Stretching of Homopolymeric RNA Reveals Single-Stranded Helices and Base-Stacking

Yeonee Seol,* Gary M. Skinner, and Koen Visscher
Department of Physics, University of Arizona, Arizona 85721, USA

Arnaud Buhot and Avraham Halperin
Structure et Propriétés d'Architectures Moléculaires (SPrAM), UMR 5819 (CEA, CNRS, UJF), CEA Grenoble, France
(Received 14 July 2006; published 12 April 2007)

We have found strong supporting evidence for the helical structures of single-stranded nucleic acids by stretching individual molecules of polyadenylic acid [poly(A)] and polycytidylic acid [poly(C)]. Analyzing the force versus extension data using a two-state elastic model in which random-coil domains alternate with rigid helical domains allows one to extract the thermodynamic and structural properties. In addition, it also yields moderate to low cooperativity of the helix-coil transition for poly(A) and poly(C), respectively.

DOI: 10.1103/PhysRevLett.98.158103 PACS numbers: 87.14.Gg, 87.15.He, 87.15.La

Single-stranded (ss) nucleic acids (NA) can form helical domains [1,2]. These are thought to arise because of stacking interactions favoring parallel orientation of adjacent bases. The evidence for stacking of NA in solution is mostly based on calorimetric measurements and spectroscopy, but the thermodynamic parameters extracted from these measurements are not consistent and vary over a wide range [2]. Attempts to use scattering techniques to probe the chain configurations yield inconclusive data because rodlike configurations are only attained at low temperatures. At room temperature, the chains behave as rod-coil multiblock copolymers with short helical domains whose hydrodynamic radius is similar to that of a flexible coil. Here, we present evidence for the occurrence of helices at room temperature by stretching ssRNA homopolymers. The resulting force versus extension (F-x) curves exhibit plateaus indicative of the presence of ss helical domains. In contrast to the traditional methods used to investigate ss helices, single-molecule measurements at fixed temperatures yield a clear signature of such helices. This method also provides an alternative route for determining the thermodynamic parameters of stacking interactions. Renewed interest in stacking effects has arisen because of a controversy concerning the interpretation of ''molecular beacon'' experiments [3–6]. These involve short ssDNA chains capable of forming stem-loop structures. Experiments by the Libchaber group revealed differences in cyclization behavior of polydeoxythymidylate [poly(dT)] and polydeoxyadenylate [poly(dA)] loops that were attributed to stacking and its effects on the rigidity of the chains [3,6,7]. In particular, it was argued that the elasticity of the two species is different because stacking is significant in poly(dA) but negligible for poly(dT). This interpretation was disputed by Ansari *et al.* who ascribed the effects to transient trapping of misfolded loops while arguing that both poly(dT) and poly(dA) behave as flexible polymers and their elastic properties are indistinguishable [4,5]. Although our results concern longer chains of poly(A) and poly(C), we present clear evidence of deviations from simple random-coil elasticity models, deviations indicative of helical structure formation. The interest in ss helices is however of wider origins. The occurrence of ss helices is expected to affect the secondary structure and elastic properties of RNA, and the thermodynamics of hairpin or loop formation can be expected to depend upon the fraction of ss helices. This also affects the elastic penalty incurred upon bending and/or stretching. From a biological point of view, it is important to note the prevalence of homopolymeric RNA where stacking interactions are of importance. For example, mRNA is generally modified to include a 3' poly(A) tail which reaches lengths of ~250 nucleotides (nts), regulating mRNA stability [8,9]. Tracts of poly(C) and poly(U) play similar roles [10,11]. Interactions of these tracts with regulatory proteins can be expected to depend upon sequence, structure, base stacking, and elastic properties.

Earlier stretching experiments of ssDNA did not reveal any signatures of ss helix formations [12,13]. Presumably, the multitude of possible stacking free energies and the distribution of different adjacent bases tends to obscure any plateaus in the F-x curve. Heteropolymers also tend to form base pairs and double-stranded hairpins. Although denaturants used to prevent secondary structure have enabled stretching experiments of ssDNA, they are likely to also affect base stacking [13]. Here, we report F-x data for poly(A) and poly(C) by stretching individual molecules using optical tweezers. To the best of our knowledge, these data are the first to show any experimental signature of a helix-coil transition in ss NAs by mechanical force. Analysis of these data yields quantitative information about cooperativity as well as structural and elastic properties.

All molecules were synthesized enzymatically by elongation of 5'-end biotinylated oligonucleotides (uridylic

TABLE I. Summary of fitting parameters for poly(A) and poly(C).

Parameters	Poly(A)	Poly(A)[a]	Poly(C)	Poly(C)[a]
a (nm)	0.624 ± 0.014	0.6	0.606 ± 0.007	0.6
b (nm)	0.370 ± 0.013	0.357 ± 0.011	0.387 ± 0.006	0.383 ± 0.007
λ (nm)	1.13 ± 0.29	1.14 ± 0.29	1.71 ± 0.26	1.81 ± 0.34
ΔG (kcal/mol)	−0.35 ± 0.10	−0.34 ± 0.10	−0.53 ± 0.04	−0.54 ± 0.05
σ	0.58 ± 0.15	0.53 ± 0.14	0.82 ± 0.08	0.81 ± 0.05
F_0 (pN)	0.53 ± 0.58	0.54 ± 0.58	0.70 ± 0.35	0.67 ± 0.39
N	2350 ± 640	2400 ± 590	1670 ± 460	1690 ± 450

[a]Fitting was performed with fixed $a = 0.6$ nm.

acid 20-mer) using polynucleotide phosphorylase (PnP) [14]. The formation of tetrads precludes stretching experiments of poly(G) [15]. Since PnP does not require a template for elongation, the resulting RNA sizes are indefinite and vary from molecule to molecule: 1500–4000 bases as estimated by gel electrophoresis. Poly(C) molecules were somewhat shorter than poly(A) as reflected in Table I. The contour length for individual molecules can therefore only be determined by fitting the force-extension data with models in which the total number of bases N appears as a fitting parameter. After 3'-end labeling with digoxigenin-dU using terminal transferase, molecules were stretched between a streptavidin-coated microscope cover glass and an antidigoxigenin coated polystyrene bead (diameter 0.49 ± 0.024 μm) held with optical tweezers identical to our previous experiments with poly(U) [16]. The dig-dU-tail is expected to be fully adsorbed on the antidig coated bead, and consistent with this, no signatures of any hairpin formation of the tail with poly(A) segments have been observed. Surfaces were blocked with 0.3% w/v acetylated BSA to prevent nonspecific surface interactions of RNA and/or beads. We did not observe any signatures of unwanted nonspecific surface interactions. Concentrations of RNA and beads were chosen such that only very few tethers were formed per flow cell, thus making it unlikely a bead would be tethered by multiple RNA molecules. F-x curves were calculated from the bead and stage displacement, correcting for the experimental geometry as previously described [17].

Figure 1 shows typical examples of the F-x curves of poly(A), poly(C) and poly(U) obtained at 500 mM Na^+ concentration where electrostatic interactions due to the charged phosphodiester backbone can be neglected [16]. We have recently shown that poly(U)'s elasticity is well described by a wormlike chain (WLC) model [16]. The F-x curves of poly(A) and poly(C) are markedly different, typified by a ''plateau'' that represents a change in the molecules' contour length, readily explained by the occurrence of helix-coil transitions. This transition differs from cooperative overstretching in dsDNA where the molecule almost doubles its length once a critical force has been reached [18]. The more gradual unfolding observed here suggests a predominantly noncooperative helix-coil transition for ss RNA [19,20]. We did not observe any hysteresis upon relaxing the molecules (data not shown), consistent with the rapid relaxation of base stacking reported in temperature-jump experiments [21].

We have analyzed F-x data according to the model by Buhot and Halperin [20,22]. In summary, it assumes that unstacked random-coil domains behave as freely-jointed chains, whereas the helical domains have an effectively infinite persistence length (Fig. 2). This assumption is justified so long as the helical domains are expected to be short because of weak cooperativity of stacking. Cooperativity is introduced using the Zimm-Bragg formalism [23] and the full distribution of helical and coil domain size is accounted for explicitly. The F-x relation is then obtained as [20,24]

$$x(f) = Na\left\{(1-\theta)L(\delta f) + \gamma\theta\coth(\gamma f) - \frac{\gamma s\theta A(f,\theta,y)}{\sinh(\gamma f)} - \frac{y}{f}\right\}, \quad (1)$$

where x is the extension; N the number of monomers; a the interphosphate distance per nt in the random coil; and b the rise per nt in the helix (Fig. 2); $\gamma = b/a$; $\delta = 2\lambda/a$ with λ

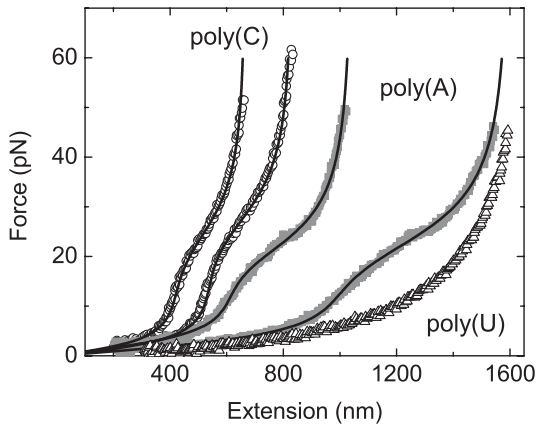

FIG. 1. Two examples of F-x curves for poly(A) (solid square) and poly(C) (open circle) and fits (solid line). A F-x curve for poly(U) is also shown (open triangle), offset by +100 nm along the horizontal axis.

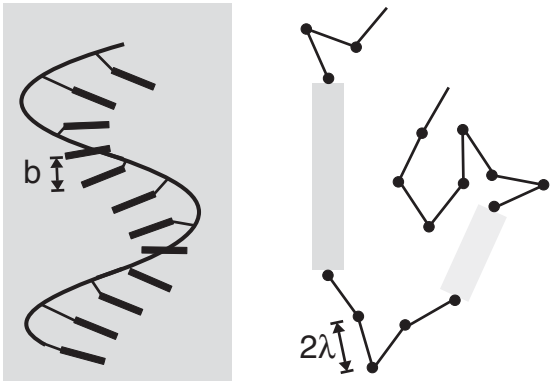

FIG. 2. A schematic of a single-stranded helix stabilized by base-stacking interactions (gray boxes). b is the rise per nucleotide in the helix. The random-coil domains are modeled as freely-jointed chains in which the base associated with each chain segment is oriented randomly with respect to its neighbors. Each Kuhn segment of the chain contains $2\lambda/a$ bases, with a the interphosphate distance per nucleotide in the random coil, and λ the persistence length.

the persistence length of the random coil; $L(z) = \ln[\sinh(z)/z]$; and $A(f, \theta, y) = [(1 - \theta - y)/(1 - \theta)] \times [\delta f/\sinh(\delta f)]^{1/\delta}$. The fraction of stacked bases θ and helices y are solutions of

$$\theta\gamma f(1 - sAe^{\gamma f})(1 - sAe^{-\gamma f}) = (1 - \theta - y)sA\sinh(\gamma f) \tag{2}$$

$$2\gamma f y = \sigma(1 - \theta - y)\ln[(1 - sAe^{-\gamma f})/(1 - sAe^{\gamma f})] \tag{3}$$

with $s = \exp(-\Delta G/\mathrm{kT})$; ΔG the free energy difference, and σ the Zimm-Bragg cooperativity factor quantifying the Boltzmann cost associated with the boundaries between stacked and unstacked domains; and $f = Fa/kT$ the normalized force with F, the applied tension. The average length of the helical domains follows from θ/y [20].

Force vs extension data were fitted using minimization of the mean square difference between the experimental and theoretical extensions. Parameters considered for the fit are N, a, b, λ , ΔG, and σ. We introduced another free parameter, a shift F_0 of the tension reflecting possible variation in the calibration of the force. This parameter improves the fit without affecting the results for poly(A) but seemed necessary for poly(C) where cooperativity is less pronounced. In the latter case, the fitting results were obtained for the 7 F-x curves with their highest tension larger than 50 pN. The high-tension regime above the plateau for the other curves was too short to extract pertinent parameters from the fits. The number of free parameters is inherent to a simple two-state model necessary to explain the plateau in the F-x curves. However, it has the advantage that it enables us to relate the parameters to known quantities. Fitting examples are shown in Fig. 1. The results for the parameters are summarized in Table I, with the errors designating the standard deviation resulting from the fits. The error for N reflects the length distribution of the molecules. We have investigated the error due to trap calibration errors by rescaling the force axis or a force offset: a 5% calibration error or an offset of 0.5 pN produces relative errors within those derived from Table I.

The length of monomers in the coil domain was found $a \sim 0.6$ nm and agrees with expected values [12,25]. Fixing this value, $a = 0.6$ nm, reduces the number of free parameters without affecting their values. Fitting of a simple wormlike chain model to the low-force and high-force regimes, i.e., excluding the force regime where most of the transitions occur, yields a 1.6-fold increase in contour length when a helix turns coil, thus estimating the rise per nucleotide in the helix form, b, as 0.38 nm for poly(A) and poly(C). Fits to our model produce comparable results (0.37 and 0.39 nm) and were slightly larger than those in the B-DNA double helix (0.34 nm) and proposed structures of poly(A) and poly(C) [26]. Poly(A) is thought to form a right-handed, A-type, ss helix with nine nts per helical rise and a rise per nucleotide of 0.28 nm [26]. For poly(C), two structures have been proposed. X-ray fiber diffraction studies are consistent with an A-type right-handed helix with 6 nts per turn and a rise of 0.31 nm per nt [26]. A NMR study, performed in a neutral aqueous environment, has yielded a left-handed helix with 8 nts per turn and a rise of 0.29 nm per nt [27]. This helix, with its bases at the outer periphery rather than on the inside, is thought to be stabilized by hydrogen bonding between the amino and carbonyl groups of adjacent cytosines. Such a finding may be consistent with the fact that the enthalpic energies for associations of cytidines and stacking of poly(C) differ, whereas in the case of poly(A) and adenosines these are similar [2]. Our data and analysis cannot sufficiently distinguish either structure for poly(C). In the case of poly(U), base-stacking interactions are weak and no stable helices form, explaining the flexible random-coil behavior observed in early hydrodynamic and recent stretching experiments [16,28,29]. These observations do not rule out stacking/unstacking much faster than accessible by stretching experiments.

Consistent with existing literature, cooperativity was found to be weak, but stronger (lower σ) for poly(A) than for poly(C). As stacking is an interaction between nearest-neighbors, it is expected to be weakly cooperative [1,2]. Thermal melting data produced a broad range of Zimm-Bragg factors ranging $\sigma = 0.1$–1.0 [21]. Combining calorimetric and spectroscopic data can further constrain σ, yielding $\sigma = 0.8$–1.0 in the case of poly(C) [21]. However, such studies have not been performed for poly(A), and conflicting values have been reported: from noncooperative [30] to modestly cooperative [21,31]. We found the stacking free energy to be larger for poly(C)

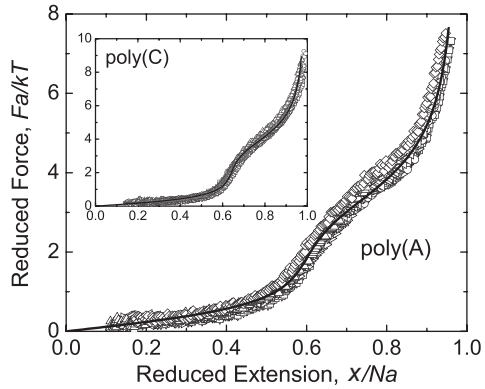

FIG. 3. Rescaled F-x curves for 7 poly(A) and 7 poly(C) chains (inset).

which compares with known results ($\Delta G = -0.02$ to -0.45 kcal/mol and $\Delta G = -0.66$ kcal/mol for poly(A) and poly(C), respectively) [2]. Considering the almost identical contour lengths per monomer for poly(A) and poly(C), these findings are consistent with the larger forces required for full unwinding of helical domains of poly(C). It is interesting to fit the low-force regime, up to 12 pN, using a simple wormlike chain (WLC) model. These fits, albeit somewhat crude due to the limited and featureless data, yield the persistence length for poly(C) larger than for poly(A), 3.1 nm vs. 1.8 nm, respectively, indicating that the helical content of poly(C) may exceed that of poly(A) in the low-force regime, and suggesting that a helix-coil transition in poly(C) requires a larger force.

The persistence lengths (λ) found for the random-coil domains of poly(A) and especially poly(C) are somewhat larger than for poly(U) [16]. However, the freely rotating junctions between helices and coil domains, as assumed in the model, tend to increase the flexibility of the chains. This effect is then compensated by a larger apparent λ.

To test for self-consistency of the model, we normalized the F-x data of all molecules and confirmed that these data fall on a single master curve (Fig. 3). This indicates that our estimates of N are reasonable, and also that the molecules show the same elastic behavior. Thus, homopolymeric ssRNA may well function as a model system for studying helix-coil transitions in nucleic acids. For example, future stretching experiments conducted as a function of temperature will be able to determine the enthalpic and entropic contributions to the free energy.

Summarizing, single-molecule stretching experiments of ss RNA molecules exhibit a plateau, the most direct evidence for the presence of helical domains in these polymers. Such experiments and subsequent data analysis using a molecular Zimm-Brag type model provide an alternative means for determining the thermodynamic and structural quantities of these block copolymers and the base-stacking interactions thought to stabilize the helical domains.

G. M. S. thanks the University of Arizona BIO5 Institute and Jane Coffin Childs Memorial Fund for support.

*Present address: JILA, University of CO, Boulder, CO 80309, USA

[1] C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry* (Freeman, New York, 1980).
[2] D. H Turner, "Conformational Changes," in *Bloomfield*, edited by V. A. Crothers and D. M. Tinoco, I. Nucleic Acids,: Structure, Properties and Functions (University Science, Sausalito, CA, 2000), Chap. 8.
[3] N. L. Goddard *et al.*, Phys. Rev. Lett. **85**, 2400 (2000).
[4] A. Ansari *et al.*, Proc. Natl. Acad. Sci. U.S.A. **98**, 7771 (2001).
[5] A. Ansari *et al.*, Phys. Rev. Lett. **88**, 069801 (2002).
[6] N. L. Goddard *et al.*, Phys. Rev. Lett. **88**, 069802 (2002).
[7] D. P. Aalberts *et al.*, Biophys. J. **84**, 3212 (2003).
[8] T. M. Hall, Curr. Opin. Struct. Biol. **12**, 82 (2002).
[9] M. Dreyfus and P. Régnier, Cell **111**, 611 (2002).
[10] H. Huang *et al.*, Nature (London) **391**, 99 (1998).
[11] A. V. Makeyev and S. A. Liebhaber, RNA **8**, 265 (2002).
[12] S. B. Smith *et al.*, Science **271**, 795 (1996).
[13] M.-N. Dessinges *et al.*, Phys. Rev. Lett. **89**, 248102 (2002).
[14] J. N. Wood and D. W. Hutchison, Nucleic Acids Res. **3**, 219 (1976).
[15] S. Arnott *et al.*, Biochem. J. **141**, 537 (1974).
[16] Y. Seol *et al.*, Phys. Rev. Lett. **93**, 118102 (2004).
[17] M. D. Wang *et al.*, Biophys. J. **72**, 1335 (1997).
[18] I. Rouzina and V. A. Bloomfield, Biophys. J. **80**, 882 (2001).
[19] M. N. Tamashiro and P. Pincus, Phys. Rev. E **63**, 021909 (2001).
[20] A. Buhot and A. Halperin, Phys. Rev. E **70**, 020902(R) (2004).
[21] S. M. Freier *et al.*, Biochemistry **20**, 1419 (1981).
[22] A. Buhot and A. Halperin, Phys. Rev. Lett. **84**, 2160 (2000).
[23] B. H. Zimm and J. K. Bragg, J. Phys. Chem. **31**, 526 (1959).
[24] A. Buhot and A. Halperin, Macromolecules **35**, 3238 (2002).
[25] J. B. Mills *et al.*, J. Mol. Biol. **285**, 245 (1999).
[26] W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1984).
[27] M. S. Broido and D. R. Kearns, J. Am. Chem. Soc. **104**, 5207 (1982).
[28] E. G. Richards *et al.*, Biopolymers **1**, 431 (1963).
[29] F. Vanzi *et al.*, Biophys. J. **89**, 1909 (2005).
[30] J. Brahms *et al.*, J. Mol. Biol. **15**, 467 (1966).
[31] T. G. Dewey and D. H. Turner, Biochemistry **18**, 5757 (1979); **19**, 1681 (1980).