Interpreting solution X-ray scattering data using molecular simulations

Jochen S Hub

Small-angle and wide-angle X-ray scattering in solution (SAXS, WAXS, SWAXS) is an increasingly accurate method for obtaining information on biomolecular structures, ensembles, and time-resolved dynamics at near-native conditions. However, the interpretation of the solution scattering data by computational methods is complicated by the low information content of the data, by scattering contributions from the hydration layer, and by unknown systematic errors. In the light of available computational methods, we first review the main computational challenges with the interpretation of SWAXS data. Molecular dynamics (MD) simulations may help to overcome these challenges and guide the interpretation of SWAXS in multiple ways. The physical information in atomistic force fields complements the low-information SWAXS data; explicit-solvent MD may be used to predict solvent scattering, and the MD-related sampling methods may guide the structure refinement against SWAXS data.

Address
University of Goettingen, Institute for Microbiology and Genetics, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany

Corresponding author: Hub, Jochen S (jhub@gwdg.de)

Current Opinion in Structural Biology 2018, 49:18–26
This review comes from a themed issue on Theory and simulation
Edited by Robert Best and Kresten Lindorf-Larsen

https://doi.org/10.1016/j.sbi.2017.11.002
0959-440X/© 2017 Elsevier Ltd. All rights reserved.

Introduction: combining experimental data and computation
The discovery of biomolecular structures and ensembles requires joint experimental and computational efforts. This is because the information content of common experimental data is too low to define all the degrees of freedom of a biomolecule. Instead, the data must be complemented with additional physical, stereo-chemical or structural information to avoid overfitting during structure determination and refinement.

A wide range of physical models have been used to complement experimental data. Popular physical models vary greatly by the number of degrees of freedom, ranging from rigid-body up to flexible all-atom models. The energy terms of the physical models (force fields) vary by accuracy, predictive power, and computational cost, ranging from simple volume-exclusion restraints up to all-atom molecular-dynamics force fields. An appropriate choice for a physical model is largely determined by the information content of the experimental data. Generally, the lower the information content of the data, the more predictive physical models are required to avoid overfitting.

Algorithms for combining experimental data with physical models may be roughly grouped into three families: first, a trivial ‘algorithm’ is to compare the experimental data with the data back-calculated from an unbiased simulation. Obviously, this approach is restricted to physical models that are sufficiently accurate to propose a good structural model. Second, unbiased Boltzmann sampling may be used to propose an approximate structural ensemble, and the ensemble may be reweighted a posteriori such that the reweighted ensemble agrees with the experimental data. Third, the experimental data may be integrated into a physical simulation as an experiment-derived energetic bias, thereby restraining the simulation into conformations that meet both the physical model and the experimental data [1].

Several excellent reviews summarized recent developments in structural modeling, with a focus on crystallography, NMR, cryo-EM, and integrative modeling, including perspectives on Bayesian methods and maximum entropy-based ensemble refinement ([2–4,5**,6**,7] and references therein). The concepts developed in these reviews are well transferrable to the interpretation of solution scattering data, so they shall not be repeated here. Instead, this review focuses on challenges specific to solution scattering, and we outline recent developments based on molecular dynamics (MD) simulation to overcome such challenges.

Small-angle and wide-angle X-ray scattering (SAXS, WAXS, or SWAXS) has developed to an increasingly accurate and predictive tool for obtaining structural information on biomolecules in solution [8–13]. The quality of SWAXS data has greatly benefited from better light sources, single-photon counting detectors, and from set-ups coupled to size-exclusion chromatography (SEC-SAXS) [14], which have led to reduced statistical noise and reduced systematic errors due to particle aggregation.
The risk of systematic errors has been further reduced by established standards for sample preparation, quality control, and data deposition [15,16]. Additional improvements are expected with the increasing availability of in-house SWAXS beamlines, which help to optimize samples before beamtime at the synchrotron. The sister method, small-angle neutron scattering (SANS), remains popular owing to the possibility of contrast variation experiments.

In practice, the structural interpretation from SWAXS and SANS is often conducted using computationally efficient algorithms and simplified physical models (such as simulated annealing, rigid-body modeling, and so on.), allowing modeling calculations on the laptop [17]. Since solution scattering data are often of low spatial resolution and of low information content, and since the data used to be noisy and at times biased by systematic errors, there was limited demand for elaborate and computationally expensive analysis methods (such as Bayesian methods, all-atom MD, etc.). Only in recent years, thanks to increasing availability of SWAXS data with low noise and reduced systematic errors, the need for elaborate methods that are capable of harvesting all the structural information in the data is more and more recognized [18].

**Challenges with the interpretation of SWAXS data**

To motivate the need for MD-based methods for SWAXS interpretation, we first review some of the challenges related to SWAXS data analysis. Recent MD-based method developments discussed further down are primarily an answer so these challenges.

**Challenge 1: low information content of SWAXS curves**

Because SWAXS curves are smooth and one-dimensional (1D), they contain quite a limited amount of information. How the information is distributed over the $q$-range is a matter of ongoing research [19], but it is generally accepted that experimental SWAXS curves do not contain more than 10–30 independent data points [8,9]. Hence, the number of backbone angles of biomolecules exceeds the number of independent data points of SWAXS curves by roughly two orders of magnitude. This precludes any straightforward fitting of protein structures against SWAXS data, but instead it leads to a high risk of overfitting.

One approach to reduce the risk of overfitting structural models is to constrain nearly all degrees of freedom of the biomolecule except for a few collective modes, leading to methods such as rigid-body modeling and normal mode refinement [20–22]. Alternatively, accurate physical and chemical knowledge may be added to the low-information SWAXS data, as provided by the potential energy function of modern biomolecular force fields.

**Challenge 2: prediction of SWAXS curves from structural models**

The interpretation of experimental data in terms of biomolecular structures requires accurate ‘forward models’, that is, theoretical predictions of the data from a given structural model. The physical basis of the SWAXS intensity is well understood; namely, the intensity is given by the orientationally averaged Fourier transform of the correlation function of the electron density contrast (compared to the buffer).

Since the quality of the experimental SWAXS data is improving, increasingly accurate forward models are required to harvest the structural information encoded in finer details of the data. Three aspects complicate the prediction of SWAXS curves of biomolecules:

(i) the density of the hydration layer of biomolecules differs from the density of bulk solvent, thus contributing the density contrast [23,24]. Simplified methods for SWAXS prediction do not predict the hydration layer density and hence require fitting of the hydration layer against the data (Table 1) [25–30]. This fitting may simultaneously adjust the radius of gyration $R_g$, a key quantity extracted from SWAXS experiments. In consequence, the fitting parameter may absorb smaller modulations of $R_g$ [31].

(ii) Computing the density contrast requires accurate knowledge of the volume of solvent that is displaced by the solute. Many methods for SWAXS prediction rely on ‘reduced atomic form factors’, which incorporate the volume displaced by each atom into the atomic form factors. The displaced atomic volumes are often taken from tables derived from high-resolution crystal structures [32] or densitometric data [33,34], and subsequently fitted against the data. This protocol may add uncertainty since estimates for atomic volumes greatly differ [25,32,35]. For instance, the volume for the backbone N–H group reported by Fraser et al., as adopted by many SWAXS tools [34], is roughly half the volume estimated from Voronoi tessellation [35]. In addition, effective atomic volumes could depend on the type of solute and atomic packing density (rigid vs. disordered protein, nucleotide, or detergent), posing additional uncertainty on the volume estimates.

(iii) Fluctuations of atoms and larger groups may influence the SWAXS curve [19,36–38], but are often neglected during the interpretation of the data, adding uncertainty to the interpretation of the data.

**Challenge 3: overfitting and lack of a cross validation set**

Because of the low information content of SWAXS data, overfitting is a major problem for structural modeling [39]. A rigorous strategy to avoid overfitting would be to
decompose the data into statistically independent training and test sets, as has become routine in crystallography and, more recently, in cryo-EM [40–42]. Accordingly, the structural model is fitted using only the training set, while overfitting is excluded by cross-validation against the test set. However, this strategy requires that the information in the test set is not used during model fitting — indeed, problems due to the implicit use of the test set, for instance by human supervision [43] or in consequence of non-crystallographic symmetry, have been discussed [44].

Since it is unclear how the structural information of SWAXS curves is distributed over the \(q\)-range, SWAXS curves were not decomposed into independent training and test sets. Instead, overfitting must be avoided by the use of sufficiently accurate physical models, thereby requiring only little model optimization against the data [45]. Further, the \(X_{\text{free}}\) measure has been proposed in analogy to the \(R_{\text{free}}\) value in crystallography [40,46], yet the significance of \(X_{\text{free}}\) has remained controversial [47].

In principle, Bayesian methods for model refinement should avoid overfitting. If the information in the data together with the physical model is insufficient to define the biomolecular structure, Bayesian methods would yield merely wide posterior distributions for the refined model, indicating large confidence intervals. Such methods have been used in NMR refinement since long [48], but related approaches have only recently been considered for SWAXS-based structure refinement [49*,50*].

**Challenge 4: account for systematic errors**

Common sources for systematic errors in SWAXS are aggregation, inter-particle repulsion, and poor buffer matching. Because modern photon-counting detectors allow data collection with low statistical errors, the overall uncertainty of the data is by now often dominated by systematic errors. This development asks for a statistically founded treatment of systematic errors. Bayesian methods that simultaneously estimate the biomolecular structures and systematic errors provide a route to accomplish this [50*], as originally proposed in the context of NMR refinement [48,51].

**Challenge 5: confidence intervals of structural models**

Students in natural science are taught that reporting quantities without at least approximate confidence intervals (CIs) is pointless. In SWAXS-based modeling, methods for the calculation of statistically founded CIs for structural models are still underdeveloped. Calculating CIs is generally complicated by unknown systematic errors, unknown errors of the physical model, and by the unclear information content of SAXS data, but also by the lack of computational tools for propagating the errors of the data and of the physical model into the error of the fitted structural model.

An important step toward such error analysis was recently suggested for *ab initio* low-resolution modeling against SWAXS data, based on the variability of repeated simulated annealing calculations [52]. However, it should be noted that the spread of repeated ‘best fits’ is not equivalent to CIs. Only very recently, a route to computing CIs based on Bayesian statistics has been shown in the

---

**Table 1**

Incomplete list of methods for predicting SWAXS curves from structural models: Fitting of hydration layer required (\(l_{\text{phys}}\), including method that ignore the hydration layer), using tabulated reduced form factors (\(f_{\text{red}}\)), resolution [atomic or coarse grained (CG)], fluctuations included, free availability [Download (D), web server (W)]. Additional software is listed in Refs. [83,84].

<table>
<thead>
<tr>
<th>ID</th>
<th>Name/authors</th>
<th>Year</th>
<th>(l_{\text{phys}}/f_{\text{red}})</th>
<th>Resol.</th>
<th>Fluct.</th>
<th>Avail.</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CRYSTOL</td>
<td>1995</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>D/W</td>
<td>[25]</td>
</tr>
<tr>
<td>2</td>
<td>ORNL-SAS</td>
<td>2007</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[65]</td>
</tr>
<tr>
<td>3</td>
<td>SoftWAXS</td>
<td>2009</td>
<td>Yes/–</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[66]</td>
</tr>
<tr>
<td>4</td>
<td>Fast-SAXS-pro</td>
<td>2009</td>
<td>Yes/yes</td>
<td>CG</td>
<td>Yes</td>
<td>D/W</td>
<td>[30,36]</td>
</tr>
<tr>
<td>5</td>
<td>FoXS</td>
<td>2010</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>W</td>
<td>[27]</td>
</tr>
<tr>
<td>6</td>
<td>PHAISTOS</td>
<td>2010</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>W</td>
<td>[68]</td>
</tr>
<tr>
<td>7</td>
<td>AquaSAXS/AquaSol</td>
<td>2011</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[70]</td>
</tr>
<tr>
<td>8</td>
<td>SASbx/Zernike</td>
<td>2012</td>
<td>Yes/–</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[71]</td>
</tr>
<tr>
<td>9</td>
<td>RISM-SAXS</td>
<td>2014</td>
<td>–/yes</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[72*]</td>
</tr>
<tr>
<td>10</td>
<td>BCL::SAXS</td>
<td>2015</td>
<td>Yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[73]</td>
</tr>
<tr>
<td>11</td>
<td>Pepsi-SAXS</td>
<td>2017</td>
<td>yes/yes</td>
<td>atom.</td>
<td>–</td>
<td>D</td>
<td>[74,75]</td>
</tr>
</tbody>
</table>

---

context of structure refinement against SAXS data [50]. The error of the physical model is so far neglected. Interesting ideas for propagating the error in the physical model by using ‘distributions of force fields’ instead of using a single force field were proposed, but the ideas were to the best of our knowledge not yet implemented [5**].

Challenge 6: interpretation of SWAXS data of heterogeneous ensembles

Algorithms for ensemble reconstructions were reviewed frequently, hence we mention these only briefly [6**,39,53–55]. The majority of work has focused on NMR data, partly complemented by SWAXS, whereas less work has focused on pure SWAXS data [21,45,56,57]. Because of the lower information content of SWAXS compared to NMR data, reconstructing the ensembles from SWAXS data is more challenging and more prone to overfitting, suggesting that the data should, if possible, be complemented by accurate physical models. Indeed, it was pointed out that reproducing the data with a fitted ensemble does by far not guarantee that the ensemble is correct; hence, great care has been advised for interpreting ensembles fitted to SWAXS data [6**].

Most algorithms for SAXS-based ensemble modeling follow an ‘ensemble reweighting’ approach [7,58], which requires exhaustive sampling of the conformational space in free a simulation. Therefore, reweighting approaches were mainly used with simplified physical models, such as coarse-grained or rigid-body models [21,57], or using an accelerated sampling scheme that disturbs the Boltzmann distribution [58]. Theoretical frameworks for refining ensembles by energetic restraints to experimental data have been developed, yet a previous work focused on test applications of reduced complexity and simplified force fields [49*,59,60,61,62**]. As such, the refinement of ensembles of complex biological systems based on SWAXS data and accurate physical models, if possible including the calculation of CIs, has remained challenging.

Method developments for SWAXS data interpretation with atomistic MD simulations

Prediction of SWAXS curves

In the pioneering SWAXS prediction method by Merzel and Smith, the hydration layer was taken for the first time from an explicit-solvent MD simulation, revealing an increased hydration layer density of lysozyme; however, the method modeled the buffer subtraction using reduced form factors (hence using tabulated atomic volumes) [24,73]. Inversely, the AXES method avoids the use of ‘reduced form factors’ by using an explicit-solvent representation for the excluded solvent; however, the hydration layer is not optimized by an MD simulation and hence requires fitting against the experimental data [26]. Recently, a range of methods have been proposed for SWAXS prediction based on explicit-solvent MD simulations of both the hydration layer and the excluded solvent [38,74,75,77,78]. Here, Orozuchi and coworkers built upon earlier work by Seki et al. [80], while Chen and Hub built upon the method by Park et al. [77], making it applicable to flexible MD simulations, and correcting it for small yet significant force field and finite-size artifacts [38], similar to thoughts by Köfinger and Hummer [78]. The method was also implemented in the web server WAXSiS [79], which runs an explicit-solvent MD simulation with an uploaded biomolecule, making the SWAXS predictions accessible to non-experts. Notably, computationally efficient approximations for modelling the hydration layer with atomic resolution were suggested, either based on integral equations (RISM) [70], or on protein-water radial distributions functions parameterized against MD simulations (HyPred) [76].

Compared to implicit-solvent methods (see Table 1), SWAXS predictions based on explicit-solvent MD are computationally more expensive but hold a number of advantages: firstly they naturally reproduce the increased radius of gyration due to the hydration layer [38], thus avoiding a fitting parameter for the hydration layer. This further avoids that the radius of gyration is fitted against the experimental data. Secondly by avoiding the use of ‘reduced form factors’, they circumvent uncertainties and fitting parameters due to the atomic volumes. Thirdly the SWAXS predictions remain valid at wide angles, where the internal structure of water may be reflected in the signals, and finally most of the methods account for thermal fluctuations of the biomolecule (see Table 1). For small solutes such as short peptides, however, explicit-solvent SWAXS predictions may converge slowly, in consequence of a weak contrast between solute and solvent. Hence, careful convergence checks are generally advised. The applied water force field hardly influences the SAXS predictions if (and only if) inaccurate bulk densities of certain water models are corrected. In the WAXS regime around the water scattering peak, in contrast, the water force field strongly influences the predictions, as expected from the fact that radial distributions functions of water are quite force field-dependent [38,81].

However, since only few studies compared implicit-solvent with explicit-solvent methods, it is not well understood under which conditions explicit-solvent models are mandatory for correctly interpreting the data, and under which conditions implicit-solvent methods offer an acceptable and computationally efficient approximation. SAXS curves computed from ensembles of disordered proteins agreed well at small angles between implicit-solvent and explicit-solvent methods [82*,83]. In the same study, in contrast, only an explicit-solvent method provided accurate fitting-free SAXS curves of the ring-shaped PCNA protein [83]. Another study investigated the effect of solvent-related fitting parameters, when
simulations of a protein-detergent complex, conducted with increasing detergent numbers $N_{\text{det}}$, are validated against SEC-SAXS data. [14,31]. The authors found that implicit-solvent methods tend to overfit the hydration layer density, absorbing the variations of $N_{\text{det}}$ into the fitting parameter, and thereby loosing the ability to differentiate between the right and wrong $N_{\text{det}}$. Explicit-solvent calculations did not face these problems. Further, the study revealed that an ensemble average of the protein-detergent complex agreed better with experiment than any single structure, highlighting the importance of conformational fluctuations even for a relatively well-defined system such as a protein-detergent complex. Clearly, more work is required to identify strengths and limitations of available SWAXS prediction methods.

**Structure refinement by energetic restraints to SWAXS data**

Multiple MD-based methods were recently published that aim at the refinement of biomolecular structures against SWAXS data using experiment-derived energetic restraints [50*,84*,85*,86]. Forces derived from energetic restraints may ‘drive’ the biomolecule into conformations that satisfy the experimental data, thereby, for instance, guiding the simulation over energy barriers associated with large-scale conformational transitions. As such, simulations are not only a tool to interpret the data; instead, SAXS data may, in turn, help the simulation to overcome limitations due to sampling problems and force field inaccuracies. This property is in contrast to methods based on ‘reweighting’ of states that require the sampling of the relevant states in a free, unbiased simulation, thereby restricting the methods to simplified and computationally efficient physical models [45,57].

Energetic restraints to the SWAXS data have been implemented with a harmonic biasing potential [85*,86] using metadynamics [84*], or following a Bayesian approach [50*]. Further, the SAXS curve prediction (or forward model) used to evaluate the energetic restraints vary between the methods: certain methods used the Debye formula, similar to pioneering work by Grishaev *et al.* on combined SAXS/NMR refinement [84*,86,87]. These two methods neglected scattering contributions from the hydration layer [84*,86]. Other methods accurately accounted for the hydration layer and excluded solvent by using explicit-solvent descriptions [50*,85*].

**Interpretation of time-resolved SWAXS data**

Time-resolved SWAXS (TR-SWAXS) experiments were used to track conformational dynamics of photoactive proteins in solution. Synchrotron-based TR-SWAXS experiments have covered the time regime between milliseconds and ~100 ps [88–90]. Conducted at X-ray free electron lasers, such experiments have reached sub-picosecond time resolution, providing a glimpse on ultrafast protein dynamics [91,92]. The interpretation of such signals in terms of real-space dynamics is complicated by

---

**Experimental data**

- SWAXS curve(s), statistical errors

**Computational considerations**

- Physical model / prior
  - Force field
  - Model constraints
  - Single structure / ensemble?

- Forward model / SWAXS prediction
  - Implicit- / explicit-solvent calculations
  - Atomic volumes
  - Fluctuations

- Coupling + sampling algorithm
  - Harmonic restraints to experiment
  - Metadynamics
  - Bayes
  - Maximum entropy
  - Account for systematic errors
  - Enhanced sampling method

**Interpretation**

- In terms of:
  - Structure(s)
  - Conformational transition
  - Protein complex
  - Ensemble
  - Confidence intervals
  - Estimated systematic errors

---

Scheme and considerations for the interpretation of SWAXS data. For the interpretation of SWAXS data, MD simulations and MD force fields may provide an accurate physical model (prior), be used for accurate SWAXS predictions (forward model), and provide algorithms for conformational sampling.
the fact that the signals reflect non-equilibrium conditions; in the picosecond regime, the signals even reflect highly dissipative processes. To interpret such data, methods for the refinement of equilibrium structures and ensembles may not be applicable, but instead realistic non-equilibrium simulations, which reproduce the data without an experiment-derived bias, are needed.

Synchrotron-based TR-SWAXS data were interpreted using available crystallographic structures, or via low-resolution bead-modeling [88,89,95]. Recently, MD simulations were used to interpret sub-picosecond TR-SWAXS data of myoglobin, providing an atomic view of the ultrafast dissipation of energy following CO-photo-dissociation [92,94], a process coined ‘protein quake’ in the 1980s [95]. The simulations revealed a strong SAXS signal as the quake propagated across the hydration layer into bulk solvent. Hence, the SAXS data were dominated by solvent and not by protein dynamics, highlighting that explicit-solvent SAXS calculations were required to correctly interpret the data.

Notably, TR-SWAXS patterns are anisotropic [96,97]. The anisotropy leads to additional structural information compared to common orientationally averaged equilibrium SWAXS data. A method for computing anisotropic TR-SWAXS patterns has been proposed [98], but applications that harvest the additional information remain elusive.

**Outlook**

A reliable atomistic interpretation of SWAXS data builds on three computational pillars (Figure 1): (a) accurate physical models, that is, accurate prior distributions for conformational states; (b) accurate and predictive SWAXS curve calculations from atomic models (forward models); and (c) algorithms founded on probability theory for integrating available experimental data into conformational sampling, for deriving structures and ensembles. All three pillars are, we believe, equally important. Most of the previous works focused on one of the three pillars, while paying less attention to the others, which is a practical strategy at the stage of method developments. Since force fields, SAXS prediction methods, and sampling methods have seen exciting developments in recent years, future efforts should try to integrate developments from the three pillars. In addition, the credibility of structural models derived from SWAXS will strongly benefit from increased efforts in computing confidence intervals, as well as from cross validation against complementary data, for instance from SAXS, FRET, or DEER. Finally, benchmark suites of high-precision experimental SWAXS data are urgently needed to reveal strengths and weaknesses of the wide range of SWAXS-related software. As high-precision experimental SWAXS data are becoming more abundant, such developments will pave the way to a truly quantitative use of SWAXS data in structural and unstructural biology [53].

**Conflict of interest statement**

Nothing declared.

**Acknowledgement**

This work was supported by the Deutsche Forschungsgemeinschaft (grant no. HU 1971/1-1, HU 1971/4-1).

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


This is an instructive perspective on using the maximum entropy principle for combining experimental data with physical models. The authors propose interesting options for propagating the error of the physical model into the error of the fitted structural model, by using ‘distributions of force fields’ instead of using a single force field.


An insightful review of methods for the fitting of structural ensembles against experimental data is being presented. Available methods are systematically classified and illustrated by a geometric model. Underlying assumptions and the correct interpretation of fitted ensembles are clarified.


Variations of the SAXS curve are correlated with fluctuations of the protein, suggesting that information on protein dynamics is not uniformly distributed over the q-range of SAXS curve.


A method for Bayesian inference on continuous protein ensembles from SAXS data is presented.


MD simulations are employed for refining a single protein structure or an ensemble of two states against SAXS data. Bayesian inference is used to simultaneously estimate structures, structural weights, and systematic errors, providing statistically founded confidence intervals for structures and weights. SAXS predictions are based on explicit-solvent calculations.

SAXS/WAXS interpretation with MD simulations

Hub


A framework for Bayesian inference of continuous ensembles is presented, based on a reference distribution and experimental data. The relation to maximum entropy refinement and ensemble refinement with a fixed number of states is clarified.


A refined force field is used to characterize the denaturant-induced unfolding of a protein. The unbiased simulations reveal good agreement with FREET and SAXS data.


Metadynamics is suggested as a sampling algorithm for driving simulations into agreement with SAXS data.


A method for structure refinement is introduced, by coupling MD simulations to SWAXS data with a differentiable energetic restraint. The SWAXS prediction (forward model) is fully based on explicit-solvent calculations.


94. Brinkmann LUL, Hub JS: Ultrafast anisotropic protein quake propagation after CO-photodissociation in myoglobin. Proc Natl Acad Sci U S A 2016, 113:10565-10570. It is shown that explicit-solvent SWAXS calculations are mandatory for the interpretation of femtosecond time-resolved SWAXS data. The first molecular movie of a protein quake is presented.


