Formation of loops in DNA under tension

Sumithra Sankararaman* and John F. Marko†
Department of Physics, University of Illinois at Chicago, 845 W Taylor Street, Chicago, Illinois 60607-7059, USA
(Received 29 June 2004; published 24 February 2005)

We study the formation of loops along a DNA molecule under applied tension, as might occur in single-DNA micromanipulation experiments with proteins which are able to simultaneously bind two DNA sites. We consider the case of "bare" DNA in the loop, which forms a "teardrop" shape, and the case where a single DNA-bending protein produces a "kink" in the middle of the loop; the presence of a right-angle kink in the loop reduces its bending energy by a factor of 3. Using the bending energy plus an estimate of the free energies associated with fluctuations and the elasticity of the extended nonlooped DNA, we obtain a probability distribution for loops as a function of loop size and force. Force strongly suppresses formation of all loops, but suppresses large loops more severely than small ones. This quenching effect of force is reduced in the presence of a kink in the loop. We also calculate the speed at which length is absorbed into loops between arbitrary positions along the DNA (i.e., for non-sequence-specific loop forming proteins). The speed of retraction of the molecule decays as a stretched exponential function of the force with characteristic force scales depending on the geometry of the loops.

DOI: 10.1103/PhysRevE.71.021911 PACS number(s): 87.14.Gg, 82.35.Pq, 87.15.Aa, 87.15.La

I. INTRODUCTION

The formation of loops in DNA is essential to a wide variety of biological processes including site-specific recombination [1], regulation of gene expression by distant sequences [2–4], and DNA packaging [5,6]. These examples involve loop-forming protein complexes which bind to two sites along the DNA chain to anchor these loops. Recently, a mechanism for motion of nucleosomes on DNA via the formation of loops on DNA has been proposed [7]. We present calculations of the free energy for forming loops on DNA molecules (a) when there is no force applied at the ends of the molecule and (b) when the DNA molecule is being stretched by force. We aim at describing single-DNA micromanipulation experiments, where the relevant forces are in the range 0.01–50 pN [8].

The loop free energy can be used to estimate a variety of quantities which are relevant to single-molecule experiments on DNA with proteins. One distribution of interest is the probability for forming a loop of a specific length as a function of the applied tension. This quantity is relevant in studying loops between specific sequences on the DNA chain [Fig. 1(b)]. Another such quantity is the probability distribution of loops of different lengths (equivalently, the rate of formation of loops of different sizes) at a given applied tension. Such a probability distribution is useful in studying "nonspecific" loop formation during which the loop-forming proteins can form a loop between two arbitrary sequences along the DNA chain [Fig. 1(c)]. We will discuss results for both these types of loops.

In addition to loop-forming proteins which stabilize loops by attaching themselves to two sites at the base of the loop, there are other DNA-bending proteins which might bind to a site at the apex of the loop and bend the DNA there. In the presence of these bending proteins, the DNA chain at the apex of the loop is not smooth, but has a kink. This situation in fact has arisen in single-DNA looping experiments, where the DNA-bending protein HU [4] was present. The Gal repressor protein (GalR) represses the transcription of the gal operon in *Escherichia coli* by forming a DNA loop which encompasses the promoters of this operon. The experiments reported in Ref. [4] indicate that the binding of HU to supercoiled DNA, subsequently bending the DNA, is essential to stabilize a loop formed by two GalR dimers. In this paper, we will discuss the effect of such DNA-bending proteins on the bending energy of the loops.

We briefly motivate our calculations and summarize our results. In Sec. II, the elastic theory of a semiflexible polymer is used to construct a model of loop formation along a DNA, including the effect of an applied force. The basic idea, and motivation for this as a problem, is that the applied force will act to destabilize loop formation; this idea has

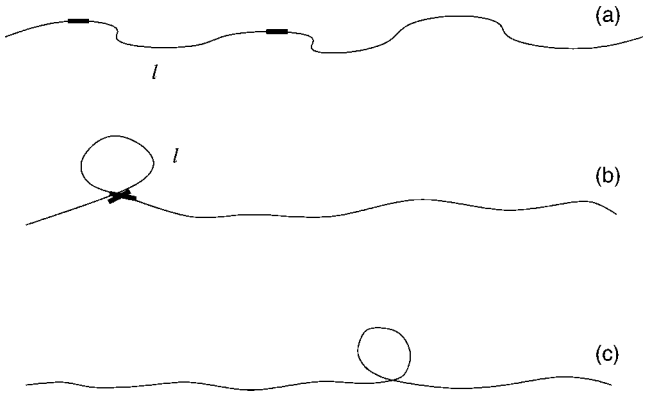

FIG. 1. Specific and nonspecific loops. (a) DNA chain without any loops. (b) A specific loop of length l formed between the two sequences marked in bold. (c) Nonspecific loops that are formed between any two points along the DNA chain.

*Email address: sumithra@uic.edu
†Email address: jmarko@uic.edu

been discussed by a few authors [9–13]. This paper focuses on the process of formation of loops, calculating the probability of arriving at DNA segments formed into loop structures. We do not focus on the lifetime of the formed loops, and instead put our attention to the problem of how the length of the segment, its possible distortion by DNA-bending proteins, and the applied tension, combine together to affect the loop formation probability. The model of Sec. II requires as input the free energy of the loop, and information about the geometry of the looped DNA. The calculations outlined in Secs. III and IV show the effect of DNA-bending proteins in reducing the bending energy of the loops. We do this by introducing a 90° "kink" in the middle of the loop, similar to the bend inserted by HU or the DNA-bending protein HMGB1 [14,15], or other DNA-bending proteins [16].

The bending energy contribution to the free energy of a loop can be calculated by minimizing the energy of a filament subject to the constraint of loop closure. In Appendix A, the resulting Euler-Lagrange equation is solved with appropriate boundary conditions to obtain loop shapes and bending energies. These bending-energy-optimizing solutions are useful for comparing relative energies of different types of loops, but for free energy calculations we will need to take into account thermal fluctuation effects. We use empirical formulas based on the analysis of Shimada and Yamakawa [17] to provide an estimate of fluctuation effects (Sec. III).

We begin by presenting results for "free" loops (loops along molecules with no end constraints, i.e., zero tension), focusing on the probability of formation of loops of different lengths l. We consider smooth "teardrop" shaped loops, recovering results of Stockmayer and Yamakawa [18] and compare them with kinked loops. We find that a 90° kink can greatly reduce the loop bending energy.

We then move to results for loop formation in molecules under tension (Sec. IV), as can be done in single-molecule experiments. Just as one example, force can be used in such experiments to keep loops from forming. We study this using our model, where the looping probability depends on work done by the applied force, as well as bending energy and entropic contributions from thermal fluctuations. To do this we first compute the minimum-energy configurations of molecules of length L containing a loop of length l, for molecules with their ends constrained as would occur in a micromanipulation experiment (Appendix B).

We then combine the resulting bending energies and the geometry of the loop region of the molecule, with the fluctuation correction for the loop region, and the free energy of the extended part of the polymer outside the loop, to arrive at a free energy model for a loop along a long DNA molecule. We use this to compute the distribution of loop sizes and therefore the most probable loop size, as a function of tension. We make quantitative predictions for how tension reduces the most probable loop size, for both teardrop and kinked loops. For all tensions, the presence of a 90° kink in the loop reduces the most probable loop length by about a factor of 3, thanks to the large saving in bending energy introduced by the kink.

In the case where nonspecific loops form, an experiment might observe gradual "folding up" of the molecule in time, and a consequent reduction in extension while constant force is applied. This kind of folding up of arbitrary sequences of DNA occurs during the packaging of DNA [5]. We calculate the rate at which the ends of a DNA molecule would be pulled in during the formation of nonspecific loops (Sec. V). Our calculation applies most simply in the regime where the displacement of the DNA against the applied force is the rate-limiting part of this process. We find that the speed at which the molecule contracts decays as a stretched exponential function of force. This is much slower than a simple exponential decay with force as is the case for loops of fixed length. Specifically, the speed of nonspecific loop formation decays as $e^{-\sqrt{(f-f^*)/f_0}}$, where f^* and f_0 are characteristic force scales depending on the geometry of the loops. These forces are factors of the basic force scale in the system, $1/(\beta A)$ and are higher for loops which have lower bending energy.

II. FREE ENERGIES OF THE OPEN AND ONE-LOOP CONFIGURATIONS OF DNA

Double-stranded DNA is quite stiff with a persistence length (denoted by A) of 50 nm. The persistence length is the length over which the tangent vectors at points along the DNA molecule remain correlated. In terms of the number of base-pairs along the sequence this length is 150 base pairs. The energy associated with DNA conformational changes is well described by the wormlike chain model. The bending modulus for the DNA polymer is just k_BTA. The Kuhn segment length (denoted by b) is $2A$. The unit of thermal energy $1k_BT=4.1$ pN nm (at 300 K). The total contour length of the DNA molecule is denoted by L.

We first summarize the free energy of a DNA molecule subjected to force at the ends [19]. Let us consider a linear piece of a DNA molecule of length L. A force f is applied to this molecule. The unit tangent vector at each point along the contour of the molecule is denoted by $\hat{u}$. The energy of gradual bends in the DNA molecule is given by

$$\beta E = \frac{A}{2}\int_0^L ds\left(\frac{d\hat{u}}{ds}\right)^2, \tag{1}$$

where s is the arclength variable measured along the contour of the loop. The extension of the molecule due to the applied force is $z=\int_0^L ds\hat{u}\cdot\hat{z}$. The partition function can therefore be written as a path integral

$$Z(f) = \int \mathcal{D}\hat{u}e^{-\beta(E-fz)} = \int \mathcal{D}\hat{u}\,\exp\left\{-\int_0^L ds\left[\frac{A}{2}\left(\frac{d\hat{u}}{ds}\right)^2 - \beta f\hat{u}\cdot\hat{z}\right]\right\}. \tag{2}$$

We define a dimensionless quantity $g(\beta Af)$ which is the free energy per unit length (in the fixed force ensemble) [12,19]. This quantity can be related to the partition function through

$$g(\beta Af) = \frac{A}{L} \ln Z. \tag{3}$$

Although this can be exactly evaluated as a function of the force [19], we give only the asymptotic forms here. In the $f \to \infty$ limit, the path integral becomes a Gaussian integral and it can be shown that $g(\beta Af) = \beta Af - \sqrt{\beta Af} + \cdots$. The high force limit corresponds to $\beta Af > 1$. In the low force limit (when $\beta Af < 1$), $g(\beta Af) = (\beta Af)^2/3 + \cdots$.

The relative probability of forming a loop of size l from the unlooped state is proportional to the exponential of the difference in free energy between these two states. The free energy of a DNA strand of length L in the unlooped state subject to an external force f [Fig. 1(a)] is given by

$$\ln Z_{open} = \frac{L}{A} g(\beta Af), \tag{4}$$

where $g(\beta Af)$ is the dimensionless free energy described in the preceding subsection. In the "one-loop" configuration [Figs. 1(b) and 1(c)], the free energy may be written as

$$\ln Z_{closed} = [-\beta E_b] + \left[\frac{L-(l+2\epsilon)}{A} g(\beta Af)\right] + \text{contributions from fluctuations}. \tag{5}$$

The first term is the curvature energy cost in bending a portion $l \ll A$ of the DNA polymer into a loop. This term is dominant for loops of small sizes (less than a persistence length). For a perfectly circular loop $\beta E_b = 2\pi^2 A/l$. When force is present, the ends of a polymer are pulled in by an additional distance 2ϵ (see Fig. 5 below) during the formation of the loop due to the necessity to bend the DNA between the loop and the force-extended regions. Therefore, the total loss in length of the molecule when the loop is formed, neglecting bending fluctuations along the extended parts of the chain, is $l+2\epsilon$. The second term in Eq. (5) computes the work done by the external force on the portion $L-(l+2\epsilon)$ of the polymer which is not part of the loop.

The final term is the contribution to the free energy due to entropic fluctuations in the loop region. For large loops, the entropic cost of loop formation is the familiar Gaussian term $(3/2)\ln l$ [20]. For small loops, it is difficult to estimate the form of the contributions from entropic fluctuations analytically. Based on their detailed numerical analysis, Yamakawa and Shimada proposed an empirical formula for the loop-closure probability, from which we extract an estimate of the entropic contribution [17]. We will discuss this in more detail below. In the next section, we fill in the unknowns in Eq. (5), for the case of a "free" loop under zero tension, where $g=0$. We will also discuss the form for the contributions from fluctuations in the loop, which will allow us to compute the zero-force loop-formation probability distribution. Then, in Sec. IV we will generalize our results to nonzero applied tension. Recently, it has been suggested that very small loops may require substantially less bending energy for their formation, due to nonlinear elastic effects not included in Eq. (1). Our model can be readily modified to include such effects, through modification of E_b for small l.

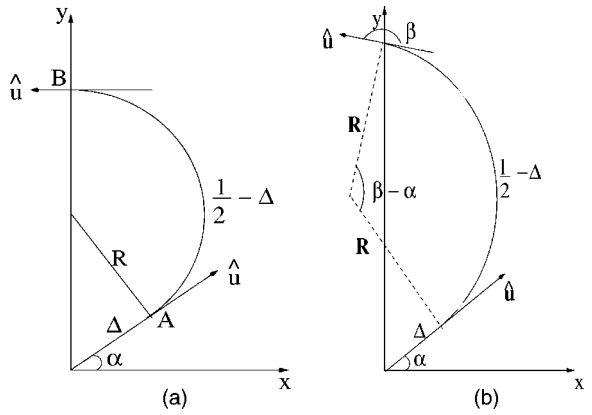

FIG. 2. (a) The teardrop loop; angle at the apex of the loop is π. (b) The kinked loop; angle at the apex of the loop is β.

III. LOOPS IN DNA WITH FREE ENDS

When no force is applied at the ends of the DNA molecule, the parts of the molecule that do not form the loop do not play a role in determining the shape and energy of the loop ($g=0$). Hence, for all the calculations in this section, we will work with only the portion of the DNA molecule that forms the loop, of length l. This is relevant to the cyclization of a long chain of DNA, or the formation of a loop in a molecule with ends that are not under tension or other constraint.

A. Simple calculation for bending energy of a loop

In this subsection we will first review how a loop which is teardrop shaped [shown in Fig. 2(a)] has substantially lower bending energy than a perfectly circular loop. Later, we will show that the presence of a kink in the loop [Fig. 2(b)] can lead to a large further reduction in the loop bending energy. For notational convenience we will refer to these two types of loops as the teardrop loop and the kinked loop, respectively. We search for configurations minimizing the bending energy in which the angle of juxtaposition between the two arms of the DNA molecule is not fixed. Alternately, we could also do the minimization to find configurations having a fixed angle of juxtaposition. We will use a simple calculation to calculate the bending energy for different types of loops based on the "circle-line" approximation introduced by Kulić and Schiessel [7].

Consider the loop shown in Fig. 2(a). In this case, the angle made by the tangent vector with the X axis at the apex of the loop is π rad. We calculate analytically the energy of the loop as a function of the angle α and then minimize the energy with respect to α to find the most favorable angle. In Fig. 2(a), AB is a portion of a circle and has a length $l/2-\Delta$, Δ being the length of the linear portion. Let the radius of curvature of the circle of which AB is a part be R. Hence, $R(\pi-\alpha) = (l/2) - \Delta$. We also know that $\Delta = R \tan \alpha$. Using this the above equation becomes $R = (l/2)/(\pi - \alpha + \tan \alpha)$. The curvature of the circle is given by the inverse of this radius. The linear part of the loop does not have any curvature energy. Thus, the bending energy of the loop is, from Eq. (1),

$$\beta E(\alpha) = \frac{2A}{l}(\pi - \alpha)(\pi - \alpha + \tan\,\alpha). \tag{6}$$

The minimum of this energy occurs at $\alpha = 0.8948$ rad and $E_{min}(0.8948) = 15.6987A/l$ (k_BT units). A perfectly circular loop of length l has a bending energy $2\pi^2 A/l$. The ratio of this energy to that of a perfect circle of the same length is 0.7953. This shows that the most favorable configuration for a loop in the absence of DNA-bending proteins is the one shown in Fig. 2(a). We note that this bending energy and loop-opening angle are those obtained by Kulić and Schiessel using their "circle-line" approximation [7].

We can generalize this argument for loops which have a kink at the apex. For example, such a loop might be formed by the presence of DNA-bending proteins and is shown in Fig. 2(b). Let the angle made by the tangent vector with the X axis at the apex of the loop be β. As before, for the circular section of the loop, $R(\beta - \alpha) = (l/2) - \Delta$. In the geometry shown above, $R = (l/2)/[\beta - \alpha + (\sin\,\alpha - \sin\,\beta)/\cos\,\alpha]$. The energy is

$$\beta E(\alpha) = \frac{2A}{l}(\beta - \alpha)\left(\beta - \alpha + \frac{\sin\,\alpha - \sin\,\beta}{\cos\,\alpha}\right). \tag{7}$$

As before, we choose the loop length to be $10A$. With $\beta = 2.356\,19$ rad (this value of β corresponds to having a 90° bend at the apex of the loop), we minimize this energy with respect to α and find that $\alpha_{min} = 1.1929$ rad and $E_{min} = 4.108\,44A/l$ (k_BT units). Expressed as a fraction of the energy of a perfect circle this is 0.2081. Thus a kink in the loop can lead to a large reduction in loop bending energy (reducing it by a factor of almost 4 from that for a teardrop shaped loop).

B. Exact bending energy for noncircular loops using the elastic filament model

We now compute the bending energy for the two different kinds of loops shown in Figs. 2(a) and 2(b). Yamakawa and Stockmayer calculated the equilibrium configuration for the teardrop loop [18]. Since the ends of the loop are not subjected to torque, the curvature at the base of the loop must vanish. This boundary condition along with the constraint of loop closure defines the equilibrium shape of the loop uniquely. Via calculations shown in Appendix A we reproduce the Yamakawa-Stockmayer (YS) result for the teardrop loop, and also find the minimum-energy configuration for the kinked loop. The calculations are essentially those of the classical "elastica" problem [21]. The influence of DNA-bending proteins on the the DNA molecule has also been studied in the context of interaction between proteins bound to DNA [10,11].

The results of Yamakawa and Stockmayer show that when no force is applied, the shape of the loop is determined by the boundary condition of zero curvature at the base of the loop [see Fig. 2(a)] and by the constraint of loop closure [18]. The angle α was 0.8604 rad. To verify the results of Yamakwa and Stockmayer, we use the values of $\beta = \pi$ rad and $\alpha = 0.8604$ rad in Eqs. (A9) and (A11) and obtain $\alpha' = 0$ and $f = 2.154/l$. The energy is $14.055A/l$ (note both f and E are reduced by a factor of k_BT). This energy is 0.712 times the energy of a perfectly circular loop of the same length. Figure 3(a) shows the YS loop shape (in the figure, the loop has length 10). The shape of the loop is unique and the energy scales as $1/l$.

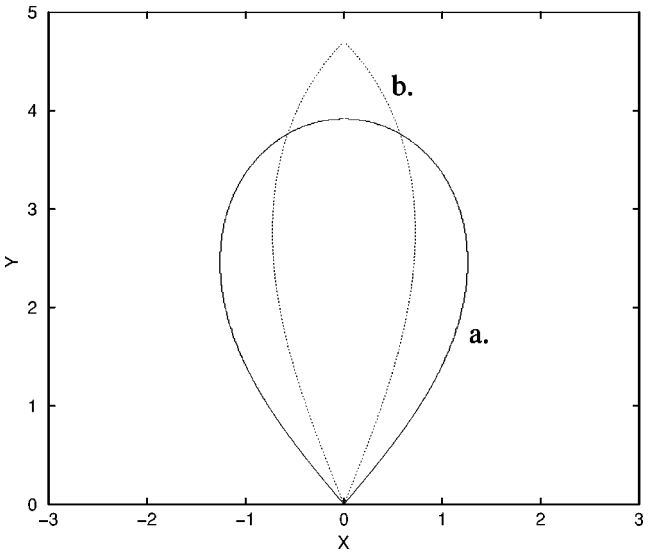

FIG. 3. (a) The teardrop loop, $\beta = \pi$ rad, and (b) the kinked loop, $\beta = 2.3562$ rad. These are the exact shapes of the loops.

We imagine also the case where a DNA-bending protein produces a strong kink at the apex of the loop. We model this by imposing a boundary condition $\beta \neq \pi$. A loop with a 90° kink between its arms at the apex has $\beta = 2.3562$ rad. We find that the lowest energy configuration of such a loop has $\alpha = 1.1874$ rad and $\alpha' = 0$. The energy of such a loop is (in units of k_BT) is $3.6567A/l$, or a fraction 0.1853 of the energy of a circular loop of the same length. Comparing this energy with that for the teardrop loop, we find that the presence of the 90° kink lowers the bending energy by almost a factor of 4. The kinked loop is shown in Fig. 3(b) (the loop in the figure has contour length 10).

C. Loop length distribution at zero force

As mentioned in Sec. II, Shimada and Yamakawa [17] proposed an empirical formula for the probability distribution including effects of thermal fluctuations in the loop region. They began by calculating the most probable configuration of a closed loop. Fluctuations around this lowest energy state were integrated to obtain the entropic correction to the bending energy of the most probable configuration. The difference in free energy of the open and closed states of the DNA molecule discussed in Sec. II appears in the exponential term of this distribution. Since there is no external force, from Eq. (5) and Eq. (4), this energy is just the sum of the bending energy and the contributions from fluctuations. The formula is

$$p(n) = \begin{cases} \pi^2 n^{-6} e^{-\beta E_b + 0.514n} & \text{for a circle,} \\ 28.01 n^{-5} e^{-\beta E_b + 0.492n} & \text{for a teardrop,} \end{cases}$$

where n is the number of segments of length b in the loop ($n = l/b = l/[2A]$), βE_b is the bending energy for the loop

(π^2/n for a circle and $7.027/n$ for a teardrop), and the second factor in the exponential gives the correction due to fluctuations, in each case. In the above formulas all lengths are measured in units of a segment length $b=2A$. We note here that this formula diverges for loop lengths greater than a few thousand kilo-base-pairs and make sure that we use it when the total length of the molecule is less than 2 kilobases. For large enough n, the free energy must approach the Gaussian limit and the probability should decay as a power law ($n^{-3/2}$). The interpolation of the above formula to the Gaussian limit is not relevant to our results and thus we do not discuss it further. The expression for the teardrop loop gives the correct value of the most probable loop size (around 500 base pairs or $n=1.6$) at zero force.

We suppose that the loop is formed only when the loop-forming sequences on the DNA polymer are nearer than some critical distance δ within a segment volume b^3 where $\delta \ll b$. To take this effect into account we multiply the probability distribution by the factor δ^3/b^3. We consider a reaction distance δ of 1 nm [4]. Our probabilities are therefore those to find the loop ends in the same 1 nm^3 volume and may be converted to mol/l by multiplying by a factor of 0.6.

We use the bending energies calculated in the previous subsection to write down the probability distribution of a single loop of size l, relative to the open state. The expressions are

$$p(n) = \begin{cases} \pi^2 \dfrac{\delta^3}{b^3} n^{-6} e^{-\pi^2/n + 0.514n} & \text{for a circle,} \\ 28.01 \dfrac{\delta^3}{b^3} n^{-5} e^{-0.712\pi^2/n + 0.492n} & \text{for a teardrop loop,} \\ 28.01 \dfrac{\delta^3}{b^3} n^{-5} e^{-0.1853\pi^2/n + 0.492n} & \text{for a kinked loop,} \end{cases} \tag{8}$$

where $n=l/b$ is the number of statistical segments in the loop (recall $b=2A$). The assumption here is that the correction for fluctuations has the same form $e^{0.492n}$ for both the teardrop loop and the kinked loop.

The above expression is central to all the calculations we present in the following sections. We therefore reiterate the meaning of this formula. This formula gives the probability distribution for the length of the loops formed at zero applied force. A loop is said to be formed if two points on the DNA polymer lie within a distance δ of each other. This expression includes the bending energy and the contributions due to fluctuations in the loops. The effects of an external force on the probability distribution are discussed in Sec. IV.

The plots in Fig. 4 show the probability distribution for the length of the loop formed when there is no tension applied to the ends of the DNA molecule. The distributions are plotted as a function of the number of base pairs along the chain; recall that each segment of length 100 nm contains 300 base pairs. We see a shift in the peak of this distribution toward smaller loops as we go from a perfectly circular loop to the kinked loop [Figs. 4(a)–4(c)]. The most probable loop sizes at zero force are around 550, 480, and 110 base pairs for a circular loop, a teardrop loop, and a kinked loop, respectively. We also observe that the probability of formation of these loops (equivalently, the rate of formation) increases dramatically when the kink is added. In the zero-force regime, the bending energy is dominant and the loop that has the lowest bending energy among the three is the most probable. At zero force, numerical calculations and empirical analysis also determine this peak to be around 500 base pairs for teardrop loops [17,22,23].

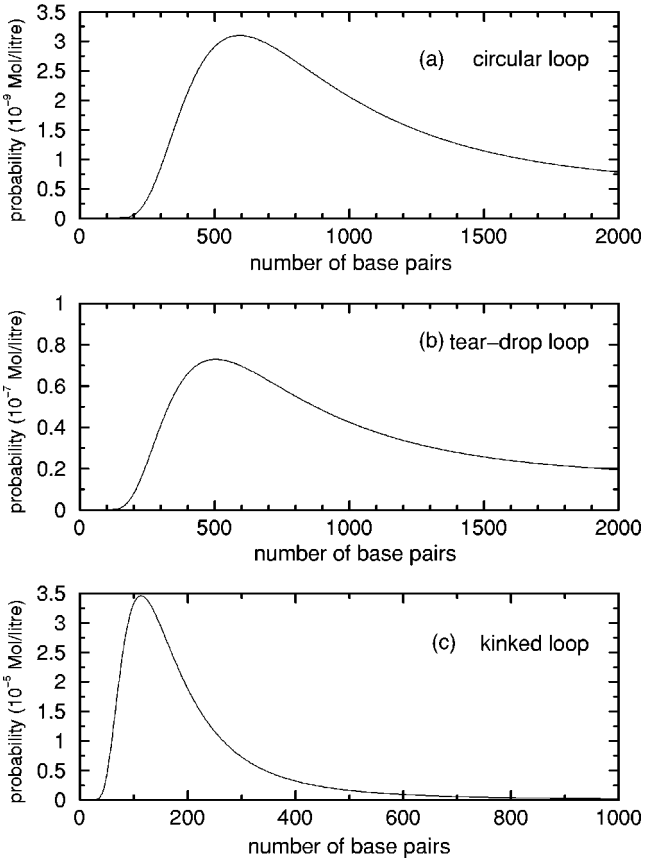

FIG. 4. Zero-force probability distribution for the length of the loop for the circular loop, the teardrop loop, and the kinked loop. Note that the most probable loop size shifts to shorter distances along the loop and the peak height increases from (a) to (c).

IV. LOOPS IN DNA WITH ENDS CONSTRAINED BY EXTERNAL FORCE

When tension is applied to the ends of the DNA molecule (as is the case in micromanipulation experiments on DNA), we need to take the full length of the molecule into account [i.e., the length l which forms the loop and the length $L-(l+\epsilon)$ which is outside the loop]. The applied force does work to pull the two ends of the molecule together during the formation of a loop and this work must be included in estimating the free energy. We also wish to estimate the bending energy stored in the part of DNA that is outside the loop region in addition to the energy stored in the loop. In this section, we calculate the bending energy, the end-to-end extension, the length of the loop, and the "opening angle" for the two types of loops. We also calculate the asymptotic

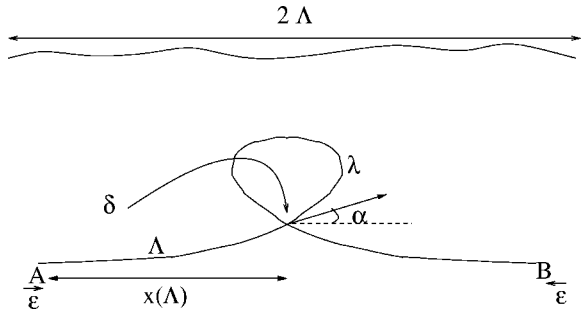

FIG. 5. Notation for the lengths in the scaled coordinates. Note that all quantities are specified for one half of the loop. Also shown is the "reaction distance" δ, which is the distance with in which the ends of the loop must approach each other in a segment volume to form a loop. The "opening angle" of the loop is $2(90°-\alpha)$. The ends of the polymer are pulled in through a distance of ϵ by the force.

values of these quantities in the long-chain limit (when $L \to \infty$) and show that the length of the loop, l, is finite in this limit. We do all the calculations for one-half of the loop. Since the loop is symmetric, all the quantities can be doubled to obtain values for the full loop. We carry out the detailed calculations in Appendix B and list only the results in this section. We note that unlike the loops described above which were all described by length-independent loop shapes, the loops of this section are members of a one-parameter family of loop shapes. The dimensionless shape parameter can be considered to be L/l [we also use a parameter called λ_0 (see Appendix B), which is more convenient as it reaches a finite limit when $L/l \to \infty$]. Our notation for the geometry is summarized in Fig. 5.

A. Teardrop loop with constrained ends

For a teardrop loop of total length l in a molecule of length L, $\lambda_0 = 1.915\,008$ (see Appendix B) and $E \approx 8\lambda_0 A/l = 15.3201 A/l$. This is slightly more than the result for the energy of a teardrop loop of the same length at zero tension of $14.0549 A/l$. This means that for an end-constrained DNA chain, the region outside the loop contributes to the bending energy and stores roughly 10% of the total energy. Hence, in estimating the bending energy for the DNA chain in the presence of tension, it is important to take the parts of the chain outside the loop into account.

We plot the probability distribution as a function of the size of the loop at different forces. The typical forces used in micromanipulation experiments are in the range 0.01–10 pN. The free energy per unit length $g(\beta A f)$ is nonzero and the second term in Eq. (5) must be included in computing the probability density. Therefore, the probability density for a loop of size $l = nb$ is given by the expression

$$p(n) = 28.01 \frac{\delta^3}{b^3} n^{-5} e^{-4\lambda_0/n + 0.492n} e^{-2ng(\beta A f)}. \tag{9}$$

The change in the end-to-end extension of the whole molecule in the looped state from its fully extended state is only 1.04 times the size of the loop. We therefore neglect the extra contribution from ϵ in calculating the work done by the force in folding the polymer into the one-loop configuration. When $\beta A f < 1$ (low force regime) we use $g(x) = x^2/3$. When $\beta A f > 1$ (high force regime), we use $g(x) = x - \sqrt{x}$. The plots in Fig. 6 show the probability distribution when the applied tension lies between 0.01 and 0.5 pN. The distributions are plotted as a function of the number of base pairs along the DNA chain (each segment of length 100 nm contains 300 base pairs). As expected, the peak of the distribution shifts toward smaller loop sizes as the force is increased. We will return to a discussion of the results in Sec. V.

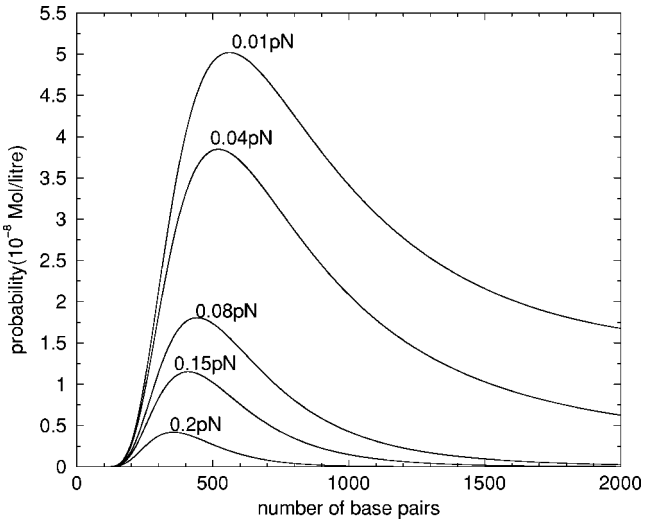

FIG. 6. Teardrop loop: Probability distribution for the length of the tear-drop loop as a function of the number of base pairs along the chain for different forces [24].

B. Kinked loop with constrained ends

Calculations for the 90° kinked loop including ends under tension are given in Appendix B 1. We simply list the results here. For the 90° kinked loop, $\beta = 2.356\,19$ rad, and the shape parameter $\lambda_0 = 1.011\,69$ (see Appendix B), and $E = 4.9963 A/l$ in the $L/l \to \infty$ limit. Compared to the energy for a teardrop loop of the same length, the energy has been reduced by a factor of almost 3.

In the kinked loop case, the distance $\epsilon = 0.22\lambda$ and the work done by the force in pulling the ends of the DNA through this distance must be taken into account. The angle between the tangent vector and the X axis at the base of the loop is asymptotically $\alpha = 0.954\,13$ rad. The "opening angle" of this loop is therefore 1.2333 rad.

The probability distribution function is

$$p(n) = 28.01 \frac{\delta^3}{b^3} n^{-5} e^{-8\lambda_0 \sin^2(\beta/4)/n + 0.492n} e^{-2\times 1.22 ng(\beta A f)}. \tag{10}$$

The factor 1.22 in the last term takes into account the change in extension due to both the formation of the loop and moving the ends through the distance ϵ. The plots in Fig. 7 show the probability distribution function as a function of the size of the loop (in base pairs) for different values of applied

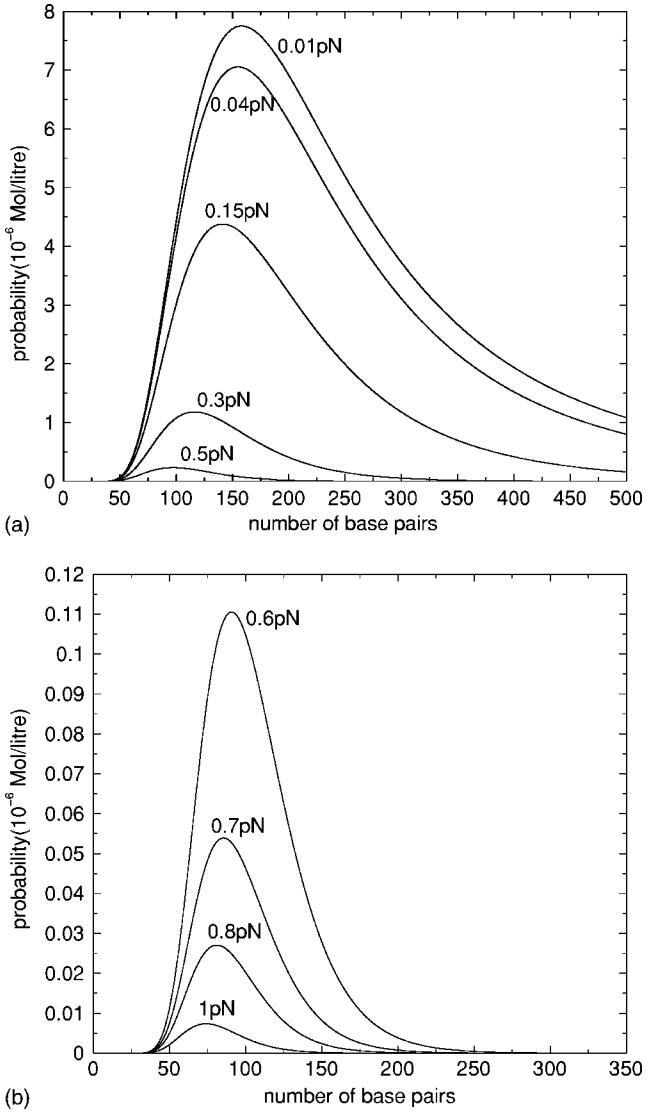

FIG. 7. Kinked loop: Probability distribution for the length of the kinked loop as a function of the number of base pairs along the chain for forces between 0.01 and 1 pN.

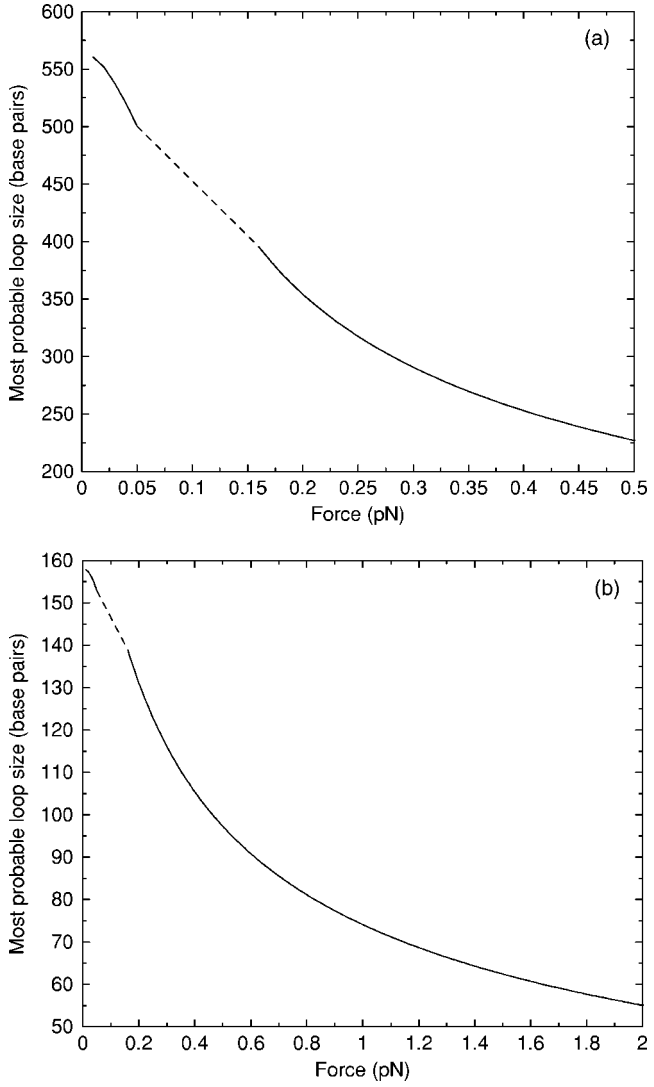

FIG. 8. Most probable loop size for (a) the teardrop loop and (b) the kinked loop as a function of the applied tension in piconewtons.

tension. The lower bending energy of the kinked loop causes the peak of the distribution function to occur at smaller loop sizes than in the case of a tear-drop loop.

V. RESULTS FOR LOOP FORMATION ALONG DNA UNDER TENSION

A. Nonspecific loops

The plots in Figs. 8 and 9 show the most probable loop size and the logarithm of the probability at the peak, as a function of the applied tension. From Fig. 8 we observe that the most probable loop size for teardrop loops is larger than that for kinked loops. The higher bending energy causes the teardrop loop to be nearly three times larger than the kinked loop in the low force (bending-energy-dominated) regime. In both cases, the loop size decreases with increasing force. As the applied force is increased, the work done by the applied force in bringing the ends of the DNA closer while forming the loop starts dominating over the bending energy.

Figure 9 shows the logarithm of the probability at the peak of the probability distributions shown in Figs. 6 and 7, as a function of force. At 0.01 pN, the most probable loop size for the teardrop loop is 550 base pairs which occurs with a probability of the order of 10^{-7} mol/l [24]. For the kinked loop at the same force the most probable loop of size 160 base pairs occurs with a probability of the order of 10^{-5} mol/l. This means that the rate of formation of a kinked loop is much higher than the rate of formation of the teardrop loop. For the kinked loop, the rate at which the value of the probability at the peak decreases is much slower than that for the teardrop loop, showing that the teardrop loop is more sensitive to force.

To see the effect of force on the formation of nonspecific loops more clearly we plot $\int_0^L p(l) l\, dl$ as a function of force. This quantity is proportional to the rate at which length is absorbed into loops during the initial folding up of a long DNA chain. This way of estimating the dynamics of loop

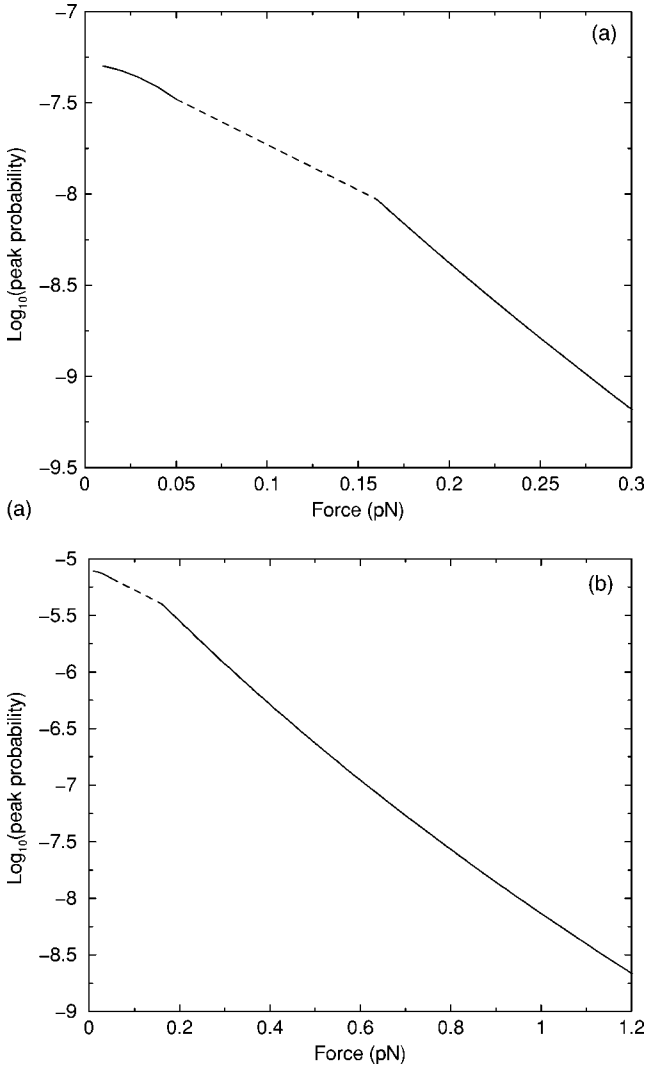

FIG. 9. Logarithm of the probability at the peak as a function of the applied tension in piconewtons: (a) teardrop loop and (b) kinked loop. The interpolation between the low and high force regimes is plotted in dotted lines.

formation is valid at large forces where the energy barrier for the formation of the loop is greater than a few k_BT. The energy barrier for the loop formation is the sum of the free energy for loop formation and the work done by the force. At forces above 0.1 pN, the energy barrier is a few k_BT. This means that at forces higher than 0.1 pN, the rate at which the ends of the polymer are retracted is dominated by the free energy and the work done by the force and the probability distribution estimated from these quantities can be used to calculate the rate.

From the graph shown in Fig. 10, we see that the speed of nonspecific loop formation decays with increasing force as a stretched exponential. This slower decay occurs during nonspecific folding because loops of different lengths can form in contrast to specific folding where loops have a fixed length. The asymptotic form for the decay is $e^{-\sqrt{(f-f^*)/f_0}}$, where f^* and f_0 are characteristic force scales of the stretched exponential (this is easily shown using a steepest descent calculation). The values of f_0 and f^* are higher for

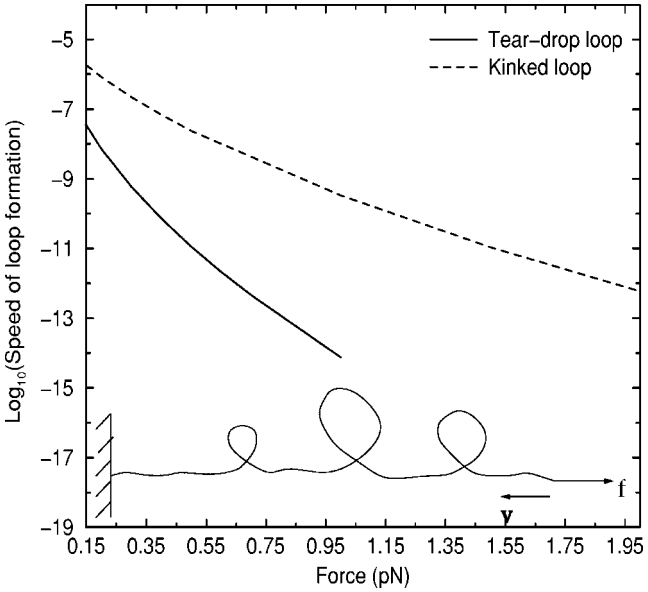

FIG. 10. Plot of the logarithm of the rate at which length is absorbed into nonspecific loops. The Y axis is $\log_{10}[\int_0^L dl\ lp(l)]$; the integral is proportional to the rate at which length is retracted during the formation of the first few loops along a long DNA molecule. Shown as an inset is a sketch of the process by which the DNA molecule is folded into nonspecific loops of different sizes when it is being stretched by force. The quantity v is the integral shown above and f is the applied tension.

loops with lower bending energy (i.e., loops that are easier to form). We find that for a teardrop loop $f_0=0.0051$ pN and $f^*=0.0012$ pN, while for a kinked loop $f_0=0.0134$ pN and $f^*=0.0372$ pN. Preliminary results for single-DNA micromanipulation experiments in the presence of proteins which nonspecifically loop the DNA show the existence of characteristic force scales for loop formation [25].

B. Specific loops

When loops are formed between specific sequences on the DNA chain, the length of the loop is fixed (it is the distance between the two sequences along the chain). An example is shown in Fig. 1(b). For such specific loops, it is useful to study the behavior of the probability distribution as a function of the applied force. These plots are shown in Figs. 11 and 12 for the two kinds of loops.

The low force regime is dominated by the bending energy. Since the bending energy of larger loops is lower, they dominate in the low force regime. As the force increases, the bending energy competes with the work done by the force in pulling the ends of the DNA together while forming the loop. The crossover from the bending-energy-dominated to the force-dominated regime is indicated by the curves crossing each other around 0.3 pN. At forces higher than 0.3 pN, the work done by the force is dominant and favors loops of smaller size (200 base pairs in the plot shown in Fig. 11). The same trend is observed for a kinked loop as shown in Fig. 12.

VI. CONCLUSIONS

This paper describes the formation of loops in a DNA molecule under tension. We can estimate the most probable

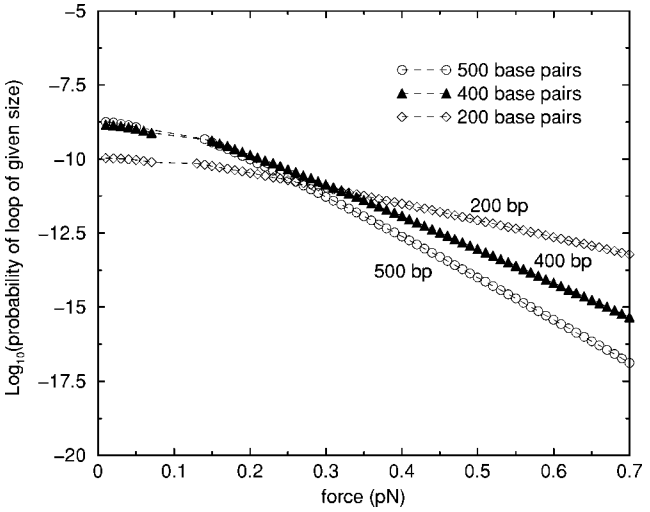

FIG. 11. Probability of forming a teardrop loop of a specific length as a function of the applied tension in piconewtons. The curves cross each other around 0.3 pN, above which the smallest loop is the most probable. The interpolation between the low and high force regimes is plotted in dotted lines.

loop size at a given force for a loop with and without a kink at its apex. It is found that the kink (which may be produced due to the presence of a DNA-bending protein) substantially reduces the bending energy of the loop. Hence, at a given force, loops with kinks are smaller than those without kinks. For example, at a force of 0.15 pN, the teardrop loop has a size of 400 base pairs while the kinked loop has a size of 130 base pairs. The value of the probability at the peak is higher for kinked loops than for teardrop loops which indicates that these loops are formed faster.

This analytic calculation is designed to obtain the correct position of the peak in the probability distribution at zero force (480 base pairs). Based on this, we are able to predict

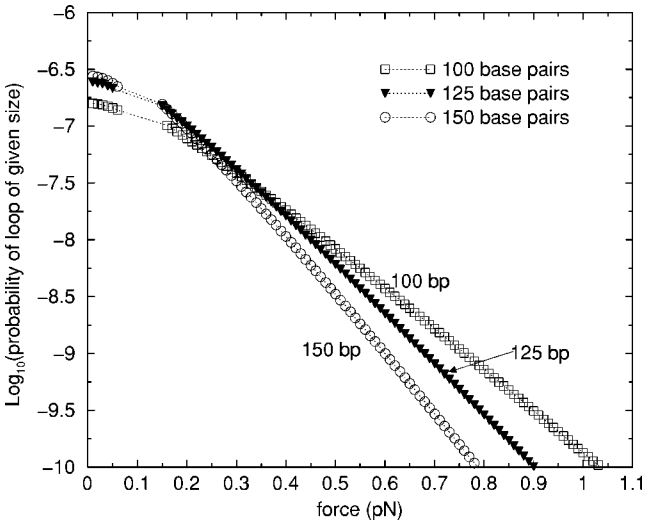

FIG. 12. Probability of forming a kinked loop of a specific length as a function of the applied tension in piconewtons. The interpolation between the low and high force regimes is plotted in dotted lines.

the distribution for kinked loops; we find that a 90° kink reduces the peak position to around 150 base pairs at zero force. Another prediction that can be made is the shift of the peak position with force and an estimate of the force range below which loop formation is favored. From Figs. 6 and 7, for nonspecific loops, we infer that the most probable loop size decreases with increasing force. The teardrop loops are more sensitive to increase in force and are suppressed at lower forces compared to kinked loops. Very recently, experiments which show anomalously large loop formation on very short (100 base pairs) DNA [26] have led to the proposal that thermally excited defects in the double-helix structure might allow similar sharp bends to occur in the absence of bending proteins [23]; our calculations might be applied to this situation as well.

We also estimate the rate at which the ends of a DNA chain are pulled in by nonspecific loop formation. The estimate for this process is valid at forces higher than 0.1 pN. At lower forces, the free energy cost of formation of loops in the few hundred base-pair range is low enough that the dynamics of the loop formation process becomes dominated by the dynamics of the two ends of the polymer randomly approaching each other within a distance of δ. The plot of the speed of retraction by nonspecific loop formation is shown as a function of force in Fig. 10. The speed of retraction for both types of loops decays as a stretched exponential function of the force. The reduction in the speed of retraction with increase in force is stronger for teardrop loops.

We also estimate the probability (equivalently, the rate) of formation of specific loops of a given length as a function of the applied force. Shorter loops are more probable at high force while the longest ones are favored at forces below 0.3 pN.

ACKNOWLEDGMENT

This research was supported by the National Science Foundation through Grant No. DMR-0203963.

APPENDIX A: BENDING ENERGY FOR LOOPS WITH FREE ENDS

Following YS [18], we start by minimizing the energy in Sec. II subject to the constraint that the loop is closed. We will minimize the energy for one half of the loop and will get the full shape of the loop by reflecting this half along the Y axis. The x and y coordinates of any point on the polymer are given by $x(s)=\int_0^s \hat{u}\cdot\hat{e}_x$ and $y(s)=\int_0^s \hat{u}\cdot\hat{e}_y$, where $\hat{e}_x$ and $\hat{e}_y$ are the unit vectors along the X and Y axes, respectively.

The condition of loop closure implies that the total distance traveled by the polymer along the X axis must be zero at the apex of the loop:

$$\int_0^{l/2} \hat{u}\cdot e_x = 0. \tag{A1}$$

The angle between the tangent vector and the X axis is $\omega(s)$. Then, $\hat{u}=\hat{e}_x\cos(\omega)+\hat{e}_y\sin(\omega)$, $(d\hat{u}/ds)^2=(d\omega/ds)^2$.

The bending energy in Eq. (1) should be minimized with respect to the constraint in Eq. (A1) using a Lagrange multiplier f. The function (including the constraint) to be minimized is

$$S = \frac{1}{2}\int_0^{l/2} ds\left[A\left(\frac{d\omega}{ds}\right)^2 - f\cos\omega\right]. \tag{A2}$$

Computing $\delta S = 0$ yields the equation of motion

$$-A\frac{d^2\omega}{ds^2} + f\sin\omega = 0. \tag{A3}$$

To get the trajectory of points on the polymer we evaluate Eq. (A3) using the initial values for angle and curvature as boundary conditions. Hence, at $s=0$,

$$\omega(s=0) = \alpha;\ \left.\frac{d\omega}{ds}\right|_{(s=0)} = \alpha'. \tag{A4}$$

Integrating the equation of motion once, we get

$$\frac{A}{2}\left(\frac{d\omega}{ds}\right)^2 = -f\cos\omega + C_1. \tag{A5}$$

C_1 is determined using the boundary conditions to be $C_1 = (A/2)\alpha'^2 + f\cos\alpha$. We change variables to $v = (\pi-\omega)/2$ and let $k^2 = 2f/(C_1+f)$. Equation (A5) becomes

$$-2\frac{dv}{ds} = \pm\left[\frac{2(C_1+f)}{A}\right]^{1/2}(1-k^2\sin^2 v)^{1/2}. \tag{A6}$$

Integration with respect to the variable v yields the trajectory

$$-2\int_{(\pi-\alpha)/2}^{[\pi-\omega(s)]/2}\frac{dv}{(1-k^2\sin^2 v)^{1/2}} = \pm\left[\frac{2(C_1+f)}{A}\right]^{1/2}\int_0^s ds. \tag{A7}$$

The sign on the right-hand side of the equation above is chosen so that $\omega(s)$ always increases with increasing s [or $v(s)$ decreases with increasing s]. In terms of elliptic integrals the trajectory is

$$F\left(\frac{\pi-\omega(s)}{2},k\right) = F\left(\frac{\pi-\alpha}{2},k\right) - \left[\frac{2(C_1+f)}{A}\right]^{1/2}\frac{s}{2}. \tag{A8}$$

Let the angle made by the tangent vector at the top of the loop be β[i.e., $\omega(s=l/2)=\beta$]. Therefore, given the initial values α and α', the angle β (as a function of f) can be determined as

$$\beta = \pi - 2\ \mathrm{am}\left[F\left(\frac{\pi-\alpha}{2},k\right) - \left(\alpha'^2 + \frac{4f}{A}\cos^2\alpha/2\right)^{1/2}\frac{l}{4},k\right], \tag{A9}$$

where "am" is the Jacobi amplitude. The "am" is the inverse of the elliptic integral of the first kind, $\mathrm{am}[F(\theta,k),k]=\theta$.

Alternatively, we could use α and β as the boundary conditions rather than α and α'. When β is specified instead of α', we treat Eq. (A9) as an equation in the two variables α' and f. For the teardrop loop, the angle $\beta=\pi$ rad. In the kinked loop, $\beta=2.3562$ rad, corresponding to producing a 90° kink between the two arms of the DNA chain.

To solve for α' and f we need one more equation in these two variables. This is the equation for the constraint given by Eq. (A1). This can be written as

$$\frac{-2}{[\alpha'^2 + (4\lambda/A)\cos^2\alpha/2]^{1/2}}\int_{F((\pi-\alpha)/2,k)}^{F((\pi-\beta)/2,k)} du\ \mathrm{sn}^2 u = \frac{l}{4}, \tag{A10}$$

where the function "sn" is the Jacobi sine amplitude. Using the identity in [27], the constraint equation is

$$\frac{2}{k^2[\alpha'^2 + (4f/A)\cos^2\alpha/2]^{1/2}}\left[E\left(\frac{\pi-\beta}{2},k\right) - E\left(\frac{\pi-\alpha}{2},k\right) + \frac{l}{4}\left(\alpha'^2 + \frac{4f}{A}\cos^2\alpha/2\right)^{1/2}\right] = l/4. \tag{A11}$$

We now solve Eqs. (A9) and (A11) for the unknown variables α' and f numerically. The energy of the loop is given by

$$\frac{A}{2}\int_0^l ds\left(\frac{d\omega}{ds}\right)^2 = -f\int_0^l ds\cos\omega(s) + C_1\int_0^l ds = C_1 l, \tag{A12}$$

using Eq. (A1).

APPENDIX B: BENDING ENERGY FOR LOOPS WITH CONSTRAINED ENDS

Before proceeding, we will rescale the length variable in our equation to absorb the Lagrange multiplier f and the persistence length A so that the equations of motion in the rescaled coordinates can be easily solved. In the rescaled coordinates, we denote the total length of the DNA chain by 2Λ and the length of the loop as 2λ. Using the transformation $\sigma = s/s_0$, where $s_0 = \sqrt{A/f}$, the equation of motion Eq. (A3) becomes

$$-A\frac{d^2\omega}{d\sigma^2} + \sin\omega = 0. \tag{B1}$$

Integrating this we get

$$\frac{1}{2}\left(\frac{d\omega}{d\sigma}\right)^2 + \cos\omega = C. \tag{B2}$$

Changing variables to $\psi = (\pi-\omega)/2$,

$$\left(\frac{d\psi}{d\sigma}\right)^2 = \frac{C+1}{2}\left(1 - \frac{2}{C+1}\sin^2\psi\right). \tag{B3}$$

Since $\omega(\sigma)$ increases as σ increases,

$$\left(\frac{d\psi}{d\sigma}\right) = -\sqrt{\frac{C+1}{2}}\sqrt{\left(1 - \frac{2}{C+1}\sin^2\psi\right)}. \tag{B4}$$

Using a second scaling transformation $\Sigma = \sqrt{[(C+1)/2]}\sigma$,

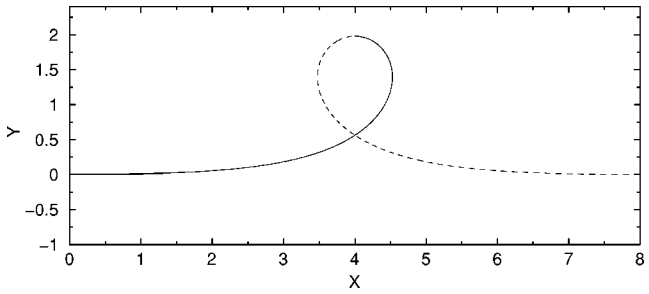

FIG. 13. Trajectory of the teardrop loop in the presence of applied tension in the asymptotic limit.

$$\left(\frac{d\psi}{d\Sigma}\right) = -\sqrt{(1-k^2\sin^2\psi)}, \tag{B5}$$

where $k^2=2/(C+1)$. The trajectory of the DNA chain is the solution of this equation with the appropriate boundary conditions. The trajectory depends only on the parameter k. The angle made by the tangent vector at $\Sigma=0$ is $\alpha=0$. Therefore,

$$F\left(\frac{\pi-\omega}{2},k\right) = F\left(\frac{\pi}{2},k\right) - \Sigma. \tag{B6}$$

This means that

$$\omega(\Sigma) = \pi - 2\ \mathrm{am}\left[F\left(\frac{\pi}{2},k\right) - \Sigma, k\right]. \tag{B7}$$

$\omega(\Lambda)=\pi$ corresponds to the teardrop loop and $\omega(\Lambda)=\beta\neq\pi$ gives the trajectory of a loop with a kink. The parameter k determines a family of solutions characterized by the ratio of the length of the loop to the total length of the molecule in each case. As $k\to 1$, the length of the polymer becomes infinite. As $k\to 0$, the length of the polymer approaches zero. These solutions are those obtained in the "elastica" problem and for the nonlinear pendulum [21,28].

1. Teardrop loop in the presence of tension

In the absence of DNA-bending proteins, the angle $\omega=\pi$ at half the total length of the chain (i.e., when $\Sigma=\Lambda$). Therefore,

$$\Lambda = F\left(\frac{\pi}{2},k\right). \tag{B8}$$

From this expression we see that as $k\to 1$, $\Lambda\to\infty$. Thus, the long-chain limit ($L\to\infty$) can be obtained by taking k close to 1. The trajectory of the loop in the asymptotic limit is shown in Fig. 13. We pause here to note that the total length of the DNA molecule in physical units is

$$L = 2\sqrt{\frac{A}{f}}k\Lambda. \tag{B9}$$

We also note that the Lagrange multiplier f can be physically interpreted as the tension in the chain, in the absence of thermal fluctuations. The effect of thermal fluctuations has been taken into account in computing $g(\beta Af)$.

The loop length in the scaled coordinates, 2λ, is determined as follows. The constraint of loop closure yields $\int_{\Lambda-\lambda}^{\Lambda} d\Sigma\ \cos\ \omega(\Sigma)=0$, to be solved for λ. This integral can be written as

$$\int_{\lambda}^{0} du\ \cos[2\ \mathrm{am}(u,k)] = 0, \tag{B10}$$

where $u=F((\pi-\omega)/2,k)=F(\pi/2,k)-\Sigma$. Using the identity in [27] we get the equation for λ as

$$\lambda(2-k^2) = 2E(\mathrm{am}(\lambda,k),k). \tag{B11}$$

In the asymptotic limit $k^2\to 1$ or $\Lambda\to\infty$. The solution to this equation gives a finite value of λ. Numerically, the asymptotic limit is for λ is $\lambda_0=1.915\ 008$. Since $E(\phi,k)$ is analytic as $k\to 1$ [29], λ behaves analytically close to its asymptotic value and can be expanded as a power series around the value at $k^2=1$. The values of the first and higher derivatives of λ with respect to k^2 can be obtained numerically at points close to $k^2=1$. This enables us to write the power series as $\lambda(k)=\lambda_0+(k^2-1)\lambda_1+(k^2-1)^2\lambda_2+\cdots$, where $\lambda_0=1.915\ 008$, $\lambda_1=0.669\ 85$, and $\lambda_2=0.8$ can be computed numerically. The length Λ can also be written as a series around $k^2=1$ [27] as

$$\Lambda(k) = \ln\frac{4}{(1-k^2)^{1/2}} + \frac{1-k^2}{4}\left[\ln\frac{4}{(1-k^2)^{1/2}} - 1\right] + \cdots. \tag{B12}$$

Therefore,

$$\frac{l}{L} = \frac{\lambda}{\Lambda} \approx \frac{\lambda_0}{\ln[4/(1-k^2)^{1/2}]}. \tag{B13}$$

This enables us to relate k^2 to the physical lengths (in the asymptotic limit) through the relation $k^2\approx 1-16e^{-2\lambda_0 L/l}$.

The end-to-end extension for one half of the loop, $x(\Lambda)$, shown in Fig. 5, can be evaluated using the identity in [27]:

$$x(\Lambda) = \int_0^{\Lambda} d\Sigma\ \cos\ \omega(\Sigma) = \left(\frac{2}{k^2}-1\right)F\left(\frac{\pi}{2},k\right) - 2E\left(\frac{\pi}{2},k\right). \tag{B14}$$

The total end-to-end extension (measured between the points A and B shown in Fig. 5) is twice the above value. The ratio of the change in extension (from the maximum of Λ) to the loop length λ is asymptotically

$$\frac{x(\Lambda)-\Lambda}{\lambda} \approx -\frac{2}{\lambda_0} + \frac{16L}{l}e^{-\lambda_0 L/l}. \tag{B15}$$

Figure 5 shows the ends of the polymer pulled in during the formation of the loop. The relative change in extension calculated above gives us an estimate of the distance ϵ defined as $x(\Lambda)-\Lambda+\lambda=-\epsilon$. For the teardrop loop, this distance is only a small fraction of the loop size, $\epsilon=0.04\lambda$. Thus, ϵ does not contribute significantly to the work done by the external force in folding the DNA into a loop. The angle between the tangent vector and the X axis at the base of the loop asymptotically approaches $\alpha=0.5852$ rad. The "opening angle" of the loop, which is defined as the angle between the two arms of the loop at the base of the loop, is 1.9712 rad.

We can evaluate the energy exactly in the scaled coordinates. The energy for half the total length of the loop is

$$E_s = \frac{1}{2}\int_0^{\Lambda} d\Sigma \left(\frac{d\omega}{d\Sigma}\right)^2$$

$$= Ck^2 \int_0^{\Lambda} d\Sigma - k^2 \int_0^{\Lambda} d\Sigma \cos\,\omega$$

$$= (2-k^2)\Lambda - k^2 x(\Lambda)$$

$$= 2E\left(\frac{\pi}{2},k\right). \tag{B16}$$

The energy for the full loop is twice the above energy. Asymptotically, we can express this energy as a function of the physical lengths L and l as

$$E_s \approx 2 + \frac{16\,\lambda_0 L}{l} e^{-2\,\lambda_0 L/l}. \tag{B17}$$

The relation between this scaled energy and the total energy E in the original coordinate system is

$$E = 4\frac{A\,\lambda}{l} E_s. \tag{B18}$$

2. Kinked loop in the presence of tension

We now extend the calculations of the previous subsection for a loop that has a kink due to the presence of DNA bending proteins. As we will show below, the presence of the kink lowers the bending energy. The angle made by the tangent vector with the X axis at the apex of the loop is $\beta = 135°$. This corresponds to having the arms of the DNA chain on either side of the apex bent through 90°.

The length of one half of the loop in the scaled coordinates is

$$\Lambda = F\left(\frac{\pi}{2},k\right) - F\left(\frac{\pi-\beta}{2},k\right). \tag{B19}$$

As $k^2 \to 1$, $\Lambda \to \infty$. The trajectory is given by Eq. (B7), where the arclength variable Σ takes values from 0 to Λ, defined above. The trajectory of the kinked loop in the asymptotic limit is shown in Fig. 14. The length λ satisfies the equation

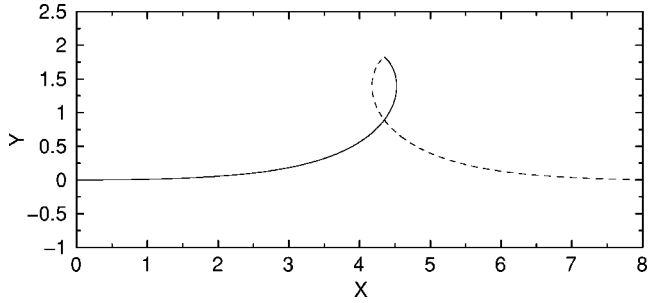

FIG. 14. Trajectory of the kinked loop in the presence of applied tension in the asymptotic limit. Note the large amount of length associated with the bend to meet the loop.

$$\lambda(k^2-2) = 2\left\{ E\left(\frac{\pi-\beta}{2},k\right) - E\left(\mathrm{am}\left[\lambda + F\left(\frac{\pi-\beta}{2},k\right),k\right],k\right)\right\} \tag{B20}$$

and has a finite value $\lambda_0 = 1.011\,69$ in the asymptotic limit. λ can be expanded as a powerseries around its asymptotic value and corrections away from $k^2 = 1$ can be calculated. We can also show that $(1-k^2)^{1/2} \approx 4\,\tan(\beta/4) e^{-\lambda_0 L/l}$ asymptotically [29].

The energy for half the total length of the loop (in the scaled units) is

$$E_s = 2\left[E\left(\frac{\pi}{2},k\right) - E\left(\frac{\pi-\beta}{2},k\right)\right]$$

$$\approx 4\,\sin^2\frac{\beta}{4} + 16\,\tan^2\frac{\beta}{4}$$

$$e^{-2\lambda_0 L/l}\left[\frac{\lambda_0 L}{l} - \frac{1}{2} + \cos(\beta/4)\right] \quad \text{as } \Lambda \to \infty. \tag{B21}$$

In the asymptotic limit, the energy in physical units for a loop of length l is

$$E \approx \frac{16\lambda_0 A}{l}\,\sin^2(\beta/4). \tag{B22}$$

The relative change in end-to-end in extension in one half of the loop is

$$\frac{x(\Lambda)-\Lambda}{\lambda} = \frac{2\Lambda\left(\frac{1}{k^2}-1\right) - \frac{2}{k^2}\left[E\left(\frac{\pi}{2},k\right) - E\left(\frac{\pi-\beta}{2},k\right)\right]}{\lambda}$$

$$\approx \frac{-4\,\sin^2\beta/4}{\lambda_0}. \tag{B23}$$

[1] I. Y. Goryshin and W. S. Reznikoff, J. Biol. Chem. **273**, 7367 (1998).
[2] E. M. Blackwood and J. T. Kadanoga, Science **281**, 60 (1998).
[3] L. Finzi and J. Gelles, Science **267**, 378 (1995).
[4] G. Lia *et al.*, Proc. Natl. Acad. Sci. U.S.A. **100**, 11373 (2003).
[5] N. J. Trun and J. F. Marko, ASM News **64**, 276 (1998).
[6] T. Strick, T. Kawaguchi, and T. Hirano, Curr. Biol. **14**, 874 (2004).

[7] I. M. Kulić and H. Schiessel, Biophys. J. **84**, 3197 (2003).
[8] J. Yan, D. Skoko, and J. F. Marko, Phys. Rev. E **70**, 011905 (2004).
[9] J. F. Marko and E. D. Siggia, Biophys. J. **73**, 2173 (1997).
[10] J. Rudnick and R. Bruinsma, Biophys. J. **76**, 1725 (1999).
[11] H. Diamant and D. Andelman, Phys. Rev. E **61**, 6740 (2000).
[12] S. Cocco, J. F. Marko, R. Monasson, A. Sarkar, and J. Yan, Eur. Phys. J. E **10**, 249 (2003).
[13] J. Yan and J. F. Marko, Phys. Rev. E **68**, 011905 (2003).
[14] D. Skoko, B. Wong, R. C. Johnson, and J. F. Marko, Biochemistry **43**, 13867 (2004).
[15] J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, and R. T. Dame, Proc. Natl. Acad. Sci. U.S.A. **101**, 6969 (2004).
[16] S. Jones, P. van Heyningen, H. M. Berman, and J. M. Thornton, J. Mol. Biol. **287**, 877 (1999).
[17] J. Shimada and H. Yamakawa, Macromolecules **17**, 689 (1984).
[18] H. Yamakawa and W. H. Stockmayer, J. Chem. Phys. **57**, 2843 (1972).
[19] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).
[20] H. Jacobson and W. H. Stockmayer, J. Chem. Phys. **18**, 1600 (1950).
[21] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity*, 4th ed. Dover Publications, New York, 1927.
[22] A. Podtelezhnikov and A. Vologodskii, Macromolecules **30**, 6638 (1997).
[23] J. Yan and J. F. Marko, Phys. Rev. Lett. **93**, 108108 (2004).
[24] The most probable loop size at 0.01 pN is slightly larger than in the zero-force case. At such low forces, the effects of the ends of the molecule are not yet significant and hence the model used to calculate the probability distribution at such low forces should not be expected to give very accurate quantitative results. However, as seen from the figures at higher forces, the trend shifts toward smaller loops with increasing force, leading us to believe that the calculation is useful in predicting loop distributions at forces greater than 0.1 pN.
[25] D. Skoko (private communication).
[26] T. E. Cloutier and J. Widom, Mol. Cell **14**, 355 (2004).
[27] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic Press, New York, 1980).
[28] R. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, MA, 1980).
[29] The asymptotic expansions for the elliptic functions of the first and second kinds (as $k^2 \to 1$) for an arbitrary angle ϕ are

$$F(\phi,k) = \frac{1}{2}\ln\frac{1+\sin\phi}{1-\sin\phi} - \frac{1-k^2}{2}\left(\frac{\sin\phi}{2\cos^2\phi} - \frac{1}{4}\ln\frac{1+\sin\phi}{1-\sin\phi}\right) + \cdots,$$

$$E(\phi,k) = \sin\phi + \frac{1-k^2}{2}\left(-\sin\phi + \frac{1}{2}\ln\frac{1+\sin\phi}{1-\sin\phi}\right) + \cdots.$$