Persistence Length Changes Dramatically as RNA Folds

G. Caliskan,[1,2] C. Hyeon,[3] U. Perez-Salas,[2,4] R. M. Briber,[3,5] S. A. Woodson,[1] and D. Thirumalai[3]

[1]*T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA*
[2]*Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*
[3]*Biophysics Program, Institute For Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA*
[4]*Department of Physiology and Biophysics, University of California at Irvine, Irvine, California 92697, USA*
[5]*Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, USA*

(Received 13 June 2005; published 29 December 2005)

We determine the persistence length l_p for a bacterial group I ribozyme as a function of concentration of monovalent and divalent cations by fitting the distance distribution functions $P(r)$ obtained from small angle x-ray scattering intensity data to the asymptotic form of the calculated $P_{\mathrm{WLC}}(r)$ for a wormlike chain. The l_p values change dramatically over a narrow range of Mg^{2+} concentration from $\sim$21 Å in the unfolded state (U) to $\sim$10 Å in the compact (I_C) and native states. Variations in l_p with increasing Na^+ concentration are more gradual. In accord with the predictions of polyelectrolyte theory we find $l_p \propto 1/\kappa^2$ where κ is the inverse Debye-screening length.

DOI: 10.1103/PhysRevLett.95.268303 PACS numbers: 82.35.Rs, 82.39.Pj, 87.14.Gg, 87.15.He

Elucidating the mechanisms by which RNA molecules self-assemble to form three-dimensional structures is a challenging problem [1–4]. Because the native state (**N**) cannot form without significantly neutralizing the negative charge on [5,6] the phosphate group, RNA folding is sensitive to the valence, size, and shape of the counterions. At low counterion concentrations (C) RNA is unfolded (**U**) in the sense that it contains isolated stretches of base-paired *stem loops* that undergo large dynamical fluctuations. When $C > C_m$, the midpoint of the transition from **U** to the **N**, RNA becomes compact as a result of formation of tertiary contacts. For many RNA molecules, such as the *Tetrahymena ribozyme* and RNase P, folding to the native state is preceded by the formation of multiple metastable kinetic intermediates (**I**) [1,2,7].

The large dynamic conformational fluctuations in the **U** and **I** states make it difficult to characterize their structures. However, small angle scattering experiments can be used to determine the shape of RNA as it folds. The conformation of RNA in the **U**, **N**, and **I** states is characterized by R_g, the radius of gyration, and l_p, the persistence length. Small angle x-ray scattering (SAXS) [2,8–11] and small angle neutron scattering [12] experiments have been used to obtain R_g as a function of counterions for a number of RNA molecules. In contrast, l_p, which is a function of C and valence and shape of counterions, is more difficult to obtain.

In this Letter, we use SAXS data and theoretical results for the wormlike chain (WLC) to obtain l_p for a 195 nucleotide group I ribozyme from pre-tRNA(Ile) of the *Azoarcus* bacterium as a function of C for monovalent and divalent counterions. The major conclusions of the present study are: (i) The experimentally determined distance distribution functions $P(r)$ can be accurately fit using the theoretical results for wormlike chains for $r/R_g > 1$ where R_g is the radius of gyration of RNA. The l_p values, which were calculated by fitting $P(r)$ to $P_{\mathrm{WLC}}(r)$) for $r > R_g$, change dramatically from $l_p \simeq 21$ Å in the **U** state to $l_p \simeq 10$ Å in the compact conformation. (ii) The large reduction in l_p occurs abruptly over a narrow concentration range in Mg^{2+} whereas the decrease of l_p in Na^+ is gradual. This result suggests that the compaction of RNA resembles a first-order transition in the presence of multivalent counterions. (iii) For both Na^+ and Mg^{2+}, the persistence length scales as $l_p \simeq l_D^2$ where l_D is the Debye-screening length. From this finding, which is in accord with the predictions of polyelectrolyte theory, we find that the intrinsic persistence length of RNA is $l_p^o \simeq 10$ Å.

The *Azoarcus* ribozyme was transcribed *in vitro* as described previously [13]. We carried out SAXS measurements at Argonne National Lab Advanced Photon Source (BIOCAT) beam line using 1.05 Å x rays that corresponds to 11.8 keV in energy. A sample to detector distance of 1.89 m allowed us to probe momentum transfer (Q) in the range from $\sim$0.007 to $\sim$0.266 $Å^{-1}$. A quartz capillary flow cell was used to minimize the radiation damage due to x-ray exposure of a given RNA chain [9]. The measurements at various flow rates showed that x-ray radiation damage is negligible. Each measurement was averaged from four separate exposures of two seconds each. The SAXS profiles were corrected for the background signal which was measured at same buffer concentrations in the absence of RNA. Using the background corrected SAXS intensity as a function of Q, $[I(Q)]$ [Fig. 1(a)] the distance distribution function, $P(r)$, was calculated by an indirect inversion algorithm [14]. The square of the radius of gyration is given by $R_g^2 = \int r^2 P(r) dr / \int 2P(r) dr$. The ribozyme is extended at low cation concentrations and is compact at elevated values of the counterion concentration [Fig. 1(b)]. The collapse transition is highly cooperative in Mg^{2+} and is much less so in Na^+ [Fig. 1(b)].

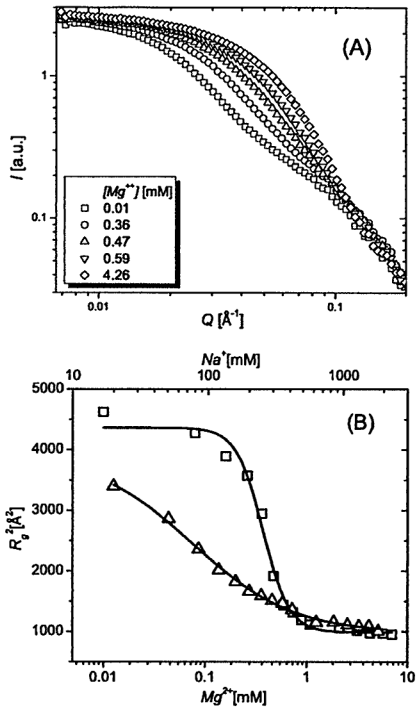

FIG. 1. (a) Scattering intensity $I(Q)$ (taken from [12(b)]) as a function of Q for 195 nucleotide *Azoarcus* ribozyme at different values of Mg^{2+} concentration in 20 mM Tris buffer, pH = 7.5 at 32 °C. The values of Mg^{2+} concentrations are given in the inset. (b) The dependance of R_g^2 on Mg^{2+} (squares) and Na^+ (triangles) concentration. Solid line is the fit using Hill equation in the form $A\{1 - [Mg^{2+}]^n/(C_m^n + [Mg^{2+}]^n)\} + y_0$, where A, n, and y_0 are adjustable parameters. We find n is 3.33 and 1.20 for Mg^{2+} and Na^+, respectively.

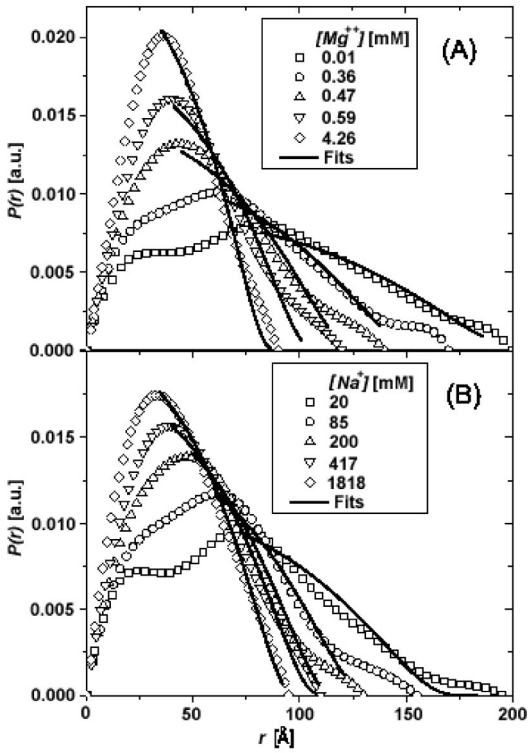

FIG. 2. Distance distribution functions. (a) $P(r)$ functions at various Mg^{2+} concentrations at 32 °C are obtained by inverting $I(q)$ by a Fourier transform with $Q_{\min} = 0.07$ Å and $Q_{\max} = 0.01$ Å. The solid lines are fits of $P(r)$ to Eq. (1). (b) Same as (a) except the counterion is Na^+. The concentrations of counterions are given in the insets.

Single molecule measurements of RNA subject to tension and our analysis of protein data bank (PDB) structure have shown that the force-extension curves can be fit using a WLC model [15]. Based on these studies we assume that RNA is a WLC for which $P(r)$ cannot be calculated analytically. However, a simple theoretical expression has been derived [16] for the end-to-end R_E distribution using a mean-field model of WLC. We expect that asymptotically ($r \gg R_g$) the behavior of $P(r)$ and the distribution of R_g or R_E should have the same functional form. Thus, for large r we predict $P(r)$ should decay as [16]

$$P_{\mathrm{WLC}}(r) \sim \exp\left(-\frac{1}{1-x^2}\right), \tag{1}$$

where $x = l_p r/R_g^2$.

At all concentrations of Na^+ and Mg^{2+}, Eq. (1) fits the data extremely accurately as long as $r/R_g > 1$ (Fig. 2). The excellent fits in Fig. 2 allow us to determine l_p as a function of the counterion concentration. When RNA is unfolded at low Na^+ or Mg^{2+} concentration, $l_p \simeq 21$ Å with $R_g \simeq 65$ Å. As the concentration of Na^+ increases from about (20–200) mM, l_p gradually decreases. There is a sharp decrease in R_g when $[Na^+] \sim 250$ mM [Fig. 1(b)] that is accompanied by a large reduction in the persistence length to $l_p \sim 10$ Å. The changes in l_p are even more dramatic in Mg^{2+} [Fig. 2(b)]. As Mg^{2+} increases from 0.01 to 0.36 mM the persistence length changes only by about 3 Å from $l_p \simeq 21$ Å (0.01 mM) to $l_p \simeq 18.3$ Å (0.36 mM). In this concentration range R_g decreases from 65 Å to 60 Å. A further increase in Mg^{2+} to 4.26 mM leads to a reduction in R_g to about 31 Å with a dramatic decrease in l_p to about 10 Å. The near discontinuous change in R_g in Mg^{2+} [Fig. 1(b)] suggests a first-order coil-globule transition in Mg^{2+}. While less common in neutral homopolymers, a discontinuous coil-globule transition has been predicted to occur in strongly charged polyelectrolytes [17].

To complement the experimental studies we calculated the $P(r)$ functions for the native three-dimensional structure of *Azoarcus* ribozyme using the coordinates from x-ray crystallography crystal structure [18] (PDB id: 1U6B) and a model based on sequence comparison [13]. The computations were done using the coordinates of only the heavy atoms (C, O, P, and N). To compare the results obtained from crystal structures and SAXS data, we used only the heavy atom coordinates for chain B (excluding nucleotides 1 and 197) from 1U6B structure to compute

$P(r)$. Similarly, the exon fragments were excluded from the Westhof model.

The $P(r)$ function from the SAXS data for the **N** state and those obtained using the x-ray structure and the Westhof model are in good agreement with each other and the SAXS data [Fig. 3(a)]. The radii of gyration for the native state calculated using $(R_g^N)^2 = \frac{1}{2N^2}\sum_i \sum_j (r_i - r_j)^2$ for the x-ray structure and the Westhof model are 31.1 and 30.7 Å, respectively. These values agree well with the results from the SAXS data ($R_g^N = 30.9$ Å). The l_p for the native state obtained by fitting the crystal structure $P(r)$ to Eq. (1) is 11 Å, while for the Westhof model we obtain $l_p \approx 10.8$ Å. The good agreement between the crystal structure (or the Westhof model) and the SAXS measurement for $P(r)$ and l_p in the I_C state suggests that the effects of complexation and interparticle interactions in the SAXS experiments are negligible.

Since RNA appears to be a charged wormlike polyelectrolyte, it is of interest to ascertain if the dependance of l_p on the Debye-screening length conforms to the theoretical predictions [19,20]. The dependence of l_p on the square of the Debye-screening length ($l_D^{-2} = \kappa^2 = 8\pi l_B I$ where $l_B = e^2/4\pi\epsilon k_B T$ is the Bjerrum length and I is the ionic strength) is linear for both Na^+ and Mg^{2+} [Fig. 3(b)]. For both flexible [21,22] and stiff polyelectrolytes [19,20] it has been shown that $l_p = l_p^o + l_p^{el}$, where l_p^o, is the intrinsic persistence length and the electrostatic contribution is $l_p^{el} \propto 1/\kappa^2$. Deviation from the Odijk-Skolnick-Fixman predictions can occur for finite-sized flexible polyelectrolytes. However, we do not expect such deviations because RNA is intrinsically stiff. Surprisingly, over the range of Na^+ and Mg^{2+} concentrations in which the *Azoarcus* ribozyme undergoes the $\mathbf{U} \rightarrow I_C$ transition, the experimental data confirm the predictions of polyelectrolyte theory. From the linear fits of l_p to κ^{-2} [Fig. 3(b)] we obtain $l_p^o \simeq 10$ Å which is similar to those found for single stranded DNA [23,24].

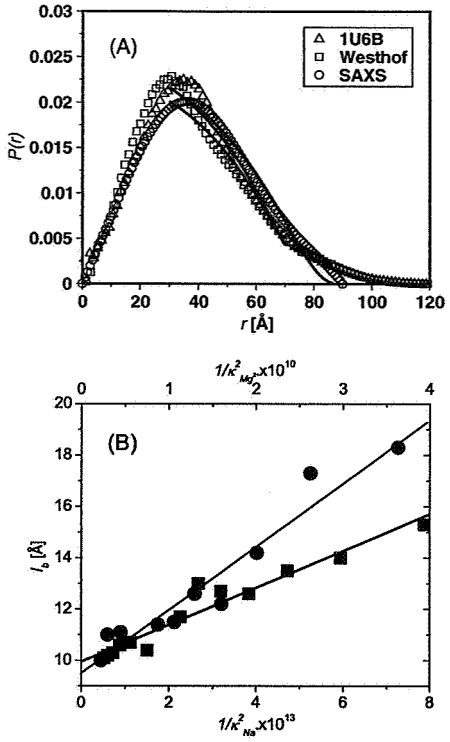

FIG. 3. (a) Calculation of $P(r)$ using the coordinates of the heavy atoms from the chain B of the crystal structure (1U6B), the Westhof model [13], and SAXS data. The symbols are given in the inset. The solid lines show the fits using $P_{\mathrm{WLC}}(r)$ in the range 30 Å to 70 Å. In this range the root mean square deviation of the $P(r)$ for the 1U6B from $P_{\mathrm{WLC}}(r)$ is 0.13 Å^{-1} and for the Westhof model [13] it is 0.10 Å^{-1}. The correlation coefficient of the fits of $P(r)$ to $P_{\mathrm{WLC}}(r)$ is 0.98. (b) Dependence of l_p on $1/\kappa^2$ in Mg^{2+} (solid circles) and in Na^+ (solid squares). Lines represent fits to the data. Note that the $1/\kappa^2$ scale for Mg^{2+} is given on top.

To assess if $P(r)$ for WLCs can be used to fit scattering measurements on other RNA molecules we used Eq. (1) and SAXS data for RNase P [9] as a function of Mg^{2+} concentration. Unlike the *Azoarcus* ribozyme, folding of RNase P is best described using three states, namely, **U**, an intermediate **I**, and the native state, **N** [9]. The **I** state is populated in the Mg^{2+} range $0.02 < Mg^{2+} < 0.2$. From the accurate fit of the SAXS data using Eq. (1) for $r/R_g > 1$ (Fig. 4), the l_p values are found to be 24.5 Å, 14.1 Å, and 11.6 Å in the **U**, **I**, and **N** states, respectively. The largest decrease in l_p and the associated R_g occurs in the $\mathbf{U} \rightleftharpoons \mathbf{I}$ transition, which is consistent with the notion that the early event in RNA collapse is initiated by counterion condensation [3].

The present work shows that the size and flexibility of RNA molecules as a function of counterion concentration

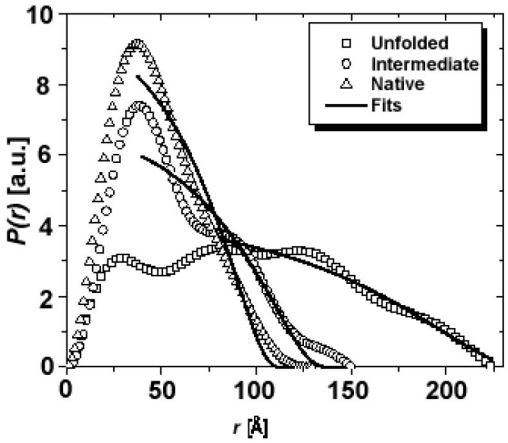

FIG. 4. $P(r)$ for RNase P from Ref. [9]. The distance distribution function was calculated by inverting $I(q)$ with a different numerical Fourier Transform method [27]. The **U**, **I**, and **N** states are for Mg^{2+} concentrations 0, 0.1, and 10 mM, respectively. The lines are fits using Eq. (1).

can be accurately obtained using scattering experiments and the WLC model. Given that RNA is a highly branched and charged polymer, it is surprising that the distance distribution functions can be described using elasticity-based polymer models for $r/R_g > 1$. Although the structural basis for such behavior is not obvious, the demonstration that single stranded DNA [24], double stranded DNA [25], and polypeptide chains [26] also behave like WLC suggests that for compatible interactions between biomolecules the local flexibility should be similar.

We are grateful to T. R. Sosnick for providing the $P(r)$ data for RNase P in tabular form. This work was supported in part by a grant from the National Science Foundation to D. T. (Grant No. 05-14056) and the National Institutes of Health to S. A. W.

[1] D. K. Treiber and J. R. Williamson, Curr. Opin. Struct. Biol. **9**, 339 (1999).
[2] T. R. Sosnick and T. Pan, Curr. Opin. Struct. Biol. **13**, 309 (2003).
[3] D. Thirumalai and C. Hyeon, Biochemistry **44**, 4957 (2005).
[4] P. Brion and E. Westhof, Annu. Rev. Biophys. Biomol. Struct. **26**, 113 (1997).
[5] V. K. Misra and D. E. Draper, Biopolymers **48**, 113 (1998).
[6] D. Thirumalai, N. Lee, S. A. Woodson, and D. K. Klimov, Annu. Rev. Phys. Chem. **52**, 751 (2001).
[7] J. Pan, D. Thirumalai, and S. A. Woodson, J. Mol. Biol. **273**, 7 (1997).
[8] K. Takamoto, R. Das, Q. He, S. Doniach, M. Brenowitz, D. Herschlag, and M. R. Chance, J. Mol. Biol. **343**, 1195 (2004).
[9] X. Fang, K. Littrell, X. J. Yang, S. J. Henderson, S. Siefert, P. Thiyagarajan, T. Pan, and T. R. Sosnick, Biochemistry **39**, 11 107 (2000).
[10] R. Russell, I. S. Millett, S. Doniach, and D. Herschlag, Nat. Struct. Biol. **7**, 367 (2000).
[11] R. Russell *et al.*, Proc. Natl. Acad. Sci. U.S.A. **99**, 4266 (2002).
[12] (a) U. A. Perez-Salas, P. Rangan, S. Krueger, R. M. Briber, D. Thirumalai, and S. A. Woodson, Biochemistry **43**, 1746 (2004); (b) S. Chauhan, G. Caliskan, R. M. Briber, P. Rangan, D. Thirumalai, and S. A. Woodson, J. Mol. Biol. **353**, 1199 (2005).
[13] P. Rangan, B. Masquida, E. Westhof, and S. A. Woodson, Proc. Natl. Acad. Sci. U.S.A. **100**, 1574 (2003).
[14] A. V. Semenyuk and D. I. Svergun, http://www.embl-hamburg.de/ExternalInfo/Research/Sax/gnom.html
[15] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, Science **292**, 733 (2001).
[16] D. Thirumalai and B.-Y. Ha, in *Theoretical and Mathematical Models in Polymer Research*, edited by A. Grosberg (Academic Press, San Diego, 1998), Chap. 1, pp. 1–35.
[17] B. Y. Ha and D. Thirumalai, Phys. Rev. A **46**, R3012 (1992).
[18] P. L. Adams, M. R. Stahley, A. B. Kosek, J. M. Wang, and S. A. Strobel, Nature (London) **430**, 45 (2004).
[19] T. Odijk, J. Polym. Sci., B Polym. Phys. **15**, 477 (1977).
[20] J. Skolnick and M. Fixman, Macromolecules **10**, 944 (1977).
[21] B. Y. Ha and D. Thirumalai, J. Chem. Phys. **110**, 7533 (1999).
[22] R. R. Netz and H. Orland, Eur. Phys. J. B **8**, 81 (1999); M. Manghi and R. R. Netz, Eur. Phys. J. E **14**, 67 (2004).
[23] C. Rivetti, C. Walker, and C. Bustamante, J. Mol. Biol. **280**, 41 (1998).
[24] B. Tinland, A. Pluen, J. Sturm, and G. Weill, Macromolecules **30**, 5763 (1997).
[25] S. B. Smith, L. Finzi, and C. Bustamante, Science **258**, 1122 (1992).
[26] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumke, and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **102**, 2754 (2005).
[27] D. I. Svergun and A. V. Semenyuk, http://www.srs.dl.ac.uk/ncd/computing/manual.gnom.html