

Sequence Dependent Rigidity of Single Stranded DNA

Noel L. Goddard,¹ Grégoire Bonnet,¹ Oleg Krichevsky,² and Albert Libchaber¹

¹Center for Studies in Physics and Biology, The Rockefeller University, New York, New York 10021

²Université Louis Pasteur, 67000 Strasbourg, France

(Received 27 March 2000)

Single stranded DNA (ssDNA) equilibrium dynamics are investigated using a fluorophore/quencher-labeled hairpin structure which thermally fluctuates between open and closed states. Temporal correlations of the fluorescence fluctuations are used to determine the energy barrier to conformational change. We find that ssDNA distortion is purely entropic for poly(*T*) but requires an additional enthalpy of $+0.5 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{base}^{-1}$ for poly(*A*), consistent with the disruption of base stacking. Such sequence dependent dynamics challenge the classical model of ssDNA as a completely flexible coil.

PACS numbers: 87.15.Kg, 87.64.Ni

Standard models of DNA structures are typically limited to local (“nearest-neighbor”) interactions and neglect the elastic contribution of the single stranded DNA (ssDNA) domains [1]. Hence, large ssDNAs are treated as highly flexible polymers with a persistence length of two bases [2–4]. To the contrary, we find that short ssDNAs exhibit strongly sequence dependent conformational dynamics, which are inconsistent with the classical model of ssDNA as a random coil with a sequence-invariant persistence length. By structurally constraining a ssDNA oligo to a hairpin-loop conformation and monitoring its thermal fluctuations, we are able to extract the entropic and enthalpic barriers associated with its deformation. We compare the conformational dynamics of two homogeneous sequences: poly(*A*) (polydeoxyadenosines) and poly(*T*) (polydeoxythymidines) (Fig. 1). Furthermore, by varying the lengths of these sequences, we can quantify the free energy penalty for deforming a ssDNA and find that poly(*A*) displays an enthalpic rigidity consistent with base stacking, whereas poly(*T*) displays a purely entropic elasticity. Finally, we introduce a single *C* (cytosine) “defect” into a poly(*A*) sequence to perturb base stacking and illustrate the sequence dependence of the hairpin dynamics.

We design short ssDNAs with ten complementary bases (five at each end) to form a hairpin loop through base pairing (Fig. 1). To detect the structure’s conformational changes, we utilize energy transfer between a fluorophore and a quencher chemically bound to the opposing ends of the hairpin stem. The fluorescence is quenched when fluorophore and quencher are within transfer range (\sim a few Å): an open hairpin fluoresces and a closed hairpin is quenched. In a previous paper, we showed that thermal fluctuations cause the hairpin to fluctuate between open and closed conformations, with two detectable limiting steps. In fact, there is an intermediate state in the closing associated with the fast zipping of the stem however undetectable with a time scale less than 100 ns [5–7]. These hairpins, developed as “molecular beacons” by Tyagi *et al.* [8], are ideal probes for studying structural dynamics due to their high fluorescence signal to noise

ratio (\sim 60) between open and closed states and their well-defined relationship to conformational changes. We use the constraints imposed by the loop upon hairpin closing to address the issue of sequence dependent rigidity of ssDNA.

In designing our hairpin sequences, the loops are varied from 8 to 30 bases in length, while the stems of all samples are identical in sequence with *TTGGG-3'* at one end and its complement *5'-CCCAA* at the other. Each hairpin is distinguished by its loop sequence, e.g., a hairpin with a loop of 21 *A* is named (*A*)₂₁. Samples are constructed from commercially available synthetic oligonucleotides (Midland Certified Reagents Company) tagged with a 4-[4-(dimethylamino)phenyl] azobenzoic acid or “dabcyl” at the 3' end and a primary amino group at the 5' end. A 6-carboxy rhodamine 6G “Rh6G” (molecular probes) is then coupled to the 5' primary amino group. The sample is purified by size exclusion filtration, reversed phase chromatography, and ethanol precipitation [5].

To investigate the conformational dynamics of a DNA hairpin, we carry out two independent measurements:

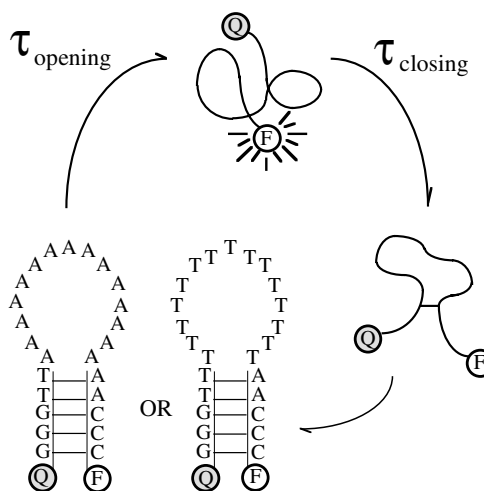


FIG. 1. The fluctuation of a hairpin loop between open and closed states where *F* is the fluorophore and *Q* the quencher. The two limiting time scales are τ_{closing} and τ_{opening} .

steady state fluorescence measurement of the structure's melting curve and temporal correlation of the fluorescence fluctuations around equilibrium. They are then combined to determine the time scales to conformational change.

We first compare hairpins of different loops by determining the thermal equilibrium between closed and open conformations (Fig. 2). The equilibrium constant $K(T)$ can be extracted from the average fluorescence. As the temperature rises the equilibrium shifts from a stable closed structure to a stable open one. We use the resulting melting curves (acquired from 96 to 10 °C) to evaluate the equilibrium constant for the following two states of the hairpin:

$$\text{Closed} \xrightleftharpoons[\tau_{\text{closing}}]{\tau_{\text{opening}}} \text{Open}, \quad \text{with } K(T) = \frac{\tau_{\text{opening}}(T)}{\tau_{\text{closing}}(T)}.$$

We fit the melting curve with a mass action law to determine the asymptotes which are then used to normalize the fluorescence [1]. $K(T)$ is then $f/(1-f)$, where f is the normalized fluorescence. The melting temperature T_m of the structure is defined as the temperature where closing and opening time scales are equal, i.e., $K(T_m) = 1$ or $f = 0.5$.

Figure 2 compares the melting curves for a series of poly(A) and poly(T) hairpins. A common trend in both poly(A) and poly(T) series is that the melting temperature of the structure decreases with increasing loop length. However, poly(A) hairpins melt at lower temperatures [e.g., T_m of $(A)_{21} = 39.2$ °C whereas $(T)_{21} = 48.4$ °C]. Moreover the spread of T_m with loop length is much larger for the poly(A)s than the poly(T)s [for loops from 12 to 30 bases, T_m drops from $54.1 \rightarrow 29.8$ °C in poly(A)s but only $58.1 \rightarrow 42.2$ °C for poly(T)s (Fig. 2 inset)]. It is our first evidence that the sequence of ssDNA can affect its structure. However we cannot distinguish between two possible scenarios (Fig. 1): either the loop destabilizes the base pairs of the stem (shorter τ_{opening}) or the loop constrains the closing of the stem (longer τ_{closing}). We thus need to measure the individual contributions of τ_{opening} and τ_{closing} in $K(T)$.

For this, we carry out fluorescence correlation spectroscopy (FCS) measurements of the thermal fluctuations around equilibrium and extract the characteristic decay time of these fluctuations through temporal correlation. FCS proves to be a powerful method for investigating the conformational fluctuations of numerous biological structures [9,10]. The collected emission signal is proportional to the number of fluorescing species in the sampling volume. We use a confocal setup to determine a sampling volume of the order $(1 \mu\text{m})^3$, and limit the hairpin concentration to five molecules/sampling volume to maximize the amplitude of fluctuations. Complete details of the setup were presented in [5].

The fluorescence fluctuations have two sources: the actual hairpin conformational fluctuation, and the diffusive contribution of molecules crossing the sampling volume.

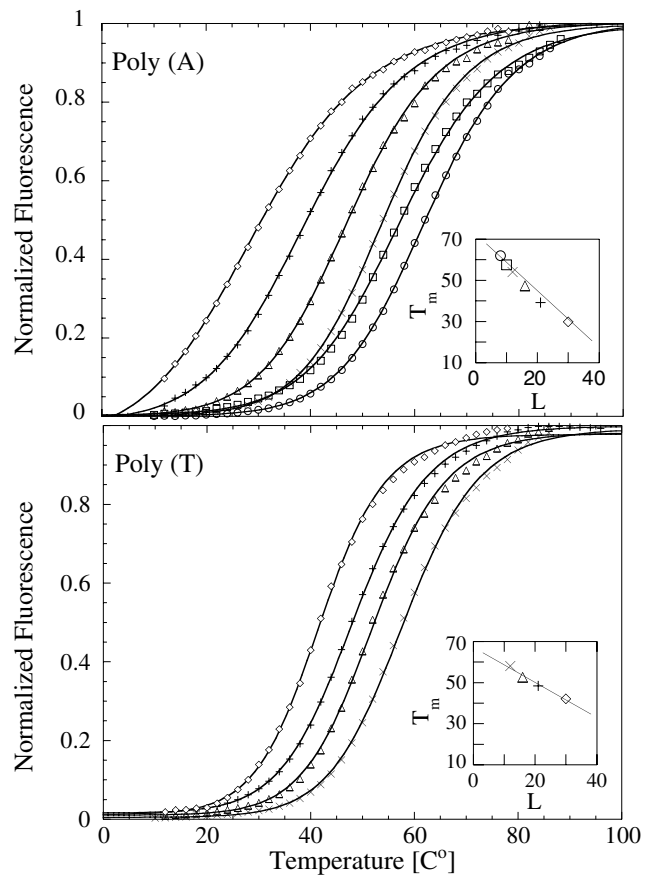


FIG. 2. Normalized melting curves of poly(A) and poly(T) (10 nm hairpin concentration in 0.25 M NaCl buffer). Loop lengths (number of bases) are described by the symbols, $\circ = 8$, $\square = 10$, $\times = 12$, $\triangle = 16$, $+ = 21$, and $\diamond = 30$. Data are fit with a single equilibrium mass action law. The inset displays the change in melting temperature with loop length for each series.

Hence the overall correlation function is the product of these two statistically independent contributions [5,11], $G_{\text{hairpin}}(t) = G_{\text{fluc}}(t)G_{\text{diff}}(t)$. To determine the hairpins diffusive contribution, a set of “control” hairpins is synthesized with identical sequences to the original hairpins, but lacking the dabcyI quencher: the correlation function acquired from the control hairpins is $G_{\text{control}}(t) = G_{\text{diff}}(t)$. By dividing $G_{\text{hairpin}}(t)$ by $G_{\text{control}}(t)$, we extract the τ_{fluc} where β is related to the equilibrium constant $K(T)$, and $\tau_{\text{fluc}}(T)$ is the relaxation time scale of the fluctuation mode. τ_{fluc} has been shown to be related to the closing and opening time scales, $1/\tau_{\text{fluc}} = 1/\tau_{\text{opening}} + 1/\tau_{\text{closing}}$ [11]. Using τ_{fluc} (from FCS measurement) and $K(T)$ from the melting curve) we can deduce the two time scales τ_{opening} and τ_{closing} .

We find that $\tau_{\text{opening}}(T)$ is independent of loop length or sequence (data not shown), thus the loop rigidity does not destabilize the base pairs of the stem, and does not explain the sequence dependence of the melting curves in Fig. 2.

Now let us focus on $\tau_{\text{closing}}(T)$. Figure 3 displays Arrhenius plots of the closing time scales, $\tau_{\text{closing}}(T)$, for a series of loop lengths (8–30 bases) of poly(A) and poly(T)

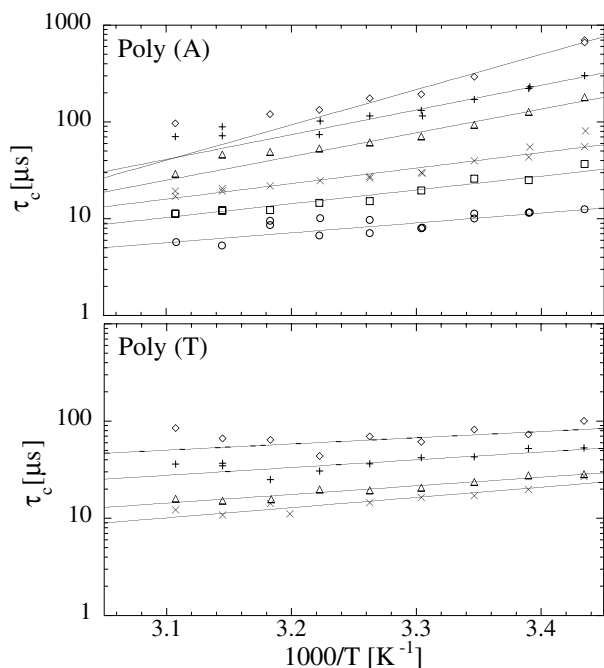


FIG. 3. Change in closing time scales with temperature for poly(A) and poly(T) series in 0.25 M NaCl buffer. Symbols represent the loop length in number of bases $\circ = 8$, $\square = 10$, $\times = 12$, $\triangle = 16$, $+$ = 21, and $\diamond = 30$ (error bars within the symbols).

hairpin loops in a 0.25 M NaCl buffer. The plots are fit by a single exponential $\tau_{\text{closing}}(T) = \tau_{\infty} \exp(+\Delta H_c/RT)$, which shows that hairpin closing involves a single energetic barrier. Fits are weighted by the error in $\tau_{\text{fluc}}(T)$ and $K(T)$. The slope of the Arrhenius plot yields the enthalpic barrier ΔH_c . At infinite temperatures, the entropic term dominates, but τ_{∞} cannot be measured reliably.

For all cases in Fig. 3 the closing time scale grows with increased loop length as we might expect that a larger molecule has a longer fluctuation time scale. By concentrating on just one series, $(T)_{12} - (T)_{30}$, we observe two key points: (1) the slopes ($\Delta H_c/R$) are practically identical, and (2) τ_{fluc} increases with increasing loop length. A different process is seen in poly(A): (1) $\Delta H_c/R$ increases with increasing loop length and (2) τ_{closing} for $T \rightarrow \infty$ (precise value for T_{∞} cannot be reasonably extrapolated, however we observe the overall trend), decreases with increased size. This sequence dependence is illustrated by directly comparing the closing times for equal loop lengths of poly(A) to poly(T). For example, we find that $\Delta H_c/R$ is much greater for A's than T's. ΔH_c of $(T)_{30}$ is ~ 2.98 kcal/mol, whereas for $(A)_{30} \sim 17.3$ kcal/mol [12]. The enthalpic barriers in Fig. 3 unravel an unexpected sequence dependence in ssDNA dynamics: Fig. 4 shows that the enthalpic barrier increases linearly with the loop length for poly(A)s ($+0.5$ kcal \cdot mol $^{-1}$ \cdot base $^{-1}$), while it is practically constant for different lengths of poly(T) (-0.1 kcal \cdot mol $^{-1}$ \cdot base $^{-1}$). Thus, the free energy of a poly(T) loop is mostly entropic whereas the free energy of

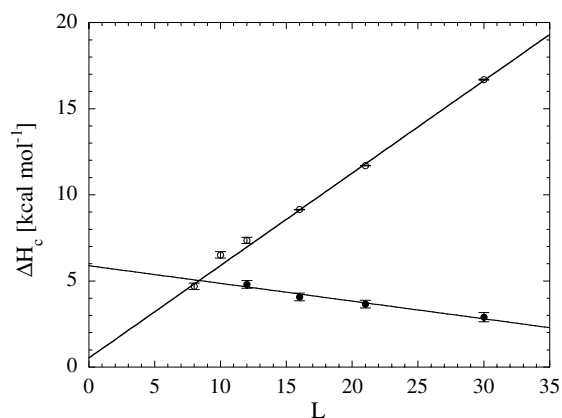


FIG. 4. Closing enthalpy vs loop lengths (number of bases) of \circ poly(A) and \bullet poly(T) series in 0.25 M NaCl buffer.

a poly(A) loop includes an additional enthalpic term. This conflicts with the standard model of ssDNA [2] as a highly flexible polymer. We propose that ssDNA molecular rigidity originates from the additional enthalpic contribution of base stacking.

Numerous experimental studies (reviewed in [13]) suggest that poly(A) has a significant helical structure stabilized through base stacking. In contrast poly(T) has been observed to have a little long range order [14]. Although a formal description of base stacking is still under debate, a recent work by Friedman and Honig [15] unifies the various contributions of electrostatic interactions, hydrophobic effects, surface tension, and dispersive interactions. Could base stacking lead to sequence dependent rigidity inferred from the hairpin closing time scale?

The energetic barrier for closing is complex. The ssDNA has to distort whichever conformation it is in to accommodate the formation of the first base pair in the stem. The neighboring base pairs in the stem are then zipped very fast (<100 ns). Thus the barrier accounts for this distortion and the nucleation of the first base pair. Since the stem is identical in all hairpins, the loop dependence for the closing barrier reflects the distortion energy of ssDNA loop sequence. We then understand the linearity of the enthalpic barrier for hairpin closing with respect to loop length. It reflects and measures the disruption of ssDNA base stacking. We get an estimate of ($+0.5$ kcal \cdot mol $^{-1}$ \cdot base $^{-1}$) comparable to predicted values [16,17] for base stacking. This sequence dependent dynamics of ssDNA is of crucial importance to understand and predict the integrity of DNA secondary structures (Fig. 2).

Let us recall that τ_{opening} is independent of the loop length of sequence. This implies that the base stacking of poly(A) has been disrupted upon closing. The closed loop is then a random coil with no influence on the stability of the stem.

To illustrate the importance of base stacking in the dynamics of single stranded domains, we show that a single

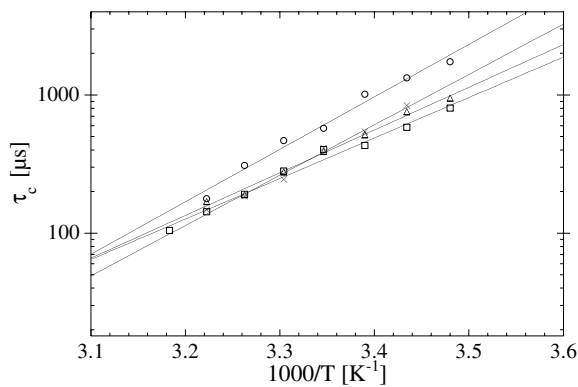


FIG. 5. Closing time scales of three poly(A) series with a single cytosine defect compared to $(A)_{21}$ (buffer: 0.1 M NaCl). Loop sequences are represented by the symbols $\circ = A_{21}$, $\square = A_{10}CA_{10}$, $\Delta = A_8CA_{12}$, and $\times = A_5CA_{15}$.

defect (a different base) in the poly(A) loop sequence significantly perturbs the stacking. $(A)_{21}$ loops containing a single cytosine (C) defect in three varying positions of the sequence were studied. Melting curves and FCS measurements were taken for the series in 0.1 M NaCl (lower salt concentration enhances the rigidity of ssDNA). The closing time scales for the series are shown in Fig. 5. We find that the closing time scales for these hairpins are shorter than for hairpins of $(A)_{21}$ with no defect. The defect induces a disruption in the base stacking, analogous to a mechanical weak point in a rigid structure. Thus depending on the sequence, a description of ssDNA as a semiflexible polymer can be more accurate than the freely jointed chain model [18].

Moreover, the enthalpic barrier for closing decreases when the defect is placed closer to the center of the loop [$\Delta H_c = 17.4$ kcal/mol for $(A)_{21}$, 16.7 kcal/mol for A_5CA_{15} , 14.2 kcal/mol for A_8CA_{12} , and 13.4 kcal/mol for $A_{10}CA_{10}$]. This is consistent with a hinge description of the mechanical fluctuations which facilitates the alignment and consecutive base pairing of the stem.

In conclusion, the measurements of steady state and fluctuations around thermal equilibrium provide a unique tool for investigating the thermodynamic and kinetic properties of molecular conformation change in ssDNA. We report for the first time the consequence of ssDNA stacking on its dynamics. Poly(T) loops display purely entropic rigidity in closing (consistent with the standard flexible coil model) whereas poly(A) loops display an enthalpic rigidity in closing consistent with base stacking. Additionally, we can estimate the enthalpy of destacking a single AA pair to $H_{\text{stacking}} = +0.5$ kcal/mol in 0.25 M NaCl. In

fact, the sequence dependence of ssDNA equilibrium dynamics is such that a single defect can significantly change the conformational fluctuations.

Our results provide strong evidence that ssDNA dynamics are so highly coupled to the sequence that it is debatable whether the persistence length is a valid physical parameter at all. We conclude that the enthalpic penalty for disrupting base stacking, as seen in poly(A), significantly modulates the stability of DNA secondary structures. These findings suggest that the standard model of DNA structures must be reconsidered to include these nontrivial effects of sequence dependent rigidity of the single stranded phase.

We acknowledge the Mathers Foundation and the Burrough-Wellcome Fund for financial support. We are grateful to H. Isambert for enlightening discussion.

- [1] C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry: the Behavior of Biological Macromolecules* (Freeman, New York, 1980), Chap. 23, Pt. III.
- [2] C. Bustamante, J. Marko, E. Siggia, and S. Smith, *Science* **265**, 1599 (1994).
- [3] S. B. Smith, Y. Cui, and C. Bustamante, *Science* **271**, 795 (1996).
- [4] B. Tinland, A. Pluen, J. Sturm, and G. Weill, *Macromolecules* **30**, 5763 (1997).
- [5] G. Bonnet, O. Krichevsky, and A. Libchaber, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8602 (1998).
- [6] D. Shore and R. L. Baldwin, *J. Mol. Biol.* **170**, 957 (1983).
- [7] W. H. Taylor and P. J. Hagerman, *J. Mol. Biol.* **212**, 363 (1990).
- [8] S. Tyagi and F. R. Kramer, *Nature Biotech.* **14**, 303 (1996).
- [9] S. Maiti, U. Haupts, and W. W. Webb, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11 753 (1997).
- [10] S. Wennmalm, L. Edman, and R. Rigler, *Proc. Natl. Acad. Sci. U.S.A.* **20**, 10 641 (1997).
- [11] D. Magde, E. Elson, and W. W. Webb, *Phys. Rev. Lett.* **29**, 705 (1972).
- [12] When $[\text{NaCl}] = 0.1$ M, $\Delta H_c/R$ is identical when compared to samples in 0.25 M NaCl for loops of 12 bases or less. For longer loops it increases by 15%.
- [13] J. Mills, E. Vacano, and P. J. Hagerman, *J. Mol. Biol.* **285**, 245 (1999).
- [14] M. Riley and B. Maling, *J. Mol. Biol.* **20**, 41 (1966).
- [15] R. A. Friedman and B. Honig, *Biophys. J.* **69**, 1528 (1995).
- [16] D. Pörschke, *Biophys. Chem.* **1**, 381 (1974).
- [17] T. G. Dewey and D. H. Turner, *Biochemistry* **18**, 5757 (1979).
- [18] S. V. Kuznetsov, Y. Shen, A. S. Benight, and A. Ansan (to be published).