
What Can We Learn from Wide-Angle Solution Scattering?

8

Yujing Wang, Hao Zhou, Emre Onuk, John Badger,
and Lee Makowski

Abstract

Extending collection of x-ray solution scattering data into the wide-angle regime (WAXS) can provide information not readily extracted from small angle (SAXS) data. It is possible to accurately predict WAXS scattering on the basis of atomic coordinate sets and thus use it as a means of testing molecular models constructed on the basis of crystallography, molecular dynamics (MD), cryo-electron microscopy or *ab initio* modeling. WAXS data may provide insights into the secondary, tertiary and quaternary structural organization of macromolecules. It can provide information on protein folding and unfolding beyond that attainable from SAXS data. It is particularly sensitive to structural fluctuations in macromolecules and can be used to generate information about the conformational make up of ensembles of structures co-existing in solution. Novel approaches to modeling of structural fluctuations can provide information on the spatial extent of large-scale structural fluctuations that are difficult to obtain by other means. Direct comparison with the results of MD simulations are becoming possible. Because it is particularly sensitive to small changes in structure and flexibility it provides unique capabilities for the screening of ligand libraries for detection of functional interactions. WAXS thereby provides an important extension of SAXS that can generate structural and dynamic information complementary to that obtainable by other biophysical techniques.

Keywords

WAXS • SAXS • Protein structure • Structural fluctuations • Molecular dynamics • Moderate throughput screening

Y. Wang • H. Zhou • E. Onuk • L. Makowski (✉)
Northeastern University, Boston, MA 02115, USA
e-mail: l.makowski@northeastern.edu

J. Badger
DeltaG Technologies, San Diego, CA, USA

8.1 Introduction

X-ray solution scattering studies of proteins produce data that can provide substantial insight into protein structure, flexibility and dynamics. Small angle scattering (SAXS) provides relatively low resolution information (~ 20 Å), whereas wide-angle scattering (WAXS) provides information about higher resolution features or motions. Scattering intensity in the WAXS regime is typically two or more orders of magnitude weaker than in the SAXS regime. However, over the past two decades, the development of SAXS/WAXS beam lines at high brilliance synchrotron sources has fostered rapid growth of solution scattering studies extending to wide angles (e.g. Allaire and Yang 2011; Fischetti et al. 2004a). The WAXS regime also extends to scattering angles where background scattering from buffer and sample chamber is considerable. Use of a synchrotron source can provide data of a quality that can overcome these challenges.

In both the SAXS and WAXS regimes, the scattered intensity distribution can be more valuable when combined with structural information generated from other techniques. Whereas x-ray crystallography and NMR produce high resolution ‘snap shots’ of protein structure and information about local motions, solution scattering has the capability of providing information about conformational changes, intermolecular interactions, large scale structural fluctuations, and slow, concerted, global motions. WAXS is particularly effective for study of large-scale motions that are difficult to characterize with other approaches. Like SAXS, it can be used to study virtually any macromolecule or molecular assembly that can be purified at concentrations of ~ 1 to 5 mg/ml. Its value is significantly enhanced when used in concert with crystallographic or NMR approaches, computational modeling and/or molecular dynamics simulations.

For purposes of this chapter, we will define the boundary between the SAXS regime and the WAXS regime as 20 Å spacing ($1/d = 0.05 \text{ \AA}^{-1} = (1/20) \text{ \AA}^{-1}$, or $q = 0.3 \text{ \AA}^{-1}$) where $q = 4\pi \sin(\theta)/\lambda = 2\pi/d$ and 2θ is the angle

between incident and scattered x-rays. We choose this boundary because beyond 20 Å spacing internal fluctuations in electron density of a protein begin to contribute substantially to scattering. This distinction alters the nature of analyses possible in the two regimes. SAXS has been utilized for decades (Luzzati and Tardieu 1980) to estimate the radius of gyration (R_g), pair-distance distribution function, $P(r)$, and oligomerization state of proteins (Putnam et al. 2007). The capability to generate three-dimensional molecular shapes directly from SAXS data (Svergun 1999; Walther et al. 2000) has dramatically increased the utility and utilization of these methods. Three-dimensional reconstruction from solution scattering data is limited to about 20 Å resolution. Beyond 20 Å resolution, solution scattering data from any protein will be consistent with multiple molecular shapes in part because of the contribution of internal structures to the observed scattering and in part due to the intrinsic limitation of information content in the measured intensities (more about this later). Thus, SAXS data can be used directly to calculate R_g and a three-dimensional shape reconstruction. The pair-distribution function, $P(r)$, can be calculated from data extending to any resolution. However, for virtually all other applications, WAXS intensities are used to test hypotheses or molecular models generated by other means such as crystallography, NMR, molecular dynamics (MD) or *ab initio* modeling. The use of WAXS for testing of molecular models of structure and dynamics is directly dependent on our ability to accurately predict WAXS data from atomic coordinate sets (Park et al. 2009).

For testing of models, WAXS may provide an advantage over SAXS data in detection of relatively small structural changes. WAXS intensities are highly sensitive to small structural changes (Fischetti et al. 2004b) and to changes in the magnitude of structural fluctuations (Makowski et al. 2011). On functional binding of a ligand, a protein may alter either its structure, its dynamics or both, and it may be a mute point to argue whether the structure or dynamics have changed. Strictly speaking it is virtually impossible to alter one without the other, so

perhaps it is most appropriate to simply state that the structural ensemble has been altered. More to the point, any interaction that alters function is almost certain to trigger a change in structure and/or dynamics – and those changes will, in many cases be detectable using WAXS.

Solution scattering methods have evolved and matured over the past decade into a suite of highly informative probes of protein structure and activity that go well beyond a simple method for determining size and shape of the molecule. They now represent an approach to detailed characterization of biochemistry in the scattering volume. As such, increased focus must be given to the state of the sample. It is critical that the sample be well defined biochemically, absent precipitates. Wide-spread adaptation of SEC-SAXS/WAXS in which scattering patterns are collected from the output of a size-exclusion column reflects this trend. In conventional, static WAXS, background scattering from a precisely matched buffer is a critical aspect of any experiment as scattering contributions from even minor buffer constituents can be important. This is even more important in WAXS than in SAXS, because below $q \sim 0.3 \text{ \AA}^{-1}$ there is little scattering from buffer or sample chamber, whereas at higher angles these contributions exceed scattering from the protein.

Although early studies were limited to structurally homogeneous samples, solution scattering is now frequently used to study the ensemble of structural forms present in solution including, for instance, enzymes undergoing catalytic cycling (Onuk et al. 2015). WAXS can be used to generate information about changes in secondary, tertiary and quaternary structures (Doniach 2001; Hirai et al. 2002; Makowski et al. 2008a); conformational changes due to ligand binding (Fischetti et al. 2004b; Rodi et al. 2007; Zhou et al. 2015), or cofactor oxidation state (Tiede et al. 2002) and by amino acid substitutions (Makowski et al. 2011; Zhou et al. 2015), or protein folding (Hirai et al. 2004). The conformational ensemble of a protein in solution can also be studied with WAXS. WAXS has proven to be highly sensitive to changes in the ensemble due to protein concentration

(Makowski et al. 2008b), mutations and ligand binding (Makowski et al. 2011; Zhou et al. 2015; Onuk et al. 2015). Time resolved (TR) studies can be carried out analogous to static studies and have been used to characterize light-triggered conformational changes occurring in nano- to milli-seconds (Cammarata et al. 2008; Cho et al. 2010). Although we will not explicitly consider TR studies in this chapter, all methods described can be applied to each diffraction pattern ‘snapshot’ of a TR data set.

8.2 WAXS Data

Collection of WAXS data simultaneously with SAXS data is challenging, even at state of the art beam lines (Zhang et al. 2000; Makowski 2010). It can be accomplished with a very small beam stop and large detector; or by using two detectors set at different sample-to-detector distances, the WAXS detector subtending only a portion of the wide angle region, but capturing enough intensity to provide good signal-to-noise ratio after merging with the SAXS data (e.g., Allaire and Yang 2011). Ideally, one would like to set a WAXS detector on axis and at relatively small sample-to-detector distance, but including a slot-shaped hole to allow passage of x-rays to a SAXS detector placed at a much higher sample-to-detector distance. Choosing a slot-shaped hole would generate a q -range in which data was collected at both detectors, providing adequate overlap for accurate scaling of data from the two detectors. This arrangement has not, as of yet, been implemented.

The scattered intensity, $I(q)$, from a protein solution can be calculated, in principle, from the position of all atoms in the protein using the Debye formula,

$$I(q) = \sum_{i=1}^n \sum_{j=1}^n f_i f_j \frac{\sin 2\pi q r_{ij}}{2\pi q r_{ij}} \quad (8.1)$$

where f_i is the scattering factor from the i th atom, and r_{ij} is the distance between atom i and atom j . As will be discussed below, direct application of this formula fails when the protein is

immersed in aqueous solution since this necessitates taking into account the impact of the shape of the region excluding solvent and the difference in water structure and density between hydration layer and bulk.

Whereas SAXS data to be used for calculation or R_g , $P(r)$, or three-dimensional shape reconstructions requires measurements to small angles dictated by the maximum spatial extent of the scattering object (see other chapters for details), data used for testing of models does not necessarily need to extend to small angles. However, collection of SAXS data simultaneously with WAXS data provides important quality assurance tests for detection of the presence of aggregates or inter-particle interference effects which may be observed at higher concentrations. Interparticle interference effects, more likely as protein concentration increases, are usually limited to the small angle regime. When present, they can distort estimates of R_g , $P(r)$, or three-dimensional shape reconstructions (see e.g. Inouye et al. 2016). They can usually be detected by comparing SAXS data collected at two or more protein concentrations. At high protein concentrations (say, >10 mg/ml) intermolecular crowding can suppress structural fluctuations in some proteins, resulting in a

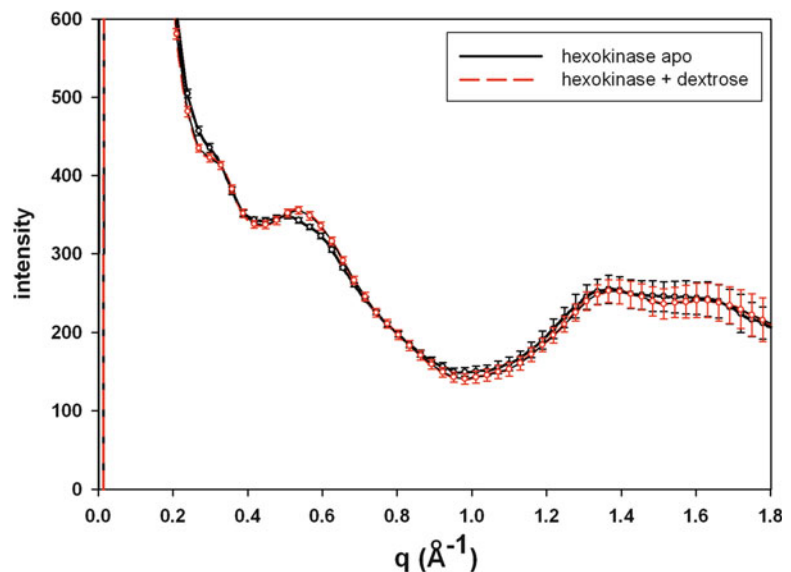
sharpening of wide-angle scattering features (Makowski et al. 2008b). More rigid proteins exhibit little reaction to changes in concentration. Amorphous aggregates (to be distinguished from multimers) can result in a sharp spike in scattering at very small angles but usually exhibit little wide-angle scattering except for potentially resulting in a small increase in diffuse background.

Figure 8.1 is an example of the impact of ligand binding on the WAXS scattering from a protein. Binding of substrate to hexokinase results in a relatively large conformational change – closing of the binding site cleft (McDonald et al. 1979). This alters the small angle scattering from the molecule, lowers R_g and induces additional intensity changes in the wide angle regime.

8.3 Predicting WAXS Data from Atomic Coordinates

The ability to accurately predict WAXS data from atomic coordinate sets is key to the utility of WAXS, making it a sensitive method of assessing the accuracy of atomic-scale models. If proteins existed in a vacuum, calculation of

Fig. 8.1 WAXS scattering from hexokinase in the presence and absence of substrate binding. Binding results in a closure of the ligand binding cleft altering R_g , as well as observed intensities in the SAXS and WAXS regimes. Error estimates increase at wide angles because of the increased intensity of buffer scatter in that regime



solution scattering would reduce to a simple application of the Debye formula (Eq. 8.1). However, proteins (or other macromolecules) are generally immersed in solvent making it essential to account for the exclusion of water in the volume occupied by the protein. One also has to model the hydration layer where the water takes on a density that may be as much as 10% greater than in bulk (Svergun et al. 1997). These effects were first taken into account in the iconic program CRY SOL (Svergun et al. 1995) that has transformed the use of SAXS for protein studies. In the WAXS regime, however, the approximations used in CRY SOL break down. In particular, CRY SOL underestimates the intensity of WAXS data by a factor of 2-3X relative to SAXS intensity when used with default parameters. This is due to the continuum representation of the hydration layer and method for representing excluded volume in CRY SOL (Bardhan et al. 2009). In the WAXS regime it is essential to utilize an explicit atom representation of water (Bardhan et al. 2009; Park et al. 2009; Grishaev et al. 2010). For precise modeling of intensity in the WAXS regime, CRY SOL may not be the most appropriate. Although CRY SOL refinement against experimental data often results in good agreement between calculated and observed, this may come as the result of non-physical values for adjustable parameters within CRY SOL (Barhan et al. 2009). Extensive experimental (Svergun et al. 1997) and computational tests indicate that the density of water in the hydration layer may be as much as 10% greater than bulk water, an amount detectable with solution scattering, and that these structural differences extend roughly 7 Å beyond the protein surface (Park et al. 2009). Once these issues are taken into account it is possible to calculate scattered intensities to within experimental error for most rigid proteins across both the SAXS and WAXS regimes (Park et al. 2009; Grishaev et al. 2010). This requires, however, MD simulation of the water in the hydration layer, a process that remains computationally laborious. Consequently, these calculations are not yet high

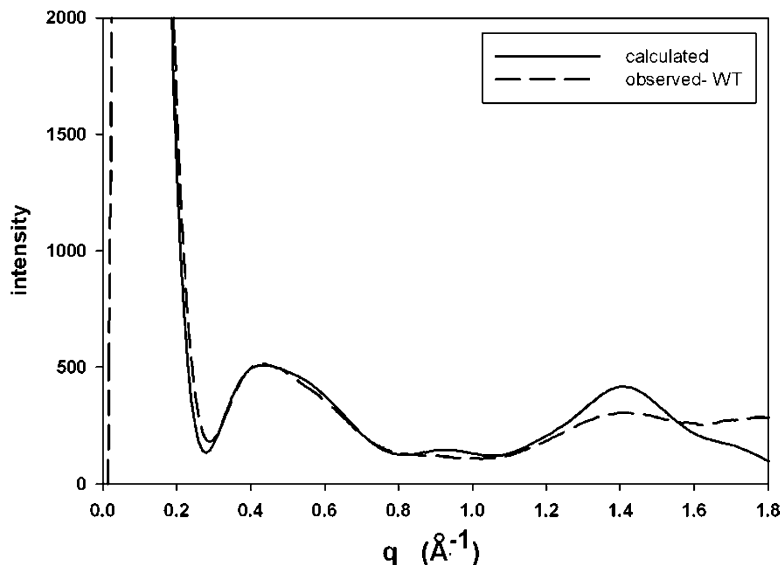
throughput and the capability of carrying them out for large ensembles of representative structures has not yet been established. For this reason, CRY SOL remains the most widely used program for estimation of solution scattering from atomic coordinates.

Computational estimates of scattering intensity presented in this chapter utilize the software package XS as described (although not named) by Park et al. (2009). In XS, water molecules are positioned around a protein surface out to ~7 Å from the protein surface, and subject to 100 ps of MD simulation during which the protein atoms are held rigid. A ‘snapshot’ of the water positions is captured once each picosecond and WAXS intensity due to the protein plus water positions in each snap shot are calculated using the Debye formula and then averaged giving I_{prot} . Simulation of a ‘droplet’ of bulk water the same shape as the protein-containing droplet (including the 7 Å -thick hydration layer) is also carried out and snapshots of this droplet are used with the Debye formula to calculate WAXS patterns that are subsequently averaged to produce I_{water} . Subtraction of this bulk water intensity from the hydrated protein intensity results in approximation of ‘excess intensity’, I_{xs} . This excess intensity corresponds closely to the difference between scattering observed from protein-solution-filled and buffer-filled sample chambers,

$$I_{\text{xs}} = I_{\text{prot}} - I_{\text{water}} \approx I_{\text{obs}} - I_{\text{buffer}} \quad (8.2)$$

Out to ~5 Å spacing, excess intensity is virtually identical to I_{prot} . At wider angles, scattering from buffer is non-negligible. At these angles I_{xs} differs from the more routinely calculated I_{prot} . Since protein usually occupies <1% of the total scattering volume, beyond 5 Å spacing ($q \sim 1.2 \text{ \AA}^{-1}$) the scattering from buffer is far more intense than that from the protein, and I_{xs} will be negative. Thus, the moniker is ‘excess intensity’ rather than ‘intensity’ which is universally considered a positive number. Figure 8.2 is a comparison of the calculated and observed WAXS from ubiquitin. Intensity calculated using XS (Park et al. 2009) with no free parameters

Fig. 8.2 Comparison of WAXS intensity distribution, I_{prot} , from ubiquitin as calculated by the software package XS (solid line) and as observed in scattering from a 10 mg/ml solution (broken line)



results in an intensity distribution indistinguishable from observed out to a $q \sim 1.2 \text{ \AA}^{-1}$. In the region $1.2 < q < 1.6 \text{ \AA}^{-1}$ the calculated intensity is greater than observed. This region corresponds to a spacing of $\sim 4.7 \text{ \AA}$ and is generated largely from the inter-strand spacings of beta strands in the molecule. The comparison suggests that the strands are undergoing small structural fluctuations, leading to observed intensity somewhat lower than that calculated for a rigid molecule.

8.4 Size and Shape

The size and shape of a protein, other macromolecule or macromolecular complex can usually be determined from SAXS data. This is a topic well covered in other chapters of this book. Although WAXS data extends to much higher resolutions (scattering angles) it cannot be used to improve the accuracy of a radius of gyration or to enhance the level of detail in three-dimensional reconstructions of molecular shape. It is worth discussing the origins of these limitations.

The radius of gyration, R_g , literally, the average radius of scattering density from the center of mass can be estimated from data in the q -range

where the Guinier approximation is valid ($qR_g < 1.3$ for most globular proteins). Intensities at higher scattering angles do not improve the estimate of R_g because the Guinier plot is not, in general, linear at wider scattering angles. In fact, extending data to smaller angles is usually more important for accuracy of the estimate of R_g than extending to higher angles. The arrangement of detector and beam stop required for WAXS data may place limits on minimum scattering angle at which data is collected. A hybrid SAXS/WAXS detector scheme tuned for collection of both simultaneously is used at a number of beam lines to overcome this problem.

SAXS data can also be used to reconstruct a three-dimensional shape of a macromolecule (Chacon et al. 1998; Svergun 1999; Walther et al. 2000; Svergun et al. 2001; Hura et al. 2009; other chapters in this book). The algorithms used to generate shape reconstructions from SAXS data implicitly assume the scattering density within the protein is roughly constant. For proteins, this is approximately true to $\sim 20 \text{ \AA}$ resolution, but no higher. At spacings greater than $(1/d) \sim 1/20 \text{ \AA}^{-1}$ ($q \sim 0.3 \text{ \AA}^{-1}$), intensities are strongly influenced by internal structural features and extending data used to higher q may result in spurious features (although inclusion of

higher angle data appears to stabilize some of the algorithms used for three-dimensional reconstructions without generating artefactual features). Combined SAXS-WAXS data should not be used for *ab initio* shape computation due to the breakdown of uniform scattering density model in the WAXS regime. It can, however, be used to test structural hypotheses. For instance, it can be used for evaluating the quality of rigid-body models derived from crystallographic structural information (Svergun et al. 2001; Zheng and Tekpinar 2011; Wen et al. 2014) or for refining the positions of (rigid body) domains. This has particular application to multi-domain proteins that may undergo large scale re-arrangements of domains in response to allosteric effectors or other interactions (Badger et al. 2016). Although most current studies utilize SAXS data, extending the approaches to WAXS has the potential to improve accuracy.

Validation of modeling efforts is not necessarily straightforward, there is potential for multiple solutions, and the calculation of uncertainty in optimized domain positions, while possible, has not usually been reported in published studies. At the very least, the use of WAXS for rigid body refinement of domain positions will produce testable hypotheses about the functional significance of domain movements.

Resolution of shape reconstructions is also limited by uniqueness (Volkova and Svergun 2003). The amount of information required for a three-dimensional reconstruction goes up roughly as q^3 . The amount of information in a solution scattering pattern goes up proportional to q . At some limiting q value, the amount of information required for unique shape determination will exceed that contained in the scattering pattern. Another way of conceptualizing this is by considering a molecular shape as a sum of spherical harmonics (Lattman 1989). At very small angles, only a small number of spherical harmonics contribute to the observed intensities. At increasing scattering angle, increasing numbers of spherical harmonics contribute. The capability of three-dimensional reconstruction is only made possible by the oversampling of the continuous (spherically averaged) intensity

distribution. At some limiting q , the amount of information required to estimate the intensity associated with all contributing spherical harmonics is greater than the amount of information within the pattern. At that point, estimation of the three-dimensional shape becomes an ill-posed problem with multiple solutions.

WAXS data can contribute to the accuracy and resolution of $P(r)$, increasing the level of detail contained in it to resolutions beyond those of the SAXS regime (Hong and Hao 2009). Among other things, this may make possible a more accurate estimate of the longest interatomic vector lengths in the protein. Intensity in a WAXS pattern is a band-limited function with the band-limit equal to the length of the longest interatomic vector in the protein, D_{\max} . Larger proteins exhibit scattering patterns with sharper features (e.g. peaks and troughs) because the patterns include more higher frequency terms – corresponding to the longest interatomic vectors (i.e., patterns from larger proteins have larger band pass). An estimate of D_{\max} can be made from the pair-distribution function, $P(r)$. Nevertheless, since the longest interatomic vectors contribute very little to the measured intensity it is often challenging to make an accurate estimate of D_{\max} . Iterative procedures may be required (e.g. Putman et al. 2007). WAXS can provide improved accuracy for $P(r)$ and consequently, D_{\max} . An accurate estimate of D_{\max} contributes to more accurate three-dimensional shape reconstructions since most algorithms require it as input. Validation of SAXS-derived structures is often difficult and, as in many biophysical approaches depends to some extent on self-consistency and consistency of models with all available data. Because of the very well defined relationship between atomic coordinates and WAXS data (Eq. 8.1), WAXS can provide a very well defined test of models constructed on the basis of multiple biophysical probes.

8.5 Secondary Structure

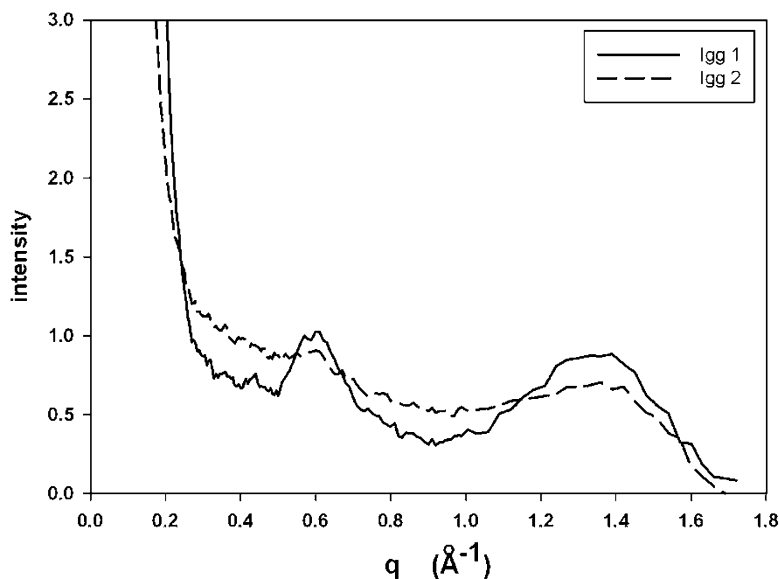
The Debye formula – Eq. 8.1 – demonstrates that solution scattering is due entirely to the

distribution of interatomic vector lengths within a sample. Secondary structures, by definition, have strong patterns of interatomic vector lengths, so we would expect them to contribute to solution scattering in distinctive ways. α -helices, for instance, pack roughly 10 Å apart and, not surprisingly, α -helical proteins have a considerable number of interatomic vectors about 10 Å in length. This usually results in a strong scattering peak at a spacing of $\sim(1/10)$ Å⁻¹ ($q \sim 0.6$ Å⁻¹). Analogously, β -sheets may also lie about 10 Å from one another, face-to-face, and may also exhibit relatively intense scattering in the 10 Å region. Furthermore, they are made up from β -strands that typically lie ~ 4.7 Å apart. This results in solution scattering patterns with a peak at a spacing of $\sim(1/4.7)$ Å⁻¹ ($q \sim 1.3$ Å⁻¹). Strong scattering in the 10 Å and 4.7 Å regions can be observed in WAXS patterns from ubiquitin as seen in Fig. 8.2. Similarly, Fig. 8.3 includes scattering from two I_{gg} molecules, one showing well-defined, strong peaks at ~ 10 Å and 4.7 Å spacing, and a second I_{gg} that, due to significant conformational flexibility, exhibits only modest peaks in these regions, an example of the impact of fluctuations on WAXS data.

8.6 Tertiary Structure

As of yet there has been no experimental demonstration that WAXS data could be used to generate information about protein tertiary structure. However, a quantitative analysis of the information embedded within a WAXS pattern was used to demonstrate that this may be possible, in principle, if accurate intensities can be measured to ~ 2.0 Å spacing (Makowski et al. 2008a). WAXS patterns computed from atomic coordinates of 498 protein domains corresponding to the known fold space at that time (Hou et al. 2003) were used to construct a multi-dimensional space of WAXS patterns ('WAXS space') corresponding to these folds. Within WAXS space, each scattering pattern is represented by a single vector. A principal components analysis (PCA) identified directions in WAXS space corresponding to the greatest discrimination among WAXS patterns. Estimates of the abundances of secondary structures were made based on training sets derived from these data. This analysis led to estimates of α -helical content with average error of 11%; and of β -sheet content with average error of $\sim 9\%$. The distribution of proteins that are members of the four global structure classes, α , β , α/β and $\alpha+\beta$, are well

Fig. 8.3 WAXS patterns from two I_{gg} molecules, I_{gg1} exhibiting strong WAXS scattering in the ~ 10 Å and 4.7 Å regions, consistent with a relatively rigid, well-formed immunoglobulin domain structure, and I_{gg2} with muted intensities in the ~ 10 Å and 4.7 Å regions ($q \sim 0.6$ and 1.35 Å⁻¹, respectively), suggestive of a structural heterogeneity derived from relatively great conformational flexibility in solution



separated in WAXS space when data extending to a spacing of 2.2 Å are used, indicating that production of highly accurate WAXS data to high resolution has the potential for producing significant information on the structural class of any protein. By contrast, data limited to ~10 Å spacing exhibits little discriminatory power for classifying proteins according to secondary or tertiary structures.

8.7 Allosteric Proteins, Domain Organization and Quaternary Structure

One of the most promising areas for x-ray solution scattering is in the study of allosteric proteins, typically multi-domain and/or multi-subunit proteins that exhibit large-scale domain motions either as part of their function or in response to allosteric effectors or regulators. These re-arrangements of domains are difficult to study by crystallography because they typically involve movements that cannot be accommodated within a crystal lattice. Their characterization may require a new search for crystallization conditions for each allosteric effector studied or each structural configuration of functional importance. By contrast, domain motion results in large changes in solution scattering often within both the SAXS and WAXS regimes. For instance, Yang et al. (2010) studied the impact of peptide ligands and amino acid substitutions on the ensemble of structures exhibited by hck tyrosine kinase, characterizing large-scale re-arrangements of SH3, SH2 and kinase domain in response to different solution conditions. Badger et al. (2016) demonstrated the power of the approach by characterizing the re-arrangements of domains in abl kinase in response to amino acid substitutions that altered the activity of the protein. Their characterization of the T315I gatekeeper mutation (that exhibits resistance to all known drugs that target bcr-abl) revealed a novel configuration of the three domains of the abl core not previously characterized, suggesting the existence of multiple levels of regulation of abl kinase activity.

8.8 Ensembles

Protein solutions are not, in general, solutions of perfectly homogeneous macromolecular structures diffusing in an ideal buffer. Yes, it is possible to find small, relatively rigid proteins that will approach this ideal. But, these are not the most interesting cases. Much more frequently, proteins of interest may be large, flexible and capable of global internal motions of functional importance. Computational approaches to be used in concert with WAXS studies are developing rapidly and they may represent one of the most important applications of WAXS since they address issues difficult to resolve by other methods. It may seem counter intuitive that a single, one-dimensional intensity distribution could provide information about the relative abundances of multiple conformations within a solution. But if the structures of conformations that may be present can be hypothesized – on the basis of crystallographic, modeling or other information – then WAXS data represents a powerful test bed for determining which structures are, in fact, present and in what proportions (Konarev et al. 2003; Bernado et al. 2007; Tsutakawa et al. 2007; Yang et al. 2010; Petoukhov and Svergun 2007; Minh and Makowski 2013; Onuk et al. 2015). There are, of course, limits. In general, WAXS data seem capable of distinguishing relative abundances of three to ten distinct conformations. Contributions from conformations that are similar are more difficult to separate; dramatic structural differences far easier to distinguish.

OLIGOMER (Konarev et al. 2003) was originally conceived to separate out scattering from monomers; dimers and higher order oligomers when present together in a mixture. It has, however, found broader utility. It estimates the relative abundances of multiple constituents by solving a set of linear equations using nonnegative or unconstrained least-squares to minimize the difference between experimental and calculated scattering. It appears to adapt to WAXS data under conditions where the CRY SOL provides accurate estimates of scattered intensities (Onuk et al. 2015). Yang

et al. (2010) introduced basis-set supported SAXS (BSS-SAXS) reconstruction, that combined solution scattering data with coarse-grained (CG) molecular dynamics to characterize the conformational states of Hck kinase in solution. In this approach, CG-MD simulations explore and sample conformational space; captured conformations are clustered into nine distinct conformational states and then used these as a basis set to analyze the scattering data. Onuk et al. (2015) took a somewhat different approach, using crystallographically determined structures of adenylate kinase as an initial basis set, clustering them into five distinct conformational classes and then used a maximum likelihood estimation (MLE) approach to generate estimates of relative abundances of these classes. It operates similar to OLIGOMER, but out-performs OLIGOMER when used with data having relatively low signal-to-noise ratio due to an accurate noise model (Onuk et al. 2015).

8.9 Priors

Addition of prior knowledge can greatly improve the accuracy or power of a calculation designed to characterize an ensemble of structures. That said, approaches to incorporation of priors are often non-trivial. Onuk et al. (2016) used a maximum a posteriori (MAP) approach to estimate the relative abundances of conformations in solutions of adenylate kinase. This enables estimates of the relative free energies of different conformations to be used to provide weights in the estimation of their relative abundances in solution (e.g., conformations with higher free energy are assigned lower weights than those with lower free energy). Computational tests indicated that prior knowledge improves estimation accuracy, and, not surprisingly, the stronger the prior constraints, the more accurate the resulting estimates of conformational abundances.

8.10 Modeling of Structural Fluctuations

It is not always convenient or informative to model structural fluctuations on the basis of an extensive ensemble of representative structures. This is particularly the case when proteins fluctuate about a single well defined conformation and characterization of the scattering can be made in terms of an average or consensus structure and the fluctuations about that structure. Increased flexibility leads to a broader structural ensemble that expresses itself in solution scattering patterns by filling in troughs in the scattered intensity and muting the intensity of peaks (Makowski et al. 2011). The range of motion of interatomic vectors can be estimated by comparison of the scattering pattern expected for a rigid protein with the observed scattering pattern (Makowski et al. 2011; Zhou et al. 2015). A formalism that makes it possible to predict the effect of these fluctuations on WAXS data has been developed and is called Vector Length Convolution (Makowski et al. 2011; Zhou et al. 2015). In this approach, the interatomic vector length of every atom pair in the protein is replaced by a distribution of vector lengths, and the breadth of that distribution is assumed to vary as a function of length. Not unreasonably, it has been found that small interatomic vectors exhibit smaller fluctuations than longer interatomic vectors (Zhou et al. 2016). Scattering from proteins undergoing this kind of fluctuation is predicted by (i) choosing a reference or consensus structure; (ii) calculating the scattering from the reference structure using XS; (iii) replacing each interatomic length in the pair correlation function, $P(r)$, of the reference structure by a distribution of vector lengths – which amounts to a generalized convolution (see below); (iv) and re-calculating the intensity from the altered $P(r)$. The resulting intensity function can then be compared with observed and parameters adjusted until a reasonable fit is achieved. Model ensembles with distinctly different properties can be generated by varying the way in which the fluctuations vary with interatomic vector length. The pair correlation function corresponding to

the model structural ensemble, $P_m(r)$, is computed from the convolution of the pair correlation function of the reference structure, $P_r(r)$, and a Gaussian of half width $\sigma(r)$ which may be a function of the interatomic vector length, r , according to

$$P_m(r) = P_r(r) * \exp\left(-\sigma(r)^2/2r^2\right)$$

The ‘*’ in the equation denotes convolution. Early applications of the method (Makowski et al. 2008b, 2011; Zhou et al. 2015) used a two parameter model for the radial variation of σ , $\sigma(r) = cr^e$, where c and e are free parameters, and varied the parameters to achieve an optimal fit to the observed data. More recently, $\sigma(r)$ has been calculated directly from MD trajectories and used, with the model $P_m(r)$ to predict scattering.

The impact of structural flexibility is to generate a heterogeneous ensemble of protein conformations. This ensemble can be modeled through vector length convolution of $P(r)$ to predict the impact of fluctuations on the WAXS pattern from a protein. Figure 8.4 includes the predicted scattering from ubiquitin assuming rigid conformation (solid line) and in the presence of fluctuations with a magnitude of

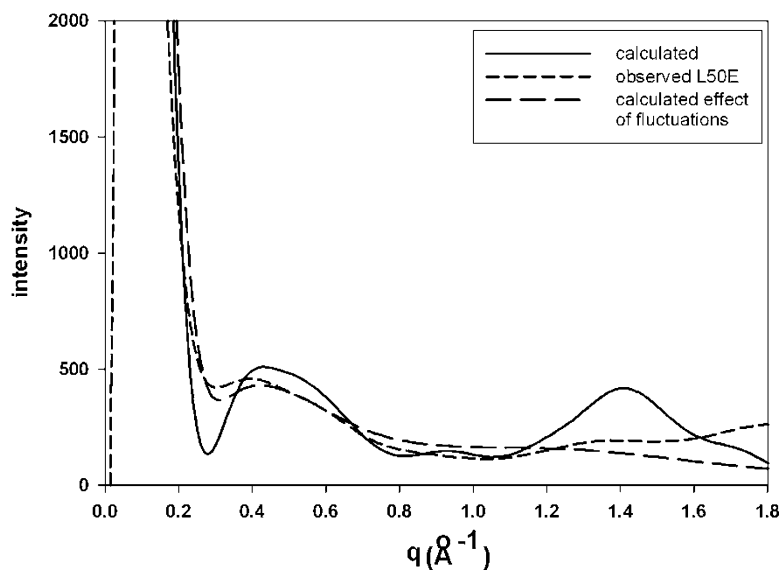
$\sigma(r) = 0.7 r^{0.5} \text{ \AA}$. These predictions are compared to the scattering from a mutant ubiquitin (L50E) in which a hydrophobic core residue is replaced by a charged residue, disrupting structure and leading to flexibility and heterogeneity. Scattering from L50E is consistent with structural fluctuations of nearly 20% in interatomic vectors 10 Å in length – corresponding to very substantial structural heterogeneity.

Note that the 4.7 Å peak, due to the arrangement of β -strands is essentially gone in scattering from the mutant, indicating the complete or nearly complete disruption of β -sheets in the structure.

8.11 Unfolding

Folding (and unfolding) of proteins in response to environmental changes results in significant alterations in WAXS scattering. During the alcohol-induced unfolding of β -lactoglobulin the largely β -structure has been reported to transform into an open α -helical structure (Hirota et al. 1997; Kumar et al. 2003). However, WAXS patterns from β -lactoglobulin in increasing concentrations of ethanol suggest the preservation of β -structure during the transformation. Figure 8.5 includes WAXS data from

Fig. 8.4 Predicted WAXS patterns for rigid ubiquitin (solid line) and a highly flexible ubiquitin molecule (long dashes) compared to observed scattering from a ubiquitin mutant (L50E) that is highly flexible (short dashes)



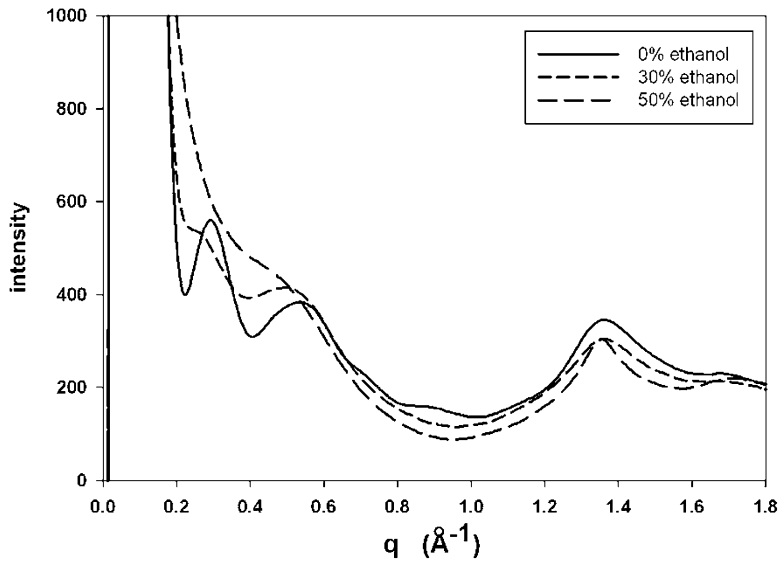


Fig. 8.5 Unfolding of β -lactoglobulin in ethanol. WAXS patterns from β -lactoglobulin in buffer; and in 30% and 50% ethanol. Strong features that correspond to tertiary structure begin to disappear in ethanol solutions $>12\%$, and are almost completely gone in 50% ethanol. However, the 4.7 \AA peak ($q \sim 1.35 \text{ \AA}^{-1}$) that corresponds to

β -strand separation remains strong in 50% ethanol, indicating the preservation of at least a part of the β -sheet structure, even in the near complete absence of tertiary structure

β -lactoglobulin in the absence of alcohol, in 30% and 50% alcohol. The scattering in 30% alcohol suggests the preservation of at least a semblance of tertiary structure since scattering features at in the $10\text{--}20 \text{ \AA}$ regime are not completely removed. In 50% alcohol, these features are gone, suggesting a complete obliteration of tertiary structure. Unexpectedly, the strong 4.7 \AA peak remains present in 50% alcohol, strongly suggesting that the β -sheets retain some structural integrity even in the virtual absence of tertiary structure.

8.12 Screening Ligand Libraries for Detection of Functional Interactions

Because modulation of function by a small molecule ligand is almost always accompanied by a structural change detectable by WAXS (Fischetti et al. 2003), WAXS is becoming a promising technology for screening of ligand libraries for

functional interactions. Target-based affinity screens may be used to screen libraries of up to 10^6 compounds, typically yielding $10^1\text{--}10^2$ candidate ligands (Stockwell 2000). Subsequent functionality tests of these candidate molecules represent a serious bottleneck in the drug discovery pipeline. *In vivo* screens may detect phenotypic changes due to ligand action but are complicated by potential for the ligand tested to bind to other targets or yield false negatives due to parallel pathways that duplicate assayed function. *In vitro* screens often require a custom, function-specific assay and these assays are not available for all functions. An alternative approach is to use a generic biophysical method to detect structural changes that almost universally accompany functional ligand binding. Unfortunately, many approaches have limited sensitivity to structural change. For instance, circular dichroism (CD) is largely sensitive to changes in secondary structure (Wallace and Janes 2003), and SAXS may be insensitive to changes that do not alter the radius of gyration

or result in large re-organization of domain structure. WAXS can be used to detect a broad range of ligand-induced alterations in secondary, tertiary, or quaternary structure. The speed of data acquisition, use of label-free targets, and adaptability to a broad range of solution conditions, make WAXS an attractive method for moderate-throughput detection and analysis of protein-ligand interactions.

How small a structural change can be detected using WAXS? Fischetti et al. (2004b) demonstrated detection of ligand binding in four proteins that had been crystallized both in the presence and absence of known ligands. Addition of ligands to transferrin, maltose binding protein (MBP), alcohol dehydrogenase (ADH) and calmodulin resulted in changes in WAXS patterns that corresponded to those predicted from atomic coordinate sets. The variation in structures triggered by these experiments ranged from ligand-induced re-folding in calmodulin, to ligand-induced domain rotation in transferrin, hinge-binding motion in MBP and change in the shape of the binding cleft in ADH. Figure 8.6 provides an example of the change in WAXS scattering for the binding of a small molecule ligand, 2,4,6 tribromophenol, to human transthyretin. This ligand is a common

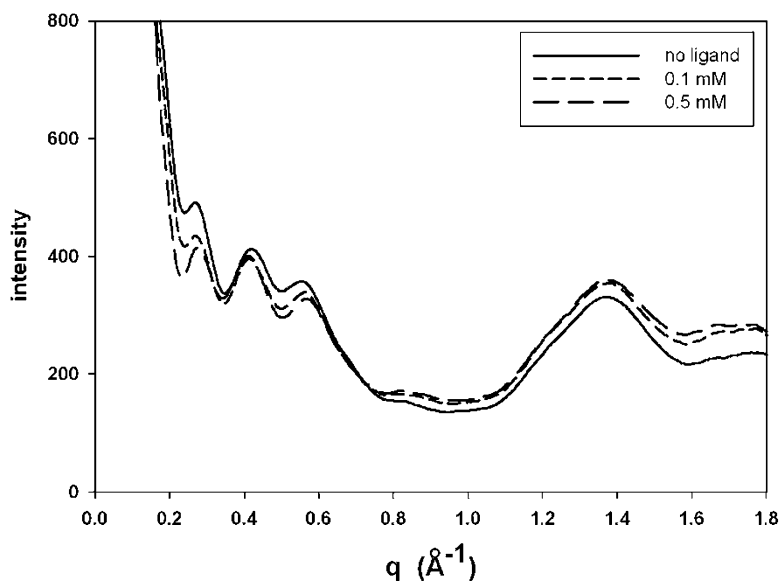
environmental contaminant that binds competitively with the natural ligand (Ghosh et al. 2000) and has been implicated in disruption of the thyroid hormone system.

It is also possible for ligand binding to alter the flexibility of a protein or the spatial extent of its structural fluctuations. These changes may also result in modulation of function. WAXS has proven unexpectedly sensitive to changes in structural fluctuations (Makowski et al. 2008b). For example, when an inhibitor is bound to HIV protease, the flaps may fold down over the inhibitor much as they do when binding substrate. However, detailed analysis of the ligand induced changes in intensity observed by WAXS indicated that the average structure does not change significantly. Rather there is a decrease in the magnitude of structural fluctuations that the protein is undergoing (Zhou et al. 2015).

8.13 Establishing the Significance of Small Intensity Changes

When ligands induce small changes in WAXS scattering that may or may not indicate a statistically significant change in structure, it is useful to have a statistical measure of the difference

Fig. 8.6 WAXS patterns from human transthyretin in solution, and in the presence of 0.1 and 0.5 mM of a 2,4,6 tribromophenol, a common environmental contaminant that competes with the natural ligand for binding to transthyretin



between two WAXS patterns. A chi-square measure has proven useful as a measure of statistical significance (Rodi et al. 2007). In our experience, a reduced chi-square (chi-square divided by number of degrees of freedom) $\chi_\nu > 1.0$ is indicative of a statistically significant difference between two WAXS patterns. The number of degrees of freedom is approximately equal to the number of independent measurements of intensity which is $\sim q_{\max}d_{\max}/\pi$, where q_{\max} is the greatest value of q for which data is used and d_{\max} is the longest interatomic vector in the structure (as estimated by a simple Shannon sampling theorem argument). However, some care needs to be taken in applying chi-square as a measure of significance of a structural change because estimates of standard deviation of intensities are difficult to make accurately and there remain questions about scaling of differences in the SAXS regime relative to differences in the WAXS regime due to the dramatic (two orders of magnitude) difference in their raw intensities.

Although it is quite easy to detect large domain motions with SAXS data, smaller motions may be detectable with WAXS. Investigation of the intensity changes generated by loop and side-chain re-arrangements (Fischetti et al. 2004b) suggested strongly that relatively minor movements can be detected with WAXS. Investigation of the impact of anomalous scattering on WAXS data (Makowski et al. 2012) suggested that differences corresponding to motion of even a few electrons can be detectable.

Establishing that two WAXS patterns are statistically distinguishable may be inadequate to address the biological question motivating the research. Comparison of scattering from several samples that represent impact of different ligands or amino acid replacements on the same molecule may require categorization or clustering of the patterns. Qualitative descriptions of differences may not elucidate the relationships among different structures or establish the structural origins of the differences observed. Use of a dimensionality reducing approach such as principal component analysis (PCA) to provide a quantitative classification of the patterns may

make possible identification of features that most distinguish the patterns from one another and generate a foundation for establishing patterns underlying functional modulation.

8.14 Discussion

These examples indicate that WAXS provides enhanced sensitivity for detection of small structural changes relative to SAXS; that it can be used to test molecular models for protein structure, and provide insights into protein flexibility both quantitatively and qualitatively.

The need for generic approaches to screening ligands for functional binding has motivated consideration of WAXS as a moderate-throughput screen. Most intermolecular interactions that give rise to significant changes in structure or dynamics will modulate the function of a protein in some way. Since WAXS is sensitive to changes in secondary, tertiary or quaternary structure or domain motions, it provides a comprehensive option for these kinds of screens. Given that it is now possible to screen 10–20 samples per hour at a synchrotron source, a screen of several hundred candidate ligands is quite feasible. As a secondary screen, focused on ligands that were originally identified by a high-throughput affinity screen, WAXS can provide an attractive addition to the drug discovery pipeline.

Although WAXS cannot be used to calculate a molecular structure (it lacks adequate information content), it can be used to test molecular models, whether generated *ab initio*, or based on crystallography, cryoEM, NMR, MD or combinations of these methods. Development of more efficient computational approaches to modeling of the hydration shell and excluded volume would contribute substantially to these calculations. This does not represent a significant bottleneck for individual calculations. However, the increased focus on ensembles, and the consequent need for calculating patterns from a large number of protein conformations as a basis for characterizing the ensembles, places high priority on improved computational tools. Recent

efforts to use solution scattering to refine structural models (Zheng and Tekpinar 2011; Roig-Solvas et al. 2017) would also benefit substantially from highly efficient computations taking into account the impact of hydration shell.

Characterization of flexibility is challenging for a number of reasons, not the least of which is the challenge of enumerating flexibility in a simple way. Utilizing the $P(r)$ function as the basis for global characterization of structural fluctuations is advantageous because (i) it is relatively intuitive (ii) can be displayed as a simple one-dimensional plot and (iii) results in estimation of the dependence of the scale of structural fluctuation on interatomic vector length, $\sigma(r)$, a function that can be estimated from WAXS data or be calculated directly from an MD simulation trajectory (Zhou et al. 2016). Characterization in terms of the most abundant conformations of a structural ensemble (e.g. Yang et al. 2010) represents a highly informative, complimentary approach.

Time-resolved (TR) WAXS studies also have significant potential for investigating tertiary and quaternary conformational changes (Cammarata et al. 2008). When those changes can be induced by a short laser pulse time resolutions in the nanoseconds are possible. The methods outlined in this chapter are entirely applicable to each ‘snap-shot’ in a time series. Ligand-induced structural changes are much more difficult to track using TR WAXS because any structural change will be convoluted by variations in diffusion times – in other words, because each ligand will take a different amount of time to find and interact with a protein, it is impossible for all proteins to change structure synchronously.

New frontiers in the method have been suggested by the examples provided here. Efforts to collect highly accurate WAXS data to the highest possible resolutions (e.g. 2.0 Å) with the highest achievable signal-to-noise ratio have the potential to drive the method to the next level where structural changes induced by binding of a small molecule or ion, or even changes in the concentrations of buffer constituents could be observable and interpretable. Used in concert with computational approaches such as MD,

these advances could increase the power of WAXS for characterization of the structure and structural fluctuations of macromolecules in solution and for comprehensive studies of biochemistry in the scattering volume.

References

- Allaire M, Yang L (2011) Biomolecular solution X-ray scattering at the National Synchrotron Light Source. *J Synchrotron Radiat* 18:41–44
- Badger J, Grover P, Panjarian SB, Engen JR, Smithgall TE, Makowski L (2016) The c-Abl tyrosine kinase adopts multiple active conformational states in solution. *Biochemistry* 55:3251–3260
- Bardhan JP, Park S, Makowski L (2009) SoftWAXS: a computational tool for modeling wide-angle x-ray solution scattering from biomolecules. *J Appl Crystallogr* 42:932–943
- 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. (2007)
- Cammarata M, Levantino M, Schotte F, Anfinrud PA, 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, Moran F, Diaz JF, Pantos E, Andreu JM (1998) Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm. *Biop J* 74:2760–2775
- Cho HS, Dashdorj N, Schotte F, Graber T, Henning R, Anfinrud P (2010) Protein structural dynamics in solution unveiled *via* 100-ps time-resolved x-ray scattering. *Proc Natl Acad Sci* 107:7281–7286
- Doniach S (2001) Changes in biomolecular conformation seen by small angle x-ray scattering. *Chem Rev* 101:1763–1778
- Fischetti RF, Rodi DJ, Mirza A, Kondrashkina E, Irving T, Makowski L (2003) Wide angle x-ray scattering of proteins: effect of beam exposure on protein integrity. *J Synch Res* 10:398–404
- Fischetti RF, Rodi DJ, Gore DB, Makowski L (2004a) Wide angle x-ray solution scattering as a probe of ligand-induced conformational changes in proteins. *Chem Biol* 11:1431–1443
- Fischetti R, Stepanov S, Rosenbaum G, Barrea R, Black E, Gore D, Heurich R, Kondrashkina E, Kropf AJ, Wang S, Zhang K, Irving TC, Bunker GB (2004b) The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and x-ray absorption spectroscopy at the advanced photon source. *J Synchrotron Radiat* 11:399–405
- Ghosh M, Meerts IATM, Cook A, Bergman A, Brouwer A, Johnson LN (2000) Structure of human

- transthyretin complexed with bromophenols: a new mode of binding. *Acta Cryst D* 56:1085–1095
- Grishaev A, Guo L, Irving T, Bax A (2010) Improved fitting of solution x-ray scattering data to macromolecular structures and structural ensembles by explicit water modeling. *J Am Chem Soc* 132:15484–15486
- Hirai M, Iwase H, Hayakawa T, Inoue K (2002) Structural hierarchy of several proteins observed by wide-angle solution scattering. *J Synchrotron Rad* 9:202–205
- Hirai M, Koizumi M, Hayakawa T, Takahashi H, Abe S, Hirai H, Miura K, Inoue K (2004) Hierarchical map of protein unfolding and refolding at thermal equilibrium revealed by wide-angle x-ray scattering. *Biochemistry* 43:9036–9049
- Hirota N, Mizuno K, Goto Y (1997) Cooperative α -helix formation of P-lactoglobulin and melittin induced by hexafluoroisopropanol. *Protein Sci* 6:416–421
- Hong X, Hao Q (2009) High resolution pair-distance distribution function $P(r)$ of protein solutions. *Appl Phys Letts* 94:083903
- Hou J, Sims GE, Zhang C, Kim S-H (2003) A global representation of the protein fold space. *Proc Natl Acad Sci* 100:2386–2390
- Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL, Tsutakawa SE, Jenney FE Jr, Classen S, Frankel KA, Hopkins RC, Yang S-J, Scott JW, Dillard BD, Adams MWW, Tainer JA (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* 6:606–612
- Inouye H, Houde D, Temel DB, Makowski L (2016) Utility of solution x-ray scattering for the development of antibody biopharmaceuticals. *J Pharm Sci* 105:3278–3289
- Konarev PV, Volkov VV, Sokolova AV, Koch MH, Svergun DI (2003) Primus: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36:1277–1282
- Kumar S, Modig K, Halle B (2003) Trifluoroethanol-Induced $\beta \rightarrow \alpha$ transition in β -Lactoglobulin: hydration and cosolvent binding studied by ^2H , ^{17}O , and ^{19}F magnetic relaxation dispersion. *Biochemistry* 42(42):13708–13716
- Lattman EE (1989) Rapid calculation of the solution scattering profile from a macromolecule of known structure. *Proteins Struct Funct Genet* 5:149–155
- Luzzati V, Tardieu A (1980) Recent developments in solution x-ray scattering. *Ann Rev Biophys Bioeng* 9:1–29
- Makowski L (2010) Characterization of proteins with Wide-angle X-ray Solution Scattering (WAXS). *J Struct Funct Genom* 11:9–19
- Makowski L, Rodi DJ, Mandava S, Devrapahli S, Fischetti RF (2008a) Characterization of protein fold using wide angle x-ray solution scattering. *J Mol Biol* 383:731–744
- Makowski L, Rodi DJ, Mandava S, Minh D, Gore D, Fischetti RF (2008b) Molecular crowding inhibits intramolecular breathing motions in proteins. *J Mol Biol* 375:529–546
- Makowski L, Gore D, Mandava S, Minh D, Park S, Rodi DJ, Fischetti RF (2011) X-ray solution scattering studies of the structural diversity intrinsic to protein ensembles. *Biopolymers* 95:531–542
- Makowski L, Bardhan J, Gore D, Rodi DJ, Fischetti RF (2012) Observation of multi-wavelength anomalous diffraction using Medium Angle X-ray solution scattering (MADMAX) from proteins in solution. *Biophys J* 102:927–933
- McDonald RC, Steitz TA, Engelman DM (1979) Yeast hexokinase in solution exhibits a large conformational change upon binding glucose or glucose 6-phosphate. *Biochemistry* 18:338–342
- Minh D, Makowski L (2013) Wide-angle x-ray solution scattering for protein-ligand binding: multivariate curve resolution with Bayesian confidence intervals. *Biop J* 104:873–884
- Onuk AE, Akcakaya M, Bardhan J, Erdogmus D, Brooks DH, Makowski L (2015) Constrained maximum likelihood estimation of relative abundances of protein conformation in a heterogeneous mixture from small angle x-ray scattering intensity measurements. *IEEE Trans Signal Process* 63:5383–5394
- Onuk E, Akcakaya M, Bardhan J, Erdogmus D, Brooks D, Makowski L (2016) Dirichlet priors for MAP inference of protein conformation abundances from SAXS. *J. Signal Processing Sys* (in press)
- Park S, Bardhan JP, Roux B, Makowski L (2009) Simulated x-ray scattering of protein solutions using explicit-solvent molecular dynamics. *J Chem Phys* 130:134114
- Petoukhov MV, Svergun DI (2007) Analysis of x-ray and neutron scattering from biomacromolecular solutions. *Curr Opin Struct Biol* 17:562–571
- Putnam CD, Hammel M, Hura GL, Tainer JA (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
- Rodi DJ, Mandava S, Gore DB, Makowski L, Fischetti RF (2007) Detection of functional ligand binding events using synchrotron x-ray scattering. *J Biomol Screen* 12:994–998
- Roig-Solvas B, Bardhan J, Erdogmus D, Brooks D, Makowski L (2017) ProSFIT: a modular optimization framework for flexible reconstruction of protein structures from SAXS measurements. (in preparation)
- Stockwell BR (2000) Chemical genetics: ligand-based discovery of gene function. *Nat Rev Genet* 1:116–125
- Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76:2879–2886
- Svergun D, Barberato C, Koch MHJ (1995) CRY SOL-a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28:768–773
- Svergun DI, Richard S, Koch MHJ, Sayers Z, Kuprin S, Zacci G (1997) Protein hydration in solution:

- experimental observation by x-ray and neutron scattering. *Proc Natl Acad Sci* 95:2267–2272
- Svergun DI, Petoukhov MV, Koch MMHJ (2001) Determination of domain structure of proteins from x-ray solution scattering. *Biop J* 80:2946–2953
- Tiede DM, Zhang R, Seifert S (2002) Protein conformations explored by difference high-angle solution x-ray scattering: oxidation state and temperature dependent changes in cytochrome C. *Biochemistry* 41:6605–6614
- Tsutakawa SE, Hura GL, Frankel KA, Cooper PK, Tainer JA (2007) Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. *J Struct Biol* 158:214–223
- Volkova VV, Svergun DI (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J Appl Crystallogr* 36:860–864
- Wallace BA, Janes RW (2003) Circular dichroism and synchrotron radiation circular dichroism spectroscopy: tools for drug discovery. *Biochem Soc Trans* 31:631–633
- Walther D, Cohen FE, Doniach S (2000) Reconstruction of low resolution three-dimensional density maps from one-dimensional small angle x-ray scattering data for biomolecules in solution. *J Appl Crystallogr* 33:350–363
- Wen B, Peng J, Zuo X, Gong Q, Zhang Z (2014) Characterization of protein flexibility using small-angle X-ray scattering and amplified collective motion simulations. *Biop J* 107:956–964
- Yang S, Blachowicz L, Makowski L, Roux B (2010) Multidomain assembled states of Hck tyrosine kinase in solution. *Proc Natl Acad Sci* 107:15757–15762
- Zhang R, Thiyagarajan P, Tiede DM (2000) Probing protein fine structures by wide angle solution x-ray scattering. *J Appl Crystallogr* 33:565–568
- Zheng W, Tekpinar M (2011) Accurate flexible fitting of high-resolution protein structures to small-angle x-ray scattering data using a coarse-grained model with implicit hydration shell. *Biophys J* 101:2981–2991
- Zhou H, Li S, Badger J, Nalivaika E, Cai Y, Foulkes-Murzycki J, Schiffer C, Makowski L (2015) Modulation of HIV protease flexibility by the T80N mutation. *Proteins Struct Funct Bioinf* 83:1929–1939. doi:10.1002/prot.24737
- Zhou H, Li S, Makowski L (2016) Visualizing global properties of a molecular dynamics trajectory proteins. *Struct Funct Bioinforma* 84:82–91