Published in final edited form as:

Biopolymers. 2011 August ; 95(8): 559–571. doi:10.1002/bip.21638.

Characterizing Flexible and Intrinsically Unstructured Biological Macromolecules by SAS using the Porod-Debye Law

Robert P. Rambo^{1,†} and John A. Tainer^{1,2,†}

¹Life Sciences Division, Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

²Department of Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Abstract

Unstructured proteins, RNA or DNA components provide functionally important flexibility that is key to many macromolecular assemblies throughout cell biology. As objective, quantitative experimental measures of flexibility and disorder in solution are limited, small angle scattering (SAS), and in particular small angle X-ray scattering (SAXS), provides a critical technology to assess macromolecular flexibility as well as shape and assembly. Here, we consider the Porod-Debye law as a powerful tool for detecting biopolymer flexibility in SAS experiments. We show that the Porod-Debye region fundamentally describes the nature of the scattering intensity decay, which captures information needed for distinguishing between folded and flexible particles. Particularly for comparative SAS experiments, application of the law, as described here, can distinguish between discrete conformational changes and localized flexibility relevant to molecular recognition and interaction networks. This approach aids insightful analyses of fully and partly flexible macromolecules that is more robust and conclusive than traditional Kratky analyses. Furthermore, we demonstrate for prototypic SAXS data that the ability to calculate particle density by the Porod-Debye criteria, as shown here, provides an objective quality assurance parameter that may prove of general use for SAXS modeling and validation.

INTRODUCTION

Our view of the nature of bio-macromolecular function has been greatly enhanced by recent observations that many complex biological processes require structural flexibility within the biological machine. Flexibility contributes to the dynamics of the macromolecular particle over large (delocalized) and small (localized) scales that can be characterized by solution methods such as NMR and SAS^{1,2}. Many cellulases, DNA repair proteins, transcription factors and extracellular matrix proteins contain long flexible linkers that tether a core domain to one or more catalytic or specificity domains^{3–5}. This type of flexibility enhances the functional dynamics of the macromolecule thereby delocalizing the functional domain over a large volume. Particularly for DNA repair proteins (*e.g.*, DNAPK, Nbs1, ligase III, RPA, Mre11 polynucleotide kinase, and XPF) flexible extensions allow the core DNA damage recognition domain to tightly bind a damaged DNA end while recruiting the requisite repair proteins to the site of DNA damage^{6–10}. In contrast, localized flexibility can functionally provide a macromolecular domain with the ability to switch between distinct conformational states or destabilize in the absence of a signaling molecule, as seen with many ATPases^{11–14}.

[†]To whom correspondence should be addressed., Telephone: 858.784.8438, Fax: 858.784.2289, jat@scripps.edu.

Objective evaluation of macromolecular flexibility is of great scientific and practical significance for biology and medicine. We know that the binding of the correct metal ion in many proteins and RNA is key for domain folding and that most metalloproteins and structured RNAs remain uncharacterized, so an efficient means to examine solution structures of biopolymers in the presence of metals is of great value^{15,16}. Aberrant flexibility from either metal ion absence or mutation can cause disease by reducing specificity of fold and assembly, as seen for reactive oxygen control enzyme superoxide dismutase and DNA repair helicase XPD^{12,17,18}. Furthermore, flexibility can be induced. The binding of the RNA chaperone DEAD-box protein Mss116 to the ai5 γ group II intron RNA assists the RNA during folding by promoting dynamic sampling of folding states thereby avoiding a kinetic trap¹⁹.

Structural information in cell biology includes shapes, flexibility and assemblies with many biopolymers mimicking the shape of an unrelated biopolymer. Such shape mimicry occurs among protein modification domains, such as SUMO and even across biopolymer classes, such as protein mimicry of DNA or RNA^{20–22}. The efficiency and robustness of SAS is creating a renaissance in structural biology by providing three-dimensional models based on the X-ray and neutron scattering data from macromolecules in solution^{1,23}. These models are shapes that can be quickly compared with known shapes to gain insights, such as mimicry. Furthermore, advanced SAXS synchrotron technologies now provide the throughput for systematic examination of macromolecular interactomes in pathways and networks²⁴. With the development of SAXS technologies combined with other results for defining accurate macromolecular structures, conformations and assemblies in solution, SAS is poised to address many important challenges for biological systems²⁵. As a solution technique, a major potential advantage of SAS is the ability to assess flexibility and to even characterize unstructured to structured transitions for biopolymers under near physiological conditions in solution. Detecting conformational switching, destabilization or long-range delocalized flexibility in solution can be made using small-angle X-ray scattering (SAXS), typically through qualitative assessments of the scattering data in a Kratky plot²⁶. Here, we analyze the behavior of flexible particles in light of what is known as the Porod-Debye law and show that application of the law provides a new and reliable means for identifying flexibility in small angle scattering (SAS) experiments.

RESULTS

Bio-SAXS FUNDAMENTALS

In most biological SAXS studies, the observed particle scattering is assumed to occur from a discrete electron density contrast between the solvent and particle²⁶. This assumption holds true for folded biopolymers such as proteins, RNA or mixed nucleic acid-protein complexes where, at low resolution, the average electron density of the particle is greater than that of the solvent. If we consider that the scattering angle, q ($\text{\AA}^{-1}$), defines a resolution ruler ($d = 2\pi / q$), then at sufficiently low resolution (as $q \rightarrow 0$) the scattering will be described by an average effective electron density contrast ($\Delta\rho$) between the particle and solvent that becomes increasingly difficult to discern at higher resolution similar to a close visual inspection of an impressionism painting (Figure 1). The existence of this contrast typically allows for the determination of $I(0)$, the scattering intensity at $q=0$, and the particle's radius-of-gyration, R_g , using the Guinier approximation²⁷ (Figure 2A). $I(0)$ relates directly to the particle's volume (V) and $\Delta\rho$ scaled by concentration, c . This relationship between $I(0)$ and V serves as the basis for determining the mass of the scattering particle against a set of known standards^{28,29}. The assumption of a discrete scattering contrast can be challenged by flexible systems, such that in the extreme case for a fully unfolded biopolymer, the scattering contrast is continuous making it difficult to clearly distinguish between the particle and solvent at higher resolution.

In the discrete contrast case (folded), G. Porod noted that the scattering data transformed as $q^2 I(q)$ vs. q , should capture a defined area²⁶ (Figure 2B), known as the Porod invariant, Q (Figure 2B). Q is a direct measurement of $\Delta\rho$ that is scaled by the particle's V and c . The determination of Q can, in theory, be used to determine several structural parameters such as the particle's V , surface-to-volume ratio and correlation length, l_c (Figure 2C)^{26,30}.

For flexible particles, the ability to derive the structural values from SAS data becomes increasingly difficult depending on the degree to which the particle behaves as a random coil. If the particle is composed of a flexible polymeric extension with a single well-folded domain, then at sufficient q , the scattering should be due to two different electron density contrasts. This behavior can be observed in a Kratky plot where the transformed curve will be described by an initial parabolic peak followed by an elevated baseline at high q (Figure 2B). Further elevation of the baseline towards a hyperbolic-like curve will be observed for particles behaving as a random coil or Gaussian chain in solution²⁶ (Figure 2B). These visual features in a Kratky plot have provided a means to qualitatively assess the degree of flexibility within the scattering particle. Furthermore, a dimensionless Kratky plot has been proposed as an alternative method for comparing macromolecules of differing masses and conformational states³¹ (Figure 3). This method normalizes the scattering data to $I(0)$ and skews the scattering angle by R_g (Figure 3C), such that the resulting SAS data in a Kratky plot is normalized for c , V and R_g thereby allowing for a comparison of scattering profiles on a common scale. The dimensionless Kratky plot will be sensitive to the accuracy in the determination of the Guinier region (Figure 3D). These observations highlight the limitations of the Kratky plot as a tool for assessing flexibility requiring SAS data to be collected to sufficiently high q from particles that are sufficiently flexible to diverge the baseline.

THE POROD-DEBYE FOURTH POWER LAW

The Porod-Debye law describes a fourth power law approximation to the relationship between q and the observed intensities, $I(q)$ ^{32,33}. Similar to the Guinier approximation, the Porod-Debye law is an approximation that occurs within a limited range of scattering angles and suggests that the scattering of a folded particle should decay as q^{-4} scaled by the particle's surface area (S), c and $\Delta\rho$ (Figure 2C). Thus, transformation of the scattering data as $q^4 I(q)$ vs. q or $q^4 I(q)$ vs. q^4 (Porod-Debye plot) should display a curve asymptotically approaching a constant value as q approaches infinity. The Porod-Debye law can be used to extract limited information from the SAXS curve, namely S , and dividing by the Porod invariant determines the surface-to-volume ratio of the particle (Figure 2C). In practice, the asymptotic relationship is rarely observed, as data at higher q contains significant shape information thereby destroying the asymptotic behavior predicted by the Porod-Debye law³⁴. In addition, experimental buffer subtracted SAXS data contains contributions from the internal fluctuations of both the particle and its excluded volume²⁶. Though insignificant at low angles, these contributions over-estimate Q such that the determination of V , l_c , or S/V ratio will be underestimated. Based on the difficulty of applying the Porod-Debye law and the evaluation of Q , it is not surprising that the aforementioned parameters are rarely reported from a SAXS study of biological particles!

If we consider that the Porod-Debye law predicts a behavior within a limited range of the SAXS data (akin to the linear relationship predicted by Guinier ($q \cdot R_g < 1.3$) for the determination of R_g), then application of the law predicts a plateau within the low resolution region of the SAXS data when transformed by q^4 . The presence of the Porod-Debye plateau validates the assumptions that the scattering particle exists with a sharp homogenous electron density contrast between the particle and solvent. Limiting the determination of Q

to the Porod-Debye region of the data (e.g. using the program PRIMUS³⁵) can now provide a reliable determination of the particle's macromolecular volume.

Experimentally, this behavior can be most readily demonstrated for glucose isomerase (GI), a nearly spherical tightly packed particle (Figure 4A). Transforming the entire SAXS dataset ($0 < q < 0.32 \text{ \AA}^{-1}$) as $q^4 I(q)$ vs. q shows a curve with significant shape-dependent oscillatory features (Figure 4B). By noting that the Porod-Debye law resides within a limited region of the data, the plateau can be identified in a truncated dataset that removes most shape dependent scattering (Figure 4C, D). SAXS measurements of GI using 8 different concentrations determine an average V of $210,000 (\pm 5,000) \text{ \AA}^3$, which imply a protein density (d_{protein}) of $1.36 (\pm 0.04) \text{ gm}\cdot\text{cm}^{-3}$, in excellent agreement with the mean empirical density of proteins taken as 1.35 to $1.37 \text{ gm}\cdot\text{cm}^{-3}$ ^{36,37}. Applying the Porod-Debye law to the appropriate range of scattering data now provides additional opportunities for interpreting SAXS experiments.

THE MACROMOLECULAR VOLUME IN SOLUTION

GI is a prototypical scattering particle that is ideal with a nearly spherical tightly packed core displaying little flexibility in solution. The macromolecular volume is simply the sum of the particle's displaced volume and associated hydration shell. For flexible particles in solution, this concept of the macromolecular volume must be expanded to include the volumetric space occupied by the flexible domain. Consequently, the observed d_{protein} is expected to decrease reflecting the particle's reduced packing density. This relationship is demonstrated using SAXS data from an intrinsically unfolded protein. Experiments on RAD51AP1 (36 kDa) confirm the protein to be intrinsically unfolded with no defined Porod-Debye plateau (Figure 5A,B). Attaching this domain to a well-folded protein, maltose-binding protein (MBP) from *E. coli*, created a biphasic 80 kDa scattering particle for SAXS. Inspection of the Kratky and Porod plots reveal a particle with a defined Porod-Debye plateau and an associated V of $145,000 \text{ \AA}^3$, consistent with a 120 kDa protein using d_{protein} of 1.37 (Figure 5B). The attachment of the intrinsically unfolded domain to MBP had the effect of increasing the apparent macromolecular volume. Therefore, using a d_{protein} of 1.37 over-estimated the mass of the scattering particle, where the actual d_{protein} is $0.92 \text{ gm}\cdot\text{cm}^{-3}$.

In addition, the determination of V for 12 previously reported proteins varying in size from 8 to 118 kDa with known flexible or 11 amino acid His-tagged N-terminal extensions (Table 1), determines a mean protein density that decreased by 23% from 1.37 to $1.09 \text{ gm}\cdot\text{cm}^{-3}$. In these protein samples, the presence of the Porod plateau and increased partial specific volume suggests the protein can be characterized by at least two scattering volumes in solution: 1) a well defined homogenous element composed of a folded domain, and 2) a diffuse scattering volume created by the flexible extension. This increase in the particle's apparent macromolecular volume would grossly overestimate the particle's true mass when assuming either the standard d_{protein} of 1.35 to $1.37 \text{ gm}\cdot\text{cm}^{-3}$ or when standardizing the $I(0)$ measurements using a protein standard such as glucose isomerase³⁸. Consequently, an overestimated mass would lead to erroneous hypotheses regarding the protein's multimeric or thermodynamic state.

DISTINGUISHING FLEXIBILITY AND DISCRETE CONFORMATIONAL CHANGE

Previous SAXS studies on the abscisic acid binding (ABA) protein PYR1 demonstrated conformational differences between the *bound* and *apo* forms of the protein in solution³⁹. In the absence of ABA, PYR1 in solution is nearly identical to the X-ray crystal structure

(PDB: 3K3K). Upon binding ABA, PYR1 experiences a decrease in R_g (23.71 to 22.72 Å) suggesting compaction of the protein (Figure 6A). Thus, it is expected that a protein undergoing a discrete conformational change must display a clearly defined Porod plateau for both states, whereas it is predicted that destabilization of the particle would subsequently invalidate the application of the Porod-Debye law resulting in a loss of the plateau. Examination of the Porod plots for PYR1 (Figure 6B) show distinct plateaus that correspond to a decreased volume (from 74,000 to 59,000 Å³) in the presence of ABA. Calculations of $d_{protein}$ reveal PYR1 goes from 0.99 to 1.18 gm·cm⁻³ consistent with compaction of the protein, where it is expected the extended loops surrounding the binding pocket collapse during the conformational rearrangement.

In contrast, SAXS studies on the S-adenosylmethionine (SAM)-1 riboswitch demonstrated conformational differences between the bound and unbound states due to an increased flexibility⁴⁰. The SAM-1 riboswitch binds the small molecule SAM and solution SAXS studies demonstrated the bound form is nearly identical to the bound crystallographic structure (PDB: 2GIS). In the absence of the metabolite SAM, the riboswitch experiences an increase in R_g (22.4 to 24.1 Å) (Figure 6C) leading to the hypothesis that the riboswitch becomes less compact⁴¹. Is this necessarily due to an alternate conformational state or an increase in flexibility? Inspection of the Porod plot shows a plateau for the bound form that is subsequently destroyed in the absence of the metabolite (Figure 6D). The loss of the plateau implies the scattering contrast has become diffuse suggesting an increase in the flexibility of the particle. More specifically, the flexibility is localized as the maximum dimension of particle changes by only a small distance (from 76 to 81 Å) between the two states⁴⁰ (Figure 6D).

Identifying flexibility by visual inspection of the Kratky plot may not always be effectively accomplished, especially with limited or weakly collected SAXS data as it may confound the baseline convergence necessary for assessing flexibility. Recent SAXS studies on the ATP dependent DNA repair complex, Mre11-Rad50, reveal distinct changes in the complex upon binding ATP⁴² (Figure 7A). Kratky analysis of the SAXS data suggested the bound and *apo* forms of Mre11-Rad50 are likely compact particles leading to the hypothesis that the particle is switching between two distinct conformational states, similar to PYR1. However, inspection of the Porod plot uncovers a different story. In the presence of ATP, the complex forms a distinct particle with a sharp scattering contrast, as evidenced by the Porod plateau (Figure 7B). The bound particle has a packing density of 1.04 gm·cm⁻³ that decreases significantly to 0.89 gm·cm⁻³ in the absence of ATP suggesting the particle becomes more flexible. In fact, inspection of the Porod-Debye region demonstrates a loss of the plateau supporting the assertion that Mre11-Rad50 is flexible in the absence of ATP.

The Porod-Debye law predicts well defined behaviors of the scattering particle that can be used to understand the nature of the structural changes observed in SAS experiments. In the first example with PYR1, the observed changes in the SAXS data were due to PYR1 switching between two distinct conformational states, as evidenced by the two distinct Porod plateaus for both the bound and free SAXS datasets. In contrast, SAXS experiments with the SAM-1 riboswitch and Mre11-Rad50 complex suggested the particles are flexible in the absence of their respective small molecules. In the bound states, a clear Porod plateau becomes established with a concomitant decrease in both R_g and V . Unlike the Kratky plot which requires SAXS data collected over a large angular range (at least $0.01 < q < 0.3 \text{ Å}^{-1}$), the Porod-Debye plot requires a limited set of data just outside the Guinier region for reliable and consistent interpretation.

PARTICLE DENSITY AS A COEFFICIENT FOR ESTIMATING SCATTERING MASS

Establishing the correct scattering mass of the particle has not been a strict prerequisite for structural modeling of SAXS data using current *ab-initio* or atomistic modeling methods. More importantly, under- or over-estimating the scattering mass will not necessarily generate worrisome values in the method's scoring function. In the most rigorous experiments, SAXS and multi-angle light scattering experiments are performed in tandem in order to accurately assess the polydispersity and mass of the scattering particle⁴¹. Such detailed analysis is especially important for multimeric particles where an accurate determination of mass can help choose the correct symmetry of the particle or guide the composition of the starting atomistic model. Since V is a SAXS derived parameter, which reflects the particle's macromolecular volume, a statistical survey of empirical SAXS-based protein densities can determine if a proposed particle mass is consistent with the observed SAXS data.

Using existing SAXS data from 31 different protein samples, a histogram of protein densities show the most commonly observed values are within 0.9 and 1.0 gm·cm⁻³, well below the canonical value of 1.37 (Figure 8). This is most likely due to the larger number of proteins with flexible or His-tagged extensions in our dataset. The next most likely observed density occurs within 1.3 and 1.4 gm·cm⁻³ encompassing the expected canonical value. With the exception that hollow-like structures may lower the packing densities below the empirically observed values, protein densities below 0.8 and above 1.5 are unlikely and may suggest under or over-modeling of the particle's expected mass.

Detecting under- and over-estimation of mass is illustrated with homomeric multi-subunit proteins GI, PYR1, WRN and catalase (Table 2). $d_{protein}$ calculations with a varying number of subunits, n , will show a set of values that fall within an acceptable range, indicating the likely number of subunits constituting the particle (Table 2). Consider GI, calculation of $d_{protein}$ using a 1 and 2 subunits result in $d_{protein}$ values that are unrealistically low suggesting the particle mass is under-estimated. Likewise, for 5 and 6 subunits, $d_{protein}$ is exceedingly high leaving the possible number of subunits at 3 or 4. Therefore, a combined *ab-initio* modeling with DAMMIN/F⁴³ using P1 symmetry and GASBOR⁴⁴ with both P3 and P4 symmetries will show one model, P3, underestimating the particle volume leaving P4 as the correctly assigned symmetry. In addition, SAXS data for IRECBP (Table 2) was previously modeled as a monomer using a previously determined X-ray crystal structure²⁴. A poor model-data agreement and further protein density calculations from the SAXS data strongly suggest the protein exists as a dimer or a mixed monomer-dimer in solution where density calculations with either the monomeric and trimeric forms of the protein produced unrealistic values. Thus, a direct check against the frequency of empirically determined values can help prevent under- or over-modeling of the SAXS data similar to the application of the Matthews coefficient in X-ray crystallographic structure determination of biological macromolecules⁴⁵.

DISCUSSION

Together with NMR, X-ray crystallography and electron microscopy (EM), SAXS is becoming a powerful complementary biophysical technique because it deals robustly with flexible particles in the solution state^{10,44,46}. A limitation to the value of SAS characterizations has been objective tools for the validation of flexibility. The Porod-Debye law is an under-appreciated but powerful tool for detecting flexibility in SAS experiments of biopolymers. We show that a useful application of this law requires a subset of SAS data for reliable analysis that is more robust and conclusive than the traditional Kratky analysis.

The Porod-Debye region fundamentally describes the nature of the scattering intensity decay. The location of this region or plateau will vary depending on the size of the particle. In general for folded to partially folded particles, the Porod-Debye region will occur after the q corresponding to the first major peak identified in a Kratky plot. For small particles (< 10 KDa), the plateau is expected to occur at moderately higher q ($> 0.2 \text{ \AA}^{-1}$) necessitating the collection of inherently weak SAXS data at higher q values ($q_{\text{max}} \approx 0.4 \text{ \AA}^{-1}$). Therefore, analysis must be made on data collected at sufficiently high concentrations with an acceptable signal-to-noise ratio. For larger particles (> 10 KDa), the plateau is expected to occur within smaller q values compatible with most synchrotron beamline configurations.

Particularly for comparative SAS experiments, application of the law can distinguish between discrete conformational changes and localized flexibility allowing for a more insightful understanding of the biological particle. Furthermore, the ability to calculate the particle density using the Porod-Debye criteria allows for a direct quality assurance check during the modeling phase of the SAXS experiment. The practical implementation of the Porod-Debye law in SAS experiments of biopolymers is a tool for assessing flexibility and for validation of SAXS models.

Many biological systems have evolved around exploiting changes in flexibility. In DNA repair, enhanced flexibility of the DNA is being used as a test for DNA damage. For some DNA repair enzymes, which exhibit open to closed conformational changes upon damaged DNA binding, DNA damage recognition is coupled to specific deformation of the DNA that promotes the damage excision pathway⁴⁷⁻⁴⁹. Conformational flexibility can regulate the XPB helix recognition of DNA damage versus DNA transcription sites¹². Further conformational changes of repair complexes utilize delocalized flexibility to capture open assemblies to load onto DNA or to remodel stalled RNA polymerases to expose damaged DNA for repair^{50,51}. In other cases, DNA repair proteins use structured surfaces to sculpt damaged DNA to create structural intersections of different DNA repair pathways⁵². SAXS has been particularly efficient at detecting and defining the types of flexibility exhibited by the DNA repair proteins⁵³. This knowledge has enhanced our mechanistic insights and expands the experimental framework for understanding these systems *in vivo*.

Knowledge of localized flexibility can be used practically to design anti-peptide antibodies that will bind intact protein targets⁵⁴⁻⁵⁶. Similarly, the ability to experimentally define the flexibility of linked functional domains has many applications in biology and medicine, as seen for the creation of linked enzymes with long serum half-lives *in vivo*. For the proper functioning of biological pathways, unstructured biopolymers can undergo disorder to order transitions that can create large, yet dynamic interfaces. Identifying localized flexibility can provide the basis for inhibitor specificity by designing compounds that can further open the flexible regions of a binding pocket^{57,58}. Thus, biopolymer flexibility ranges across the spectrum from unstructured proteins to localized flexibility, which creates challenges for most biophysical methods, including crystallography and electron microscopy. As molecular modeling seeks to provide increasingly accurate models of protein or RNA domains at the atomic level, the concept of the structural model must include in solution information regarding conformational states and flexibility⁵⁹.

Looking forward, validation of SAS characterizations using parameters, such as the one presented here, add to our supramolecular understanding that links proteins or RNA to assemblies and pathways to biological outcomes. We can expect synergistic integration of SAXS structural results in solution with biochemical, genetic, and bioinformatics approaches to effectively create a structural interpretation of how shape and conformation join macromolecules in biological networks.

Materials and Methods

Except for glucose isomerase, all SAXS data used in this study was taken from BioIsis.net. The SAXS data in BioIsis.net defines q as the momentum transfer function: $(4\pi \sin\theta)/\lambda$ with λ and θ describing the wavelength (Å) and angle of the scattered radiation. For glucose isomerase, sample was prepared by dissolving 33 uL of protein crystals (Hampton Research, Aliso Viejo, CA) in 1400 uL of buffer A containing 40 mM HEPES (pH 8.0), 1% glycerol, 6 mM MgCl₂, 50 mM KCl, and 50 mM NaCl. Sample was incubated at 37 °C on a nutator for 30 minutes and concentrated to 180 uL before injecting on a KW-803 size exclusion column (Shodex, San Diego, CA) equilibrated in buffer A. Sample purification was monitored inline with an 18-angle light scattering instrument (Wyatt Technology, Santa Barbara, CA) where detector 12 was replaced with a quasi-elastic light scattering detector. The light scattering was connected in tandem to the Optilab rEX refractive index detector (Wyatt Technology, Santa Barbara, CA). Peak fraction at 2.3 mg/mL was taken for direct SAXS analysis (Beamline 12.3.1, Advanced Light Source, Berkeley, CA).

The determination of the Porod-Debye region and Porod volume was performed in several steps using the software package PRIMUS³⁵. For the determination of the Porod volume, it is essential that the initial analysis starts with the accurate determination of the Guinier region since the determination of the particle volume relies on $I(0)$. Furthermore, SAXS profiles should be truncated of erroneous data from both the beginning of the lowest q region and highest q region ($q \cdot R_g < 1.3$) such that visual inspection of the residuals during Guinier analysis produce an unbiased distribution. Any “smiling” or “frowning” of the residuals is indicative of an improper fit. Following the definition of the Guinier region, higher q data is added back incrementally and visualized in the Porod plot ($q^4 I(q)$ vs. q) of PRIMUS until an asymptotic region is found. Verification of the asymptotic region can be achieved by plotting the selected data in SASPLOT³⁵. An asymptotic plateau will be observed for folded particles in a $q^4 I(q)$ vs. q^4 plot and for completely flexible particles in a $q^2 I(q)$ vs. q^2 plot. In addition, for partially folded particles, an asymptote may be visible in a $q^3 I(q)$ vs. q^3 plot. The maximum q value that maintains the asymptotic region in the Porod plot defines the upper limit for which the determination of the particle volume should be made in PRIMUS. For further information regarding the determination of the particle’s volume from SAXS data we refer to Glatter, O., and Kratky, O.²⁶, Konarev, P.V., *et al*³⁵, and Feigin L.A. *et al*⁶⁰.

The particle volume (Å³) from homogeneous SAXS sample of known mass can be

$$d_{particle} = \frac{m_{particle}}{V_{particle}} \cdot \frac{10^{24} \text{ Å}^3}{\text{cm}^3} \cdot \frac{1}{N_a}$$

used to calculate the particle’s corresponding density (in gm·cm⁻³) using: where N_a is Avogadro’s number at 6.022×10^{23} molecules per mole. The ratios on the

$$d_{particle} = \frac{m_{particle}}{V_{particle}} \cdot 1.66$$

right reduce to a convenient representation: All graphs were prepared with Kaleidagraph (Synergy Software, Reading, PA).

Acknowledgments

The SIBYLS SAXS beamline (BL12.3.1) at the Advanced Light Source is supported by United States Department of Energy program Integrated Diffraction Analysis Technologies DE-AC02-05CH11231. SAXS research on microbial systems and pathways in the Tainer laboratory is supported by supported by the ENIGMA Program of the Department of Energy, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory. SAXS efforts on eukaryotic DNA repair complexes is supported by the National Institutes of Health (NIH) Structural Cell Biology of DNA Repair Machines P01 grant CA92584. We thank SIBYLS staff members G. Williams, G. Hura and M. Hammel for discussions.

REFERENCES

1. Rambo RP, Tainer JA. *Curr Opin Struct Biol.* 2010; 20:128–137. [PubMed: 20097063]
2. Dyson HJ, Wright PE. *Nat Rev Mol Cell Biol.* 2005; 6:197–208. [PubMed: 15738986]
3. Dunker AK, Silman I, Uversky VN, Sussman JL. *Curr Opin Struct Biol.* 2008; 18:756–764. [PubMed: 18952168]
4. Wells M, Tidow H, Rutherford TJ, Markwick P, Jensen MR, Mylonas E, Svergun DI, Blackledge M, Fersht AR. *Proc Natl Acad Sci U S A.* 2008; 105:5762–5767. [PubMed: 18391200]
5. Receveur V, Czjzek M, Schulein M, Panine P, Henrissat B. *J Biol Chem.* 2002; 277:40887–40892. [PubMed: 12186865]
6. Williams RS, Moncalian G, Williams JS, Yamada Y, Limbo O, Shin DS, Grocock LM, Cahill D, Hitomi C, Guenther G, Moiani D, Carney JP, Russell P, Tainer JA. *Cell.* 2008; 135:97–109. [PubMed: 18854158]
7. Pretto DI, Tsutakawa S, Brosey CA, Castillo A, Chagot ME, Smith JA, Tainer JA, Chazin WJ. *Biochemistry.* 49:2880–2889. [PubMed: 20184389]
8. Williams RS, Dodson GE, Limbo O, Yamada Y, Williams JS, Guenther G, Classen S, Glover JN, Iwasaki H, Russell P, Tainer JA. *Cell.* 2009; 139:87–99. [PubMed: 19804755]
9. Cotner-Gohara E, Kim IK, Hammel M, Tainer JA, Tomkinson AE, Ellenberger T. *Biochemistry.* 49:6165–6176. [PubMed: 20518483]
10. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S, Abolfath RM, Chen DJ, Lees-Miller SP, Tainer JA. *J Biol Chem.* 2009; 285:1414–1423. [PubMed: 19893054]
11. Yamagata A, Tainer JA. *EMBO J.* 2007; 26:878–890. [PubMed: 17255937]
12. Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, Cooper PK, Tainer JA. *Cell.* 2008; 133:789–800. [PubMed: 18510924]
13. Shin DS, Pellegrini L, Daniels DS, Yelent B, Craig L, Bates D, Yu DS, Shivji MK, Hitomi C, Arvai AS, Volkman N, Tsuruta H, Blundell TL, Venkitaraman AR, Tainer JA. *EMBO J.* 2003; 22:4566–4576. [PubMed: 12941707]
14. Hopfner KP, Karcher A, Shin DS, Craig L, Arthur LM, Carney JP, Tainer JA. *Cell.* 2000; 101:789–800. [PubMed: 10892749]
15. Cvetkovic A, Menon AL, Thorgersen MP, Scott JW, Poole FL 2nd, Jenney FE Jr, Lancaster WA, Praissman JL, Shanmukh S, Vaccaro BJ, Trauger SA, Kalisiak E, Apon JV, Siuzdak G, Yannoni SM, Tainer JA, Adams MW. *Nature.* 2010; 466:779–782. [PubMed: 20639861]
16. Takamoto K, Das R, He Q, Doniach S, Brenowitz M, Herschlag D, Chance MR. *J Mol Biol.* 2004; 343:1195–1206. [PubMed: 15491606]
17. DiDonato M, Craig L, Huff ME, Thayer MM, Cardoso RM, Kassmann CJ, Lo TP, Bruns CK, Powers ET, Kelly JW, Getzoff ED, Tainer JA. *J Mol Biol.* 2003; 332:601–615. [PubMed: 12963370]
18. Shin DS, DiDonato M, Barondeau DP, Hura GL, Hitomi C, Berglund JA, Getzoff ED, Cary SC, Tainer JA. *J Mol Biol.* 2009; 385:1534–1555. [PubMed: 19063897]
19. Karunatilaka KS, Solem A, Pyle AM, Rueda D. *Nature.* 467:935–939. [PubMed: 20944626]
20. Prudden J, Perry JJ, Arvai AS, Tainer JA, Boddy MN. *Nat Struct Mol Biol.* 2009; 16:509–516. [PubMed: 19363481]

21. Selmer M, Al-Karadaghi S, Hirokawa G, Kaji A, Liljas A. *Science*. 1999; 286:2349–2352. [PubMed: 10600747]
22. Putnam CD, Tainer JA. *DNA Repair (Amst)*. 2005; 4:1410–1420. [PubMed: 16226493]
23. Petoukhov MV, Svergun DI. *Curr Opin Struct Biol*. 2007; 17:562–571. [PubMed: 17714935]
24. Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL 2nd, Tsutakawa SE, Jenney FE Jr, Classen S, Frankel KA, Hopkins RC, Yang SJ, Scott JW, Dillard BD, Adams MW, Tainer JA. *Nat Methods*. 2009; 6:606–612. [PubMed: 19620974]
25. Putnam CD, Hammel M, Hura GL, Tainer JA. *Q Rev Biophys*. 2007; 40:191–285. [PubMed: 18078545]
26. Glatte, O.; Kratky, O. *Small angle x-ray scattering*. London; New York: Academic Press; 1982.
27. Guinier A. *Ann Phys*. 1939; 12:76.
28. Mylonas E, Svergun DI. *Journal of Applied Crystallography*. 2007; 40:s245–s249.
29. Fischer H, de Oliveira Neto M, Napolitano HB, Polikarpov I, Craievich AF. *Journal of Applied Crystallography*. 2009; 43:101–109.
30. Koch MH, Vachette P, Svergun DI. *Q Rev Biophys*. 2003; 36:147–227. [PubMed: 14686102]
31. Durand D, Vives C, Cannella D, Perez J, Pebay-Peyroula E, Vachette P, Fieschi F. *J Struct Biol*. 1993; 113:45–53. [PubMed: 19723583]
32. Porod G. *Kolloid Zeitschrift*. 1951; 124:31.
33. Debye P, Anderson JHR, Brumberger H. *Journal of Applied Physics*. 1957; 28:679–683.
34. Ruland W. *Journal of Applied Crystallography*. 1971; 4:70–73.
35. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. *Journal of Applied Crystallography*. 2003; 36:1277–1282.
36. Quillin ML, Wingfield PT, Matthews BW. *Proc Natl Acad Sci U S A*. 2006; 103:19749–19753. [PubMed: 17179045]
37. Squire PG, Himmel ME. *Arch Biochem Biophys*. 1979; 196:165–177. [PubMed: 507801]
38. Kozak M. *Journal of Applied Crystallography*. 2005; 38:4.
39. Nishimura N, Hitomi K, Arvai AS, Rambo RP, Hitomi C, Cutler SR, Schroeder JI, Getzoff ED. *Science*. 2009; 326:1373–1379. [PubMed: 19933100]
40. Stoddard CD, Montange RK, Hennelly SP, Rambo RP, Sanbonmatsu KY, Batey RT. *Structure*. 2010; 18:787–797. [PubMed: 20637415]
41. Rambo RP, Tainer JA. *RNA*. 2010; 16:638–646. [PubMed: 20106957]
42. Williams GJ, Williams RS, Williams JS, Moncalian G, Arvai A, Limbo O, Guenther G, SilDas S, Hammel M, Russel P, Tainer J. *A. Nat Struct Biol*. 2011 accepted.
43. Franke D, Svergun D. *Journal of Applied Crystallography*. 2009; 42:342–346.
44. Svergun DI, Petoukhov MV, Koch MH. *Biophys J*. 2001; 80:2946–2953. [PubMed: 11371467]
45. Matthews BW. *J Mol Biol*. 1968; 33:491–497. [PubMed: 5700707]
46. Mareuil F, Sizun C, Perez J, Schoenauer M, Lallemand JY, Bontems F. *Eur Biophys J*. 2007; 37:95–104. [PubMed: 17522855]
47. Parikh SS, Walcher G, Jones GD, Slupphaug G, Krokan HE, Blackburn GM, Tainer JA. *Proc Natl Acad Sci U S A*. 2000; 97:5083–5088. [PubMed: 10805771]
48. Huffman JL, Sundheim O, Tainer JA. *Mutat Res*. 2005; 577:55–76. [PubMed: 15941573]
49. Tsutakawa SE, Classen S, Chapados BR, Arvai A, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA. *Cell*. 2011 accepted.
50. Sarker AH, Tsutakawa SE, Kostek S, Ng C, Shin DS, Peris M, Campeau E, Tainer JA, Nogales E, Cooper PK. *Mol Cell*. 2005; 20:187–198. [PubMed: 16246722]
51. Tainer JA, McCammon JA, Ivanov I. *J Am Chem Soc*. 1991; 113:7372–7378. [PubMed: 20455582]
52. Tubbs JL, Latypov V, Kanugula S, Butt A, Melikishvili M, Kraehenbuehl R, Fleck O, Marriott A, Watson AJ, Verbeek B, McGown G, Thorncroft M, Santibanez-Koref MF, Millington C, Arvai AS, Kroeger MD, Peterson LA, Williams DM, Fried MG, Margison GP, Pegg AE, Tainer JA. *Nature*. 2009; 459:808–813. [PubMed: 19516334]
53. Perry JJ, Cotner-Gohara E, Ellenberger T, Tainer JA. *Curr Opin Struct Biol*. 2010; 20:283–294. [PubMed: 20439160]

54. Tainer JA, Getzoff ED, Alexander H, Houghten RA, Olson AJ, Lerner RA, Hendrickson WA. *Nature*. 1984; 312:127–134. [PubMed: 6209578]
55. Fieser TM, Tainer JA, Geysen HM, Houghten RA, Lerner RA. *Proc Natl Acad Sci U S A*. 1987; 84:8568–8572. [PubMed: 2446325]
56. Hallewell RA, Laria I, Tabrizi A, Carlin G, Getzoff ED, Tainer JA, Cousens LS, Mullenbach GT. *J Biol Chem*. 1989; 264:5260–5268. [PubMed: 2647749]
57. Schnecke V, Swanson CA, Getzoff ED, Tainer JA, Kuhn LA. *Proteins*. 1998; 33:74–87. [PubMed: 9741846]
58. Garcin ED, Arvai AS, Rosenfeld RJ, Kroeger MD, Crane BR, Andersson G, Andrews G, Hamley PJ, Mallinder PR, Nicholls DJ, St-Gallay SA, Tinker AC, Gensmantel NP, Mete A, Cheshire DR, Connolly S, Stuehr DJ, Aberg A, Wallace AV, Tainer JA, Getzoff ED. *Nat Chem Biol*. 2008; 4:700–707. [PubMed: 18849972]
59. Das R, Baker D. *Annu Rev Biochem*. 2008; 77:363–382. [PubMed: 18410248]
60. Svergun, DI.; Fe*igin, LA.; Taylor, GW. *Structure analysis by small-angle x-ray and neutron scattering*. New York: Plenum Press; 1987.

FIGURE 1.

At low resolution (standing far from picture), differing contrasts contribute to the identification of objects within the picture. The hat, dress and grassy surroundings are clear features of the paintings with differing contrasts. Under close inspection (high resolution), only high contrast differences, such as the blue dress, allow for individual features to be distinguishable whereas differences with low contrast (boxed region) become difficult to discern. The difficulty of identifying the hat against the grassy background is analogous to the electron density contrast becoming continuous between the particle and solvent for flexible systems. This image was adapted from *Matin à Eragny* by Camille Pissarro and kindly prepared by Mike Pique at the Scripps Research Institute.

FIGURE 2.

Canonical SAXS scattering equations. A) Guinier approximation describing the linear relationship between the observed scattering intensities, $I(q)$, and scattering angle, q . The approximation determines the radius-of-gyration, R_g , and scattering at $q=0$, $I(0)$. $I(0)$ is directly related to the particle's volume, V , and electron density contrast, $\Delta\rho$. B) Kratky plot of scattering data illustrating changes in the behavior of the curve for folded (sphere), partially folded (sphere-random coil) and completely unfolded particles (random coil). For a folded particle, the integrated area under the curve determines the Porod invariant, Q , and is scaled by concentration, c . C) SAXS derived structural parameters. The ratio of $I(0)$ to Q determines the volume of the scattering particle sometimes known as the Porod volume, V_p .

Application of the Porod-Debye law determines the surface area, S , of the scattering particle that is scaled by concentration and $\Delta\rho$ that can be normalized using Q .

FIGURE 3.

Dimensionless Kratky analysis. A) SAXS data of a 70 kDa protein collected at two concentrations (open gray circles represent a 70% dilution of the sample in red). B) Kratky plot of the data in A where the data in gray was scaled by $1.4 \times$ dilution factor. C) Transformation equations for the dimensionless Kratky plot using Guinier parameters: R_g and $I(0)$. D) Dimensionless Kratky transformation using the experimental Guinier parameters for each respective curve in A.

FIGURE 4.

S. rubiginosus glucose isomerase. A) Experimental SAXS intensity plot with the tetrameric structure (blue). B) Porod plot of SAXS data transformed as $q^4 I(q)$ vs. q . Boxed region in gray describes the Porod-Debye region used in C and D. Examination of the boxed region in C) shows the Porod-Debye plateau that is proportional to $2\pi \cdot (\Delta\rho)^2 \cdot S$. D) Porod-Debye plot of SAXS data transformed as $q^4 I(q)$ vs. q^4 . Both plots in C and D aid in defining the Porod-Debye region for the determination of V using PRIMUS.

FIGURE 5.

SAXS with an exemplary intrinsically disordered domain Rad51 AP1. Data was collected for both rad51 AP1 (red) and a fusion construct with *E. coli* maltose binding protein (MBP) (black). A) Kratky plot demonstrating the parabolic shape for a partially folded particle (black) and hyperbolic shape for a full unfolded particle (red). B) Porod-Debye plot demonstrating a clear plateau for the partially folded MBP-Rad51 AP1 hybrid particle. In the absence of MBP, the fully unfolded domain is devoid of any discernible plateau. Data was kindly provided by Gareth Williams at the Lawrence Berkeley National Laboratory.

FIGURE 6.

Metabolite induced structural changes monitored by SAXS. A) SAXS data for the 42 kDa abscisic acid binding domain PYR1 *bound* (black) and *apo* (red). B) Porod-Debye plot demonstrating discrete plateaus in *bound* and *apo* states suggesting each state can be characterized by discretely folded particles with sharp scattering contrasts. X-ray crystal structure (PDB: 3K3K) characterizes the *apo* state where it is expected the loop regions reorganize collapsing the structure in the *bound* state. C) SAXS data for the SAM-1 riboswitch *bound* (black) and *apo* (red). Crystal structure of the *bound* form (blue) changes in the *apo* state. D) Porod-Debye plot illustrating loss of the Porod plateau in the unbound state suggesting enhanced flexibility of the riboswitch. Transforming *apo* SAXS data by $q^3 I(q)$ vs. q^3 (inset) verifies the intensity decay is no longer q^{-4} but q^{-3} . Data adapted from Nishimura, N.; Hitomi, K.; Arvai, A. S.; Rambo, R. P.; Hitomi, C.; Cutler, S. R.; Schroeder, J. I.; Getzoff, E. D. *Science* 2009, 326, 1373–1379.

FIGURE 7.

Detecting conformational flexibility. A) SAXS data for the Mre11-Rad50 complex in the presence (black) and absence (red) of ATP. SAXS scattering profiles transformed as a Kratky plot does not confidently demonstrate flexibility (inset). B) Porod-Debye plot illustrating changes in the Porod-Debye region. Loss of the plateau suggests Mre11-Rad50 complex becomes more flexible in the absence of ATP. Transforming *apo* SAXS data by $q^3 I(q)$ vs. q^3 (inset) verifies the intensity decay is q^{-3} . Data was adapted from Williams, G.J., Williams, R.S., Williams, J.S., Moncalian, G., Arvai, A., Limbo, O., Guenther, G., SilDas, S., Hammel, M., Russell, P., and Tainer, J. A., (2011) ABC ATPase signature helices in Rad50

link its nucleotide state to the Mre11 interface for DNA double-strand break repair, Nature SMB, accepted.

FIGURE 8.

Histogram analysis of protein densities calculated from SAXS data using 31 different proteins and conditions. For each protein, V was calculated using PRIMUS limiting the data to the Porod-Debye region. Distribution shows the most likely density is 0.9 to 1.0 $\text{gm}\cdot\text{cm}^{-3}$. The equation can be used to calculate a test density where n is the number of subunits in multimeric complexes. For hetero-complexes, n is 1. Units of M , V_{Porod} and 1.66 are in Da, $\text{\AA}^3$ and $\text{\AA}^3\cdot\text{gm}\cdot\text{cm}^{-3}\cdot\text{Da}^{-1}$, respectively. SAXS data was taken from <http://www.bioisis.net>.

SAXS Derived Particle Volumes Using the Porod-Debye Law

Table 1

	BID	V (Å ³)	mass (kDa)	R _g (Å)	d _{protein} (gm·cm ⁻³)
glucose isomerase	n.d.	210000	172.2	33.1	1.36
	PF0699	50333	32.1	23	1.06
NADH oxidase	INADHP	52829	42.7	22.9	1.34
	1ACSHP	129680	80.4	34.2	1.03
Acyl-CoA synthetase	Y/S phosphatase	27990	19.7	16.9	1.17
	phosphoribosyltransferase	77047	51.6	25.3	1.11
arsR transcriptional regulatory protein	PF1372	61972	39.5	23.4	1.06
	PF1291	174720	117.5	36.1	1.12
hypotheical	1HYP0P	13945	8.34	19.5	0.99
	1ARSDP	26217	18	20.2	1.14
amidotransferase	1AMIGP	38165	22.8	19.9	0.99
	PF0863	73405	42.6	27.4	0.96

BID: Biois ID code

n.d. not deposited in Biois

Table 2

Calculated Protein Densities Using Empirically Derived Volumes

n	GI	PYRI	WRN	CAT	IRECBP
1	0.34	0.49	0.45	0.32	0.62
2	0.68	0.99	0.9	0.63	1.24
3	1.02	1.48	1.34	0.95	1.85
4	1.36	1.98	1.79	1.27	2.47
5	1.7	2.47	2.24	1.58	3.09
6	2.04	2.97	2.69	1.90	3.71

n: number of subunits

GI: glucose isomerase (43 kDa)

PYRI: abscisic acid binding protein (21 kDa)

WRN: Werner's exonuclease domain (47 kDa)

CAT: human catalase (57 kDa)

IRECBP: recombinationase (27 kDa)