FEBS Letters 589 (2015) 2570–2577

FEBS Letters

journal homepage: www.FEBSLetters.org

Review

A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins

Alexey G. Kikhney, Dmitri I. Svergun *

European Molecular Biology Laboratory, Hamburg Outstation, Notkestr. 85, Geb. 25a, 22607 Hamburg, Germany

ARTICLE INFO

Article history:
Received 26 July 2015
Revised 14 August 2015
Accepted 15 August 2015
Available online 29 August 2015

Edited by Wilhelm Just

Keywords:
Small angle scattering
Small-angle X-ray scattering
Flexibility
Disorder
Intrinsically disordered protein

ABSTRACT

Small-angle X-ray scattering (SAXS) is a biophysical method to study the overall shape and structural transitions of biological macromolecules in solution. SAXS provides low resolution information on the shape, conformation and assembly state of proteins, nucleic acids and various macromolecular complexes. The technique also offers powerful means for the quantitative analysis of flexible systems, including intrinsically disordered proteins (IDPs). Here, the basic principles of SAXS are presented, and profits and pitfalls of the characterization of multidomain flexible proteins and IDPs using SAXS are discussed from the practical point of view. Examples of the synergistic use of SAXS with high resolution methods like X-ray crystallography and nuclear magnetic resonance (NMR), as well as other experimental and *in silico* techniques to characterize completely, or partially unstructured proteins, are presented.

1. Introduction

Small angle X-ray scattering (SAXS) is a powerful method for the structural characterization of both ordered and disordered proteins in solution. It provides information about the sizes and shapes of proteins and complexes in a broad range of molecular sizes ranging from a few kDa to GDa and under various experimental conditions varying from extreme (e.g. high pressure or cryo-frozen) to nearly native [1–3]. SAXS, especially using high intensity synchrotron sources, is a rapid technique and time-resolved studies yield unique information about kinetics of processes and interactions [4–6]. SAXS is one of the few methods that allow one to quantitatively characterize conformational polydispersity, in particular, of completely or partially disordered macromolecules, including multi-domain proteins with flexible linkers and intrinsically disordered proteins (IDPs). Advances in SAXS instrumentation in the last years allowed for high-throughput structural sample analysis with increasing speed of both data collection and sample characterization [7,8]. It is now possible not only to determine a few simple geometric parameters of flexible ensembles, but also to assess their probable configurations in solution [9,10].

It has been predicted that more than 35% of human proteins have significant regions of disorder [11] and about 25% are likely to be completely disordered, i.e. lack a stable tertiary structure in solution [12]. These proteins are functionally important for many cellular regulatory processes, and may also be involved in pathological processes associated with protein misfolding or aggregation [13,14]. Flexible regions can serve as connectors between folded protein domains or as docking regions for binding partners [15–18]. Some regions lose their flexibility upon binding, others remain flexible [19]. Under physiological conditions these proteins constantly fluctuate between different structural states, resulting in a dynamic mixture of conformations in a polydisperse solution. Quantitative characterization of such heterogeneous systems is a difficult task, and the use of SAXS often provides unique information on the structural properties of the flexible macromolecules. In the following, an overview of SAXS is given putting a special emphasis on its practical applications to flexible proteins and IDPs.

1.1. Technical aspects: how does SAXS work?

The following section gives a general overview of SAXS; for more detailed information refer to recent reviews or textbooks [20–23]. The setup of a SAXS experiment is conceptually simple: a solution of particles usually placed in a quartz capillary is illuminated by a collimated monochromatic X-ray beam, the intensity of the scattered X-rays is recorded by an X-ray detector (Fig. 1). The scattering pattern of the pure solvent is collected as well and subtracted from the sample solution scattering leaving only the

* Corresponding author. Fax: +49 40 89902149.
E-mail address: d.svergun@embl-hamburg.de (D.I. Svergun).

http://dx.doi.org/10.1016/j.febslet.2015.08.027
0014-5793/

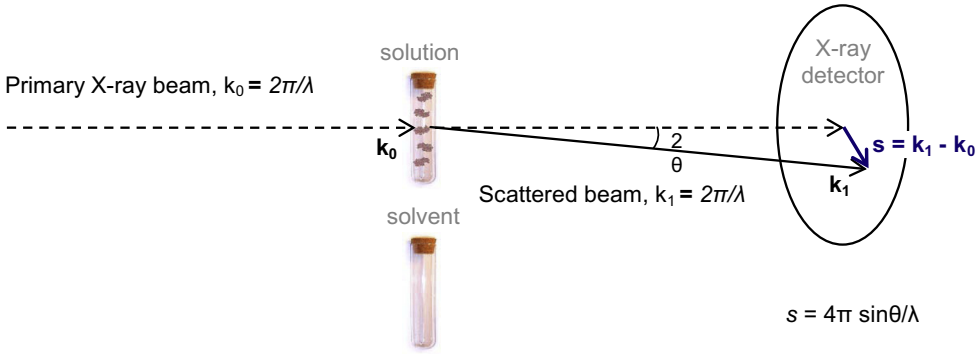

Fig. 1. Schematic representation of a SAXS experiment.

signal from the particles of interest. The resulting scattering pattern is related to the overall shape and size of the particles under investigation [22,23].

Due to the random orientations of the particles in solution the scattering pattern is isotropic, and thus, the scattering pattern recorded usually by a two-dimensional detector can be radially averaged. The scattering intensity I is represented as a function of momentum transfer $s = 4\pi\sin\theta/\lambda$, where λ is the beam wavelength and 2θ is the scattering angle:

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

here the scattering amplitude $A(s)$ is a Fourier transform of the excess electron density:

$$A(\mathbf{s}) = \Im[\rho(\mathbf{r})] = \int \Delta\rho(\mathbf{r}) \exp(i\mathbf{sr}) d\mathbf{r}, \quad (2)$$

where $\Delta\rho(r) = \rho(r) - \rho_s$, $\rho(r)$ and ρ_s being the electron density of the particle and of the solvent, respectively, and $\langle\ \rangle_\Omega$ stands for the spherical average. These isotropic scattering patterns are plotted as radially averaged one-dimensional curves $I(s)$ (Fig. 2A).

The advanced biological SAXS experiments are usually conducted on synchrotron sources providing high brilliance X-rays. All major synchrotrons like e.g. ESRF (Grenoble, France), DESY (Hamburg, Germany), Diamond (Oxford, Great Britain), ANL (Argonne, USA), SSRL (Stanford, USA) and Spring-8 (Himeji, Japan) have beamlines optimized for biological SAXS experiments. Often, good results can also be obtained with laboratory home X-ray sources (e.g. fabricated by Rigaku or Bruker), which, although yielding much lower flux than the synchrotrons, still have low background allowing one to reliably measure the low angle scattering signal. On synchrotrons the exposure times range from fractions of a second to minutes and on home sources they can be up to hours, whereas the amount of sample required for both experiments is approximately the same.

In SAXS, one typically needs 10–100 μl of sample per measurement, or a total of 1–2 mg of purified protein including a compulsory concentration series measurement (e.g. 1, 2, 5 and 10 mg/ml).

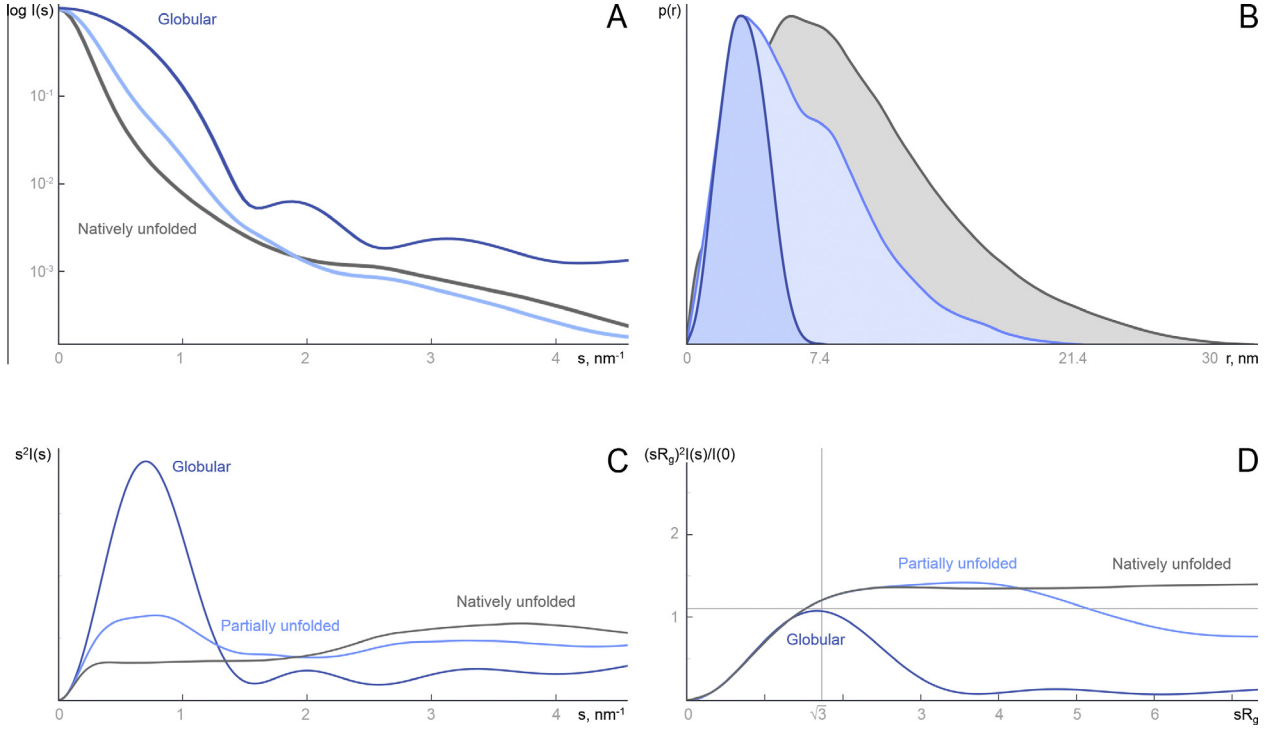

Fig. 2. Data simulated from three 60 kDa proteins: globular (dark blue), 50% unfolded (light blue) and fully disordered (gray). (A) Logarithmic plot of the scattering intensity $I(s)$ (in arbitrary units) *vs.* s (in inverse nanometres). (B) Distance distribution functions $p(r)$ (in arbitrary units) *vs.* r (in nanometres). (C) Kratky plot $s^2I(s)$ *vs.* s. (D) Normalized (or "dimensionless") Kratky plot $(sR_g)^2I(s)/I(0)$ *vs.* sR_g.

The sample scattering intensity is proportional to the concentration: the higher the concentration the better the signal-to-noise ratio of the solvent-subtracted data. However at higher concentrations the distances between individual particles come to the same order of magnitude as the intra-particle distances and therefore are contributing to the scattering pattern. A decrease of intensity at low angles usually indicates repulsive inter-particle interactions whereas a sharp increase of intensity points to attractive interactions, which may lead to unspecific aggregation of the sample. To minimize this contribution the low angle data measured at lower protein concentrations are normally merged with the high angle data from the higher concentration to yield the final composite scattering curve. Assuming that the inter-particle interference is linearly dependent on concentration it is also possible to extrapolate the scattering pattern to infinite dilution [24].

On synchrotrons the intense X-ray beam may cause radiation damage to the samples. This is usually controlled by repeated short exposures of the same sample solution during measurement: comparison of the resulting scattering patterns quickly detects differences due to radiation damage. Various means are further used to limit the radiation damage: flowing of the sample during data collection, attenuation of the beam, solution additives like glycerol [25]. For a meaningful analysis of the overall shape but also of the flexibility of a given protein, solutions containing single molecular species without aggregates, i.e. monodisperse solutions, are usually required. Typically, monodispersity better than 95% is required for successful structural analysis, which must be verified by methods like native gel filtration, dynamic light scattering (DLS), or analytical ultracentrifugation (AUC).

1.2. What information is provided by SAXS?

Several characteristic parameters of the investigated sample can be obtained directly from the experimental scattering pattern including molecular weight, excluded particle volume, maximum dimension D_{max} and the radius of gyration R_g. For monodisperse systems, where all particles in solution are identical and the experimental intensity is given by the spherical average of the single particle scattering following Eqs. (1) and (2), these parameters directly reflect the overall characteristics of the particle. For polydisperse systems, the average is also performed over the types of particles present in solution, and therefore the experimentally determined values do not correspond to a single particle, but rather to an average over the entire ensemble. Still, even for polydisperse systems the overall parameters can provide useful information about the particle sizes and structural properties. As we shall see below, for flexible objects, the analysis of the distributions of the overall parameters like R_g reconstructed from the scattering data, is a useful tool for a quantitatively characterization of flexibility.

The classical and the most known parameter directly extracted from the SAXS data, R_g provides a measure of the overall size of the macromolecule. The R_g is the average root-mean- square distanced to the centre of density in the molecule weighted by the scattering length density. R_g is smaller for proteins with a compact shape as compared to extended proteins with identical number of amino acids. It can be estimated using the Guinier approximation [26], which states that for very small angles ($s < 1/R_g$) the intensity depends only on two parameters: $I(s) = I(0)\exp(s^2R_g^2/-3)$. Practically, this means that the scattering intensity plotted as $\ln I(s)$ vs. s^2 should be a linear function for a particle of any shape. From the slope of the linear fit one can determine the R_g and the intercept gives the forward scattering $I(0)$, which is proportional to the molecular weight and the concentration of the protein. For the scattering patterns having a limited number of experimental data points the accuracy of the Guinier fit can be improved by extending the range of the linear fit up to $s < 1.3/R_g$. Attractive interactions between particles and nonspecific aggregation result in overestimation of both R_g and $I(0)$, whereas repulsive interactions lead to an underestimation of R_g and $I(0)$. Typically these effects can be removed by measurements of a series of SAXS curves at different concentrations and extrapolation to infinite dilution as described above.In some cases, Debye's approximation can be used:

$$\frac{I(s)}{I(0)} = \frac{2}{x^2}(x - 1 + e^{-x}) \tag{3}$$

where $x = (sR_g)^2$ [27]. This representation may be advantageous over the Guinier fit as the validity of Debye's approximation extends to larger momentum transfer ranges, but the applicability of the latter approximation is limited to extended polymer-line chains.

In real space, the particle sample can be conveniently described by the distance distribution function $p(r)$ which is a histogram of distances between all possible pairs of atoms within a particle. The scattering pattern of a particle $I(s)$ is a Fourier transform of its $p(r)$ function and vice versa, being related to each other by the equation:

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

By definition the distance distribution function equals zero at $p(0)$, is non-negative, terminates smoothly at the maximum dimension D_{max} and is zero for $r > D_{max}$. The $p(r)$ function can be obtained from the experimental scattering data using indirect Fourier transformation [28,29]. In some cases it is more intuitive to interpret the structural properties analysing the $p(r)$ rather than the scattering data itself. In particular, globular compact particles have a symmetric bell-shaped $p(r)$, whereas unfolded particles have an extended tail (Fig. 2B). The radius of gyration can be calculated from the $p(r)$ function using the equation:

$$R_g^2 = \frac{\int_0^{D_{max}} r^2 p(r)dr}{2\int_0^{D_{max}} p(r)dr} \tag{5}$$

which often yields a more reliable estimate compared to the Guinier approximation, in particular for unstructured systems [30]. The maximum dimension D_{max} is not only an important characteristic of the particle under investigation but also sets an important limitation on the used data range: $s_{min} < \pi/D_{max}$. This means that more extended particles require the experimental data collected at smaller angles, i.e. closer to the primary beam.

2. SAXS detects protein flexibility

The SAXS data on their own do provide several indicators of the presence of protein flexibility, which are listed below.

Traditionally, the Kratky plot ($s^2I(s)$ as a function of s, Fig. 2C) is employed to qualitatively identify disordered states and to distinguish them from globular proteins. The Kratky representation is able to visualize particular features of the scattering profiles for an easier identification of the folding state and flexibility [12,31]. The scattering intensity from a solid body decays at high angles approximately as $I(s) \sim 1/s^4$ conferring a bell-shaped Kratky plot with a well-defined maximum. Conversely, an ideal Gaussian chain has a $1/s^2$ asymptotic of $I(s)$ and therefore presents a plateau at large s values. In the case of an extended thin chain, the Kratky plot also presents a plateau over a specific range of s, which is followed by a monotonic increase. The latter case is normally observed experimentally for unfolded proteins; chemical or thermal unfolding experiments monitored by SAXS are good examples on how the decrease in protein compactness is translated into the Kratky plots [30]. In order to compare the folding states of different proteins it is convenient to normalize the data such that $I(0) = 1$ and to

multiply s by R_g prior to plotting. This way the information about the size of the protein is removed but the information about the shape is kept. The Kratky plot of such normalized data is called dimensionless Kratky plot (Fig. 2D).

Unstructured proteins, due to the presence of highly extended conformations, are characterized by large average sizes compared to globular proteins with a tightly packed core. The comparison of the experimentally measured R_g and D_{max} with these predicted by theoretical models is used to diagnose the unstructured nature of a protein.

An important metric holding for the unstructured proteins is given by Flory's equation describing how the radius of gyration an unfolded protein correlates to its length [32]

$$R_g = R_0 N^{\nu}, \tag{6}$$

here N is the number of amino acid residues, for chemically denatured proteins R_0 is 1.33 ± 0.076 and ν is 0.598 ± 0.028 [1]. The agreement between the ν value obtained experimentally and the theoretical models suggest the random coil nature of the chemically denatured proteins, at least in terms of the R_g parameter. However, the conformational sampling of a chemically denatured state differs from that found for IDPs in the native state: for intrinsically disordered proteins R_0 is 2.54 ± 0.01 and ν is 0.522 ± 0.01 [33] indicating that IDPs are in general more compact than chemically denatured proteins.

3. SAXS of flexible and intrinsically disordered proteins: what is different?

Flexible systems are, strictly speaking, polydisperse systems, where the solution contains different types of particles. In general, polydispersity makes the analysis of the particle structure by SAXS difficult or impossible, because, contrary to monodisperse systems, the measured intensity is no longer related to the scattering from a single particle. Instead, each individual protein structure present in the investigated sample adds to the scattering pattern, resulting in an averaged intensity. Given K distinct particle types (components) in a mixture, SAXS intensity is a linear combination of their contributions:

$$I(s) = \sum_{k=1}^{K} \nu_k I_k(s), \tag{7}$$

where ν_k and $I_k(s)$ are the volume fraction and the scattering intensity from the k-th component, respectively.

Perhaps the worst case of a polydisperse system in practical applications is unspecific aggregation, which often prevents meaningful analysis of the SAXS data. However, SAXS is extremely useful for the studies of specific aggregates (oligomers) and equilibrium systems like oligomeric protein mixtures. Here, the number of components is relatively low (e.g. two components for a monomer–dimer equilibrium) and their computed scattering patterns are often available from the high resolution structures or models, such that SAXS can provide the volume fractions of the components by solving Eq. (7) with respect to unknown ν_k [34].

Both IDPs and modular multi-domain proteins with flexible linkers can be described as mixtures of different conformations of the same macromolecule, and their scattering is therefore described by Eq. (7). However, these mixtures contain very large numbers of configurations ($K \gg 1$), so that plain decomposition, as in the case of oligomeric mixtures, is not feasible.

3.1. Shape reconstruction and modelling

SAXS allows for *ab initio* and rigid body modelling: It is possible to build a low resolution model either without any *a priori* information or by using higher resolution structures from complementary methods such as X-ray crystallography and NMR, and to validate this model against the SAXS data [22]. In case of interpreting the data from a monodisperse solution the theoretical scattering pattern computed from a single model should fit the experimental data. Classical rigid body modelling approaches utilize the scattering from the domains/subunits and may even account for the missing, potentially flexible loops or linkers [35], but these methods always search for a single configuration fitting the experimental data. Clearly, a disordered protein cannot be represented by a single model and alternative modelling approaches are required. In case of a flexible or intrinsically disordered protein we can construct a multitude of models representing different states of the protein in solution and compute the scattering from a mixture of such models.

The first step in this approach, adequate generation of a manifold of models exploring the possible conformational space of IPDs, requires dedicated computational tools. The program Flexible-Meccano (FM) assembles peptide units, considered as rigid entities, in a consecutive way [36]. This algorithm models the residue-specific Ramachandran space of the amino acids and roughly their side-chains. It has been tested on a large number of IDPs and has successfully described several NMR observations and SAXS data [37]. Another FM application example is the biophysical characterization of the transactivation domain of p53 [38]. Importantly, the same structural model was able to simultaneously describe the NMR results, which mainly report on the conformational (local) properties, and the SAXS curves that report on size and shape, (global) characteristics, of the proteins. This underlines the synergy between SAXS and NMR, as will be discussed in the section "Complementary methods".

3.2. Ensemble approach

Proteins feature a large spectrum of motional modes encompassing fast fluctuations around the average globular structure, slow large-scale molecular reorganizations, and the inherent disorder observed in IDPs. The development of realistic ensemble models of unstructured states of proteins has been an important subject of research for many years [39]. In recent years, this challenge has been addressed by ensembles of reliable conformations guided by experimental data that represent average values for the complete ensemble of conformations [40]. For SAXS applications several approaches have been developed to characterize protein mobility: The Ensemble Optimization Method (EOM) [41,42], the Minimal Ensemble Search (MES) [43], Basis-Set Supported SAXS (BSS-SAXS) [44], the Ensemble Refinement of SAXS (EROS) [45]. These methods are based on a common strategy containing two consecutive steps: (i) Computational generation of a large amount (thousands) of static random models of various possible conformations of a given protein, and (ii) selection of a sub-ensemble of conformations that together describe the experimental profile utilizing Eq. (7). The implementations and the optimization methods to select the ensemble may vary for different methods; here, the EOM approach is presented, which is the first ensemble analysis tool developed for SAXS, and it is well suited for the structural characterization of IDPs.

In the EOM, the experimental SAXS profile of a given flexible protein is assumed to derive from an (undetermined *a priori*) number of coexisting conformational states following Eq. (7). A sub-ensemble of 10–50 conformations is selected by a genetic algorithm from a large pool of scattering patterns (typically ten thousand different conformers, the scattering curves from each conformer being pre-computed using the program CRYSOL [46]). The scattering data from each ensemble is calculated by summing up the individual scattering patterns of each conformation, where

it is assumed that all conformers equally populate the subset. The sub-ensemble with the best agreement to the experimental data is collected for further structural analysis. Of course, an ensemble of 10–50 structures is far too small to explain the conformational behaviour of a flexible protein in solution but these subpopulations are usually sufficient to describe the global properties of the protein in terms of size and shape (i.e. R_g, D_{max} and anisometry), see Fig. 3.

The initial pool may not just consist of random models of single-chain unstructured proteins, and it is possible to follow for various scenarios accounting for the available *a priori* information. The generated models may include folded regions or domains or even nucleic acids if needed; for oligomeric structures, asymmetric or symmetric multimers can be generated. The latter approach was employed in the study of a flexible protein gephyrin which consists of two folded domains (of which one forms a trimer) connected by a 150 amino acid residues linker [48]. The possibility of combining multiple independent pools further expands the range of applicability of EOM: in a recent study of a full-length mitochondrial glutaminase C [49], pools consisting of dimers, trimers and octamers were combined to explain how disordered regions influence enzymatic activity and the oligomeric state.

3.3. Model validation

Importantly, SAXS provides the means for an *a posteriori* validation of the assumptions under which the flexible model is constructed. Using e.g. EOM as described above a model of flexibility is assumed in the pool generation, and the pool may or may not lead to an adequate fit to the experimental data. If the flexible model fits the data the properties of this model can be further analysed to make conclusions about the flexibility. If it does not fit then the data cannot be described by the model, for example because the model is not sufficiently flexible or the size/oligomeric state of the protein in question is incorrect. In this case, alternative assumptions about the system and/or its flexibility are required.

The quality of *in silico* models of IDP structures is directly correlated to their ability to reproduce the experimental data from disordered proteins. Several types of models with different levels of structural resolution have been compared to radii of gyration for denatured proteins [10,50–56]. In these studies a remarkable agreement with the experimental data was obtained. However, for most of the models an exclusion or solvation term had to be included, or the conformational sampling at residue level was required.

3.4. Detection of structural transitions in IDPs upon changing environments

IDPs are often involved in cellular signalling processes, since they structurally adapt to changes in their environment within the cell in order to bind or release binding partners. SAXS is a well-suited tool to rapidly monitor large-scale structural transitions in proteins upon such environmental changes. These global alterations influence the apparent R_g, D_{max}, and the appearance of the Kratky plots. Several studies have been reported to monitor structural transition upon changes in temperature [57,58], pH [59], ionic strength [60], presence of specific ions [61,62], additives [63,64], reducing agents [65,66], and post-translational modifications such as phosphorylation [53,67], sumoylation [68], and glycosilation [69]. Also, the effect of point mutations on the overall shape of proteins has been explored by SAXS [70,71].

An excellent example of a protein studied by SAXS under different environmental conditions is given by the analysis of prothymosin α at nearly neutral pH 7.5, and at acidic pH 2.5 [72]. The protein is predicted to have a global charge of -54 at neutral pH. Although the Kratky plots indicated that the protein was unstructured at both pH values, a dramatic R_g reduction by 10 Å was observed upon pH decrease. Interestingly, a similar level of compaction was observed upon the addition of zinc ions at neutral pH [73]. These changes in protein size can be explained by a reduced electrostatic repulsion within the chain at low pH, or by bound cations.

Two other interesting studies investigated the effect of temperature on the IDP structure where SAXS has been combined with other biophysical tools. Kjaergaard et al. [73] found a decreased compactness as shown in the R_g values of two IDPs, human NHE1 and ACTR, with increasing temperature from 5 to 45 °C. CD and NMR experiments in the same experimental conditions demonstrated that transient α-helices partially unfold upon increasing temperature. The effect of temperature on the structure of the protein Tau has been studied by Shkumatov et al. [74]. SAXS data on wild-type Tau and a mutant that mimics a phosphorylated state have shown no differences at 10 or 50 °C. However, after a

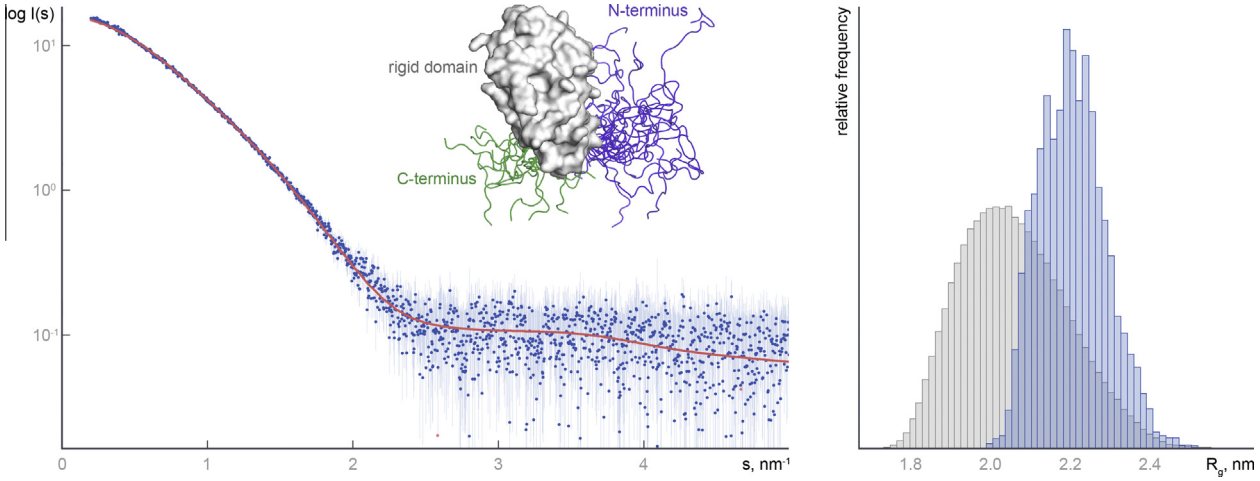

Fig. 3. *Left:* The theoretical scattering curve obtained for an ensemble of $BR_{187-385}$ monomer models (red) fits the SAXS profile of $BR_{187-385}$ (blue dots with error bars) with a χ^2-value of 0.58, SASBDB: SASDAR3. The radius of gyration obtained from the Guinier approximation was 2.27 ± 0.12 nm. *Inset:* Structural representations of the 18 models from the ensemble obtained by EOM analysis indicate high flexibility of the N-terminus (25 residues) and C-terminus (10 residues), which are drawn as thin lines in violet and green, respectively. The flexible terminal regions extend from the globular structure of $BR_{187-385}$ (170 residues), which is displayed as white molecular surface. Right: The initial pool of 10,000 models with random configurations for the N and C-termini extending the rigid domain of $BR_{187-385}$ has a broad distribution of R_g values with a peak maximum at 2.0 nm (grey bars). The selected ensemble is clearly biased towards extended structures with a mean R_g value of 2.2 nm (blue bars). © Schulte et al. [47].

fast temperature shift, from 10 to 50 °C or vice versa, the R_g values were notably smaller than in equilibrium. This observation indicated an increased compactness of Tau and its mutant that was corroborated by DLS experiments. Interestingly, the structural compaction was preserved for several hours after the temperature shift until it reached the conformational equilibrium. The authors speculate that this intriguing effect might be due to a structural memory of the protein related to the more compact nature of hyperphosphorylated Tau involved in Alzheimer's disease.

3.5. SAXS for the detection of IDP biomolecular interactions

IDPs play crucial roles in the cellular process regulation as binding partners for many regulatory switches (including proteins and nucleic acids). These interactions can be investigated by SAXS and complementary methods which is highlighted by two application examples.

The complex of Msh2 and Msh6 recognizes mismatched bases in DNA during mismatch repair. Shell et al. showed that the intrinsically disordered N-terminal region of Msh6 binds PCNA, a homotrimeric protein that controls the progress of DNA polymerases [21]. However, comparison of the R_g, Kratky plots and $p(r)$ functions of the isolated proteins and the complex showed that upon PCNA binding Msh6 remains mainly disordered and proteolytically accessible. Next, the authors investigated the interaction of the Msh2–Msh6 complex with PCNA by SAXS. This complex could be described as a highly flexible dumbbell where both globular domains are tethered by the N-terminal Msh6 fragment that acts as a molecular leash. A biologically active deletion mutant of Msh6 with a notably shorter N-terminal tail confirmed important size changes upon binding which was easily monitored by the $p(r)$ and D_{max} derived from the SAXS profiles.

The tumour suppressor p53 is a multifunctional protein that plays a crucial role in processes like apoptosis and DNA repair. P53 is a homotetramer containing unstructured regions that represent 37% of the whole sequence and can thus be considered as a highly flexible protein. Rigid-body modelling of p53 SAXS data suggests that the protein is an open cross-like tetramer, which collapses to tightly embrace DNA upon complex formation [66]. Interestingly, the SAXS model of the nucleoprotein without the disordered N-terminal transactivation domains excellently fits an independent electron microscopy map of the same complex. A higher resolution model of the isolated disordered transactivation domain and together with the full-length protein upon complexation was obtained by the joint use of SAXS with NMR data [38].

4. Complementary techniques

Conformational variation is a general characteristic of different types of proteins. This implies that biologically meaningful structural analyses should comprise investigation of both flexibility, as seen by SAXS, and detail, as determined by X-ray crystallography and NMR. For the study of intrinsically disordered and flexible proteins several techniques provide valuable complementary information to SAXS:

(i) *In silico* analysis of the primary sequence may predict structural disorder, since disordered regions feature more specific disorder-promoting amino acids, such as glycine, proline and charged residues, while lacking order-promoting, mostly hydrophobic, amino acids [75,76].

(ii) Protein X-ray crystallography on folded subdomains in flexible and intrinsically disordered proteins yields building blocks for the SAXS modelling. The combination of SAXS and protein crystallography is of particular help, since SAXS may shed light on the regions that are invisible in the electron density maps. In a recent study from Schulte et al. [47] the crystal structure of the KRT10-binding region domain of the pneumococcal serine-rich repeat protein (PsrP) was determined. It was demonstrated that 30 terminal residues (out of 200) missing from the crystal structure are flexible and the experimental SAXS data could be fitted only with an ensemble of extended models.

(iii) SAXS and solution nuclear magnetic resonance (NMR) is a very effective combination to analyse the structure of multi-domain proteins [77] and biomolecular complexes [78,79]. In particular, the program ENSEMBLE derives ensembles of IDPs that collectively describe SAXS curves in addition of several NMR observables [80,81]. The EOM approach of SAXS can also be synergistically combined with NMR, since SAXS provides information about the relative interdomain position in flexible multidomain proteins, whereas NMR parameters such as residual dipolar couplings (RDC) and pseudo-contact shifts (PCS) are sensitive to the relative orientation of the domains [82].

(iv) Other complementary techniques include circular dichroism (CD), fluorescence spectroscopy, and hydrodynamic techniques: size exclusion chromatography (SEC), AUC, DLS. Since the pioneering study of prothymosin α [83], several IDPs have been biophysically characterized in combination with scattering techniques, including viral proteins [84], or partially folded proteins [9,36,58,64,67,68,70,85–91]. In all these cases, SAXS data revealed rather large R_g values and expanded $p(r)$ functions were obtained with the D_{max} values much larger than those for folded particles. Also, in the case of partially folded proteins, the Kratky plots featured a dual behaviour with a clear maximum, corresponding to the folded part of the protein, and a continuous rise at higher angles due to the presence of disordered regions, which indicate highly flexible regions attached to globular domains. The overall appearance of the Kratky plots is expectedly defined by the proportion of the amino acids belonging to the structurally stable domains and to the unstructured domains. Using synthetic data it was shown that decreased structural features in SAXS profiles in combination with a single broad maximum in the Kratky plot indicate loosening of the structure, when moving from static scenarios to highly dynamic ones [92]. These observations are in agreement with several experimental studies of multi-domain proteins [58,93].

5. Conclusions and perspectives

Flexible molecules like IDPs are very difficult objects to be studied by high resolution methods. SAXS, a rapid method applicable under a broad variety of experimental conditions, is a valuable tool to characterize the low resolution structure of various flexible systems in solution. In the past, only average parameters like radius of gyration were extracted from the SAXS data of IDPs, accompanied by qualitative assessment of disorder by the Kratky plots. Although useful information was provided even with this basic analyses, novel approaches exploring the configurational space of flexible systems and explicitly allowing for a co-existence of multiple conformers in solution, pawed the way for a much more comprehensive use of SAXS. As demonstrated in the above examples, the concept of ensemble analysis makes SAXS an even more effective tool for the characterization of structure and dynamics of IDPs in solution. Especially useful is the joint application of SAXS with molecular dynamics simulations and also with other techniques providing local structural information, like NMR, CD, FRET or, when

applicable, protein crystallography on the folded domains. Importantly, the SAXS results including ensembles for flexible systems can now be deposited in publically available databases like SASBDB [94] or PED [95]. This opens the way for a more active dissemination of the results and for the use of SAXS-generated data and models by a broad biological community.

6. Summary points

- The overall parameters of proteins in solution, such as the radius of gyration, volume, molecular mass, and folding state, can be directly computed from SAXS data.
- SAXS can provide quantitative structural information about flexible, unfolded and intrinsically disordered proteins (IDPs).
- SAXS can be synergistically combined with complementary techniques such as *in silico* sequence data, X-ray crystallography and, in particular for flexible systems, NMR.
- Advances in SAXS instrumentation in the last years allow for high-throughput structural sample analysis with increasing speed of both data collection and sample characterization.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgement

This work was supported by the EU FP7 Infrastructures Grant BioStructX (contract 283570).

References

[1] Kohn, J.E. et al. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. U.S.A. 101, 12491–12496.
[2] Wilkins, D.K., Grimshaw, S.B., Receveur, V., Dobson, C.M., Jones, J.A. and Smith, L.J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. Biochemistry 38, 16424–16431.
[3] Stumpe, M.C. and Grubmuller, H. (2007) Interaction of urea with amino acids: implications for urea-induced protein denaturation. J. Am. Chem. Soc. 129, 16126–16131.
[4] Galantini, L., Leggio, C., Konarev, P.V. and Pavel, N.V. (2010) Human serum albumin binding ibuprofen: a 3D description of the unfolding pathway in urea. Biophys. Chem. 147, 111–122.
[5] Kojima, M., Tanokura, M., Maeda, M., Kimura, K., Amemiya, Y., Kihara, H. and Takahashi, K. (2000) PH-dependent unfolding of aspergillopepsin II studied by small-angle X-ray scattering. Biochemistry 39, 1364–1372.
[6] Pollack, L. and Doniach, S. (2009) Time-resolved X-ray scattering and RNA folding. Methods Enzymol. 469, 253–268.
[7] Round, A.R. et al. (2008) Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. J. Appl. Crystallogr. 41, 913–917.
[8] Hura, G.L. et al. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods 6, 606–612.
[9] Aslam, M. and Perkins, S.J. (2001) Folded-back solution structure of monomeric factor H of human complement by synchrotron X-ray and neutron scattering, analytical ultracentrifugation and constrained molecular modelling. J. Mol. Biol. 309, 1117–1138.
[10] Jha, A.K., Colubri, A., Freed, K.F. and Sosnick, T.R. (2005) Statistical coil model of the unfolded state: resolving the reconciliation problem. Proc. Natl. Acad. Sci. U.S.A. 102, 13099–13104.
[11] Fukuchi, S., Hosoda, K., Homma, K., Gojobori, T. and Nishikawa, K. (2011) Binary classification of protein molecules into intrinsically disordered and ordered segments. BMC Struct. Biol. 11, 29.
[12] Uversky, V.N. and Dunker, A.K. (2010) Understanding protein non-folding. Biochim. Biophys. Acta 1804, 1231–1264.
[13] Chiti, F. and Dobson, C.M. (2006) Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. 75, 333–366.
[14] Uversky, V.N., Oldfield, C.J. and Dunker, A.K. (2008) Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu. Rev. Biophys. 37, 215–246.
[15] Dyson, H.J. and Wright, P.E. (2005) Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 6, 197–208.
[16] Tompa, P. (2002) Intrinsically unstructured proteins. Trends Biochem. Sci. 27, 527–533.
[17] Tompa, P. (2011) Unstructural biology coming of age. Curr. Opin. Struct. Biol. 21, 419–425.
[18] Dunker, A.K., Silman, I., Uversky, V.N. and Sussman, J.L. (2008) Function and structure of inherently disordered proteins. Curr. Opin. Struct. Biol. 18, 756–764.
[19] Tompa, P. and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem. Sci. 33, 2–8.
[20] Yang, C., van der Woerd, M.J., Muthurajan, U.M., Hansen, J.C. and Luger, K. (2011) Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2–nucleosome complexes. Nucleic Acids Res. 39, 4122–4135.
[21] Shell, S.S., Putnam, C.D. and Kolodner, R.D. (2007) The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA. Mol. Cell 26, 565–578.
[22] Mertens, H.D. and Svergun, D.I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. J. Struct. Biol. 172, 128–141.
[23] Jacques, D.A. and Trewhella, J. (2010) Small-angle scattering for structural biology–expanding the frontier while avoiding the pitfalls. Protein Sci. 19, 642–657.
[24] Franke, D., Kikhney, A.G. and Svergun, D.I. (2012) Automated acquisition and analysis of small angle X-ray scattering data. Nuclear Instrum. Methods Phys. Res. A 689, 52–59.
[25] Jeffries, C.M., Graewert, M.A., Svergun, D.I. and Blanchet, C.E. (2015) Limiting radiation damage for high-brilliance biological solution scattering: practical experience at the EMBL P12 beamline PETRAIII. J. Synchrotron Radiat. 22, 273–279.
[26] A. Guinier, La diffraction des rayons X aux tr'es petits angles: application 'a l'étude de phénom'enes ultramicroscopiquesed, Univ. de Paris, Paris, 1939, pp. 3, p.l., 80, 2, p. 1 l.
[27] Calmettes, P., Durand, D., Desmadril, M., Minard, P., Receveur, V. and Smith, J.C. (1994) How random is a highly denatured protein? Biophys. Chem. 53, 105–113.
[28] Glatter, O. (1977) Data evaluation in small-angle scattering – calculation of radial electron-density distribution by means of indirect fourier transformation. Acta Phys. Aust. 47, 83–102.
[29] Svergun, D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495–503.
[30] Perez, J., Vachette, P., Russo, D., Desmadril, M. and Durand, D. (2001) Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering. J. Mol. Biol. 308, 721–743.
[31] Bernado, P. and Blackledge, M. (2009) A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. Biophys. J. 97, 2839–2845.
[32] Flory, P.J. (1953) Principles of Polymer Chemistry, Cornell University Press, Ithaka, NY.
[33] Bernado, P. and Svergun, D.I. (2012) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol. Biosyst. 8, 151–167.
[34] Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS – a Windows-PC based system for small-angle scattering data analysis. J. Appl. Crystallogr. 36, 1277–1282.
[35] Petoukhov, M.V. and Svergun, D.I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys. J. 89, 1237–1250.
[36] Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W. and Blackledge, M. (2005) A structural model for unfolded proteins from residual dipolar couplings and small-angle X-ray scattering. Proc. Natl. Acad. Sci. U.S.A. 102, 17002–17007.
[37] Jensen, M.R., Markwick, P.R., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernado, P. and Blackledge, M. (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. Structure 17, 1169–1185.
[38] Wells, M. et al. (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc. Natl. Acad. Sci. U.S.A. 105, 5762–5767.
[39] Zhou, H.X. (2004) Polymer models of protein stability, folding, and interactions. Biochemistry 43, 2141–2154.
[40] Bernado, P. and Blackledge, M. (2010) Structural biology: proteins in dynamic equilibrium. Nature 468, 1046–1048.
[41] Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M. and Svergun, D.I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J. Am. Chem. Soc. 129, 5656–5664.
[42] Tria, G., Mertens, H.D.T., Kachala, M. and Svergun, D.I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. Iucrj 2, 207–217.
[43] Pelikan, M., Hura, G.L. and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen. Physiol. Biophys. 28, 174–189.
[44] Yang, S., Blachowicz, L., Makowski, L. and Roux, B. (2010) Multidomain assembled states of Hck tyrosine kinase in solution. Proc. Natl. Acad. Sci. U.S.A. 107, 15757–15762.
[45] Rozycki, B., Kim, Y.C. and Hummer, G. (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. Structure 19, 109–116.
[46] Svergun, D.I., Barberato, C. and Koch, M.H.J. (1995) CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. 28, 768–773.
[47] Schulte, T. et al. (2014) The basic keratin 10-binding domain of the virulence-associated pneumococcal serine-rich protein PsrP adopts a novel MSCRAMM fold. Open Biol. 4, 130090.

[48] Sander, B., Tria, G., Shkumatov, A.V., Kim, E.Y., Grossmann, J.G., Tessmer, I., Svergun, D.I. and Schindelin, H. (2013) Structural characterization of gephyrin by AFM and SAXS reveals a mixture of compact and extended states. Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 2050–2060.

[49] Moller, M. et al. (2013) Small angle X-ray scattering studies of mitochondrial glutaminase C reveal extended flexible regions, and link oligomeric state with enzyme activity. PLoS One 8, e74783.

[50] Zhou, R. and Silverman, B.D. (2002) Detecting native protein folds among large decoy sets with hydrophobic moment profiling. Pac. Symp. Biocomput., 673–684.

[51] Goldenberg, D.P. (2003) Computational simulation of the statistical properties of unfolded proteins. J. Mol. Biol. 326, 1615–1633.

[52] Fitzkee, N.C. and Rose, G.D. (2004) Reassessing random-coil statistics in unfolded proteins. Proc. Natl. Acad. Sci. U.S.A. 101, 12497–12502.

[53] Tran, H.T., Wang, X. and Pappu, R.V. (2005) Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins. Biochemistry 44, 11369–11380.

[54] Ding, F., Jha, R.K. and Dokholyan, N.V. (2005) Scaling behavior and structure of denatured proteins. Structure 13, 1047–1054.

[55] Wang, Z., Plaxco, K.W. and Makarov, D.E. (2007) Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins. Biopolymers 86, 321–328.

[56] Wang, Y., Trewhella, J. and Goldenberg, D.P. (2008) Small-angle X-ray scattering of reduced ribonuclease A: effects of solution conditions and comparisons with a computational model of unfolded proteins. J. Mol. Biol. 377, 1576–1592.

[57] Baker, J.M., Hudson, R.P., Kanelis, V., Choy, W.Y., Thibodeau, P.H., Thomas, P.J. and Forman-Kay, J.D. (2007) CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. Nat. Struct. Mol. Biol. 14, 738–745.

[58] Pretto, D.I., Tsutakawa, S., Brosey, C.A., Castillo, A., Chagot, M.E., Smith, J.A., Tainer, J.A. and Chazin, W.J. (2010) Structural dynamics and single-stranded DNA binding activity of the three N-terminal domains of the large subunit of replication protein A from small angle X-ray scattering. Biochemistry 49, 2880–2889.

[59] Konno, T., Tanaka, N., Kataoka, M., Takano, E. and Maki, M. (1997) A circular dichroism study of preferential hydration and alcohol effects on a denatured protein, pig calpastatin domain I. Biochim. Biophys. Acta 1342, 73–82.

[60] Munishkina, L.A., Fink, A.L. and Uversky, V.N. (2004) Conformational prerequisites for formation of amyloid fibrils from histones. J. Mol. Biol. 342, 1305–1324.

[61] He, G., Ramachandran, A., Dahl, T., George, S., Schultz, D., Cookson, D., Veis, A. and George, A. (2005) Phosphorylation of phosphophoryn is crucial for its function as a mediator of biomineralization. J. Biol. Chem. 280, 33109–33114.

[62] Binolfi, A., Rasia, R.M., Bertoncini, C.W., Ceolin, M., Zweckstetter, M., Griesinger, C., Jovin, T.M. and Fernandez, C.O. (2006) Interaction of alpha-synuclein with divalent metal ions reveals key differences: a link between structure, binding specificity and fibrillation enhancement. J. Am. Chem. Soc. 128, 9893–9901.

[63] Hong, D.P., Fink, A.L. and Uversky, V.N. (2008) Structural characteristics of alpha-synuclein oligomers stabilized by the flavonoid baicalein. J. Mol. Biol. 383, 214–223.

[64] Hammel, M., Yu, Y., Fang, S., Lees-Miller, S.P. and Tainer, J.A. (2010) XLF regulates filament architecture of the XRCC4.ligase IV complex. Structure 18, 1431–1442.

[65] Lanza, D.C., Silva, J.C., Assmann, E.M., Quaresma, A.J., Bressan, G.C., Torriani, I.L. and Kobarg, J. (2009) Human FEZ1 has characteristics of a natively unfolded protein and dimerizes in solution. Proteins 74, 104–121.

[66] Tidow, H. et al. (2007) From the cover: quaternary structures of tumor suppressor p53 and a specific p53 DNA complex. Proc. Natl. Acad. Sci. U.S.A. 104, 12324–12329.

[67] Paoletti, F. et al. (2009) Intrinsic structural disorder of mouse proNGF. Proteins 75, 990–1009.

[68] Nardini, M. et al. (2006) The C-terminal domain of the transcriptional corepressor CtBP is intrinsically unstructured. Protein Sci. 15, 1042–1050.

[69] Kingston, R.L., Hamel, D.J., Gay, L.S., Dahlquist, F.W. and Matthews, B.W. (2004) Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc. Natl. Acad. Sci. U.S.A. 101, 8301–8306.

[70] Kim, H.S. et al. (2011) Different modes of interaction by TIAR and HuR with target RNA and DNA. Nucleic Acids Res. 39, 1117–1130.

[71] Li, J., Uversky, V.N. and Fink, A.L. (2002) Conformational behavior of human alpha-synuclein is modulated by familial Parkinson's disease point mutations A30P and A53T. Neurotoxicology 23, 553–567.

[72] Uversky, V.N. et al. (1999) Natively unfolded human prothymosin alpha adopts partially folded collapsed conformation at acidic pH. Biochemistry 38, 15009–15016.

[73] Uversky, V.N. et al. (2000) Zn(2+)-mediated structure formation and compaction of the "natively unfolded" human prothymosin alpha. Biochem. Biophys. Res. Commun. 267, 663–668.

[74] Shkumatov, A.V., Chinnathambi, S., Mandelkow, E. and Svergun, D.I. (2011) Structural memory of natively unfolded tau protein detected by small-angle X-ray scattering. Proteins 79, 2122–2131.

[75] Dunker, A.K. et al. (2001) Intrinsically disordered protein. J. Mol. Graph. Model. 19, 26–59.

[76] Uversky, V.N., Gillespie, J.R. and Fink, A.L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins 41, 415–427.

[77] Grishaev, A., Wu, J., Trewhella, J. and 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.

[78] Wang, J. et al. (2009) Determination of multicomponent protein structures in solution using global orientation and shape restraints. J. Am. Chem. Soc. 131, 10507–10515.

[79] Schwieters, C.D., Suh, J.Y., Grishaev, A., Ghirlando, R., Takayama, Y. and Clore, G.M. (2010) Solution structure of the 128 kDa enzyme I dimer from *Escherichia coli* and its 146 kDa complex with HPr using residual dipolar couplings and small- and wide-angle X-ray scattering. J. Am. Chem. Soc. 132, 13026–13045.

[80] Marsh, J.A., Neale, C., Jack, F.E., Choy, W.Y., Lee, A.Y., Crowhurst, K.A. and Forman-Kay, J.D. (2007) Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. J. Mol. Biol. 367, 1494–1510.

[81] Marsh, J.A. and Forman-Kay, J.D. (2009) Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. J. Mol. Biol. 391, 359–374.

[82] Bertini, I., Giachetti, A., Luchinat, C., Parigi, G., Petoukhov, M.V., Pierattelli, R., Ravera, E. and Svergun, D.I. (2010) Conformational space of flexible biological macromolecules from average data. J. Am. Chem. Soc. 132, 13553–13558.

[83] Gast, K. et al. (1995) Prothymosin alpha: a biologically active protein with random coil conformation. Biochemistry 34, 13211–13218.

[84] Jensen, M.R., Houben, K., Lescop, E., Blanchard, L., Ruigrok, R.W. and Blackledge, M. (2008) Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: application to the molecular recognition element of Sendai virus nucleoprotein. J. Am. Chem. Soc. 130, 8055–8061.

[85] Schmidt, C.Q. et al. (2010) The central portion of factor H (modules 10–15) is compact and contains a structurally deviant CCP module. J. Mol. Biol. 395, 105–122.

[86] Moncoq, K., Broutin, I., Craescu, C.T., Vachette, P., Ducruix, A. and Durand, D. (2004) SAXS study of the PIR domain from the Grb14 molecular adaptor: a natively unfolded protein with a transient structure primer? Biophys. J. 87, 4056–4064.

[87] Lima, L.M. et al. (2006) Structural insights into the interaction between prion protein and nucleic acid. Biochemistry 45, 9180–9187.

[88] Shi, X. et al. (2010) Identification and biophysical assessment of the molecular recognition mechanisms between the human haemopoietic cell kinase Src homology domain 3 and ALG-2-interacting protein X. Biochem. J. 431, 93–102.

[89] Garcia-Pino, A. et al. (2010) Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. Cell 142, 101–111.

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

[91] Elliott, P.R., Goult, B.T., Kopp, P.M., Bate, N., Grossmann, J.G., Roberts, G.C., Critchley, D.R. and Barsukov, I.L. (2010) The structure of the talin head reveals a novel extended conformation of the FERM domain. Structure 18, 1289–1299.

[92] Bernadó, P. (2010) Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. Eur. Biophys. J. 39, 769–780.

[93] Edwards, R.A., Lee, M.S., Tsutakawa, S.E., Williams, R.S., Nazeer, I., Kleiman, F. E., Tainer, J.A. and Glover, J.N. (2008) The BARD1 C-terminal domain structure and interactions with polyadenylation factor CstF-50. Biochemistry 47, 11446–11456.

[94] Valentini, E., Kikhney, A.G., Previtali, G., Jeffries, C.M. and Svergun, D.I. (2015) SASBDB, a repository for biological small-angle scattering data. Nucleic Acids Res. 43. D357–D363.

[95] Varadi, M. et al. (2014) PE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. Nucleic Acids Res. 42, D326–D335.