21 Structural Studies of Proteins and Nucleic Acids in Solution Using Small Angle X-Ray Scattering (SAXS)

R. DAS
Stanford University
Stanford, CA, USA

S. DONIACH
Stanford University
Stanford, CA, USA

1 Introduction

Over the last 20 years, the application of small angle x-ray scattering (SAXS) to the structural investigation of biological macromolecules has made tremendous progress. Historically used to measure molecule size, SAXS has evolved into a versatile technique for quantitatively characterizing the features and forces that govern the behavior of biomolecules in solution. This progress has been driven by the dramatically improving brightness of synchrotron radiation x-ray sources combined with the wide availability of new computational resources for interpreting SAXS measurements. This chapter uses contemporary examples to illustrate how solution SAXS provides insights into the chemical physics of proteins and nucleic acids that are difficult – in some cases, impossible – to obtain by other techniques.

The presented examples are ordered from the most straightforward analyses and applications of SAXS experiments to the most sophisticated, and include practical notes on each analysis type and experimental setup. An outline of the sections follows. After a brief introduction to the simple equations underlying SAXS (› *Section 2*), a description follows of the model-independent analysis of data at the lowest scattering angles to find the size and molecularity of biomolecules randomly oriented and randomly positioned in solution. Recent applications of this classic procedure include a wide-ranging survey of the denatured states of proteins (› *Section 3*) and an unexpected result for the molecularity of the ribonucleoprotein ribonuclease P (› *Section 4*). As the concentration is increased, SAXS data at lowest angles become sensitive to intermolecular attractive or repulsive forces, as illustrated with examples from investigations of attractive interactions for crystallizing proteins and of Coulomb-repulsion ordering of DNA molecules (› *Section 5*).

SAXS data at intermediate to high scattering angles provide significant information beyond the overall size of the biomolecule. For well-structured molecules, recently developed fitting algorithms permit *ab initio* three-dimensional reconstructions at low resolution from the full scattering profiles, and examples are given from a large protein-folding chaperone and from an RNA subdomain of the ribosome (› *Section 6*). Even for unstructured biomolecules, modeling of the SAXS profiles can effectively distinguish between models for the conformational ensembles populated by the molecule; such analysis of a simple tethered-DNA-duplex model system leads to a measurement of the fundamental electrostatic forces that dominate nucleic acid folding (› *Section 7*).

In the final sections, a description is provided of more specialized SAXS implementations that continue to have a strong impact on biomolecular structural studies. Anomalous (resonant) SAXS permits the dissection of the scattering

profile into subparts reflecting the distribution of individual elements within a biomolecule, as illustrated by a recent study of a mixed counterion atmosphere around nucleic acids (Section 8). Finally, time-resolved SAXS incorporates special mixers and high-flux beamlines to probe biomolecular conformational changes, such as the folding of RNAs and proteins, with time resolutions extending to milliseconds and faster (Section 9). The conclusion provides some final notes on the advantages and disadvantages of pursuing SAXS studies of biomolecular systems (Section 10).

2 What Does SAXS Measure?

Small angle x-ray scattering from biomolecules in solution at low concentration results in a scattering pattern that is circularly symmetric around the direction of the incident beam and is read out by a linear proportional counter or a charge-coupled detector (Figure 21-1).

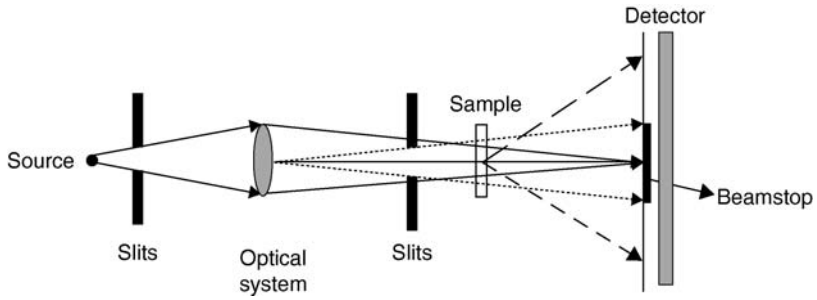

Figure 21-1
Schematic of a typical small-angle x-ray scattering setup. The optical system (monochromator/mirrors) selects the appropriate wavelength (~1 Å) from the synchrotron source and focuses the beam on the detector, usually a linear proportional counter or a charge-coupled detector. The lowest scattering angle of observation is determined not only by the focus size but also by the dimensions of the last aperture, which defines a region of high background (due to the optical system and slits). The biomolecule sample is contained in a cell about 1 mm thick; and the sample-detector distance is usually set between 1 and 10 m. Figure reproduced from [1]

The resulting scattering profiles are generally plotted as the intensity of scattered photons $I(q)$ *versus* the scattering wavevector $q = (2\pi/\lambda)\sin(\theta/2)$, where θ is the scattering angle and λ is the x-ray wavelength. (Many authors display the intensity *versus* $s = q/2\pi$, a representation used in some of the figures in this chapter.) After subtracting a background profile acquired on a sample with

the solvent buffer alone, the observed scattering profile is related to the Fourier transform of the electron density of the scattering molecules relative to solvent, $\Delta\rho(\mathbf{r})$:

$$I(\mathbf{q}) = \left| \int \mathbf{d^3 r} \, \mathbf{e^{iq \cdot r}} \Delta\rho(\mathbf{r}) \right|^2 \tag{1}$$

If the N molecules in the probed sample are oriented randomly and positioned randomly, the scattering intensity can be expressed more simply as a sum of scattering contributions from all pairs of atoms within each molecule, given by the Debye formula [2]:

$$I(q) = N \sum_{i,j} f_i f_j \frac{\sin qr_{ij}}{qr_{ij}}, \tag{2}$$

where r_{ij} is the distance between two atoms and f_i is the number of electrons in each atom minus the number of electrons in the solvent displaced by the atom. If intermolecular forces become significant (e.g., at high molecule concentrations), additional scattering contributions due to pairs of atoms in different molecules arise and can be treated separately (*Section 5*).

The immersion of the biomolecules in water requires extra practical considerations compared to scattering off objects *in vacuo*, with respect to the evaluation of (1) and (2) and to the signal-to-noise of the measurement. First, the scattering is dependent on the contrast in electron density of the hydrated biomolecule relative to the electron density from an equivalent volume and shape of the background solvent (the so-called "ghost" scattering). An algorithm CRYSOL [3] to calculate this effect, along with scattering contributions from the biomolecule's hydration shell, is available from EMBL, Hamburg [4]. The main consequence of contrast reduction by solvent is to diminish the overall SAXS signal measured in solution. Since the average density of a protein molecule is of order $\rho_p = 0.42$ $e^-/Å^3$ while that of the water solvent is $\rho_s = 0.33$ $e^-/Å^3$, the scattered intensity is only $(\rho_p - \rho_s)^2/{\rho_p}^2 \approx 5\%$ of the intensity that would be calculated for the molecule *in vacuo*. For this reason, the acquisition of SAXS data is quite strongly subject to signal-to-noise problems; in particular, time dependent drifts in background scattering that can occur at synchrotron radiation sources must be minimized. Since the background profile also contains spurious signals resulting from parasitic x-ray scattering from beamline components, any small fluctuations in the background measurement can lead to strong distortions of the true scattering profile.

The question of how to relate the scattered intensity to a real space physical model of the molecule is complicated by the limited range of scattering wavevectors

$q_{min} < q < q_{max}$ over which data can be collected with acceptable signal-to-noise. For example, (3) relates the scattering intensity to a useful physical characteristic of a molecule, a histogram $P(r)$ of its electron-density-weighted intramolecular distances r:

$$I(q) = N \int_0^\infty P(r) \frac{\sin qr}{qr} dr \tag{3}$$

However, the inversion of (3) to find a unique distance distribution $P(r)$ from the observed scattering profile $I(q)$ can only be accomplished if scattering data are acquired over scattering wavevectors q from zero to infinity. For data sets with a limited q range, constraints on the smoothness of $P(r)$ and the maximum diameter of the molecule allow acceptable solutions of $P(r)$, though there may still be multiple reasonable results that require further data or physical knowledge to allow them to be distinguished [5–7].

We recommend that the first analysis of biomolecule SAXS profiles include model-independent procedures to determine basic parameters like the radius of gyration and the molecularity of the system (Section 3 and 4). Then, further information on molecular shape can be obtained through comparison of the data to predictions from well-defined models (Section 5–8). While these comparisons can be visualized as overlaid intermolecular distance distributions $P(r)$, uncertainties in the transformations that invert (3) can sometimes render such comparisons misleading. Therefore, a comparison of models to the (untransformed) scattering profiles $I(q)$ is advocated rather than to the $P(r)$ curves when assessing the quality of fits, and this presentation style is used throughout this chapter.

3 The Size of a Biomolecule: Radius-of-Gyration Measurements

The most common application of small-angle x-ray scattering to the investigation of biomolecules has been the determination of the radius of gyration R_g, the root-mean-squared extent of a molecule relative to its center-of-mass. This model-independent measurement depends only on the data at lowest scattering angles through the use of the Guinier formula, derived by expansion of (2):

$$\ln I(q) = \ln I(0) - \frac{1}{3} q^2 R_g^2 + O(q^4), \tag{4}$$

where

$$R_g^2 = \frac{\sum_{i,j} f_i f_j r_{ij}^2}{2\sum_{i,j} f_i f_j}, \tag{5}$$

The classic Guinier fit of the low q region of a scattering profile is illustrated in *Figure 21-2* for one of the workhorses of protein folding studies, the lysozyme molecule.

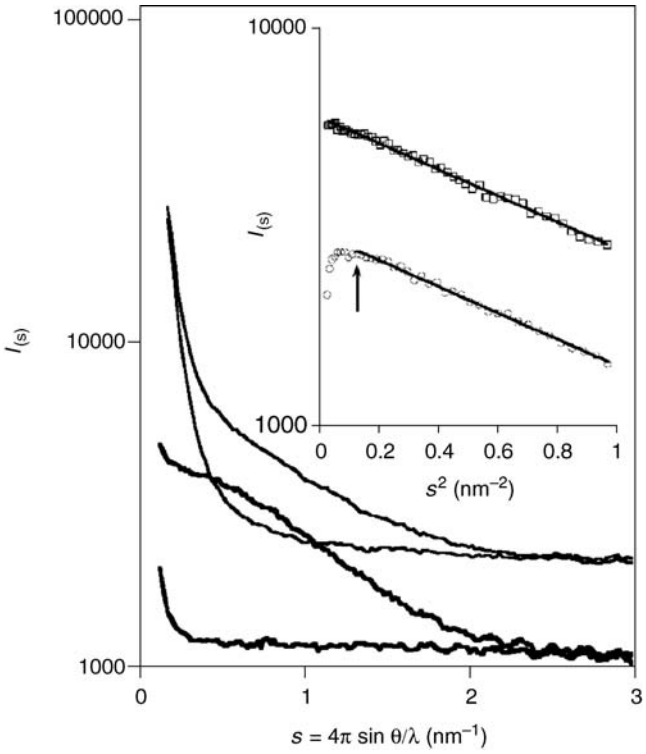

Figure 21-2
Scattering profiles of lysozyme at concentration 5 mg/ml and of its buffer, acquired in a standard capillary tube (*thin lines*) and, for better signal-to-noise, in an evacuated cell (*lower* pair of thick lines). The *insets* display the corresponding Guinier plots for the capillary (*circles*) and the evacuated cell (*squares*). The profiles have been displaced along the ordinate for better visualization. Figure reproduced from [1]

On a practical note, the Guinier formula (4) is not applicable if the concentration is so high that interparticle scattering becomes important or that the molecule aggregates (see *Section 5*). The standard consistency requirements for Guinier analysis are linearity of the Guinier plot (*Figure 21-2*) and

independence of the observed profile with respect to increases in biomolecule concentration and in radiation-damaging x-ray exposure, both of which can induce aggregation. Careful purification (including filtering and centrifugation of samples) and testing a range of buffers can prevent contamination of the sample with aggregates or large contaminants that can invalidate radius-of-gyration measurements.

Radius-of-gyration measurements from SAXS have long played important roles in constraining theoretical models of biomolecules, since R_g values are straightforward to calculate for different proposed structures. A contemporary example is the SAXS study of the chemically denatured state of proteins, with comparison to the R_g predictions for a "random coil" state. The investigation of denatured proteins has been rejuvenated by several nuclear magnetic relaxation studies that indicate significant residue-residue correlations in the denatured state; see the review by Millett et al. [8] and references therein. Understanding the folding process will likely depend on characterizing such initial correlations that vary from protein to protein and that are not taken into account by a random coil picture. However, SAXS studies appear to contradict the NMR observations. The dependence of R_g on residue number was determined for a set of twenty-eight cross link-and prosthetic group- free chemically denatured proteins and polypeptides spanning the range from 16 to 549 residues [9]. Strikingly, these data (*Figure 21-3*) agree well with Flory's classic prediction of a power law for the radius of gyration of a random coil polymer:

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

with exponent $\nu \cong 3/5$.

Can the disagreement [8] between the NMR-observed correlations and the random coil behavior seen by SAXS be reconciled? Enlightening Monte Carlo simulations based on artificial manipulations of the native structure [10] and on NMR data for spin-labeled proteins [11] have shown how significant short-range and long-range interactions can be present in "denatured" conformational ensembles that still give radii of gyration consistent with a random-coil value. Similarly, SAXS experiments on the cytochrome *c* protein find indistinguishable R_g values at 3.5 M and 5 M concentrations of the denaturant guanidinium hydrochloride, while the higher scattering angle data exhibit correlations that indicate rather different overall ensembles under the two conditions (a partially ordered globule and a random coil, respectively) [12]. The relatively weak dependence of the radius of gyration on the degree of denaturation likely explains how the R_g data are consistent with random coil values while NMR measurements show considerable residue correlations under the same conditions.

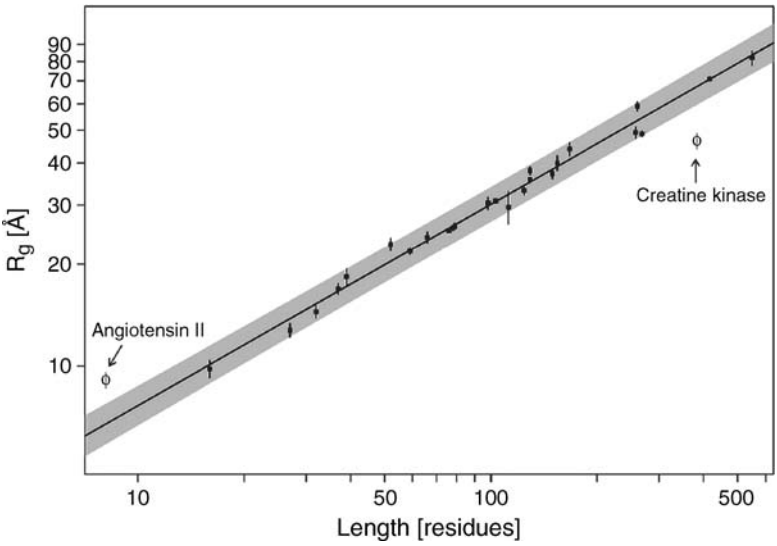

Figure 21-3
The radius-of-gyration values of two dozen chemically denatured proteins scale with polymer length, *N*, via the power-law relationship $R_g = R_o N^{\nu}$. The solid line, which is the least *squares* fit ignoring the two indicated outliers, produces an exponent, $\nu = 0.598 \pm 0.029$ (95% confidence interval), that is indistinguishable from the 0.588 predicted for an excluded volume random coil. Figure reproduced from [9]

These studies on the denatured state of proteins exemplify the utility of Guinier analysis, allowing remarkable relationships like that depicted in *Figure 21-3* to be visualized, as well as its limitation as a low resolution measurement. Radius-of-gyration determinations provide a single measure of global size and should be complemented with other "local" probes of structural correlations like NMR spectroscopy as well as careful analysis of the higher angle SAXS data (see *Section 6–9*).

4 Monomer, Dimer, or Multimer?

In addition to giving a global measurement of the overall size of a biomolecule, SAXS data at the lowest scattering angles provides a model-independent determination of its molecularity, i.e., whether the molecule is a monomer, dimer, or multimer. For a scatterer with contrast $\overline{\Delta\rho}$ averaged over the molecular volume V, the forward scattering intensity at $q = 0$ is given by a simple expression, derivable from (1):

$$I_0 = N\left(\overline{\Delta\rho}V\right)^2 \tag{7}$$

If the molecule is present as a dimer rather than a monomer, the number of scatterers N is halved but the molecular volume V per scatterer doubles, resulting in an overall doubling of I_0. Thus, the molecularity of a sample is proportional to the observed forward scattering intensity and can be determined by comparison to the I_0 value measured for a well-characterized standard known to be monomeric. In practice, the beamstop that blocks the unscattered x-ray beam (*Figure 21-1*) precludes a direct measurement of I_0, but a Guinier fit (4) of the low q data allows an accurate extrapolation to $q = 0$.

As an example of this model-independent measurement, a recent surprising result is described from a SAXS investigation by Fang et al. [13] of the universally conserved enzyme ribonuclease P, one of the first systems shown to have a catalytic RNA component in 1983. Functional studies had established the RNA-protein stochiometry of this ribonucleoprotein complex to be 1:1, and it had been widely assumed that each enzyme molecule contained one RNA subunit and one protein subunit. SAXS data comparing the forward scattering of the *B. subtilis* enzyme's RNA subunit to a control molecule, unmodified $tRNA^{phe}$ from yeast, verified that the RNA subunit alone is a monomer. However, addition of the protein subunit to form the ribonuclease P holoenzyme resulted in a doubling of the forward scattering (*Figure 21-4*), consistent with a dimerization of the RNA domains induced by two low-molecular-weight protein subunits. This simple SAXS measurement has forced the re-evaluation of the interpretations

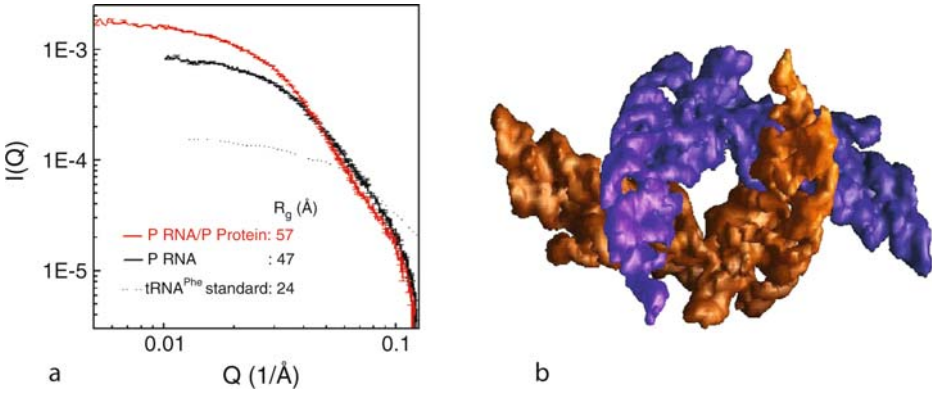

Figure 21-4

SAXS measurements on ribonuclease P show that the RNA-protein enzyme is a dimer. (a) Scattering profile of the ribonuclease P RNA shows twice the forward scattering intensity with protein (red) than without protein (black). The dotted line shows scattering of $tRNA^{(phe)}$ from yeast at the same concentration, as a monomer standard. (b) Model of the two RNA subunits in the RNA-protein holoenzyme based on fits to the full SAXS profile. Figures reproduced from [13]

of structure mapping and functional studies that for more than a decade assumed the holoenzyme to be a monomer.

5 Probing Intermolecular Forces Between Biomolecules

As the biomolecule concentration of a sample is increased, an interparticle interference signal due to x-ray scattering from two biomolecules in close proximity appears in the forward direction. For molecules of approximate spherical shape, the intermolecular potential may be treated as isotropic, and the effect of this SAXS interference signal can be related to the intermolecular potential.

The scattering from a solution of biomolecules may be approximated in terms of the configurational average over the electron density, by analogy with the treatment of scattering from liquids. The scattering profile may be written in terms of the product of a form factor for a given biomolecule with a term for the pairwise intermolecular distribution,

$$I(q) = I_{\text{biomolecule}}(q) S_{\text{intermolecular}}(q), \tag{8}$$

where $I_{\text{biomolecule}}(q)$ is the form factor, corresponding to the scattering profile for the biomolecule in the limit of zero concentration. The new term $S_{\text{intermolecular}}(q)$ represents the effects of interparticle interference and is related to the intermolecular distribution function $g(r)$ and the biomolecule concentration C:

$$S_{\text{intermolecular}}(q) = 1 + C \int d^3\mathbf{r}[g(r) - 1]e^{i\mathbf{q}\cdot\mathbf{r}}. \tag{9}$$

The distribution function $g(r) = \exp(-U(r)/k_B T)$ may be compared to the predictions of the interparticle potential $U(r)$ from analytical treatments such as the hypernetted chain equation or from numerical calculations. Thus, once the form factor of the biomolecule is known from measurements at low concentration, a quantitative measure of the molecular interaction may be obtained by expanding in powers of the concentration.

Physically, the interpretation of interparticle interference is straightforward – if the biomolecules attract, then the interparticle correlation function $g(r)$ will contain contributions in which the probability of a pair of molecules being close to each other is larger than that that obtained in a model system with only hard-sphere repulsion. Since the forward scattering of a given pair of molecules is proportional to the square of the molecular weight, it is enhanced by a factor of two relative to scattering of the two separate molecules [(9); see also (7)]. Thus, a system of molecules in which the intermolecular potential is attractive at

short distances will lead to SAXS profiles in which increasing the concentration leads to an increase of scattering at small angles. The increase of the forward scattering by an attractive interaction is thus a generalization of the effect seen from the formation of dimers or multimers upon increasing biomolecule concentration (see, e.g., *Section 4*). As an extreme case, very large increases in forward scattering result from formation of oligomers or large aggregates. Analogous arguments show that if the potential is repulsive, the forward scattering will decrease as concentration is increased.

The effects of interparticle forces are often summarized in terms of a single parameter, the second virial coefficient, A_2, defined as an integral over the interparticle correlation function:

$$A_2 = \frac{2\pi N_a}{M^2} \int_0^\infty [1 - g(r)] r^2 dr, \tag{10}$$

where N_a is Avogadro's number and M is the molecular weight. Based on the relation (9), the second virial coefficient can be directly extracted from the SAXS data at $q = 0$:

$$\frac{1}{S_{\text{intermolecular}}(0)} = 1 + 2M A_2 c + O(c^2), \tag{11}$$

Here, c is the molecule concentration (in g/ml). A net repulsive force between particles corresponds to a positive value of the measured coefficient A_2, and an attractive force corresponds to negative A_2.

A practical application of this interparticle effect is the investigation of solvent conditions that optimize the biomolecular crystallization process. X-ray crystallography is the pre-eminent tool for detailed structural characterization of biomolecules, but this technology has been limited by the extensive time required to screen salt, pH, cosolvent, and other buffer conditions to allow crystal formation. Building on previous light scattering studies, Bonneté and Vivarès have used SAXS to postulate general rules for how attractive the intermolecular force needs to be to promote crystallization [14, 15]. They have measured the second virial coefficient for four biological systems with sizes ranging from 14 kDa to 4,600 kDa under varying concentrations of salt and the copolymer polyethylene glycol; for example, *Figure 21-5* shows concentration-dependent SAXS profiles displaying evidence for a polyethylene-glycol-induced attractive intermolecular force. These results allow the proposal of a universal, tightly defined "crystallization zone" for all the studied particles in terms of a dimensionless version of the second virial coefficient, normalized to the molecular volume [14]. In addition to the $I(0)$ values giving the potential strength via the

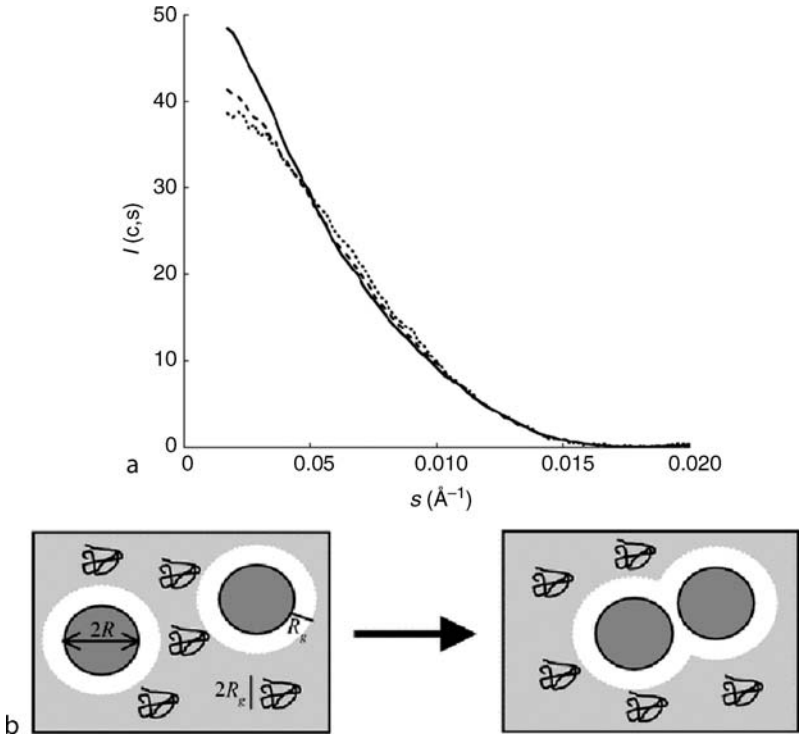

Figure 21-5

An increase in forward scattering with protein concentration indicates an attractive force between urate oxidase molecules. (a) SAXS profiles in the presence of 8% polyethylene glycol (PEG) 3350 at various protein concentrations: 4 mg/ml (*dotted line*), 16 mg/ml (*dashed line*), and 32 mg/ml (*solid line*). (b) The interparticle interference signal is consistent with a simple mechanism of PEG-induced attraction. The depletion zone around each protein where the centers of the PEG molecule cannot enter is depicted in white. When the depletion zones overlap, the volume accessible to the polymer increases, giving an entropic drive for bringing the proteins together. Figures reproduced from [15]

A_2 measurement, the full SAXS profiles $I(q)$ provide estimates of the *range* of the potential $U(r)$ through (9) and comparison to Monte Carlo calculations. The attractive potentials for proteins under rather different salt-induced and polyethylene-glycol-induced crystallizing conditions have experimentally derived ranges similar to the diameters of the biomolecules, suggesting a general scaling rule for interactions that lead to crystallization.

In contrast to the weakly attractive forces described for crystallizing proteins, intermolecular forces can be strongly repulsive for charged biomolecules. For example, two 20 base-pair DNA double helices have a massive Coulomb repulsion energy of >10 kT for interhelical separations smaller than 100 Å, in the limit of low salt concentrations. SAXS measurements on DNA by Skibinska et al. [16] illustrate this repulsive force, displaying a sharp dip in the intensity

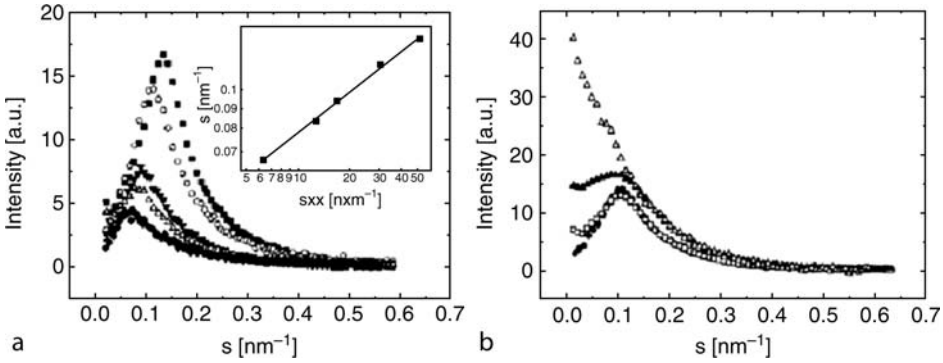

Figure 21-6

A strong negative interference signal at low scattering angles demonstrates the strong Coulomb repulsion between 20 base-pair DNA double helices. (a) SAXS profiles of samples with no added salt with DNA concentrations of 6.3, 12.6, 16.7, 30, and 50 mg/ml (from bottom to top). The inset shows the scaling of peak position with DNA concentration. (b) The interparticle signal disappears as salt is added to screen Coulomb repulsion; added NaCl concentrations are 0 mM (solid circles), 10 mM (open squares), 100 mM (solid triangles), and 500 mM (open triangles). Figures reproduced from [16]

at lowest scattering angles (a positive second virial coefficient; *Figure 21-6a*). The resulting profiles show a characteristic "peak" whose position scales with DNA concentration as $c^{1/3}$ (*Figure 21-6a*), consistent with Bragg scattering from locally ordered lattices of repelling DNA molecules. The interparticle signal disappears as salt is added to the sample, screening Coulomb repulsion (*Figure 21-6b*). Although the non-spherical shape of the molecules and many-body effects complicate quantitative comparison of this data to theoretical predictions of the electrostatic force, these results have inspired the development of a simple model system to measure this fundamental force between nucleic acids by observing intramolecular correlations rather than intermolecular ordering (see below, *Section 8*).

6 Three-Dimensional Reconstruction of Molecule Shapes

By going beyond the lowest scattering angle data discussed in the previous sections, shape information beyond size and molecularity can be extracted from SAXS data. Although the high angle regions of scattering curves look relatively featureless, without sharp maxima or oscillations, they can yield remarkable three-dimensional reconstructions of molecule shapes. The amount of information present in SAXS profiles has only been fully appreciated in recent years with the development of *ab initio* reconstruction algorithms [17–20]. Previously, simple

calculations based on numerical inversion of (3) indicated that the number of independent parameters $N_{channel}$ encoded in the SAXS profiles is rather small:

$$N_{channel} = (q_{\max} - q_{\min}) \times d_{\max}/\pi \tag{12}$$

where $d_{\max}$ is the maximum chord length of the intramolecular distribution function and $q_{\min}$ and $q_{\max}$ are the minimum and maximum scattering wavevectors for which data are acquired. The number $N_{channel}$ is the Shannon information content of the data, representing the maximum number of independent parameters that may be extracted from the data by a linear transformation, and is typically no more than 30.

However, since the work of Svergun and Stuhrmann [21] in the 1970s, it has become apparent that such calculations underestimate the information available from SAXS data. The electron density function is not arbitrarily discontinuous (as would be represented in Shannon's theorem, based on a series of delta functions) but obeys strong physical constraints due to the close packing of amino acid residues or nucleotides within each subdomain of a biomolecule. This observation has led to the development of a number of algorithms that appear to beat the "Shannon limit." By searching through compact bead models to represent the electron density at modest resolution (down to 5 Å), these algorithms successfully mimic the physical constraints of close packing of the amino acid residues or nucleotides in the constituent domains of the biomolecules. The important breakthrough was the demonstration by Chacon et al. [17], further developed by Svergun [19, 20] and by Walther et al. [18], that unique three-dimensional bead representations could be obtained from the scattering data for many different molecule shapes, despite the angular averaging inherent to SAXS profiles.

One observed limitation of the SAXS reconstruction algorithms is that highly symmetric molecules can lead to artifacts in which comparably excellent fits are given by bead models that differ enormously from the known structure, as determined by crystallography [1]. This high symmetry problem can be overcome by algorithms which take into account the pseudo-rotational symmetry expected for a given molecule, based on other data. For a more comprehensive overview of recent developments, the interested reader is directed to a recent review by Koch et al. [1], which provides a detailed discussion of three-dimensional reconstruction and other contemporary applications of SAXS to the study of biological macromolecular structure; see also a recent comparison of three available reconstruction algorithms [22].

An example of SAXS three-dimensional reconstruction is a conformational change induced by adenosine triphosphate (ATP) in a large protein chaperone, exemplifying the large structural changes induced by ATP and other nucleotides in many enzyme systems. Hydrolysis of ATP is a necessary step in the reaction

cycle of the chaperonin TRiC/CTT, an essential and universal constituent of eukaryotic and archaeal cells that assists in the refolding of misfolded proteins like actin. A "lid" that encloses a misfolded substrate into the interior cavity of the molecule consists of subdomains from the eight constituent proteins that make up the barrel of TRiC. The timing of when the lid opens or closes through TRiC's ATP reaction cycle has recently been elucidated by SAXS [23]. The algorithm SAXS3D enables clear determination of the size and shape of the molecule's central cavity and also of the changes in molecule length upon closure of the flexible lid segments, initiated by the addition of ATP and of transition state analogs (*Figure 21-7*). The use of SAXS here avoids possible problems associated with techniques such as crystallography and cryoelectron microscopy, where

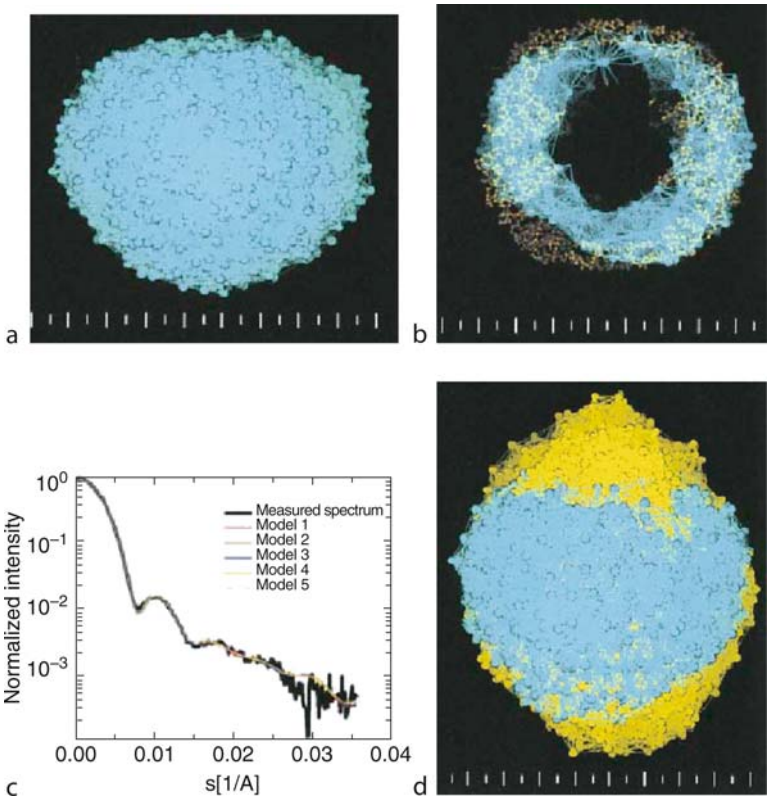

Figure 21-7

Three dimensional SAXS-reconstructed bead model for the large TRiC/CCT complex, a protein folding chaperone. (a–c) The best-fit bead model for the chaperone in the presence of ATP exhibits a hollow cavity, better seen in the cross-section (b), and excellent agreement with the experimental scattering profile (c). (d) Superposition of models with (*turquoise*) and without (*yellow*) ATP reveals that compaction occurs along the longitudinal axis of the complex, suggesting movement of the lid segments. The ruler at the bottom of the model images marks 10 Å increments. Figures reproduced from [23]

distortions of the flexible segments from crystal packing forces and vitrification can complicate interpretation. More generally, SAXS is straightforward to apply to biomolecules for which crystallization of stable intermediates is difficult or unsuccessful. The picture quality is comparable to that of low resolution electron density maps from cryoelectron tomography [24], while avoiding particle selection bias.

In addition to its applications to proteins, SAXS three-dimensional reconstruction offers exciting possibilities for visualizing the native and metastable folds of structured DNAs and RNAs. Large nucleic acids have been far more difficult to crystallize than proteins and are difficult targets for NMR structure determination due to a relative lack of protons and chemical shift dispersion. However, the high electron density of nucleic acids, relative to proteins, make them ideal candidates for scattering experiments. An early example of three-dimensional reconstruction from nucleic acid SAXS data was the solution structure of the 5S rRNA [25]. an essential component of the ribosome that had evaded high resolution structural characterization by x-ray crystallography. The ab initio reconstruction reveals a Y-shaped structure with one short arm and two long arms, in contrast to previous SAXS-based models, and allows the tentative positioning of atomic models of the 5S rRNA's subdomains within the low resolution shape (*Figure 21-8*). The resulting low-resolution model of the isolated domain is indistinguishable from the structures of the 5S rRNA solved

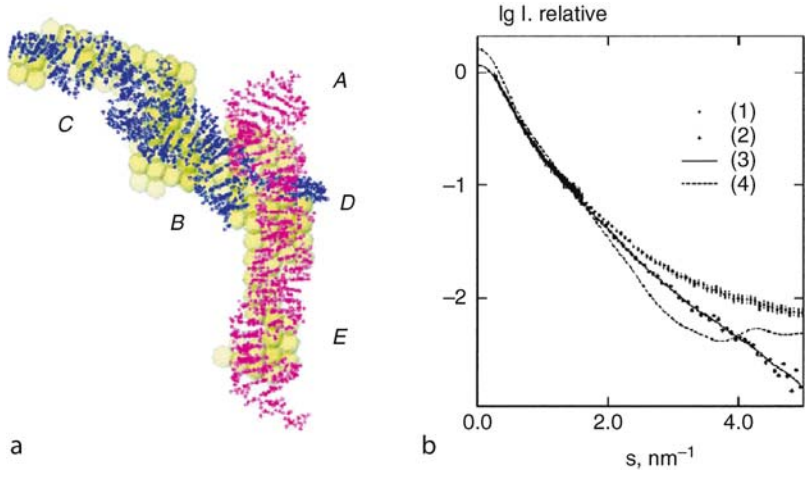

Figure 21-8

Three-dimensional reconstruction of a folded RNA molecule, the 5S ribosomal RNA. (a) Bead model (*yellow spheres*) of the RNA with atomic models of the separately crystallized domains docked inside. (b) The SAXS profile predicted from the reconstructed bead model (*solid line*) but not a previous molecular model (*dotted line* [28]) displays an excellent fit to the data (symbols)

in complex with the 50S ribosomal subunit from the landmark high-resolution crystallographic studies [26, 27] published soon after the SAXS reconstruction.

As a final caveat, the use of 3D reconstruction algorithms requires an assumption that the biomolecules are in a fixed, well-defined configuration. Any domains of the molecule that sample multiple conformations will be "solved" in a single intermediate conformation by the reconstruction algorithms [22]. Such ambiguities in the reconstruction are likely to be present for the open-lid conformations of the nucleotide-free TRiC chaperone (*Figure 21-7*) and for the counterion atmosphere (see *Section 8* below) that surround the RNA helices in the 5S rRNA (*Figure 21-8*). For systems where quantitative modeling of these dynamic biomolecule components is possible, the SAXS data can be usefully compared to the predictions for conformational ensembles of the domain arrangements or counterion distributions, as is described in the following sections.

7 Modeling States with Conformational Diversity

SAXS is a useful tool for investigating poorly structured molecules, as has already been noted for the denatured states of proteins in *Section 3*. Indeed, because such systems cannot be crystallized and because their spectroscopic signatures are often difficult to interpret, the abilities to accurately calculate full scattering profiles with (2) and to compare these predictions to experimental data make SAXS one of the premiere tools for distinguishing different models of conformational ensembles.

To illustrate this type of modeling, the application of full SAXS profiles to understanding the unfolded states of nucleic acids is described. The "unfolded" state for most studies of RNAs and DNAs is not a single-stranded random coil but rather a state with secondary structure (double helices) preformed. Thus, the question arises as to how these interlinked helices are arranged in the absence of the sequence motifs or divalent ions that promote specific tertiary contacts between helices. In such unfolded conditions, the negatively charged helices may repel each other to form a highly extended unfolded state or take on a random conformational ensemble. Recently, sophisticated theoretical work on DNA condensates [29–31] has advanced another possibility: a net counterion-induced attractive force between helices may lead to a *compact* denatured state for nucleic acids, with helices non-specifically held together by fluctuating salt bridges in which Mg^{2+} counterions attract pairs of negative charges from different helices. All three of these scenarios for the unfolded state of RNA have been proposed as appropriate descriptions of the large *Tetrahymena* ribozyme [32, 33] (see also *Section 9* below).

An experiment that clearly distinguishes whether the unfolded states of nucleic acids are extended, random, or compact has recently been carried out

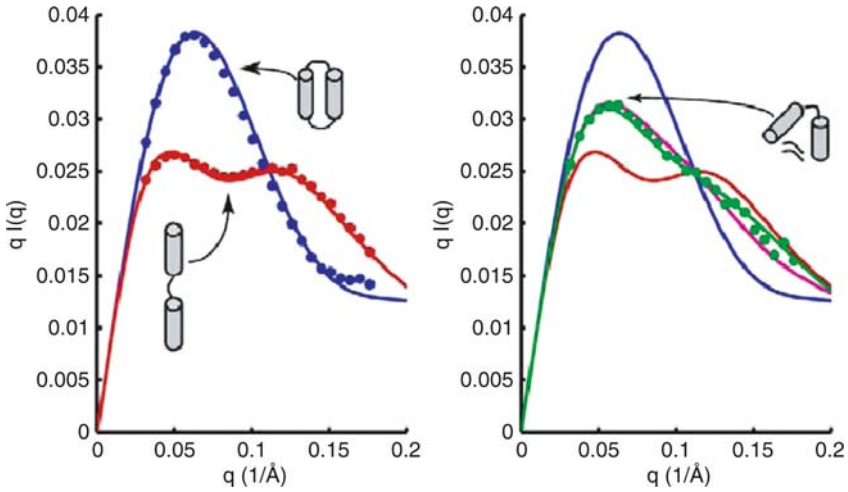

Figure 21-9

SAXS profiles reveal the forces between DNA double helices [34]. (a) Measurements on a pair of singly tethered 12 base-pair double helices in low salt (16 mM Na^+; red symbols) match the prediction for a conformation extended by Coulomb repulsion (red curves). Data for a model system where both ends of each double helix are tethered together (blue symbols) match the prediction for a compact conformation of the helices. (b) Data for the singly tethered duplex in 100 mM $MgCl_2$ (*green symbols*) show no evidence for an attractive force between helices. The full SAXS profile agrees with the prediction for a random ensemble (*green line*) but not a linear combination of the extended state (*red*) and compact state (*blue*) with matching radius of gyration

on an extremely simple model system. To avoid the specific tertiary contacts, metal-ion binding sites, and other complications present in natural RNAs, two 12 base-pair DNA double helices were connected by a flexible hexaethyleneglycol tether (*Figure 21-9*) [34]. Under the low salt conditions typical of *in vitro* experiments, the SAXS profile of the model system indicates a highly extended state, consistent with Coulomb repulsion pushing apart the tethered helices. Such repulsion was also detected in the intermolecular ordering of a concentrated solution of (untethered) DNA helices described in *Section 5*. However, the isolation of two DNA helices in the tethered model system makes quantitative comparison to theoretical predictions straightforward (red line in *Figure 21-9a*).

At high concentrations of monovalent, divalent, and higher valence salt (below the aggregation threshold), the tethered duplex exhibits a radius-of-gyration that is intermediate between an extended state and a compact state (mimicked by a double-tethered DNA; blue symbols and line in *Figure 21-9a*), but the full scattering profile is inconsistent with a linear combination of the two extreme states (magenta lines, *Figure 21-9b*). Rather, the data in high salt (green symbols, *Figure 21-9b*) agrees with the prediction for a random conformational

ensemble, as would be expected if the Coulomb repulsion is screened (green line, *Figure 21-9b*). There is no evidence for a counterion-mediated attractive force inducing a compact denatured state of the model system under any of the tested conditions, indicating that such a force is negligible on the length scale of this small model system in ion concentrations far exceeding typical *in vivo* and *in vitro* conditions. To make these conclusions, it was crucial to use SAXS as a quantitative assay of the global conformation of the tethered helix model system. For example, other assays of molecular conformation based on fluorescence-resonant energy transfer are complicated by nonlinear weighting of different conformational states and by the fortuitous similarity of timescales of probe physics to the motional timescales of the nucleic acid system [35].

8 Anomalous Small-Angle X-Ray Scattering of Biomolecules

A powerful but largely untapped extension of SAXS is the dissection of a specific atom type's contribution to the scattering profile by tuning the x-ray energy near its atomic resonance, also known as its anomalous absorption edge. Although the resonances of typical biomolecular elements (carbon, nitrogen, oxygen, sulfur) are at energies too low to be accessible in synchrotron beamlines, many transition metals and lanthanides that play special roles in biomolecular function can be probed at beamlines equipped with tunable monochromators with tight bandwidths (5 eV or less). The anomalous difference signal is obtained by differencing a scattering profile acquired at an x-ray energy away from the anomalous absorption edge and a profile acquired just a few electron volts below the edge, where the element's scattering is decreased by a small number of electrons f′:

$$I_{\mathrm{anom}}(q) = I_{\mathrm{off-edge}}(q) - I_{\mathrm{on-edge}}(q) \propto f' \sum_{\substack{i = anom.\,element \\ j = molecule}} f_j \frac{\sin qr_{ij}}{qr_{ij}} \quad (13)$$

Thus the anomalous difference signal reports on the distribution of distances r_{ij} between atoms of the probed element (i) and of the whole biomolecule (j); cf. (2). (Smaller terms corresponding to distances between atoms of the probed element itself are usually negligible and are not shown.) Due to the low signal-to-noise of biomolecule scattering, anomalous SAXS (ASAXS) applications to biomolecules are generally far more difficult than its typical applications to metal alloys and other strongly scattering solid-state systems. In particular, anomalous absorption as well as fluorescence by the probed element can appear in the measurements near the anomalous edge; these effects can contaminate the

small anomalous difference scattering signal from biomolecules if they are not corrected with precise and accurate measurements of the beam intensity.

A recent application to nucleic acids illustrates the power of ASAXS. All nucleic acids are surrounded by a diffusely bound "atmosphere" of counterions that help neutralize the backbone's massive negative charge and strongly influencing folding energetics and catalysis [36]. However, the distribution of this atmosphere is invisible to the standard high-resolution structural techniques of crystallography and NMR. Solution scattering has provided a much-needed window into this ubiquitous but fluctuating atmosphere. The spatial extent of the counterions around DNA and transfer RNA molecules have been probed in x-ray and neutron scattering studies since the 1960s [37–40]. More recently, the increased signal-to-noise of anomalous SAXS has made it possible to dissect spatial information for different types of counterion around a DNA, allowing unprecedented tests of the understanding of nucleic acid electrostatics [40, 41].

Figure 21-10 shows anomalous SAXS difference signals measured for a sample of DNA double helices immersed in a solution of monovalent and divalent ions; because of their accessible atomic resonances, the non-physiological ions Rb^+ and Sr^{2+} were used as stand-ins for the more common ions K^+ and Mg^{2+}. Under the probed conditions, both Rb^+ and Sr^{2+} are present in the DNA's ion cloud and yield anomalous signals that can be measured with high precision. Remarkably, the shapes of these two components within the mixed ion cloud are different. Compared to the monovalent ions, the divalent ions exhibit wider ASAXS profiles, reflecting their tighter spatial localization to the DNA and consistent with quantitative predictions based on the widely used non-linear Poisson-Boltzmann model for the counterion distribution (*Figure 21-10b*).

9 Time-Resolved SAXS

With the advent of high-brilliance synchrotron sources, it is now possible to obtain high signal-to-noise SAXS profiles of biomolecule samples with x-ray exposures of milliseconds, or even less with continuous flow mixing [42–44]. Using these beamline setups and appropriate mixers, SAXS allows the time-resolved characterization of the size and shape of systems ranging from small peptides to whole viruses as they undergo changes on time scales much faster than a second. This section describes two examples, for a small protein and for a large RNA, that together illustrate what time scales can be probed and what fundamental insights can be gained in time-resolved SAXS studies. The interested reader is referred to more complete descriptions in recent reviews by Doniach [45] and Koch et al. [1].

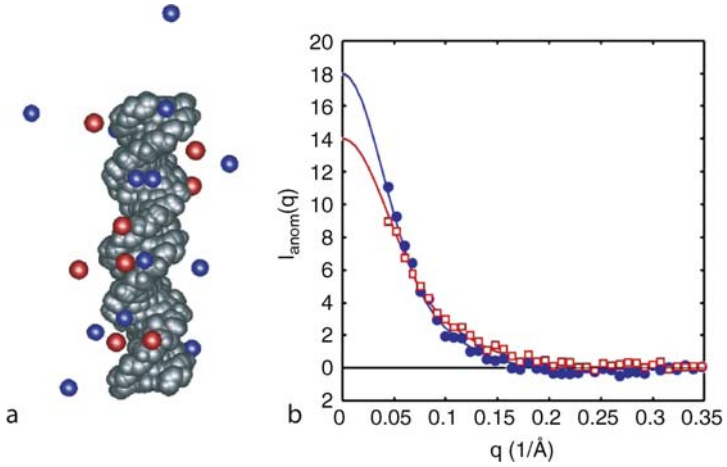

Figure 21-10

Anomalous SAXS dissects the scattering contributions from the counterion cloud enveloping DNA [41]. (a) Simulated snapshot of the dynamic counterion cloud bound to a 25 base-pair DNA double helix (*gray*). The cloud is a mixture of monovalent (*blue*) and divalent (*red*) cations, distributed according to a nonlinear Poisson-Boltzmann calculation for the electrostatic potential, with concentrations in bulk solution (far from the DNA) set to 90 mM monovalent and 10 mM divalent ions. (b) Anomalous SAXS data for the spatial correlations of DNA with Rb^+ ions (*blue, closed symbols*) and with Sr^{2+} ions (*red, open symbols*) in the mixed ion cloud compared to Poisson-Boltzmann predictions. The two measurements were made on the same sample, with bulk cation concentrations of 90 mM Rb^+ and 10 mM Sr^{2+}. The tighter spatial localization of the divalent ions than the monovalent ions is reflected in the wider ASAXS profile for Sr^{2+}

In protein folding, the simplest time-resolved folding processes have been observed for the so-called "two-state folders." When followed by spectroscopic signatures of folding such as circular dichroism and fluorescence, these small proteins exhibit timecourses that follow single exponentials, with the same rate constant measured by different techniques. Through SAXS studies of a two-state folder, the 62-residue protein L, Plaxco et al. have addressed a long-standing controversy in protein folding: whether chemically denatured proteins transferred to native solvent conditions necessarily undergo a "burst-phase" hydrophobic collapse before completing the folding process, or whether the collapse can be concomitant with folding [46]. The folding timecourse of protein L, initiated by dilution of 6.0 M guanidinium hydrochloride (GuHCl) to 1.4 M GuHCl, showed no significant radius-of-gyration decrease in the first 100 ms (*Figure 21-11*). Rather the radius of gyration decreases with a time constant of 8 ± 3 s, consistent with the folding rate observed in fluorescence measurements. These SAXS observations and more precise recent measurements on two other proteins by Sosnick and colleagues [47] demonstrate that rapid chain collapse is not an obligatory feature of protein folding reactions. As discussed in *Section 3*,

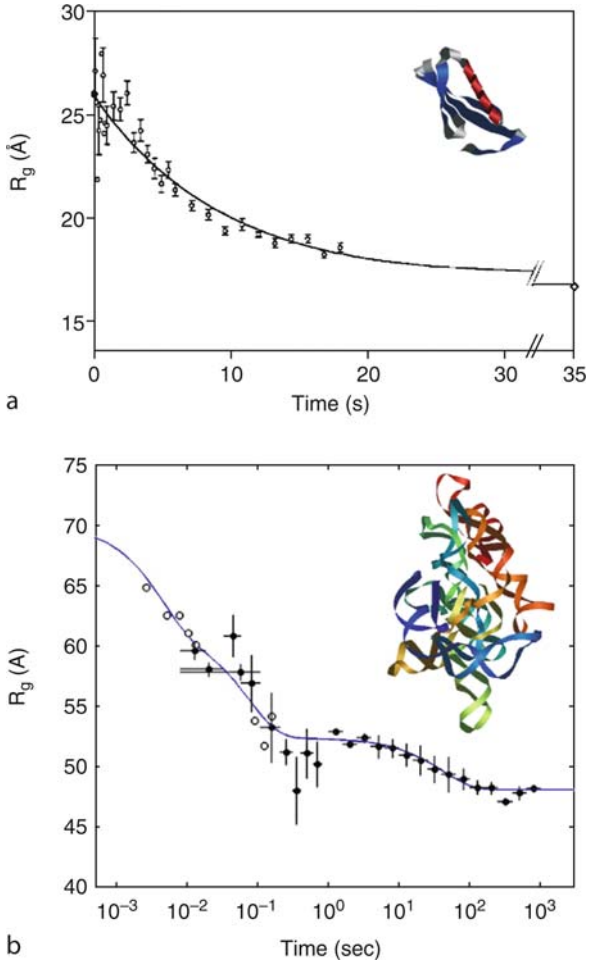

Figure 21-11

Time-resolved SAXS unveils contrasting folding behaviors for two biomolecules of rather different sizes and natures. (a) The small "two-state folder" protein L exhibits an exponential timecourse for compaction consistent with its folding monitored by spectroscopic probes. Figure reproduced from Plaxco et al. [46]. (b) A large RNA, the *Tetrahymena* ribozyme, displays multiple phases of folding (note logarithmic scale for time axis), including early compaction phases with time constants of tens of milliseconds, faster than those expected for stable tertiary contact formation (seconds) and attainment of a catalytically active fold (minutes) [33]

the observations do not rule out the presence of subtle correlations in the initial arrangements of the denatured state that do not impact overall chain dimensions but may still influence the folding pathway.

Time-resolved SAXS can reveal much more intricate behavior than the simple two-state kinetics exhibited by protein L, as has been illustrated by a series of studies on the *Tetrahymena* ribozyme, a 390-nucleotide catalytic RNA [32, 33, 48].

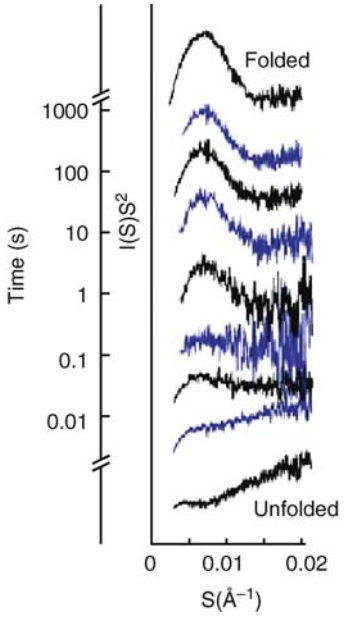

Figure 21-12
Shape changes in the *Tetrahymena* ribozyme folding pathway shown in Figure 11(b) are clearly seen in timecourses displaying the full SAXS profiles. The rising peak in the Kratky profile (s^2I *versus* s; see [45]) indicate transitions from an initially extended conformation to a shape consistent with a random ensemble (10 ms) and then to a compact globule (>100 ms). Each Kratky plot is positioned by its folding time, which increases logarithmically from *bottom* to *top*. Figure reproduced from Russell et al. [32]

The Mg^{2+}-induced folding of this RNA differs from proteins in that its secondary structure of double helices is preformed in the initial conditions. Despite this simplification, the folding is more complex than the behavior seen for small proteins; the ribozyme's radius-of-gyration decreases in at least three phases of folding with timescale of a few milliseconds, ~100 ms, and ~1 min (See *Figure 21-11b* and *Figure 21-12*) [32]. The first two collapse phases were unexpected, because time-resolved synchrotron footprinting studies had suggested that the formation of stable tertiary contacts would occur on timescales of 1 s and slower [49]. A follow-up SAXS study that includes further analysis of full scattering profiles and investigation of mutants indicates that the fastest phase occurs without the aid of tertiary contacts. The first phase is consistent with the simple relaxation of a highly extended unfolded state to a random conformational ensemble, promoted by the screening of initial charge repulsion by Mg^{2+} [33]. Such "electrostatic relaxation" is analogous to the equilibrium conformation change seen in the model tethered-DNA-duplex system described in *Section 7*.

On a practical note, the fast timepoints in the experiments described above require appropriate mixers and triggering systems. Without such equipment,

time resolutions of minutes can still be achieved at most SAXS beamlines by manual mixing [48]. Faster time resolution, down to milliseconds, can be attained by attaching x-ray cuvettes or capillary tubes to commercial stopped-flow mixers, such as the Biologic SFM-400 system adapted by Sosnick and Thiyagarajan for use at the Advanced Photon Source [33, 47]. Finally, sub-millisecond time resolution can now be achieved with specially fabricated continuous flow mixers coupled with extremely high-intensity x-ray beams [32, 43, 44], though this technology is not yet widely available. The experiments with the fastest time resolutions generally require large quantities (many milligrams) of protein or RNA.

10 Final Notes

Recent developments in x-ray beamlines and software have led to substantial improvements in applications of small angle x-ray scattering (SAXS) to determine sizes and shapes of biomolecules in solution. Researchers considering its use should be aware of the amounts of sample required for good signal-to-noise data – these amounts are higher than standard biochemical assays but similar to the requirements of nuclear magnetic resonance experiments – and the sometimes-limited availability of synchrotron SAXS beamlines. Nevertheless, SAXS data can offer fundamental information on biomolecules that is invisible to other techniques: low-resolution reconstructions of biomolecules that evade crystallization, structural characterizations of the partially ordered components of molecules, and time-resolved portraits of molecular size and shape at millisecond resolution. SAXS provides a valuable complement to spectroscopic, computational, and biochemical methods for the structural study of biomolecules and is therefore a highly recommended technique for inclusion in the toolkits of soft matter researchers.

References

1. Koch, M.H., Vachette, P., and Svergun, D.I. (2003) *Q. Rev. Biophys.,* **36**, 147.
2. Feigin, L.A. and Svergun, D.I. (1987) *Structure Analysis by Small-Angle X-ray and Neutron Scattering.* Plenum Press, New York.
3. Svergun, D.I., Richard, S., Koch, M.H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) *Proc. Natl. Acad. Sci. USA,* **95**, 2267.
4. Svergun, D., Barberato, C., and Koch, M.H.J. (1995) *J. Appl. Cryst.,* **28**, 768.
5. Moore, P.B. (1980) *J. Appl. Cryst.,* **13**, 168.
6. Svergun, D.I., Semenyuk, A.V., and Feigin, L.A. (1988) *Acta Cryst.,* **A44**, 244.
7. Svergun, D.I. (1992) *J. Appl. Cryst.,* **25**, 495.
8. Millett, I.S., Doniach, S., and Plaxco, K.W. (2002) *Adv. Protein Chem.,* **62**, 241.

9. Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., Hasan, M.Z., Pande, V.S., Ruczinski, I., Doniach, S., and Plaxco, K.W. (2004) *Proc. Natl. Acad. Sci. USA,* **101**, 12491.
10. Fitzkee, N.C. and Rose, G.D. (2004) *Proc. Natl. Acad. Sci. USA,* **101**, 12497.
11. Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C.M., Poulsen, F.M., and Vendruscolo, M. (2004) *J. Am. Chem. Soc.,* **126**, 3291.
12. Segel, D.J., Fink, A.L., Hodgson, K.O., and Doniach, S. (1998) *Biochemistry,* **37**, 12443.
13. Fang, X.W., Yang, X.J., Littrell, K., Niranjanakumari, S., Thiyagarajan, P., Fierke, C.A., Sosnick, T.R., and Pan, T. (2001) *RNA,* **7**, 233.
14. Bonnete, F. and Vivares, D. (2002) *Acta Cryst.,* **D58**, 1571.
15. Vivares, D. and Bonnete, F. (2002) *Acta Cryst.,* **D58**, 472.
16. Skibinska, L., Gipinski, J., Liu, H., Patkowski, A., Fischer, E.W., and Pecora, R. (1999) *J. Chem. Phys.*, 1794.
17. Chacon, P., Moran, F., Diaz, J.F., Pantos, E., and Andreu, J.M. (1998) *Biophys. J.,* **74**, 2760.
18. Walther, D., Cohen, D., and Doniach, S. (2000) *J. Appl. Cryst.,* **33**, 350.
19. Svergun, D. (1999) *Biophys. J.,* **76**, 2879.
20. Svergun, D.I., Petoukhov, M.V., and Koch, M.H.J. (2001) *Biophys. J.,* **80**, 2946.
21. Stuhrmann, H.B. (1970) *Acta Cryst.,* **A26**, 297.
22. Takahashi, Y., Nishikawa, Y., and Fujisawa, T. (2003) *J. Appl. Cryst.,* **36**, 549.
23. Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003) *Cell,* **113**, 369.
24. Gutsche, I., Essen, L.O., and Baumeister, W. (1999) *J. Mol. Biol.,* **293**, 295.
25. Funari, S.S., Rapp, G., Perbandt, M., Dierks, K., Vallazza, M., Betzel, C., Erdmann, V.A., and Svergun, D.I. (2000) *J. Biol. Chem.,* **275**, 31283.
26. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) *Cell,* **107**, 679.
27. Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000) *Science,* **289**, 905.
28. Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.P., Ehresmann, C., and Ehresmann, B. (1989) *J. Mol. Biol.,* **207**, 417.
29. Grosberg, A.Y., Nguyen, T.T., and Shklovskii, B.I. (2002) *Rev. Mod. Phys.,* **74**, 329.
30. Bloomfield, V.A. (1997) *Biopolymers,* **44**, 269.
31. Oosawa, F. (1968) *Biopolymers,* **6**, 134.
32. Russell, R., Millett, I.S., Tate, M.W., Kwok, L.W., Nakatani, B., Gruner, S.M., Mochrie, S.G., Pande, V., Doniach, S., Herschlag, D., and Pollack, L. (2002) *Proc. Natl. Acad. Sci. USA,* **99**, 4266.
33. Das, R., Kwok, L.W., Millett, I.S., Bai, Y., Mills, T.T., Jacob, J., Maskel, G.S., Seifert, S., Mochrie, S.G., Thiyagarajan, P., Doniach, S., Pollack, L., and Herschlag, D. (2003) *J. Mol. Biol.,* **332**, 311.
34. Bai, Y., Das, R., Millett, I.S., Herschlag, D., and Doniach, S. (2005) Probing counterion modulated repulsion and attraction between nucleic acid duplexes in solution. *Proceedings of the National Academy of Sciences USA*, **102**, 959–960.
35. Murphy, M.C., Rasnik, I., Cheng, W., Lohman, T.M., and Ha, T. (2004) *Biophys. J.,* **86**, 2530.
36. Draper, D.E. (2004) *RNA,* **10**, 335.
37. Bram, S. and Beeman, W.W. (1971) *J. Mol. Biol.,* **55**, 311.
38. Zaccai, G. and Xian, S.-Y. *Biochemistry,* **27**, 1316.
39. Zakharova, S.S., Egelhaaf, S.U., Bhuiyan, L.B., Outhwaite, C.W., Bratko, D., and Maarel, van der J.R.C. (1999) *J. Chem. Phys.,* **111**, 10706.
40. Das, R., Mills, T.T., Kwok, L.W., Maskel, G.S., Millett, I.S., Doniach, S., Finkelstein, K.D., Herschlag, D., and Pollack, L. (2003) *Phys. Rev. Lett.,* **90**, 188103.
41. Andresen, K., Das, R., Park, H.Y., Smith, H., Kwok, L.W., Lamb, J.S., Kirkland, E.J., Herschlag, D., Finkelstein, K.D., and Pollack, L. (2004) Spatial distribution of competing ions around DNA in

solution. *Physical Review Letters* **93**, 248013.
42. Segel, D.J., Bachmann, A., Hofrichter, J., Hodgson, K.O., Doniach, S., and Kiefhaber, T. (1999) *J. Mol. Biol.,* **288**, 489.
43. Pollack, L., Tate, M.W., Darnton, N.C., Knight, J.B., Gruner, S.M., Eaton, W.A., and Austin, R.H. (1999) *Proc. Natl. Acad. Sci. USA,* **96**, 10115.
44. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y., and Fujisawa, T. (2002) *Proc. Natl. Acad. Sci. USA,* **99**, 1329.
45. Doniach, S. (2001) *Chem. Rev.,* **101**, 1763.
46. Plaxco, K.W., Millett, I.S., Segel, D.J., Doniach, S., and Baker, D. (1999) *Nat. Struct. Biol.,* **6**, 554.
47. Jacob, J., Krantz, B., Dothager, R.S., Thiyagarajan, P., and Sosnick, T.R. (2004) *J. Mol. Biol.,* **338**, 369.
48. Russell, R., Millett, I.S., Doniach, S., and Herschlag, D. (2000) *Nat. Struct. Biol.,* **7**, 367.
49. Sclavi, B., Sullivan, M., Chance, M.R., Brenowitz, M., and Woodson, S.A. (1998) *Science,* **279**, 1940.