

Compaction of a Bacterial Group I Ribozyme Coincides with the Assembly of Core Helices[†]

Ursula A. Perez-Salas,^{*,#,\\$} Prashanth Rangan,^{#,§} Susan Krueger,[‡] R. M. Briber,^{*,||} D. Thirumalai,[⊥] and Sarah A. Woodson^{*,§}

Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8562, T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218-2685, and Department of Materials Science and Engineering and Institute for Physical Sciences and Technology, University of Maryland, College Park, Maryland 20472

Received September 11, 2003; Revised Manuscript Received December 5, 2003

ABSTRACT: Counterions are critical to the self-assembly of RNA tertiary structure because they neutralize the large electrostatic forces which oppose the folding process. Changes in the size and shape of the *Azoarcus* group I ribozyme as a function of Mg²⁺ and Na⁺ concentration were followed by small angle neutron scattering. In low salt buffer, the RNA was expanded, with an average radius of gyration (R_g) of $53 \pm 1 \text{ \AA}$. A highly cooperative transition to a compact form ($R_g = 31.5 \pm 0.5 \text{ \AA}$) was observed between 1.6 and 1.7 mM MgCl₂. The collapse transition, which is unusually sharp in Mg²⁺, has the characteristics of a first-order phase transition. Partial digestion with ribonuclease T1 under identical conditions showed that this transition correlated with the assembly of double helices in the ribozyme core. Fivefold higher Mg²⁺ concentrations were required for self-splicing, indicating that compaction occurs before native tertiary interactions are fully stabilized. No further decrease in R_g was observed between 1.7 and 20 mM MgCl₂, indicating that the intermediates have the same dimensions as the native ribozyme, within the uncertainty of the data ($\pm 1 \text{ \AA}$). A more gradual transition to a final R_g of approximately 33.5 \AA was observed between 0.45 and 2 M NaCl. This confirms the expectation that monovalent ions not only are less efficient in charge neutralization but also contract the RNA less efficiently than multivalent ions.

Like proteins, certain RNA molecules assemble into unique three-dimensional structures that are essential for their biological activity. How these folded structures form from the denatured state has become the subject of intense investigation (e.g., refs 1–3). In contrast to proteins, where hydrophobic interactions drive the collapse of the polypeptide chain, RNA folding requires cations to neutralize the electrostatic repulsion between phosphate groups. The collapse of RNA chains to intermediate (non-native) structures in the presence of counterions is of fundamental importance because it determines the probability of forming biologically active structures in a short time.

Theoretical and experimental studies of DNA and RNA show that counterion condensation around nucleic acids

reduces the effective phosphate charge by 75–90% (ref 4 and references cited therein). Theoretical models of polyelectrolytes suggest that counterion condensation initially produces an ensemble of compact forms that contain both native and non-native interactions that slowly diffuse to the native state (2).

Experimentally, the presence of collapsed intermediates in RNA folding was indicated by biochemical (5, 6) and structural (7) studies on group I ribozymes. Small angle X-ray scattering (SAXS)¹ experiments on the *Tetrahymena* group I ribozyme (7), yeast tRNA^{Phe}, and RNase P (8) demonstrated that counterions induce compact structures at concentrations below what is required to stabilize the native structure. In accord with theoretical predictions (2), UV absorption (9) and stopped-flow SAXS (10) showed that the initial collapse occurs in 1–10 ms, which is a much shorter time than required to form the native ribozyme.

An important question is the extent to which native interactions stabilize these compact folding intermediates. We addressed this question using small angle neutron scattering (SANS) to measure changes in the global dimensions of a 195 nt group I ribozyme from pre-tRNA^{Ile} of the *Azoarcus* bacterium (11) (Figure 1). Like SAXS, SANS is a

[†] Supported by the National Research Council Postdoctoral Associateship Program (U.A.P.-S.), the National Institute of Standards and Technology, and the National Institutes of Health (S.A.W.). The Center for High-Resolution Neutron Scattering at the NIST is sponsored by the NSF (DMR-9986442).

* Corresponding authors: U.A.P.-S. (tel, 301-975-8395; fax, 301-921-9847; e-mail, ursula.perez-salas@nist.gov), R.M.B. (tel, 301-405-7313; fax, 301-314-9601; e-mail, rbriber@umd.edu), and S.A.W. (tel, 410-516-2015; fax, 410-516-4118; e-mail, swoodson@jhu.edu).

The first two authors contributed equally to this work.

‡ Center for Neutron Research, NIST.

§ T. C. Jenkins Department of Biophysics, Johns Hopkins University.

|| Department of Materials Science and Engineering, University of Maryland.

⊥ Institute for Physical Sciences and Technology, University of Maryland.

¹ Abbreviations: 3D, three dimensional; Hepes, *N*-(2-hydroxyethyl)piperazine-*N'*-2-ethanesulfonic acid; R_g , radius of gyration; RNase, ribonuclease; SANS, small angle neutron scattering; SAXS, small-angle X-ray scattering; SLD, scattering length density.

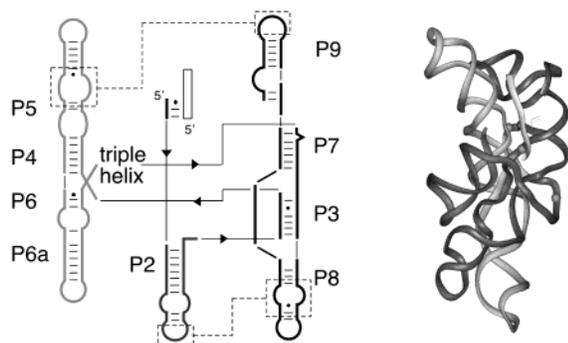


FIGURE 1: Structure of the *Azoarcus* group I ribozyme. The secondary (22, 42) and tertiary structures (14) were modeled from comparative sequence analysis. Base-paired (P) regions in the ribozyme are indicated.

powerful technique for measuring changes in the size and shape of noninteracting particles in solution, with a resolution on the order of angstroms (12, 13). Neutron scattering can also provide information about solvation or dynamics.

The collapse transition detected by neutron scattering was compared with two conformational phase transitions previously defined by biochemical probes of RNA structure (14): a transition from unfolded (U) RNA to a more ordered intermediate (I_C) in low Mg^{2+} concentration that involves the assembly of helices in the core of the ribozyme and a second transition from I_C to native tertiary structure (N) in higher Mg^{2+} concentrations that coincides with the appearance of catalytic activity. The reduction in R_g correlates well with the transition from U to I_C as assayed by RNase T1 protection and precedes the appearance of catalytic activity. The results suggest that the ensemble of collapsed structures contains many native interactions, implying that the *Azoarcus* ribozyme collapses to native-like structures upon counterion condensation.

MATERIALS AND METHODS²

RNA Preparation. The L-9 *Azoarcus* ribozyme (195 nt) was transcribed in vitro from pAz-IVS digested with *EarI* (14). Transcription reactions (10 mL) were carried out in disposable 50 mL centrifuge tubes as previously described (15), with 5 μ g/mL plasmid DNA, 50000 units of T7 RNA polymerase, and 2 mM each nucleotide triphosphate. Reactions were incubated 6 h at 37 °C. Reactions were pooled and concentrated to 3 mL with Centriprep-30 concentrators (Millipore) before purification from a denaturing 4% polyacrylamide gel. RNA was extracted overnight in 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 250 mM NaCl, and the eluate was concentrated and exchanged several times with deionized water as above. The RNA was resuspended at 10 mg/mL, as estimated from absorption at 260 nm. Greater than 95% of the RNA was full-length, as judged from analytical 4% PAGE and staining with ethidium bromide. For SANS, RNA was diluted to 2 mg/mL in 25 mM Na-Hepes, pH 7.0, and 0–20 mM $MgCl_2$ or 0–2 M NaCl and

then incubated for 10 min at 50 °C before placing it in the beam. This is sufficient time for folding to reach equilibrium (14).

Partial RNase T1 Digestion. 5'-³²P-labeled ribozyme was incubated with 2 mg/mL unlabeled RNA in 25 mM Na-Hepes, pH 7.0, plus 0–5 mM $MgCl_2$ for 10 min at 50 °C. RNase T1 (0.1 unit) was added to each sample, and the resultant solution was incubated for 1 min at 50 °C. The reactions were stopped with equal volumes of formamide loading dye. The products were separated on a 6% sequencing gel and quantified as previously described (14). Data were normalized to the minimum and maximum extent of cleavage for each residue and fit to the Hill equation, $\bar{Y} = [Mg^{2+}]^n / ([Mg^{2+}]^n + C_m^n)$, where \bar{Y} is the fractional saturation, C_m is the midpoint of the transition, and n is the Hill constant.

Self-Splicing Assays. Splicing reactions were carried out as previously described (14, 16), except that ³²P-labeled pre-tRNA was mixed with an additional 2 mg/mL unlabeled *Azoarcus* ribozyme in 25 mM Na-Hepes, pH 7.0, plus 0–30 mM $MgCl_2$ to mimic SANS conditions. Samples were equilibrated for 10 min at 50 °C and 2 min at 32 °C before the splicing was initiated with 100 μ M GTP. Spliced products were analyzed as previously described (14), and progress curves were fit to first-order rate equations.

SANS Measurements. SANS experiments were carried out on the 195 nt *Azoarcus* ribozyme in aqueous (100% H₂O) 25 mM Na-Hepes, pH 7.0, supplemented with NaCl or $MgCl_2$. Measurements were performed using the NG3 30 m SANS instrument at the Center for High-Resolution Neutron Scattering at the National Institute of Standards and Technology (NIST) in Gaithersburg, MD (17). RNA samples prepared as above were placed in quartz cuvettes (0.3 mL) with a 1 mm path length. Sample holders were maintained at 32 °C throughout the experiment.

In small angle scattering, the scattering from the sample is measured over the angular range from a fraction of a degree to a few degrees, which corresponds to size scales of 10–500 Å for 6 Å wavelength neutrons. This scattering is due to the interaction of the neutrons with the variations in neutron scattering length density (SLD) in the sample. For a sample consisting of particles in a solvent, these variations are characterized by the particles' relative SLD function, $\rho(\vec{r})$, resulting from differences in the atomic and isotopic composition and density of the particles relative to the solvent. The scattered intensity is measured as a function of the momentum transfer, \vec{Q} , and is related to the scattering angle 2θ by $|\vec{Q}| = Q = (4\pi/\lambda) \sin \theta$, where λ is the neutron wavelength.

The scattered intensity was obtained by subtracting from the raw scattering data the background scattering due to the buffer and cuvette and the incoherent scattering from hydrogen atoms. The data sets were placed on an absolute scale by normalizing the scattered intensity to the incident beam flux. Data were radially averaged to produce the scattering intensity $I(Q)$ as a function of Q . The wavelength λ was fixed at 6 Å with $\Delta\lambda/\lambda = 26.4\%$ over the entire range of momentum transfer measured. The spatial range obtained from scattering measurements is given to first order by $\langle r \rangle = \pi/Q$. In our experiments Q ranged from 0.01 to 0.15 Å⁻¹, which corresponds to a spatial range of 314 Å < $\langle r \rangle$ < 20.9 Å. More details of the SANS instrument can be found in ref 17.

² Certain commercial materials, instruments, and equipment are identified in this paper in order to specify the experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Data Analysis. The distance distribution function $P(r)$ for the *Azoarcus* ribozyme in solution was estimated from the scattering data using an inversion algorithm, GNOM, developed by Semenyuk and Svergun (<http://www.emblham-burg.de/ExternalInfo/Research/Sax>) (12, 18). The $P(r)$ was calculated using

$$P(r) \cong \frac{1}{2\pi^2} \int_{Q_{\min}}^{Q_{\max}} I(Q) Q r \sin(Qr) dQ \quad (1)$$

and normalized to $I(0)$ using $\int_0^\infty P(r) dr = I(0)(4\pi)^{-1}$. By definition, $P(r)^3$ is related to the density–density correlation function as follows:

$$P(r) = r^2 \int_{\Omega} \int_V \rho(\vec{u}) \rho(\vec{u} + \vec{r}) d\vec{u} d\Omega \quad (2)$$

where Ω is a solid angle, \vec{u} is any vector in the volume V defining the space occupied by a particle, and $\rho(\vec{r})$ is the relative SLD function of the particles. Because of its direct relation to $\rho(\vec{r})$, $P(r)$ can be used to calculate the size and shape of the particles. The radius of gyration was obtained from

$$R_g^2 = \int_{D_{\min}}^{D_{\max}} P(r) r^2 dr / 2 \int_{D_{\min}}^{D_{\max}} P(r) dr \quad (3)$$

where $P(r) = 0$ for $r < D_{\min}$ and $r > D_{\max}$, D_{\min} is the smallest distance within the particle (generally zero), and D_{\max} is the largest dimension within the particle. GNOM calculates $P(r)$ from small angle scattering data using eq 1. For a given D_{\min} and D_{\max} the calculated $I(Q)$, $I_{\text{calc}}(Q)$, is

$$I_{\text{calc}}(Q) = \int_{D_{\min}}^{D_{\max}} \frac{1}{r^2} P(r) e^{-i\vec{Q}\vec{r}} d\vec{r} \quad (4)$$

Theoretical Models. A theoretical $P(r)$ distribution was computed from the 3D atomic model of the *Azoarcus* ribozyme (14) using the program CRYSON (19) by D. Svergun (<http://www.embl-hamburg.de/ExternalInfo/Research/Sax/cryson.html>). The $P(r)$ curves were compared to standard polymer models of single chains. Treating the unfolded RNA as a random coil, we obtained $P(r)$ by an inverse Fourier transform of the Debye scattering function. Letting $x = r/2R_g$, $P(r)$ for a random coil is

$$\frac{P(r)}{I(0)} = \frac{2}{\sqrt{\pi}R_g} [\sqrt{\pi}x(x^2 + 1/4)\text{erfc}(x) - x^2 e^{-x^2}] \quad (5)$$

where

$$\text{erfc}(x) = 1 - \frac{2}{\pi} \int_0^x e^{-t^2} dt \quad (6)$$

Asymptotically ($x \gg 1$), $P(r) \approx [1/(2\sqrt{\pi}R_g)]e^{-x^2}$. The native RNA can be considered maximally compact and was approximated by a sphere of radius R . For a sphere with $R_g^2 = 3R^2/5$, $P(r)$ is

$$\frac{P(r)}{I(0)} = \frac{3r^2}{5R^3} \left[1 - \frac{3r}{4R} + \frac{r^3}{16R^3} \right] \quad (7)$$

Both these models characterize $P(r)$ in terms of a single length, R_g .

RESULTS

Change in RNA Conformation by SANS. To monitor conformational transitions in the RNA, SANS experiments were carried out on 2 mg/mL ribozyme solution over a range of MgCl_2 and NaCl concentrations. The scattered intensity measured on samples containing 1, 2, and 3 mg/mL RNA in 10 mM MgCl_2 scaled with concentration and showed no evidence of particle–particle interactions. In Hepes buffer alone, some aggregation was apparent in 3 mg/mL RNA. All further SANS experiments were performed on 2 mg/mL solutions of the ribozyme to minimize interactions between molecules.

Figure 2a shows the scattering intensity data as the Mg^{2+} concentration is increased from 0 to 20 mM. The scattering curves fall into two distinct classes: One set of curves, corresponding to samples where the concentrations of Mg^{2+} varied from 0 to 1.6 mM, has scattered intensity which decreases rapidly as a function of Q , indicating particles with a relatively large R_g . The second set, corresponding to samples with Mg^{2+} concentrations from 1.7 to 20 mM, represents structures that are more compact. In these samples, the scattered intensity decreases more slowly as a function of Q . The data are similar within each group, and the transition between the two types of scattering behavior occurs abruptly between 1.6 and 1.7 mM MgCl_2 .

Similarly, the scattering curves for the ribozyme in increasing concentration of NaCl are split into two types with a transition at 450 mM NaCl , as shown in Figure 2b. The transition, however, is reached more gradually as the salt concentration varies between 0 and 450 mM. The higher concentrations of Na^+ required to condense the RNA are consistent with less efficient charge neutralization by the monovalent counterions and the ability of multivalent metal ions to bridge two phosphates (ref 6 and references cited therein).

Metal Ion Induced Collapse. Inverting the scattering data to the real space density correlation function $P(r)$ has been useful for the analysis of macromolecules because it gives a model-independent measure of the size and shape of the molecules (13). The scattering curves in Figure 2 were indirectly inverted to obtain the real space correlation function $P(r)$ (Figure 3). It is evident from Figure 3 that the two classes of $P(r)$ functions relate to two distinct particle shapes: an extended shape at low cation concentrations and a significantly more compact state at high cation concentrations. The variation in the maximum extension of the particles, D_{\max} , for the extended shape or the compact shape shows that the precision of D_{\max} is within 7% (Table 1). The R_g computed from $P(r)$ decreases from an average of 53 ± 1 Å below 1.6 mM MgCl_2 to 31.5 ± 0.5 Å above 1.7 mM MgCl_2 and 33.4 ± 0.2 Å in 2 M NaCl (Table 1).

To evaluate the unfolded state in low counterion concentrations, we compared the distribution functions $P(r)$ to a standard Gaussian chain model (random coil) with an equivalent R_g (see Materials and Methods). The $P(r)$ func-

³ $P(r)$ is the chord (secant) distribution in an object of uniform density representing the particle $\rho(\vec{r})$, as obtained by eq 1, which is then averaged over all angles following eq 2. Equivalently, $P(r)$ is the distribution of line segments formed by the joining of two points on the surface of the object representing the particle.

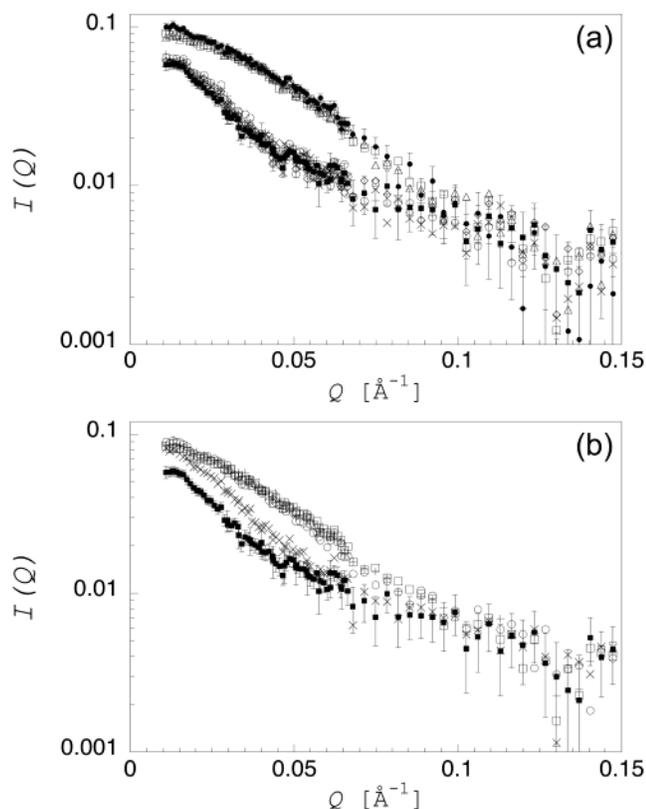


FIGURE 2: SANS data for the *Azoarcus* ribozyme. (a) Mg^{2+} concentration dependence: 25 mM Na-Hepes, pH 7.0 at 32 °C (■), plus 1 mM Mg^{2+} (◇), 1.3 mM Mg^{2+} (×), 1.5 mM Mg^{2+} (○), 1.6 mM Mg^{2+} (+), 1.7 mM Mg^{2+} (□), 4 mM Mg^{2+} (△), and 20 mM Mg^{2+} (●). (b) Na^+ concentration dependence: 25 mM Na-Hepes, pH 7.0 (■), plus 100 mM Na^+ (×), 450 mM Na^+ (○), 750 mM Na^+ (+), and 2 M Na^+ (□).

tions show that even with an equivalent R_g , the mass in the RNA in 25 mM Na-Hepes is distributed over shorter distances than predicted by the random coil model (Figure 4). The experimentally obtained value of $P(r)$ is greater over distances of 90–130 Å than expected for a random coil. This suggests that, on intermediate scales, RNA is more rigid than a Gaussian chain. This local stiffness is presumably due to double helical segments in the unfolded RNA.

We also compared the experimental $I(Q)$ and $P(r)$ curves for the 20 mM Mg^{2+} sample, where the ribozyme is in its native conformation, with the predicted scattering curve $I_{3D}(Q)$ and corresponding $P_{3D}(r)$ computed from a 3D model of the ribozyme (14) (Figure 1) and $P_{\text{sphere}}(r)$ for a sphere (see Materials and Methods). As shown in the insert of Figure 4, the SANS curve predicted by the 3D model is similar to the data but yields a smaller R_g of 30 Å. The difference between the most probable values of r of the 3D model and the experimental curve could be due either to conformational fluctuations in the native state or to errors in the model, which is based on comparative sequence analysis (14). As expected, $P(r)$ obtained from the scattering data more closely resembles $P_{3D}(r)$ than $P_{\text{sphere}}(r)$, especially for $r < R_g$.

Comparison of Folding in Mg^{2+} and Na^+ . Differences in the size of the compact states formed in Mg^{2+} and Na^+ are small, with a deviation in R_g at the largest salt concentrations of about 2 Å (Figure 3 and Table 1). A slightly less compact shape is attained in 2 M NaCl ($R_g = 33.4 \pm 0.2$ Å) than in 20 mM MgCl_2 ($R_g = 31.5 \pm 0.6$ Å). The difference in R_g

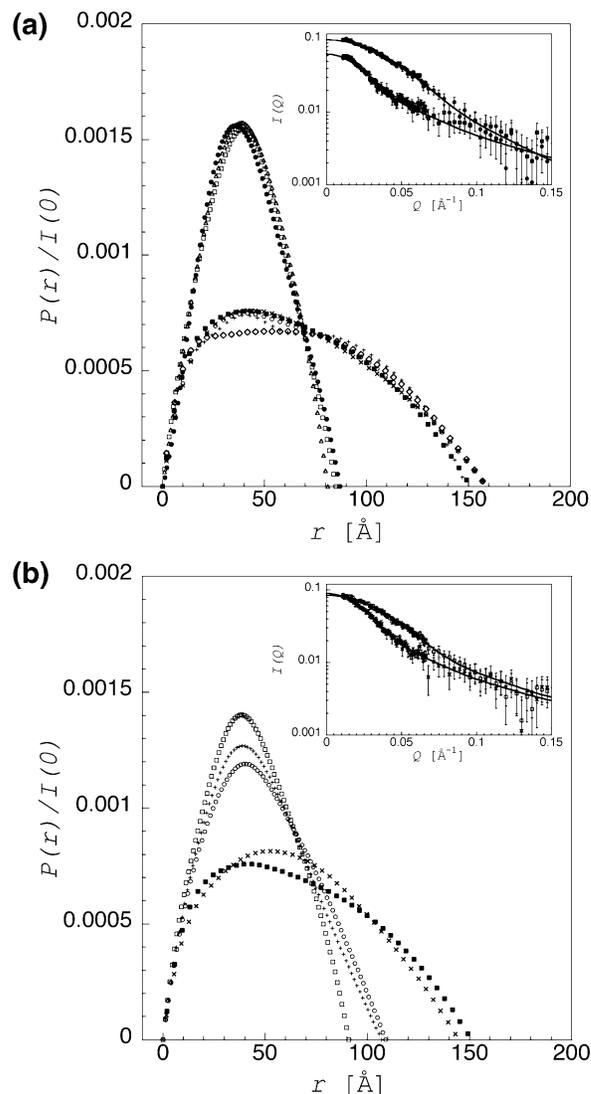


FIGURE 3: Distance distribution functions. $P(r)$ distributions were fit to the SANS data in Figure 2 according to eq 1 and scaled by $I_{\text{calc}}(0)^{-1}$. For clarity, error bars are not shown. Symbols as in Figure 2. (a) Mg^{2+} titration. (b) Na^+ titration. Insets: Scattering curves computed from $P(r)$ distribution functions are compared with neutron scattering data from samples containing 0 and 20 mM MgCl_2 or 0 and 2 M NaCl, respectively.

may be smaller, as the Na^+ transition may not be fully saturated even at 2 M NaCl. The similarity of R_g values obtained in Mg^{2+} and Na^+ is consistent with biochemical results showing that the *Azoarcus* ribozyme forms many tertiary interactions in monovalent salts, lacking only a few within the active site (20). Because many different metal ions support the formation of I_C (20), it seems unlikely that the formation of compact intermediates requires site-specific coordination of metal ions. Instead, collapse is driven by nonspecific electrostatic interactions. Previous biochemical data showed that native-like intermediates of the *Tetrahymena* ribozyme are also stabilized by a variety of counterions (6, 21).

Collapse Correlates with Helix Assembly. An important question is whether the metal ion induced decrease in the radius of gyration (R_g) of the ribozyme correlates with the formation of native secondary or tertiary interactions. We previously observed two macroscopic conformational transitions in the *Azoarcus* ribozyme with increasing Mg^{2+}

Table 1: Summary of SANS Results^a

salt concn	D_{\max} (Å)	R_g (Å) ^b
buffer only	151 ± 2^c	52.3
1 mM MgCl ₂	159	55.3
1.3 mM MgCl ₂	159	54.1
1.5 mM MgCl ₂	159	54.3
1.6 mM MgCl ₂	149	53.9
1.7 mM MgCl ₂	85	31.6
4 mM MgCl ₂	81	31.1
20 mM MgCl ₂	87	31.9
100 mM NaCl	144 ± 1^c	50.4
450 mM NaCl	109	38.1
750 mM NaCl	107	36.6
2 M NaCl	91	33.4

^a All samples contain 2 mg/mL ribozyme in 25 mM Na-Hepes, pH 7.0 at 32 °C. Parameters are derived from fits to $P(r)$ functions as described in the text. ^b For samples containing Mg²⁺, the uncertainty in R_g values due to random error can be estimated from the standard deviations of the mean values for the unfolded and compact conformations, which were 53 ± 1 and 31.5 ± 0.5 Å, respectively. Statistical errors of individual fits computed with GNOM ranged from 0.2 to 0.7 Å. ^c Error reflects variation in D_{\max} among equivalent “best” fits for these samples. Data for other samples yielded a single best fit.

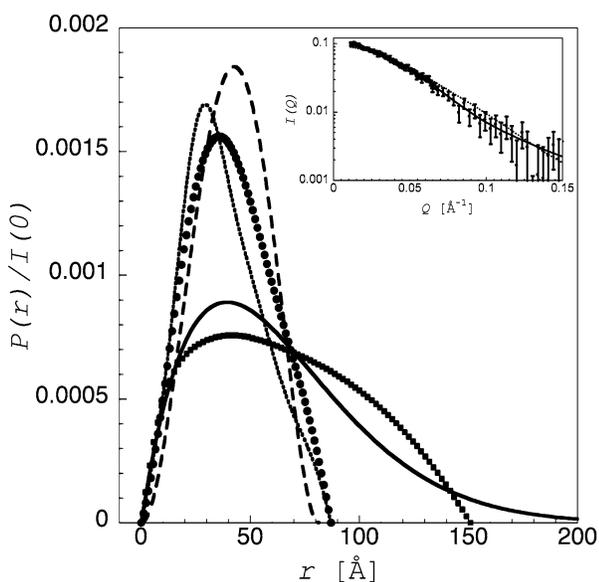


FIGURE 4: Comparison of distance distribution functions of unfolded and native RNA with models. $P(r)$'s obtained from SANS data for 25 mM Na-Hepes, pH 7.0 at 32 °C: (■) with no added salt; (●) plus 20 mM Mg²⁺. Curves represent the following: (—) $P_{\text{random coil}}(r)$ for a random coil, $R_g = 53$ Å; (---) $P_{\text{sphere}}(r)$ for a uniform sphere, $R_g = 31.5$ Å; (- · -) $P_{3D}(r)$ for the 3D atomic model, $R_g = 30$ Å. Inset: SANS data for RNA in 25 mM Na-Hepes, pH 7.0, plus 20 mM Mg²⁺. The continuous curves correspond to $I_{\text{calc}}(Q)$ computed from the experimental $P(r)$ (—) and $I_{3D}(Q)$ computed from the 3D model (14) using CRYSON (19).

concentration (14). In low ionic strength buffer (25 mM Na-Hepes, pH 7.0), the P2, P4, P5, and P6a stem-loops are formed, but the RNA is otherwise unfolded (U). In the first transition, moderate concentrations of monovalent or divalent salt stabilize an intermediate (I_C) containing native double helices in the core of the ribozyme. These interactions include the P3/P7 pseudoknot and a triple helix that mediates interactions between the P4–P6 and P3–P9 domains (Figure 1). In the second transition, higher Mg²⁺ concentrations stabilize the native tertiary structure (N), which coincides with the appearance of catalytic activity (14). Similar transitions were observed in Na⁺ and other monovalent and

divalent counterions, except that 1 M NaCl produced a folded but inactive structure (I_F) (20).

To determine which of these transitions (U to I_C or I_C to N) correlates with compaction of the RNA, the secondary structure of the ribozyme was probed by partial digestion with ribonuclease (RNase) T1 under the conditions of the SANS experiments (see Supporting Information). Addition of Mg²⁺ resulted in the protection of G's throughout the core of the ribozyme, with midpoints for individual residues ranging from 1.2 ± 0.4 to 1.9 ± 0.1 mM MgCl₂ under these conditions (Figure 5a). These values are higher than reported previously (14), because the high RNA concentrations used for SANS (6 mM in nucleotides) require higher numbers of counterions. The midpoints of base-pairing transitions monitored by RNase T1 were close to the MgCl₂ concentration at which we observed collapse of the RNA (1.7 mM). Protection of core nucleotides from RNase T1 in Na⁺ also coincided with the decrease in R_g measured by SANS (Figure 5a).

By contrast, higher Mg²⁺ concentrations were required for self-splicing activity under these conditions (Figure 5b and Supporting Information), which reflects the formation of native tertiary structure (14). The midpoint of the transition to the native structure was 4.5 mM Mg²⁺, with maximal activity above 20 mM Mg²⁺. Thus, global collapse of the RNA chain correlates with the assembly of core helices during the transition to I_C and occurs at lower Mg²⁺ concentrations than the formation of native tertiary structure (N) as reported by self-splicing activity (Figure 5b).

DISCUSSION

Small angle neutron scattering measurements reveal a highly cooperative transition to a compact intermediate for the 195 nt *Azoarcus* group I ribozyme in MgCl₂. Parallel RNase T1 assays under identical conditions showed that this transition correlates with the assembly of the ribozyme's core helices. Catalytic activity is not observed until the Mg²⁺ concentration is increased 3–5-fold above the concentration where the initial collapse occurs. Yet, even after a 5-fold increase in Mg²⁺ concentration, no further compaction of the ribozyme is observed by SANS (Table 1), indicating that changes in the size and shape of the ribozyme during the transition from the intermediates to the native state are smaller than the 1 Å error of these measurements. Together, these experiments provide strong support for the idea that counterions induce the collapse of polynucleotide chains but that collapse alone is not sufficient to produce the native state.

Long-range interactions in the catalytic center of group I ribozymes, such as the P3/P7 pseudoknot and the central triple helix, bring together nucleotides that are distant in the primary sequence (22). Consequently, assembly of helices in the core of the ribozyme was expected to result in compaction of the RNA (14), consistent with our observation that reduction in R_g coincides with the U to I_C transition. That the assembly of helices in the ribozyme core occurs at low Mg²⁺ concentrations and precedes tertiary folding is also suggested by recent experiments on group I introns from yeast mitochondria and *Candida albicans* (23, 24).

For the *Azoarcus* ribozyme, as for the *Tetrahymena* group I ribozyme (388 nt), the catalytic domain of *Bacillus subtilis*

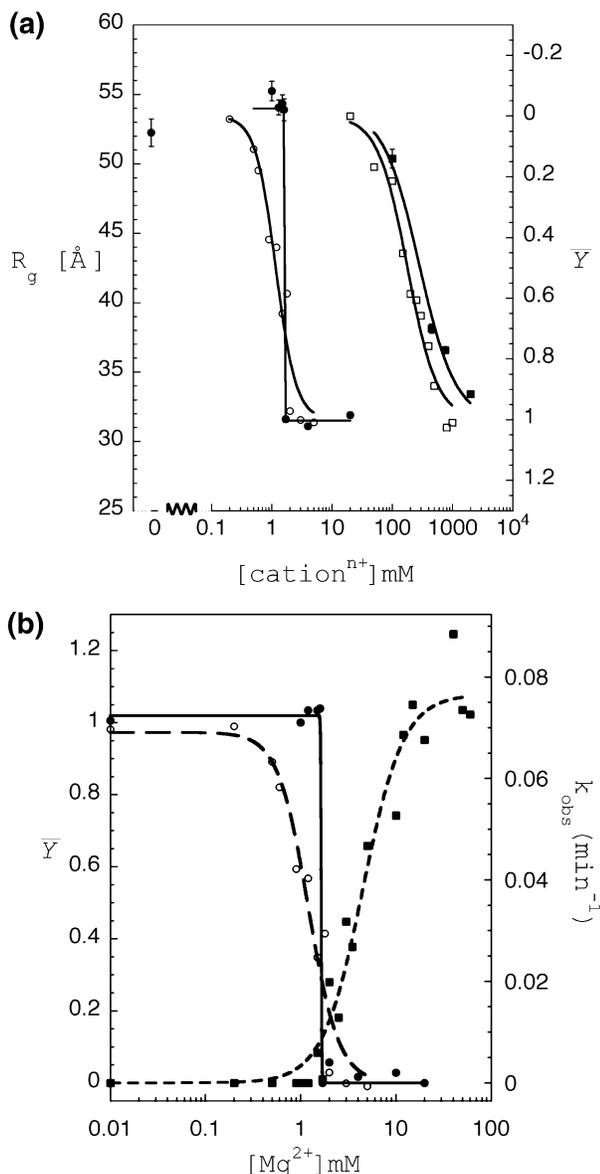


FIGURE 5: Comparison of R_g with ribozyme folding. (a) The change in R_g measured by SANS (filled symbols) is compared with the fractional saturation \bar{Y} for RNase T1 protection (open symbols) in MgCl_2 (circles) or NaCl (squares). Data were fit to the Hill model: (○) $C_m = 1.2 \pm 0.1$ mM, $n = 2.7 \pm 0.6$; (●) $C_m \approx 1.7 \pm 0.1$ mM, $n > 100$; (□) $C_m = 180 \pm 11$ mM, $n = 1.8 \pm 0.2$; (■) $C_m = 282 \pm 22$ mM, $n = 1.5 \pm 0.1$. RNase T1 data shown are for G121, G122, and G124. Similar results were obtained for other residues (see Supporting Information). (b) Comparison of folding and activity in MgCl_2 : R_g (●); RNase T1 (○); rate of self-splicing (■). Data were fit to the Hill equation as in A: $C_m = 4.5 \pm 0.5$; $n = 2.0 \pm 0.3$. R_g was normalized to the maximum and minimum values obtained. Biochemical assays were conducted in the same conditions as SANS to permit direct comparison of the results.

RNase P (250 nt), and yeast tRNA^{Phe} (76 nt), the major collapse transition occurs in the early stages of the salt titration (7, 8), during which the effective charge on the nucleic acid is most sensitive to changes in metal ion concentration. Unlike the *Azoarcus* ribozyme, the *Tetrahymena* ribozyme and RNase P C domain form intermediates that are 5–15% less compact than the native structure. More expanded intermediates could reflect the increased presence of non-native interactions or greater structural dynamics of these folding intermediates, compared with the intermediates

formed by the *Azoarcus* ribozyme. Indeed, recent stopped-flow SAXS studies provide evidence of an early nonspecific collapse followed by more specific compaction in the *Tetrahymena* ribozyme (25). As we argue below, these differences may reflect differences in the collapse mechanisms of these RNAs.

A First-Order Mg^{2+} -Induced Collapse Transition? The transition to the collapsed state, especially in Mg^{2+} , is unexpectedly sharp (Figure 5). It is generally assumed that the collapse transition should be continuous, with R_g changing smoothly as the counterion concentration is increased. By contrast, the U to I_C transition in Mg^{2+} has the characteristics of an “all-or-none” transition. Because this transition is induced by both Mg^{2+} and Na^+ , it is likely to reflect nonspecific condensation of counterions around the nucleic acid rather than specific metal ion interactions. The transition in Mg^{2+} may appear particularly steep because of the large numbers of Mg^{2+} ions needed to satisfy the phosphates at the high RNA concentrations used for SANS (6 mM nucleotides). However, a sharp (but less abrupt) collapse transition is also observed at lower concentrations of the *Azoarcus* ribozyme (G. Caliskan, U. Perez-Salas, P. Rangan, D. Thirumalai, R. M. Briber, and S. A. Woodson, unpublished results) and the *Tetrahymena* ribozyme (6).

Biochemical data show that the discontinuous change in R_g is accompanied by an increase in intrachain interactions and a decrease in free energy over the U to I_C transition. Together, these observations suggest that the collapse transition in Mg^{2+} is first order. If true, this implies that U and I_C are separated by a free energy barrier. A first-order U to I_C transition would be consistent with theoretical predictions of the generic coil to globule transition in polyelectrolytes (26). Further experiments will be required to test this idea.

Compact Intermediates in RNA Folding. The similarity in size between the intermediate and the native forms of the *Azoarcus* ribozyme suggests that the I_C intermediate state is structured, even though native tertiary contacts are barely detectable under these conditions by hydroxyl radical footprinting (14, 20) or by self-splicing activity (Figure 5). This suggests that I_C represents an ensemble of compact and native-like states. An alternative hypothesis is that I_C is not a distinct structural ensemble but simply represents a rapid equilibrium between U and N. This explanation is less satisfactory, because the 10% of native RNA in 1.7 mM MgCl_2 deduced from self-splicing assays is too low to account for the native-like $P(r)$ distribution of this sample.

In folded RNAs, interactions between double helices are mediated by “bridging” counterions (27), hydrogen-bonding networks (28), and van der Waals interactions along the minor groove (29). The extent to which counterions and hydrophobic interactions also stabilize nonspecific helix packing is not known. Recent results from single molecule studies of the hairpin ribozyme suggest that Mg^{2+} ions associate with compact forms of the ribozyme before the RNA molecules cross the major folding transition state (30). The presence of compact but non-native intermediates was also inferred from the folding kinetics of the P4–P6 domain of the *Tetrahymena* ribozyme (31).

Counterion-Mediated Collapse. As in DNA condensation (32), multivalent cations drive the compaction of RNA chains more efficiently than monovalent cations (6). In our experiments, a 100-fold lower concentration of Mg^{2+} than Na^+ is

required to induce compaction of the ribozyme. Because electrostatic repulsion of the phosphates is a major force opposing RNA folding, the R_g of an approximately spherical folded RNA is expected to correlate with the residual net charge of the polynucleotide after counterion condensation, $N\bar{\nu} \approx -\ln \phi(R_g/Zl_B)$, where N is the number of residues, $\bar{\nu}$ is the residual charge per phosphate, ϕ is the volume fraction of the counterions, l_B is the Bjerrum length, and Z is the valence of the counterion (6). Taking the value of $l_B = 7.1$ Å in water, we find that $\bar{\nu} \approx 0.1$ near the midpoints of the collapse transitions (1.7 mM $MgCl_2$ and 280 mM NaCl). Hence, the RNA folds in the presence of counterions when approximately 90% of the phosphate charge is neutralized, which is qualitatively consistent with previous work (refs 33 and 34 and references cited therein). The net charge per phosphate is roughly similar whether the RNA is folded in monovalent or divalent salts. By contrast, counterion condensation theory predicts $\bar{\nu} = 0.24$ for monovalent cations and 0.12 for divalent cations associated with double-stranded DNA (35, 36), which cannot form such tightly packed structures.

Because of nearly complete neutralization of the backbone charges due to nonspecific counterion condensation, the native structures can become maximally compact, so that $R_g^o \approx l_p N^{1/3}$, where R_g^o is the size of the native state, N is the number of nucleotides, and l_p is an effective persistence length. Using experimental R_g^o values, we obtain $l_p \approx 5.4$ – 5.8 Å for the native state, depending on the counterion. Using SAXS data on the *Tetrahymena* ribozyme (7), we previously estimated $l_p \approx 7$ Å.

CONCLUSION

Theoretical arguments (2) and the present and previous (7, 8) scattering experiments have shown that the transition to the native state must be preceded by counterion-mediated collapse of the RNA chain. A key unresolved question is the nature of the collapsed states. The I_C state, which we take to be an ensemble of compact structures, could be produced by either specific collapse to native-like structures, nonspecific collapse to an ensemble of native and non-native structures, or both. Structures resulting from specific collapse have large numbers of fluctuating native contacts, whereas those arising from nonspecific collapse are stabilized by many non-native contacts.

The well-studied *Tetrahymena* ribozyme produces a collection of non-native metastable intermediates that require several minutes or longer to reach the native state (37–40). For this RNA, only a small fraction of the folding population folds within 1 s. The prevalence of misfolded structures is consistent with nonspecific counterion-mediated collapse.

By contrast, collapse of the *Azoarcus* ribozyme coincides with the formation of native helices in the core of the ribozyme. That the unusually sharp collapse transition of the *Azoarcus* ribozyme produces an I_C state with near-native R_g values suggests that Mg^{2+} favors specific collapse, in agreement with biochemical assays (14). Because the search for the native state occurs among compact native-like structures (41), it should require less time. The observation that a large fraction of the *Azoarcus* ribozyme population folds in ~ 100 ms also lends support to the assertion that I_C represents an ensemble of specifically nucleated “native-like”

structures. Additional experiments will be needed to further characterize the nature of collapsed structures in this and related RNAs.

ACKNOWLEDGMENT

We thank Robert Moss for purification of T7 RNA polymerase and for help with RNA preparation.

SUPPORTING INFORMATION AVAILABLE

Two figures showing ribonuclease T1 digestion of ribozyme RNA and self-splicing assays and one table listing the Mg^{2+} dependence of RNase T1 digestion of 10 mg/mL ribozyme. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Brion, P., and Westhof, E. (1997) Hierarchy and dynamics of RNA folding, *Annu. Rev. Biophys. Biomol. Struct.* 26, 113–137.
- Thirumalai, D., Lee, N., Woodson, S. A., and Klimov, D. (2001) Early events in RNA folding, *Annu. Rev. Phys. Chem.* 52, 751–762.
- Sosnick, T. R., and Pan, T. (2003) RNA folding: models and perspectives, *Curr. Opin. Struct. Biol.* 13, 309–316.
- Misra, V. K., and Draper, D. E. (1998) On the role of magnesium ions in RNA stability, *Biopolymers* 48, 113–135.
- Buchmueller, K. L., Webb, A. E., Richardson, D. A., and Weeks, K. M. (2000) A collapsed, non-native RNA folding state, *Nat. Struct. Biol.* 7, 362–366.
- Heilman-Miller, S. L., Thirumalai, D., and Woodson, S. A. (2001) Role of counterion condensation in folding of the *Tetrahymena* ribozyme. I. Equilibrium stabilization by cations, *J. Mol. Biol.* 306, 1157–1166.
- Russell, R., Millett, I. S., Doniach, S., and Herschlag, D. (2000) Small angle X-ray scattering reveals a compact intermediate in RNA folding, *Nat. Struct. Biol.* 7, 367–370.
- Fang, X., Littrell, K., Yang, X., Henderson, S. J., Siefert, S., Thiagarajan, P., Pan, T., and Sosnick, T. R. (2000) Mg^{2+} -dependent compaction and folding of yeast tRNA(Phe) and the catalytic domain of the *B. subtilis* RNase P RNA determined by small-angle X-ray scattering, *Biochemistry* 39, 11107–11113.
- Pan, T., and Sosnick, T. R. (1997) Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity, *Nat. Struct. Biol.* 4, 931–938.
- 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) Rapid compaction during RNA folding, *Proc. Natl. Acad. Sci. U.S.A.* 99, 4266–4271.
- Reinhold-Hurek, B., and Shub, D. A. (1992) Self-splicing introns in tRNA genes of widely divergent bacteria, *Nature* 357, 173–176.
- Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small Angle X-ray and Neutron Scattering*, Plenum Press, New York.
- Glatter, O., and Kratky, O. (1982) *Small Angle X-ray Scattering*, Academic Press, San Diego, CA.
- Rangan, P., Masquida, B., Westhof, E., and Woodson, S. A. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme, *Proc. Natl. Acad. Sci. U.S.A.* 100, 1574–1579.
- Davanloo, P., Rosenberg, A. H., Dunn, J. J., and Studier, F. W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.* 81, 2035–2039.
- Tanner, M., and Cech, T. (1996) Activity and thermostability of the small self-splicing group I intron in the pre-tRNA(Ile) of the purple bacterium *Azoarcus*, *RNA* 2, 74–83.
- Glinka, C. J., Barker, J. G., Hammouda, B., Krueger, S., Moyer, J. J., and Orts, W. J. (1998) The 30 m small-angle neutron scattering instruments at the National Institute of Standards and Technology, *J. Appl. Crystallogr.* 31, 430–445.
- Semenyuk, A. V., and Svergun, D. I. (1991) GNOM—a Program Package for Small-Angle Scattering Data- Processing, *J. Appl. Crystallogr.* 24, 537–540.

19. Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) Protein hydration in solution: experimental observation by X-ray and neutron scattering, *Proc. Natl. Acad. Sci. U.S.A.* 95, 2267–2272.
20. Rangan, P., and Woodson, S. A. (2003) Structural requirement for Mg²⁺ binding in the group I intron core, *J. Mol. Biol.* 329, 229–238.
21. Uchida, T., Takamoto, K., He, Q., Chance, M. R., and Brenowitz, M. (2003) Multiple monovalent ion-dependent pathways for the folding of the L-21 *Tetrahymena thermophila* ribozyme, *J. Mol. Biol.* 328, 463–478.
22. Michel, F., and Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis, *J. Mol. Biol.* 216, 585–610.
23. Chamberlin, S. I., and Weeks, K. M. (2003) Differential helix stabilities and sites pre-organized for tertiary interactions revealed by monitoring local nucleotide flexibility in the bI5 group I intron RNA, *Biochemistry* 42, 901–909; Buchmueller, K. L., and Weeks, K. M. (2003) Near native structure in an RNA collapsed state, *Biochemistry* 42, 13869–13878.
24. Xiao, M., Leibowitz, M. J., and Zhang, Y. (2003) Concerted folding of a *Candida* ribozyme into the catalytically active structure posterior to a rapid RNA compaction, *Nucleic Acids Res.* 31, 3901–3908.
25. Das, R., Kwok, L. W., Millett, I. S., Bai, Y., Mills, T. T., Jacob, J., Maskel, G. S., Seifert, S., Mochrie, S. G. J., Thiyagarajan, P., Doniach, S., Pollack, L., and Herschlag, D. (2003) The Fastest Global Events in RNA Folding: Electrostatic Relaxation and Tertiary Collapse of the *Tetrahymena* Ribozyme, *J. Mol. Biol.* 332, 311–319.
26. Ha, B. Y., and Thirumalai, D. (1992) Conformations of a Polyelectrolyte Chain, *Phys. Rev. A* 46, R3012–R3015.
27. Rouzina, I., and Bloomfield, V. A. (1996) Macroion attraction due to electrostatic correlation between screening counterions 0.1. Mobile surface-adsorbed ions and diffuse ion cloud, *J. Phys. Chem.* 100, 9977–9989.
28. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing, *Science* 273, 1678–1685.
29. Doherty, E. A., Batey, R. T., Masquida, B., and Doudna, J. A. (2001) A universal mode of helix packing in RNA, *Nat. Struct. Biol.* 8, 339–343.
30. Bokinsky, G., Rueda, D., Misra, V. K., Rhodes, M. M., Gordus, A., Babcock, H. P., Walter, N. G., and Zhuang, X. (2003) Single-molecule transition-state analysis of RNA folding, *Proc. Natl. Acad. Sci. U.S.A.* 100, 9302–9307.
31. Deras, M. L., Brenowitz, M., Ralston, C. Y., Chance, M. R., and Woodson, S. A. (2000) Folding mechanism of the *Tetrahymena* ribozyme P4–P6 domain, *Biochemistry* 39, 10975–10985.
32. Bloomfield, V. A. (1997) DNA condensation by multivalent cations, *Biopolymers* 44, 269–282.
33. Misra, V. K., and Draper, D. E. (1999) The interpretation of Mg-(2+) binding isotherms for nucleic acids using Poisson–Boltzmann theory, *J. Mol. Biol.* 294, 1135–1147.
34. Misra, V. K., and Draper, D. E. (2002) The linkage between magnesium binding and RNA folding, *J. Mol. Biol.* 317, 507–521.
35. Manning, G. S. (1977) A field-dissociation relation for polyelectrolytes with an application to field-induced conformational changes of polynucleotides, *Biophys. Chem.* 7, 189–192.
36. Bloomfield, V. A., Crothers, D. M., and Tinoco, I. J. (2000) *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito, CA.
37. Pan, J., Thirumalai, D., and Woodson, S. A. (1997) Folding of RNA involves parallel pathways, *J. Mol. Biol.* 273, 7–13.
38. Pan, J., and Woodson, S. A. (1998) Folding intermediates of a self-splicing RNA: mispairing of the catalytic core, *J. Mol. Biol.* 280, 597–609.
39. Rook, M. S., Treiber, D. K., and Williamson, J. R. (1998) Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape, *J. Mol. Biol.* 281, 609–620.
40. Russell, R., Zhuang, X., Babcock, H. P., Millett, I. S., Doniach, S., Chu, S., and Herschlag, D. (2002) Exploring the folding landscape of a structured RNA, *Proc. Natl. Acad. Sci. U.S.A.* 99, 155–160.
41. Thirumalai, D. (1995) From the minimal models to real proteins: time scales for protein folding kinetics, *J. Phys.* 15, 1457–1467.
42. Damberger, S. H., and Gutell, R. R. (1994) A comparative database of group I intron structures, *Nucleic Acids Res.* 22, 3508–3510.

BI0356420