REVIEW

Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls

David A. Jacques and Jill Trewhella*

School of Molecular and Microbial Biosciences, The University of Sydney, Sydney, New South Wales 2006, Australia

Received 23 December 2009; Revised 17 January 2010; Accepted 19 January 2010
DOI: 10.1002/pro.351
Published online 21 January 2010 proteinscience.org

Abstract: The last decade has seen a dramatic increase in the use of small-angle scattering for the study of biological macromolecules in solution. The drive for more complete structural characterization of proteins and their interactions, coupled with the increasing availability of instrumentation and easy-to-use software for data analysis and interpretation, is expanding the utility of the technique beyond the domain of the biophysicist and into the realm of the protein scientist. However, the absence of publication standards and the ease with which 3D models can be calculated against the inherently 1D scattering data means that an understanding of sample quality, data quality, and modeling assumptions is essential to have confidence in the results. This review is intended to provide a road map through the small-angle scattering experiment, while also providing a set of guidelines for the critical evaluation of scattering data. Examples of current best practice are given that also demonstrate the power of the technique to advance our understanding of protein structure and function.

Keywords: Small-angle scattering; neutron scattering; X-ray scattering; protein structure; protein complexes; structural modeling; SAXS; SANS; contrast variation

Introduction

The small-angle scattering of X-rays from macromolecules in solution and its relationship to basic shape information has been understood for over 80 years.[1] The earliest biological small-angle X-ray scattering experiments, performed in the 1950s, necessarily used easily purified proteins such as hemoglobin and ovalbumin. Although data interpretation at that time was restricted to relatively simple parameters, such as radius of gyration, it was believed even then that the most important application of small-angle scattering would be to the study of biological macromolecules.[2] In the 1970s and 1980s, the molecular biology revolution and advances in instrumentation saw some growth in applications of the technique. A significant advance in this period was in the use of small-angle neutron scattering with contrast variation to characterize the shapes and dispositions of components within bimolecular complexes. Neutron

Abbreviations BME, β-mercaptoethanol; CD, circular dichroism spectropolarimetry; D_{max}, maximum linear dimension; DTT, dithiothreitol; EM, electron microscopy; HSQC, heteronuclear single quantum coherence; NMR, nuclear magnetic resonance; R_g, radius of gyration; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TCEP, tris(carboxyethyl) phosphine.

**Correspondence to:* Jill Trewhella, Building G08, Butlin Ave, University of Sydney, Sydney, NSW 2006, Australia. E-mail: jill.trewhella@sydney.edu.au.

Grant sponsor: Australian Research Council (ARC) Discovery Project; Grant number: DP09984536; Grant sponsor: Australian Institute of Nuclear Science and Engineering (AINSE) Postgraduate Research Award.

contrast variation revealed that DNA was wrapped around the outside of nucleosome core particles providing an understanding of how DNA was packaged in chromosomes[3] 22 years before the first high-resolution crystal structure of the nucleosome appeared.[4] Through the 1970s and 1980s, two groups, Engelman and Moore and coworkers, and May et al., used neutron contrast variation to systematically map the positions of protein and RNA subunits within the ribosome leading to a comprehensive map of the 30S subunit[5] and a partial map of the larger 50S subunit.[6] It would take another 8 years before the first high-resolution crystal structure of the large subunit of the ribosome would appear.[7]

With the sophistication and, at the time, unique nature of these experiments, the main challenge lay in data interpretation, and experts in theoretical and computational methods were required for data analysis and modeling. These difficulties have largely been overcome by the ready availability of powerful desk top computers and the development and continual improvement of data interpretation and modeling tools; today the most widely used is the ATSAS suite of programs from the Hamburg EMBL (European Molecular Biology Laboratory) group led by Svergun and coworkers[8] as it has been developed explicitly for structural biology applications and is readily downloaded from the web complete with user manuals. These developments have made small-angle scattering accessible to the broader structural biology community. Consequently, the scope of questions that can now be answered by small-angle scattering is growing rapidly, and its potential application to biology has never been greater.

As yet, no community-accepted criteria or standards have been agreed to for the publication of small-angle scattering data. In the case of crystallography and NMR, the pressure to develop such standards grew as the techniques became more widely used, and so the development of broadly accepted standards must also evolve as the use of small-angle scattering continues to grow. As small-angle scattering yields data that are inherently one-dimensional, and the user invariably is looking to support a three-dimensional model, overinterpretation of the data will always be a risk if not given careful consideration. Importantly, the absence of quality control in sample characterization and data reduction can mislead the experimenter. In this review, we describe the small-angle scattering experiment, its range of application, and the necessary quality control measures. A particular emphasis is given to the importance of standards for demonstrating sample quality, and the power of other biophysical and biochemical methods in providing constraints for modeling that will reduce the likelihood of overinterpretation of the scattering data. It is intended that this review be a guide to the nonspecialist small-angle scattering experimenter, so that they may avoid potential pitfalls as the biological applications continue to expand into new territory.

The Basics of Small-Angle Scattering

It is important for the small-angle scattering experimenter to have a basic understanding of the underlying physics of the technique to avoid being misled. We therefore provide a short description of the essential elements here.

X-rays and neutrons have properties of plane waves, that is, amplitude and wavelength, and as they pass through matter, secondary wavelets are generated by interactions with individual atoms, and the resulting coherent scattering can constructively or destructively interfere. The diffraction patterns from proteins in crystals that are exploited in crystallography are the result of such interference, between wavelets scattered by atoms within protein molecules that are regularly spaced in the crystal lattice. Small-angle scattering arises from the coherent secondary wavelets that are scattered by atoms within a single molecule, and as a result it is observed for molecules in crystals or in solution. Although small-angle solution scattering is often described as a low-resolution technique (as it does not provide information on atomic coordinates), it is more appropriate to describe it as a technique capable of providing high-precision information with respect to size and shape.[9] It is the rotational averaging of the molecules in solution that limits the information content of small-angle scattering more than the resolution limits of the experiment. Indeed, although crystallographers refer to resolution in terms of the highest angle data measured, as it is a measure of the smallest distances between scattering centers that can be resolved, the small-angle scattering specialist will refer to the resolution limits more often in terms of the smallest angles for which data can be measured as it is this resolution limit that determines the longest distances that can be characterized by the data. This effect is a direct result of the reciprocal nature of the scattering data with respect to real space dimensions, and the condition for the minimum angle required to characterize a scattering particle with a given maximum dimension is given below.

To interpret scattering data in terms of accurate structural parameters, the scattering signal must be measured from a sample of monodisperse, identical particles. Sample preparation is therefore a critically important step [Fig. 1(A)] and will be discussed more completely later.

Although there are several technical challenges associated with the development of small-angle scattering instrumentation, schematically the experimental setup is relatively simple [Fig. 1(B)]. A highly collimated X-ray or neutron beam is used to

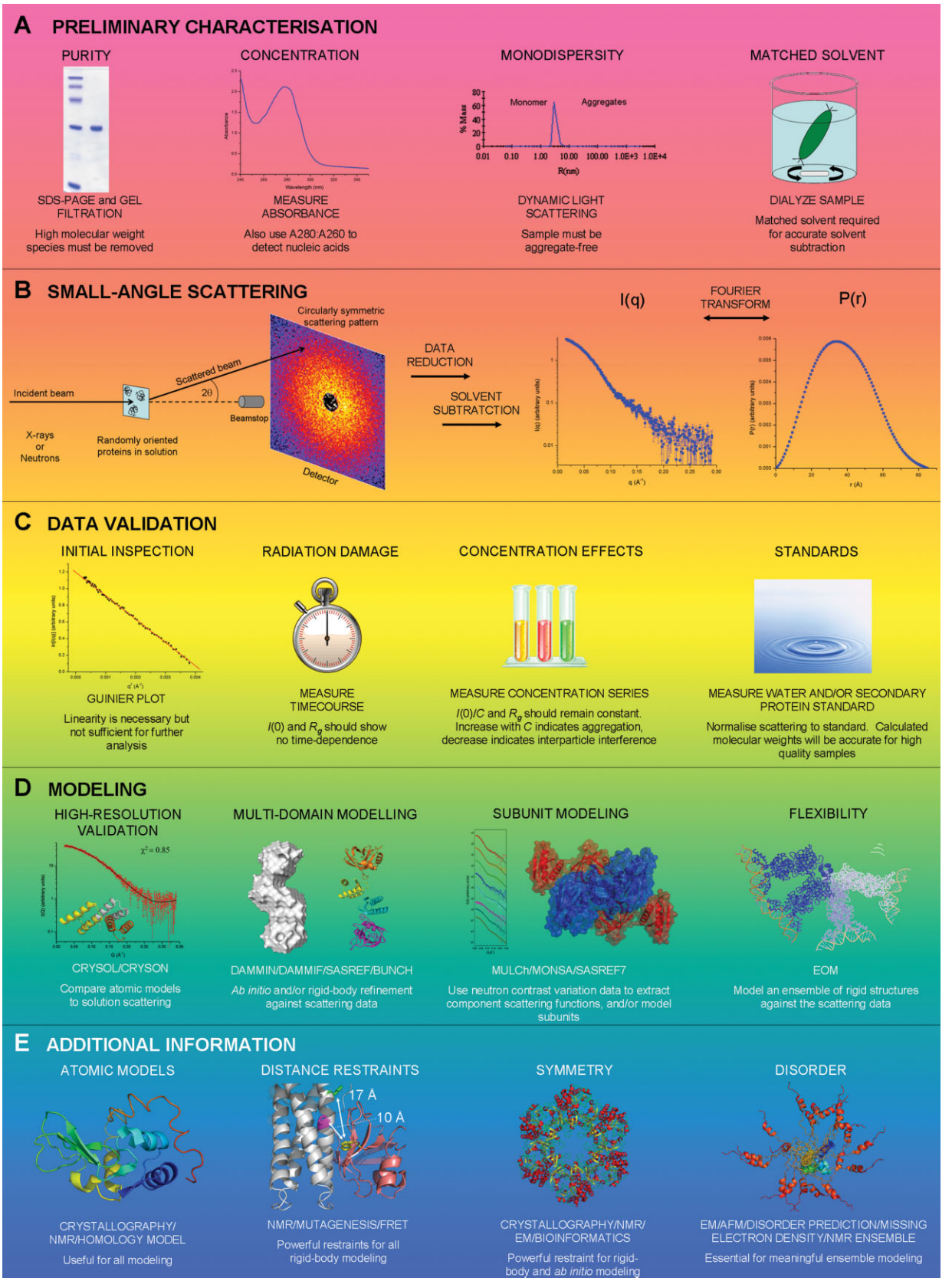

Figure 1. Roadmap through the small-angle scattering experiment.

illuminate the sample, usually a protein or macromolecular complex in solution (typically >1 mg mL^{-1} in 5–30 μL for X-ray scattering and >3 mg mL^{-1} in 150–300 μL for neutron scattering). Traditionally, the radiation is of a single wavelength (or narrow band of wavelengths), although the development of pulsed neutron sources has led to time-of-flight neutron scattering instruments that can use white radiation to maximize flux. The scattered radiation is recorded on a detector, while the direct beam is usually absorbed by a beam stop; the size and position of which are key factors determining the minimum angle measured in an experiment. New developments in solid-state devices have led to detectors that can absorb an intense direct X-ray beam without incurring damage.

For reasons of mathematical convenience, the scattering pattern is normally described by intensity (I) as a function of the amplitude of the scattering vector or momentum transfer, q:

$$q = \frac{4\pi \sin\theta}{\lambda}, \tag{1}$$

where λ is the wavelength of the incident radiation, and θ is half the angle between the incident and scattered radiation [Fig. 1(B)]. As λ is fixed and θ is small (typically <3°), $I(q)$ versus q is essentially the intensity as a function of scattering angle. The relationship between a scattering particle and its contribution to the scattering profile can be mathematically expressed as follows:

$$I(q) = \left\langle \int |\, (\rho(\vec{r}) - \bar{\rho}_s) e^{i\vec{q}\cdot\vec{r}} \mathrm{d}\vec{r}\, |^2 \right\rangle, \tag{2}$$

where $\langle\ \rangle$ refers to the rotational average and $\rho(\vec{r}) - \bar{\rho}_s$ is the difference in scattering density between the volume element at position $\vec{r}$ within the scattering particle and that of the solvent. The mean difference between the particle and solvent scattering density is termed the "contrast" and is represented as Δρ (Fig. 2, discussed in detail later).

Although $I(q)$ versus q is related to the shape of the macromolecule in solution, the profile is not intuitively informative, and so to interpret a scattering profile in terms of a structure, it is useful to Fourier transform the scattering profile to obtain the interatomic distance distribution function, $P(r)$, of the scattering particle [Fig. 1(B)] ($P(r)$ is also referred to as the pair distance or vector length distribution function). Just as a diffraction image can be Fourier transformed to obtain the Patterson function that is used to locate heavy atom positions in X-ray crystallography, so too can an $I(q)$ profile be Fourier transformed to give $P(r)$ (the radial Patterson function), which describes the probable frequency of interatomic vector lengths (r) within a protein. The $P(r)$ profile is sensitive to the symmetry and domain structure within proteins and as such can yield information concerning the shape and volume occupied by a protein or protein complex. An excellent demonstration of the effect of protein shape on the $P(r)$ curve is reported in Figure 5, page 1746 of Svergun and Koch.[10] Because scattering data can only be measured over a finite q range, $P(r)$ calculations depend upon indirect Fourier transform methods and as such they depend upon assumptions, such as the fact that $P(r)$ is zero at $r = 0$ and at the maximum linear dimension, D_{max}. As a result, D_{max} is a "soft" or model parameter in the interpretation of scattering data, and its uncertainty is highly dependent on the quality of the scattering data. Further, data must be measured to q values $<\pi/D_{max}$ to accurately characterize the longest dimensions of the particle.[11] It is generally not valid to assign an error to D_{max}.

From a high-quality small-angle scattering profile, there are two parameters relating to the size and shape of the scattering particle that can be readily calculated with high precision; the forward (or zero-angle) scattered intensity, $I(0)$, and the radius of gyration, R_g. $I(0)$ is the intensity of radiation scattered through zero angle ($2\theta = 0°$). This value cannot be measured, as it cannot be distinguished from the radiation that passes through the sample unscattered (i.e., the direct beam). However, $I(0)$ can be determined by extrapolation. $I(0)$ is related to the number of scattering particles per unit volume (N) and the particle volume (V) squared. Alternatively, it can be expressed in terms of the mass per unit volume (C), molecular weight (MW), the contrast (Δρ), and the partial specific volume (ν) of the particle:

$$I(0) = N(\Delta\rho V)^2 = \frac{C\Delta\rho^2 v^2 \mathrm{MW}}{N_A}, \tag{3}$$

where N_A is Avagadro's number.

The R_g is defined as the root-mean-squared distance of all elemental scattering volumes from their centre of mass weighted by their scattering densities. Objects with the same volume but with different shapes have different R_g values. Thus, R_g provides information as to the mass distribution within a particle.

A useful means for quickly estimating values for R_g and $I(0)$ is the Guinier approximation. In 1939, Guinier showed that for sufficiently small values of q:

$$I(q) = I(0) e^{\frac{-q^2 R_g^2}{3}}. \tag{4}$$

Thus, a linear fit of $\ln[I(q)]$ versus q^2 yields R_g and $I(0)$ from the slope and y-intercept, respectively. For globular proteins, this relationship is

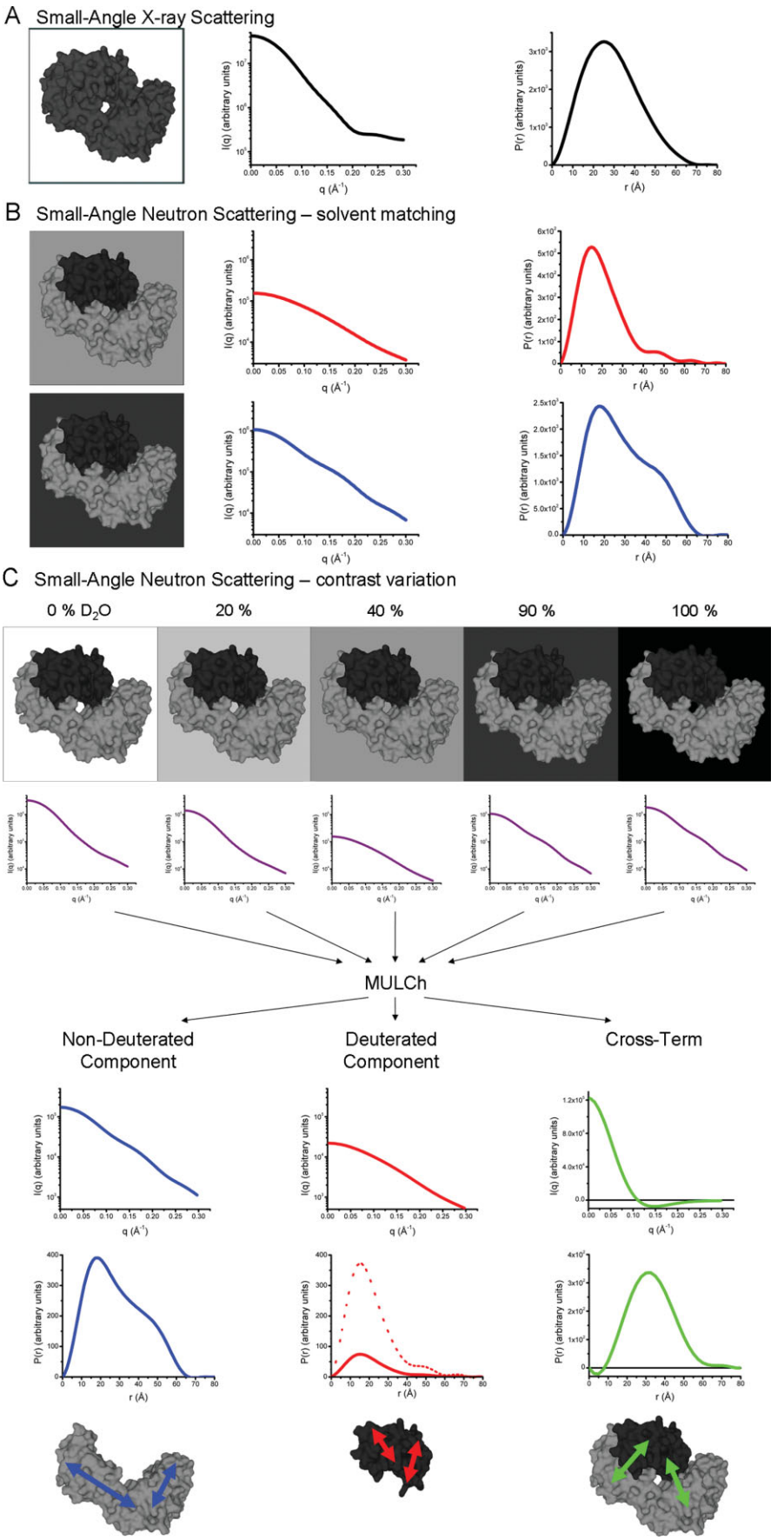
A Small-Angle X-ray Scattering
I(q) (arbitrary units)
q (Å-1)
P(r) (arbitrary units)
r (Å)
B Small-Angle Neutron Scattering – solvent matching
C Small-Angle Neutron Scattering – contrast variation
0 % D2O
20 %
40 %
90 %
100 %
MULCh
Non-Deuterated Component
Deuterated Component
Cross-Term

Figure 2.

true only for $qR_g < 1.3$, while for elongated shapes the upper limit is smaller, increasingly smaller for increasingly extended molecules.[12] The necessity that a Guinier plot be linear for globular proteins is a rapid and powerful diagnostic of sample quality.

A method for determining $I(0)$ and R_g that uses the entire scattering profile, and which therefore can be more precise, is to calculate them from the *P(r)* function. R_g is the second moment of the *P(r)* function:

$$R_g^2 = \frac{\int P(r) r^2 \mathrm{d}r}{2 \int P(r) \mathrm{d}r}, \quad (5)$$

while $I(0)$ is the zeroth moment, which corresponds simply to the area under $P(r)$:

$$I(0) = 4\pi \int_0^{D_{\max}} P(r) \mathrm{d}r. \quad (6)$$

The relative ease with which R_g and $I(0)$ can be calculated provides a rapid means by which structural information can be obtained from scattering data. Also, the fact that they can be calculated from several different mathematical models (Guinier or $P(r)$) allows an assessment to be made regarding the internal consistency and robustness of the data.

Contrast and Contrast Variation

For the purposes of small-angle scattering data analysis in terms of molecular shape, when the internal density fluctuations within a protein are small with respect to the difference between the mean scattering density of the protein ($\bar{\rho}_p$) and that of its solvent ($\bar{\rho}_s$), one can treat the protein as a uniform scattering density object with a mean contrast $\Delta\rho = \bar{\rho}_p - \bar{\rho}_s$. Equations (2) and (3) then imply that the intensity of the small-angle scattering signal from a protein in solution is proportional to the square of this contrast, that is,

$$I(q) \propto \Delta\rho^2, \quad (7)$$

where $\Delta\rho$ is readily calculated from the atomic compositions of the protein and its solvent. If the mean scattering density of the protein is the same as the solvent ($\bar{\rho}_p - \bar{\rho}_s = 0$), in practical terms no small-angle scattering from the protein is observed, as the contributions to the small-angle scattering arising from the internal density fluctuations are generally insignificant. Contrast variation experiments involving the systematic manipulation of protein and solvent scattering densities can be used to great advantage in studies of protein complexes.

As electromagnetic radiation, X-rays interact with the electron clouds surrounding each atom. Thus, the more electrons about an atom, the more likely it is that X-rays will be scattered. Proteins generally have similar elemental compositions and therefore two different proteins will usually have similar X-ray scattering density, and there is no easy and benign way to manipulate the scattering density of the solvent.

Neutrons, on the other hand, are neutral particles and have no associated electric field. As such they penetrate the electron cloud and are scattered by atomic nuclei. The scattering powers of nuclei are determined by the properties of the neutron–nucleus interaction and vary in a nonsystematic manner with atomic number and isotopes of the same element can exhibit very different neutron scattering powers. One of the largest differences in neutron scattering power is that between naturally abundant hydrogen (^{1}H) and its stable isotope deuterium (^{2}H or D). It is therefore possible, in a small-angle neutron scattering experiment, to have different contrasts for two interacting proteins if one protein is deuterated. A contrast series is obtained by

Figure 2. Small-angle scattering from a complex between a deuterated and nondeuterated protein: the contrast variation experiment. The figure illustrates contrast variation using a structure (PDB: 3GMR) of T-cell surface glycoprotein CD1d1 in complex with beta-2 microglobulin. Theoretical scattering profiles were generated for the protein complex in which the beta-2 microglobulin component had its nonexchangeable hydrogens deuterated to 60% *in silico*.[66] (A) The proteins each have the same electron density and hence X-ray scattering density and the small-angle X-ray scattering profile therefore yield information on the shape of the entire complex as a uniform contrast (depicted as all black) particle. (B) The neutron scattering contrast for the deuterated (90% gray) and nondeuterated (40% gray) components is distinct, and it is therefore possible to measure scattering from each individual component by solvent matching the scattering density of the other by H/D substitution in the solvent. The nondeuterated protein will be solvent matched around 40% D_2O, whereas the deuterated protein's solvent match point will be 90% D_2O, this value depending upon the deuteration level of the protein (60% in this example). (C) A theoretical set of data acquired for different %D_2O in the solvent (a contrast series) yields information on the shapes and dispositions of the deuterated and nondeuterated proteins. From the contrast series, it is possible to extract scattering profiles corresponding to the deuterated and nondeuterated components (red and blue, respectively) and a cross-term (green) that is related to their relative dispositions. Note that because of the smaller size of the deuterated component, the corresponding *P(r)* needs to be multiplied by a factor of five to observe the curve clearly. Also, the cross-term contains negative values for *I(q)* and therefore must be plotted on a linear scale.

recording SANS data in solvents with different ratios of H_2O:D_2O; in this manner, the contributions of the two components of the complex to the scattering can be systematically varied. Information concerning the shape of the individual proteins and their relative dispositions can then be extracted from the contrast series (Fig. 2). Greater detail on this subject can be found in Whitten *et al.*,[13] which describes MULCh, a useful set of tools for calculating contrasts, extracting component scattering functions, and other analyses for the extraction of structural information from contrast variation data.

The Nature of Small-Angle Scattering Data

In other structural techniques, interpretable data are usually only obtained when the sample is properly prepared. In crystallography, a diffraction pattern is only obtained from a good quality crystal, while in NMR a simple inspection of an HSQC can immediately tell the experimenter if the protein is folded and hence if it would be worth while proceeding with further experiments. Small-angle scattering patterns can *always* be obtained from *any* quality of sample, and, as such, verification that the scattering particles are monodisperse and identical is essential before data analysis in terms of a structural model can proceed. When adequate quality control is not demonstrated, it is not possible to have confidence in the structural models derived from small-angle scattering data.

Preliminary Sample Characterization

The first step in quality control is to ensure that the sample is thoroughly characterized before the scattering experiment is performed. The stringent criteria outlined later, and summarized in Figure 1(A), must be met before interpretable scattering data can be expected.

Requirements on sample purity

The sample to be measured must be highly pure. Macromolecular impurities (protein or nucleic acid) need to be removed, especially if they are of higher MW than the macromolecule of interest. The scattering signal is proportional to the square of the MW [Eq. (3)], and therefore small amounts of contamination by high-MW species will contribute disproportionately to the scattering signal and bias the data to larger structural parameters than are true for the molecule of interest. The largest effects will be at the lowest q values, but removal of these data from the analysis is not sufficient to ensure accurate structural parameters. Purity is best assessed by SDS-PAGE as well as 280:260 nm absorbance ratio to avoid nucleic acid contamination (where appropriate).

Establishing that samples contain monodispere, identical particles

As stated earlier, contaminant high-MW species must be removed. Similarly, sample aggregation must also be strictly avoided. In general, aggregation is not observed by SDS-PAGE and is best monitored by dynamic light scattering and removed by size-exclusion chromatography. A special case of aggregation is disulfide-mediated aggregation, which can appear over time in protein samples carrying reduced cysteine residues. The presence of inappropriate disulfide formation can be detected by SDS-PAGE run with and without β-mercaptoethanol (BME). The nonreducing gel will contain higher MW species corresponding to dimers, trimers, and so forth, of the sample protein if disulfide-mediated aggregation is present. Proteins that contain reduced cysteines must be rigorously handled under reducing conditions, and the effects of pH carefully considered as pH can dramatically affect the behavior of S—S linkages.

Importance of protein concentration determination

The small-angle scattering signal is proportional to the concentration and MW of the macromolecule being measured [Eq. (3)]. If an accurate sample concentration can be measured, then the MW of the macromolecule can be estimated from the scattering data. This parameter (along with partial specific volume and $\Delta\rho$) can be critical for data evaluation (see later) as an accurate MW determination can be used to demonstrate sample monodispersity. Accurate concentrations are also required if the oligomerization state of the sample is to be investigated. Typically, concentrations are determined spectrophotometrically by measuring the absorbance at 280 nm, using calculated extinction coefficients.[14] The accuracy of this method is related to the extinction coefficient of the protein, its purity, and the absorption properties buffer constituents. Results obtained from proteins carrying no tryptophan residues must be treated with skepticism as errors significantly greater than 10% have been observed. Inaccuracies in protein concentration can mislead the experimenter, especially if one is depending on $I(0)$ for the determination of the MW or volume of the scattering particle. Great care must be taken to ensure sample purity (small amounts of highly absorbing contaminants such as nucleic acids can result in overestimation of sample concentration). Solvent blanks can also be a source of error. For example, when using DTT to maintain reducing conditions or as a free radical absorber to minimize radiation damage (see later), one needs to be aware that there can be different rates of DTT oxidation between the protein-containing samples and buffers. These differences can result in a net difference in the measured UV absorbance at 280 nm of a sample against its solvent blank as DTT absorbs strongly in this region as its absorbance is dependent on its oxidation state.

Alternatives to DTT as a reducing agent are BME or TCEP, the latter being preferred as its low volatility maintains the reducing environment for extended periods of time and does not lead to a UV absorbance mismatch between sample and solvent blank. Ascorbate buffers can also be used as free radical absorbers.

Solvent blank measurements

The solvent scattering must be subtracted from the protein plus solvent scattering to obtain the scattering from the protein molecule. Perfectly matched solvent blanks are important for accurate absorbance measurements and for scattering data analysis. Inaccuracies in solvent subtraction manifest mostly in the high-q region of the scattering data and can interfere with modeling calculations.[15,16] A good test for the adequacy of the solvent subtraction is to release the requirement for $P(r)$ to come to zero at $r = 0$. If $P(0)$ is positive or negative, then the solvent has been under- or oversubtracted, respectively.

It can be difficult to match the solvent perfectly, and it is often unavoidable that a small additive correction to the solvent scattering needs to be made. This effect can be especially problematic in neutron scattering experiments because of the large incoherent scattering from ^{1}H that results in an isotropic background contribution that can be mismatched as a result of subtle differences between sample and solvent measurement, such as very small changes in the path length. Solvent mismatch is often best minimized by using a last step dialysate as the solvent blank. Often, samples need to be concentrated before the scattering experiment, and it is tempting to use the filtrate from a centrifugal device as the solvent blank. Our experience suggests that even after extensive prewashing of the membrane, the filtrate can be imperfectly matched to the sample (possibly because of the presence of glycerol on the membrane). Therefore, concentrated samples are best dialyzed against the appropriate buffer before measurement wherever possible.

Scattering Data Acquisition and Evaluation

Once a "scattering-quality" sample has been obtained, certain control experiments are required to assess the validity of the scattering data and exclude certain sources of interference [Fig. 1(C)]. The following describes how to interrogate scattering data and the methods used for demonstrating their robustness. Useful software from the ATSAS suite includes PRIMUS[17] for data inspection and GNOM[18] for performing $P(r)$ analysis.

Detecting radiation damage

X-ray-induced radiation damage arises from bond breakage and free radical formation in the samples and can be a source of radiation-induced aggregation.[19] For this reason, it is often a good idea, especially at synchrotron intensities, to have a free radical absorber, such as DTT or ascorbate, in the solvent. Radiation-induced aggregation must be assessed on a case-by-case basis as it depends upon characteristics such as molecular mass, composition, and structure of each individual protein. As the effects are dose-dependent, a useful strategy for X-ray scattering experiments is to collect several exposures and compare the resultant scattering profiles. Radiation-induced aggregation will be evident in increasing R_g and $I(0)$ values with time. Multiple exposures allow the experimenter to determine how long an exposure the sample can withstand before the data become unusable. If no time dependence for the R_g and $I(0)$ values is observed, the data from the multiple exposures can be combined into one dataset for analysis. For some laboratory-based X-ray instruments, it is impractical to collect data for multiple periods of sufficiently short duration for this approach. In these cases, performing SDS-PAGE before and after sample irradiation is good practice for each new protein under study. Because of the low energy of cold neutrons (equivalent to infrared electromagnetic energies), radiation damage is rarely seen in neutron scattering. The term 'cold neutrons' refers to the fact that the longest wavelength (4–6 Å) neutrons are used for small-angle neutron scattering to maximize resolution (q_{min}) and these wavelengths are obtained by passing thermal neutrons through a cold source, such as liquid hydrogen.

Initial inspection of the scattering data

Once the scattering data have been reduced to $I(q)$ versus q and solvent subtraction performed to leave only the scattering from the molecules of interest, a simple and powerful way to initially evaluate the solvent subtracted scattering data is to view them as $\log[I(q)]$ versus $\log(q)$. If the sample contains monodisperse, identical particles, and the data are measured to low enough q to reliably characterize their largest dimensions, then the plot will have a discernable flat region in the lowest q regime.[20]

The next analysis should be a Guinier plot, which is most sensitive to the longest distances present in the sample. A source of bias for the long-distance information in the scattering signal is protein aggregation, which can cause an upturn in the Guinier plot (which can be dramatic or simply appear as a "smiling" Guinier) and associated increases in the observed R_g and $I(0)$ values. A downturn in the Guinier plot (the "frowning" Guinier) is symptomatic of interparticle interference—an effect due to correlated distances between particles in solution caused by Coulombic repulsion[21] (discussed later). Examples of Guinier plots for lysozyme data are shown in Figure 3(D–F). Importantly, if aggregation (including oligomerization) or interparticle interference is sufficiently subtle, the Guinier plot can be perfectly linear, but the R_g and $I(0)$ values will not be accurate,

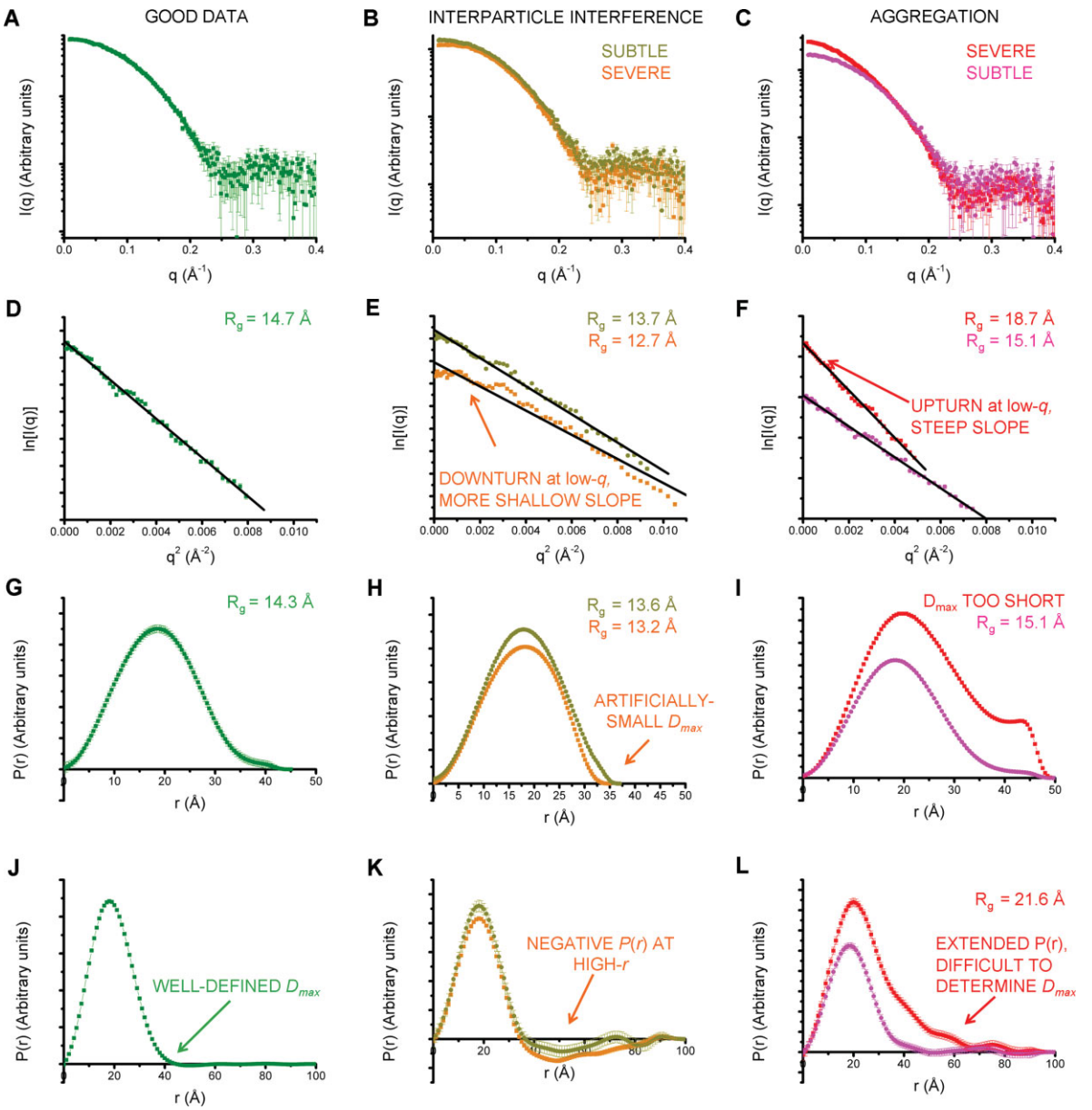

Figure 3. Effects of interparticle interference and aggregation on small-angle scattering data. It is often difficult to assess sample quality from $I(q)$ plots alone (A–C). Guinier plots (D–F) can be instructive, as deviations from linearity are clear indicators of interparticle interference (E, orange) or aggregation (F, red). The $P(r)$ presented with a single chosen D_{max} (G–I) can be deceptive, and it is possible that only severe aggregation is detected (I, red). Behavior of the $P(r)$ at extended D_{max} values (J–L) gives the clearest indication of sample quality. It is important to note that the subtle cases (K, dark yellow and L, magenta) may escape detection by these initial data inspection methods, and as such comparison of the data to secondary standards is essential. In evaluating these differences of around 1 Å in R_g, it is important to note that such a difference is well outside the precision with which R_g can be determined (generally to within a few tenths of Å for a small- to medium-sized protein), and that this degree of inaccuracy will significantly bias any 3D modeling performed against the scattering data.

being either too large or too small, respectively, and thus bias any 3D modeling performed using the scattering data.[22] A linear Guinier plot is therefore necessary, but not sufficient, for the accurate interpretation of globular proteins, and good practice is to always show it along with the secondary standards that confirm the scattering particle has the expected volume or mass (see below).

Another way to detect aggregation is in the behavior of the $P(r)$ function. Proteins do not have hard surfaces, and as such, the distribution of longest dimensions is expected to approach zero (at D_{max}) in a concave fashion as viewed from above the horizontal axis [Fig. 3(G,J)]. An abrupt arrival of $P(r)$ at zero suggests that the chosen D_{max} is artificially short. Badly aggregated samples may give small peaks in

$P(r)$ at high r values or the $P(r)$ curve may never approach zero appropriately [Fig. 3(I,L)], whereas samples exhibiting interparticle interference will reach D_{max} at artificially small values [Fig. 3(H,K)]. In some cases, especially where the scattering molecules have flexibility, it may be difficult to choose D_{max}. In these cases, Guinier-determined R_g and $I(0)$ values may be used as a guide. Importantly, one must always keep in mind that D_{max} is a soft parameter in $P(r)$ modeling; for high quality data with accurate background subtractions D_{max} can be chosen with reasonable reliability, but confidence in the choice of D_{max} depends upon understanding how the $P(r)$ transform behaves for a range of chosen D_{max} values. The emergence of easy to apply 3D modeling methods has led to a practice of skipping or failing to report $P(r)$ and Guinier analyses, yet these are important steps in evaluating data quality and sample integrity.

The importance of measuring data at multiple concentrations

As described in Eq. (2), $I(0)$ is proportional to the sample concentration (C); however, aggregation or interparticle interference effects can complicate this relationship. A decrease in $I(0)/C$ with increasing C is indicative of interparticle interference and will be accompanied by an apparent decrease in R_g. Such effects are often linearly dependent on concentration and can therefore be eliminated by extrapolation to infinite dilution.[23] Alternatively, charge screening (by increasing the ionic strength of the solvent) or adjusting the pH to nearer the isoelectric point of the protein can perturb the Coulombic repulsion and reduce the interference effect.[24] An increase in $I(0)/C$, and accompanying increase in R_g with C, indicates concentration-dependent aggregation or oligomerization. In the latter case, this can be instructive, as monomer–multimer equilibria may be present in solution (an example of such a study is the measurement of the monomer–dimer equilibrium constant for the SARS CoV main proteinase[25]). It may be possible to find a low concentration range where the aggregation is eliminated; alternatively, changes in pH or ionic strength may be helpful. A concentration series should always be measured to demonstrate that the data are free of effects that would bias any structural interpretation. The subtle example of interparticle interference shown in Figure 3 may not have been detected had a concentration series not been performed.

Standards

Calibrating $I(0)$ to a scattering standard allows the MW or volume of the scattering molecule to be determined providing its concentration is known. Verification that the scattering particle has the expected MW or volume is perhaps the most important measure to ensure that the structural parameters and interpretation are accurate for the molecule of interest. Few published small-angle scattering experiments report calibrated $I(0)$s and those that do often do not provide sufficient information to demonstrate the accuracy of this calibration. This lack of standard practice is a significant issue for those who want to assess the reliability of the data interpretation. Typically, the $I(0)$s are calibrated in one of two ways: by comparison with a known particle of similar composition such as a protein standard or by placing the data on an absolute scale using the scattering from water. Both methods require accurate protein concentration determination of the molecule of interest. When the secondary standard is a protein, the standard will typically have a large extinction coefficient at 280 nm and is measured under conditions in which it is known to be monodisperse and free of interparticle interference effects (e.g., glucose isomerase[26] or lysozyme[27]). The measured $I(0)$ values, normalized to concentration (in mg/mL) and expected MW, should be constant, provided that the molecule of interest and the standard are of similar mean contrast and the same partial specific volume [Eq. (3)]. If the molecule (or complex) has a different scattering density than the standard (e.g., proteins and polynucleotides), the differences in contrast and partial specific volume must be accounted for. Because $I(0)$ is proportional to the square of the partial specific volume and contrast, small errors or differences in these parameters can result in quite large errors in the MW determinations. Hence, great care must be taken to understand the assumptions behind these calculations and the accuracy of each parameter.

The isotropic (q-independent) scattering of water can be exploited to place scattering data on an absolute scale, which allows one to estimate the molecular mass of the scattering molecule. The scattering of pure water over a wide range of temperatures is precisely known.[28] The ratio between the experimentally determined and theoretical water scattering intensity is used to apply a multiplicative correction to the data, which then describes I in units of cm^{-1}. The absolute scaling allows the MW to be calculated directly by rearranging Eq. (3):

$$\mathrm{MW} = \frac{I(0) \cdot N_A}{C(\Delta\rho \cdot v)^2}, \tag{8}$$

where the correct units are essential: C (g cm^{-3}), $\Delta\rho$ (cm^{-2}, which can be calculated using MULCh[13]), and ν (cm^3 g^{-1}, which can be calculated using NucProt[29]).

The methods described earlier to determine molecular mass or volume each depend on accurate measurement of the lowest angle data. Recently, Fischer *et al.*[15] described a method for determining molecular volume from the high-angle data, which has been shown to be quite powerful and provides

a way to evaluate data quality independent of sample concentration. The method relies on accurate high-q data and, as such, poor background subtractions and flexible systems can be significant sources of error.

Best practice in scattering experiments is to always put the data on an absolute scale (requiring only a water and empty cell measurement), measure a secondary protein standard, and perform the Fischer method. The reason one wants to do all three is that each method has different sources of error and so agreement between each method combined with a linear Guinier and a "well-behaved" $P(r)$ function can provide confidence that the scattering data are from samples of monodisperse identical particles of the expected molecular mass and hence will yield accurate structural parameters.

An example of how one can be misled by inaccurate concentration or partial specific volume information is from our own work that used the NMR structure of the bacterial sporulation inhibitor, Sda,[30] to interpret solution scattering data. We reported that Sda was dimeric in solution for the conditions of our measurements based on an $I(0)$ analysis that turned out to use insufficiently accurate concentration values for Sda.[31] Sda is a small (46-residue) protein with no tryptophan residues and an unusually weak UV extinction coefficient. Subsequent to our scattering experiments, we solved the crystal structure of Sda and observed a tight hexamer that caused us to test our earlier conclusions from the scattering data.[32] Multiangle laser light scattering combined with a reexamination of the NMR data suggested that the trimer was weakly associated and probably an artifact of the high protein concentration required for structural analysis. Further reexamination of the scattering data revealed that the trimer was a significantly better fit to the scattering data than the dimer; the χ^2 for the fit dropping to 0.85 (with no steric clashes) from 1.08 (with some steric clashes). This example serves as an important cautionary tale demonstrating the need not only for accurate data evaluation and a good understanding of error propagation in the data but also the value of having multiple techniques available for structure analysis.

Data Analysis and Complementary Information

The specific analysis undertaken with small-angle scattering data largely depends on the nature of the problem being investigated and the additional information available [Fig. 1(D,E)]. As stated above, rotational averaging limits the information content in a small-angle scattering experiment. As such, the data are susceptible to overinterpretation with the increasing parameterization of any modeling calculation. This problem is not unique to small-angle scattering. In crystallographic refinement, overparameterization is indicated by a significant drop in the R factor without a concomitant drop in the free-R factor.[33] In this situation, restraints are applied (such as noncrystallographic symmetry and geometric restraints) to reduce the degrees of freedom in the refined parameters. The limited information content of the small-angle scattering experiment prevents the calculation of a free-R factor equivalent, and so an understanding of the nature of the modeled parameters is essential to avoid overinterpretation. To that end, an "Occam's razor" approach is recommended, that is, that additional parameters be introduced only when simpler models fail to adequately explain the data. Additionally, data from complimentary techniques can be a powerful source of restraints to greatly improve the confidence in the uniqueness of best-fit models. Conversely, inconclusive modeling can indicate a need for restraints and hence direct further experimentation. Here, we discuss some of the applications of small-angle scattering, the type of analyses that are appropriate, and the nature of additional information that can be used to provide restraints.

High-resolution structure validation and refinement

One of the simplest and most popular uses for small-angle scattering data is structure validation. The theoretical scattering from atomic models of proteins can be easily calculated using CRYSOL[34] or CRYSON[35] and evaluated against measured scattering data. Atomic models for proteins can be obtained by crystallography or NMR, or theoretical models can be built by homology modeling (e.g., using SWISS-MODEL[36]). X-ray crystallography is the dominant method of protein structure determination but because of the nature of crystal packing, the solution structure or oligomeric state may not be clear. In the case of release factor 1 (RF1, responsible for termination of translation in *E. coli*), a more compact form of the protein was observed by crystallography compared with that observed in complex with the ribosome by electron microscopy (EM).[37] Small-angle X-ray scattering was able to resolve these differences by demonstrating that the more extended (EM) form was present in solution, and that the compaction was an artifact of crystallization.

Similarly, ambiguities in generated NMR structures have been resolved by comparison against scattering data. NMR provides short distance restraints, and hence scattering data can provide additional long distance restraints to improve solution structures. Methods have recently been developed to corefine NMR and scattering data, thereby increasing accuracy of structure determination of large proteins by NMR.[38–41] In these corefinement methods, it is essential to have precise and accurate

data if the objective is to increase the accuracy of the derived solution structures.

The protein data bank requires that the relevant "biological assembly" be reported during structure deposition. Yet, crystallography and NMR frequently yield ambiguous results as to the oligomeric state of the protein. Small-angle scattering can be a simple means for determining the biological assembly and is becoming increasingly utilized in structure reports. Recent examples include the functional dimer of superoxide dismutase from *Alvinella pompejana*[42] and the autoinhibitory dimer of the human MAP kinase, MEK6.[43]

Although scattering data are significantly challenged in proving any model correct, they can unequivocally prove a model wrong or incomplete, providing that the data are of high quality and demonstrated to be from the molecule of interest. A classic example of a case when a single crystal structure did not tell the complete story is the original scattering work on calmodulin and troponin C[44] that showed distinct differences between the crystal and average solution forms that were subsequently shown to be important in the functioning of these intracellular calcium receptors.

Characterization of multidomain proteins

One of the major advances in small-angle scattering in recent years is the ability to reconstruct a molecular shape from the scattering data alone. Previously thought to be impossible, modern *ab initio* algorithms such as DAMMIN,[45] DAMMIF,[46] and MONSA (http://www.embl-hamburg.de/ExternalInfo/Research/Sax/monsa.html) are able to represent the particle by finite volume elements (or dummy atoms) and employ simulated annealing to fit the experimental data. The calculation is constrained to produce a "physically sound" model (i.e., compact and interconnected), and symmetry constraints may also be applied.

The central problem in *ab initio* shape restoration is that of uniqueness, that is, there may be several models with equally good fit to the data. The dummy atom approach allows multiple models to be generated. The program DAMAVER[47] may be used to place models into structurally similar classes and then generate an average structure from each class. It is tempting to use the number of models generated in each class as an assessment of the likelihood that average structure is correct. However, this probability argument is insufficient grounds for dismissing a particular class, and additional information (such as structure predictions, domain models, or EM data) is required. On the other hand, one class of models having a significantly greater goodness-of-fit measure to the experimental data can be grounds for favoring that averaged model.

It is unusual for *ab initio* structures to be presented alone as they are highly parameterized models and seldom yield biologically relevant information. As small-angle scattering provides information on the large-scale features, it is often possible to model multisubunit proteins based on high-resolution structures of domains, either of the protein in question or from structural homologues. Although it is possible to superpose the atomic structures on the *ab initio* model using SUPCOMB,[48] it is often more useful, and more revealing, to perform rigid-body refinement against the scattering data using SASREF.[49] As one or more molecules are free to move in space, rigid-body refinement can also suffer from overparameterization, and multiple nonunique solutions may be obtained. To reduce this problem, distance and symmetry constraints (obtained from mutagenesis, FRET, or residual dipolar coupling from NMR) can be included in the modeling calculation. SASREF is also capable of modeling against multiple datasets, such as those derived from a neutron contrast variation series or X-ray data from subcomponents of a modular protein or complex. This increase in the number of input scattering curves again helps in addressing the problem of the uniqueness of possible solutions. Here again, however, care needs to be taken to demonstrate that all the scattering curves relate to the same structure; for example, in the case of a neutron contrast series, $I(0)$ analysis is needed to demonstrate that there is no D_2O-induced aggregation; in the case of X-ray scattering from subunits or truncated proteins, one needs to demonstrate that the truncation or disassociation has not induced conformational changes.

Accurate rigid-body modeling against scattering data requires complete structures. In practice, model structures are often missing loops and/or interdomain linkers that may be disordered. The program BUNCH[49] is capable of performing the rigid-body refinement while modeling missing loops as dummy residues. BUNCH is especially useful for multidomain proteins where X-ray scattering has been performed on various partial constructs, such as polypyrimidine tract binding protein.[50]

The use of a combination of *ab initio* shape restoration and rigid-body refinement is demonstrated by the study of the N-terminal domain organization of cardiac myosin binding protein C (cMyBP-C).[51] DAMMIN revealed an elongated structure consistent with four folded domains. With that information, it was possible to construct homology models (consistent with analysis by circular dichroism spectropolarimetry) for the two unknown domains. Combined with NMR structures for the known domains, a rigid-body model was refined with BUNCH to yield a model consistent with the original DAMMIN structure. This model was then used in a small-angle neutron scattering study to demonstrate that

cMyBP-C decorates actin filaments with implications for the regulation of heart muscle contraction.[52] Other demonstrations of *ab initio* modeling of such modular proteins include the transcriptional regulator PUR-α[53] and the protein filamin,[54] which cross-links the cell membrane with the cytoskeleton. In the latter example, the small-angle scattering results drove further NMR studies to reveal a previously unseen immunoglobulin domain-pairing interaction.

Modeling multisubunit complexes

Small-angle X-ray scattering can be used to study protein complexes and is particularly informative if it is expected that one component of a complex undergoes a conformational change upon binding. Once again, the dumbbell-shaped calmodulin provides a classic example of small-angle scattering being used to detect the dramatic conformational collapse of the protein upon binding its target peptide and advancing our understanding of how this calcium receptor functions.[55] Larger complexes can also be studied by this method, such as that formed between proliferating cell nuclear antigen (PCNA) and DNA ligase[56] where a combination of *ab initio* and rigid-body modeling was used to model the complex, which clearly showed a 1:1 ratio of PCNA:DNA ligase, the latter adopting a more open conformation.

In the calmodulin example, the peptide binding partner was sufficiently small as to not contribute significantly the scattering, whereas in the PCNA-DNA ligase example, the two proteins are so large, and with such different shapes that determining their relative dispositions was possible. However, accurate determination of the shapes and dispositions of proteins in complexes generally requires using small-angle neutron scattering with contrast variation.

As seen in Figure 2(B), one can "solvent match" a component of a biomolecular complex by adjusting the D_2O composition of the solvent so that the contrast of one component is zero. If two components of a complex have different scattering densities (e.g., protein–nucleic acid or protein–deuterated protein complexes) solvent matching one component yields neutron scattering only from the other. This approach has been successfully employed for the study of protein–detergent complexes such as lipid transport protein apolipoprotein B-100[57] and for protein–deuterated protein complexes such as the synaptic proteins neuroligin and neurexin.[58] The latter case is especially pertinent to demonstrating the quality of modeling SANS data as it was soon followed by a crystal structure,[59] which revealed the same subunit arrangement.

Although solvent matching is feasible, it is often better to perform a contrast variation experiment, that is, to collect data at multiple $\%D_2O$ solvent conditions [Figs. 1(D) and 2(C)]. This strategy allows the extraction of scattering functions describing each component and the cross-term relating their positions in space.[13] Neutron contrast variation studies also obviate the need for perfectly matching the component and solvent scattering densities. A good example of such a scenario is the complex formed between the sporulation histidine kinase, KinA, and its inhibitor, Sda, from *Bacillus subtilis*.[31] By deuterating Sda and performing small-angle neutron scattering with contrast variation on the complex, it was possible to extract structural information for each component. The binding of the inhibitor caused a compaction of the dimeric KinA (seen as a reduction in R_g and D_{max} compared with the X-ray scattering), and an increase in the apparent dimension of the Sda component resulting from the two Sda molecules binding on opposite sides of the kinase dimer. A structural homologue of KinA and the NMR structure of Sda were used in rigid-body modeling against the entire contrast series data (including X-ray scattering) using SASREF7 (which incorporates the contrast information). A similar experiment was performed for the complex between KinA and its second inhibitor KipI [pictured in Fig. 1(D), third panel from left].[60] The results of these two experiments showed that the inhibitors bind the base of the KinA dimerization domain and not the flexible hinge between the dimerization and catalytic domains as had been previously proposed.[30] With this information, redirected mutagenesis and biochemical studies[61] led to new understanding of the mechanisms involved in bacterial signaling pathways; particularly, that KipI and Sda inhibited not only the autokinase reaction but also the phosphotransferase reaction. The crystal structure of a homologous kinase, KinB, from *Geobacillus stearothermophilus*, also was solved in complex with Sda,[62] confirming that the Sda binding site was correctly predicted by the neutron scattering model, albeit with a 180° rotation of Sda—information that was beyond the resolution of the small-angle scattering experiment.

Flexible systems

As small-angle scattering is a solution technique, the data are time and ensemble-averaged. Thus, it is possible to measure data for flexible systems. If a system cannot be adequately described by simple rigid-body modeling, it is possible to consider conformational flexibility using an ensemble model with the ensemble optimization method[63] or BIBLIOMOD.[64] Because of the very high parameterization of ensemble modeling, extreme care must be taken to ensure that data are not being overinterpreted. Ideally, evidence for flexibility should be obtained to demonstrate that ensemble modeling is a valid

approach, such as disorder predictions, EM data, NMR ensembles, or missing electron density in crystal structures.

An excellent example of a flexible system studied by small-angle scattering is protein kinase R (PKR), a key component of the interferon antiviral pathway.[20] PKR is known to carry three folded domains (whose structures have been solved) demonstrated by atomic force microscopy to be separated by two long disordered linkers.[65] Linear Guinier plots and $I(0)$ analysis demonstrated sample monodispersity, yet it is not possible to adequately describe the structure with a single BUNCH-derived model. Good fits to the data require an ensemble of multiple conformations to account for the intrinsic flexibility of the protein. These ensembles support the view that there is no interaction between any of the domains in solution.

Concluding Remarks

Small-angle scattering experiments must demonstrate sample quality and data validation by reporting the results of the kinds of quality control measures described here. Especially important are the detailed reporting of methods for protein concentration determination, of Guinier and $P(r)$ analyses for demonstrating monodispersity, a concentration series for the discounting of interparticle interference and concentration-dependent aggregation, and normalization to at least one, preferably two secondary standards (water and a protein standard) for correct volume (and hence MW) determination. We also recommend that the "Occam's razor" approach be taken with any 3D modeling, and that additional experimental data be used to avoid overparametrization and to substantiate assumptions.

The small-angle scattering studies described here represent examples of the systems to which the technique has been applied in recent years, highlighting the broad applicability of the technique, along with its limitations. The utilization of the technique is increasing rapidly and it appears inevitable that small-angle scattering will be called upon to answer even more biological questions in the future.

Acknowledgment

The authors thank Dr. Don Parkin and Dr. Yanling Lu for their assistance collecting lysozyme data in Figure 3.

References

1. Guinier A (1938) The diffusion of X-rays under the extremely weak angles applied to the study of fine particles and colloidal suspension. Comptes Rendus Hebdomadaires Des Seances De L Acad Des Sci 206: 1374–1376.
2. Guinier A, Fournet G (1955) Small-angle scattering of X-rays (structure of matter series). New York: Wiley.
3. Pardon JF, Worcester DL, Wooley JC, Tatchell K, Vanholde KE, Richards BM (1975) Low-angle neutron-scattering from chromatin subunit particles. Nucleic Acids Res 2:2163–2176.
4. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. Nature 389: 251–260.
5. Capel MS, Engelman DM, Freeborn BR, Kjeldgaard M, Langer JA, Ramakrishnan V, Schindler DG, Schneider DK, Schoenborn BP, Sillers IY, Yabuki S, Moore PB (1987) A complete mapping of the proteins in the small ribosomal-subunit of *Escherichia coli*. Science 238: 1403–1406.
6. May RP, Nowotny V, Nowotny P, Voss H, Nierhaus KH (1992) Inter-protein distances within the large subunit from *Escherichia coli* ribosomes. EMBO J 11:373–378.
7. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. Science 289: 905–920.
8. Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006) ATSAS 2.1, a program package for small-angle scattering data analysis. J Appl Crystallogr 39:277–286.
9. Neylon C (2008) Small angle neutron and X-ray scattering in structural biology: recent examples from the literature. Eur Biophys J Biophys Lett 37:531–541.
10. Svergun DI, Koch MHJ (2003) Small-angle scattering studies of biological macromolecules in solution. Rep Prog Phys 66:1735–1782.
11. Moore PB (1980) Small-angle scattering. Information content and error analysis. J Appl Crystallogr 13:168–175.
12. Hjelm RP (1985) The small-angle approximation of X-ray and neutron scatter from rigid rods of non-uniform cross-section and finite length. J Appl Crystallogr 18: 452–460.
13. Whitten AE, Ca SZ, Trewhella J (2008) MULCh: modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. J Appl Crystallogr 41:222–226.
14. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788.
15. Fischer H, Neto MO, Napolitano HB, Polikarpov I, Craievich AF (2010) The molecular weight of proteins in solution can be determined from a single SAXS measurement on a relative scale. J Appl Crystallogr 43: 101–109.
16. Rubinson KA, Stanley C, Krueger S (2008) Small-angle neutron scattering and the errors in protein structures that arise from uncorrected background and intermolecular interactions. J Appl Crystallogr 41:456–465.
17. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr 36:1277–1282.
18. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 25:495–503.
19. Kuwamoto S, Akiyama S, Fujisawa T (2004) Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. J Synchrotron Radiat 11:462–468.
20. VanOudenhove J, Anderson E, Krueger S, Cole JL (2009) Analysis of PKR structure by small-angle scattering. J Mol Biol 387:910–920.

21. Chen SH, Bendedouch D (1986) Structure and interactions of proteins in solution studied by small-angle neutron-scattering. Methods Enzymol 130:79–116.
22. Blechner SL, Olah GA, Strynadka NCJ, Hodges RS, Trewhella J (1992) $4Ca^{2+}$.troponin-C forms dimers in solution at neutral pH that dissociate upon binding various peptides—small-angle X-ray-scattering studies of peptide-induced structural-changes. Biochemistry 31: 11326–11334.
23. Miyamoto Y, Nishimura S, Inoue K, Shimamoto S, Yoshida T, Fukuhara A, Yamada M, Urade Y, Yagi N, Ohkubo T, Inui T (2010) Structural analysis of lipocalin-type prostaglandin D synthase complexed with biliverdin by small-angle X-ray scattering and multidimensional NMR. J Struct Biol 169:209–218.
24. Niebuhr M, Koch MHJ (2005) Effects of urea and trimethylamine-N-oxide (TMAO) on the interactions of lysozyme in solution. Biophys J 89:1978–1983.
25. Graziano V, McGrath WJ, Yang L, Mangel WF (2006) SARS CoV main proteinase: the monomer-dimer equilibrium dissociation constant. Biochemistry 45: 14632–14641.
26. Kozak M (2005) Glucose isomerase from *Streptomyces rubiginosus*—potential molecular weight standard for small-angle X-ray scattering. J Appl Crystallogr 38: 555–558.
27. Krigbaum WR, Kugler FR (1970) Molecular conformation of egg-white lysozyme and bovine alpha-lactalbumin in solution. Biochemistry 9:1216–1223.
28. Orthaber D, Bergmann A, Glatter O (2000) SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. J Appl Crystallogr 33:218–225.
29. Voss NR, Gerstein M (2005) Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly. J Mol Biol 346:477–492.
30. Rowland SL, Burkholder WF, Cunningham KA, Maciejewski MW, Grossman AD, King GF (2004) Structure and mechanism of action of Sda, an inhibitor of the histidine kinases that regulate initiation of sporulation in *Bacillus subtilis*. Mol Cell 13:689–701.
31. Whitten AE, Jacques DA, Hammouda B, Hanley T, King GF, Guss JM, Trewhella J, Langley DB (2007) The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition. J Mol Biol 368:407–420.
32. Jacques DA, Streamer M, Rowland SL, King GF, Guss JM, Trewhella J, Langley DB (2009) Structure of the sporulation histidine kinase inhibitor Sda from *Bacillus subtilis* and insights into its solution state. Acta Crystallogr D Biol Crystallogr 65:574–581.
33. Brunger AT (1992) Free R-value—a novel statistical quantity forassessing the accuracy of crystal-structures. Nature 355:472–475.
34. Svergun D, Barberato C, Koch MHJ (1995) CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 28:768–773.
35. Svergun DI, Richard S, Koch MHJ, Sayers Z, Kuprin S, Zaccai G (1998) Protein hydration in solution: experimental observation by X-ray and neutron scattering. Proc Natl Acad Sci USA 95:2267–2272.
36. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201.
37. Vestergaard B, Sanyal S, Roessle M, Mora L, Buckingham RH, Kastrup JS, Gajhede M, Svergun DI, 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.
38. 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.
39. 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.
40. Grishaev A, Tugarinov V, Kay LE, Trewhella J, Bax A (2008) Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. J Biomol NMR 40:95–106.
41. Grishaev A, Ying J, Canny MD, Pardi A, Bax A (2008) Solution structure of tRNA(Val) from refinement of homology model against residual dipolar coupling and SAXS data. J Biomol NMR 42:99–109.
42. Shin DS, Didonato M, Barondeau DP, Hura GL, Hitomi C, Berglund JA, Getzoff ED, Cary SC, Tainer JA (2009) Superoxide dismutase from the eukaryotic thermophile *Alvinella pompejana*: structures, stability, mechanism, and insights into amyotrophic lateral sclerosis. J Mol Biol 385:1534–1555.
43. Min X, Akella R, He H, Humphreys JM, Tsutakawa SE, Lee SJ, Tainer JA, Cobb MH, Goldsmith EJ (2009) The structure of the MAP2K MEK6 reveals an autoinhibitory dimer. Structure 17:96–104.
44. Heidorn DB, Trewhella J (1988) Comparison of the crystal and solution structures of calmodulin and troponin-C. Biochemistry 27:909–915.
45. Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76:2879–2886.
46. Franke D, Svergun DI (2009) DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. J Appl Crystallogr 42:342–346.
47. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. J Appl Crystallogr 36:860–864.
48. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. J Appl Crystallogr 34:33–41.
49. Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys J 89:1237–1250.
50. Petoukhov MV, Monie TP, Allain FHT, Matthews S, Curry S, Svergun DI (2006) Conformation of polypyrimidine tract binding protein in solution. Structure 14: 1021–1027.
51. Jeffries CM, Whitten AE, Harris SP, Trewhella J (2008) Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. J Mol Biol 377:1186–1199.
52. Whitten AE, Jeffries CM, Harris SP, Trewhella J (2008) Cardiac myosin-binding protein C decorates F-actin: implications for cardiac function. Proc Natl Acad Sci USA 105:18360–18365.
53. Graebsch A, Roche S, Niessing D (2009) X-ray structure of Pur-alpha reveals a Whirly-like fold and an unusual nucleic-acid binding surface. Proc Natl Acad Sci USA 106:18521–18526.
54. Heikkinen OK, Ruskamo S, Konarev PV, Svergun DI, Iivanainen T, Heikkinen SM, Permi P, Koskela H, Kilpelainen I, Ylanne J (2009) Atomic structures of two novel immunoglobulin-like domain pairs in the actin cross-linking protein filamin. J Biol Chem 284: 25450–25458.

55. Heidorn DB, Seeger PA, Rokop SE, Blumenthal DK, Means AR, Crespi H, Trewhella J (1989) Changes in the structure of calmodulin induced by a peptide based on the calmodulin-binding domain of myosin light chain kinase. Biochemistry 28:6757–6764.
56. Pascal JM, Tsodikov OV, Hura GL, Song W, Cotner EA, Classen S, Tomkinson AE, Tainer JA, Ellenberger T (2006) A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. Mol Cell 24:279–291.
57. Johs A, Hammel M, Waldner I, May RP, Laggner P, Prassl R (2006) Modular structure of solubilized human apolipoprotein B-100—low resolution model revealed by small angle neutron scattering. J Biol Chem 281: 19732–19739.
58. Comoletti D, Grishaev A, Whitten AE, Tsigelny I, Taylor P, Trewhella J (2007) Synaptic arrangement of the neuroligin/β-neurexin complex revealed by X-ray and neutron scattering. Structure 15:693–705.
59. Chen X, Liu H, Shim AHR, Focia PJ, He X (2008) Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions. Nat Struct Mol Biol 15:50–56.
60. Jacques DA, Langley DB, Jeffries CM, Cunningham KA, Burkholder WF, Guss JM, Trewhella J (2008) Histidine kinase regulation by a cyclophilin-like inhibitor. J Mol Biol 384:422–435.
61. Cunningham KA, Burkholder WF (2009) The histidine kinase inhibitor Sda binds near the site of autophosphorylation and may sterically hinder autophosphorylation and phosphotransfer to Spo0F. Mol Microbiol 71: 659–677.
62. Bick MJ, Lamour V, Rajashankar KR, Gordiyenko Y, Robinson CV, Darst SA (2009) How to switch off a histidine kinase: crystal structure of *Geobacillus stearothermophilus* KinB with the inhibitor Sda. J Mol Biol 386:163–177.
63. Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129:5656–5664.
64. Pelikan M, Hura GL, Hammel M (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen Physiol Biophys 28: 174–189.
65. Lemaire PA, Tessmer I, Craig R, Erie DA, Cole JL (2006) Unactivated PKR exists in an open conformation capable of binding nucleotides. Biochemistry 45: 9074–9084.
66. Schiefner A, Fujio M, Wu D, Wong CH, Wilson IA (2009) Structural evaluation of potent NKT cell agonists: implications for design of novel stimulatory ligands. J Mol Biol 394:71–82.