis of flexible systems, and is highly complementary to the high resolution methods of X-ray crystallography and NMR. At present, SAXS analysis methods have reached an advanced state, allowing for automated and rapid characterization of protein solutions in terms of low-resolution models, quaternary structure and oligomeric composition. In this communication, main approaches to the characterization of proteins and protein complexes using SAXS are reviewed. The tools for the analysis of proteins in solution are presented, and the impact that these tools have made in modern structural biology is discussed.

1. Introduction

Small-angle scattering (SAS) of X-rays (SAXS) and neutrons (SANS) is a powerful method for the analysis of biological macromolecules in solution. Great progress has been made over the years in applying this technique to extract structural information from non-crystalline samples in the fields of physics, materials science and biology (Feigin and Svergun, 1987). Over the last decade, major advances in instrumentation and computational methods have led to new and exciting developments in the application of SAXS to structural biology. Active research is now being conducted by an increasing number of laboratories on advancing *ab initio* and rigid body modeling methods, the calculation of theoretical scattering curves from atomic models and the characterization of quaternary structure and intrinsic flexibility. In addition, advances in the automation of data collection and analysis make high throughput applications of SAXS experiments tractable (Hura et al., 2009; Round et al., 2008). Such developments have generated a renewed interest in the wider applications of the technique in the structural biology community. The present review is focused on the characterization of protein structure and complex formation, but the method is widely used for other macromolecular structures, e.g. RNA (Doniach and Lipfert, 2009; Rambo and Tainer, 2010).

Production of good quality samples is a prerequisite for a successful structural study by any method, and modern approaches to protein expression and purification used in structural biology laboratories help to facilitate this. Like the high resolution methods X-ray crystallography and nuclear magnetic resonance (NMR), SAS requires milligram amounts of highly pure, monodisperse protein that remains soluble at high concentration. However, while sample requirements are similar for the three methods (noting that an additional crystallization step is not required for the solution methods) a distinct advantage of SAXS is the speed of both data collection and sample characterization. On a modern synchrotron, scattering data can be collected in seconds, allowing an almost immediate characterization of the sample and the sample quality through the extraction of several overall parameters from the radially averaged scattering pattern. SAXS can thus be used as a method for the rapid screening of samples in various aqueous solvents/additives, including e.g. identification and optimization of crystallization conditions (Bonnete et al., 1999; Hamiaux et al., 2000).

In this review the discussion of SAS focusses on the elastic scattering of X-rays, SAXS, where dissolved macromolecules are exposed to a collimated and (for synchrotrons) focussed X-ray beam and the scattered intensity *I* is recorded by a detector as a function of the scattering angle (Fig. 1A). For an in-depth review of the theory behind SAXS the reader is directed to text-books and recent reviews (Feigin and Svergun, 1987; Koch et al., 2003; Putnam et al., 2007; Svergun, 2007; Svergun and Koch, 2003; Tsuruta and Irving, 2008), here some of the basic concepts will be presented with a focus on the characterization of proteins and protein complexes.

* Corresponding author. Fax: +49 40 89902 149.
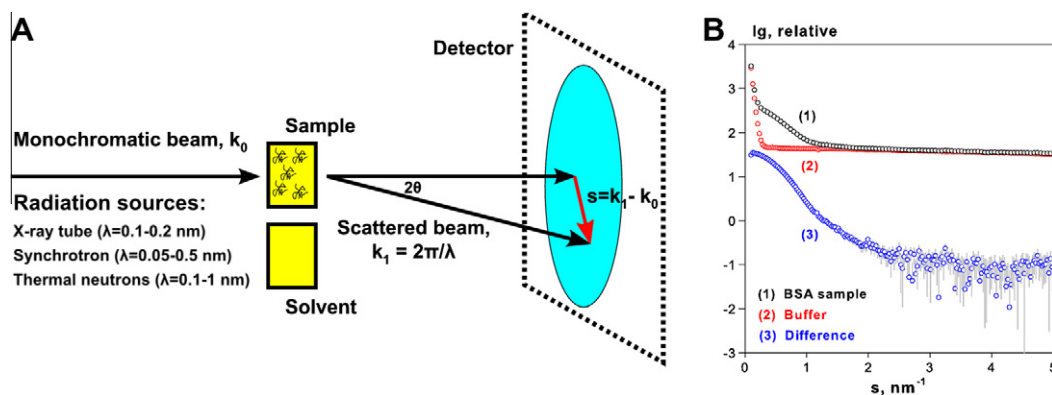E-mail address: Svergun@EMBL-Hamburg.DE (D.I. Svergun).

Fig. 1. Schematic representation of a typical SAS experiment and radially averaged data. (A) Standard scheme of a SAS experiment. (B) X-ray scattering patterns from a solution of BSA measured at X33 (DORIS, Hamburg) in 50 mM HEPES, pH 7.5, solvent scattering and the difference curve (containing the contribution from the protein alone, scaled for the solute concentration, 5 mg/ml).

Scattering of X-rays by a solution of biomolecules is dependant on the number of biomolecules in the illuminated volume (i.e. to the solute concentration) and the excess scattering length density (often also called the contrast). For X-rays, the excess scattering length density, $\Delta\rho(r)$, comes from the difference in the electron density of the solute and solvent which, for biomolecules in aqueous solutions is very small. Consequently, synchrotron SAXS beamlines and laboratory sources must be optimized for the minimization of the contribution of background.

Dilute aqueous solutions of proteins, nucleic acids or other macromolecules give rise to an isotropic scattering intensity, which depends on the modulus of the momentum transfer s ($s = 4\pi\sin(\theta)/\lambda$, where 2θ is the angle between the incident and scattered beam):

$$I(s) = \langle \mathbf{I}(\mathbf{s}) \rangle_\Omega = \langle \mathbf{A}(\mathbf{s})\mathbf{A}^*(\mathbf{s}) \rangle_\Omega \qquad (1)$$

where the scattering amplitude $\mathbf{A}(\mathbf{s})$ is a Fourier transformation of the excess scattering length density, and the scattering intensity is average over all orientations (Ω). Following subtraction of the solvent scattering, the background corrected intensity $I(s)$ is proportional to the scattering of a single particle averaged over all orientations.

The scattering patterns generated from a dilute solution of macromolecules are typically presented as radially averaged one-dimensional curves (Fig. 1B). From these curves several overall important parameters can be directly obtained providing information about the size, oligomeric state and overall shape of the molecule. However, advances in computational methods have now made it possible to not only extract these simple parameters, but to also determine reliable three-dimensional structures from scattering data. Low-resolution (1–2 nm) SAXS models can be determined *ab initio* or through the refinement of available high-resolution structures and/or homology models. While the former analysis provides a low-resolution shape of the molecule in question and often adds insight to the biological problem at hand, the latter combination of SAXS and complementary data is a powerful method for the determination of the organisation of macromolecular complexes. In addition to structure determination SAXS is routinely used for the validation of structural models, the quantitative analysis of oligomeric state and the estimation of volume fractions of components in mixtures/polydisperse systems. While SAXS has been readily employed for the analysis of flexible systems including solutions of intrinsically unfolded proteins, methods were often restricted to the determination of simple geometric parameters. A renaissance in the study of such systems by structural biologists over the last 5–10 years has led to the development of novel approaches for the analysis of flexible systems including multi-domain and intrinsically unfolded proteins (Bernado et al., 2005, 2007; Obolensky et al., 2007).

SAXS is a technique that can probe structure on an extremely broad range of macromolecular sizes (Feigin and Svergun, 1987). Small proteins and polypeptides in the range of 1–10 kDa, macromolecular complexes and large viral particles up to several hundred MDa can all be measured with modern instrumentation under near native conditions. It is often attractive to laboratory based researchers as the amount of material required for a complete study is relatively low (typically 1–2 mg protein), and almost any biologically relevant sample conditions can be used. The effect of changes to sample environment (pH, temperature, salt concentration and ligand/co-factor titration) can be easily measured and, moreover, at high-brilliance synchrotron beamlines time-resolved experiments can be conducted (Lamb et al., 2008a,b; Pollack and Doniach, 2009; West et al., 2008).

It should be noted here that the elastic scattering of neutrons (SANS) is also widely used to characterize macromolecular solutions. Moreover, many approaches described below for SAXS are also applicable for SANS, where the excess scattering length density (contrast) is due to the nuclear (and sometimes spin) scattering length density instead of the electron density. In SANS, samples highly absorbing to X-rays (e.g. solvents containing high salt) can be measured, and the samples will not suffer from radiation damage. Most importantly, contrast variation by hydrogen/deuterium exchange can be used yielding precious additional information about the structure of macromolecular complexes. The disadvantages to SANS are that it usually requires more material than is required for SAXS, buffer subtraction is often difficult due to the high incoherent hydrogen scattering and that the measurements cannot be done on a laboratory source. Overall, SANS is a powerful complementary tool to SAXS (Ibel and Stuhrmann, 1975; Petoukhov and Svergun, 2006; Wall et al., 2000; Whitten and Trewhella, 2009).

2. Overall SAXS parameters and rapid sample characterization

Although sophisticated approaches have now been developed for the determination of three-dimensional structure from scattering data (see the following sections), several overall invariant shape and weight parameters can be extracted directly from scattering curves enabling fast sample characterization. These parameters include: the molecular mass (MM), radius of gyration (R_g), hydrated particle volume (V_p) and maximum particle diameter (D_{max}). The Guinier analysis developed by A. Guinier in the 1930s (Guinier, 1939) is still the most straightforward method for the extraction of the forward scattering intensity $I(0)$ and the radius of gyration, R_g. For a monodisperse solution of globular macromolecules the Guinier equation is defined as:

$$I(s) = I(0) \exp\left(-\frac{1}{3}R_g^2 s^2\right) \qquad (2)$$

In principle, $I(0)$ and R_g can be extracted from the y-axis intercept and the slope of the linear region of a Guinier plot ($\ln[I(s)]$ versus s^2), respectively (Fig. 2A and B). However, the range (s_{min} to s_1) over which the Guinier approximation is valid for each measured scattering curve must be considered. The lower limit of this range, s_{min}, is usually restricted by the experimental set-up, and for an ideal sample is taken to be the minimum angle for which intensity is recorded. The Guinier approximation is based on a power law expansion, used to describe the linear dependence of $\ln[I(s)]$ on s^2 (Guinier, 1939). When extended to larger values of s, the higher order terms in the expansion begin to significantly contribute to the scattering intensity, breaking this linear dependence. Given that the power value $(sR_g)^n$ decreases with n for $sR_g < 1$ and increases for $sR_g > 1$, $s_1 < 1/R_g$ is a reasonable estimate for the upper limit of the Guinier fit. However, it is often the case that the range $s_{min} < s < 1/R_g$ contains too few points, especially in the case of very large macromolecules. It is common in biological SAXS to extend this range up to $s_1 < 1.3/R_g$, so that a sufficient number of data points are available for the estimation of $I(0)$ and R_g. Practice shows that $1.3/R_g$ is a safe estimate for s_1, which does not introduce systematic deviations from linearity.

A non-linear Guinier plot is a strong indicator of poor sample quality. Improper background subtraction, the presence of attractive or repulsive inter-particle effects and sample polydispersity result in deviations from linearity (Fig. 2A and B). For example, samples that contain a significant proportion of non-specific aggregates yield scattering curves and Guinier plots with a sharp increase in intensity at very small values of s, while samples containing significant inter-particle repulsion yield curves and Guinier plots that show a decrease in intensity at small values of s. Note that it might still be possible to obtain "linear fits" even to data with significant concentration/aggregation effects (Fig. 2B, fits 1 and 3), albeit in rather short ranges, and care must be taken to avoid wrong results caused by these effects. Moreover, even a "long" linear Guinier plot does not always guarantee that a sample is monodisperse and researchers are advised to check sample monodispersity using methods such as dynamic light scattering before conducting SAXS measurements.

Samples often contain molecules that interact strongly with each other in a concentration dependent manner. In non-ideal solutions, strong attractive or repulsive inter-particle interactions modulate the recorded scattering intensity particularly at low angles ($s < 1$ nm^{-1}) and influence the parameters extracted from the SAXS curve. For example, attractive interactions can result in the overestimation of $I(0)$ (and thus MM) and R_g and repulsive interactions can result in these parameters being underestimated. The contribution of these interactions to the scattering intensity can be separated from that derived from the shape of the particles by
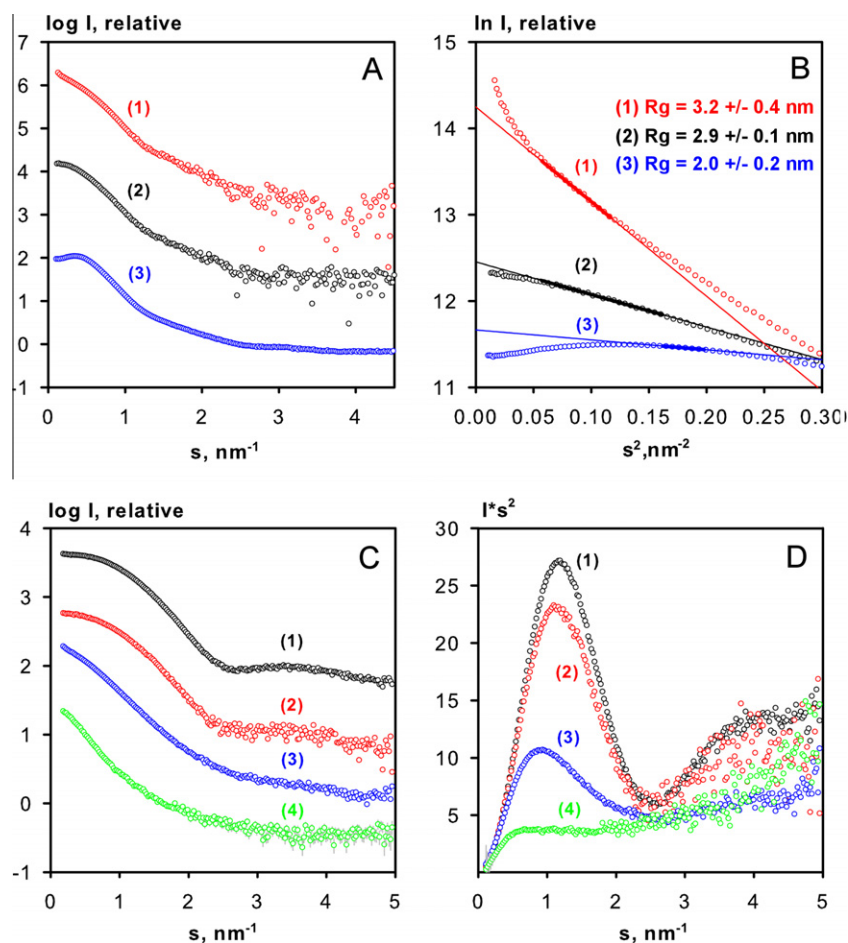

Fig. 2. Standard plots for characterization by SAXS. (A and B) SAXS curves and Guinier plots for BSA samples measured at X33 (DORIS, Hamburg) in different buffers showing (1) aggregation, (2) good data and (3) inter-particle repulsion. The Guinier fits for estimation of R_g and $I(0)$ are displayed, with the linear regions defining s_{min} and s_{max} used for parameter estimation indicated by the thick lines. (C and D) SAXS curves and Kratky plots for lysozyme samples measured at X33 (DORIS, Hamburg) showing (1) folded lysozyme, (2) partially unfolded lysozyme (in 8 M urea), (3) partially unfolded lysozyme at 90 °C and (4) unfolded lysozyme (in 8 M urea at 90 °C). Plots are arbitrarily displaced on the vertical axis for clarity with the exception of (D), where all curves have been scaled to the same forward scattering intensity, $I(0)$.

recording scattering patterns at several concentrations. From a concentration series it is usually possible to identify a sample concentration where inter-particle interactions are negligible or to extrapolate the data to infinite dilution to obtain an "ideal" curve for further structural analysis. In addition, characterization of the concentration dependent behavior of the sample provides an important source of information that has been used to determine interaction potentials between macromolecules in solution (Tardieu et al., 1999). This can help to define conditions for crystallization, which typically require weakly attractive interactions (for detailed reviews on the determination of interaction potentials from SAXS data see (Koch et al., 2003) and (Finet et al., 2004).

With knowledge of the limitations of the Guinier approximation in mind, the method is an essential first step in sample characterization by SAXS. From visual inspection of the Guinier plot samples can be screened for non-specific aggregation (Fig. 2A and B, curve 1), the presence of inter-particle repulsion (Fig. 2A and B, curve 3), the prevalent oligomerisation state in solution estimated and the invariant shape and weight parameters, R_g and $I(0)$ extracted. In the past, Guinier analysis was always done interactively; recently, automated procedures have become available. In particular, the program *AUTORG* (Petoukhov et al., 2007) employs statistical methods to optimize the range of s, to detect possible aggregation or repulsive interactions and, based on this, to evaluate the quality and reliability of the extracted parameters.

Following on from the Guinier analysis the *MM* can be estimated based on knowledge of the forward scattering intensities and concentrations of both the macromolecule of interest and a standard such as bovine serum albumin. This estimate requires normalization against the solute concentrations for the two measurements, and the accuracy of the *MM* estimate is limited (Mylonas and Svergun, 2007). An alternative approach to *MM* estimation that is more suited to solutions with significant lipid, carbohydrate or nucleic acid content (including for example, glycoproteins or protein–lipid complexes) involves the use of water as a standard (Orthaber et al., 2000).

Independent from the Guinier analysis, the hydrated particle volume (V_p) can be obtained from the data on a relative scale, avoiding inaccuracies in parameter estimation caused by errors in concentration measurement. Assuming a uniform electron density inside the particle, V_p is estimated following Porod's equation (Porod, 1982):

$$V_p = 2\pi^2 I(0)/Q, \quad Q = \int_0^\infty s^2 I(s) \cdot ds \tag{3}$$

where Q is the so-called Porod invariant.

For real macromolecules the electron density is of course not uniform, however, at sufficiently high *MM* (>30 kDa), the subtraction of an appropriate constant from the scattering data generates a reasonable approximation to the scattering of the corresponding homogenous body. The particle volume, V_p allows one to make an alternative estimate of the *MM* with the added advantage that this estimate is independent of errors in the sample concentration. Typically, for a globular protein V_p (in nm^3) is 1.5–2 times the *MM* (in kDa). While this estimate is a crude approximation only, with prior knowledge of the expected size/state of the sample, a rough indication of the homogeneity of the sample can in this way be made available directly following measurement. A Web server has recently been made available allowing one to estimate the *MM* from SAXS data based on this approach (Fischer et al., 2010).

Due to the limitations of the Guinier approximation, the extraction of R_g and $I(0)$ from scattering data is also routinely done through the use of indirect Fourier transform methods. Fourier transformation of the scattering intensity yields the distance distribution function, $p(r)$:

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty s^2 I(s) \frac{\sin sr}{sr} ds \tag{4}$$

The distance distribution function is a real space representation of the scattering data and allows one to graphically display the peculiarities of the particle shape (Fig. 3). For example, globular particles yield bell-shaped profiles with a maximum at approximately $D_{max}/2$ and multi-domain particles often yield profiles with multiple shoulders and oscillations corresponding to intra and inter-subunit distances. Computation of $p(r)$ is not straightforward as a limited range of $I(s)$ is available (from s_{min} to s_{max}), and direct Fourier transformation of the scattering curve from this finite number of points is not possible. A solution to this problem is the indirect Fourier transformation method first proposed by O. Glatter in the 1970s (Glatter, 1977). In this approach $p(r)$ is represented as a linear combination of K orthogonal functions φ_k, in the range [0, D_{max}], with D_{max} being a user defined variable:

$$p(r) = \sum_{k=1}^K c_k \phi_k(s_i) \tag{5}$$

The optimal coefficients c_k are sought through minimization of the functional:

$$\Phi = \chi^2 + \alpha P(p) \tag{6}$$

where the first term, χ^2 is the goodness of fit between the experimental data and that calculated by the direct transform of the $p(r)$ function (Eq. (7)), and the second (penalty) term, $P(p)$ ensures the smoothness of the $p(r)$ function (Eq. (8)).

$$\chi^2 = \frac{1}{N-1} \sum_{j=1}^N \left[\frac{I_{exp}(S_j) - cI_{calc}(S_j)}{\sigma(S_j)} \right]^2 \tag{7}$$

$$P(p) = \int_0^{D_{max}} [p']^2 dr \tag{8}$$

where N, σ and c are the number of data points, the standard deviations and scaling factor respectively. The regularizing multiplier α balances between the fit to the data and the smoothness of the $p(r)$. One should note here that fitting the experimental data using Eq. (6) is often employed in physical experiments including SAXS data analysis and interpretation. As will be presented below, *ab initio* and rigid body modeling methods will use a similar notion of a penalty term to ensure that physically sensible solutions are obtained.

In the indirect transform program GNOM (Svergun, 1992), the solution yielding the $p(r)$ function is evaluated using perceptual criteria, providing the user with the means to easily identify reliable solutions and to obtain an optimal value of D_{max}. This procedure has also been automated in the program *AUTOGNOM* (Petoukhov et al., 2007), where multiple GNOM runs are performed without user intervention across a range of D_{max} values. The parameters estimated from the indirect Fourier transform approach, $I(0)$ and R_g are typically more accurate than those obtained from a Guinier analysis as the entire scattering curve is used for their estimation.

For the study of protein folding the Kratky plot ($s^2 I(s)$ vs s) can be used as an indication of the folded/unfolded state (Doniach, 2001). Folded globular proteins typically yield a prominent peak at low angles (Fig. 2C and D, curve 1), whereas unfolded proteins show a continuous increase in $s^2 I(s)$ with s (Fig. 2C and D, curve 4). Flexible multi-domain proteins can also potentially be identified from the Kratky plot, displaying a mixture of characteristic features of both folded and unfolded proteins similar to that observed for a partially unfolded state (Fig. 2C and D, curves 2 and 3). However, recent studies (Bernadó, 2009) show that flexible proteins can yield Kratky plots that do not indicate the presence of flexibility
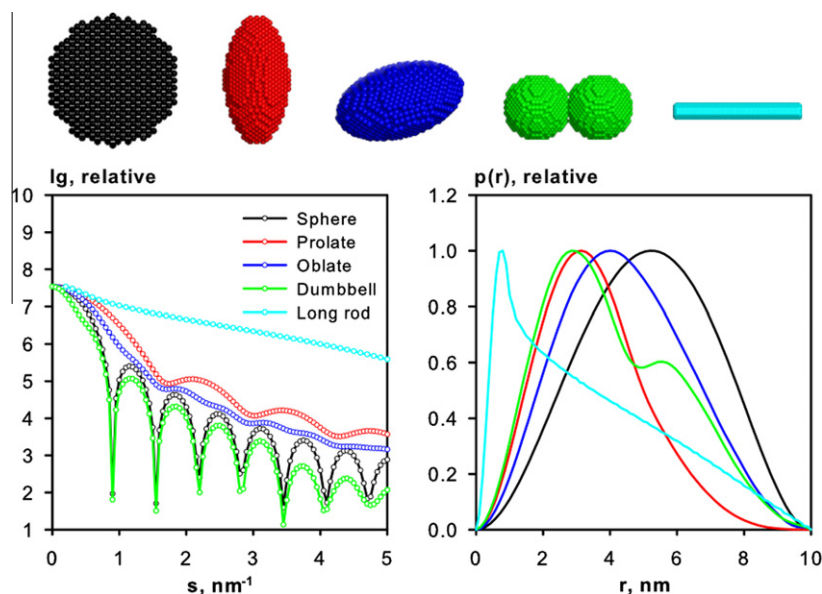
Fig. 3. Scattering intensities and distance distribution functions, $p(r)$ calculated for typical geometric shapes: Solid sphere (black), prolate ellipsoid (red), oblate ellipsoid (blue), two-domain (green) and long rod (cyan). The bead models used for the calculation of scattering intensities are shown above the plots.

but suggest relatively rigid compact globular structures (see Section 6).

Rapid sample characterization under near native solution conditions is one of the major advantages of SAXS over other structural techniques. Overall parameters describing size, shape and volume can be extracted from SAXS patterns almost immediately following measurement and used to answer important biological questions. However, in addition to sample characterization from extracted parameters only, characterization of the three-dimensional structure of the macromolecule or complex from the scattering curves is also possible. One approach is to use the SAXS profile to search a database of calculated scattering curves based on deposited structures in the PDB, with the aim of finding models that describe the measured data. In a recent high throughput study the *DARA* database (Sokolova et al., 2003) was successfully used to find closely matching structures of a series of proteins measured by SAXS (Hura et al., 2009). A more direct approach is to take advantage of modern methods of *ab initio* reconstruction from scattering data and combine this information with other complementary structural and biochemical data. The following two sections focus on the recent developments in the determination of structure from SAXS.

3. *Ab initio* methods

The reconstruction of low-resolution 3D models from SAXS data alone is now a standard procedure and as such can also be considered a rapid characterization tool. The basic principles behind shape determination from 1D SAXS data were established in the 1960s, where scattering patterns were computed from different geometrical shapes and compared with experimental data. These trial-and-error methods were superseded in the 1970s through the introduction of a spherical harmonics representation by Stuhrmann (1970b). In this representation a multipole expansion is used in order to derive a simple expression for the SAXS intensity $I(s)$:

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} |A_{lm}(s)|^2 \qquad (9)$$

This expression, which is a sum of independent contributions from the substructures corresponding to different spherical harmonics (where $A_{lm}(s)$ are the partial scattering amplitudes), al-

lows for rapid analytical computation of scattering patterns from known structures. This expression is readily incorporated into algorithms for the minimization of discrepancy χ between experimental and calculated scattering curves (Eq. (7)). The spherical harmonics formalism is heavily exploited by most advanced modeling programs.

In the initial *ab initio* approach, the shape of particles was described by an angular envelope function, the latter developed into a series of spherical harmonics (Stuhrmann, 1970a). This method was further developed into the first publicly available program SASHA (Svergun et al., 1996). While this method was a major breakthrough in the determination of low-resolution structure from SAXS it was restricted to particles without internal cavities.

More detailed *ab initio* reconstructions became possible through the development of automated bead-modeling. This approach was first proposed by Chacon et al. (1998)) and further implemented in different variations by other authors (Bada et al., 2000; Chacon et al., 2000; Svergun, 1999; Vigil et al., 2001; Walther et al., 2000). The most popular *ab initio* bead-modeling program in current use is DAMMIN (Dummy Atom Model Minimisation) (Svergun, 1999). The algorithm represents a particle as a collection of M ($\gg 1$) densely packed beads inside a constrained (usually spherical) search volume, with a maximum diameter defined by the experimentally determined D_{max} (Fig. 4). (2) Each bead is randomly assigned to the solvent (index = 0) or solute (index = 1), and the particle structure is described by a binary string X of length M. The shape reconstruction is conducted starting from a random initial approximation by simulated annealing (SA) (Kirkpatrick et al., 1983), minimizing the goal function as defined in Eq. (6). The discrepancy (χ^2) is evaluated in Eq. (7) between the experimental and calculated scattering intensities (the latter being rapidly computed using spherical harmonics). At each step in the SA procedure the assignment of a single bead is randomly changed leading to a new model X'. The solution is constrained by the penalty term, $P(X)$, requiring that the beads must be connected and the model compact to ensure that physically feasible low-resolution structures are generated. A multiphase version of DAMMIN bead-modeling is implemented in the program MONSA (Svergun and Nierhaus, 2000), also widely used when contrast variation data from SANS and/or multiple SAXS curves from components of a complex (e.g. a nucleoprotein complex) are available.
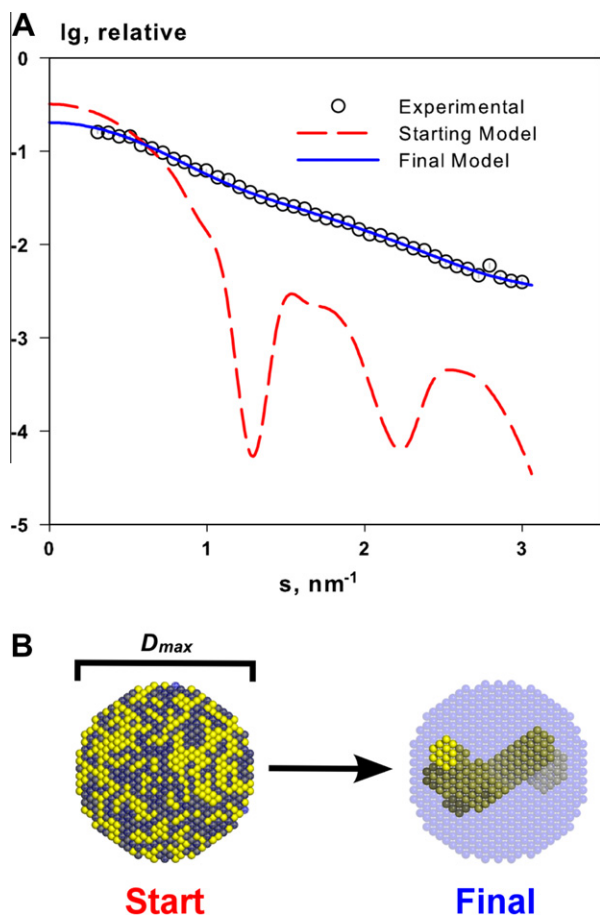
Fig. 4. *Ab initio* modeling procedure using DAMMIN. (A) Starting from a spherical search volume a fitting procedure is conducted until a final model is generated satisfying not only a fit to the experimental data, but also forming a compact and connected model of dummy atoms/beads. (B) The spherical search volume with beads assigned to the particle (yellow) and solvent (blue).

Although DAMMIN is reasonably fast, typically taking several minutes to several hours on modern computers, the widespread use of the program for high-throughput data analysis necessitates that, where possible, improvements in speed and performance should be sought. Thus DAMMIN has recently been re-written and optimized for performance. The new implementation is the program DAMMIF (Franke and Svergun, 2009) where F refers to *fast*.

DAMMIF differs in several important respects to DAMMIN. (1) The bounded search volume has been replaced by an unrestricted volume that can grow in size as needed during the SA procedure. This improvement helps to avoid artefactual boundary effects that may occur when using a search volume restricted by a slightly underestimated D_{max}. (2) Prior to the calculation of scattering amplitudes models that are disconnected are immediately rejected, and the calculation is performed only for interconnected models. (3) The necessary scattering amplitudes in terms of spherical harmonics are intelligently pre-computed for each bead contributing to the total scattering at least once. All these measures accelerate the modeling procedure by 25–40 times compared to the original DAMMIN, and already a number of publications are appearing in the literature citing the use of DAMMIF (Cheng et al., 2009; Heikkinen et al., 2009; Schmidt et al., 2010; Yadavalli et al., 2009).

The resolution of shape determination from bead models and envelope functions is limited by the assumption of a uniform electron density distribution within the particle, and consequently scattering curves can only be fit within a restricted range (typically up to s ~2.5 nm^{-1}, i.e. a resolution of 2.5 nm, where resolution d is defined as $d = 2\pi/s$). An alternative approach for proteins is to represent the molecule not as a collection of uniformly distributed beads, but as an assembly of dummy residues (DR). In this simplification of the atomic structure a globbic approximation is employed, where the atomic scattering densities for the atomic groups of each amino acid are combined to generate an effective scattering density of an average amino acid in water (Guo et al., 1995, 1999; Harker, 1953). Such an approach is used in the program GASBOR (Svergun et al., 2001), where the DR scattering represents the scattering from an amino acid averaged over the abundance of amino acids in proteins. The program starts from a randomly distributed "gas" of DRs in a spherical search volume defined by D_{max}. The number of DRs is equal to the number of residues in the protein sequence, and their positions are determined by SA driven minimization using Eq. (6), with a penalty term requiring that the DRs form a *protein-like* or folded chain-compatible structure. The latter conditions are ensured by requiring, in particular, that the average distance histogram of neighboring DRs in the model is similar to that for globular proteins. As the limitation of particle homogeneity is abrogated the data can be fit to much higher angles than for the bead-modeling programs (up to $s < 10$ nm^{-1}). GASBOR is routinely used by structural biologists to determine the low-resolution structures of proteins and protein complexes (Chen et al., 2007; De Marco et al., 2009; Pavkov et al., 2008; Schmidt et al., 2010; Trindade et al., 2009). In the field of structure based development of protein therapeutics GASBOR was recently used to generate low-resolution *ab intio* models of PEGylated Haemoglobin (Svergun et al., 2008). In combination with multiphase modeling using MONSA the structures determined help explain how the increased vascular retention of such therapeutics is likely a result of the surface shielding and intermolecular repulsion associated with PEG conjugation.

One must note that *ab initio* methods, seeking to reconstruct a 3D shape from a 1D scattering pattern cannot provide a unique solution and, when running the programs multiple times, somewhat different models are obtained. A comparison of these models ensures that the most persistent features are identified from a number of models that fit the data equally well but show variation in shape. Thus the intrinsic ambiguity of SAXS data interpretation can be reduced and reliable average models obtained. The average *ab initio* SAXS model is conceptually similar to a mean average structure from an NMR ensemble, albeit at a much lower resolution. The programs SUPCOMB (Kozin and Svergun, 2001) and DAMAVER (Volkov and Svergun, 2003) have been developed for this purpose. SUPCOMB performs the rapid superposition of ensembles of models and calculates the degree of structural similarity. Further, SUPCOMB identifies the most probable/representative ensemble member and DAMAVER averages the superposed models over the ensemble yielding a smoothed model containing the most persistent features in the reconstructions. The use of these methods allows for the assessment of the uniqueness of the *ab initio* models determined from scattering data, but a researcher must still critically evaluate the chosen model based on the expected size and hydrated volume (with possible multimers also in mind). Finally, one must always remember that the scattering patterns are invariant to handedness such that an enantiomorphic model must always be considered (SUPCOMB/DAMAVER do allow for this enantiomorphism).

The *ab initio* analysis of proteins and protein complexes with SAXS is very usefully complemented by information from other methods. High-resolution crystal and NMR structures can be docked into the low-resolution SAXS models, and the shapes provided by electron microscopy (EM) can be used as starting volumes for bead-modeling (Svergun, 1999). Information regarding

expected symmetry and anisometry can be particularly useful for obtaining reliable *ab initio* models, and can also significantly speed up the computations. Symmetry restrictions associated with the space groups P2–P12, P222–P62, cubic and icosahedral symmetry can be explicitly defined in DAMMIN and GASBOR, and most of these symmetries are also applicable in DAMMIF and MONSA. The use of correct symmetry allows one to further restrain the solution to yield more detailed *ab initio* models; however, symmetry must always be employed with caution and models in P1 should also be calculated for control and comparison.

4. Computation of scattering from high-resolution models

Another method for the rapid characterization of proteins and complexes if high-resolution structures or homology models are known is the computation of scattering curves from atomic models, and the comparison of these predicted curves with the experimentally determined SAXS profiles (Hough et al., 2004; King et al., 2005; Vestergaard et al., 2005). Given a model, the theoretical scattering curve can be computed and fit to the measured data, with this computation taking into account the atomic scattering *in vacuo*, the excluded volume (the volume occupied by the biomolecule in solution that is inaccessible to solvent) and scattering from the hydration layer. A model that provides a good fit to the data is considered a valid description of the structure under the solution conditions used for the measurement.

A number of methods exist for the calculation of theoretical scattering curves from atomic models, these methods generally differ in their approach to the calculation of the atomic scattering intensity, the subtraction of solvent excluded volume and the way in which a hydrated surface with a solvent density higher than that of the bulk solvent is approximated. Traditionally, the Debye formula (Debye, 1915) is used for the calculation of atomic scattering, however, the time required for computation scales quadratically with the number of atoms and makes the method less useful for large proteins and complexes. Recent programs employ the globbic approximation (see Section 3) to speed up computation using the Debye formula (Yang et al., 2009). Arguably, the most efficient approach to the calculation of the scattering intensity from atomic models is the spherical harmonics approximation used in the programs CRYSOL (Svergun et al., 1995) for SAXS and CRYSON (Svergun et al., 1998) for SANS. In this approach the computation time scales linearly with the size of the model and it is highly accurate up to $s < 5.0$ nm^{-1}, but has been successfully used up to much higher resolutions (up to $s = 10$–15 nm^{-1}). CRYSOL/CRYSON employ a gaussian sphere approximation for the calculation of the excluded volume (Fraser et al., 1978) and use spherical harmonics to calculate an envelope at the surface of the atomic model to approximate the hydration layer. For prediction of the scattering intensity of theoretical models default parameters for the excluded volume and the excess scattering density of the hydration layer are used. It is assumed that the scattering density of the hydration layer is ~10% greater than that of the bulk, and this assumption has been verified experimentally in a combined SAXS/SANS study (Svergun et al., 1998). When used in fitting mode, the excess scattering density of the hydration layer is used as a fitting parameter and adjusted to best fit the scattering data up to a resolution of about 0.5 nm.

Computation of scattering from high-resolution models is often used to identify the biologically active conformations of crystal structures and help distinguish between alternative crystallographic dimers and/or higher oligomers (De Marco et al., 2009; Santiago et al., 2009). For example, a new crystallographic form of the protein complex of Cdt1 and Geminin was validated in solution by SAXS (De Marco et al., 2009). In this study a heterohexameric structure was observed in the crystal whereas a previous crystallographic study had shown only the existence of a heterotri-

mer. From the fit of the theoretical scattering curves computed with the program CRYSOL (Svergun et al., 1995) to the SAXS data (Fig. 5), the heterohexamer was identified as the correct model in solution.

New approaches are currently being developed, particularly for the accurate prediction of wide-angle scattering data (WAXS) where information on both tertiary and secondary structure can potentially be extracted (Bardhan et al., 2009; Park et al., 2009). These approaches also include alternative methods for the approximation of the hydration layer, including the addition of an explicit layer of water molecules in a solvent density matching that of the bulk solution (Yang et al., 2009).

5. Rigid body modeling

The assembly of macromolecular complexes can be studied through the docking of individual components into *ab initio* shapes (Wriggers and Chacón, 2001). However, as the resolution of SAXS derived shapes is low, it is more reliable to model the assembly of such complexes through direct refinement against the scattering data. A number of interactive and automated approaches have been developed using SAXS to determine the positions and orientations of subunits within macromolecular complexes (Boehm et al., 1999; Konarev et al., 2001; Petoukhov and Svergun, 2005; Sun et al., 2004). From the known structures (or homology models) of subunits, the theoretical scattering of a complex can be rapidly calculated using an application of the spherical harmonics formalism mentioned in the Section 3. This formalism, the basis for the above described programs CRYSOL (Svergun et al., 1995) and CRYSON (Svergun et al., 1998) is applied to automated rigid body modeling in the programs SASREF and BUNCH (Konarev et al., 2006; Petoukhov and Svergun, 2005).

SASREF is a comprehensive automated rigid body modeling program for SAS, allowing for the simultaneous fitting of multiple scattering curves (e.g. when multiple constructs such as deletion mutants have been measured and also for contrast variation data from SANS). The use of symmetry, orientational constraints (e.g. from residual dipolar couplings measured by NMR), inter-residue
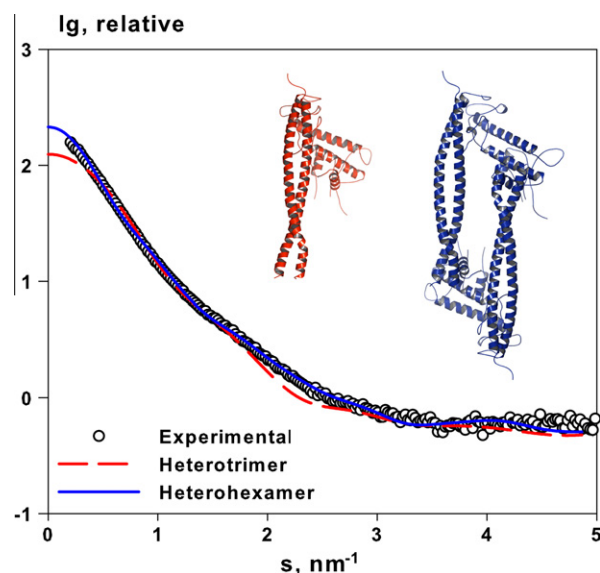

Fig. 5. Identification of heterohexameric solution state of the human Cdt1–Geminin complex by SAXS (De Marco et al., 2009). Experimental scattering pattern for the Cdt1–Geminin complex in solution and the fits calculated from the crystal structures of the heterotrimer (red broken line) and heterohexamer (blue broken line). It is clear that the heterohexamer (blue structure) fits the experimental data while the heterotrimer (red structure) does not.

contacts (e.g. from mutagenesis or cross-linking experiments) and inter-subunit distances (e.g. from FTIR and FRET) are fully supported. Starting from an arbitrary positioning of subunits, subject to any user defined constraints SASREF conducts a series of random rigid body movements and rotations, using SA to search for a best fit of the computed complex scattering to the experimental data. The target to be minimized has the form of Eq. (6) combining the discrepancy (possibly calculated over multiple scattering patterns) and penalty term. The latter may include restraints from other methods such as those mentioned above, but always incorporates constraints making sure that the models generated are interconnected and have no main-chain/backbone steric clashes (side-chains of proteins are ignored). SASREF is actively used by many groups; a recent example is the determination of the quaternary structure of the *Drosophila* neuronal adhesion protein Amalgam (Zeev-Ben-Mordehai et al., 2009). From this study a V-shaped dimer with a parallel arrangement of monomers was identified, and along with the identification of the dimerization interface provided a structural mechanism for the observed adhesion of this protein to neuronal cells. In another example, the low-resolution structure of the N-terminal region of the enzyme poly(ADP-ribose) polymerase-1 (PARP-1) was determined by SAXS (Lilyestrom et al., 2010). In this study, both *ab initio* modeling using DAMMIN and rigid body modeling with SASREF were used to identify the extended modular s-shaped structure of the N-terminal region of PARP-1. This model, combined with DNA binding studies were used to propose that the mechanism for activation of PARP-1 involves a conformational rearrangement upon the binding of damaged DNA and not dimerization.

SASREF requires that complete high-resolution models (or reliable homology models) of all of the subunits are available for the rigid body modeling of a protein complex. When the structures of linkers or entire domains are unknown an alternative approach combining both rigid body modeling and *ab initio* methods is used. The program BUNCH (Konarev et al., 2006; Petoukhov and Svergun, 2005) uses DRs to model missing regions in both protein complexes and multi-domain proteins connected by flexible linkers. As in SASREF a SA minimization is used to locate a best fitting arrangement of the rigid bodies and also the optimal local conformation of DRs. This approach has been used to successfully determine the structures of multi-domain proteins with flexible linkers (Clantin et al., 2010; Gut et al., 2009; Kozlov et al., 2009; Nemeth-Pongracz et al., 2007; Schmidt et al., 2010), and also for the addition of missing portions to crystal structures (Gut et al., 2009).

Rigid body modeling using SAXS data is very actively employed by structural biologists for the analysis of macromolecular complexes. For example, the combined *ab initio* and rigid body modeling approach in BUNCH was successfully used to determine the compact architecture of the central portion of the human complement factor H (fH) (Schmidt et al., 2010). fH is a modular multi-domain protein composed of 20 complement control protein modules (CCPs; each ~60 residues) connected by short linkers and is a major component of complement regulation in the human innate immune system. High-resolution structures of 12 terminal fH CCPs have been previously determined in isolation or as two to four domain fragments (reviewed in (Schmidt et al., 2008)), but the structure of the full length protein is unknown. Previous studies by transmission electron microscopy, analytical ultra centrifugation (AUC) and SAXS indicate that full length fH is not fully extended in solution and may contain highly flexible linkers, allowing the molecule to fold back upon itself (Aslam and Perkins, 2001). SAXS data was collected on several deletion mutants consisting of the central CCPs 10–15, and the NMR structure of CCPs 12–13 (fH12–13) determined in the same study (Schmidt et al., 2010). The constructs fH12–13, fH11–14 and fH10–15 were determined to be monomeric in solution and rigid body modeling with BUNCH

applied to the larger two constructs. The NMR ensemble of conformers determined for fH12–13 was independently validated through a good fit to the SAXS data and the near-perfect superposition of the *ab initio* structure with the ensemble (Fig. 6A and B). The rigid body models of fH11–14 and fH10–15 were highly reproducible, forming zig-zag arrangements, and demonstrated that the core of fH is compact (Fig. 6A and B).

6. Flexible systems

In the last example of the preceding section the scattering data from a multi-domain protein were successfully analyzed in terms of rigid body models, assuming therefore that all linkers were rigid such that constructs displayed no flexibility in solution. It was indeed the case for this particular protein (and it was possible to demonstrate this, see below), but very often in practice one deals with systems, which possess significant flexibility in solution.

SAXS was proven to be a powerful technique for the analysis of flexible systems (Bernadó et al., 2007, 2005; von Ossowski et al., 2005). However, a recent study by Bernadó (2009) clearly presents the difficulties associated with meaningful interpretation of SAXS curves for highly flexible modular proteins. This study demonstrates that proteins may be wrongly identified as rigid from dynamically averaged SAXS profiles and that several indicators for inter-domain flexibility should be monitored. Typically, *ab initio* models produced from dynamically averaged scattering data display a decrease in resolution or structural detail, while rigid body models are generated with highly extended conformations and a paucity of inter-domain contacts. A recently developed ensemble optimisation method (EOM) (Bernadó et al., 2007) provides a useful approach for assessing the flexibility of the system under study when high-resolution structures of domains are available.

Both intrinsically unfolded proteins and modular multi-domain proteins with flexible linkers can be represented as ensembles of structures. In EOM a large pool of random configurations is generated and ensembles are selected from this pool using a genetic algorithm, such that the average computed scattering over the ensemble fits the experimental scattering data (Bernadó et al., 2007). If the R_g distribution of the models in the selected ensembles is as broad as that in the initial random pool, the protein is likely to be flexible; obtaining a narrow R_g peak suggests that the system may be rigid. The EOM results provide a useful guidance but to reach definitive conclusions regarding the flexibility of the system these data should be correlated with complementary techniques including e.g. NMR relaxation studies. In fact, the above data from the human complement factor H was analyzed using EOM to suggest the relatively rigid nature of the constructs studied (Fig. 6C), supporting the hypothesis that the central domains of this protein do not provide a flexible tether between N and C-terminal ligand binding sites in agreement with the NMR data.

A recent application of the ensemble representation of flexible structures is the combined NMR and SAXS study of inter-domain flexibility in full-length matrix metalloproteinase-1 (MMP-1) (Bertini et al., 2009). In this study it was shown that a dynamic equilibrium of open and closed conformations of MMP-1 are present in solution, with the EOM analysis supporting the NMR data. The R_g distribution of the generated random pool of conformations is broad, covering a range of compact and extended structures (~20 to 45 Å). However, the R_g distribution of the selected ensemble of structures displays a relatively sharp peak at ~25 Å (Fig. 7), suggesting that the most highly populated conformations of MMP-1 are compact. An extended tail is observed in the distribution at higher values of R_g, suggesting that the MMP-1 is inherently flexible in solution and that a population of extended conformations
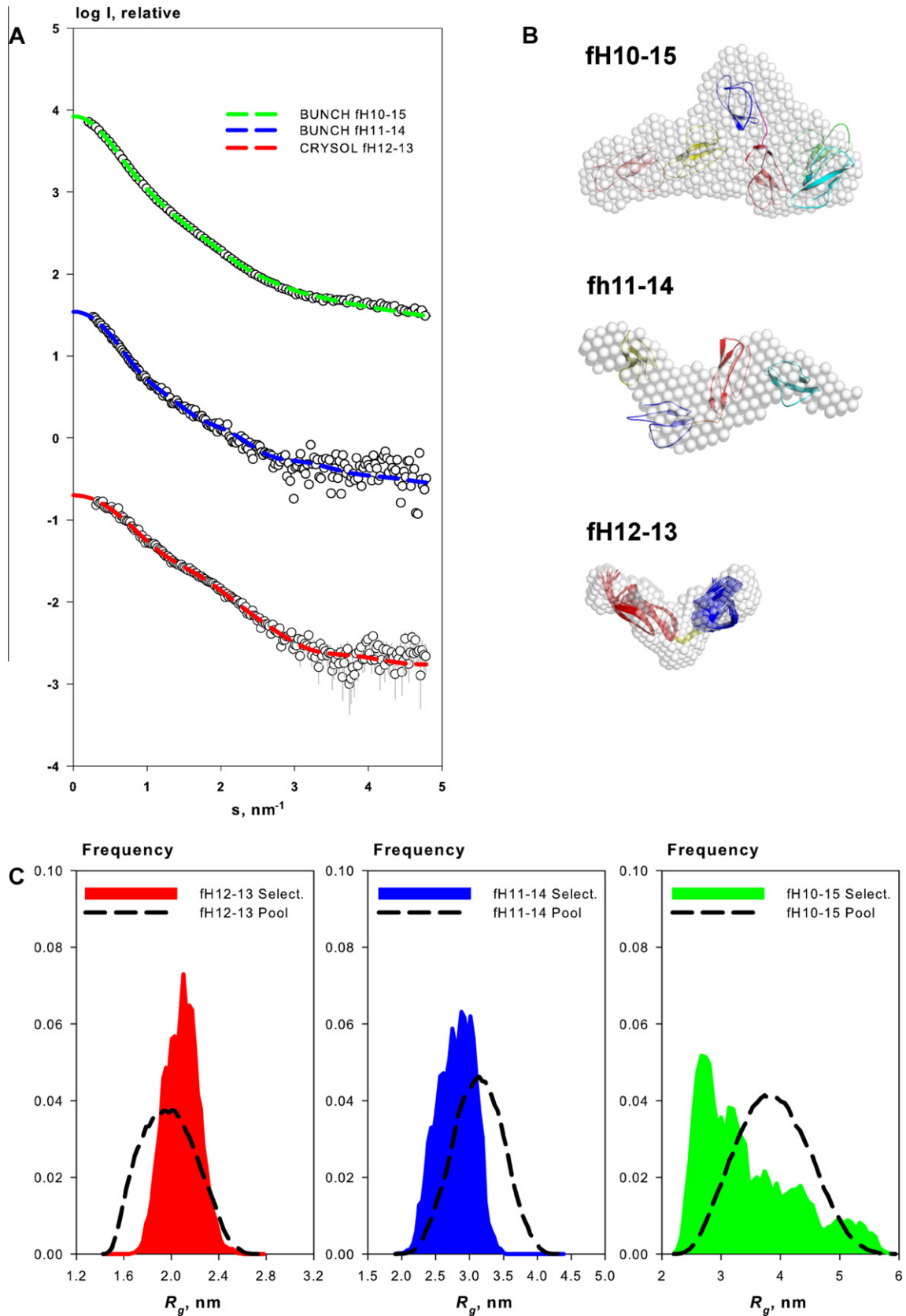
Fig. 6. SAXS study of human factor-H (Schmidt et al., 2010). (A) Scattering curves for three constructs of the central 2–5 CCP modules of factor-H. Broken lines represent fits obtained by CRYSOL for the best fH12–13 NMR model, or by rigid body modeling (BUNCH) for fH11–14 and fH10–15; curves have been arbitrarily displaced along the logarithmic axis for clarity. (B) Overlay of the NMR ensemble of fH12–13, and the rigid body models (BUNCH) for fH11–14 and fH10–15, with *ab initio* shape envelopes produced with DAMMIF. (C) Radius of gyration distributions of pools (broken lines) and selected structures (coloured areas) for the EOM analysis of fH12–13, fH11–14 and fH10–15. Each distribution is skewed toward either extended (fH12–13) or compact (fH10–14 and fH10–15) structures.
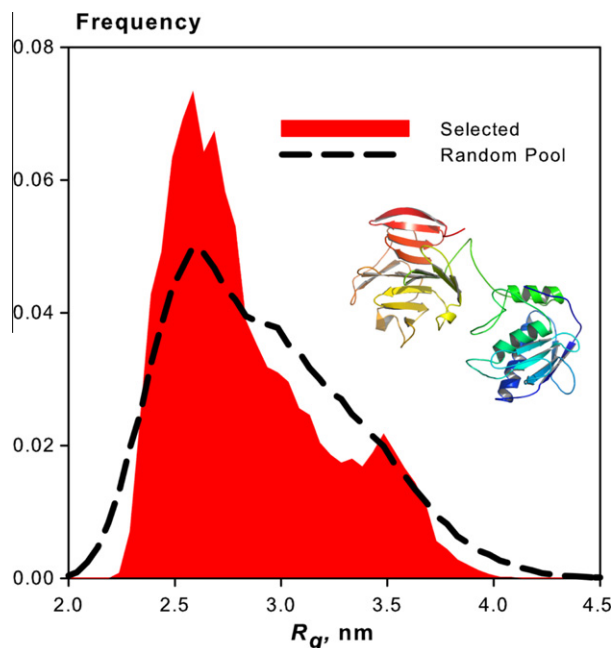
Fig. 7. EOM analysis of MMP-1 (Bertini et al., 2009). Radius of gyration distribution of the pool (broken line) and selected (coloured area) structures for the EOM analysis of MMP-1. The crystal structure of MMP-1 is shown inset.

also exist. In another example, tetrameric Flavorubredoxin (FDP), the rigid body models generated using SASREF and defined P222 symmetry suggested that a highly flexible linker between the core structure and c-terminal rubredoxin (Rd) domains was likely (Petoukhov et al., 2008). To investigate the degree of flexibility and to quantitatively describe the preferential arrangement of the Rd domains an EOM analysis was conducted. Indeed, through the addition to the random pool of conformations generated with constraints maintaining extended linkers, an excellent fit to the data was obtained. In all selected models no contacts between the FDP core and Rd domains were observed and all linkers were significantly extended and peripherally located.

As a word of caution it should be mentioned that EOM analysis of polydisperse samples containing mixtures of oligomers or aggregates may provide an erroneous indication of sample flexibility and lead to false conclusions. Similar to the *ab initio* and rigid body modeling methods, careful sample characterization prior to using this method is essential (see Section 2).

7. Analysis of mixtures

Another important application of SAXS to rapidly characterize protein solutions is the quantitative description of mixtures (e.g. oligomeric equilibria and assembly processes). For mixtures and polydisperse solutions of non-interacting particles the resulting scattering pattern is a sum of the contributions from each component of the mixture $I_k(s)$, weighted by the volume fraction v_k of that component:

$$I(s) = \sum_{k=1}^{K} v_k I_k(s),\qquad(10)$$

Several methods have been developed to help simplify the analysis of equilibrium mixtures using SAXS (Feigin and Svergun, 1987; Fowler et al., 1983; Koenig et al., 1992; Konarev et al., 2003). If the number of components K in the mixture is not known but a series of measurements is available from the samples containing different amounts of the components (e.g. at different stages of an assembly process), a model-independent estimate of K can be obtained using singular value decomposition (SVD) (Golub and Reinsh, 1970; Konarev et al., 2006). In the program OLIGOMER (Konarev et al., 2003), volume fractions are readily computed provided the scattering patterns of the components are known (e.g. provided by CRYSOL from known structures of components). OLIGOMER has been successfully used to characterize oligomeric equilibria and complex formation (Bernadó et al., 2009; Niemann et al., 2008; Paravisi et al., 2009), and has been applied to the study of self fibrillating proteins (Vestergaard et al., 2007). Recently, a multivariate curve resolution (MCR-ALS) method (Blobel et al., 2009) was proposed to determine scattering patterns from components in oligomeric mixtures.

An exciting development in the analysis of mixtures from scattering data is the determination of low-resolution *ab initio* models of protein–protein complexes that exist as minority species in solution. In a recent study reliable models of a major and minor component from a monomer–dimer equilibrium could be reconstructed from the deconvolution of SAXS data (Blobel et al., 2009). In this work, Blobel et al. analyzed a solution of low molecular weight phosphatase (lmwPTP) using a multivariate curve resolution method (MCR-ALS) to characterize the monomer–dimer equilibrium. From the extracted SAXS contributions corresponding to the monomeric and dimeric components (the dimer component constituting only ~15% of the total protein concentration), DAMMIN models were generated and were in very good agreement with the corresponding monomer and dimer crystal structures (Fig. 8).

Another powerful method for the analysis of mixtures is time-resolved SAXS (TR-SAXS). TR-SAXS has been used for the study of protein and RNA folding (Cammarata et al., 2008; Lamb et al., 2008a), the formation and dissociation of complexes and for kinetic analyses of conformational change stimulated by some external stimulus (Cammarata et al., 2008; Lamb et al., 2008b). While manual mixing and a series of static measurements can be conducted for slow (minutes to hours) molecular processes, studies of sub-millisecond to millisecond processes (for example protein folding under near native conditions of pH and temperature) require devices for rapid mixing and a high-brilliance X-ray beam coupled with a state-of-the-art detector. Third-generation synchrotrons provide the necessary flux for recording SAXS data with good signal to noise from very short (millisecond) exposures, and modern fast read-out detectors are now in use or a being installed
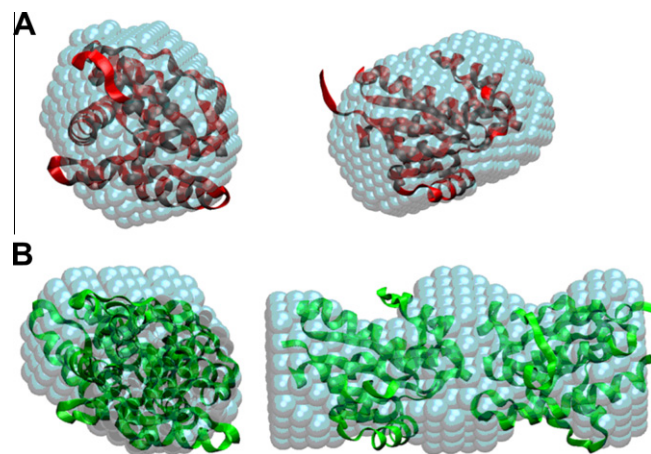

Fig. 8. Superposition of the crystal structures of lmwPTP monomer (A) and dimer (B) with the *ab initio* DAMMIN models generated from analysis of the oligomeric equilibrium (Blobel et al., 2009). Note that the *ab initio* model of the dimer is affected by boundary effects caused by the limited search volume. Such artifacts are absent in DAMMIF, where an adaptable search volume can be used (see Section 3).

at many SAXS beamlines, making time-resolved scattering studies more available to researchers. The recent time-resolved wide-angle scattering (WAXS) study of the conformational changes of haemoglobin highlights the power of the technique (Cammarata et al., 2008). In this study, nanosecond time resolution was made possible through the laser-induced photolysis of the carbon monoxide ligand, allowing the transition of the protein from the *R* to the *T* state to be followed in solution. TR-SAXS/WAXS has a bright future in modern structural biology as the study of kinetic processes can help to link structure with biological function, especially for complex systems of interacting molecules.

8. Combining NMR, crystallography and SAXS

The combination of X-ray crystallography and SAXS as complementary methods is very well established. The use of SAXS to investigate the solution properties of crystal structures was pioneered in the 1970s and early 1980s, with the development of sophisticated methods for the prediction of theoretical scattering from crystal structures and initial attempts at rigid body modeling (Fedorov and Denisyuk, 1978; McDonald et al., 1979; Pavlov, 1985). These methods have been actively developed and now provide powerful tools used by structural biologists to compliment crystallographic studies. Beamlines dedicated to biological SAXS have been constructed and help to consolidate the complementarity of SAXS and crystallography (Putnam et al., 2007). More recently, the NMR community has embraced SAXS as a technique that is not only complementary to high resolution solution structure analysis but that can be incorporated directly in structure determination (Gabel et al., 2008; Grishaev et al., 2005).

In addition to model validation (discussed in Section 4), SAXS can be used in crystallography as a tool for molecular replacement. The program *FSEARCH* (Hao, 2006) uses *ab initio* shape envelopes and bead or dummy residue models from EM or SAXS for the determination of low-resolution phases. Following correct positioning of the molecular envelope or bead/dummy model within the unit cell the low-resolution phases are extended to crystallographic resolution. This promising method has been used successfully for several proteins (Kollman and Quispe, 2005; Liu et al., 2003) and continues to be developed (Hong and Hao, 2009).

A major obstacle in the determination of large (>30 kDa) proteins and complexes by NMR spectroscopy is the increased difficulty of extracting useful structural information as the size of the system to be studied increases. As resonance overlap becomes a significant problem with the increased size of macromolecules, unambiguous sequence specific assignment of both the backbone and side-chains becomes difficult. This, combined with the disappearance of peaks due to relaxation processes leads to a reduction in the number of distance constraints typically used for structure calculation. Orientation constraints derived from residual dipolar couplings help to overcome the size limitations by providing long-range information on the relative orientation of distant parts of the structure (Bax et al., 2001; Tjandra et al., 1997; Tolman et al., 1995). However, they do not provide translational information and by themselves cannot be used to generate an unambiguous solution. SAXS data, providing information on the global shape of the macromolecule can be introduced into the structure calculation and reduce this ambiguity. The inclusion of potentials for the refinement of NMR structures against SAXS data have been introduced into several popular structure calculation packages, for example the programs Xplor-NIH (Schwieters et al., 2003) and CNS (Brunger, 2007; Brunger et al., 1999; Gabel et al., 2008; Grishaev et al., 2005), and have been shown to significantly improve the accuracy of calculated structures (Grishaev et al., 2005, 2008a). These methods have also been shown to be applicable to the determination of DNA and RNA structures by NMR (Grishaev et al., 2008b; Schwieters and Clore, 2007).

SAXS data can be used to complement many studies in structural biology including validation of high-resolution models, solving the phase problem in crystallography and the refinement of solution structures. New approaches to structural modeling are currently being developed which seek to incorporate as much complementary data as possible. The integrated modeling platform (IMP) is one such project using SAXS data as an additional spatial restraint (Forster et al., 2008).

9. Future developments

The study of biological systems using solution SAXS is increasingly gaining momentum, with many research groups looking to incorporate this technique into their research programs. As most of the SAXS analysis tools have now reached a mature state, their application is straightforward and can even be performed automatically. Therefore, not only evaluation of the overall parameters, but also shape determination, analysis of the oligomeric composition and to some extent rigid body modeling of quaternary structure can be considered rapid characterization tools for proteins and complexes (Fig. 9). Having said that, one should keep in mind that X-rays provide a more powerful and demanding tool than standard biophysical equipment. Therefore synchrotrons should not be employed as analytical tools to characterize unknown, perhaps poorly behaving samples, and preliminary characterization by light scattering, gel-filtration or analytical ultracentrifugation is recommended.

The many novel and exciting biological questions brought by new users of the technique require that laboratories specializing in SAXS continue to push the boundaries in methods development. A number of research groups from around the world are actively involved in the development of advanced computational methods for the characterization of biomolecules using SAXS. The tools made available by these groups have had a significant impact on the field of structural biology, with automation of data collection, data reduction and analysis in particular making SAXS more accessible to the non-expert (Petoukhov et al., 2007; Round et al., 2008). The availability of beamlines for biological SAXS has also improved over the last few years, with the high-brilliance beamlines BL45XU (Spring-8, Japan), ID14–3 (ESRF, France), and SAXS/WAXS (Australian Synchrotron, Australia) now in operation, with BioSAXS (PETRA III, Hamburg) expected to be operational in 2011. The beamlines X33 (DORIS, Hamburg) (Roessle et al., 2007; Round et al., 2008), SYBILS (ALS, Berkley) (Hura et al., 2009), BL4–2 (SSRL, USA) and SWING (Soleil, Orsay) (David and Perez, 2009) offer automated sample changers, the latter one also combined with on-line HPLC purification and a UV–vis absorption monitoring system. Other complementary techniques are also employed on-line, e.g. resonant Raman spectroscopy at the ESRF ID-13 beamline (Davies et al., 2009).

It should be mentioned that for many biological SAXS applications the use of modern laboratory instruments (for example those built by Bruker, Rigaku, Anton Paar and Hecus) permit one to collect data of sufficient quality. The laboratory SAXS experiment takes hours instead of seconds or minutes at a synchrotron, but still provides data for computation of overall parameters, *ab initio* shapes and rigid body analysis leading in some cases to exciting results (Cramer et al., 2010; Hamley et al., 2010).

The programs mentioned in this review belonging to the ATSAS suite (Konarev et al., 2006; Petoukhov et al., 2007) are publicly available for download by academic use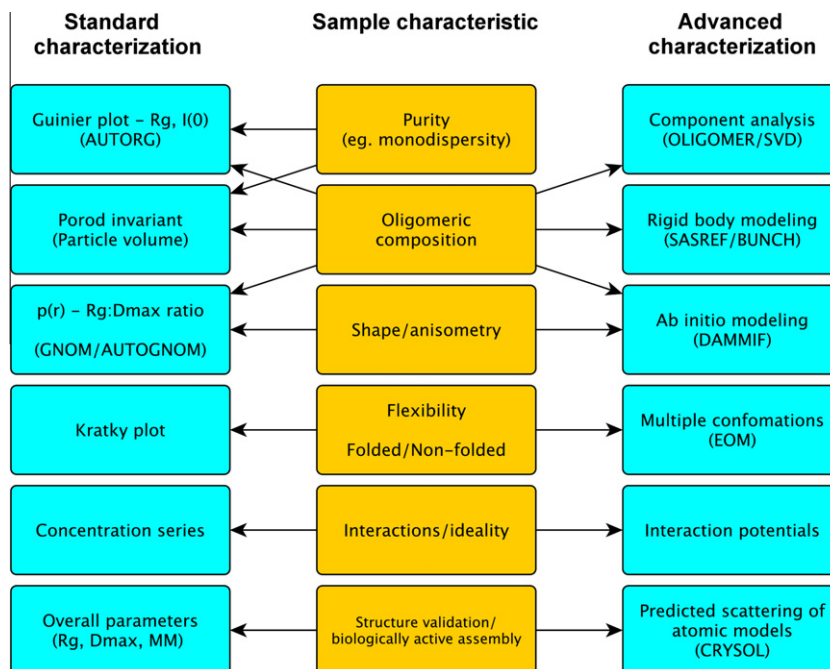rs and for on-line access from the EMBL-Hamburg web-site: http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html/.

Fig. 9. A summary of the standard and more advanced tools for characterization of protein samples using SAXS. For detailed descriptions of the tools refer to the text. Several suggested programs from the ATSAS package are indicated in capital letters.

In the present review we largely mentioned the SAXS applications, which emerged from the collaborative projects at the EMBL X33 beamline, but many extremely interesting SAXS applications are now being published by numerous groups worldwide (Fetler et al., 2007; Hoiberg-Nielsen et al., 2009; West et al., 2008; Whitten et al., 2009). Especially powerful is the use of SAXS together with other structural and biochemical techniques, in a streamlined approach to characterize the structure and dynamical properties of proteins and protein complexes in solution.

Acknowledgments

References

Aslam, M., Perkins, S.J., 2001. Folded-back solution structure of monomeric factor H of human complement by synchrotron X-ray and neutron scattering, analytical ultracentrifugation and constrained molecular modelling. J. Mol. Biol. 309, 1117–1138.

Bada, M., Walther, D., Arcangioli, B., Doniach, S., Delarue, M., 2000. Solution structural studies and low-resolution model of the *Schizosaccharomyces pombe* sap1 protein. J. Mol. Biol. 300, 563–574.

Bardhan, J., Park, S., Makowski, L., 2009. SoftWAXS: a computational tool for modeling wide-angle X-ray solution scattering from biomolecules. J. Appl. Cryst. 42, 932–943.

Bax, A., Kontaxis, G., Tjandra, N., 2001. Dipolar couplings in macromolecular structure determination. Methods Enzymol. 339, 127–174.

Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W., Blackledge, M., 2005. A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. Proc. Natl. Acad. Sci. USA 102, 17002–17007.

Bernadó, P., 2009. Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. Eur. Biophys. J. Oct 21. [Epub ahead of print].

Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., Svergun, D.I., 2007. Structural characterization of flexible proteins using small-angle X-ray scattering. J. Am. Chem. Soc. 129, 5656–5664.

Bernadó, P., Pérez, Y., Blobel, J., Fernández-Recio, J., Svergun, D.I., Pons, M., 2009. Structural characterization of unphosphorylated STAT5a oligomerization equilibrium in solution by small-angle X-ray scattering. Protein Sci. 18, 716–726.

Bertini, I., Fragai, M., Luchinat, C., Melikian, M., Mylonas, E., Sarti, N., Svergun, D.I., 2009. Interdomain flexibility in full-length matrix metalloproteinase-1 (MMP-1). J. Biol. Chem. 284, 12821–12828.

Blobel, J., Bernadó, P., Svergun, D.I., Tauler, R., Pons, M., 2009. Low-resolution structures of transient protein–protein complexes using small-angle X-ray scattering. J. Am. Chem. Soc. 131, 4378–4386.

Boehm, M.K., Woof, J.M., Kerr, M.A., Perkins, S.J., 1999. The fab and fc fragments of IgA1 exhibit a different arrangement from that in IgG: a study by X-ray and neutron solution scattering and homology modelling. J. Mol. Biol. 286, 1421–1447.

Bonnete, F., Finet, S., Tardieu, A., 1999. Second virial coefficient: variations with lysozyme crystallization conditions. J. Cryst. Growth 196, 403–414.

Brunger, A.T., 2007. Version 1.2 of the crystallography and NMR system. Nat. Protoc. 2, 2728–2733.

Brunger, A.T., Adams, P.D., Rice, L.M., 1999. Annealing in crystallography: a powerful optimization tool. Prog. Biophys. Mol. Biol. 72, 135–155.

Cammarata, M., Levantino, M., Schotte, F., Anfinrud, P.A., Ewald, F., Choi, J., Cupane, A., Wulff, M., Ihee, H., 2008. Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. Nat. Methods 5, 881–886.

Chacon, P., Diaz, J.F., Moran, F., Andreu, J.M., 2000. Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. J. Mol. Biol. 299, 1289–1302.

Chacon, P., Moran, F., Diaz, J.F., Pantos, E., Andreu, J.M., 1998. Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. Biophys. J. 74, 2760–2775.

Chen, B., Doucleff, M., Wemmer, D.E., De Carlo, S., Huang, H.H., Nogales, E., Hoover, T.R., Kondrashkina, E., Guo, L., Nixon, B.T., 2007. ATP ground- and transition states of bacterial enhancer binding AAA+ ATPases support complex formation with their target protein, [sigma]54. Structure 15, 429–440.

Cheng, C.Y., Yang, J., Taylor, S.S., Blumenthal, D.K., 2009. Sensing domain dynamics in pk-alpha complexes by solution X-ray scattering. J. Biol. Chem. 284, 35916–35925.

Clantin, B., Leyrat, C., Wohlkonig, A., Hodak, H., Ribeiro Ede Jr., A., Martinez, N., Baud, C., Smet-Nocca, C., Villeret, V., Jacob-Dubuisson, F., Jamin, M., 2010. Structure and plasticity of the peptidyl–prolyl isomerase Par27 of Bordetella pertussis revealed by X-ray diffraction and small-angle X-ray scattering. J. Struct. Biol. 169, 253–265.

Cramer, J.F., Gustafsen, C., Behrens, M.A., Oliveira, C.L.P., Pedersen, J.S., Madsen, P., Petersen, C.M., Thirup, S.S., 2010. GGA autoinhibition revisited. Traffic 11, 259–273.

David, G., Perez, J., 2009. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray

scattering experiments at the synchrotron SOLEIL SWING beamline. J. Appl. Cryst. 42, 892–900.

Davies, R.J., Burghammer, M., Riekel, C., 2009. A combined microRaman and microdiffraction set-up at the European Synchrotron Radiation Facility ID13 beamline. J. Synchrotron Radiat. 16, 22–29.

De Marco, V., Gillespie, P.J., Li, A., Karantzelis, N., Christodoulou, E., Klompmaker, R., van Gerwen, S., Fish, A., Petoukhov, M.V., Iliou, M.S., Lygerou, Z., Medema, R.H., Blow, J.J., Svergun, D.I., Taraviras, S., Perrakis, A., 2009. Quaternary structure of the human Cdt1–Geminin complex regulates DNA replication licensing. Proc. Natl. Acad. Sci. USA 106, 19807–19812.

Debye, P., 1915. Zerstreuung von Roentgenstrahlen. Ann. Physik. 46, 809–823.

Doniach, S., 2001. Changes in biomolecular conformation seen by small angle X-ray scattering. Chem. Rev. 101, 1763–1778.

Doniach, S., Lipfert, J., 2009. Use of small angle X-ray scattering (Saxs) to characterize conformational states of functional RNAs. Methods Enzymol. Biophys. Chem. Funct. Probes RNA Struct. Interact. Folding 469 (Pt. B), 237–251.

Fedorov, B.A., Denisyuk, A.I., 1978. Large-angle X-ray diffuse scattering, a new method for investigating changes in the conformation of globular proteins in solutions. J. Appl. Cryst. 11, 473–477.

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

Fetler, L., Kantrowitz, E.R., Vachette, P., 2007. Direct observation in solution of a preexisting structural equilibrium for a mutant of the allosteric aspartate transcarbamoylase. Proc. Natl. Acad. Sci. USA 104, 495–500.

Finet, S., Skouri-Panet, F., Casselyn, M., Bonnete, F., Tardieu, A., 2004. The Hofmeister effect as seen by SAXS in protein solutions. Curr. Opin. Colloid Interface Sci. 9, 112–116.

Fischer, H., Neto, M.D., Napolitano, H.B., Polikarpov, I., Craievich, A.F., 2010. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. J. Appl. Cryst. 43, 101–109.

Forster, F., Webb, B., Krukenberg, K.A., Tsuruta, H., Agard, D.A., Sali, A., 2008. Integration of small-angle X-ray scattering data into structural modeling of proteins and their assemblies. J. Mol. Biol. 382, 1089–1106.

Fowler, A.G., Foote, A.M., Moody, M.F., Vachette, P., Provencher, S.W., Gabriel, A., Bordas, J., Koch, M.H., 1983. Stopped-flow solution scattering using synchrotron radiation: apparatus, data collection and data analysis. J. Biochem. Biophys. Methods 7, 317–329.

Franke, D., Svergun, D.I., 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Cryst. 42, 342–346.

Fraser, R.D.B., MacRae, T.P., Suzuki, E., 1978. An improved method for calculating the contribution of solvent to the X-ray diffraction pattern of biological molecules. J. Appl. Cryst. 11, 693–694.

Gabel, F., Simon, B., Nilges, M., Petoukhov, M., Svergun, D., Sattler, M., 2008. A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. J. Biomol. NMR 41, 199–208.

Glatter, O., 1977. A new method for the evaluation of small-angle scattering data. J. Appl. Cryst. 10, 415–421.

Golub, G.H., Reinsh, C., 1970. Singular value decomposition and least squares solution. Numer. Math. 14, 403–420.

Grishaev, A., Wu, J., Trewhella, J., Bax, A., 2005. Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. J. Am. Chem. Soc. 127, 16621–16628.

Grishaev, A., Tugarinov, V., Kay, L.E., Trewhella, J., Bax, A., 2008a. Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. J. Biomol. NMR 40, 95–106.

Grishaev, A., Ying, J., Canny, M.D., Pardi, A., Bax, A., 2008b. Solution structure of tRNA(Val) from refinement of homology model against residual dipolar coupling and SAXS data. J. Biomol. NMR 42, 99–109.

Guinier, A., 1939. La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. Ann. Phys. (Paris) 12, 161–237.

Guo, D.Y., Smith, G.D., Griffin, J.F., Langs, D.A., 1995. Use of globic scattering factors for protein structures at low resolution. Acta Cryst. A51, 945–947.

Guo, D.Y., Blessing, R.H., Langs, D.A., Smith, G.D., 1999. On 'globbicity' of low-resolution protein structures. Acta Cryst. D55, 230–237.

Gut, H., Dominici, P., Pilati, S., Astegno, A., Petoukhov, M.V., Svergun, D.I., Grütter, M.G., Capitani, G., 2009. A common structural basis for pH- and calmodulin-mediated regulation in plant glutamate decarboxylase. J. Mol. Biol. 392, 334–351.

Hamiaux, C., Perez, J., Prange, T., Veesler, S., Ries-Kautt, M., Vachette, P., 2000. The BPTI decamer observed in acidic pH crystal forms pre-exists as a stable species in solution. J. Mol. Biol. 297, 697–712.

Hamley, I.W., Castelletto, V., Moulton, C., Myatt, D., Siligardi, G., Oliveira, C.L.P., Pedersen, J.S., Abutbul, I., Danino, D., 2010. Self-assembly of a modified amyloid peptide fragment: pH-responsiveness and nematic phase formation. Macromol. Biosci. 10, 40–48.

Hao, Q., 2006. Macromolecular envelope determination and envelope-based phasing. Acta Crystallogr. D Biol. Crystallogr. 62, 909–914.

Harker, D., 1953. The meaning of the average of |F|2 for large values of the interplanar spacing. Acta Crystallogr. 6, 731–736.

Heikkinen, O.K., Ruskamo, S., Konarev, P.V., Svergun, D.I., Iivanainen, T., Heikkinen, S.M., Permi, P., Koskela, H., Kilpelaeinen, I., Ylaenne, J., 2009. Atomic structures of two novel immunoglobulin-like domain pairs in the actin cross-linking protein filamin. J. Biol. Chem. 284, 25450–25458.

Hoiberg-Nielsen, R., Westh, P., Skov, L.K., Arleth, L., 2009. Interrelationship of steric stabilization and self-crowding of a glycosylated protein. Biophys. J. 97, 1445–1453.

Hong, X.G., Hao, Q., 2009. Combining solution wide-angle X-ray scattering and crystallography: determination of molecular envelope and heavy-atom sites. J. Appl. Cryst. 42, 259–264.

Hough, M.A., Grossmann, J.G., Antonyuk, S.V., Strange, R.W., Doucette, P.A., Rodriguez, J.A., Whitson, L.J., Hart, P.J., Hayward, L.J., Valentine, J.S., Hasnain, S.S., 2004. Dimer destabilization in superoxide dismutase may result in disease-causing properties: structures of motor neuron disease mutants. Proc. Natl. Acad. Sci. USA 101, 5976–5981.

Hura, G.L., Menon, A.L., Hammel, M., Rambo, R.P., Poole, F.L., Tsutakawa, S.E., Jenney, F.E., Classen, S., Frankel, K.A., Hopkins, R.C., Yang, S.J., Scott, J.W., Dillard, B.D., Adams, M.W.W., Tainer, J.A., 2009. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods 6, 606–612.

Ibel, K., Stuhrmann, H.B., 1975. Comparison of neutron and X-ray scattering of dilute myoglobin solutions. J. Mol. Biol. 93, 255–265.

King, W.A., Stone, D.B., Timmins, P.A., Narayanan, T., von Brasch, A.A.M., Mendelson, R.A., Curmi, P.M.G., 2005. Solution structure of the chicken skeletal muscle troponin complex via small-angle neutron and X-ray scattering. J. Mol. Biol. 345, 797–815.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680.

Koch, M.H., Vachette, P., Svergun, D.I., 2003. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. Q. Rev. Biophys. 36, 147–227.

Koenig, S., Svergun, D., Koch, M.H.J., Hubner, G., Schellenberger, A., 1992. Synchrotron radiation solution X-ray scattering study of the pH dependence of the quaternary structure of yeast pyruvate decarboxylase. Biochemistry (Moscow) 31, 8726–8731.

Kollman, J.M., Quispe, J., 2005. The 17 angstrom structure of the 420 kDa lobster clottable protein by single particle reconstruction from cryoelectron micrographs. J. Struct. Biol. 151, 306–314.

Konarev, P.V., Petoukhov, M.V., Svergun, D.I., 2001. MASSHA-a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. J. Appl. Cryst. 34, 527–532.

Konarev, P.V., Petoukhov, M.V., Volkov, V.V., Svergun, D.I., 2006. ATSAS 2.1, a program package for small-angle scattering data analysis. J. Appl. Cryst. 39, 277–286.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., Svergun, D.I., 2003. PRIMUS—a Windows-PC based system for small-angle scattering data analysis. J. Appl. Cryst. 36, 1277–1282.

Kozin, M.B., Svergun, D.I., 2001. Automated matching of high- and low-resolution structural models. J. Appl. Cryst. 34, 33–41.

Kozlov, G., Maattanen, P., Schrag, J.D., Hura, G.L., Gabrielli, L., Cygler, M., Thomas, D.Y., Gehring, K., 2009. Structure of the noncatalytic domains and global fold of the protein disulfide isomerase ERp72. Structure 17, 651–659.

Lamb, J., Kwok, L., Qiu, X.Y., Andresen, K., Park, H.Y., Pollack, L., 2008a. Reconstructing three-dimensional shape envelopes from time-resolved small-angle X-ray scattering data. J. Appl. Cryst. 41, 1046–1052.

Lamb, J.S., Zoltowski, B.D., Pabit, S.A., Crane, B.R., Pollack, L., 2008b. Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering. J. Am. Chem. Soc. 130, 12226–12227.

Lilyestrom, W., van der Woerd, M.J., Clark, N., Luger, K., 2010. Structural and biophysical studies of human PARP-1 in complex with damaged DNA. J. Mol. Biol. 395, 983–994.

Liu, Q., Weaver, A.J., Xiang, T., Thiel, D.J., Hao, Q., 2003. Low-resolution molecular replacement using a six-dimensional search. Acta Crystallogr. D Biol. Crystallogr. 59, 1016–1019.

McDonald, R.C., Engelman, D.M., Steitz, T.A., 1979. Small angle X-ray scattering of dimeric yeast hexokinase in solution. J. Biol. Chem. 254, 2942–2943.

Mylonas, E., Svergun, D.I., 2007. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. J. Appl. Cryst. 40, s245–s249.

Nemeth-Pongracz, V., Barabas, O., Fuxreiter, M., Simon, I., Pichova, I., Rumlova, M., Zabranska, H., Svergun, D., Petoukhov, M., Harmat, V., Klement, E., Hunyadi-Gulyas, E., Medzihradszky, K.F., Konya, E., Vertessy, B.G., 2007. Flexible segments modulate co-folding of dUTPase and nucleocapsid proteins. Nucleic Acids Res. 35, 495–505.

Niemann, H.H., Petoukhov, M.V., Härtlein, M., Moulin, M., Gherardi, E., Timmins, P., Heinz, D.W., Svergun, D.I., 2008. X-ray and neutron small-angle scattering analysis of the complex formed by the met receptor and the listeria monocytogenes invasion protein InlB. J. Mol. Biol. 377, 489–500.

Obolensky, O.I., Schlepckow, K., Schwalbe, H., Solov'yov, A.V., 2007. Theoretical framework for NMR residual dipolar couplings in unfolded proteins. J. Biomol. NMR 39, 1–16.

Orthaber, D., Bergmann, A., Glatter, O., 2000. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. J. Appl. Cryst. 33, 218–225.

Paravisi, S., Fumagalli, G., Riva, M., Morandi, P., Morosi, R., Konarev, P.V., Petoukhov, M.V., Bernier, S., Chênevert, R., Svergun, D.I., Curti, B., Vanoni, M.A., 2009. Kinetic and mechanistic characterization of Mycobacterium tuberculosis glutamyl-tRNA synthetase and determination of its oligomeric structure in solution. FEBS J. 276, 1398–1417.

Park, S., Bardhan, J.P., Roux, B., Makowski, L., 2009. Simulated X-ray scattering of protein solutions using explicit-solvent models. J. Chem. Phys. 130, 134114–134118.

Pavkov, T., Egelseer, E.M., Tesarz, M., Svergun, D.I., Sleytr, U.B., Keller, W., 2008. The structure and binding behavior of the bacterial cell surface layer protein SbsC. Structure 16, 1226–1237.

Pavlov, M., 1985. Determination of the relative position of the domains in 2-domain proteins based on diffuse X-ray scattering data. Dokl. Akad. Nauka SSSR 281, 458–462.

Petoukhov, M.V., Svergun, D.I., 2005. Global rigid body modelling of macromolecular complexes against small-angle scattering data. Biophys. J. 89, 1237–1250.

Petoukhov, M.V., Svergun, D.I., 2006. Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. Eur. Biophys. J. 35, 567–576.

Petoukhov, M.V., Konarev, P.V., Kikhney, A.G., Svergun, D.I., 2007. ATSAS 2.1—towards automated and web-supported small-angle scattering data analysis. J. Appl. Cryst. 40, s223–s228.

Petoukhov, M.V., Vicente, J.B., Crowley, P.B., Carrondo, M.A., Teixeira, M., Svergun, D.I., 2008. Quaternary structure of flavorubredoxin as revealed by synchrotron radiation small-angle X-ray scattering. Structure 16, 1428–1436.

Pollack, L., Doniach, S., 2009. Time-resolved X-ray scattering and RNA folding. Methods Enzymol. Biophys. Chem. Funct. Probes RNA Struct. Interact. Folding 469 (Pt. B), 253–268.

Porod, G., 1982. General theory. In: Glatter, O., Kratky, O. (Eds.), Small-angle X-ray Scattering. Academic Press, London, pp. 17–51.

Putnam, C.D., Hammel, M., Hura, G.L., Tainer, J.A., 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. 40, 191–285.

Rambo, R.P., Tainer, J.A., 2010. Improving small-angle X-ray scattering data for structural analyses of the RNA world. Rna-a Publication of the Rna Society 16, 638–646.

Roessle, M.W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, T., Konarev, P., Round, A., Fiedler, S., Hermes, C., Svergun, D., 2007. Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. J. Appl. Cryst. 40, s190–s194.

Round, A.R., Franke, D., Moritz, S., Huchler, R., Fritsche, M., Malthan, D., Klaering, R., Svergun, D.I., Roessle, M., 2008. Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. J. Appl. Cryst. 41, 913–917.

Santiago, J., Dupeux, F., Round, A., Antoni, R., Park, S.-Y., Jamin, M., Cutler, S.R., Rodriguez, P.L., Márquez, J.A., 2009. The abscisic acid receptor PYR1 in complex with abscisic acid. Nature 462, 665–668.

Schmidt, C.Q., Herbert, A.P., Hocking, H.G., Uhrin, D., Barlow, P.N., 2008. Translational mini-review series on complement factor H: structural and functional correlations for factor H. Clin. Exp. Immunol. 151, 14.

Schmidt, C.Q., Herbert, A.P., Mertens, H.D.T., Guariento, M., Soares, D.C., Uhrin, D., Rowe, A.J., Svergun, D.I., Barlow, P.N., 2010. The central portion of factor H (modules 10–15) is compact and contains a structurally deviant CCP module. J. Mol. Biol. 395, 105–122.

Schwieters, C.D., Clore, G.M., 2007. A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. Biochemistry (Moscow) 46, 1152–1166.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., Clore, G.M., 2003. The Xplor-NIH NMR molecular structure determination package. J. Magn. Reson. 160, 65–73.

Sokolova, A.V., Volkov, V.V., Svergun, D.I., 2003. Prototype of database for rapid protein classification based on solution scattering data. J. Appl. Cryst. 36, 865–868.

Stuhrmann, H.B., 1970a. New method for determination of surface form and internal structure of dissolved globular proteins from small-angle X-ray measurements. Z. Phys. Chem. 72, 177–182.

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

Sun, Z., Reid, K.B.M., Perkins, S.J., 2004. The dimeric and trimeric solution structures of the multidomain complement protein properdin by X-ray scattering, analytical ultracentrifugation and constrained modelling. J. Mol. Biol. 343, 1327–1343.

Svergun, D.I., 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. 25, 495–503.

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–2886.

Svergun, D.I., 2007. Small-angle scattering studies of macromolecular solutions. J. Appl. Cryst. 40, s10–s17.

Svergun, D.I., Nierhaus, K.H., 2000. A map of protein-rRNA distribution in the 70 S Escherichia coli ribosome. J. Biol. Chem. 275, 14432–14439.

Svergun, D.I., Koch, M.H.J., 2003. Small-angle scattering studies of biological macromolecules. Rep. Prog. Phys. 66, 1735–1782.

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

Svergun, D.I., Petoukhov, M.V., Koch, M.H.J., 2001. Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. 80, 2946–2953.

Svergun, D.I., Volkov, V.V., Kozin, M.B., Stuhrmann, H.B., 1996. New developments in direct shape determination from small-angle scattering. 2. Uniqueness. Acta Crystallogr. A 52, 419–426.

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

Svergun, D.I., Ekström, F., Vandegriff, K.D., Malavalli, A., Baker, D.A., Nilsson, C., Winslow, R.M., 2008. Solution structure of poly(ethylene) glycol-conjugated hemoglobin revealed by small-angle X-ray scattering: implications for a new oxygen therapeutic. Biophys. J. 94, 173–181.

Tardieu, A., Le Verge, A., Riès-Kautt, M., Malfois, M., Bonneté, F., Finet, S., Belloni, L., 1999. Proteins in solution: from X-ray scattering intensities to interaction potentials. J. Cryst. Growth 196, 193–203.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M., Bax, A., 1997. Use of dipolar 1H–15N and 1H–13C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat. Struct. Biol. 4, 732–738.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A., Prestegard, J.H., 1995. Nuclear magnetic dipole interactions in field-oriented proteins—information for structure determination in solution. Proc. Natl. Acad. Sci. USA 92, 9279–9283.

Trindade, D.M., Silva, J.C., Navarro, M.S., Torriani, I.C.L., Kobarg, J., 2009. Low-resolution structural studies of human Stanniocalcin-1. BMC Struct. Biol. 9, 57.

Tsuruta, H., Irving, T., 2008. Experimental approaches for solution X-ray scattering and fiber diffraction. Curr. Opin. Struct. Biol. 18, 601–608.

Vestergaard, B., Sanyal, S., Roessle, M., Mora, L., Buckingham, R.H., Kastrup, J.S., Gajhede, M., Svergun, D.I., Ehrenberg, M., 2005. The SAXS solution structure of RF1 differs from its crystal structure and is similar to its ribosome bound cryo-EM structure. Mol. Cell 20, 929–938.

Vestergaard, B., Groenning, M., Roessle, M., Kastrup, J.S., van de Weert, M., Flink, J.M., Frokjaer, S., Gajhede, M., Svergun, D.I., 2007. A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. PLoS Biol. 5, 1089–1097.

Vigil, D., Gallagher, S.C., Trewhella, J., Garcia, A.E., 2001. Functional dynamics of the hydrophobic cleft in the N-domain of calmodulin. Biophys. J. 80, 2082–2092.

Volkov, V.V., Svergun, D.I., 2003. Uniqueness of ab initio shape determination in small angle scattering. J. Appl. Cryst. 36, 860–864.

von Ossowski, I., Eaton, J.T., Czjzek, M., Perkins, S.J., Frandsen, T.P., Schuelein, M., Panine, P., Henrissat, B., Receveur-Bréchot, V., 2005. Protein disorder: conformational distribution of the flexible linker in a chimeric double cellulase. Biophys. J. 88, 2823–2832.

Wall, M.E., Gallagher, S.C., Trewhella, J., 2000. Large-scale shape changes in proteins and macromolecular complexes. Annu. Rev. Phys. Chem. 51, 355–380.

Walther, D., Cohen, F.E., Doniach, S., 2000. Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray scattering data for biomolecules. J. Appl. Cryst. 33, 350–363.

West, J.M., Xia, J.R., Tsuruta, H., Guo, W.Y., O'Day, E.M., Kantrowitz, E.R., 2008. Time evolution of the quaternary structure of Escherichia coli aspartate transcarbamoylase upon reaction with the natural substrates and a slow, tight-binding inhibitor. J. Mol. Biol. 384, 206–218.

Whitten, A.E., Trewhella, J., 2009. Small-angle scattering and neutron contrast variation for studying bio-molecular complexes. Methods Mol. Biol. 544, 307–323.

Whitten, A.E., Smith, B.J., Menting, J.G., Margetts, M.B., McKern, N.M., Lovrecz, G.O., Adams, T.E., Richards, K., Bentley, J.D., Trewhella, J., Ward, C.W., Lawrence, M.C., 2009. Solution structure of ectodomains of the insulin receptor family: the ectodomain of the type 1 insulin-like growth factor receptor displays asymmetry of ligand binding accompanied by limited conformational change. J. Mol. Biol. 394, 878–892.

Wriggers, W., Chacón, P., 2001. Using situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. J. Appl. Cryst. 34, 773–776.

Yadavalli, S.S., Klipcan, L., Zozulya, A., Banerjee, R., Svergun, D., Safro, M., Ibba, M., 2009. Large-scale movement of functional domains facilitates aminoacylation by human mitochondrial phenylalanyl-tRNA synthetase. FEBS Lett. 583, 3204–3208.

Yang, S.C., Park, S., Makowski, L., Roux, B., 2009. A rapid coarse residue-based computational method for X-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes. Biophys. J. 96, 4449–4463.

Zeev-Ben-Mordehai, T., Mylonas, E., Paz, A., Peleg, Y., Toker, L., Silman, I., Svergun, D.I., Sussman, J.L., 2009. The quaternary structure of Amalgam, a Drosophila neuronal adhesion protein, explains its dual adhesion properties. Biophys. J. 97, 2316–2326.