A Semiflexible Polymer Model Applied to Loop Formation in DNA Hairpins

Serguei V. Kuznetsov,* Yiqing Shen,* Albert S. Benight,† and Anjum Ansari*‡
Departments of *Physics, †Chemistry, and ‡Bioengineering, University of Illinois at Chicago, Chicago, Illinois 60607 USA

ABSTRACT A statistical mechanical "zipper" model is applied to describe the equilibrium melting of short DNA hairpins with poly(dT) loops ranging from 4 to 12 bases in the loop. The free energy of loop formation is expressed in terms of the persistence length of the chain. This method provides a new measurement of the persistence length of single-stranded DNA, which is found to be ~1.4 nm for poly(dT) strands in 100 mM NaCl. The free energy of the hairpin relative to the random coil state is found to scale with the loop size with an apparent exponent of ≳7, much larger than the exponent of ~1.5–1.8 expected from considerations of loop entropy alone. This result indicates a strong dependence of the excess stability of the hairpins, from stacking interactions of the bases within the loop, on the size of the loop. We interpret this excess stability as arising from favorable hydrophobic interactions among the bases in tight loops and which diminish as the loops get larger. Free energy profiles along a generalized reaction coordinate are calculated from the equilibrium zipper model. The transition state for hairpin formation is identified as an ensemble of looped conformations with one basepair closing the loop, and with a lower enthalpy than the random coil state. The equilibrium model predicts apparent activation energy of ~−11 kcal/mol for the hairpin closing step, in remarkable agreement with the value obtained from kinetics measurements.

INTRODUCTION

Regions of DNA that are prone to strand separation and hairpin formation are believed to play a key role in recombination processes (Roth et al., 1992), in DNA transposition (Kennedy et al., 1998; Bhasin et al., 1999), and are also likely to affect the superhelicity of DNA that is implicated in global regulation of gene expression (Chen et al., 1998). Hairpin loops are an intrinsic and ubiquitous part of all native conformations of RNA molecules. Loop formation is most likely a first step in the folding of the RNA molecule (Alberts et al., 1989). Hairpin loops also serve as interaction sites for proteins (Yuan et al., 1979), and to bind to loops of other RNA molecules (Marino et al., 1995). Information on the stability of these hairpins and the energetic costs of loop formation are necessary for a complete physical understanding of these mechanisms. Hairpins also serve as model systems to test theories of polymer cyclization under conditions where the length of the polymer is only a few statistical segments long., i.e., not in the infinite chain limit (Winnik, 1986).

The stability of hairpin loops and the thermodynamics and kinetics of hairpin-to-coil transitions have been a subject of intense investigations dating back more than three decades (Elson et al., 1970; Scheffler et al., 1970; Gralla and Crothers, 1973; Uhlenbeck et al., 1973; Hilbers et al., 1985; Wemmer et al., 1985; Haasnoot et al., 1986; Senior et al., 1988; Zuker, 1989; Paner et al., 1990, 1992; Rentzeperis et al., 1993; Serra and Turner, 1995; Bonnet et al., 1998; Vallone et al., 1999). It is well known that the free energy of small loops in hairpins deviates from the simplest theory, which assumes only an entropic cost of bringing two ends of a Gaussian chain together (Elson et al., 1970; Scheffler et al., 1970; Uhlenbeck et al., 1973; Hilbers et al., 1985; Haasnoot et al., 1986; Zuker, 1989; Paner et al., 1990, 1992, 1996; Serra and Turner, 1995; Vallone et al., 1999). Previous attempts at estimating the free energy of small loops did not explicitly consider the dependence on the size of the loop. Baldwin and co-workers, in their pioneering studies of hairpin loop formation in oligomers with alternating base sequence . . .ATAT. . ., assumed a simplified partition function in which microstates with loop sizes larger than the minimum loop were not included (Elson et al., 1970; Scheffler et al., 1970). The free energy of loop formation for this minimum size loop was determined empirically from a fit of the melting temperature as a function of chain length. Benight and co-workers (Paner et al., 1990, 1992) used a more complete statistical mechanical description including all possible microstates. In their model, the free energy cost of forming loops of varying sizes were varied independently to get the best fit to the melting profiles (Paner et al., 1992, 1996).

In this paper, we use a semiflexible polymer model to describe the end-loop weighting function. Earlier attempts to describe the scaling of the molecular dimensions of single-stranded polynucleotides with increasing lengths under theta-solvent conditions (Eisenberg and Felsenfeld, 1967; Inners and Felsenfeld, 1970), as well as more recent attempts to describe the elastic properties of single-stranded DNA (ssDNA) using single-molecule force–extension measurements (Smith et al., 1996), have shown that a semiflexible polymer model works reasonably well for long ssDNA. At very low forces (≲10 pN), the force–extension curves on ssDNA exhibit deviations from a simple polymer model, which have been attributed to the "stickiness" of the bases and possible local hairpin formation (Maier et al., 2000). A theoretical model that includes the possibility of hairpin

Received for publication 9 March 2001 and in final form 24 July 2001
Address reprint requests to Anjum Ansari, Department of Physics, 845 W. Taylor St., University of Illinois at Chicago, Chicago, IL 60607. Tel.: 312-996-8735; Fax: 312-996-9016; E-mail: ansari@uic.edu.

0006-3495/01/11/2864/12 $2.00

formation reproduces the force–extension measurements of Maier et al. down to ~1 pN (Montanari and Mezard, 2001).

In kinetics measurements of hairpin formation in ssDNA as a function of the length L of the loop, we found that the characteristic times for forming hairpins scale as $L^{2\pm0.2}$ near the melting temperature of the hairpins for both poly(dT) and poly(dA) loops ranging from $N = 4$ to $N = 12$ bases (Ansari et al., 2001b; Shen et al., 2001). Our kinetics results show that, even at these rather small length scales of less than 10 bases, the probability of loop-closure in ssDNA is in remarkable agreement with what is expected from polymer theories with excluded volume, which predict a scaling exponent of 1.8 (DeGennes, 1979). Therefore, although a semiflexible polymer model may be too idealized a description of ssDNA at very short length scales of a few bases, it adequately describes the flexibility of ssDNA segments longer than a few persistence lengths, under conditions for which the self-interactions are not very strong.

Despite the reasonable success of semiflexible polymer models to describe the range of conformations accessible to ssDNA, and to describe the scaling of the loop closure times with the length of the loop, previous attempts to describe the thermodynamic stability of loops in hairpins and dumbbells as a function of the loop size have revealed that smaller loops are significantly more stable than what is expected from considerations of loop entropy alone (Hilbers et al., 1985; Haasnoot et al., 1986; Paner et al., 1992, 1996). This excess stability has been attributed to favorable interactions of the bases within the loop and between the loop and the stem. Here we use a semiflexible polymer model to describe the end-loop weighting function with an additional dependence on loop size to parameterize the enhanced stability of smaller loops. Our model provides a new approach to measuring a characteristic persistence length for ssDNA. One of the motivations of this work is to write down a statistical mechanical description in which the statistical weight of each microstate in the ensemble of partially melted hairpins with varying loop sizes can be enumerated. The model then allows us to calculate free energy profiles along an effective reaction coordinate as a function of temperature and to make predictions about the dynamics of hairpin formation. A key finding of the analysis is that the enthalpy of the transition state along this reaction coordinate, calculated from the equilibrium model, agrees quantitatively with the value obtained from kinetics measurements on these hairpins (Ansari et al., 2001a).

MATERIALS AND METHODS

Preparation and characterization of hairpin oligonucleotides

DNA oligomer strands were purchased from Oligos Etc. (Wilsonville, OR). All samples were HPLC purified. The oligomers were dialyzed against buffer (100 mM NaCl, 10 mM sodium phosphate, and 0.1 mM EDTA, pH 7.5). The strand concentrations (calculated using an extinction coefficient $\epsilon = 9631.25\ M^{-1}\ cm^{-1}$ per base at 266 nm) were about 80 μM, well within the concentration range of up to 500–800 μM possible for unimolecular hairpin formation (Senior et al., 1988; Rentzeperis et al., 1993). Some samples had 2-aminopurine (2AP), a fluorescent analog of the adenine base, substituted at various sites along the hairpin stem (see Fig. 1).

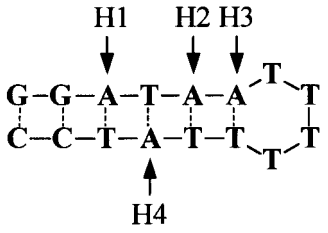

FIGURE 1 One of the hairpins used in this study in which 2AP is substituted instead of A at the sites indicated. H1: substitution at site 1 only; H2: substitution at site 2 only, etc. H0 in the text refers to the hairpin with no substitutions.

Optical melting profiles from static absorbance measurements

Optical melting curves were obtained by measuring changes in the absorbance at 266 nm and at 330 nm (where 2AP absorbs) as a function of temperature. Absorbance curves were acquired using a Hewlett-Packard 8452 (Palo Alto, CA) diode array single-beam spectrophotometer equipped with a thermostated temperature controller. Samples were equilibrated for 20 min at the beginning temperature of each heating–cooling cycle and the rate of temperature change was ~28°C/hour. For each sample, at least three complete (heating–cooling) melting curves were collected and averaged together. A buffer baseline was subtracted from the sample absorbance versus temperature curves to yield the corrected absorbance A(T) for each sample.

Equilibrium zipper model for hairpin melting

The statistical mechanical algorithm used to describe the melting curves is based on the simplest one-dimensional Ising model that allows for only two states for each basepair, broken or intact (Poland and Scheraga, 1970). Benight and co-workers have modified the model to include loop formation and nearest-neighbor (n-n) sequence dependence in the stacking interactions (Wartell and Benight, 1985; Paner et al., 1990, 1992). Their formulation requires four parameters to describe the partition function: s_i, the statistical weight for each basepair; σ, the cooperativity parameter; $f(m)$, the statistical weight for forming an internal loop of size m; and $w_{loop}(n)$, the end-loop weighting function for a loop consisting of n bases. Here we use the zipper version of the model, in which partially melted microstates corresponding to internal loops inside the stem are not allowed, i.e., $f(m) = 0$. The number of states reduces from 2^{N_s} to $N_s(N_s + 1)/2 + 1$ where N_s is the number of basepairs in the stem.

The statistical weight corresponding to each basepair formation, s_i, depends upon the type of basepair i (A-T or G-C) and interactions with its neighbors, and includes contributions to a basepair's stability from hydrogen bonding as well as stacking interactions with n-n base pairs on either side,

$$s_i = \exp[-\Delta G_i/RT], \tag{1}$$

where

$$\Delta G_i = \Delta H_i - T\Delta S_i + \frac{\delta G_{i-1,i} + \delta G_{i,i+1}}{2}, \tag{2}$$

TABLE 1 Nearest-neighbor stacking interactions

MN Doublet 5′-3′ (M_pN)	$\delta G_{i,i+1}$* (cal mol^{-1})	MN Doublet 5′-3′ (M_pN)	$\delta G_{i,i+1}$* (cal mol^{-1})	MN Doublet 5′-3′ (M_pN)	$\delta G_{i,i+1}$† (cal mol^{-1})
G_pG	178	C_pG	562	A_pX	369 ± 167
G_pA	44	G_pC	−682	G_pX	1398 ± 352
A_pT	−179	A_pA	−54	T_pX	1275 ± 397
T_pA	334	A_pG	149	X_pA	−950 ± 476
A_pC	−272	C_pA	45	X_pT	−666 ± 419

*Averaged over the two sets of values reported in Wartell and Benight (1985).
†Varied as parameters in the model described in the text. The values of these parameters are poorly determined from one random search to another. The error bars are estimated from the deviations in the parameters obtained from 20 independent random searches.

ΔH_i and ΔS_i are the enthalpy change and the entropy change, respectively, associated with basepair formation. The index i increases from the free end of the stem to the loop with $i = 1$ corresponding to the first basepair at the free end. $\delta G_{i,i\pm1}$ terms in Eq. 2 are the deviations from the average n-n stacking free-energy for each type of n-n stack and their values are taken from the work of Benight and co-workers (Wartell and Benight, 1985; Paner et al., 1990) and summarized in Table 1. ΔS_i for each basepair is assumed identical and assigned a value of −25.2 cal mol^{-1} K^{-1} (Klump and Ackermann, 1971). $\Delta H_i = T_i \Delta S_i$, with $T_i = T_{AT}$ or T_{GC}, the average melting temperatures of A-T or G-C basepairs. In 100 mM NaCl, the solution conditions of our sample, $T_{AT} = 64.3$°C and $T_{GC} = 107.4$°C (Frank-Kamenetskii, 1971).

The cooperativity parameter, σ, depends upon the specific type of basepairs at a junction between an intact and broken basepair. For example, if basepair i is broken and basepair $i + 1$ is intact, the cooperativity parameter assigned to the junction has the form

$$\sigma_{i,i+1} = \langle\sigma\rangle^{1/2} \exp(\delta G_{i,i+1}/2RT), \tag{3}$$

where $\langle\sigma\rangle$ is the average of the 10 different stacking interactions and is assigned a value of 4.5×10^{-5} (Wartell and Benight, 1985). Examples of some partially melted microstates of a hairpin with six basepairs in the stem and four bases in the loop are shown in Fig. 2. When the fully melted random coil is assumed to be the free energy reference state, the statistical weights z_j of the microstates in Fig. 2 are

$$z_a = \sigma_{end}\left(\prod_{i=1}^{Ns} s_i\right) w_{loop}(N), \tag{4a}$$

$$z_b = \sigma_{1,2}\left(\prod_{i=2}^{Ns} s_i\right) w_{loop}(N), \tag{4b}$$

$$z_c = \sigma_{end}\left(\prod_{i=1}^{Ns-2} s_i\right) w_{loop}(N+4), \tag{4c}$$

where N is the number of bases in the loop of a fully intact hairpin and $\sigma_{end} = \langle\sigma\rangle^{1/2}$ is the cooperativity parameter assigned to the free end of the stem of an intact hairpin. The partition function $Q(T)$ is obtained by summing the statistical weights of all microstates $\{j\} = 1, \ldots, N_s(N_s + 1)/2 + 1$. The probability that the ith basepair is intact is

$$p_i = \sum_{\{j\}} z_j(\text{intact})/Q(T), \tag{5}$$

where the sum is over the statistical weights z_j(intact) of every microstate $\{j\}$ for which the basepair i is intact. If we define n_j as the number of intact basepairs in the jth microstate, the fraction of intact basepairs in that microstate is (n_j/N_s). The fraction of intact basepairs summed over all the microstates, $\theta_I(T)$, is

$$\theta_I(T) = \frac{1}{N_s} \sum_{\{j\}} n_j (z_j/Q(T)), \tag{6}$$

where z_j is the statistical weight of the jth microstate.

The end-loop weighting function

The statistical weights of the loop $w_{loop}(n)$ are written as

$$w_{loop}(n) = \left(\frac{3}{2\pi b^2}\right)^{3/2} V_r g(n) \sigma_{loop}(n), \tag{7}$$

where n is the number of bases in the loop, $b = 2P$ is the statistical segment length (also known as the Kuhn's length), V_r is a characteristic reaction volume within which the bases at the two ends of the loop can form hydrogen bonds, $g(n)$ is proportional to the probability of forming a loop with n bases in the loop, and $\sigma_{loop}(n)$ parameterizes the stabilizing interactions of the bases within the loop and between the loop and the stem. We use a wormlike chain description for $g(n)$ as derived by Yamakawa and Stockmayer (1972),

$$g(n) = \frac{1}{N_b^{3/2}}\left[1 - \frac{5}{8N_b} - \frac{79}{640N_b^2}\right] \quad N_b > 1, \tag{8a}$$

$$g(n) = \left(\frac{C_0}{N_b}\right)\exp\left(-\frac{A}{N_b}\right)[1 + C_1 N_b] \quad N_b \le 1, \tag{8b}$$

where $N_b = (n + 1)h/b$ is the number of statistical segments in a loop with n bases, and h is the internucleotide distance. $C_0 = 1.51 \times 10^3$, $C_1 = -0.81242$, and $A = 7.0266$ are constants derived by Yamakawa and Stockmayer to give a smooth function over all values of n. The reaction

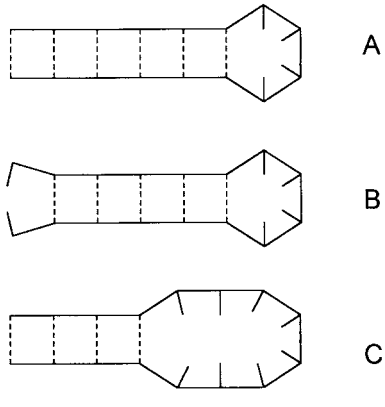

FIGURE 2 Schematic representation of some of the microstates of the zipper model described in the text. The corresponding statistical weights are given in Eqs. 4a—4c.

volume V_r ($= 4/3\pi r^3$) in Eq. 7 is calculated with $r = 1$ nm. The internucleotide distance is assigned a value of $h = 0.52$ nm (Mills et al., 1999).

The main features of $g(n)$, and which are important for our analysis, are that $g(n)$ scales as $n^{-1.5}$ in the infinitely long chain limit (Jacobsen and Stockmayer, 1950; DeGennes, 1979) and diminishes sharply for loop sizes smaller than the Kuhn's statistical segment length (Yamakawa and Stockmayer, 1972; Shore et al., 1981; Hagerman, 1985; Camacho and Thirumalai, 1995). Therefore, for loops smaller than the optimal size, given by $(n + 1)h \lesssim b$, the hairpins are substantially destabilized because basepairs closest to the loop have to break to accommodate the loop (Haasnoot et al., 1986; Paner et al., 1992, 1996).

Stabilizing interactions of the bases in the loop

It is well known that the free energy of DNA hairpins is not simply the sum of the free energy of the stem and the free energy from the entropic cost of forming a loop, and that favorable stacking interactions among the bases in the loop and between the loop and the stem contribute to an additional favorable free energy term in the stability of these hairpins (Hilbers et al., 1985; Haasnoot et al., 1986; Paner et al., 1992, 1996). What is less well appreciated is that the extent of this excess stability is strongly dependent on the size of the loop and diminishes gradually as the size of the loop increases (Paner et al., 1992, 1996). In our description, we have used two different functional forms for $\sigma_{loop}(n)$ that incorporate two important features as dictated by the experimental data. One is the decrease in $\sigma_{loop}(n)$ as the size of the loop increases and the other is the assumption that $\sigma_{loop}(n)$ approaches the cooperativity parameter assigned to the free end of the stem ($\approx\langle\sigma\rangle^{1/2}$) in the limit of very large loops. The two functional forms we tried are

$$\sigma_{loop}(n) = \langle\sigma\rangle^{1/2} + \frac{C_{loop}}{N_b^{\gamma}} \tag{9}$$

and

$$\sigma_{loop}(n) = \langle\sigma\rangle^{1/2}\exp\left[-\frac{D_{loop}/N_b^{\kappa}}{RT}\right]. \tag{10}$$

The free energy contribution of the loop to the excess stability of the hairpin, for the most optimal loop ($N_b \approx 1$) is $-RT\ln(C_{loop})$ in Eq. 9 and D_{loop} in Eq. 10. The dependence of $\sigma_{loop}(n)$ on the size of the loop is parameterized by γ in Eq. 9 or κ in Eq. 10.

Analysis of equilibrium melting profiles

The absorbance melting profiles at 266 nm can be expressed as $A(T) = \theta(T)[A_U(T) - A_L(T)] + A_L(T)$. Here, $\theta(T)$ is the net fraction of broken basepairs, and $A_U(T)$ and $A_L(T)$ are the limiting baselines at high and low temperatures, respectively. $\theta(T)$ is calculated from Eq. 6 as $\theta(T) = 1 - \theta_1(T)$; the limiting baselines are parameterized as straight lines for each absorbance profile. Similarly, the absorbance melting profiles at 330 nm can be expressed as $A(T) = (1 - p_i(T))[A_U(T) - A_L(T)] + A_L(T)$, where $(1 - p_i(T))$, the probability that the ith basepair is broken, is calculated from Eq. 5. The parameters that are varied to get the best fit to the absorbance profiles are the persistence length P, the loop-stabilizing parameters C_{loop} (or D_{loop}), and γ (or κ).

To ensure that the parameters obtained from the fits are robust and represent a global minimum in the residual sum-of-squares, a simulated annealing procedure was used to sample the parameter space to minimize the residual sum-of-squares, starting from 20 independent randomly chosen sets of parameters for each of the two functional forms for $\sigma_{loop}(n)$ (Metropolis et al., 1953; Ansari et al., 1994). Each set of parameters obtained by this procedure was then used as starting values for minimization using a Marquardt–Levenberg algorithm (Marquardt, 1963). The parameters reported in the text correspond to the best fit from 20 independent random searches, and the error bars in the parameters are standard deviations obtained from a subset of these independent searches for which the residual sum-of-squares increased by less than a factor of 3.

RESULTS AND DISCUSSION

The unwinding of hairpin molecules formed from the self-complementary oligomer 5′-GGATAA(T_4)TTATCC-3′ (6 basepairs in the stem, see Fig. 1) and 5′-CGGATAA(T_N)TTATCCG-3′ (7 basepairs in the stem and for $N = 4$, 8, and 12) were investigated using the equilibrium zipper model as described in the Methods section. Four analogs of the 6-basepair hairpin (Fig. 1) had 2AP substituted at four different sites along the stem to probe both local and global melting (Nordlund et al., 1989; Bloom et al., 1993). Melting profiles were monitored using absorbance changes at 266 nm that monitor changes in the helical state of the molecule as probed by the normal nucleotides, and absorbance changes at 330 nm that monitor changes in the helical state as probed by the 2AP nucleotide.

Evaluation of loop parameters

Theoretical melting profiles calculated using the equilibrium zipper model were fit simultaneously to the experimentally obtained absorbance profiles for the hairpins with 7 basepair stems and varying loop sizes (Fig. 3 *A*). The model was also used to fit the melting temperatures of a different set of hairpins, 5′-ATCCTA(T_N)TAGGAT-3′ with N ranging from 2 to 7 (in 200 mM Na^+), taken from the work of Hilbers et al. (1985). Our model fails to describe their melting temperatures for sequences with $N < 2$ (Fig. 3 *B*). In these hairpins, one or two basepairs closest to T_N are known to break apart to maintain a loop of about 4 bases (Haasnoot et al., 1986). The parameters describing the end-loop weighting function for both sets of hairpins, along with the variation in the parameters from independent random searches, for each parameterization of $\sigma_{loop}(n)$ (Eq. 9 or 10), are summarized in Table 2. We also tried different sets of n–n stacking interactions parameters, summarized by Blake and Delcourt (1998). The final parameters for the end-loop weighting function were within the error bars reported in Table 2. A recent comparison of n–n interaction parameters obtained from different groups shows that the different sets of interaction parameters are, in fact, in good agreement when corrected for the variations in salt concentration (SantaLucia, 1998).

There are three notable results from our analysis. First, the persistence length parameter is well determined for both sets of hairpins and over the various random searches in parameter space, and is not very sensitive to the exact functional form for $\sigma_{loop}(n)$ (Eq. 9 or Eq. 10). The fits yield a value of $P \approx 1.3$–1.5 nm for poly(dT) strands in 100–200 mM NaCl. Second, for both sets of hairpins, there is a

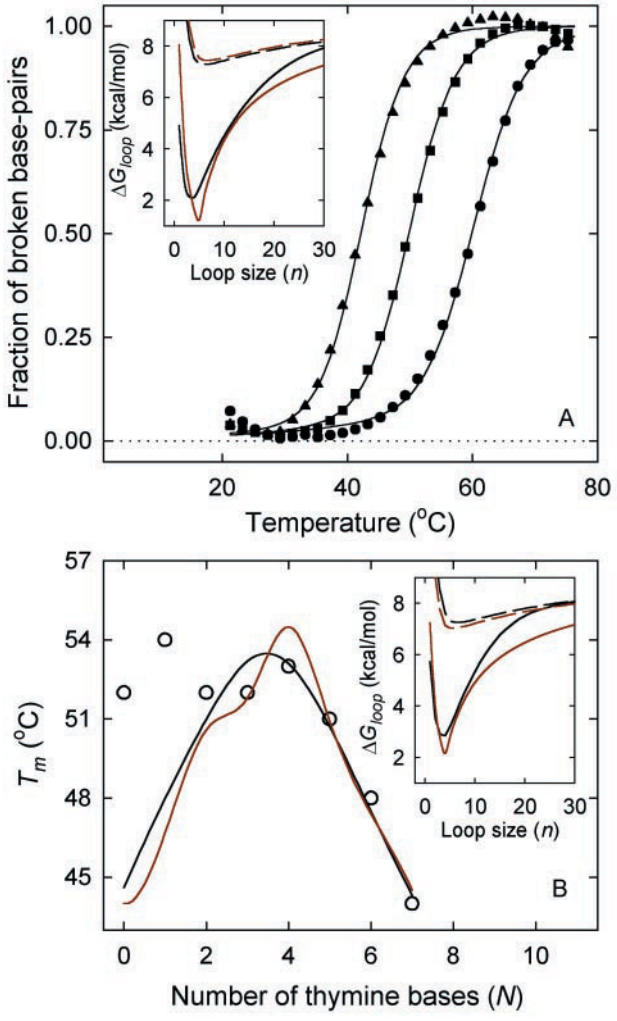

FIGURE 3 (*A*) Fits to the equilibrium melting profiles of 5′-CGGATA-A(T_N)TTATCCG-3′ using the model described in the text; ●, $N = 4$; ■, $N = 8$; ▲, $N = 12$. The symbols are normalized absorbance at 266 nm $[A(T) - A_L(T)]/[(A_U(T) - A_L(T)]$ and represent the average of 20 data points obtained with a temperature resolution of 0.1°C; the lines are the fraction of broken basepairs $\theta(T)$ and are calculated using Eq. 6. (*Inset*) The free energy of forming a loop closed by an A-T basepair, $\Delta G_{loop}(n) = -RT \ln(\sigma_{end}\, s w_{loop}(n))$, versus the number of bases in the loop. Parameters used to calculate $w_{loop}(n)$ are with $\sigma_{loop}(n)$ defined as in Eq. 9 (*black curve*) and as in Eq. 10 (*red curve*); the dashed lines are corresponding curves with $\sigma_{loop}(n) = \langle\sigma\rangle^{1/2}$. (*B*) Fit to the melting temperatures (T_m) as a function of the number of thymine bases (N) in the sequence 5′-ATCCTA(T_N)TAG-GAT-3′. The open circles are data points digitized from Hilbers et al. (1985). The fit was obtained using Eq. 9 (*black curve*) and Eq. 10 (*red curve*) for values of N ranging from 2 to 7. The theoretical curves are calculated for all values of N. These measurements were done in 200 mM Na^+, and the corresponding values of $T_{AT} = 69.61$°C and $T_{GC} = 110.53$°C are used to calculate the theoretical curves (Frank-Kamenetskii, 1971). (*Inset*) Same as the inset in *Panel A* calculated for the hairpin of Hilbers et al. (1985).

significant dependence of the stabilizing interactions in the loop on the size of the loop. The apparent exponent α that describes the scaling of the free energy contribution from the end-loop, $\Delta G_{loop}(n)$, with the length of the loop, $L \approx (n + 1)h$, is $\alpha \approx \gamma + 1.5 \gtrsim 7$ (see Eqs. 7–9), much larger than the exponent of 1.5–1.8 expected from purely entropic losses upon loop formation (see Fig. 3, *insets*). And third, our assumption that, for very large loops $\sigma_{loop}(n)$ approaches $\langle\sigma\rangle^{1/2}$ is borne out by theoretical modeling of the melting temperatures of hairpins with larger loops $N \approx$ 12–30 (Shen et al. 2001). The primary difference between the two representations of $\sigma_{loop}(n)$ is in how the free energy cost of loop formation approaches the large loops limit (see Fig. 3, *insets*).

An earlier study attributed much of the contribution of the increased stability of hairpins to stacking interactions between the bases in the loop and in the stem (Hilbers et al., 1985; Haasnoot et al., 1986). However, this simple picture fails to provide an adequate explanation for the strong dependence of the stabilizing free energy on the size of the loop, with loops of $N = 4$ stabilized by free energies $\gtrsim$3 kcal/mol in comparison with loops of $N = 12$ and by $\gtrsim$5 kcal/mol in comparison with loops of $N = 20$ (see Fig. 3, *insets*). These values are in good agreement with the free energy of loop formation estimated by Benight and co-workers for various size loops from their studies of the melting of DNA dumbbells (Paner et al., 1992, 1996). In contrast, the entropic cost of loop formation accounts for $<$0.5 kcal/mol difference between $N = 4$ and $N = 20$. Haasnoot et al. (1986) had concluded from their thermodynamic analysis of the melting of hairpins with varying loop sizes that the enthalpy of the proposed loop–stem stacking interactions was independent of the number of bases in the loop. However, their analysis did not include any temperature dependence to the enthalpy changes, and their constant enthalpy results may well be accidental.

Benight and co-workers have argued that a significant contribution to this excess stability is from the strong hydrophobic interactions of the bases within the loop, resulting in a negative free energy contribution that stabilizes hairpins with tighter loops in which water is excluded from inside the loops (Vallone and Benight, 1999). Larger loops are expected to have weaker intraloop hydrophobic interactions and a correspondingly higher free energy. The experiment that supports this hypothesis is the increased stability (by ~1 kcal/mol) of hairpins when the four thymine bases in the loop are replaced by a synthetic base 5-nitroindole (5-NI) (Vallone and Benight, 1999), and which are reported to be more hydrophobic than the natural bases (Loakes et al., 1997). If the decrease in stability of hairpins with increasing loop size is indeed attributable to the diminishing hydrophobic interactions within the loop, then the stability of hairpins with loops of varying lengths of 5-NI bases should show an even stronger dependence on the size of the loop than hairpins with poly(dT) bases, and changing the solvent to a less polar solvent should diminish the dependence on the loop size. Both these predictions can be tested and will be important to establish the physical origin of the increased stability of smaller loops.

TABLE 2 Parameters describing the end-loop weighting function

5′-stem of hairpin	P (nm)*	$-RT\ln(C_{loop})$* (kcal/mol)	γ*	P (nm)†	D_{loop}† (kcal/mol)	κ†
5′-CGGATAA-3′	1.4 ± 0.2‡	−2.2§ ± 0.4	5.4 ± 0.2	1.5 ± 0.3	−6.3 ± 3	1.11 ± 0.1
5′-ATCCTA-3′	1.5 ± 0.1	−1.3 ± 0.1	6 ± 0.8	1.3 ± 0.2	−5 ± 0.3	1 ± 0.4

*σ_{loop} is defined as in Eq. 9.
†σ_{loop} is defined as in Eq. 10.
‡The values of the parameters are the best out of 20 independent random searches in parameter space. The error bars are estimated from the deviations in the parameters from a subset of the independent searches for which the residual sum-of-squares increase by less than a factor of 3.
§T = 25°C for all values in this column.

It should be noted that the excess stability of hairpins with smaller loops is also observed in RNA hairpins in 10 mM Na^+, with an increase in the melting temperature of the hairpins by almost 15°C and a decrease in the free energy by ~2.5–3 kcal/mol when the loop size is decreased from N = 9 to N = 4, with $N \approx 3$–4 as the most optimal loop size (Groebe and Uhlenbeck, 1988). Increasing the salt concentration to 1 M NaCl shifts the most optimal loop size in RNA hairpins to $N \approx 4$–7, and appears to diminish the loop-size dependence (Groebe and Uhlenbeck, 1988; Serra et al., 1997).

Persistence length of ssDNA

Our measurements of the stability of hairpins as a function of loop size yields a persistence length of $P \approx 1.4$ nm for ssDNA with poly(dT) strands in 100 mM NaCl. Previous estimates of P from a number of different experiments range widely from 0.75 to 7.5 nm (Eisenberg and Felsenfeld, 1967; Inners and Felsenfeld, 1970; Smith et al., 1996; Tinland et al., 1997; Rivetti et al., 1998; Mills et al., 1999). Experiments based on random coil dimensions from light scattering and sedimentation measurements under theta-solvent conditions for long single-stranded poly(rU) chains yield a value of $P \approx 1.4$ nm (Eisenberg and Felsenfeld, 1967; Inners and Felsenfeld, 1970). Bustamante and coworkers obtain $P \approx 0.75$ nm from measurements of the elastic response of ssDNA molecules using optical tweezers (Smith et al., 1996), and $P \approx 1.3$–2.8 nm from measurements of end-to-end distances of DNA duplexes with poly(dT) segments inserted in the middle (Rivetti et al., 1998). Measurements of the self-diffusion coefficient of single-stranded chains from fluorescence recovery after photobleaching yield $P \approx 1.3$ nm (Tinland et al., 1997). The decay of transient electric birefringence for duplexes with single-stranded segments yields $P \approx 2.5$–3.5 nm for poly(dT) and $P \gtrsim 7.5$ nm for poly(dA) (Mills et al., 1999). Therefore, the basic issue of the persistence lengths of single-stranded polynucleotides is far from settled. The deviations from different experiments reflect, to some extent, the variation of persistence length values on ionic strength and sequence. They may also, in large part, be due to the tendency of the hydrophobic bases in single-stranded polynucleotides to stick together, resulting in deviations from a random coil model assumed in the analysis of many of these measurements. Deviations from a semiflexible polymer model have been observed in some force–extension measurements on ssDNA, especially in the limit of low ($\lesssim$10 pN) forces (Maier et al., 2000). Montanari and Mezard (2001) have modeled the force–extension measurements of Maier et al. by including the possibility of interactions and hairpin formation; they have been able to describe the experimental curves down to ~1 pN with a persistence length value about three bases long, which corresponds to $P \approx 1.6$ nm.

In our theoretical modeling of the equilibrium melting profiles, we exploit the secondary-structure formation in ssDNA by explicitly including all possible microstates of the system. Furthermore, the melting temperatures of hairpins, especially in the vicinity of the optimal loops, are very sensitive to the free energy cost of forming loops, which are predicted to go up sharply for $L < 2P$. Note that this behavior is expected to hold for any description of semiflexible polymers and is not specific to the model used in our calculations (Yamakawa and Stockmayer, 1972; Shore et al., 1981; Hagerman, 1985; Camacho and Thirumalai, 1995). Our persistence length values are in very good agreement with the light-scattering and sedimentation measurements that were done in theta-solvent conditions (Eisenberg and Felsenfeld, 1967; Inners and Felsenfeld, 1970), with the measurements of the self-diffusion coefficient (Tinland et al., 1997), and with the analysis of the force–extension measurements, which include hairpin formation at low forces (Maier et al., 2000; Montanari and Mezard, 2001). Our values of the persistence length are also consistent with a simple estimate. The most optimal size of poly(dT) loops in DNA hairpins is found to be around four bases, which corresponds to a loop length $L \approx (n + 1)h \approx 2.6$ nm. Because the optimal loop size is $L \approx 2P$, we can estimate $P \approx 1.3$ nm, in close agreement with the values obtained from analyzing the equilibrium melting profiles.

Recently, we reported a result of surprising simplicity on the dynamics of ssDNA; the characteristic time for loop formation that nucleates the zipping of hairpins scales with the length L of the loop as $L^{2\pm0.2}$ for hairpins with poly(dT) or poly(dA) loops, near the melting temperatures of the

hairpins (Ansari et al., 2001b; Shen et al., 2001). This result is in close agreement with the predictions of the theories of dynamics of semiflexible polymers with excluded volume considerations, for which the scaling exponent is expected to be ~1.8. The hairpins used in the dynamical studies were identical to the hairpins of Fig. 3 *A*. Because the analysis of the melting curves of hairpins is dominated by the data near the melting temperatures, the persistence length values that we report are most characteristic of the semiflexible polymer properties of ssDNA under conditions for which the self-interactions are not very strong. These results suggest that, when the interactions are reduced, the ssDNA does behave like a random polymer even at these short length scales. Therefore, if force–extension experiments on ssDNA were done under conditions of high temperatures or less polar solvents, the results may be more consistently reproduced by the elastic response expected for simple polymers.

Local and global melting of hairpins

A stringent test of the zipper model comes from the ability of the model to describe both the average fraction of broken basepairs (monitored using absorbance changes at 266 nm) and the probability that any particular site is broken (monitored using absorbance changes of 2AP at 330 nm). To describe the melting curves of the 2AP-substituted hairpins, six additional parameters describing the destabilizing effects of the 2AP modification need to be included: five parameters corresponding to changes in stacking interactions from GpX, XpT, TpX, XpA, ApX (X is used to indicate 2AP), and one parameter corresponding to the melting temperature T_{XT}. The values of $P = 1.4$ nm and $\gamma = 5.4$ (Eq. 9) are fixed from the fits to the data shown in Fig. 3 *A*. The theoretical curves corresponding to the best fit to the 2AP-substituted hairpins (Fig. 4) yield a value of $T_{XT} = 45.3 \pm 5.6$°C and $RT\ln(C_{loop}) = 1.2 \pm 0.1$ kcal/mol at $T = 25$°C. The parameters describing the modifications in the stacking interactions from 2AP insertion are summarized in Table 1. The key result here is that the melting profiles obtained at 266 and 330 nm overlap for each 2AP-substituted analog (Fig. 5), confirming that the hairpin loses its helicity in a two-state cooperative transition.

Free energy profile versus an effective reaction coordinate

We use the results of the equilibrium model to calculate free energy profiles versus an effective one-dimensional reaction coordinate that can be used to model the dynamics of hairpin formation (Ansari et al., 2001a; Shen et al., 2001). In the theoretical description of protein folding kinetics, the multidimensional potential energy surface is collapsed into a free energy profile as a function of a single reaction coordinate, an effective order parameter, defining the degree of "correctness" for any configuration (Bryngelson and Wolynes, 1987; Zwanzig, 1995). Analogous to the protein

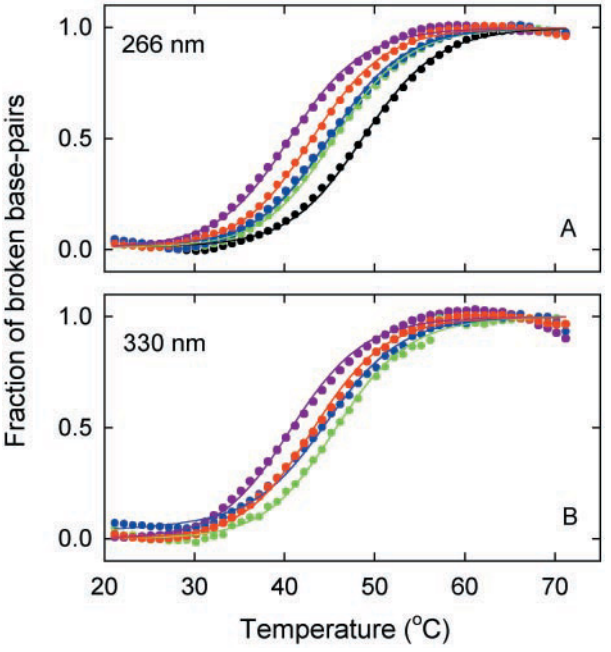

FIGURE 4 Fits to the equilibrium melting profiles of 5′-GGATAA(T_4)TTATCC-3′ and its 2AP analogs (H0–H4) using the model described in the text: *black*, H0; *purple*, H1; *green*, H2; *blue*, H3; *red*, H4. The lines are fits to the data with $P = 1.4$ nm, $\gamma = 5.4$, $RT\ln(C_{loop}) = 1.2$ kcal/mol at $T = 25$°C (see Eq. 9), and $T_{XT} = 45.3$°C. (*A*) *Symbols*, normalized absorbance at 266 nm $[A(T) - A_L(T)]/[(A_U(T) - A_L(T)]$; *lines*, the fraction of broken basepairs $\theta(T)$, calculated using Eq. 6. (*B*) *Symbols*, normalized absorbance at 330 nm; *lines*, the probability that the 2AP site is broken, calculated using Eq. 5.

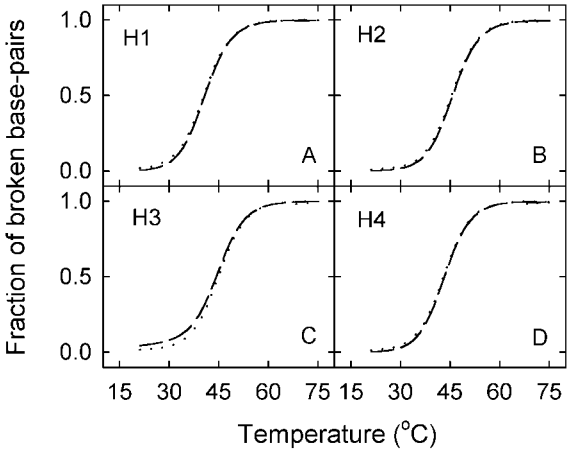

FIGURE 5 Equilibrium melting profiles for (*A*) H1, (*B*) H2, (*C*) H3, and (*D*) H4 calculated using the same parameters as in Fig. 4. *Dotted lines*, fraction of broken basepairs (calculated using Eq. 6) fits the normalized absorbance at 266 nm (see Fig. 4 *A*). *Dashed lines*, probability that the 2AP site is broken (calculated using Eq. 5) fits the normalized absorbance at 330 nm (see Fig. 4 *B*).

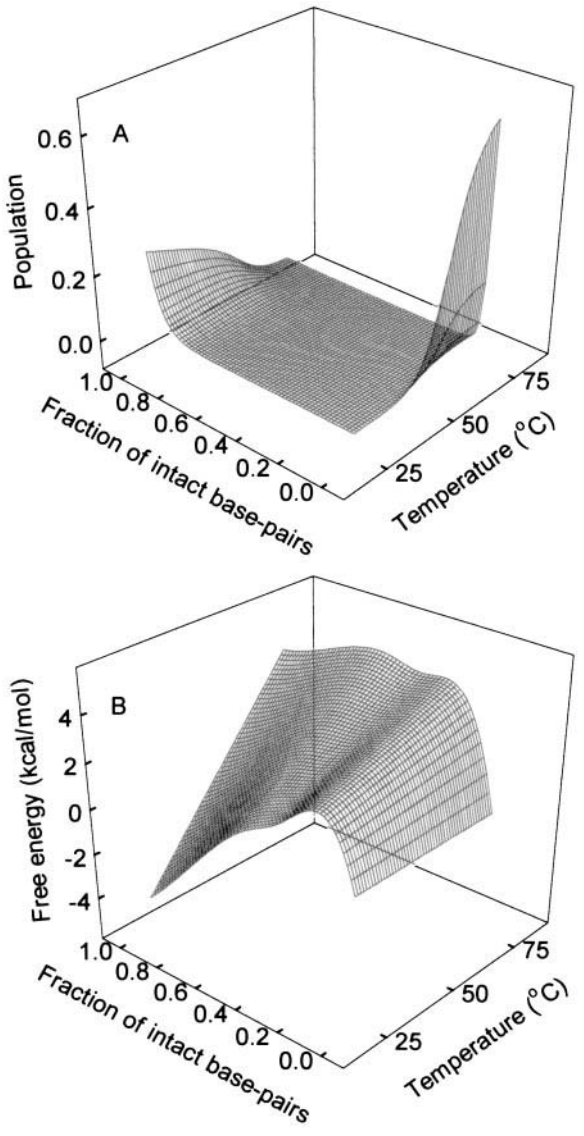

FIGURE 6 (*A*) Fractional population of hairpin conformations and (*B*) the corresponding free energy profiles as a function of the order parameter defined as the fraction of intact basepairs (θ_I) and temperature. (*A*) The fractional populations $p(\theta_I, T)$, were calculated by summing up the statistical weights of all possible microstates with a given value of θ_I, normalized by the partition function Q. The discrete values of $\ln(p)$ were interpolated on a finer grid using a cubic-spline along the direction θ_I and renormalized to obtain the continuous surface plotted in the figure. (*B*) The free energy profiles $G(\theta_I, T) = -RT \ln(pQ)$ with the random coil chosen as the reference state. The discrete values of $G(\theta_I, T)$ were interpolated on a finer grid as in Fig. 6 *A*.

folding studies, we choose θ_I, the fraction of intact basepairs, as our order parameter. We can readily calculate the fractional population $p(\theta_I, T)$ of all microstates with a specific value of θ_I, and the free energy $G(\theta_I, T)$ associated with that ensemble with the random coil chosen as the reference state (Fig. 6).

The two-state nature of the melting transition is also revealed in the free energy surface, which shows essentially

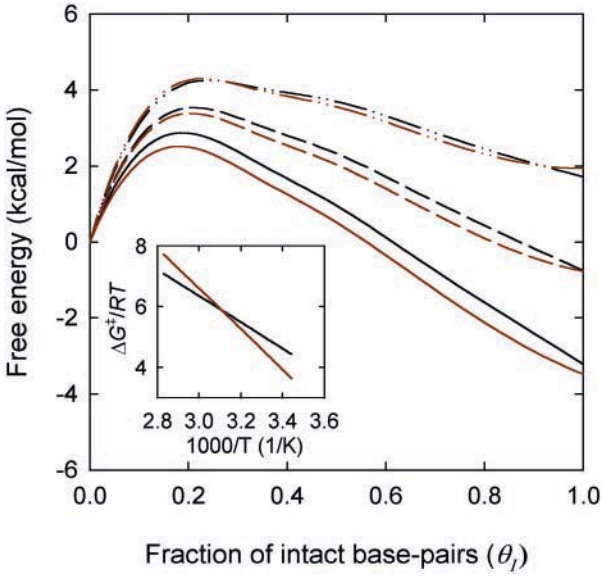

FIGURE 7 The free energy profiles $G(\theta_I, T)$ for the hairpin with 2AP substituted at site 3 (hairpin H3) plotted at three temperatures and calculated from parameters corresponding to Eq. 9 (*black curves*) and Eq. 10 (*red curves*); *continuous line*, 25°C; *dashed line*, 40°C; *dashed-dot-dot*, 55°C. (*Inset*) $\Delta G^{\ddagger}/RT$ is plotted versus inverse temperature; $\Delta G^{\ddagger}$ is the free energy barrier for the hairpin formation step, calculated from $\Delta G^{\ddagger} = G(\theta_I \approx \text{⅙})$. The black and red lines are calculated from parameters corresponding to Eq. 9 and Eq. 10, respectively. The corresponding values of $\Delta H^{\ddagger}$ are −8.6 kcal/mol and −13.3 kcal/mol, respectively.

two states corresponding to the hairpin and the random coil conformations separated by a free energy barrier. A local minimum near the transition state in the free energy profile suggests metastable states that are populated with a very low probability at equilibrium. Evidence for deviations from a strict two-state transition has been presented in earlier studies of the melting transition of hairpins (Benight et al., 1988; Senior et al., 1988; Paner et al., 1990). Transiently populated conformations of double-stranded DNA, with distorted or locally unwound structures as a result of thermal fluctuations, have been implicated in the recognition of DNA by proteins (Benight et al., 1995; Leger et al., 1998).

Transition state for hairpin formation

The free energy barrier encountered in the hairpin-formation step at $T = 25$°C is ~2.5–2.9 kcal/mol (Fig. 7), which is approximately the free energy of forming the most optimal loop with one basepair closing the loop (see Fig. 3, *insets*). The order parameter corresponding to the location of the barrier along the reaction coordinate is ⅙ for hairpins with 6 basepairs in the stem (Fig. 7) and ⅐ for hairpins with 7 basepairs in the stem (data not shown). Therefore, the transition state for these hairpins is identified as an ensemble of looped conformations stabilized by the formation of one basepair.

We can estimate the apparent enthalpy change $\Delta H^{\ddagger}$ between the transition state and the random-coil state from the temperature dependence of the free energy profiles. In a simple two-state analysis, the closing rates k_c are proportional to $\exp(-\Delta G^{\ddagger}/RT)$ where $\Delta G^{\ddagger}$ is the free energy barrier encountered in the hairpin formation step and which can be estimated from the free energy profiles of Fig. 7 as $\Delta G^{\ddagger} = G(\theta_I \approx 1/6)$. A plot of $\Delta G^{\ddagger}/RT$ versus inverse temperature yields a slope equal to $\Delta H^{\ddagger}/R$ (see Fig. 7, *inset*) with $\Delta H^{\ddagger} = -(11 \pm 2.3)$ kcal/mol. This analysis is equivalent to estimating apparent activation energies from slopes of logarithm of rates versus inverse temperature, as is done in kinetics measurements. Thus, our value of $\Delta H^{\ddagger}$ can be compared directly with the apparent activation energy obtained for the hairpin formation step from direct kinetics measurements.

A negative value of $\Delta H^{\ddagger}$ implies that the transition state along our effective reaction coordinate has a lower enthalpy than the random coil state. Negative apparent activation enthalpies are, of course, possible if the reaction that one is considering is not a single elementary step but, instead, consists of a series of steps with at least one intermediate. In a two-state analysis, the intermediate with the highest free energy is the effective transition state. Our equilibrium modeling shows that this state is an ensemble of nucleating loops stabilized by a single basepair, thus explaining the lower enthalpy of the transition state. The estimate of $\Delta H^{\ddagger}$ obtained from our equilibrium model is in good agreement with the enthalpy contribution of ~ -6 kcal/mol to -10 kcal/mol to the stability of the hairpin from the loop, as estimated by the enthalpy difference between a hairpin and a duplex with identical sequence as the stem of the hairpin (Haasnoot et al., 1986). This comparison further supports the hypothesis that the looped conformations in the transition state ensemble have at most one basepair closing the loop.

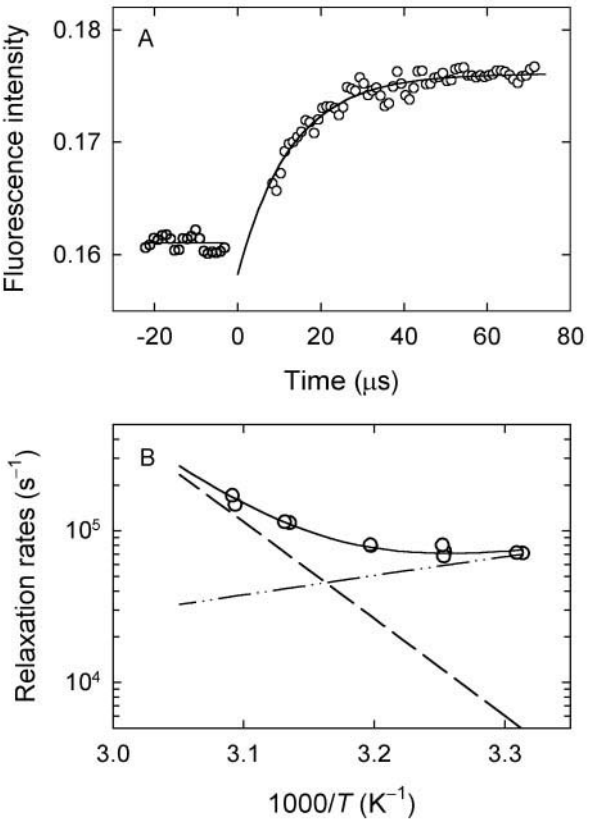

FIGURE 8 (*A*) Kinetics of unwinding of hairpin H3 (with 2AP substituted at site 3) following a laser T-jump from 30°C to 37°C. The fluorescence of 2AP is plotted as a function of time. The continuous line is an exponential fit to the data with a characteristic relaxation rate $k_r = 1/(12.8\ \mu s)$. The amplitude of the fluorescence extrapolated to zero time is smaller than the prelaser baseline because of the intrinsic change in the quantum yield of 2AP in response to the T-jump. (*B*) The relaxation rates k_r versus inverse temperature. The temperatures are the final temperatures after the T-jump. The continuous line is calculated assuming a two-state approximation, as described in the text, and an Arrhenius temperature dependence for the closing rates k_c (*dashed-dot-dot line*) and the opening rates k_o (*dashed line*).

Measurements of the kinetics of hairpin formation

The most critical test of our equilibrium statistical mechanical model and the validity of the parameters is a comparison of the predictions of the model with direct kinetics measurements. We have monitored the unwinding of the 2AP-substituted hairpins using time-resolved fluorescence changes of 2AP following a laser temperature jump (T-jump) to initiate the unwinding. The T-jump apparatus and kinetics measurements are described in detail elsewhere (Ansari et al., 2001a). Here, we present a characteristic response of 2AP fluorescence for the hairpins with 2AP substituted at site 3 (H3) and the temperature dependence of the measured relaxation rates k_r (Fig. 8). Assuming a simple two-state system, we can write $k_r = k_c + k_o$ where k_c and k_o are the characteristic rate coefficients for hairpin closing and opening steps, respectively, and k_c/k_o is the equilibrium constant K_{eq} for forming hairpins, and which can be calculated directly from the absorbance melting profiles as $K_{eq} = (1 - \theta)/\theta$. We further assume an Arrhenius temperature dependence for the closing rate coefficient k_c,

$$k_c = k_{c0}(T_0)\,\frac{\eta(T_0)}{\eta(T)}\exp\left[-\frac{E_a}{R}\left(\frac{1}{T}-\frac{1}{T_0}\right)\right], \qquad (11)$$

where k_{c0} is the rate coefficient at the reference temperature T_0 (chosen to be 25°C), E_a is the apparent activation energy, and $\eta(T)$ is the temperature-dependent viscosity of the solvent. Therefore, the temperature dependence of k_r can be described in terms of two free parameters, k_{c0} and E_a. The best fit parameters yield $k_{c0} = 1/(12.8\ \mu s)$ and $E_a = -9.5$ kcal/mol. The value for the apparent activation energy E_a obtained from these kinetics measurements is in close agreement with the

value of $\Delta H^{\ddagger}$ obtained from the temperature dependence of the free energy barrier calculated using the equilibrium model.

An assumption underlying the assignment of the slopes of Arrhenius plots to an enthalpy change is that the enthalpy and entropy changes are temperature independent. This assumption is most likely not valid for DNA, where the melting is accompanied by a significant heat capacity increase (Chalikian et al., 1999; Rouzina and Bloomfield, 1999). However, a comparison between $\Delta H^{\ddagger}$ obtained from the temperature dependence of free energy barriers (Fig. 7) and E_a obtained from direct kinetics measurements (Fig. 8), both of which are a measure of slopes on Arrhenius plots, still serves as a useful comparison between the predictions of the model and the kinetics measurements, and as a check for the validity of the parameters obtained from the equilibrium model, which were used to generate the free energy profiles. The close agreement between the predictions of the equilibrium model and the kinetics measurements also suggests that our estimate of the effective reaction coordinate and the corresponding free energy profiles are appropriate for modeling the dynamics of hairpin formation (Ansari et al., 2001a; Shen et al., 2001).

Negative activation enthalpies are also observed for the association step in the kinetics of duplex formation in short oligonucleotides (Porschke and Eigen, 1971; Craig et al., 1971). Our identification of the nature of the transition state in hairpins is, however, in apparent contradiction with the conclusions based on this earlier work on duplex formation. Using a kinetic zipper model to analyze the kinetics of duplex formation, these papers showed that the minimum nucleus size for dimer formation consists of about three basepairs. Their results are not directly applicable to hairpin formation for the following very important reasons. First, in duplex formation, two single strands have to make contact, resulting in a significant loss of translational entropy that is not a factor in unimolecular hairpin formation. Second, the first basepair that is formed in the duplex is not stabilized by stacking interactions on either side. Therefore, in duplexes, the first basepair melts faster than the next one can be formed. In hairpins, however, there are substantial stabilizing interactions arising from intraloop and loop–stem stacking interactions. Our results then suggest that, in a hairpin, the formation of the first basepair is sufficiently stable and can lead to rapid "zipping" of the stem once the correct nucleating loop is formed.

Finally, it is illuminating to draw analogies with the folding studies of β-hairpins and α-helices in polypeptides, which also yield a negative activation enthalpy for the folding step (Munoz et al., 1997; Lednev et al., 1999). In β-hairpins, the transition state is postulated to be stabilized by the formation of hydrogen bonds upon loop formation (Eaton et al., 1998).

Some caveats

The equilibrium model presented here has several limitations. First, there is no sequence dependence assumed for the persistence length or for the stacking interactions within the loop. Second, electrostatic effects are not included explicitly. Sequence dependence, both in the loop composition and in the stem sequence closing the loop, affect the stability of the hairpin (Vallone et al., 1999). The persistence length of poly(dA) is reported to be larger than that for poly(dT) (Mills et al., 1999), and the persistence length of poly(rA) is found to exhibit a significant temperature dependence (Eisenberg and Felsenfeld, 1967). Both the persistence length values and the stacking interactions are also expected to vary with salt concentration (Tinland et al., 1997). Furthermore, for low salt concentrations, the repulsion between different segments of the chain are likely to affect the loop-closure probability that should be accounted for in a more complete description. In the analysis presented in this paper, the electrostatic effects are included only implicitly in the dependence of the melting temperature on salt concentration (Frank-Kamenetskii, 1971).

SUMMARY AND CONCLUSIONS

We have used an equilibrium zipper model to describe the dependence of the melting profiles of DNA hairpins on the size of the loop and to calculate free energy profiles that are appropriate for describing the dynamics of hairpin formation along an effective one-dimensional reaction coordinate. The principal results from this study are: 1) The free energy of the hairpin scales with the loop size with an apparent exponent of $\gtrsim 7$, much larger than the exponent of ~ 1.5–1.8 expected from considerations of loop entropy alone. 2) The equilibrium model yields a very reasonable value for the persistence length of ssDNA, $P \approx 1.4$ nm for poly(dT) strands in 100 mM NaCl. 3) The transition state corresponding to the largest free energy barrier for hairpin formation along the effective reaction coordinate is identified as an ensemble of looped conformations stabilized by one basepair closing the loop. 4) The equilibrium model predicts a lower enthalpy of the transition state relative to the random coil state. 5) The predictions from the equilibrium model agree quantitatively with apparent activation energies obtained from kinetics measurements.

Our results and analysis show that the end-loop weighting function describing the variation in the stability of hairpins for varying loop sizes depends both on the entropic cost of forming loops and stabilizing interactions within the loop; these stabilizing interactions depend strongly on the size of the loop. We propose that the increased stability of tighter loops arises from hydrophobic interactions of the bases within the loop, and that larger loops are significantly destabilized from exposure of the loop bases to the aqueous environment. Two experiments are proposed to test this hypothesis. One is to replace the thymine bases in the loop with the more hydrophobic 5-nitroindole base, which should exhibit an increased dependence on the stability of hairpins with loop size. The other is to change the solution to a less polar solvent, which should exhibit a decreased

dependence on the stability of hairpins with loop size. It will also be important to check the salt dependence of these stabilizing interactions in DNA hairpins to determine whether the interactions are diminished at high salt concentrations, as appears to be the case for RNA hairpins (Groebe and Uhlenbeck, 1988; Serra et al., 1997).

The parameters obtained from the equilibrium model were used to calculate free energy profiles along a generalized one-dimensional reaction coordinate describing the hairpin-to-coil transition. The location of the free energy barrier along this coordinate, and which is identified as an effective transition state, corresponds to an ensemble of looped conformations with one basepair closing the loop. In support of this result, our analysis shows that the enthalpy of the transition state is lower than that of the random coil by an amount that is approximately the enthalpy gain upon forming a loop in hairpins. Perhaps the most remarkable finding of this work is that we can quantitatively predict, from the equilibrium model, the magnitude of the apparent activation energy observed in kinetics measurements.

Accurate measurements of the flexibility of single-stranded polynucleotides and a deeper understanding of the specific interactions that stabilize local secondary-structure formation is an important step toward developing improved physical models for the interpretation of many biophysical measurements ranging from single-molecule manipulation of single-stranded polynucleotides (Smith et al., 1996; Maier et al., 2000; Clausen-Schaumann et al., 2000; Liphardt et al., 2001; Williams et al., 2001; Montanari and Mezard, 2001) to a more fundamental understanding of the thermodynamics of protein–DNA interactions in general. In particular, accurate measurements of the temperature, salt, and sequence dependence of the intrinsic flexibility of ssDNA will be important to understand the differences observed between the binding of poly(dA), poly(dT), and poly(dC) strands to single-strand binding proteins such as SSB (Kozlov and Lohman, 1999). Finally, a physical understanding of the dynamics of hairpin formation and the underlying free energy landscape is key to elucidating the nucleation events that subsequently lead to the correct folding of RNA molecules.

We thank John F. Marko and Mark Schlossman for helpful comments and Peter M. Vallone for his assistance with equilibrium melting measurements.

This work was supported by National Science Foundation grants MCB-9707480 and MCB-9722295, and by the donors of The Petroleum Research Fund administered by the American Chemical Society through grant ACS-PRF 32099-AC4. A.A. is the recipient of a CAREER award from the National Science Foundation.

REFERENCES

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1989. Molecular Biology of the Cell. Garland Publishing, Inc., New York.

Ansari, A., C. M. Jones, E. R. Henry, J. Hofrichter, and W. A. Eaton. 1994. Conformational relaxation and ligand binding in myoglobin. *Biochemistry.* 33:5128–5145.

Ansari, A., S. V. Kuznetsov, and Y. Shen. 2001a. Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci. U.S.A.* 98:7771–7776.

Ansari, A., Y. Shen, and S. V. Kuznetsov. 2001b. Misfolded loops decrease the effective rate of DNA hairpin formation. *Phys. Rev. Lett.* (in press).

Benight, A. S., F. J. Gallo, T. M. Paner, K. D. Bishop, B. D. Faldasz, and M. J. Lane. 1995. Sequence context and DNA reactivity: application to sequence-specific cleavage of DNA. *Adv. Biophys. Chem.* 5:1–55.

Benight, A. S., J. M. Schurr, P. F. Flynn, B. R. Reid, and D. E. Wemmer. 1988. Melting of a self-complementary DNA minicircle. Comparison of optical melting theory with exchange broadening of the nuclear magnetic resonance spectrum. *J. Mol. Biol.* 200:377–399.

Bhasin, A., I. Y. Goryshin, and W. S. Reznikoff. 1999. Hairpin formation in Tn5 transposition. *J. Biol. Chem.* 274:37021–37029.

Blake, R. D., and S. G. Delcourt. 1998. Thermal stability of DNA. *Nucleic Acids Res.* 26:3323–3332.

Bloom, L. B., M. R. Otto, J. M. Beechem, and M. F. Goodman. 1993. Influence of 5′-nearest neighbors on the insertion kinetics of the fluorescent nucleotide analog 2-aminopurine by Klenow fragment. *Biochemistry.* 32:11247–11258.

Bonnet, G., O. Krichevsky, and A. Libchaber. 1998. Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. U.S.A.* 95:8602–8606.

Bryngelson, J. D., and P. G. Wolynes. 1987. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 84: 7524–7528.

Camacho, C. J., and D. Thirumalai. 1995. Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* 92:1277–1281.

Chalikian, T. V., J. Volker, G. E. Plum, and K. J. Breslauer. 1999. A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques. *Proc. Natl. Acad. Sci. U.S.A.* 96:7853–7858.

Chen, D., S. Bachellier, and D. M. Lilley. 1998. Activation of the leu-500 promoter by a reversed polarity tetA gene. Response to global plasmid supercoiling. *J. Biol. Chem.* 273:653–659.

Clausen-Schaumann, H., M. Rief, C. Tolksdorf, and H. E. Gaub. 2000. Mechanical stability of single DNA molecules. *Biophys. J.* 78:1997–2007.

Craig, M. E., D. M. Crothers, and P. Doty. 1971. Relaxation kinetics of dimer formation by self complementary oligonucleotides. *J. Mol. Biol.* 62:383–401.

DeGennes, P. G. 1979. Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, NY.

Eaton, W. A., V. Munoz, P. A. Thompson, E. R. Henry, and J. Hofrichter. 1998. Kinetics and dynamics of loops, alpha-helices, beta-hairpins, and fast-folding proteins. *Acc. Chem. Res.* 31:745–753.

Eisenberg, H., and G. Felsenfeld. 1967. Studies of the temperature-dependent conformation and phase separation of polyriboadenylic acid solutions at neutral pH. *J. Mol. Biol.* 30:17–37.

Elson, E. L., I. E. Scheffler, and R. L. Baldwin. 1970. Helix formation by d(TA) oligomers. 3. Electrostatic effects. *J. Mol. Biol.* 54:401–415.

Frank-Kamenetskii, F. 1971. Simplification of the empirical relationship between melting temperature of DNA, its GC content and concentration of sodium ions in solution. *Biopolymers.* 10:2623–2624.

Gralla, J., and D. M. Crothers. 1973. Free energy of imperfect nucleic acid helices. II. Small hairpin loops. *J. Mol. Biol.* 73:497–511.

Groebe, D. R., and O. C. Uhlenbeck. 1988. Characterization of RNA hairpin loop stability. *Nucleic Acids Res.* 16:11725–11735.

Haasnoot, C. A., C. W. Hilbers, G. A. van der Marel, J. H. van Boom, U. C. Singh, N. Pattabiraman, and P. A. Kollman. 1986. On loop folding in nucleic acid hairpin-type structures. *J. Biomol. Struct. Dyn.* 3:843–857.

Hagerman, P. J. 1985. Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers.* 24: 1881–1897.

Hilbers, C. W., C. A. Haasnoot, S. H. de Bruin, J. J. Joordens, G. A. van der Marel, and J. H. van Boom. 1985. Hairpin formation in synthetic oligonucleotides. *Biochimie.* 67:685–695.

Inners, L. D., and G. Felsenfeld. 1970. Conformation of polyribouridylic acid in solution. *J. Mol. Biol.* 50:373–389.

Jacobsen, H., and W. H. Stockmayer. 1950. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* 18:1600–1606.

Kennedy, A. K., A. Guhathakurta, N. Kleckner, and D. B. Haniford. 1998. Tn10 transposition via a DNA hairpin intermediate. *Cell.* 95:125–134.

Klump, H., and T. Ackermann. 1971. Experimental thermodynamics of the helix–random coil transition. IV. Influence of the base composition of DNA on the transition enthalpy. *Biopolymers.* 10:513–522.

Kozlov, A. G., and T. M. Lohman. 1999. Adenine base unstacking dominates the observed enthalpy and heat capacity changes for the *Escherichia coli* SSB tetramer binding to single-stranded oligoadenylates. *Biochemistry.* 38:7388–7397.

Lednev, I. K., A. S. Karnoup, M. C. Sparrow, and S. A. Asher. 1999. Nanosecond UV resonance Raman examination of initial steps in alpha-helix secondary structure evolution. *J. Am. Chem. Soc.* 121:4076–4077.

Leger, J. F., J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko. 1998. RecA binding to a single double-stranded DNA molecule: a possible role of DNA conformational fluctuations. *Proc. Natl. Acad. Sci. U.S.A.* 95:12295–12299.

Liphardt, J., B. Onoa, S. B. Smith, I. J. Tinoco, and C. Bustamante. 2001. Reversible unfolding of single RNA molecules by mechanical force. *Science.* 292:733–737.

Loakes, D., F. Hill, D. M. Brown, and S. A. Salisbury. 1997. Stability and structure of DNA oligonucleotides containing non-specific base analogues. *J. Mol. Biol.* 270:426–435.

Maier, B., D. Bensimon, and V. Croquette. 2000. Replication by a single DNA polymerase of a stretched single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* 97:12002–12007.

Marino, J. P., R. S. Gregorian, G. Csankovszki, and D. M. Crothers. 1995. Bent helix formation between RNA hairpins with complementary loops. *Science.* 268:1448–1454.

Marquardt, D. W. 1963. An algorithm for least square estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11:431–441.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.

Mills, J. B., E. Vacano, and P. J. Hagerman. 1999. Flexibility of single-stranded DNA: use of gapped duplex helices to determine the persistence lengths of poly(dT) and poly(dA). *J. Mol. Biol.* 285:245–257.

Montanari, A., and M. Mezard. 2001. Hairpin formation and elongation of biomolecules. *Phys. Rev. Lett.* 86:2178–2181.

Munoz, V., P. A. Thompson, J. Hofrichter, and W. A. Eaton. 1997. Folding dynamics and mechanism of beta-hairpin formation. *Nature.* 390:196–199.

Nordlund, T. M., S. Andersson, L. Nilsson, R. Rigler, A. Graslund, and L. W. McLaughlin. 1989. Structure and dynamics of a fluorescent DNA oligomer containing the *Eco*RI recognition sequence: fluorescence, molecular dynamics, and NMR studies. *Biochemistry.* 28:9095–9103.

Paner, T. M., M. Amaratunga, and A. S. Benight. 1992. Studies of DNA dumbbells. III. Theoretical analysis of optical melting curves of dumbbells with a 16 base-pair duplex stem and Tn end loops (n = 2, 3, 4, 6, 8, 10, 14). *Biopolymers.* 32:881–892.

Paner, T. M., M. Amaratunga, M. J. Doktycz, and A. S. Benight. 1990. Analysis of melting transitions of the DNA hairpins formed from the oligomer sequences d[GGATAC(X)4GTATCC] (X = A, T, G, C). *Biopolymers.* 29:1715–1734.

Paner, T. M., P. V. Riccelli, R. Owczarzy, and A. S. Benight. 1996. Studies of DNA dumbbells. VI. Analysis of optical melting curves of dumbbells with a sixteen-base pair duplex stem and end-loops of variable size and sequence. *Biopolymers.* 39:779–793.

Poland, D., and H. A. Scheraga. 1970. Theory of Helix–Coil Transitions in Biopolymers; Statistical Mechanical Theory of Order–Disorder Transitions in Biological Macromolecules. Academic Press, New York.

Porschke, D., and M. Eigen. 1971. Co-operative non-enzymic base recognition. 3. Kinetics of the helix–coil transition of the oligoribouridylic–oligoriboadenylic acid system and of oligoriboadenylic acid alone at acidic pH. *J. Mol. Biol.* 62:361–381.

Rentzeperis, D., K. Alessi, and L. A. Marky. 1993. Thermodynamics of DNA hairpins: contribution of loop size to hairpin stability and ethidium binding. *Nucleic Acids Res.* 21:2683–2689.

Rivetti, C., C. Walker, and C. Bustamante. 1998. Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility. *J. Mol. Biol.* 280:41–59.

Roth, D. B., J. P. Menetski, P. B. Nakajima, M. J. Bosma, and M. Gellert. 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. *Cell.* 70:983–991.

Rouzina, I., and V. A. Bloomfield. 1999. Heat capacity effects on the melting of DNA. 1. General aspects. *Biophys. J.* 77:3242–3251.

SantaLucia, J., Jr.1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.* 95:1460–1465.

Scheffler, I. E., E. L. Elson, and R. L. Baldwin. 1970. Helix formation by d(TA) oligomers. II. Analysis of the helix–coil transitions of linear and circular oligomers. *J. Mol. Biol.* 48:145–171.

Senior, M. M., R. A. Jones, and K. J. Breslauer. 1988. Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc. Natl. Acad. Sci. U.S.A.* 85:6242–6246.

Serra, M. J., T. W. Barnes, K. Betschart, M. J. Gutierrez, K. J. Sprouse, C. K. Riley, L. Stewart, and R. E. Temel. 1997. Improved parameters for the prediction of RNA hairpin stability. *Biochemistry.* 36:4844–4851.

Serra, M. J., and D. H. Turner. 1995. Predicting thermodynamic properties of RNA. *Methods Enzymol.* 259:242–261.

Shen, Y., S. V. Kuznetsov, and A. Ansari. 2001. Loop dependence of the dynamics of DNA hairpins. *J. Phys. Chem. B.* (in press).

Shore, D., J. Langowski, and R. L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. U.S.A.* 78:4833–4837.

Smith, S. B., Y. Cui, and C. Bustamante. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science.* 271:795–799.

Tinland, B., A. Pluen, J. Sturm, and G. Weill. 1997. Persistence length of single-stranded DNA. *Macromolecules.* 30:5763–5765.

Uhlenbeck, O. C., P. N. Borer, B. Dengler, and I. Tinoco, Jr. 1973. Stability of RNA hairpin loops: A 6 -C m -U 6. *J Mol Biol.* 73:483–496.

Vallone, P. M., and A. S. Benight. 1999. Melting studies of short DNA hairpins containing the universal base 5-nitroindole. *Nucleic Acids Res.* 27:3589–3596.

Vallone, P. M., T. M. Paner, J. Hilario, M. J. Lane, B. D. Faldasz, and A. S. Benight. 1999. Melting studies of short DNA hairpins: influence of loop sequence and adjoining base pair identity on hairpin thermodynamic stability. *Biopolymers.* 50:425–442.

Wartell, R. M., and A. S. Benight. 1985. Thermal denaturation of DNA molecules: a comparison of theory with experiment. *Phys. Rep.* 126: 67–107.

Wemmer, D. E., S. H. Chou, D. R. Hare, and B. R. Reid. 1985. Duplex-hairpin transitions in DNA: NMR studies on CGCGTATACGCG. *Nucleic Acids Res.* 13:3755–3772.

Williams, M. C., J. R. Wenner, I. Rouzina, and V. A. Bloomfield. 2001. Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting. *Biophys. J.* 80:874–881.

Winnik, M. A. 1986. Cyclic Polymers. Elsevier, NY.

Yamakawa, H., and W. H. Stockmayer. 1972:Statistical mechanics of worm-like chains. II. Excluded volume effect. *J. Chem. Phys.* 57:2843–2854.

Yuan, R. C., J. A. Steitz, P. B. Moore, and D. M. Crothers. 1979. The 3′ terminus of 16S rRNA: secondary structure and interaction with ribosomal protein S1. *Nucleic Acids Res.* 7:2399–2418.

Zuker, M. 1989. On finding all suboptimal foldings of an RNA molecule. *Science.* 244:48–52.

Zwanzig, R. 1995. Simple model of protein folding kinetics. *Proc. Natl. Acad. Sci. U.S.A.* 92:9801–9804.