APS physics

CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering

Adelene Y. L. Sim, Jan Lipfert, Daniel Herschlag, and Sebastian Doniach

Phys. Rev. E **86**, 021901 — Published 1 August 2012

DOI: 10.1103/PhysRevE.86.021901

Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering

Adelene Y.L. Sim[1], Jan Lipfert[2], Daniel Herschlag[3], Sebastian Doniach[1,4,5]

[1]*Applied Physics Department, Stanford University*

[2]*Department of Bionanoscience, Kavli Institute of Nanoscience, Delft Institute of Technology, The Netherlands*

[3]*Biochemistry Department, Stanford University, Stanford, California 94305, USA*

[4]*Physics Department, Stanford University, Stanford, California 94305, USA*

[5]*Biophyics Program, Stanford University, Stanford, California 94305, USA*

Abstract

Short single-stranded nucleic acids are ubiquitous in biological processes and understanding their physical properties provides insights to nucleic acid folding and dynamics. We used small angle x-ray scattering to study 8-100 residue homopolymeric single-stranded DNAs in solution, without external forces or labeling probes. Poly-T's structural ensemble changes with increasing ionic strength in a manner consistent with a polyelectrolyte persistence length theory that accounts for molecular flexibility. For any number of residues, poly-A is consistently more elongated than poly-T, likely due to the tendency of A residues to form stronger base-stacking interactions than T residues.

I Introduction

Nucleic acids play a central role in the storage, expression, and regulation of genetic information. In the cell, RNA and DNA are confined, packed, twisted, and pulled on, and many of their properties can be understood from their polymeric nature and from the basic physical principles governing the behavior of charged polymers (polyelectrolytes). Understanding these polymer properties, therefore, can inform us about fundamental physical constraints underlying nucleic acid function in the cell.

Due to the highly negatively charged backbone, nucleic acid conformation, flexibility, and folding strongly depend on ionic solution conditions. Electrostatic repulsion tends to disfavor compaction and folding of RNA and DNA. Conversely, sequence-specific interactions like base-pairing and base-stacking promote folding, and understanding these interactions, and the corresponding properties of the unfolded states under different ionic environments will further our ability to predict stable secondary [1] and tertiary structures [2, 3] and, ultimately, to develop quantitatively accurate energetic models.

Single-stranded nucleic acids (ssNA) play a number of fundamental biological roles. In RNA, single-stranded regions are ubiquitous, e.g. in mRNA, and in the single-stranded regions linking base paired regions of functional RNAs such as ribozymes or riboswitches [4, 5]. While genomic DNA exists mostly as a double-stranded helix in the cell, the DNA helix is commonly unwound as part of DNA replication and repair, thereby exposing short segments of single stranded DNA. In addition, long ssDNA stretches occur in telomeres [6] and in ssDNA

viruses [7]. Nucleic acids also are increasingly used in engineered nanostructures [8, 9] and the properties of ssNA can affect the flexibility and yield of these assemblies [10].

We employed small-angle x-ray scattering (SAXS) to probe 8-100 residue homopolymeric ssDNA molecules in solution. SAXS directly probes ssDNA conformations under a range of conditions in solution and in the absence of external perturbations such as pulling forces [11-13], fluorescent labels [14, 15], or terminal base-pairing contacts [16-19]. The homopolymeric nature of our samples minimizes the formation of secondary structure that could complicate the interpretations of intrinsic ssDNA flexibility.

II Sampling preparation and measurements

SAXS data were taken on purified ssDNA samples over a wide range of Na^+ concentrations (12.5 mM – 1 M). Measurements employed 25 mM Tris•HCl buffer, pH 8.3, and were carried out as previously described [20]. The radius of gyration (R_g) is a model-free measure of the global size of a polymer that can be directly determined from SAXS data. R_g fitting was conducted using the Debye function that describes the form factor of an unfolded polymer [21, 22] at low scattering angles (Fig. 1(a)):

$$\frac{I(q)}{I(0)} = \frac{2}{(qR_g)^4}\left((qR_g)^2 - 1 + e^{-(qR_g)^2}\right) \quad (1)$$

where $I(q)$ is the scattering intensity and $I(0)$ is the forward scattering intensity; $q = 4\pi \sin\theta / \lambda$, with 2θ and λ the scattered angle and wavelength of the x-ray respectively. This expression is valid for small q ($0 \le q \le 3R_g^{-1}$) and can be

approximated by $I(0)/I(q) = 1 + 0.359\left(qR_g\right)^{2.206}$ [21], facilitating a linear fit. The Debye approximation is the most robust approach to obtain R_g for unfolded polymers. We also tried obtaining R_g by Guinier analysis [23, 24], but this approach was less reliable due to the small range of validity for Guinier fitting because of the non-globular nature of ssDNA (c.f. the case for proteins [21]). Using regularized inversion of the data [25], we obtained R_g values similar to those obtained from the Debye fits, within experimental error [26].

For the lowest salt concentrations used in this study, we found that the normalized scattering profiles obtained at different DNA concentrations were not superimposable after rescaling by DNA concentration and showed a systematic reduction in forward scattering with increasing concentration (Fig. 1(b) and [26]). The small but systematic changes in the shape of the scattering profile at low *q* are likely due to inter-particle repulsion, i.e. due to a solution structure factor caused by the repulsion of the DNA molecules in solution, as expected for negatively charged DNA at low counterion concentration [27, 28]. For the salt concentrations where scattering profiles were not superimposable after rescaling by DNA concentration, we determined R_g values at each measured DNA concentration and extrapolated the measured R_g values linearly to zero DNA concentration to obtain the R_g values in the absence of inter-particle interference, i.e. in the infinite dilution limit (Fig. 1(b) and [26]), as was previously done for proteins [29].

III Scaling law of the radius of gyration for ssDNA

R_g decreases with increasing salt concentration for poly-A and poly-T (Fig. 2 and [26]), consistent with the DNA adopting more compact conformations when electrostatic repulsion of the backbone is screened at higher salt concentrations. In each instance, the Kratky representation ($I \cdot q^2$ versus q) of the SAXS profile (Fig. 2, bottom insets) did not show a distinct peak –a shape expected for a globular sample due to the q^{-4} dependence of I on q for a well-folded polymer [22]. Instead, the Kratky plots at low Na^+ concentrations show a linear increase at high q, characteristic of unfolded, random coil polymers. At higher salt concentrations, the shape of the profiles changes, indicative of somewhat more compact conformations, but neither poly-A nor poly-T collapses into globular forms at the salt concentrations used in this study.

The dependence of R_g on the number of bases (N) is well-described by a general scaling law of the form $R_g = A_0 N^{\nu}$ [30], and we determined A_0 and ν for poly-A and poly-T at different Na^+ concentrations from fits to these data (Fig. 2, solid lines, and Fig. 3). In general, the magnitude of ν is a measure of the flexibility of the molecule [31]. In the extreme case where ν = 1, molecular size scales linearly with the number of monomers, suggesting that monomers in the polymer are rigidly connected, as in the case of a double stranded DNA on short length scales [11, 32]. A low ν indicates greater molecular flexibility; in the limiting case where the polymer behaves like a self-avoiding random walk (SAW) chain, ν equals 0.588 for large N [33], as was observed for denatured proteins [34].

We found experimental scaling exponents around 0.7 (Fig. 3 and [26]) at the lowest investigated salt concentrations, strongly suggesting that charge-

charge repulsion make ssDNA more rigid than a SAW chain at low salt, independent of base identity. For both poly-A and poly-T, ν decreases with increasing salt concentration, when more counterions are present to more effectively screen the charge repulsion of the DNA backbone. Because there is no simple theory that predicts the SAW limit of ν for small N (finite size effects [35]), we simulated spherical beads of different radii on a string ("beads on a string" model) for small N (8-100, as was experimentally probed), and found the limiting scaling exponent to be around 0.55, depending on the radius used [26]. Additionally, since adenine and thymine are not perfect spheres and asymmetry of the monomers could further alter the limiting ν, we also carried out "sterics only" torsional angle simulations of all-atom DNA chains (using the program MOSAICS [36]) to isolate the steric effects on ν (Fig. 2, dashed lines and [26]). Our "sterics only" simulations gave a limiting ν of ~0.63 (independent of monomer type), similar to previous independent simulations of ssDNA molecules of the same size range (ref. [16]; ν = 0.62 ± 0.01). This calculated limit is a consequence of both small N and asymmetry of the monomers in the absence of electrostatic effects.

The scaling coefficient ν for poly-T at moderate salt concentrations (100 – 200 mM [Na^+]) is consistent with the simulated SAW-like behavior. At greater than 500 mM [Na^+], ν falls below the "sterics only" simulation limit (Fig. 3), which might indicate some inter-base attraction that causes poly-T to compact more than expected for a non-interacting, neutral polymer. Nonetheless, the "sterics only" simulated R_g values fall within error of the measured poly-T R_gs at 1 M [Na^+]

(Fig. 2), suggesting that poly-T on average behaves similar to an ideal SAW polymer at intermediate-to-high monovalent ion concentrations.

In contrast, the scaling exponent ν for poly-A derived from our SAXS data remains above the "sterics only" simulated limit and the R_g of each measured poly-A is consistently higher than that obtained from the "sterics only" simulations (Fig. 2); poly-A does not behave like a SAW polymer even up to 1 M [Na^+]. Poly-A is systematically stiffer than poly-T, as judged by the consistently higher values for ν (Fig. 3 and [26]). Since A and T polymers have the same overall charge and because the difference in ν persists regardless of salt concentration, it is unlikely that the observed difference in poly-A and poly-T flexibility is dominated by electrostatics. The similar scaling behavior of poly-A and poly-T in our "sterics only" simulations of DNA chains further argue against an origin related to the different sizes of adenine and thymine bases (see insets in Fig. 2). Consequently, it is likely that the disparities in behavior of poly-A from poly-T result from the differences in chemical properties of adenine (purine) and thymine (pyrimidine). In general, purine bases have a propensity to stack, while less base-stacking occurs with pyrimidine bases [37, 38]; it is likely that base-stacking interactions give poly-A strands in the range of 8-100 bases a significantly larger stiffness compared to their poly-T counterparts. These results qualitatively agree with previous reports that found larger stiffness for poly-A than poly-T on different length scales: very long length scales (N > 500) measured by atomic force microscopy [39] or short length scales ($N \leq 30$) measured by hairpin folding [17].

IV Persistence length of ssDNA

An alternative quantitative measure of flexibility is the persistence length (L_p), which is a measure of the length along the polymer chain where monomer backbone orientations are correlated. There are a variety of predictions for the dependence of L_p on salt concentration for polyelectrolytes; our experimental results allow us to test theories on short length scales ($N \leq 100$ compared to $N \rightarrow \infty$ that are often discussed in theories). Poly-T behaves similar to an ideal polyelectrolyte, presumably due to the negligible stacking interactions of pyrimidine, so we focused on poly-T for comparison of L_p to electrostatic theories.

L_p was first estimated from the R_g scaling data [26] using the worm-like chain model [40, 41]:

$$R_g^2 = \left(\frac{lL_P}{3}\right) - L_P^2 + \left(\frac{2L_P^3}{l}\right) - \left(\frac{2L_P^4}{l^2}\right)\left(1 - e^{-l/L_P}\right) \quad (2)$$

where $l = N{\cdot}a$ is the contour length, with N the number of bases and a the effective monomer length. We found that L_p decreased from about 32 Å to 10 Å as [Na^+] increased from 12.5 mM to ~1 M ([26] and Fig. 4). The effective monomer length, a showed little systematic dependence on [Na^+] in the same range [26], and on average was 6.5 ± 0.7 Å, within range of prior reported values from ssDNA-protein crystal structures (6.3 ± 0.8 Å [19]), other experimental methods (~5.2 Å [18]; 4.0 - 4.5 Å [42]) and simulations (6.7 ± 0.7 Å [16]).

A second approach to obtain L_p was to fit the full scattering profile (extrapolated to infinite dilution if necessary) to the form factor ($I_{\mathrm{WLC}}(q)$) derived for a worm-like chain model without excluded volume effects [43]. The finite cross-section of the nucleic acid was accounted for with the mean squared cross-

sectional radius of gyration (R_{cs}^2) so that $I(q) = I(0)I_{WLC}(q)\exp(-q^2R_{cs}^2/2)$ [15]. Each scattering profile was fit to this model, setting the contour length as $N{\cdot}a$ with fixed a = 6.5 Å to reduce the number of free fitting parameters [26].

The L_p values determined from the two fitting approaches agreed within experimental error ([26] and Fig. 4) and fall within the rather wide range of values (7.5 – 78 Å) previously found using a variety of different experimental methods (refs. [11, 14, 18, 19, 44, 45]; see [26]).

L_p is often separated into two components: $L_p = L_0 + L_e$, where L_0 is the intrinsic persistence length (due to bond flexibilities) that is independent of salt conditions, while L_e is the electrostatic persistence length arising from repulsion between like charges within the polyelectrolyte [46, 47]. Fitting to the form $L_p = L_0 + m[\mathrm{Na}^+]^c$ to our poly-T data derived from R_g gives $L_e \propto [Na^+]^{-0.44}$ and an L_0 = 9.8 Å (or $L_e \propto [Na^+]^{-0.69}$ and L_0 = 12.5 for the L_p estimates from fitting the full scattering profiles).

Two main theoretical relationships between L_e and [Na^+] have been proposed, namely the Odijk and Skolnick-Fixman (OSF) [46, 47] relation of $L_e \propto [Na^+]^{-1}$ and the Barrat and Joanny (BJ) model ($L_e \propto [Na^+]^{-0.5}$) [48, 49]. The main difference between these theories is how the flexibility of the polyelectrolyte is modeled. The OSF theory assumes that small angular fluctuations within the polymer (due to chain flexibility) are negligible compared to electrostatic effects. Therefore this theory is expected to break down for flexible chains and weakly charged polyelectrolytes; using variational calculations, chain flexibility is

incorporated into the BJ model making it applicable for flexible polyelectrolytes [48, 49].

Our L_p results show a weaker dependence of L_p on [Na^+] than OSF theory suggests, and instead appears to be consistent with BJ theory (Fig. 4). Recently, Chen *et al.* observed that the persistence length of poly-T varied with [Na^+]$^{-1}$ in accordance with OSF theory [15]. In contrast, measurements based on single-molecule Föster resonance energy transfer [42] and hydrodynamic radius measurements [16] determined much weaker salt dependences of $L_e \propto$ [Na^+]$^{-0.2\pm0.05}$ and $L_e \propto$ [Na^+]$^{-0.22\pm0.01}$, respectively. However, the two studies found widely different values for L_0: Laurence *et al.* obtained a negative value for L_0 [42], while Doose et al. used an estimate of L_0 = 17 Å derived from "sterics only" ssDNA simulations [16]. Finally, it was shown that L_p for denatured ssDNA under tension follows the BJ scaling law, with an L_0 of about 6.2 Å [13], in approximate agreement with our solution scattering results. The apparent discrepancies between different experimental results could be due to differences in ssDNA sequence, experimental techniques, and/or the assumptions entering the analyses of the data and should inspire future work using common sequences over multiple techniques.

Based on further theoretical work by Ha et al. [50], the expected scaling of L_e on [Na^+] depends on the value of the parameter $(L_0L_B)/A^2 \approx (L_0L_B)/a^2$ (L_B is the Bjerrum length; A is the average distance between charges that we estimate as effective monomer length, or a as defined previously in Equation (2)): $(L_0L_B)/a^2 << 1$ or $>> 1$, the OSF scaling prevails; the BJ theory is valid for intermediate values

of $(L_0L_B)/a^2$. In addition, the authors found that weaker or more complex dependence of L_e on [Na^+] can also exist in this regime. Since we have $(L_0L_B)/a^2 \approx$ 1.7, our measurements are around the intermediate range of $(L_0L_B)/a^2$, and therefore $L_e \propto [Na^+]^{-0.44}$ or $L_e \propto [Na^+]^{-0.69}$ is consistent with these theoretical expectations.

V Summary

In conclusions, our SAXS studies of the polymer properties of short poly-T and poly-A under different salt conditions provide a baseline for understanding nucleic acid folding, can guide theoretical developments of polyelectrolyte behavior under finite-size limits, and can serve as a tractable model system for testing the accuracy of nucleic acid simulations.

Acknowledgements

We thank Andy Spakowitz and Michael Levitt for useful discussions and Sönke Seifert for help with SAXS measurements. This work was supported by the Netherlands Organization for Scientific Research (NWO) and the Agency for Science, Technology and Research (Singapore) and in part by National Institutes of Health Grant GM 49243 to D.H. This work was also partially supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, under contract DE-AC02-76SF00515. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No.

DE-AC02-06CH11357. Simulations were done on the Stanford's Bio-X^2 computers [National Science Foundation (NSF) award CNS-0619926].

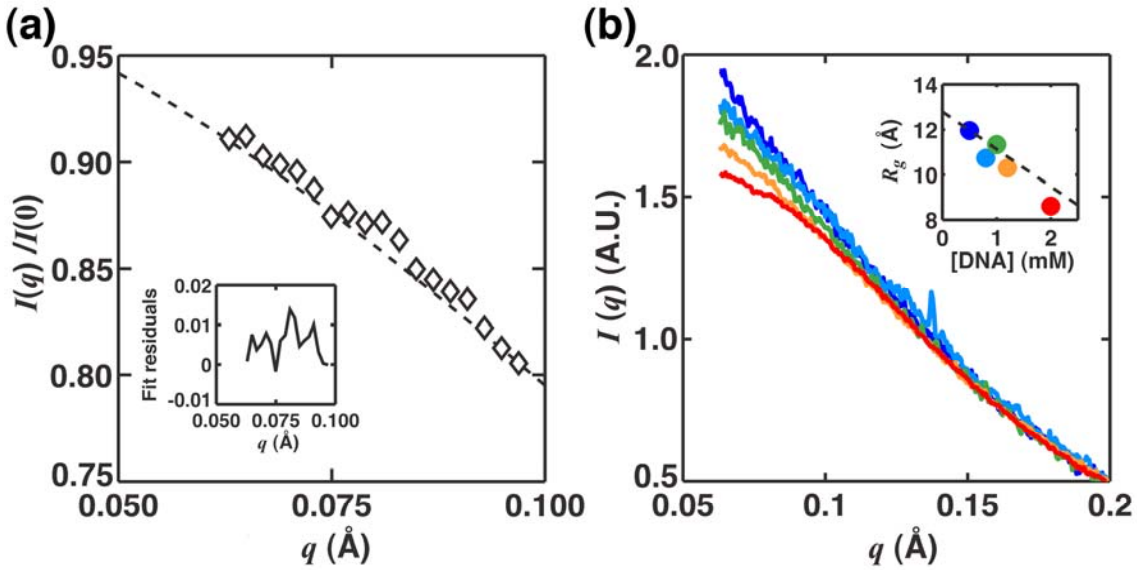

FIG. 1 (color): Determination of R_g from experimental data. (a) Fitting the Debye function to the experimental scattering profile. The residuals of the fit are shown in the inset (for 2 mM of DNA). (b) To determine the effects of inter-particle interference, measurements were taken at different DNA concentrations ([DNA]). If a systematic trend is observed in the scattering profiles (each color indicates a different DNA concentration; color scheme goes from blue to red as [DNA] increases), the trend in R_g is extrapolated to zero [DNA] to estimate the R_g in the absence of inter-particle interference (inset). In both panels, the results for poly-A8 in 25 mM Na^+ are shown.

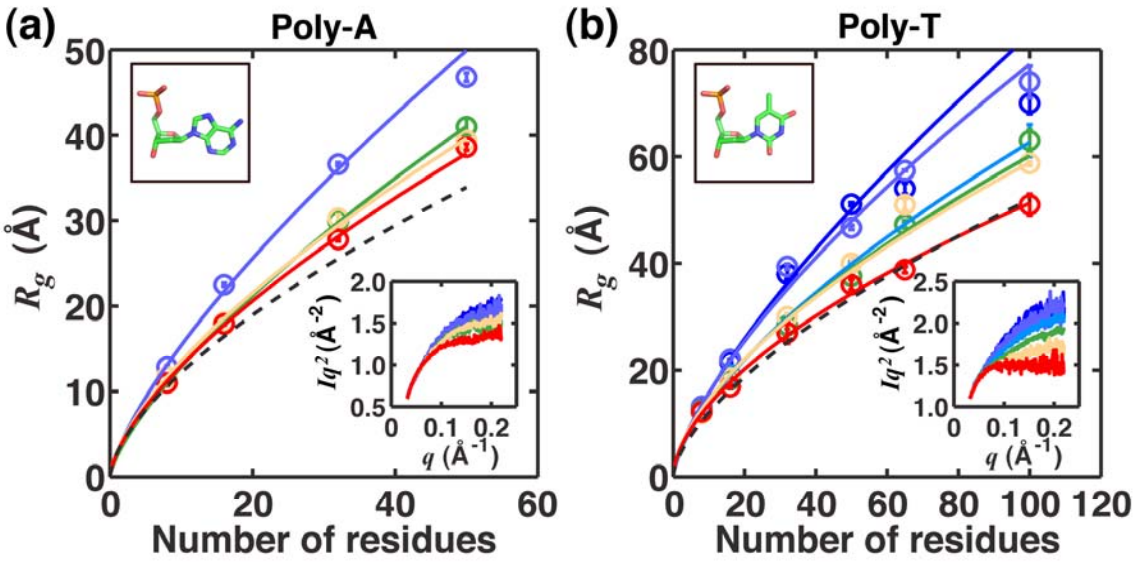

FIG. 2 (color): Radii of gyration (R_g) as a function of number of bases for single-stranded DNA homopolynucleotides poly-A (a) and poly-T (b) in the presence of different concentrations of Na^+: 12.5 mM (dark blue), 25 mM (light blue), 125 mM (cyan), 225 mM (green), 525 mM (yellow), 1025 mM (red). Fits of the scaling law to the experimental data (see main text) are shown as corresponding colored lines. The predictions from the "sterics only" simulations (see main text) are shown as dashed black lines. Idealized structures of adenine and thymine bases are shown in top insets in either panel. The small angle x-ray scattering profiles (in Kratky representations) for poly-A50 and poly-T50, respectively, are shown as bottom insets with the same color scheme for Na^+ concentration. The scattering profiles for poly-T50 show more variation with salt concentration than those for poly-A50.

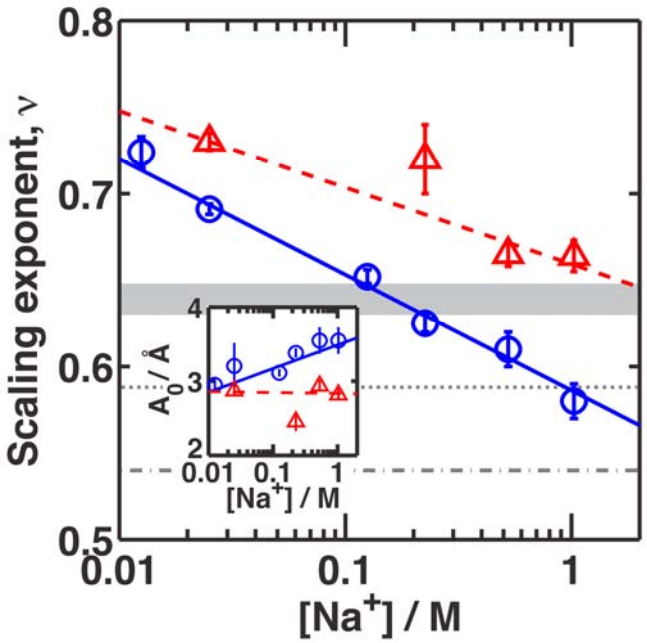

FIG. 3 (color online): Dependence of the scaling exponent (ν) on Na^+ concentration. The scaling exponent of poly-T (blue circles) and poly-A (red triangles) decrease linearly with log([Na^+]) with slopes of -0.067 and -0.044 respectively. For comparison, the scaling exponent for a self-avoiding polymer in the approximation of large number of monomers (ν = 0.588) is shown as a dotted line while that for a "beads on a string" model to simulate finite size effects (ν = 0.54) is shown as a dot-dash line. The gray region indicates the range of ν found from "sterics only" simulations of poly-A and poly-T. Inset: The scaling prefactor (A_0) increases slowly with [Na^+] for poly-T, but remains approximately constant for poly-A across about two orders of magnitude changes in [Na^+].

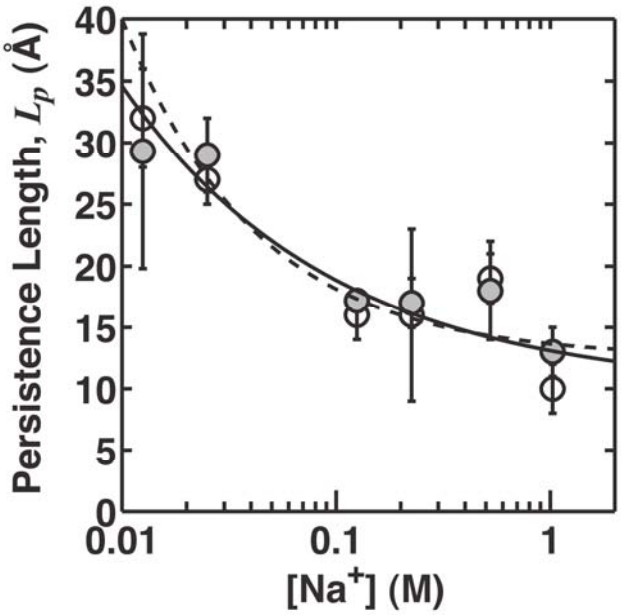

FIG. 4: The effects of monovalent salt concentration on persistence length of poly-T determined using two fitting schemes (worm-like chain fitting to R_g scaling in black and fitting individual scattering profiles in gray filled circles respectively). The persistence length of poly-T varies sharply with [Na^+]. The best-fit curves (solid lines and dashed line for the different L_p fitting protocols respectively) yield a dependence of [Na^+]$^{-0.44}$ and [Na^+]$^{-0.69}$ respectively.

References

[1] M. Zuker, Curr Opin Struct Biol 10, 303 (2000).
[2] R. Das, J. Karanicolas, and D. Baker, Nat Methods 7, 291 (2010).
[3] Z. J. Tan, and S. J. Chen, Biophys J 101, 176 (2011).
[4] P. Brion, and E. Westhof, Annu Rev Biophys Biomol Struct 26, 113 (1997).
[5] V. B. Chu *et al.*, RNA 15, 2195 (2009).
[6] J. E. Croy, and D. S. Wuttke, Trends Biochem Sci 31, 516 (2006).
[7] D. P. Martin *et al.*, Viruses 3, 1699 (2011).
[8] P. W. Rothemund, Nature 440, 297 (2006).
[9] N. C. Seeman, Annu Rev Biochem 79, 65 (2010).
[10] D.-N. Kim *et al.*, Nucl Acids Res (2012).
[11] S. B. Smith, Y. Cui, and C. Bustamante, Science 271, 795 (1996).
[12] Y. Seol, G. M. Skinner, and K. Visscher, Phys Rev Lett 93 (2004).
[13] O. A. Saleh *et al.*, Phys Rev Lett 102, 068301 (2009).
[14] B. Tinland *et al.*, Macromolecules 30, 5763 (1997).
[15] H. Chen *et al.*, Proc Natl Acad Sci USA 109, 799 (2012).
[16] S. Doose, H. Barsch, and M. Sauer, Biophys J 93, 1224 (2007).
[17] N. L. Goddard *et al.*, Phys Rev Lett 85, 2400 (2000).
[18] J. B. Mills, E. Vacano, and P. J. Hagerman, J Mol Biol 285, 245 (1999).
[19] M. C. Murphy *et al.*, Biophys J 86, 2530 (2004).
[20] J. Lipfert *et al.*, Rev Sci Instrum 77, 046108 (2006).
[21] P. Calmettes *et al.*, Biophys Chem 53, 105 (1994).
[22] S. Doniach, Chem Rev 101, 1763 (2001).
[23] J. Lipfert, and S. Doniach, Annu Rev Biophys Biomol Struct 36, 307 (2007).
[24] A. Guinier, Ann Phys 12, 161 (1939).
[25] D. I. Svergun, J Appl Crystallogr 25, 495 (1992).
[26] See supplemental material at [URL will be inserted by publisher] for additional figures and tables.
[27] Y. Bai *et al.*, Proc Natl Acad Sci USA 102, 1035 (2005).
[28] X. Y. Qiu *et al.*, Phys Rev Lett 96, 138101 (2006).
[29] B. Zagrovic *et al.*, Proc Natl Acad Sci USA 102, 11698 (2005).
[30] P. G. de Gennes, *Scaling concepts in polymer physics* (Cornell University Press, 1979).
[31] A. V. Dobrynin, R. H. Colby, and M. Rubinstein, Macromolecules 28, 1859 (1995).
[32] C. Bustamante *et al.*, Science 265, 1599 (1994).
[33] M. Doi, and M. Edwards, *The Theory of Polymer Dynamics* 1986).
[34] J. E. Kohn *et al.*, Proc Natl Acad Sci USA 101, 12491 (2004).
[35] D. C. Rapaport, J Phys A 18, 113 (1985).
[36] P. Minary.
[37] J. Isaksson *et al.*, Biochemistry 43, 15996 (2004).
[38] W. Saenger, *Principles of nucleic acid structure* (Springer-Verlag, 1984).
[39] C. Ke *et al.*, Phys Rev Lett 99, 018302 (2007).

[40] J. B. Hays, M. E. Magar, and B. H. Zimm, Biopolymers 8, 531 (1969).
[41] H. Benoit, and P. Doty, J Phys Chem 57, 958 (1954).
[42] T. A. Laurence *et al.*, Proc Natl Acad Sci USA 102, 17348 (2005).
[43] J. S. Pedersen, and P. Schurtenberger, Macromolecules 29, 7602 (1996).
[44] S. V. Kuznetsov *et al.*, Biophys J 81, 2864 (2001).
[45] C. Rivetti, C. Walker, and C. Bustamante, J Mol Biol 280, 41 (1998).
[46] T. Odijk, J Polym Sci 15, 447 (1977).
[47] J. Skolnick, and M. Fixman, Macromolecules 10, 944 (1977).
[48] J. L. Barrat, and J. F. Joanny, Europhys Lett 24, 333 (1993).
[49] J. L. Barrat, and J. F. Joanny, Adv Chem Phys 94, 1 (1996).
[50] B. Y. Ha, and D. Thirumala, J Chem Phys 110, 7533 (1999).