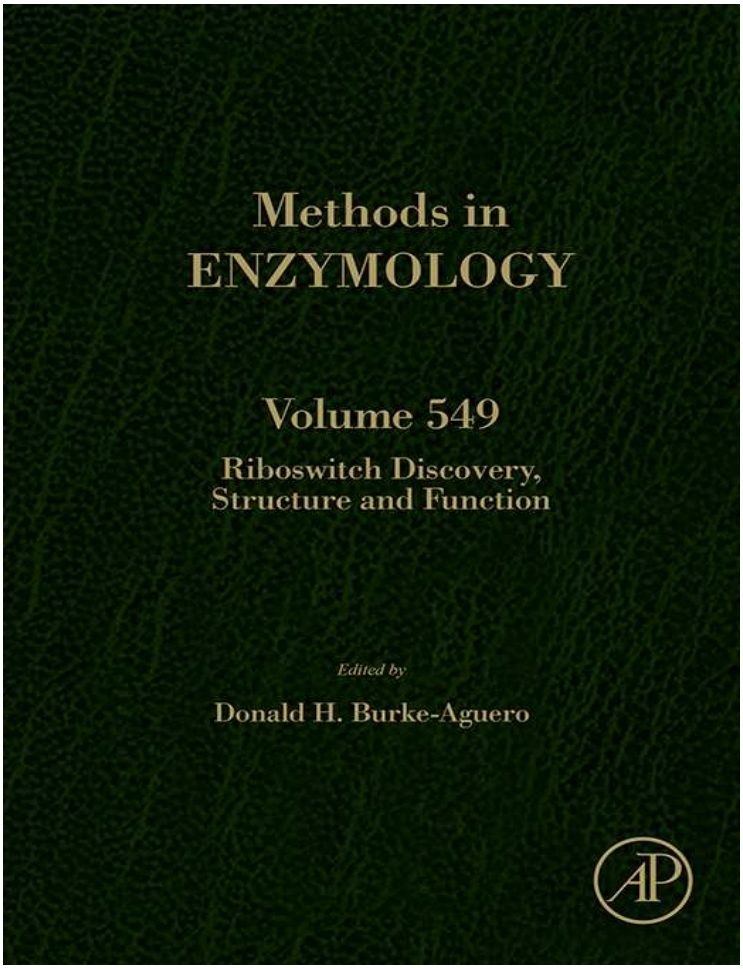
Methods for Using New Conceptual Tools and Parameters to Assess RNA Structure by Small-Angle X-Ray Scattering

ARTICLE *in* METHODS IN ENZYMOLOGY · NOVEMBER 2014

Impact Factor: 2.09 · DOI: 10.1016/B978-0-12-801122-5.00011-8 · Source: PubMed



From Francis E. Reyes, Camille R. Schwartz, John A. Tainer and Robert P. Rambo, Methods for Using New Conceptual Tools and Parameters to Assess RNA Structure by Small-Angle X-Ray Scattering. In: Donald H. Burke-Aguero, editor, *Methods in Enzymology, Vol. 549*, Burlington: Academic Press, 2014, pp. 235-263.
ISBN: 978-0-12-801122-5

CHAPTER ELEVEN

Methods for Using New Conceptual Tools and Parameters to Assess RNA Structure by Small-Angle X-Ray Scattering

Francis E. Reyes*,[1], Camille R. Schwartz*, John A. Tainer†,‡, Robert P. Rambo*,[2,3]

*Physical Bioscience Division Lawrence Berkeley National Lab, Berkeley, California, USA
†Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
‡The Scripps Research Institute, La Jolla, California, USA
[1]Current address: Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia, USA
[2]Current address: Diamond Light Source Ltd, Harwell Science & Innovation Campus, Didcot, United Kingdom
[3]Corresponding author: e-mail address: robert_p_rambo@hotmail.com

Contents

Abstract

Understanding the biological activities of riboswitches and of RNA in general requires a thorough analysis of both the spatial arrangement of the residues and the dynamics of

Methods in Enzymology, Volume 549
ISSN 0076-6879
http://dx.doi.org/10.1016/B978-0-12-801122-5.00011-8

the structural ensemble. Specifically, evaluating the structural basis for riboswitch function requires analyses of many relevant states that include ligand-bound and -free, high Mg^{2+}, and quite possibly, the active transcription state, which is challenging to achieve by most methods. Small angle X-ray scattering (SAXS) is an enabling technique for comprehensive analyses of RNA structures in solution. Here, we describe recent SAXS tools and technologies that substantially improve the potential for accurate and comprehensive analyses of flexibility, unstructured elements, conformational selection, and induced fit in RNA function. We note equipment needed plus appropriate annealing and purification procedures. We describe key model-independent parameters (SAXS invariants) which can be used to monitor changes in a particle's thermodynamic state: the Guinier-based R_g, the volume-of-correlation (V_c), the Porod–Debye exponent (P_E), and the power-law parameter, Q_R, that determines mass directly from the SAXS data. We also consider the value of real-space parameters and of multiphase modeling with MONSA to locate secondary structure elements within SAXS volumetric envelopes. For conformation changes, experiments with nanogold-labeled RNA analyzed using the SAXS structural comparison map and volatility ratio difference metric enable high-throughput evaluation of solution-state conformations. Collectively, the described tools and procedures enable quantitative and comprehensive measures of riboswitch structures with general implications for our views and strategies of RNA structural analysis.

1. INTRODUCTION

X-ray crystallographic studies of riboswitch aptamer domains with their cognate ligands have yielded tremendous insight into how RNA can be harnessed to recognize small molecule ligands specifically. The current repertoire of small molecule ligands includes amino acids, nucleobases, nucleotides, metals, and cofactors (Peselis & Serganov, 2014). Naturally, these studies have been extended to understand how riboswitches discriminate among related molecules in a complex cellular environment (Gilbert & Batey, 2009; Johnson, Reyes, Polaski, & Batey, 2012; Serganov, Huang, & Patel, 2008; Trausch et al., 2014). While "equilibrium" studies of riboswitches are easy to perform and are useful for guiding virtual screening of alternative ligands (Daldrop et al., 2011), a number of studies have suggested that such conditions are far from physiological.

Biochemical and structural investigations of full-length riboswitches, as opposed to solely their aptamer domains, have reinforced the notion that a complete understanding of riboswitch function requires a consideration of its behavior in the context of its free state or active transcription. For instance, the add riboswitch from *Vibrio vulnificus* displays a three-state behavior in limiting amounts of adenine (Reining et al., 2013). Several

studies have observed multiple conformations in ligand-free aptamers (Chen, Zuo, Wang, & Dayie, 2012; Stoddard et al., 2010; Vicens, Mondragon, & Batey, 2011), suggesting that some riboswitch aptamer domains are preorganized in the absence of ligand, sampling a number of binding-competent states that may be different than the bound state, otherwise known as "conformational selection" (Zhang, Jones, & Ferre-D'Amare, 2014). Strikingly, transcriptional pausing by the *Escherichia coli* btuB riboswitch can have profound effects on RNA folding (Perdrizet, Artsimovitch, Furman, Sosnick, & Pan, 2012). Conceivably, a balancing act among speed of transcription, RNA folding, and ligand is carefully maintained for efficient regulation (Frieda & Block, 2012; Garst & Batey, 2009). Taken together, these studies reveal the importance of ligand-free studies of riboswitches toward understanding their structure and function.

Conformational heterogeneity of riboswitches in the free state precludes the use of X-ray crystallography for structural studies. Instead, solution-state experiments such as nuclear magnetic resonance, hydroxyl-radical probing, selective 2′-hydroxyl acylation analyzed by primer extension (John, Merino, & Weeks, 2004) and small angle X-ray scattering (SAXS) (Perry & Tainer, 2013; Petoukhov & Svergun, 2013; Rambo & Tainer, 2010a, 2013b; Sibille & Bernado, 2012) can be used for comprehensive structural analyses. SAXS is a solution-state measurement that requires minimal sample (<17 μL) and time (Hura et al., 2009). An entire experiment can be performed in less than 4 minute making the experiment highly efficient and economical. SAXS is a robust technique that can be performed on a wide range of RNA masses and solution conditions (Baird & Ferre-D'Amare, 2010; Gopal, Zhou, Knobler, & Gelbart, 2012; Rambo & Tainer, 2010b). The development of modern SAXS synchrotron beamlines, as typified by SIBYLS at the Advanced Light Source (Classen et al., 2013) and of commercial laboratory SAXS instruments, has made SAXS measurements readily available for routine structural investigations. However, SAXS is limited by sample quality and requires extensive evaluations of sample heterogeneity for reliable interpretation (Perez & Nishino, 2012; Rambo & Tainer, 2010b). We present general considerations for the robust application of SAXS to RNA and riboswitches in particular. In this chapter, we present specific recommendations when considering a structural study of RNAs by SAXS and, using the SAM-I, adenosylcobalamin, and lysine riboswitches illustrate the role SAXS can play in elucidating new mechanistic and structural insights These riboswitch systems exemplify general issues in structural biology regarding unstructured components, flexibility, and interactions, in

addition to conformational switching, selection, and induced fit. The equipment, procedures, parameters, and modeling described here provide the means to enable efficient quantitative and comprehensive measures of RNA structures.

2. SPECIALIZED EQUIPMENT

We refer the reader to the review on *in vitro* transcription (Beckert & Masquida, 2011). All analysis was performed with the JAVA-based program ScÅtter available at www.bioisis.net. We suggest the following equipment for performing the analysis discussed in this chapter (Tables 11.1–11.3).

3. PREPARATION OF THE RNA FOR A SAXS STUDY

3.1. Assessing the folded state of the RNA

SAXS offers the opportunity to perform a structural analysis of a wide variety of macromolecular RNA types (Gajda, Martinez Zapien, Uchikawa, & Dock-Bregeon, 2013; Rambo & Tainer, 2010b; Yang, Parisien, Major, & Roux, 2010) that not only includes the canonical classes of structured RNA (e.g., introns, riboswitches, ribozymes) but also the classes of seemingly unstructured RNA described as either noncoding or single stranded (Gopal et al., 2012; Ribitsch et al., 1985). However, as SAXS is a solution-scattering technique, care must be used to ensure homogenous samples. In most cases, the RNA of interest is *in vitro* transcribed, followed by a denaturing purification and then renatured through an annealing procedure (Fig. 11.1). The denaturing purification typically leads to chemically homogenous RNA (Fig. 11.1B) observed as a single migrating band on an 8 *M* urea denaturing polyacrylamide gel electrophoresis (PAGE).

Table 11.1 Silica hydrophilic polymer-based size exclusion column (SEC) withstand high operating pressures at high flow rates without sacrificing column efficiency

Shodex SEC column	Exclusion limit (kDa)	Column volume (mL)
KW402.5-4F	150	4.8
KW403-4F	600	4.8
KW404-4F	1000	4.8

The columns are ideal for screening and for the KW series, a typical SEC run will take 16 minutes Columns are not stable to 1 *M* NaOH washes.

Table 11.2 High-performance liquid chromatography (HPLC) instruments for analytical scale preparation of RNA samples for SAXS

HPLC	Note
AKTA Ettan microLC[a]	Can equilibrate and perform several runs with 50-mL of running buffer, ideal for including ligands in buffer during SEC separations
Agilent Infinity 1260[a]	Operates in continuous flow mode with an autoinjector. Ideal for inline light scattering and SAXS instrumentation
BioRad BioLogic	Can be used with larger (15 mL) KW 800 series Shodex columns for processing larger sample quantities. Required flow rates are 1–1.5 mL/min

[a]These HPLCs are specific to ultrasmall flow rates such as 50 μL/min and must be able to withstand tubing diameters of 100–150 μm. Large tubing diameters should be avoided to reduce sample dilution during purification. The instruments are analytical scale and need to be restricted to nonpreparative workloads in order to maintain a clean system for RNA studies. To reduce background from buffer contaminants, we recommend installation of an inline 0.1-μm filter after the HPLC pump but prior to the sample loading valve.

Table 11.3 Thorough characterization of the hydrodynamic state of an RNA sample can be reliably achieved using light scattering analysis during SEC purification

Accessories	Source	Note
Multiangle light scattering (MALS) instrument	Wyatt	18-angle detector is preferred for studies of large macromolecular assemblies
Quasi-elastic light scattering (QELS) detector	Wyatt	Time-resolved light scattering measurements that can determine radius-of-hydration
Refractive index detector	Wyatt	Conformation-independent assessment of RNA or protein concentration
Inline buffer filter (<0.1 μm)	Millipore	Reduces baseline noise in light scattering detectors
Inline buffer vacuum degasser	Rheodyne	Reduces baseline noise in light scattering and refractive index detector

MALS provides absolute mass measurements across the SEC elution peak assessing mass heterogeneity, whereas QELS provides relative estimates of the radius-of-hydration assessing conformational heterogeneity. Unlike fixed wavelength absorbance-based concentration detectors that typically monitor proteins or nucleic acids, refractive index (rI) detectors monitor optical differences between a reference buffer and sample. rI detectors can reveal small differences between a sample and buffer background due to salts and dissolved gases.

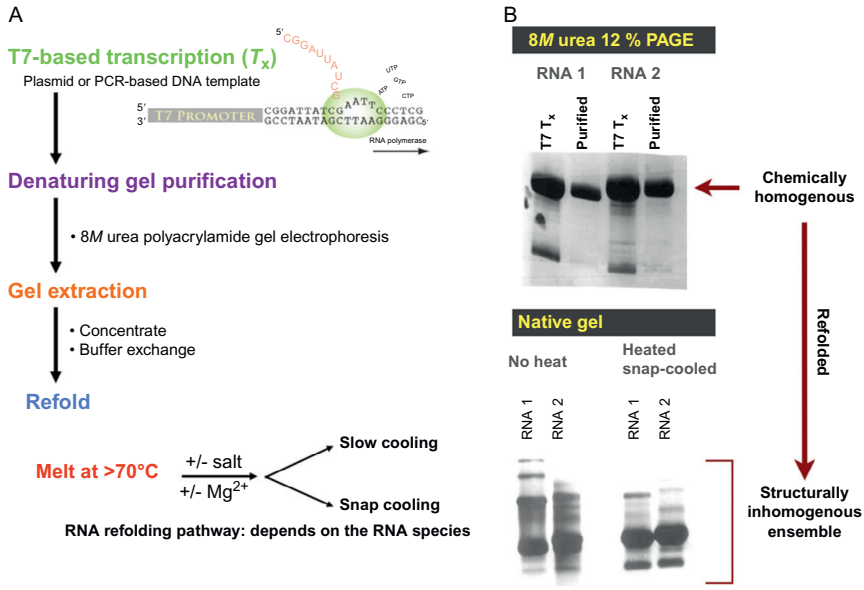


Figure 11.1 General outline for purification of *in vitro* synthesized RNA for structural purposes. (A) T7 RNA polymerase-based methods for synthesizing large amounts of template-directed RNA using either plasmid or PCR amplified DNA templates. Following enzymatic synthesis, RNA is purified by electrophoresis through an 8 *M* urea polyacrylamide denaturing gel (slab) to remove truncated transcription products and enzyme. The RNA of interest (ROI) is visualized by UV shadowing and physically excised as a gel slice. The ROI is extracted from the gel, exchanged into an appropriate storage buffer, and concentrated. Since the RNA was purified by denaturing methods, an annealing step is taken to refold the RNA. (B) Denaturing purification of SAM-I riboswitch (94 nucleotides). Ethidium bromide staining clearly demonstrates product heterogeneity before gel extraction (T7 T_x lanes), whereas gel extraction and purification produces a single mass species (purified lanes). Native gel purification of refolded, purified RNA leads to a polydisperse sample and ambiguity regarding mass of each band. Snap cooling of the RNA (annealing) reduces higher mass species (no heat) but not completely. *Denaturing and native PAGE of SAM-I courtesy of Robert Batey, University of Colorado at Boulder.*

Subsequently, the chemically homogenous RNA is renatured and analyzed using native PAGE (Fig. 11.1B).

Native PAGE has been tremendously useful for understanding RNA structure and folding (Woodson & Koculi, 2009); however, the method is inadequate for evaluating the quality of the RNA sample for SAXS studies due to the inability to assess the mass of the RNA reliably or to detect the presence (<5%) of misfolded species (Rambo & Tainer, 2010b). The presence of misfolded species can confound SAXS interpretation and limit the

usefulness of the SAXS data. As an alternative method for accessing RNA sample quality, we suggest using size exclusion chromatography (SEC) coupled inline with a multiangle light scattering (MALS) instrument (Rambo & Tainer, 2010b). We direct the reader to an excellent review on the theory and application of MALS (Wyatt, 1993). Furthermore, entire SEC runs can be performed in less than 15 minute using modern high precision silica-based size exclusion columns. These columns can withstand high pressures without sacrificing column resolution (Table 11.1).

The MALS instrument is used principally for determining the molecular mass (MM) of the eluting peaks during an SEC separation. Since the scattering process depends only on the square of the particle mass, the MALS instrument can observe non-RNA or -protein species such as residual polyacrylamide during the SEC separation. Residual polyacrylamide carried over from the denaturing purification step is invisible at wavelengths between 260 and 280 nm but is readily visible by MALS and the associated refractive index (rI) detector. Ideally, the MALS instrument will have greater than 16 detectors with an integrated quasi-elastic light scattering (QELS) detector. The additional MALS detectors provide angular scattering information for an independent determination of a particle's radius-of-gyration, R_g, whereas the QELS information will determine the radius-of-hydration, R_h. In the absence of SAXS information, knowing the mass, R_h, and R_g can immediately suggest the shape of the RNA. Furthermore, changes in RNA compaction can be monitored directly by following R_h.

Mass determination from SEC-MALS requires an accurate concentration measurement across the profile of the elution peak. This is reliably achieved by simultaneous monitoring of the rI during the chromatographic separation. In the solution state, the RNA rI increment (the change in rI per change in RNA concentration) is independent of particle size, shape or level-of-compactness and can be used to determine RNA concentration. However, rI does depend on buffer composition: therefore, the experimental rI must be determined from a standard such as the P4–P6 domain (Rambo & Doudna, 2004; Rambo & Tainer, 2010b). Figure 11.2A illustrates information obtained from an SEC-MALS run. Based on the concentration detector alone (i.e., rI or UV monitor), the SEC profile would be described by two major peaks between 14 and 22 minutes. However, inspection of the light scattering signal during the SEC separation shows three peaks. Since the light scattering intensity is proportional to particle mass, we can conclude that peak 1 is a massive aggregate at exceptionally low concentration that would contribute a scattering signal nearly equal

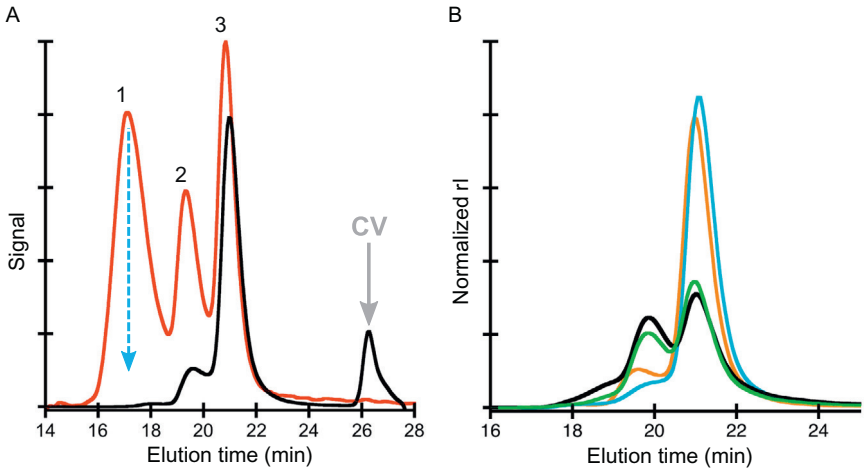


Figure 11.2 Size exclusion chromatographic (SEC) separation of SAM-I riboswitch under varying annealing conditions. (A) Elution profile of SAM-I riboswitch monitored by refractive index (concentration detector, black) and light scattering (mass detector, red). Cyan and gray arrows mark the void and column volumes, respectively. Peak 1 is an aggregate peak given by the high light scattering at very low concentration. Peak 2 is an RNA dimer. Peak 3 corresponds to the RNA monomer. (B) Elution profile (refractive index) of SAM-I riboswitch annealed using four different refolding strategies. RNA was heated (85 °C) and either snap- or slow cooled in the presence and absence of Mg^{2+} under dilute conditions (~0.1 mg/mL). In the absence of Mg^{2+}, snap cooled using pre-chilled buffer (cyan) performed better than slow cooling (orange) which performed better than snap cooling (green) and slow cooling (black) in the presence of Mg^{2+}. Annealing in the presence of Mg^{2+}, results in severe material loss of monomeric RNA. RNA was concentrated to 5 mg/mL using a 5-kDa cut-off spin concentrator at 18 °C.

to peak 3. The aggregate peak would corrupt the observed SAXS profile, thereby limiting the scope and interpretation of the experiment.

Based on the SEC profile and analysis of the MALS data in Fig. 11.2A, an optimal SAXS sample can be obtained by sampling the eluting peak directly. In this case, peak 3 is monodisperse (symmetric elution peak containing a single mass species) and corresponds to the monomeric form of the RNA. If the SAXS instrument cannot be directly connected inline to the SEC instrument, then a sample must be taken manually by attaching a small (<1 μL) outlet tube directly to the UV cell. It can be anticipated that injecting a 25–50 μL sample will distribute the sample elution peak over a volume of ~500 μL. For a sample with ~10% heterogenous contamination, we observe a five- to sixfold dilution based on the UV absorbance at peak elution. Therefore, a sample injected at 5–10 mg/mL will yield ~1 mg/mL

SAXS sample at peak elution, which is an ideal concentration for modern SAXS instruments. However, if the sample contains significant heterogenous contamination, then there will be a significant material loss of the RNA of interest due to the RNA being distributed over the heterogenous state (Fig. 11.2B). In these situations, it may be best to optimize the refolding strategy of the RNA or seek a native purification method (Batey & Kieft, 2007; Gopal et al., 2012; Kazantsev et al., 2011).

Through our many collaborations with RNA crystallography research groups, we have examined a wide range of RNAs that include the SAM-I, lysine, and B_{12} riboswitches, P4–P6 RNA domain, HDV and CPEB3 ribozyme, tRNA, tRNA-like RNAs, RNase P, and pre-Q1 (Costantino, Pfingsten, Rambo, & Kieft, 2008; Garst, Heroux, Rambo, & Batey, 2008; Hammond, Rambo, Filbin, & Kieft, 2009; Hammond, Rambo, & Kieft, 2010; Johnson et al., 2012; Kazantsev et al., 2011; Rambo & Tainer, 2010b; Stoddard et al., 2010; Strulson, Yennawar, Rambo, & Bevilacqua, 2013). In each case, after refolding, the RNA sample always contained contaminating species that required optimization of the refolding protocol and SEC purification. From our experience, we suggest a refolding strategy outlined in Fig. 11.3. We have found that annealing the RNA under dilute conditions (~0.1 mg/mL) in the absence of Mg^{2+} has been critical to minimizing intermolecular multimerization. As demonstrated by the SAM-I riboswitch in Fig. 11.2B, the optimal refolding strategy is characterized by the absence of Mg^{2+} during annealing, with snap cooling (using prechilled buffer) performing better than slow cooling (judged by the larger leading elution peak). Annealing the RNA in the presence of Mg^{2+}, under dilute conditions produced a heterogenous sample with nearly 50% of the RNA lost to a heterogenous state. In some cases (e.g., B_{12} riboswitch and RNase P), a refolding strategy that adequately promoted the RNA into a monodisperse, homogenous state could not be determined. For RNase P, a native purification method that optimized monodispersity and catalytic activity was developed. Here, both A- and B-type RNase P were transcribed in a high-salt condition that was optimal for RNA-only catalyzed activity. Interestingly, this RNA was more compact than either the refolded or low salt (Kazantsev et al., 2011), suggesting that cotranscriptional folding yields greater natively folded RNA.

For RNAs in the unstructured class, purification methods will need to be developed that maintain native-like conditions. It can be anticipated that these RNAs will have local structures that form during transcription, and attempts to refold the RNA will likely create a heterogenous kinetically

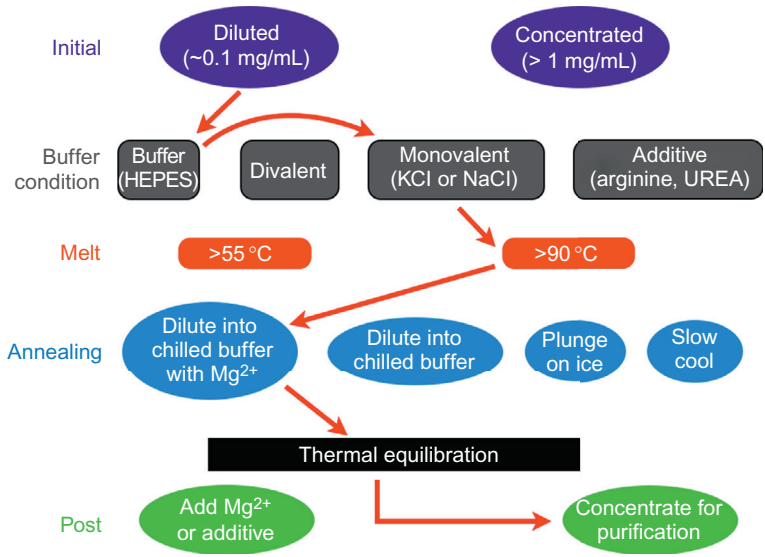


Figure 11.3 Template for generating and testing annealing strategies. A good annealing strategy should minimize material loss of the RNA, i.e., minimize formation of dimer or aggregated forms of the RNA. For "buffer condition" select one or more followed by choosing a melting temperature. The annealing step can be performed using prechilled buffer with or without Mg^{2+} followed by a thermal equilibration at room temperature. The RNA will be dilute and will require concentration step prior to SEC separation.

trapped state. Gopal et al. (2012) examined three single-stranded RNAs from 1000 to 2800 nucleotides (nt) by SAXS and used a mild treatment (phenol–chloroform extraction) to remove transcription-related proteins followed by an ethanol precipitation.

3.2. Importance of buffer subtraction

A solution-state biological SAXS curve is a difference measurement taken as the difference in scattering intensities between the sample and corresponding buffer. Here, buffer refers to the mixture of salts, additives (e.g., glycerol, sucrose, etc.) and weak acid or base used to suspend the biological molecule in a water environment and sample refers to the RNA of interest suspended in the buffer. The observed intensity differences between the sample and buffer are small (<1%); therefore, to minimize systematic errors in the difference, it is necessary that the buffer be as closely matched as possible to the sample buffer. Buffer matching can be readily accomplished using (1) spin concentrator, (2) microdialysis, or (3) SEC. Spin concentrators

are effective but must be washed with buffer at least three times prior to buffer exchanging the sample. Here, the flow through buffer can be used as the matching buffer for SAXS. Dialysis is also effective for concentrated samples but requires much longer equilibration times and may not be suitable for samples that are unstable at high concentrations. The prior two methods are suitable for sample concentrations >1 mg/mL; however, for RNA samples <0.5 mg/mL, exact matching of the buffer is critical to the accuracy of the intensity difference. For such dilute samples, it is recommended to use SEC purified samples and the corresponding buffer. As seen in Fig. 11.2A, the rI detector shows a small peak occurring at the column volume and corresponds to the differences in small molecules between the injected sample and buffer (e.g., salts, dissolved gases, glycerol). Therefore, to ensure a closely matched buffer, it is recommended the SAXS buffer be taken after 1.2 column volumes.

4. INTERPRETATION OF THE X-RAY SCATTERING CURVE

X-rays are scattered by electrons and the excess intensity that gives the observed SAXS difference will be due to the collection of electrons comprising the RNA and associated bound waters. The three-dimensional collection of electrons forming the RNA structure defines the molecular form factor, $A(q)$, and we assume the RNAs are noninteracting (negligible interparticle interference) forming a monodisperse sample.

There are many methods for calculating a SAXS profile from a set of atomic coordinates. These methods can be divided into two classes based on their explicit use of atomic scattering form factors. Algorithms such as SASTBX (Liu, Morris, Hexemer, Grandison, & Zwart, 2012), CRYSOL (Svergun, Barberato, & Koch, 1995), Aqua-SAXS (Poitevin, Orland, Doniach, Koehl, & Delarue, 2011), and AXES (Grishaev, Guo, Irving, & Bax, 2010) use atomic scattering form factors directly, whereas FOXS (Schneidman-Duhovny, Hammel, Tainer, & Sali, 2013) and DALAI-GA (Chacon, Diaz, Moran, & Andreu, 2000) use the $P(r)$-distribution (histogram of the set of interatomic distance vectors) scaled by an approximation function. The $P(r)$-distribution is a resolution-limited distribution that fully describes the structural state of the RNA. Since we are examining the solution state of the RNA, a SAXS measurement will scatter from ~10,000 billion RNAs during a single exposure. Therefore, the $P(r)$-distribution will be the sum of each of the RNAs and is a description of the thermodynamic

state. In general, SAXS intensity calculations from $P(r)$-distributions is given by Eq. (11.1):

$$I_{\text{molecule}}(q) = \int P(r) \cdot \frac{\sin(q.r)}{q.r} \mathrm{d}r \tag{11.1}$$

where q is the scattering momentum vector with units of Å^{-1}, r is the interatomic distance vector (Å), and $I(q)$ is the scattered intensity at the specified q. The integration is limited to the maximum dimension, $d_{\max}$, of the RNA. A more detailed explanation of scattering theory and applications can be found in several comprehensive reviews (Glatter & Kratky, 1982; Rambo & Tainer, 2013b; Svergun, Feĭgin, & Taylor, 1987).

4.1. Quantitating compactness

Approximations to the $P(r)$-distribution can be used to provide resolution-limited information regarding macroscopic properties of the RNA. For a polymer exhibiting a Gaussian distribution between all intersegmental distances, Debye (Glatter & Kratky, 1982) derived an analytical expression (Eq. 11.2) showing that $I(q)$ is dependent on the product $(q \cdot R_g)^2$.

$$I(q) = \frac{2\left(\mathrm{e}^{-R_g^2 \cdot q^2} + R_g^2 \cdot q^2 - 1\right)}{\left(R_g^2 \cdot q^2\right)^2} \tag{11.2}$$

For RNA, intersegmental units refer to the smallest rigid units (nucleotides or small base-paired regions) displaying the aforementioned Gaussian distribution. The expression in Eq. (11.2) shows that within a limited q-range, the intensity approaches an asymptotic limit.

$$\lim_{q \to \infty} I(q) \cdot q^2 = \frac{2}{R_g^2}\left(1 - \frac{1}{q^2 \cdot \mathrm{R}_g^2}\right)$$

The asymptotic limit is readily evident as a plateau (Rambo & Tainer, 2011) in a $q^2 \cdot I(q)$ versus q^2 plot (Debye–Kratky plot).

In contrast, both Porod and Debye (Debye, Anderson, & Brumberger, 1957; Glatter & Kratky, 1982) showed for particles with a discrete surface and electron density contrast with the surrounding buffer, $I(q)$ is directly proportional to q^{-4} scaled by a constant k (Eq. 11.3).

$$I(q) = k \cdot \frac{1}{q^4} \tag{11.3}$$

Similar to the Debye approximation from above, the q^{-4} relationship is valid within a limited q-range and is evident by a plateau (Rambo & Tainer, 2011) in a $q^4 \cdot I(q)$ versus q^4 plot (Porod–Debye plot).

The Debye and Porod–Debye approximations define a q-limited power-law relationship (Fig. 11.4) that can be used to characterize the thermodynamic state of the RNA. The approximations are low resolution where in the case of the Gaussian-like coil, $I(q)$ depends only on the macroscopic property R_g. Likewise, for the Porod–Debye approximation, the scattering will depend on the particle's surface area, a well-defined quantity for a compact particle. Transitions of the RNA from a compact particle with a well-defined surface area to a Gaussian-like chain can be visually assessed by creating three types of plots (Fig. 11.5): Debye–Kratky plot, $q^3 \cdot I(q)$ versus q^3 (SIBYLS) plot and Porod–Debye plot. Here, the plot that demonstrates the most hyperbolic-like asymptotic plateau in the low q-range data suggests the power of the exponent. The asymptotic plateau occurs immediately after the Guinier region ($q \cdot R_g > 1.3$) and is observed after the major inflection point in one of the aforementioned plot types. It should be noted that particles can have more than one Porod–Debye region (Ciccariello,

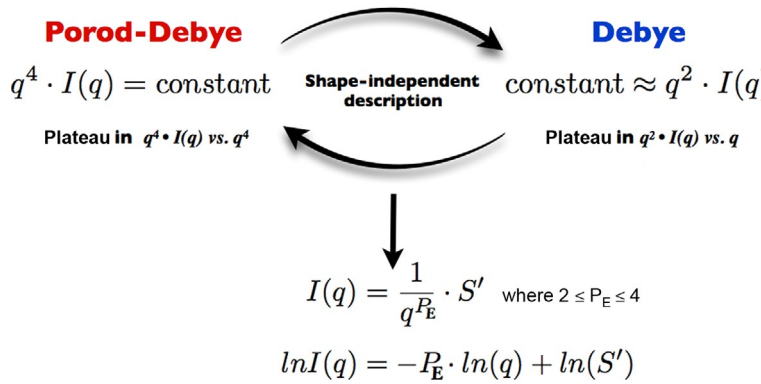


Figure 11.4 SAXS power-law relationships as indicators of RNA compactness. At low resolution, the SAXS profile of an RNA can be approximated by either its surface area (Porod–Debye approximation) or its R_g (Debye approximation). For the Porod–Debye approximation, the SAXS curve will decay as q^{-4}, whereas for the Debye approximation of a Gaussian chain, the SAXS curve will decay as q^{-2}. The exponential decay (Porod–Debye exponent, P_E) can be determined directly by fitting a line to the appropriate scattering range in a log–log plot. P_E will be bounded between 2 and 4 and can be determined using the program ScÅtter available at www.bioisis.net.

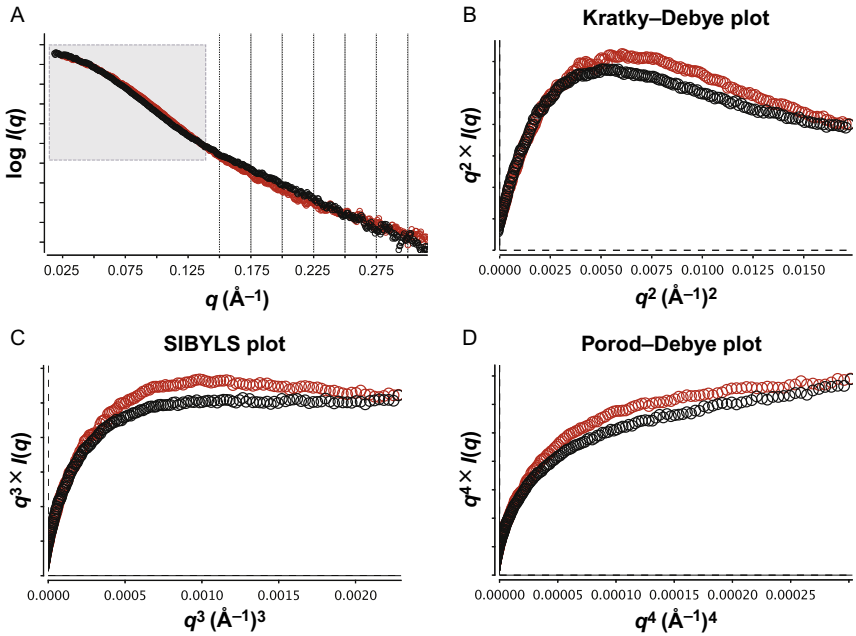


Figure 11.5 Flexibility plots for the SAM-I riboswitch in the presence (red) and absence (black) of ligand. (A) SAXS data plot as log $I(q)$ versus q. Boxed region defines limits for the plots in B–D. (B) Kratky–Debye plot ($q^2 \cdot I(q)$ vs. q^2). A plateau would indicate the data could be approximated using the Debye approximation for a Gaussian chain with a P_E near 2. (C) SIBYLS plot ($q^3 \cdot I(q)$ vs. q^3). A plateau would indicate a P_E near 3. (D) Porod–Debye plot ($q^4 \cdot I(q)$ vs. q^4). A plateau would indicate a P_E near 4 and that the data could be approximated using the Porod–Debye approximation for particles with a defined surface and contrast. For the SAM-I riboswitch, the (+) ligand state has a P_E near 4, whereas the (−) ligand state would be near 3. These plots clearly demonstrate a change in state of the riboswitch in the presence of ligand. Figures were prepared using the program ScÅtter available at www.bioisis.net.

Goodisman, & Brumberger, 1988) and the above discussion refers to the primary Porod–Debye region occurring immediately after the Guinier region.

Alternatively, a log–log transform of the SAXS data and subsequent fitting of the best line within the Porod–Debye region will determine the power-law Porod–Debye exponent, P_E. The exponent is bounded between 2 and 4 and we can expect for an RNA that undergoes a transition from a compact-folded state to a multiconformer, less compact state, P_E will decrease. The Porod–Debye exponent is a unique, quantitative descriptor of the thermodynamic ensemble that can be used for following gross changes in the folded state of the RNA.

4.2. SAXS invariants

The Porod–Debye exponent is an example of a model-independent parameter that can be derived directly from the SAXS data. Additional model-independent parameters (Glatter & Kratky, 1982; Svergun et al., 1987) are the Guinier-based R_g, volume (V_P), and correlation length (l_c). If the SAXS intensities are calibrated to an absolute scale (Orthaber, Bergmann, & Glatter, 2000), then the MM of the particle can be determined. The Guinier R_g is determined from a linear approximation of the low q data ($q \cdot R_g < 1.3$), whereas V_P and l_c are determined using the Porod invariant, Q. The Porod invariant is readily defined for compact particles but is undefined (Rambo & Tainer, 2013a) for particles behaving as Gaussian-like chains (flexible). Instead, we can examine the ratio of V_P to l_c, a term known as the volume-of-correlation (V_c) (Rambo & Tainer, 2013a). V_c is defined for both compact and flexible particles and can be used to monitor changes in a particle's thermodynamic state. Furthermore, the ratio of V_c and R_g^2 defines a power-law parameter (Rambo & Tainer, 2013a), Q_R, that determines the mass of a particle directly from a SAXS curve.

4.3. Real-space parameters

The $P(r)$-distribution is a resolution-limited histogram of all electron pair distances within the RNA. The shape of the distribution can be determined using indirect Fourier transform (IFT) methods where an orthogonal set of basis functions is used to parameterize the empirical SAXS dataset (Glatter & Kratky, 1982; Liu & Zwart, 2012; Moore, 1980; Svergun, 1992).

In practice, determining the $P(r)$-distribution by IFT is an iterative process that requires several trial d_{max} values to be evaluated by the experimenter. The goal is to find the smallest d_{max} that supports the following: (1) the $P(r)$ is greater than zero for $0 < r < d_{max}$, and (2) maximally smooth (minimize the sum of the second derivative $P(r)$ commensurate with the given SAXS resolution). In addition, the lowest and highest q-values defining the useable range of the SAXS dataset should be adjusted to test the robustness of the chosen d_{max} to the aforementioned criteria.

The first and second moments of $P(r)$-distribution define the average distance ($r_{average}$) and real space R_g. It should be noted, the Guinier R_g is a low-resolution approximation of the SAXS intensities that uses a small subset of the data ($q \cdot R_g < 1.3$), whereas the real space R_g is determined from all the available SAXS data but depends on the choice of d_{max}. The chosen d_{max} should produce real space values for R_g and $I(0)$ that are in close agreement

to the Guinier determined values. Difficult IFT solutions and large differences between real and reciprocal space values (>5%) can be indicative of poor data and sample quality (Putnam, Hammel, Hura, & Tainer, 2007) (e.g., aggregation or poor buffer subtraction).

Comparing SAXS data in real space as normalized $P(r)$-distributions is an excellent and recommended method for detecting conformational changes. It can be expected that small structural differences will not produce significant observable differences in $I(q)$ for $q < 2||{d_{\max}}^{-1}$ unless SAXS data are collected to sufficiently high resolution. In addition, for small changes where intensities can be corrupted by poor buffer subtractions and parasitic scattering near the beam stop, IFT methods can mitigate minor scattering issues thereby better demonstrating small differences in real space.

4.4. Dimensionless Kratky plot

The Kratky plot ($q^2 \cdot I(q)$ vs. q) is a standard qualitative method for asserting the compact (flexible) state of the RNA. The Kratky plot visualizes the Debye formalism for a Gaussian-like chain and should demonstrate a hyperbolic plateau within a limited q-range of the SAXS data (Rambo & Tainer, 2011; Receveur-Brechot & Durand, 2012). In the case of a compact particle, the Kratky plot will converge to baseline at higher q-values with a single maximum occurring after the Guinier region ($q \cdot R_g > 1.3$). Furthermore, the magnitude and location of the maximum (plateau) will vary depending on the concentration, MM and degree of compactness of the particle (Durand et al., 2010). This can make comparisons of the same RNA under different conditions problematic. The Kratky plot can be normalized by using $(q \cdot R_g)^2 \cdot I(q)/I(0)$ instead of $q^2 \cdot I(q)$ and plotting the data against $q \cdot R_g$ (Durand et al., 2010). Dividing by $I(0)$ normalizes the data for particle concentration and V_P, whereas plotting against $q \cdot R_g$ scales the q-axis to particle size. The plot is dimensionless and for globular particles, the plot will show a peak value at $q \cdot R_g = \sqrt{3}$ with a maximum of 1.104 (Fig. 11.6A) (Durand et al., 2010). In contrast, a Gaussian-like chain will have a peak maximum of 2 that is shifted right of $\sqrt{3}$. For nonglobular but compact particles, the peak will shift slightly right and upward. The use of a R_g-based dimensionless Kratky plot provides a semi-quantitative method for asserting changes in the compact state of the RNA by noting the location and height of the maximum.

Alternatively, a dimensionless Kratky plot can be made using V_c (Fig. 11.6B) by plotting $(q^2 \cdot V_c) \cdot I(q)/I(0)$ against $q^2 \cdot V_c$. Here, peak (plateau)

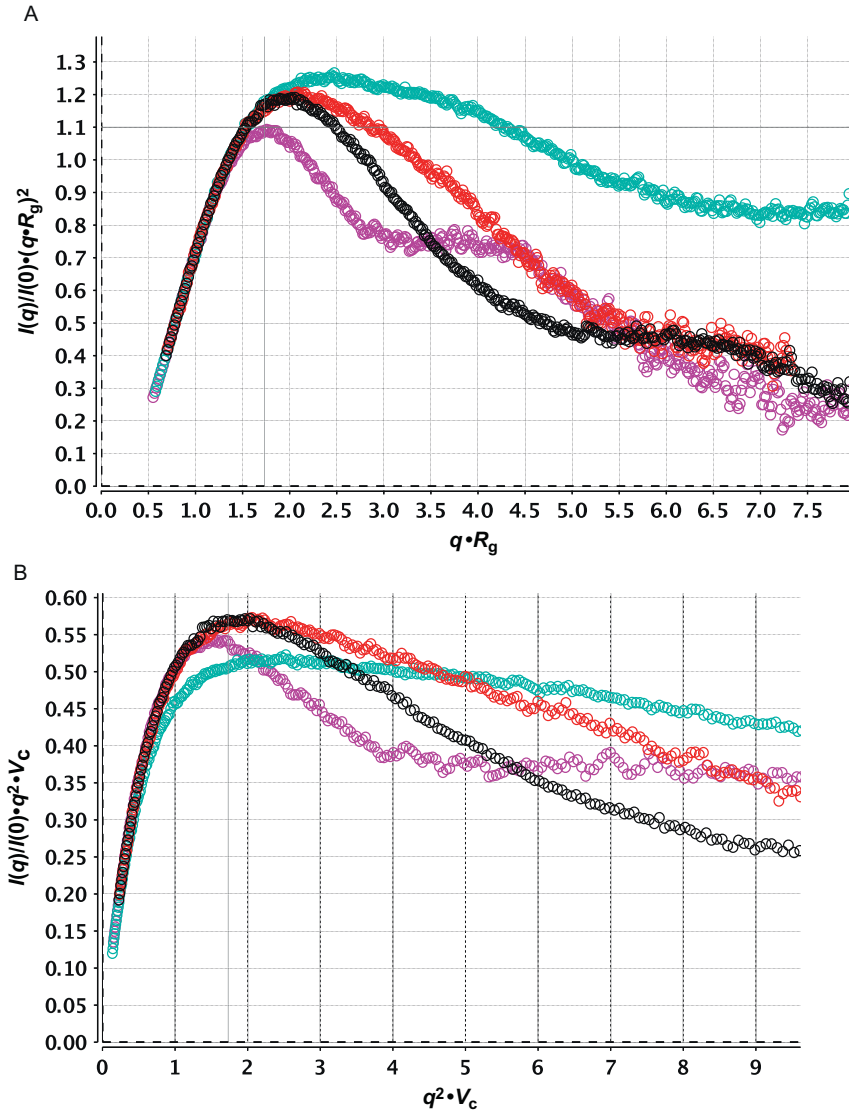


Figure 11.6 Dimensionless Kratky plots. P4–P6 domain (black), $tRNA^{phe} + Mg^{2+}$ (red), $tRNA^{phe} - Mg^{2+}$ (cyan), and TyMV UDP (purple). (A) R_g-based plot, cross hair (gray) marks the position of the peak for an ideal, compact, globular particle at $q \cdot R_g = \sqrt{3}$. Less compact or nonglobular particles will have a peak shifted upward and to the right. (B) V_c-based plot, vertical line represents $\sqrt{3}$. Peak height is inversely proportional to the particle's surface-to-volume ratio with the maximum occurring for a sphere at 0.85. Changes in peak height reflect changes to the particles surface to volume ratio. Similar peak heights for P4–P6 domain and $tRNA^{phe} + Mg^{2+}$ V_c-based plots suggest similar surface-to-volume ratios. Dimensionless plots will be sensitive to the accuracy of the Guinier region or real space transforms. Figures were prepared using the program ScÅtter available at www.bioisis.net.

height will be inversely proportional to the particle surface–volume ratio. The plot will have a maximum of 0.85 (perfect sphere) and can readily illustrate condition specific changes in the surface-to-volume ratio of an RNA (Fig. 11.6B). We suggest a comprehensive analysis using both dimensionless Kratky plots. In Fig. 11.6A, SAXS data for $tRNA^{phe}$ (76 nt), TyMV-UDP (109 nt), and P4-P6 RNA domain (160 nt) are plotted using both dimensionless Kratky plots. First, comparison of folded $tRNA^{phe}$ with P4–P6 RNA domain in the R_g-based plot show the tRNA peak is shifted to the right, suggesting tRNA is less globular than P4–P6. However, the peak heights are nearly coincident in the V_c-based plot, suggesting the two RNAs have similar surface-to-volume ratios. It has been firmly established that the P4–P6 domain from the Group I intron is a compact, folded domain in solution (Murphy & Cech, 1993), whereas tRNA are multiconformer displaying a wide range of rigid and nonrigid body motions (Wang & Jernigan, 2005). In a Mg^{2+}-free state, the $tRNA^{phe}$ peak height increases in the R_g-based plot, suggesting a less compact state. It can be expected that a less compact state will have a comparatively greater surface-to-volume ratio, and in the V_c-based Kratky plot for the Mg^{2+}-free $tRNA^{phe}$, we observe the expected decrease in peak height. Finally, TyMV-UDP RNA is considered a tRNA-like structure from Turnip yellow mosaic virus. The R_g-based plot shows the peak for TyMV-UDP occurs at $\sqrt{3}$ with a peak height $\sim$1.1 implying the RNA is globular in solution unlike $tRNA^{phe}$. However, the V_c-based plot shows TyMV-UDP has a greater surface-to-volume ratio than $tRNA^{phe}$. This can be reconciled by noting the P_E for TyMV-UDP is less than $tRNA^{phe}$ ($3 < 3.5$) suggesting TyMV-UDP exists as a globular multiconformer state in solution.

5. CASE STUDIES

The SAXS-derived parameters (R_g, V_c, V_P, P_E, d_{max}, and $r_{average}$) can be used to provide a comprehensive evaluation of the state of an RNA in solution. Furthermore, these parameters, including the dimensionless Kratky plots and $P(r)$-distribution, can be used to comparatively uncover mechanisms in folding or binding (e.g., conformational capture) by an RNA. Here, we will use reported SAXS data of the SAM-I (Stoddard et al., 2010) and LYS riboswitches (Garst et al., 2008) from the SAXS database Bioisis.net to examine two different riboswitch mechanisms.

5.1. SAM-I riboswitch

The SAM-I riboswitch is a 94 nt RNA aptamer that binds the small molecule *S*-adenosylmethionine (SAM). X-ray crystallographic studies of the riboswitch in the presence (+) and absence (−) of SAM revealed nearly identical riboswitch structures, suggesting that the unbound state of the RNA has a preformed binding site that is blocked by an internal adenosine. SAM (+) SAXS data showed an excellent agreement with the bound X-ray crystal structure. However, the apo-structure implies the solution state would be occupied by a single distinct conformation but paradoxically poses the problem of how substrate recognition could proceed when the binding pocket is blocked. Fitting the apo-structure to SAM (−) SAXS data immediately showed that the apo-structure from the crystal could not explain the solution state. The SAXS invariants R_g, V_c, and V_P (Table 11.4) all increased in the absence of SAM, suggesting that the riboswitch is becoming less compact. Does this apo solution state represent a conformational change to a single distinct species or a change in state of the riboswitch? Examining the Porod–Debye exponent, P_E, and apparent volume, V_P, shows a significant decrease in P_E (3.6–3.0) with a 36% increase in V_P. The drop in P_E is consistent with the riboswitch adopting a multiconformer, flexible state. The

Table 11.4 SAXS-derived parameters describing the solution (thermodynamic) state of RNA

	SAM(−)	SAM(+)	SAM(EDTA)	LYS(−)	LYS(+)	LYS(EDTA)
$R_g^{Guinier}$ (Å)	23.45	22.58	32.08	30.10	31.35	41.35
$R_g^{realspace}$ (Å)	24.77	22.59	34.60	30.72	30.27	43.86
$V_c^{Guinier}$ (Å^2)	295	277	354	399	423	454
$V_c^{realspace}$ (Å^2)	294	264	n/a	373	394	n/a
Volume[a] (Å^3)	75,900	56,300	162,000	105,000	106,000	269,000
P_E	3.0	3.6	2.1	3.4	3.4	1.9
d_{max} (Å)	79	74	118	101	99	163
$r_{average}$ (Å)	31.8	29.2	41	39.1	38.6	54.5

[a]Volume calculations are based on Guinier $I(0)$.
The set of parameters can be used comparatively to suggest changes in RNA compaction or flexibility. Compaction is readily identified by simultaneous decreases in R_g, V_c, d_{max}, $r_{average}$ and particle volume, whereas changes in the flexible state of the RNA can be implied by changes in the Porod–Debye exponent, P_E.

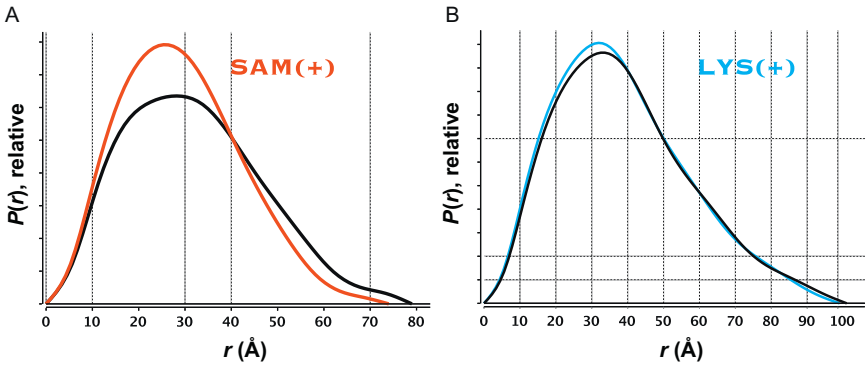


Figure 11.7 Normalized $P(r)$-distributions demonstrating conformational differences. Each distribution is scaled by $I(0)^{-1}$ and normalizes for RNA concentration and volume. (A) SAM-I riboswitch in the presence (red) and absence (black) of ligand. Large increase in peak height and narrowing of the distribution suggests significant compaction of the RNA in the presence of ligand. (B) LYS riboswitch in the presence (blue) and absence (black) of ligand. Overlay shows small but significant differences suggesting a small conformational change occurs upon binding.

multiconformer state will occupy a larger apparent volume suggesting the riboswitch goes from a collapsed single species to a multiconformer state. In fact, the SAXS data of the free state could not be modeled without using an ensemble of structures (Stoddard et al., 2010). The ensemble illustrated an opening of the SAM binding site akin to a conformational capture mechanism. Furthermore, the delocalization of the SAM-riboswitch in the free state is also apparent in the $P(r)$ distribution (Fig. 11.7A). Overlaying and normalizing the $P(r)$ distributions of the (+) and (−) states to their respective $I(0)$ values shows a large decrease in peak height and spreading of the SAM (−) $P(r)$ distribution.

5.2. LYS riboswitch

In contrast, the LYS riboswitch SAXS experiments detail a much different binding mechanism. The X-ray crystal structure of the (+) lysine state showed the ligand to be completely buried within the RNA. Furthermore, crystallization in the (−) lysine state illustrated a nearly identical arrangement of nucleotides comprising the binding pocket, suggesting a conformational change must be necessary to promote substrate recognition before encapsulation. Solution-state experiments using chemical probing showed the (−) and (+) lysine states to be nearly identical, thus arguing that the recognition and binding of the substrate occurs as a small perturbation to the RNA structure. As

mentioned above, the most reliable method for detecting small differences by SAXS will be through a direct comparison of the normalized $P(r)$ distributions (Fig. 11.7A). Here, comparison of the $P(r)$ distributions ($q_{max} = 0.31$ Å^{-1}) shows a small (~2 Å) decrease in d_{max} as the riboswitch transitions to the (+) state with slight but significant changes throughout the $P(r)$ distribution. Illustrated by the transition to the (+) lysine state, compaction of objects of the same mass and composition will show an increase in the $P(r)$ peak height with a corresponding decrease (1.5%) in R_g (Table 11.4). However, comparing the Guinier R_g values show an opposite effect, a 4% increase during the transition to the (+) state. The inconsistencies between real and reciprocal space R_g highlights the sensitivity and reliability of the Guinier region for detecting small differences. Nonetheless, we see no change in P_E between the two states (Table 11.4), suggesting the riboswitch maintains the same thermodynamic state of compactness and conclude that switching occurs through a small lysine stabilized conformational change.

6. MULTIPHASE VOLUMETRIC MODELING

Volumetric modeling of RNA SAXS data using DAMMIN/F (Svergun, 1999) provides a low-resolution three-dimensional shape of the thermodynamic state. Since this shape will be the average of the thermodynamic ensemble, the structural interpretation of the shape must be informed by the SAXS parameters mentioned above. Unrealistic particle volumes (V_P), low P_E values or significant differences between real and reciprocal space values suggest aggregation or a flexible unfolded state. Volumetric models in these cases of aggregation or flexible particles will contain artifacts and loose details in the averaged model. Regardless, volumetric models will be difficult to interpret at the secondary structure level without performing additional SAXS experiments using derivatives of the RNA. For the VS ribozyme (Lipfert, Ouellet, Norman, Doniach, & Lilley, 2008) and HIV REV response element (Fang et al., 2013), comparative DAMMIN/F modeling of various helical deletion or extension constructs inferred the location of secondary structure elements such that an all atom model of the RNA could be constructed. This inferred approach required significant structural features to exist in the base model to facilitate the alignment of the various constructs.

A more reliable method for locating secondary structure elements within SAXS volumetric models can be achieved with the program MONSA (Svergun & Nierhaus, 2000; Svergun, Petoukhov, & Koch, 2001). MONSA is a multiphase modeling algorithm that can integrate all the input SAXS

information during a single modeling run. Helical extensions or deletions are modeled as separate volume elements (phases) with their relative arrangements in three-dimensional space unambiguously determined notwithstanding mirror symmetry or severe changes to the base structure. Consider a SAXS experiment on a structured RNA with a 12-bp extension. MONSA would require two separate SAXS curves (wild-type and wild-type plus extension) with estimates of the volume elements specifying the wild-type RNA and the extension. The volumes can be determined from V_P with V_P (extension) taken as the difference from the two curves. As with DAMMIN/F, 15–20 independent modeling runs should be performed and averaged.

6.1. B_{12} riboswitch

To test MONSA in RNA modeling, we performed modeling on synthetic data created from the *E. coli* B_{12} riboswitch (PDB 4GMA). Here, a deletion of the P1,3 regulatory stem-loop was made (Fig. 11.8), and simulated SAXS data for the full-length and deletion constructs were used in modeling. In this example, we only specified the volume of the (−) P1,3 (117,000 $Å^3$) construct, leaving the volume for the P1,3 stem-loop constrained by the SAXS data. MONSA produced two volumetric models (Fig. 11.8E) in their relative arrangement clearly identifying the location of P1,3. Though the P1,3 volume was unconstrained, the averaged volume shows a density consistent with the P1,3 stem-loop (Fig. 11.8F). The P1,3 stem-loop is 27 nucleotides representing 12.8% of the riboswitch; in our experience, we have used MONSA to identify differences as small as 6 kDa in the volumetric models. In practice, the volume for each phase should be specified but multirounds of MONSA can be performed where one of the phases is left unconstrained. Leaving a phase unconstrained is a useful strategy to evaluate the quality of the samples as the final volume should be consistent with expectations. In addition, each MONSA run must be initiated with a different random number seed to insure independence for the final model averaging step by DAMAVER.

7. GOLD LABELS AND COMPREHENSIVE CONFORMATIONS

A major strength of SAXS is the ability to be high-throughput and therefore to provide comprehensive analysis of macromolecules under many biologically relevant conditions. Although not yet utilized for RNA, the sample preparations and analytical tools have been developed for DNA using nanogold labels (5-nm gold balls) (Hura, Tsai, et al., 2013). This technique

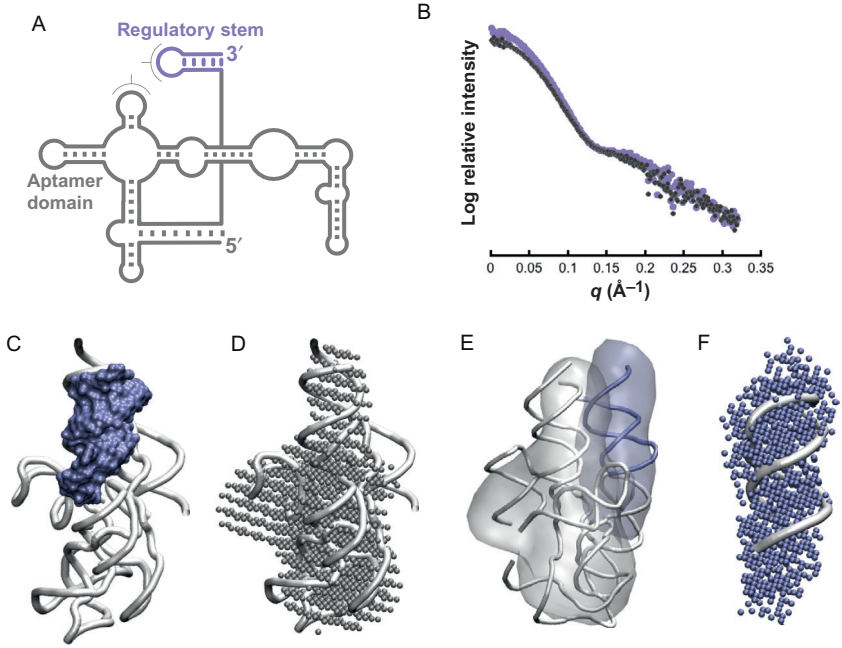


Figure 11.8 Secondary structure identification within volumetric model using MONSA. (A) Secondary structure of B_{12} riboswitch (Johnson et al., 2012). Full-length RNA (WT) is divided into two independent domains: (1) aptamer domain (APT, gray) and (2) regulatory P1,3 stem. (B) Simulated SAXS data with noise for WT (purple) and APT (gray). Data were simulated with CRYSOL (Svergun et al., 1995) using default parameters. Noise was transposed onto the simulated datasets from real SAXS data using a Gaussian sampling algorithm. (C) Atomistic model (PDB 4GMA) showing orientation of regulatory stem (purple) and APT (gray). (D) Averaged volumetric model (beads) of the APT volume calculated using MONSA from the SAXS data in B. (E) Averaged volumetric model of the APT domain (gray) and regulatory stem (purple). Relative orientation of the two volumes was determined by MONSA during the *ab initio* modeling using the SAXS data in B. (F) Alignment of the regulatory stem within the MONSA volume. In this example, only the APT volume (V_P^{APT}) was constrained during the MONSA modeling. V_P^{APT} was determined as the Porod volume using the program ScÅtter available at www.bioisis.net. In practice, the final model will be an average of 15–20 MONSA runs with the volume of each phase (APT domain and P13 stem) constrained during the MONSA modeling. All atomistic models were superimposed using SUPCOMB (Kozin & Svergun, 2001).

can examine short or long sequences of nucleic acids in most solution conditions, including those closest to cellular environments. Gold atoms have an electron (e^-) density greater than protein and RNA ($4.6e^-/Å^3$ for gold nanocrystals vs. 0.43 and $0.55e^-/Å^3$ for protein or nucleic acid, respectively). In water, the measured scattering of a gold nanocrystal with 50 Å

diameter was 5400-fold higher than that of 31-bp dsDNA (Hura, Tsai, et al., 2013). Thus, measured scattering in mixed systems containing gold nanocrystals, nucleic acids, and proteins is dominated by the gold signal. Selective labeling methods can be used to identity specific secondary structure elements or follow small conformational changes. In practice, the gold is in the form of a nanocrystal coated with a neutral PEG ligand. Experiments with the protein ATPase MutS and dsDNA labeled at both ends provided end-to-end distance measurements under a variety of salt concentrations (0–200 m*M* NaCl). The nanogold labels did not precipitate protein, are stable, and may be useful in designing experiments to understand RNA and RNA–protein mechanisms.

Other developments relevant to detecting conformation changes, are the SAXS structural comparison map (SCM) and volatility of ratio (V_R) difference metric (Hura, Budworth, et al., 2013), which together provide quantitative and superposition-independent evaluation of solution-state conformations. SCMs are diagonally symmetric matrices that contain the set of pairwise agreements from a collection of SAXS curves. SCMs can be color mapped to illustrate the degree of similarity among multiple conformations. V_R, which is visually displayed in an SCM, is given by the ratio between two experimental SAXS profiles. This ratio is normalized such that the average across a specified resolution range is 1 and V_R is determined from a binned partition of the ratio. For $d_{max} < 400$ Å and $q < 0.2$ Å^{-1}, we suggest using 25 bins.

SAXS experiments on human MSH2-MSH3 ATPase in complex with mismatch or normal DNA in the presence of various nucleotide ligands resulted in a large number of curves to analyze. V_R calculations and creation of an SCM efficiently revealed multiple unique conformational states showing how the different nucleotide ligands drive the sculpting of the DNA conformation (Hura, Budworth, et al., 2013). Furthermore, V_R-based SCM approaches were successful at resolving various discrete translation states of the ribosome (Fig. 11.9). It will be interesting to see V_R applied to riboswitches and to RNA in general where high-throughput SAXS experiments that explore condition space together with SCMs can help uncover novel functional states.

8. CONSIDERATIONS

Macromolecular crystallography provides unmatched precision in the analysis of specific RNA riboswitch structural states. However, the methods and results presented here show that flexibility, disorder, conformational

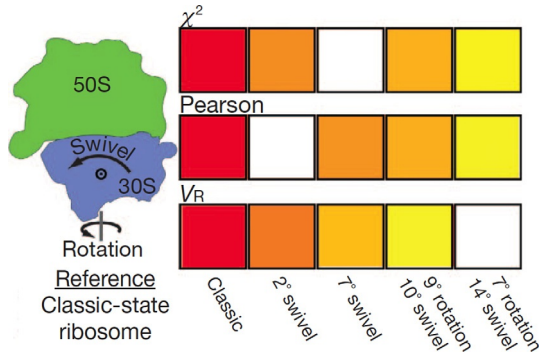


Figure 11.9 SAXS applied to measure structural similarity. χ^2, Pearson correlation coefficient and V_R scores were assigned a gradient color from red (high similarity) to white (low similarity). Comparison of ribosomal structures solved with the 30S subunit rotated about two axes (annotated by rotation and swivel degree) relative to the 50S subunit. Left to right: Protein Data Bank reference structures 3R8T:4GD2, 3I21:3I22, 3I1O:3I1P, 3R8S:4GD1, and 3UOQ:3UOS. *Adapted from Hura, Budworth, et al. (2013).*

variation, and other aspects of dynamic states, defined by biological SAXS experiments, are critical to an accurate description of RNA biology, activity, and mechanisms. Fortunately, SAXS data interpretation tools and technologies are evolving at a rapid pace due to the involvement of a growing group of developers and users (Gajda et al., 2013; Kofinger & Hummer, 2013; Moore, 2014; Pollack, 2011; Rambo & Tainer, 2013a; Schneidman-Duhovny, Hammel, & Sali, 2011; Schneidman-Duhovny et al., 2013; Yang et al., 2010). A critical component for these developments has been the accessibility of SAXS experimental data for both for structural validation and for software development.

As we outlined here, SAXS is a multidimensional structural technique that can both inform on sample quality and mechanisms of activity and therefore should be considered in any structural analyses of RNA. Besides complementing X-ray crystallography, electron microscopy, and NMR, SAXS provides fundamental information about macromolecular interfaces and their solvent interactions. This information can be used to drive biological research into areas previously unavailable and to suggest improvements in the analysis of data from other methods, such as NMR and crystallography. New and updated SAXS beamlines and laboratory equipment are expanding the user community and the types of biological problems amendable to SAXS (Perez & Nishino, 2012). The growth of SAXS is stimulating experts to develop improved experimental and computational methods.

Yet, despite many seminal SAXS advances, the information in the comprehensive set of electron pair distances that is specified by the $P(r)$-distribution remains greatly underappreciated and underutilized. Therefore, there exists great scope for continued advances of SAXS over the next decade. Ultimately, a more complete description of the thermodynamic ensemble from SAXS experiments may help reduce the increasing complexity of macromolecular structural biology by contributing to general principles of interaction, conformation, and mechanism.

ACKNOWLEDGMENTS

We are grateful to Michal Hammel, Greg Hura, and Robert T. Batey for insightful discussions. This work is supported in part by funding to foster collaboration with Bruker and Lawrence Berkeley National Laboratory on Novel Technology for Structural Biology. The SIBYLS beamline (BL12.3.1) facility and team at the ALS is supported by United States Department of Energy program Integrated Diffraction Analysis Technologies DEAC02-05CH11231 and by the National Institute of Health Grant R01GM105404.

REFERENCES

Baird, N. J., & Ferre-D'Amare, A. R. (2010). Idiosyncratically tuned switching behavior of riboswitch aptamer domains revealed by comparative small-angle X-ray scattering analysis. *RNA*, *16*, 598–609.

Batey, R. T., & Kieft, J. S. (2007). Improved native affinity purification of RNA. *RNA*, *13*, 1384–1389.

Beckert, B., & Masquida, B. (2011). Synthesis of RNA by in vitro transcription. *Methods in Molecular Biology*, *703*, 29–41.

Chacon, P., Diaz, J. F., Moran, F., & Andreu, J. M. (2000). Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. *Journal of Molecular Biology*, *299*, 1289–1302.

Chen, B., Zuo, X., Wang, Y. X., & Dayie, T. K. (2012). Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy. *Nucleic Acids Research*, *40*, 3117–3130.

Ciccariello, S., Goodisman, J., & Brumberger, H. (1988). On the Porod law. *Journal of Applied Crystallography*, *21*, 117–128.

Classen, S., Hura, G. L., Holton, J. M., Rambo, R. P., Rodic, I., McGuire, P. J., et al. (2013). Implementation and performance of SIBYLS: A dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the Advanced Light Source. *Journal of Applied Crystallography*, *46*, 1–13.

Costantino, D. A., Pfingsten, J. S., Rambo, R. P., & Kieft, J. S. (2008). tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nature Structural & Molecular Biology*, *15*, 57–64.

Daldrop, P., Reyes, F. E., Robinson, D. A., Hammond, C. M., Lilley, D. M., Batey, R. T., et al. (2011). Novel ligands for a purine riboswitch discovered by RNA-ligand docking. *Chemistry & Biology*, *18*, 324–335.

Debye, P., Anderson, H. R., & Brumberger, H. (1957). Scattering by an inhomogeneous solid. II. The correlation function and its application. *Journal of Applied Physics*, *28*, 5.

Durand, D., Vives, C., Cannella, D., Perez, J., Pebay-Peyroula, E., Vachette, P., et al. (2010). NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *Journal of Structural Biology*, *169*, 45–53.

Fang, X., Wang, J., O'Carroll, I. P., Mitchell, M., Zuo, X., Wang, Y., et al. (2013). An unusual topological structure of the HIV-1 Rev response element. *Cell*, *155*, 594–605.
Frieda, K. L., & Block, S. M. (2012). Direct observation of cotranscriptional folding in an adenine riboswitch. *Science*, *338*, 397–400.
Gajda, M. J., Martinez Zapien, D., Uchikawa, E., & Dock-Bregeon, A. C. (2013). Modeling the structure of RNA molecules with small-angle X-ray scattering data. *PloS One*, *8*, e78007.
Garst, A. D., & Batey, R. T. (2009). A switch in time: Detailing the life of a riboswitch. *Biochimica et Biophysica Acta*, *1789*, 584–591.
Garst, A. D., Heroux, A., Rambo, R. P., & Batey, R. T. (2008). Crystal structure of the lysine riboswitch regulatory mRNA element. *The Journal of Biological Chemistry*, *283*, 22347–22351.
Gilbert, S. D., & Batey, R. T. (2009). Monitoring RNA-ligand interactions using isothermal titration calorimetry. *Methods in Molecular Biology*, *540*, 97–114.
Glatter, O., & Kratky, O. (1982). *Small angle X-ray scattering*. London; New York: Academic Press.
Gopal, A., Zhou, Z. H., Knobler, C. M., & Gelbart, W. M. (2012). Visualizing large RNA molecules in solution. *RNA*, *18*, 284–299.
Grishaev, A., Guo, L., Irving, T., & Bax, A. (2010). Improved fitting of solution X-ray scattering data to macromolecular structures and structural ensembles by explicit water modeling. *Journal of the American Chemical Society*, *132*, 15484–15486.
Hammond, J. A., Rambo, R. P., Filbin, M. E., & Kieft, J. S. (2009). Comparison and functional implications of the 3D architectures of viral tRNA-like structures. *RNA*, *15*, 294–307.
Hammond, J. A., Rambo, R. P., & Kieft, J. S. (2010). Multi-domain packing in the aminoacylatable 3′ end of a plant viral RNA. *Journal of Molecular Biology*, *399*, 450–463.
Hura, G. L., Budworth, H., Dyer, K. N., Rambo, R. P., Hammel, M., McMurray, C. T., et al. (2013). Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. *Nature Methods*, *10*, 453–454.
Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., II, Tsutakawa, S. E., et al. (2009). Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nature Methods*, *6*, 606–612.
Hura, G. L., Tsai, C. L., Claridge, S. A., Mendillo, M. L., Smith, J. M., Williams, G. J., et al. (2013). DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 17308–17313.
John, D. M., Merino, E. J., & Weeks, K. M. (2004). Mechanics of DNA flexibility visualized by selective 2′-amine acylation at nucleotide bulges. *Journal of Molecular Biology*, *337*, 611–619.
Johnson, J. E., Jr., Reyes, F. E., Polaski, J. T., & Batey, R. T. (2012). B12 cofactors directly stabilize an mRNA regulatory switch. *Nature*, *492*, 133–137.
Kazantsev, A. V., Rambo, R. P., Karimpour, S., Santalucia, J., Jr., Tainer, J. A., & Pace, N. R. (2011). Solution structure of RNase P RNA. *RNA*, *17*, 1159–1171.
Kofinger, J., & Hummer, G. (2013). Atomic-resolution structural information from scattering experiments on macromolecules in solution. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, *87*, 052712.
Kozin, M. B., & Svergun, D. I. (2001). Automated matching of high- and low-resolution structural models. *Journal of Applied Crystallography*, *34*, 33–41.
Lipfert, J., Ouellet, J., Norman, D. G., Doniach, S., & Lilley, D. M. (2008). The complete VS ribozyme in solution studied by small-angle X-ray scattering. *Structure*, *16*, 1357–1367.
Liu, H., Morris, R. J., Hexemer, A., Grandison, S., & Zwart, P. H. (2012). Computation of small-angle scattering profiles with three-dimensional Zernike polynomials. *Acta Crystallographica. Section A*, *68*, 278–285.

Liu, H., & Zwart, P. H. (2012). Determining pair distance distribution function from SAXS data using parametric functionals. *Journal of Structural Biology*, *180*, 226–234.
Moore, P. (1980). Small-angle scattering. Information content and error analysis. *Journal of Applied Crystallography*, *13*, 168–175.
Moore, P. B. (2014). The effects of thermal disorder on the solution-scattering profiles of macromolecules. *Biophysical Journal*, *106*, 1489–1496.
Murphy, F. L., & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the Tetrahymena ribozyme. *Biochemistry*, *32*, 5291–5300.
Orthaber, D., Bergmann, A., & Glatter, O. (2000). SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *Journal of Applied Crystallography*, *33*, 218–225.
Perdrizet, G. A., II, Artsimovitch, I., Furman, R., Sosnick, T. R., & Pan, T. (2012). Transcriptional pausing coordinates folding of the aptamer domain and the expression platform of a riboswitch. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 3323–3328.
Perez, J., & Nishino, Y. (2012). Advances in X-ray scattering: From solution SAXS to achievements with coherent beams. *Current Opinion in Structural Biology*, *22*, 670–678.
Perry, J. J., & Tainer, J. A. (2013). Developing advanced X-ray scattering methods combined with crystallography and computation. *Methods*, *59*, 363–371.
Peselis, A., & Serganov, A. (2014). Themes and variations in riboswitch structure and function. *Biochimica et Biophysica Acta*, *1839*, 908–918.
Petoukhov, M. V., & Svergun, D. I. (2013). Applications of small-angle X-ray scattering to biomacromolecular solutions. *The International Journal of Biochemistry & Cell Biology*, *45*, 429–437.
Poitevin, F., Orland, H., Doniach, S., Koehl, P., & Delarue, M. (2011). AquaSAXS: A web server for computation and fitting of SAXS profiles with non-uniformally hydrated atomic models. *Nucleic Acids Research*, *39*, W184–W189.
Pollack, L. (2011). Time resolved SAXS and RNA folding. *Biopolymers*, *95*, 543–549.
Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics*, *40*, 191–285.
Rambo, R. P., & Doudna, J. A. (2004). Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry*, *43*, 6486–6497.
Rambo, R. P., & Tainer, J. A. (2010a). Bridging the solution divide: Comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Current Opinion in Structural Biology*, *20*, 128–137.
Rambo, R. P., & Tainer, J. A. (2010b). Improving small-angle X-ray scattering data for structural analyses of the RNA world. *RNA*, *16*, 638–646.
Rambo, R. P., & Tainer, J. A. (2011). Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, *95*, 559–571.
Rambo, R. P., & Tainer, J. A. (2013a). Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*, *496*, 477–481.
Rambo, R. P., & Tainer, J. A. (2013b). Super-resolution in solution X-ray scattering and its applications to structural systems biology. *Annual Review of Biophysics*, *42*, 415–441.
Receveur-Brechot, V., & Durand, D. (2012). How random are intrinsically disordered proteins? A small angle scattering perspective. *Current Protein & Peptide Science*, *13*, 55–75.
Reining, A., Nozinovic, S., Schlepckow, K., Buhr, F., Furtig, B., & Schwalbe, H. (2013). Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature*, *499*, 355–359.

Ribitsch, G., De Clercq, R., Folkhard, W., Zipper, P., Schurz, J., & Clauwaert, J. (1985). Small-angle X-ray and light scattering studies on the influence of Mg^{2+} ions on the structure of the RNA from bacteriophage MS2. *Zeitschrift für Naturforschung. Section C*, *40*, 234–241.

Schneidman-Duhovny, D., Hammel, M., & Sali, A. (2011). Macromolecular docking restrained by a small angle X-ray scattering profile. *Journal of Structural Biology*, *173*, 461–471.

Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., & Sali, A. (2013). Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophysical Journal*, *105*, 962–974.

Serganov, A., Huang, L., & Patel, D. J. (2008). Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature*, *455*, 1263–1267.

Sibille, N., & Bernado, P. (2012). Structural characterization of intrinsically disordered proteins by the combined use of NMR and SAXS. *Biochemical Society Transactions*, *40*, 955–962.

Stoddard, C. D., Montange, R. K., Hennelly, S. P., Rambo, R. P., Sanbonmatsu, K. Y., & Batey, R. T. (2010). Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure*, *18*, 787–797.

Strulson, C. A., Yennawar, N. H., Rambo, R. P., & Bevilacqua, P. C. (2013). Molecular crowding favors reactivity of a human ribozyme under physiological ionic conditions. *Biochemistry*, *52*, 8187–8197.

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, *25*, 495–503.

Svergun, D. I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical Journal*, *76*, 2879–2886.

Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*, *28*, 768–773.

Svergun, D. I., Feĭgin, L. A., & Taylor, G. W. (1987). *Structure analysis by small-angle X-ray and neutron scattering*. New York: Plenum Press.

Svergun, D. I., & Nierhaus, K. H. (2000). A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. *The Journal of Biological Chemistry*, *275*, 14432–14439.

Svergun, D. I., Petoukhov, M. V., & Koch, M. H. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal*, *80*, 2946–2953.

Trausch, J. J., Xu, Z., Edwards, A. L., Reyes, F. E., Ross, P. E., Knight, R., et al. (2014). Structural basis for diversity in the SAM clan of riboswitches. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 6624–6629.

Vicens, Q., Mondragon, E., & Batey, R. T. (2011). Molecular sensing by the aptamer domain of the FMN riboswitch: A general model for ligand binding by conformational selection. *Nucleic Acids Research*, *39*, 8586–8598.

Wang, Y., & Jernigan, R. L. (2005). Comparison of tRNA motions in the free and ribosomal bound structures. *Biophysical Journal*, *89*, 3399–3409.

Woodson, S. A., & Koculi, E. (2009). Analysis of RNA folding by native polyacrylamide gel electrophoresis. *Methods in Enzymology*, *469*, 189–208.

Wyatt, P. (1993). Light scattering and the absolute characterization of macromolecules. *Analytica Chimica Acta*, *272*, 1–40.

Yang, S., Parisien, M., Major, F., & Roux, B. (2010). RNA structure determination using SAXS data. *The Journal of Physical Chemistry. B*, *114*, 10039–10048.

Zhang, J., Jones, C. P., & Ferre-D'Amare, A. R. (2014). Global analysis of riboswitches by small-angle X-ray scattering and calorimetry. *Biochimica et Biophysica Acta*, *1839*, 1020–1029.