Force and kinetic barriers to initiation of DNA unzipping

Simona Cocco,[1,*] Rémi Monasson,[2,†] and John F. Marko[1]
[1]*Department of Physics, The University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607*
[2]*The James Franck Institute, The University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637*
(Received 28 August 2001; published 28 March 2002)

A semimicroscopic model of the binding of the two nucleotide strands in a double-stranded DNA is used to describe the effects of applied tension on strand unpairing. We show that the model describes strand separation by elevated temperature, applied torque, and applied force. In particular, we show how the interactions responsible for stabilizing the double helix against thermal denaturation determine the ≈12 pN force threshold for DNA strand separation. The larger rigidity of the strands when formed into double-stranded DNA, relative to that of isolated strands, gives rise to a potential barrier to unzipping. We show that this barrier results in a ≈250 pN force barrier opposing the beginning of strand separation. The thermal-fluctuation-assisted ''tunnelling'' through the barrier is then analyzed using instanton calculations. The resulting kinetics of unzipping initiation is shown to be consistent with solution-phase strand dissociation experiments, and can explain results of two recent unzipping experiments done using atomic-force microscopy.

DOI: 10.1103/PhysRevE.65.041907 PACS number(s): 87.15.−v, 05.70.Ln, 64.60.Qb

I. INTRODUCTION

Genetic information in cells is stored in double-helical *B*-form double-stranded DNAs (dsDNA). The hydrogen bonded and stacked bases are well protected in the interior of dsDNA. However, separation of the two strands is a key part of both the ''reading'' of DNA sequences (transcription) and DNA replication [1]. During transcription, a transient ''bubble'' of single-stranded DNA (ssDNA) is formed, to allow the enzyme that makes a RNA copy of the DNA sequence to access the DNA bases. During replication, the two DNA strands are permanently separated, each strand then serving as a template for the synthesis of a new strand.

In biochemistry experiments, dsDNAs are routinely converted to separated ssDNAs by ''melting'' at elevated temperature ≈80 °C; this approach uses thermal fluctuations to simply overwhelm the k_BT scale base pairing and base-stacking interactions that stabilize the double helix at room temperature. This is not how strand separation is driven in cells. Instead, DNA strands are separated by the application of force, or in chemical terms, by enzymes whose interactions with DNA make strand separation thermodynamically favorable at ambient temperature.

Therefore, it is of basic biophysical interest to analyze the separation of DNA strands by force. A further motivation is provided by recent micromanipulation experiments, which accomplish precisely this feat. Essevaz-Roulet, Bockelman, and Heslot have shown that the two strands of a dsDNA can be pulled apart if a force ≈12 pN is applied [2]. Fluctuations of the ''unzipping'' force about this mean corresponded to DNA sequence: slightly higher forces were shown to correspond to DNA regions with higher *GC* densities. This result was in accord with the fact that *GC*-rich sequences have stronger base-pairing interactions than *AT*-rich sequences, causing them to thermally melt at higher temperatures [3].

The micromanipulation experiments of Essevaz-Roulet, Bockelman, and Heslot use optical microscopy, detecting force via observation of the deflection of a glass fiber to which one of the DNA strands is attached. That experiment does not observe the beginning of base unpairing since the distance between the two single strands extremities cannot be controlled on the Angström scale. However, similar experiments using atomic-force microscope (AFM) cantilevers are, in principle, able to probe the unzipping of the first few base pairs. The initial opening of a ssDNA bubble can be expected to involve a large free energy barrier due to the large cooperativity of DNA strand separation known from thermal studies of melting. Thus, one can anticipate that the double helix will be stable against unzipping forces larger than the macroscopic unzipping threshold force of 12 pN, for some finite amount of time.

In this paper, we present a semimicroscopic theory that allows these effects to be described [4]. We begin by reviewing the current experimental situation (Sec. II), and by discussing our main theoretical results relevant to those experiments (Sec. III). We then discuss the basic features of DNA unzipping using a simple mean-field theory (Sec. IV), including prediction of the modification of the 12 pN unzipping threshold by application of torque.

In Sec. V, we present a theory to describe unzipping at the base-pair scale, based on previous work [5–7] that studied strand separation by temperature and torque. In Sec. VI, we analyze the equilibrium properties of this model, and show that the ≈12 pN threshold for complete unzipping emerges in our detailed model with no fitting or calibration beyond that done previously for temperature-torque unzipping [5], and that the basic unzipping behavior of our detailed model matches that of the simple mean-field theory.

In Sec. VII, we focus on the use of this model to analyze the force-distance behavior of the initial stages of unzipping; we find a large force barrier opposing initial double helix unzipping of ≈250 pN height. Although this barrier is high,

*Permanent address: CNRS, LDFC–Institut de Physique, 3 rue de l'Université, 67000 Strasbourg, France.
†Mailing address: CNRS–Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75231 Paris Cedex 05, France.

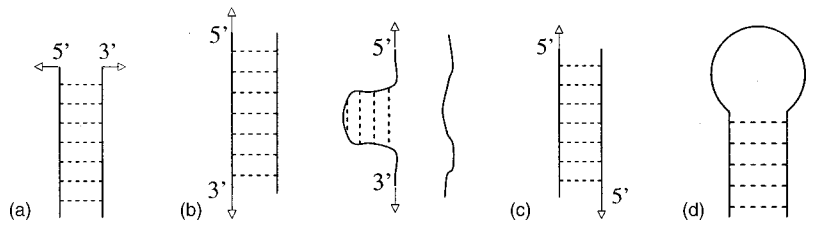

FIG. 1. Sketch of unzipping and stretching experiments; arrows indicate applied force. (a) Adjacent $5'$-$3'$ unzipping experiment of Bockelmann, Essevaz-Roulet, and Heslot [2]. (b) Untethered strand removal, followed by annealing and subsequent unzipping of hairpins as studied by Rief, Classen-Schaumann, and Gaub [11]. (c) Opposite $5'$-$5'$ unzipping experiment of Strunz *et al.* [10]. (d) Spontaneous thermal opening of hairpin under zero force as studied by Bonnet, Krichevsky, and Libchaber [9].

it is also of base-pair dimensions in width, and can be crossed by thermal fluctuation on experimentally relevant timescales. The basic kinetic theory for this is presented in Sec. VIII, and then the calculations of kinetics of unzipping initiation are presented in Sec. IX. The connection of these results to studies of spontaneous strand dissociation observed for molecules in solution [8,9], and to the kinetics of unzipping as probed by AFM experiments [10,11], are discussed in Sec. X. Conclusions are presented in Sec. XI.

II. DNA UNZIPPING EXPERIMENTS

In the experiment of Essevaz-Roulet, Bockelmann, and Heslot [2], the strands of a λ-phage DNA (the chromosome of a virus that infects the bacterium Escherichia coli, of contour length ≈ 16 μm, $\approx 48\,500$ base pairs (bp)) in aqueous buffer [phosphate buffered saline (PBS) at pH 7, 150 mM NaCl, $T = 25\,^{\circ}\mathrm{C}$] are pulled apart [Figure 1(a)]. One strand was attached to a glass microscope slide and the other to a ≈ 3-μm-diameter polystyrene bead. The tip of a glass microneedle is attached to the bead, and serves the force-measuring cantilever. Bending of the needle is simply observed using the microscope on which the experiment is assembled; calibration of the needle stiffness allows the observed deflection to be converted to force.

The DNA is then forced to open by a lateral displacement of the microscope slide, and the needle deflection is obtained by analysis of video frames. The stiffness of the microneedle, $k = 1.7$ pN/micron, and the translational velocity, $v = 40$ nm/s, result in a loading rate (rate of force increase) $\lambda = kv = 0.06$ pN/s. When the force reaches ≈ 12 pN the unzipping of the DNA begins. The force is measured as a function of the displacement, and varies between 10 and 15 pN depending on sequence. The experimenters estimate that they are able to resolve variations of base-pair sequence on the $\approx 10^2$ base-pair scale.

In AFM experiments of Rief, Clausen-Schaumann, and Gaub [11], λ-DNA (48 502 bp), poly(dA-dT) (5100 bp), and poly(dG-dC) (1260 bp) DNAs were studied, at room temperature in aqueous buffer (10 mM Tris 150 mM NaCl, 1 mM ethylenediamine tetra-acetic acid (EDTA), pH 8). These molecules were attached to a gold substrate and to an AFM tip, and then stretched longitudinally, at loading rates in the range of 200 to 6500 pN/s [Fig. 1(b)]. The molecules were observed to go through two structural transitions. The molecules with *GC* content [λ and poly(dG-dC)] underwent the well known sharp *B*-DNA to *S*-DNA transition [12,13] at ≈ 65 pN. They then went through a second highly loading-rate dependent and hysteretic strand-separation transition, at roughly 150 pN for λ and at roughly 250 to 350 pN for poly(dG-dC). Only the strand-separation transition was observed for the poly(dA-dT) molecule, at much lower forces ≈ 35 pN.

Thus Rief, Clausen-Schaumann, and Gaub found that ds-DNA could be converted to ssDNA by application of large longitudinal forces, with a force threshold correlated with *GC* content. Moreover, they studied the unzipping of the repeated-sequence poly(dA-dT) and poly(dG-dC) molecules, as follows. After force melting, tension on the molecules was relaxed, allowing them to reanneal into ''hairpin'' structures. These hairpins could occur prodigiously due to the repeated sequence. Then, the hairpins could be unzipped by reapplication of force. This was observed to occur at forces of 20 ± 3 pN for poly(dG-dG) and 9 ± 3 pN for poly(dA-dT). To date, these experiments provide the only calibrations for unzipping repeated sequences.

AFM experiments on much shorter molecules have been done by Strunz, Oroszlan, Schafer, and Guntherodt [10]. This group studied the unbinding forces between short complementary strands [Fig. 1(c)], including the dependence of unbinding on the force loading rate. Molecules of 10, 20, or 30 bp lengths with about 60% *GC* content were studied in PBS buffer. One $5'$, end was attached to the surface, and the other $5'$ end to the AFM tip. The first step of an experiment was to move the tip near the surface, so as to hybridize (bind) the two complementary strands. When a binding event was observed, the tip was then moved away from the surface to drive unbinding. Thus, this experiment studied unbinding using force applied to the $5'$ ends of the two strands at opposite ends of the molecule. The probability distribution of the rupture force is obtained for loading rates ranging from 16 to 4000 pN/s.

In addition to these mechanical experiments, experiments on molecules in free solution can also give information about unzipping kinetics at zero applied force. These experiments are usually on short (≈ 10 bp) dsDNAs, and usually focus on spontaneous, thermally activated unzipping [Fig. 1(d)]. The unzipping rate ν_- and closing rate ν_+ of a 5 bp DNA ''hairpin'' ($5'$-CCCAA-loop-TTGGG-$3'$) was investigated using a combination of fluorescence energy transfer and fluorescence correlation spectroscopy by Bonnet, Krichevsky, and Libchaber [9]. Use of a hairpin allows the study of unzipping

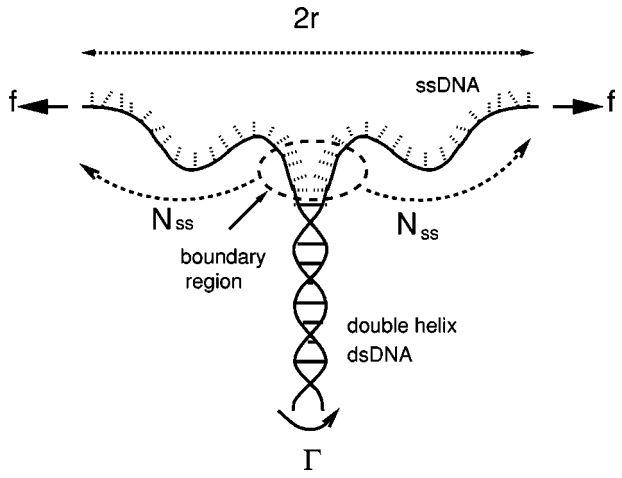

FIG. 2. A schematic view of a DNA molecule when unzipped by a force. The double-stranded DNA has paired bases, and a large strand rigidity. As a result of the mechanical stress, N_{ss} base pairs may separate, or unzip. The distance between the two single strand ends is defined as $2r$. The unzipped region of the molecule is two single-stranded DNAs with low rigidity and unpaired bases. The junction defines a boundary extending over $N^* \simeq 4$ partly opened base pairs with high strand rigidity, at the origin of the free energy barrier opposing the initiation of unzipping. Denaturation may also be the result of the application of torque. A positive torque Γ overtwists the double helix, and $\Gamma < 0$ untwists it.

in a unimolecular system, and is convenient for single-molecule spectroscopy. The loop of this hairpin was varied in length from 12 to 21 bases, and was either a T or A polymer. The buffer contained 0.1 M NaCl, 50 μM EDTA, 5 mM cacodylic acid at pH 7. The opening time (or lifetime of the closed state) was found to be $t_- = 1/\nu_- \simeq 0.5 \times 10^{-3}$ s for $N = 5$ bp, essentially independent of the loop length.

The recent single-molecule experiments of Bonnet, Krichevsky, and Libchaber are in good accord with results of older spectroscopic measurements on ensembles of molecules in solution. Of particular relevance to us, Pörschke [8] found the dissociation rate for poly(dA)-poly(dU) molecules to vary with molecular length N as $\nu_- = 10^{\alpha - \delta N}$ with $\delta = 0.5$ and $\alpha = 8$, for molecules of length $N = 8$ to 18. The method was a combination of temperature jump and time-resolved spectroscopy. This work also suggested that strand separation starts with a nucleation ''bubble'' $\approx$3 bp long.

III. PHYSICAL PICTURE AND OVERVIEW OF RESULTS

Unzipping can be described in terms of increasing the distance $2r$ between nucleotides originally belonging to the same base pair of a double helix, beyond the equilibrium value in the double helix of $2R = 20$ Å. This increase may be the result of a change in temperature or chemical conditions, or of direct mechanical action. In the case of mechanical unzipping studied in this paper, the control parameters may be the applied torque Γ or force f, the loading rate df/dt, or the half-distance r itself (Fig. 2). In this paper, we will discuss unzipping driven by changes in each of these parameters, using first macroscopic (essentially thermodynamic) and then microscopic (statistical-mechanical) perspectives.

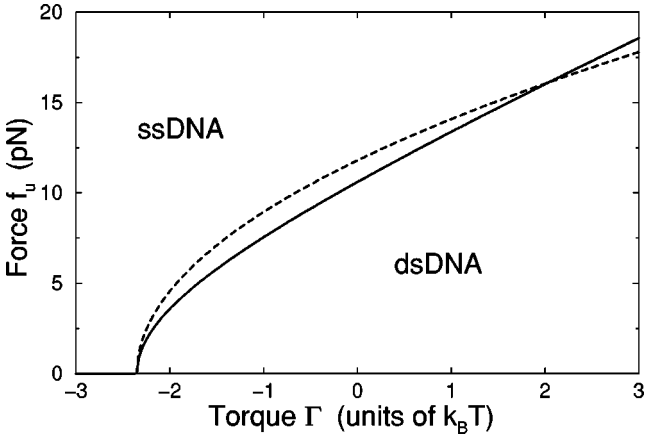

FIG. 3. DNA unzipping phase diagram as a function of torque Γ in units of k_BT and force f in pN. The solid line shows the result for the FJCL model of ssDNA elasticity, while the dashed line shows the result for its Gaussian approximation (see text); both results are in close agreement. For zero torque, the unzipping force is $f_u \approx 12$ pN, negative (unwinding) torques reduce f_u until at about $\Gamma = -2.4 k_B T$, $f_u \to 0$.

In macroscopic terms, a DNA may exist either in dsDNA or ssDNA conformations, characterized by the free energies g_{ds} and g_{ss} per bp. The denaturation free energy, $\Delta g \equiv g_{ss} - g_{ds}$ depends on sequence and on control parameters such as force, torque, and temperature. These free energies can be computed using models for DNA inferred from experimental data. Algorithms such as MFOLD [14] provide base-pairing free energies g_{ds}, and fits of experimental data to polymer models such as the freely jointed chain (FJC) give the elastic free energy g_{ss} of ssDNA as a function of force [13]. We review in Sec. IV how important quantities, such as the dependence of critical force f_u required to unzip large sequences upon the applied torque (Fig. 3), can be easily and accurately extracted from an essentially thermodynamic approach. The unzipping experiments of Bockelmann, Essevaz-Roulet, and Heslot, and Rief, Clausen-Schaumann, and Gaub are accurately described at this macroscopic level.

However, other experimentally observable aspects of unzipping are intrinsically related to the microscopic structure of DNA, and cannot be understood thermodynamically. Of particular interest to us is the barrier to initiation of unzipping, which we find to have strong consequences for the unzipping kinetics of short molecules. We find the force for initiation of unzipping to have a strong dependence on the timescale and rates of application of force. We present a semimicroscopic model of DNA, accounting for the base-pairing interactions (hydrogen bonds and stacking forces) in sufficient detail to understand these kinetic aspects of unzipping. By adopting a relatively simple semimicroscopic approach, we obtain a model that can be theoretically analyzed in detail, and which allows a precise calculation of barrier effects.

The physical origin of this barrier is sketched in Fig. 2. The single- and double-stranded portions of the molecule are separated by a boundary region. In this boundary region, the bases are unpaired so that their bonding enthalpy is lost. However, the bases are still partially stacked, so that they are unable to fluctuate as much as completely denatured ssDNA.

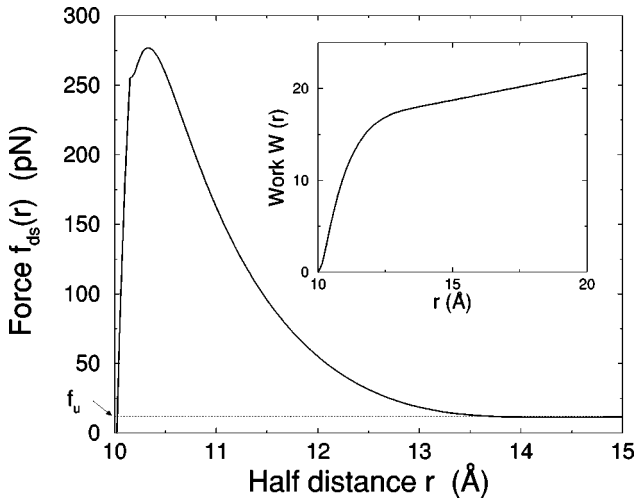

FIG. 4. Unzipping force $f_{ds}(r)$ in pN required to be applied to the ssDNA ends to maintain their half-separation r in Å. The minimal separation $r=10$ Å corresponds to the *B*-DNA-double helix structure. For sufficiently large $r(>13$ Å$)$, $f_{ds}(r)$ approaches the thermodynamic value $f_u \approx 12$ pN. Inset shows the work done $W(r)$ in k_BT for fixed r, i.e., the integral of $f(r)$. For $r>13$ Å the slope of $W(r)$ approaches its "thermodynamic" value of $f_u \approx 0.3 k_BT/$Å ≈ 12 pN.

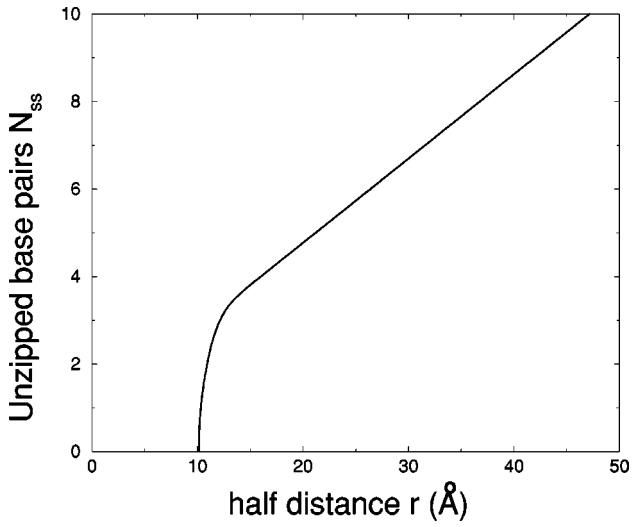

FIG. 5. Average number of unzipped based pairs N_{ss} as a function of the half-distance r between extremities. Due to the very large rigidity of dsDNA, bases pairs are initially unzipped with a very small increase of the separation r by its double helix structure value, $r=10$ Å, see Fig. 2. For sufficiently large $r(>13$ Å$)$, $N_{ss}(r)$ increases linearly with r, with a slope $1/d_u$ predicted from the thermodynamical study of Sec. IV.

Thus, the bases in the boundary region lack both the favorable free energy contributions of base-pairing enthalpy associated with dsDNA, and the fluctuation entropy associated with ssDNA. The simultaneous presence of unpaired bases with large strand rigidity is responsible for the free energy barrier.

We are able to compute the force necessary to keep the ends of the two unzipped ssDNAs at a fixed distance $2r$ (Fig. 4), as well as the shape of the opening fork (Fig. 5). Due to the free energy barrier, this force for small r is an order of magnitude larger than the asymptotic unzipping force of 10 to 20 pN depending on the sequence.

The presence of barriers makes strand dissociation an activated dynamical process that can be understood using nucleation theory. Unzipping occurs through the nucleation of a denatured "bubble," whose length depends on control parameters and is about 4 bp when located at the terminus of a DNA and 8 bp when located in the middle. We calculate the shape and free energy of this critical bubble ($\simeq 16.5 k_BT$ at the critical force), and the kinetic rates of dissociation, for different forces and molecule lengths (Fig. 6). The results are compared with the experiments of Bonnet, Krichevsky, and Libchaber and of Pörschke. Adapting Evans' theory for the breaking of single bonds [29] to the case of a one-dimensional polymer [30], we have also calculated the most probable rupture force f^* when the DNA molecule is subjected to a constant loading rate λ (Fig. 7). The dependence of f^* upon λ and N is a quantitative prediction that could be tested by AFM experiments on unzipping. Moreover it shed lights on the AFM experiments for DNA stretching of Struntz *et al.* and of Rief, Clausen-Schaumann, and Gaub.

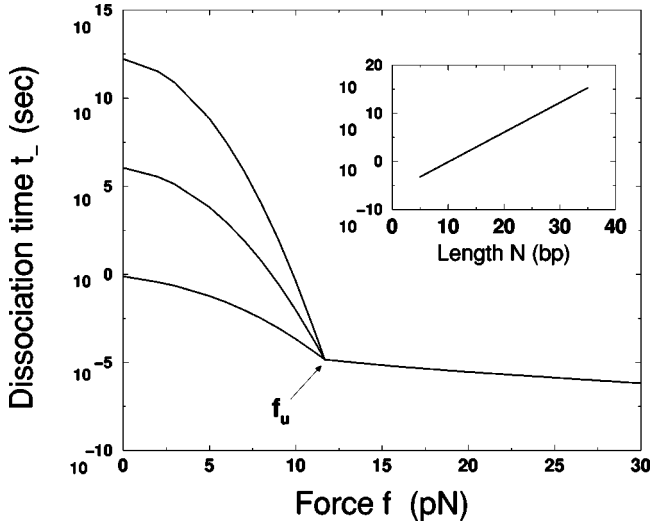

FIG. 6. Dissociation time (inverse rate) as a function of force in pN, showing three regimes of behavior. First, for $f<f_u=12$ pN, dissociation time depends on molecule length N (curves are for $N=10$, 20, 30, from bottom to top, respectively). Second, for $f_u<f<f_b=230$ pN, dissociation time decreases as the applied force gradually reduces the barrier. Finally, for $f>f_b$, no barrier is left and dissociation is immediate. Inset shows the dissociation time at zero applied force as a function of molecule length N in base pairs. The logarithm of the dissociation time is approximately a linear function of N (see text). Note that over the range $N=10$ to 20 base pairs, the zero-force dissociation time increases from a fraction of a second to hours. Also note that $f_b \approx 230$ pN is not in the range of the main figure.

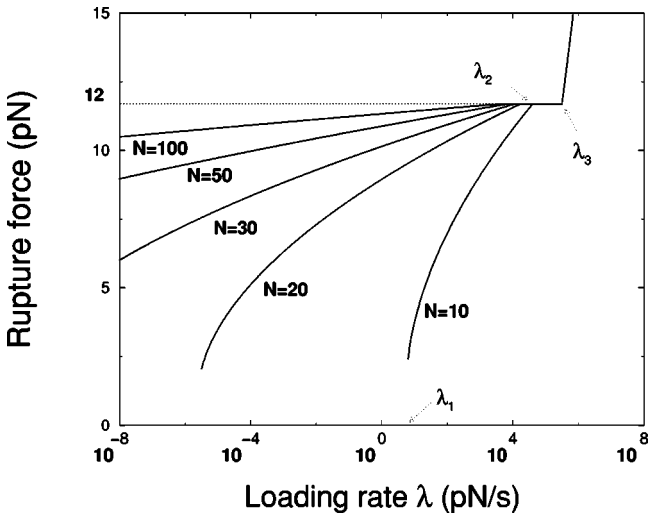

FIG. 7. Most probable rupture force f^* (pN) as a function of the loading rate λ (pN/s) for $b=1.75$. Results are shown for five different lengths ($N=10$, 20, 30, 50, and 100). Arrows indicate the critical loading rates for the $N=10$ case. Values of these critical rates (in pN/s) and associated forces (in pN) are: for $N=10$, $\log_{10}\lambda_1=0.77$, $f^*=2.3$, and $\log_{10}\lambda_2=4.6$; for $N=20$, $\log_{10}\lambda_1=-5.6$, $f^*=1.5$, and $\log_{10}\lambda_2=4.2$. The two remaining critical rates are sequence length independent with $\log_{10}\lambda_3=5.5$, $\log_{10}\lambda_4=12$.

IV. MACROSCOPIC DESCRIPTION OF DNA UNZIPPING BY FORCE AND TORQUE

DNAs can be be either base paired to form the Watson-Crick double helix (dsDNA), or the two strands can be separated (ssDNA). As shown in Fig. 2, a partially unzipped molecule consists of regions of these two phases, separated by a ''fork.'' Under conditions comparable to those *in vivo* (roughly aqueous solution at temperature 20–37 °C, pH 7 to 8, Na^+ concentration between 10 mM and 1 M), and in the absence of external stress, dsDNA is stable, with a free energy per bp less than that of ssDNA: $g_{ss}-g_{ds}=-g_0>0$. The free energy g_0 of course depends on base sequence, but in this paper we do not consider sequence effects, and, therefore, take an intermediate value of $g_0=-1.4k_BT$. This value corresponds to the denaturation free energy for a repeated poly(AC)-poly(GT) sequence, which is comparable to that of the *AT*-rich region of λ-phage used in many experiments [15,2].

A. Unwinding of dsDNA by torque

We first consider the action of an external torque Γ. Since an unperturbed dsDNA has one right-handed twist each 10.4 bp, the twist per base is $\theta_0\simeq 2\pi/10.4$; the twist per base for separated single strands is zero. The free energy of ssDNA relative to dsDNA is, therefore,

$$\Delta g(\Gamma)=-g_0+\theta_0\Gamma. \tag{1}$$

The last term represents work done by the torque Γ converting dsDNA into ssDNAs. When Δg is negative, ssDNA becomes stable, and, therefore, there is a critical unwinding torque $\Gamma_u=g_0/\theta_0\simeq -2.4k_BT$ beyond which dsDNA unwinds to form ssDNA. This is in good accord with results of single-molecule unwinding experiments [15].

B. Unzipping of dsDNA by force

The static aspects of force-induced unzipping can be largely understood using a simple Flory-like calculation [16,17]. As a simple starting point, we suppose the ssDNA regions to be flexible polymers of monomer length $l=7$ Å, with stretching free energy

$$F_{\text{stretch}}=\frac{C_{ss}r_0^2}{2N_{ss}}, \tag{2}$$

where r_0 is the end-to-end extension of a chain of N_{ss} bases. The force constant $C_{ss}=3k_BT/(l^2)$ is of the form of that of a flexible, Gaussian polymer [16]. The unzipped ssDNA region of Fig. 2 can be considered as one such chain of $2N_{ss}$ monomers with its ends $r_0=2r$ apart.

1. Unzipping as a function of extension

If the ssDNA ends are separated a distance $2r$ (Fig. 2), some average number of bases N_{ss} will open. The total free energy is a sum of chain stretching and denaturation contributions,

$$F(r,N_{ss})=\frac{Cr^2}{2N_{ss}}-N_{ss}g_0, \tag{3}$$

where $C=2C_{ss}=6k_BT/(l^2)$. The value of N_{ss} is obtained by minimization of this free energy. Solution of $dF(r,N)/dN=0$ yields a linear relation between the number of unzipped bases N_{ss} and r,

$$N_{ss}(r)=\frac{r}{d_u}, \tag{4}$$

where $d_u=\sqrt{2(-g_0)/C}$ is the projection of the monomer length along the force direction. The free energy is

$$F(r)=2\left(\frac{C}{2}(-g_0)\right)^{1/2}r=-2N_{ss}(r)g_0. \tag{5}$$

The tension in the chain is just the derivative of this with respect to $2r$,

$$f_u=\left(\frac{C}{2}(-g_0)\right)^{1/2}, \tag{6}$$

and is, therefore, a constant as unzipping proceeds. The excess free energy per unpaired bp at fixed r is double the free energy of denaturation because work done extending the ssDNAs adds to the work done opening the molecule.

2. Unzipping as a function of force

For fixed unzipping force f the excess free energy of the molecule relative to dsDNA is

$$\Delta G(r,N_{ss}f)=\frac{Cr^2}{2N_{ss}}-N_{ss}g_0-2rf. \tag{7}$$

The half-distance r between the ends adjusts to minimize ΔG and equals $r=2fN_{ss}/C$. The resulting free-energy difference is

$$\Delta G(N_{ss},f)=N_{ss}\Delta g(f)=N_{ss}\left(-g_0-\frac{2f^2}{C}\right). \tag{8}$$

This effective denaturation free-energy density $\Delta g(f)$ is lower than $-g_0$, as expected, and vanishes at the critical unzipping force f_u. For $f<f_u$, dsDNA is stable, and if $f>f_u$, the double helix unzips.

C. Improving the description of ssDNA elasticity

The free energy per base pair at fixed force (8) is

$$\Delta g(f)=-g_0+2g_s(f), \tag{9}$$

where $g_s(f)$ is the stretching free energy at fixed force, per base of ssDNA. This form of our Flory-like unzipping theory is independent of the precise details of g_s. Above, we assumed the simple Gaussian form

$$g_s^{\text{gaus}}(f)=-\frac{f^2}{C}, \tag{10}$$

which is appropriate to describe a flexible polymer that is not stretched to near its maximum extension.

Although this formula provides an empirically accurate description of ssDNA elastic response in the force range where unzipping experiments are done, the actual elastic response of ssDNA is far more complex than that of a simple flexible polymer. The reason for this complexity is that ssDNA conformation is determined by a balance between self-attractive interactions between the exposed bases, and repulsive interactions between the (PO_4^-) charges along the backbone (for a rather complete theoretical discussion of this, see Ref. [18]).

Under the conditions of the experiments described above, ssDNA actually collapses at small extensions, and requires a threshold force of about 1.5 pN to begin to extend. Then, over forces from 1.5 pN to 25 pN ssDNA elastic response is reasonably well described by the Gaussian free energy used above, provided an appropriate value of C is taken. However, strictly speaking, the value $l=7$ Å should not be interpreted as the ''true'' segment length of ssDNA. This has been a source of confusion in some interpretations of ssDNA stretching data (e.g., the factor-of-two conflict between values of ''segment length'' determined in Refs. [13] and [11]).

The collapse effect means that the finite extensibility of ssDNA is difficult to globally fit to simple flexible polymer models, for example the FJC model free energy,

$$g_s^{\text{FJC}}(f)=-k_BT\ln\left[\frac{k_BT}{lf}\sinh\left(\frac{lf}{k_BT}\right)\right]. \tag{11}$$

This model cannot be fit to ssDNA elasticity data at both high and low forces.

However, Cui *et al.* found that their data for ssDNA elasticity in 150 mM NaCl was well described by a slightly modified version of Eq. (11),

$$g_s^{\text{FJCL}}(f)=-k_BT\frac{l_{ss}}{d}\ln\left[\frac{k_BT}{df}\sinh\left(\frac{df}{k_BT}\right)\right], \tag{12}$$

which we call a ''freely-jointed-chain-like'' model (note that Cui *et al.*'s model for force versus extension has been integrated to obtain $g_s(f)$; we ignore the helix stretching elastic contribution that is not important at the low forces we are considering). This model effectively has two separate ''segment length'' parameters: the contour length per base pair $l_{ss}=27\ \mu$m/48.5 kb$\simeq$5.6 Å; and the segment length $d=15$ Å. Strictly speaking, neither of these should be regarded as anything more than fit parameters. The free energy for forces up to $\approx$20 pN is shown in Fig. 8 over the range $0\leqslant f\leqslant 20$ pN. The free energy (12) gives an unzipping force of $f_u=11$ pN and a segment projection of $d_u=4$ Å.

D. Effective segment length for the Gaussian approximation to ssDNA elasticity

The quadratic expansion of $g_s^{\text{FJCL}}(f)$ for small forces reads

$$g_s(f)=-\frac{f^2}{2C_{ss}^{\text{FJCL}}}, \tag{13}$$

where $C_{ss}^{\text{FJCL}}=3k_BT/(dl_{ss})$. This result coincides with the free energy of the Gaussian model with $l=\sqrt{dl_{ss}}\simeq 9$ Å. We show in Fig. 8 the results of this expansion. The choice $l=7$ Å gives a better approximation of the FJCL with a $0.15k_BT$ accuracy over the whole range $0\leqslant f\leqslant 20$ pN. Using the Gaussian approximation with $l=7$ Å and $g_0=-1.4k_BT$ gives $f_u=12$ pN, $d_u=5$ Å, i.e., essentially the same result as if a nonlinear ssDNA elasticity model such as Eq. (12) were used.

E. Unzipping force-torque phase diagram

Combining Eqs. (1) and (9), we obtain the denaturation free energy as a function of both torque and applied force,

$$\Delta g(\Gamma,f)=-g_0+\theta_0\Gamma+2g_s(f). \tag{14}$$

The predicted phase diagram as a function of torque and unzipping force is shown in Fig. 3. The two curves shown are for different forms of g_s: the FJCL model (12) with $l_{ss}=5.6$ Å and $d=15$ Å, and the Gaussian approximation (10) with $l=7$ Å. The details of this phase diagram are not highly sensitive to the replacement of the FJCL model with a Gaussian approximation. Existing experiments have probed only two points on this phase boundary, the zero torque [2] and zero-force [15] intercepts. It would be interesting to carry out controlled-torque experiments to verify the shape of our predicted phase boundary.

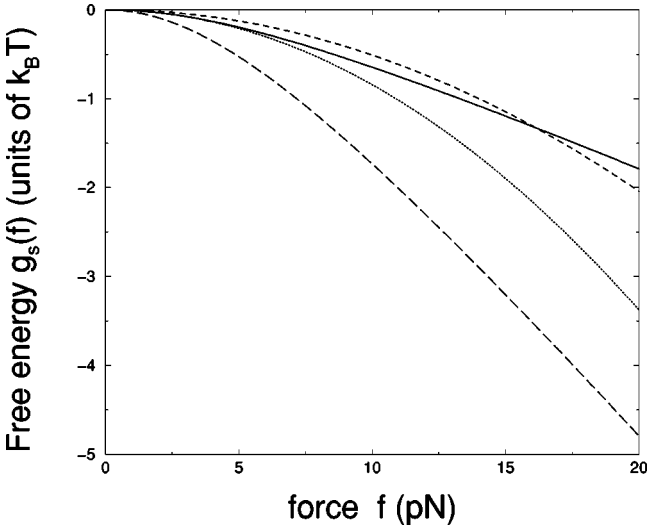

FIG. 8. Single strand free energy per base $g_s(f)$ in units of k_BT as a function of force f in pN. Solid curve is the FJCL model fit by Smith, Cui, and Bustamante [13] to their experimental stretching data; long dashed curve is the FJC model with Kuhn length $d=15$ Å. Dotted line is the Gaussian approximation with $d=9$ Å (the best quadratic approximation to the FJCL at zero force); dashed curve is the Gaussian approximation with $d=7$ Å (the best quadratic approximation to the FJCL at forces $\approx$12 pN).

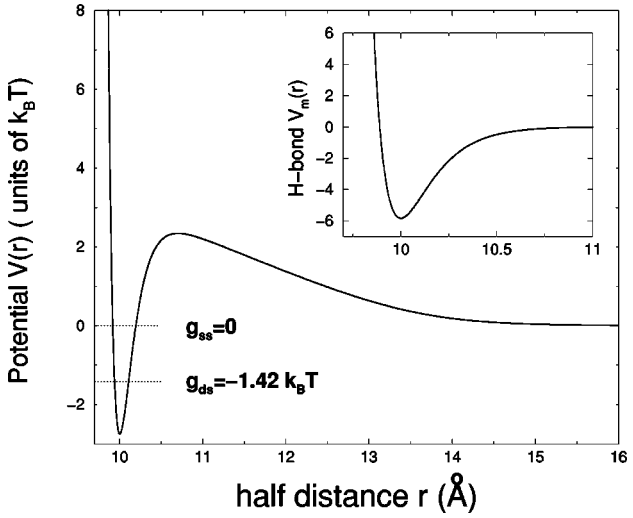

FIG. 9. Inter-base-pair potentials in units of k_BT, as a function of base-pair half-separation r in Å. Inset shows the Morse potential $V_m(r)$ describing the short-ranged hydrogen-bond interaction (the energetic part of the base-pairing interaction in our theory). Main figure shows the total potential $V(r)$ for zero torque, including ssDNA conformational entropy (see text). Once the entropic contribution is included, an energy barrier appears. Note the difference of distance scales on the main figure and inset. Zero force and torque free energies of separated ssDNAs and strands bound into the *B*-DNA double helix are indicated.

In conclusion, simple Flory-like models give a satisfying overview of DNA unzipping experiments at large ($>$100 base) scales. But, thermodynamic models of the sort discussed in this section cannot give much insight into either initiation of unzipping (i.e., the alteration of f_u during the opening of the first few bases), or into the kinetics of opening. These questions are central to understanding the initial steps of opening of the double helix essential to a wide range of cellular machinery, and they form the focus of the remainder of this paper.

We note that the Gaussian approximation (10) to ssDNA elasticity discussed above will become an essential simplification in the base-pair-scale theory that follows. We emphasize that the Gaussian approximation can be used because the forces we are interested in, $f\approx f_u=12$ pN, are both well above the $\approx$2 pN threshold for pulling a collapsed ssDNA apart, and well below the $\approx$50 pN forces required to fully extend a ssDNA.

V. SEMIMICROSCOPIC MODEL OF DNA UNZIPPING

A. Model with discrete base-pair degrees of freedom

We now develop a more detailed model of DNA unzipping based on our previous work on base-pairing interactions [5,6,19]. To begin with, we number the base pairs along a dsDNA with a discrete index n. The microscopic degrees of freedom that we consider are the half-distances between bases $r(n)$, and the twist angle $\theta(n)$ and axial distance $z(n)$ between bases n and $n-1$. Gaussian thermal fluctuations of the angular and axial distance degrees of freedom can be integrated out, leaving an effective energy $\mathcal{H}[\mathcal{R}]$ that depends only on the set of inter-base-pair radii $\mathcal{R}=\{r(0),r(1),\ldots,r(n)\}$ [6], see Appendix A.

In the dsDNA state, the radii r fluctuate around their average value $R=10$ Å, in a potential well resulting from the balance of attractive hydrogen-bond attraction with repulsive hardcore repulsion. This potential well can be described using the short-range Morse potential

$$U_H(r)=D[(e^{-a(r-R)}-1)^2-1], \tag{15}$$

as shown in Fig. 9. The Morse potential has a strong repulsive barrier at small radius $r<R$, and a narrow attractive well with minimum near $r=R$. Parameters of the Morse potential suitable to describe DNA base-pairing interactions are $D=5.84k_BT$, $a=6.3$ Å^{-1} [20,5].

In addition to base-pairing interactions acting across the double helix, stacking interactions between successive base pairs favor the successive radii r and r' to be similar in value. This attractive interaction is strongest in dsDNA [i.e., when $r(n)\approx R$] and is greatly reduced as the bases are pulled apart. We account for these effects with an enhancement of the intrinsic ssDNA rigidity C discussed in Sec. IV, parametrized by a strength E and a range $1/b$,

$$E(r,r')=E\exp\{-b[\tfrac{1}{2}(r+r')-R]\}. \tag{16}$$

The interaction strength is determined from Raman spectroscopy measurements of base-pair radii vibrational modes of dsDNA [21] to be $E\simeq 58\ k_BT/$Å^2 [6].

The total energy for configuration $\mathcal{R}$ reads

$$\mathcal{H}[\mathcal{R}]=\sum_{n=1}^{N}\frac{1}{2}m(r(n),r(n-1))(r(n)-r(n-1))^2 + \sum_{n=0}^{N}U(r(n)). \tag{17}$$

The net rigidity m includes both the intrinsic ssDNA rigidity C and its enhancement by stacking interactions for dsDNA,

$$m(r,r')=E(r,r')+tC. \tag{18}$$

This quantity interpolates between dsDNA and ssDNA values. The reduced temperature $t=T/T_0$, where $T_0=298$ K, appears in front of C, due to the largely entropic origin of the intrinsic ssDNA chain rigidity (see Sec. IV) in the regime of extension of interest to us.

The second term in Eq. (17) includes the hydrogen-bond energy $U_H(r)$ (15) and the elastic torque energy,

$$U(r)=U_H(r)-\frac{\Gamma}{t}\frac{R_1}{r}. \tag{19}$$

We show in Appendix A how the total energy (17) and the elastic torque energy in Eq. (19) come from the microscopic model of Ref. [5], once the twist angles are integrated out. As calculated in Appendix A, $R_1=6$ Å. Thus, for dsDNA, the torsional energy is roughly equal to $-\Gamma\ R_1/R$ from Eq. (19), in agreement with the expression $-\Gamma\ 2\pi/10.4$ of Eq. (1).

The partition function at inverse temperature $\beta=1/k_BT$ is equal to

$$Z=\int d\mathcal{R}\, e^{-\beta\mathcal{H}[\mathcal{R}]}, \tag{20}$$

where the measure over the chain configurations is

$$d\mathcal{R}=\prod_{j=1}^{N}dr(j), \tag{21}$$

B. Conversion of the discrete transfer matrix to a continuum Schrödinger equation

A number of thermodynamical results were derived in Refs. [5], [6] by diagonalizing the transfer matrix of the discrete model. The dynamical analysis of this paper is facilitated by treating the base-pair index n as a continuous variable [22–24]. A configuration of the chain is then described by radii $\mathcal{R}=\{r(n)\}$, with $0\leqslant n\leqslant N$. The main difficulty in taking the continuum limit comes from the source of main new and interesting results, the r dependence of the ''mass'' or stiffness (18). Our approach will be to take the discrete-model transfer matrix, and convert it to a continuum version, which takes the form of a Schrödinger equation. We will present an outline of this calculation, with some details deferred to Appendix B.

The first step is to write the discrete-model transfer matrix as

$$T_1^{\mathrm{ass}}(r,r')=\exp\{-\beta[\tfrac{1}{2}m(r)(r-r')^2+U(r)]\}. \tag{22}$$

Here, the dependence of m has been made local in r, which amounts to keeping only derivative terms in r at the Gaussian level. Thus, $m(r)=E(r)+tC$ where $E(r)=E\exp[-b(r-R)]$

Bouchiat [32] has communicated to us an alternate calculational scheme that allows an exact calculation at this step, which we describe in Appendix B. It turns out that the corrections that our truncation approximation ignores are simply a small shift in the net potential. We use the form (22) in what follows since it leads to a substantial simplification of the calculations with only small quantitative changes in the results.

In the continuum limit, the transfer matrix and its eigenfunction are replaced by the differential equation

$$\left(-\frac{(k_BT)^2}{2m(r)}\frac{\partial^2}{\partial r^2}+V(r)\right)\psi_k(r)=g_k\psi_k(r). \tag{23}$$

where the eigenfunctions $\psi_k(r)$ and eigenvalues g_k are indexed by the integer k. This has the form of a one-dimensional quantum-mechanical Schrödinger equation, where the base-pair index n plays the role of time, and the Planck constant h is replaced by k_BT.

A unique feature of Eq. (23) is that its ''mass'' m is position dependent,

$$m(r)=E\exp[-b(r-R)]+tC. \tag{24}$$

This feature results in the base-pair potential picking up a contribution from $m(r)$,

$$V(r)=U(r)+\frac{t}{2}k_BT_0\ln[m(r)/m(\infty)]. \tag{25}$$

This additional contribution comes from the measure factor required to carry out the usual Feynman-Hibbs derivation of Eq. (23). For convenience, we set the offset of the net potential so that $V\to 0$ as $r\to\infty$. This choice sets the absolute free energy of separated ssDNAs under no force or torque, to zero.

The net potential $V(r)$ is shown on Fig. 9 for $b=1.75$ Å^{-1}, a value chosen in light of the results for kinetics in Sec. VIII. As a result of the r dependence of the rigidity, a barrier separating ds-DNA from ss-DNA has appeared. This barrier has little effect on macroscopic aspects of equilibrium unzipping, but has large consequences for the denaturation kinetics.

The most negative eigenvalue g_0 is related to the free energy per base pair for a long molecule, by $g_0=-(k_BT\ln Z)/N$. The corresponding ''wave function'' $\psi_0(r)$ determines the equilibrium probability distribution of r. Far from the ends of a chain, the probability to open a given pair of bases to have radius r is $|\psi_0(r)|^2$. Alternately, at the chain ends, the probability to open either the first or the last base pair by r is $\psi_0(r)$.

C. Energy for path-integral formulation of continuous model

From the Schrödinger equation we obtain the continuous counterpart of the discrete energy $\mathcal{H}$ (17),

$$\mathcal{G}[\mathcal{R}]=\int_0^N dn\left\{\frac{1}{2}m(r(n))\left(\frac{dr}{dn}\right)^2+V(r(n))\right\}, \qquad (26)$$

To obtain the same potential $V(r)$ in the path integral $\mathcal{G}$ as in the Schrödinger equation the path integral measure must be

$$d\mathcal{R}=\lim_{\epsilon\to 0}\prod_{j=0}^{N/\epsilon}\sqrt{\frac{\beta m(r(j))}{2\pi\epsilon}}\,dr(j), \qquad (27)$$

where ϵ is the ''lattice constant,'' equal to unity in the discrete version of the model, and taken to zero in the continuum model. Note that when $\epsilon=1$ the $\ln m$ term of the exponential cancels the integration measure factor. Including the inverse temperature $\beta=1/t$, the partition function is formally

$$Z=\int d\mathcal{R}\,e^{-\beta\mathcal{G}[\mathcal{R}]}. \qquad (28)$$

Below, we will use the continuum energy (26) to calculate shapes of domain-wall structures, and as a basis for a Langevin equation for unzipping dynamics.

VI. EQUILIBRIUM STATES OF THE SEMIMICROSCOPIC MODEL

A. Zero applied force

1. ssDNA

The ''free particle'' continuum states of Eq. (23), where the strands are far away from one another ($r\gg R$), correspond to isolated ssDNAs under zero force and torque. This situation is described by

$$-\frac{(k_BT)^2}{2C}\frac{\partial^2}{\partial r^2}\psi_{ss}(r)=g_{ss}\psi_{ss}(r). \qquad (29)$$

Thus, ssDNA is treated as a sugar-phosphate polymer with Gaussian entropic elasticity, with rigidity $C=6k_BT/l^2$ (Sec. IV C). The eigenfunctions. are plane waves, and the lowest-energy ssDNA states have free energy per base pair $g_{ss}=0$.

2. dsDNA

The double helix is the bound state of the potential V (Fig. 9). This bound state can be found numerically, but an accurate analytical calculation of its free energy per base pair g_{ds} can be obtained by considering $\psi_{ds}(r)$ to be essentially confined inside the Morse well. The only potential that variates significantly for $r\simeq R$ is the Morse potential representing the hydrogen bonds; as $a/b\gg 1$ the entropic and torque terms vary relatively slowly.

We, therefore, obtain the Morse equation

$$\left(-\frac{(k_BT)^2}{2E}\frac{\partial^2}{\partial r^2}+V_{ds}(r)\right)\psi_{ds}(r)=g_{ds}\psi_{ds}(r), \qquad (30)$$

where

$$V_{ds}(r)=U_H(r)+\frac{t}{2}k_BT_0\ln[m(R)/m(\infty)]-\frac{\Gamma}{t}\frac{R_1}{R}. \qquad (31)$$

The ground state of this equation is exactly known [25], giving the zero-force free energy,

$$g_{ds}(f=0,\Gamma)=g_0-\Gamma\theta_0, \qquad (32)$$

with θ_0 defined in Eq. (1), and

$$g_0\simeq -D+\frac{t}{2}a\sqrt{2D/E}-(ta)^2/(8E)+\frac{t}{2}\ln[1+E/(tC)]. \qquad (33)$$

At room temperature ($t=1$) and zero applied force, $g_0\simeq -1.42k_BT$; therefore, at zero force the critical unwinding torque is $\Gamma_c=-2.35k_BT$, in good agreement with the experimental results obtained by the group of Bensimon and Croquette [15].

3. Thermal denaturation

Our expression for $\Delta g=-g_{ds}$ (32) can, in principle, be used to determine the temperature T_m at which dsDNA ''melts'' into isolated ssDNAs in the absence of mechanical stress ($f=0,\Gamma=0$). The resulting value, $T_m=410$ K, is, however, too large with respect to thermodynamical measurements of melting temperatures. The main reason is that our model as formulated above does not account for the change in the hydrogen-bonding potential energy with temperature.

The correct melting temperature can be obtained by making D, the depth of the Morse potential, depend explicitly on the temperature T. Assuming that the depth behaves as $D=D_0-D_1(t-1)$, with $D_0=5.84k_BT$, and requiring melting temperature $T_m=350$ K, we have $D_1=5\ k_BT/\mathrm{K}$. When T ranges from 298 to 350 K, $D(T)$ thus decreases from $5.84k_BT$ to $5k_BT$. The order of magnitude of this variation caused by temperature is similar to that discussed by Prohofsky *et al.* [20].

B. Unzipping by external force

1. Schrödinger equation

In presence of an unzipping force f applied on the last base pair $n=0$, the discrete model transfer matrix (22) needs an additional term $2fr(0)/k_BT$ included in the exponential. This leads to the one-dimensional Schrödinger equation (see Appendix B),

$$\left[-\frac{(k_BT)^2}{2m(r)}\frac{\partial^2}{\partial r^2}+\frac{2f(k_BT)}{m(r)}\frac{\partial}{\partial r}+\hat{V}(r)\right]\psi(r)=g\psi(r). \qquad (34)$$

Here the potential $\hat{V}$ is just

$$\hat{V}(r)=V(r)-\frac{2f^2}{m(r)}. \qquad (35)$$

The potential (35) accounts for the decrease of ssDNA free energy relative to that of dsDNA by an amount that tends to the value $-2f^2/C$ at large radii (8), (9), (10). The path integral effective energy corresponding to Eq. (34) is

$$\mathcal{G}[\mathcal{R}]=\int_0^N dn\left\{\frac{1}{2}m(r(n))\left(\frac{dr}{dn}\right)^2+V(r(n))\right\}-2fr(0). \tag{36}$$

This expression can be rewritten as

$$\mathcal{G}[\mathcal{R}]=\int_0^N dn\left\{\frac{1}{2}m(r(n))\left(\frac{dr}{dn}-\frac{2f}{m(r(n))}\right)^2+V(r(n))-\frac{2f^2}{m(r(n))}\right\}+2fr(N). \tag{37}$$

Below we will assume that deep inside the dsDNA phase, the base-pair radius does not fluctuate much around its equilibrium value R, so that the last term on the right-hand side (rhs) of Eq. (37) can be considered to be an additive constant, which we will not write.

2. Phase diagram in the force-torque plane

From Eq. (35), the potential at nonzero force is lowered by $2f^2/C$ at large radii, and essentially left unchanged at small r. As a result, the free-energy g_{ss} of ssDNA is decreased by $2f^2/C$, while g_{ds} is unaffected, relative to their zero-force values. This allows us to reobtain Eq. (14), and therefore the ''thermodynamic'' phase diagram (Fig. 3).

VII. EQUILIBRIUM FORCE BARRIER TO INITIATION OF UNZIPPING

In the previous section we saw that the semimicroscopic model was consistent with the simple thermodynamic model of Sec. IV. We now move on to examine the initial stages of unzipping, when the two ssDNA ends are just beginning to be unpaired. We carry out these calculations in the ensemble of fixed extension, since at fixed-force equilibrium one will simply pass from dsDNA to ssDNA at the unzipping force.

A. Work done initiating unzipping

The work done in pulling the last base pairs to a given separation follows simply from the equilibrium probability distribution of end radius r. This is determined in turn by the dsDNA ''wave function'' $\psi_{ds}(r)$. As mentioned in Sec. VI, in the base-pairing potential well $\psi_{ds}(r)$ is well approximated by the exactly known Morse ground state wave function. In turn, the average force that must be applied to keep the ssDNA ends a certain fixed distance apart will be just the derivative of the work done with end separation.

For radius values outside of the Morse well, i.e., $r-R \gg 1/a$, this expression is not valid anymore. However, the wave function can in this case be calculated using the standard quantum-mechanical WKB approximation [26]. Defining r_{ds} as the value of the radius for which $V(r_{ds})=g_{ds}$ we obtain

$$\psi_{ds}(r)=\frac{A}{\sqrt{p(r)}}\exp\left(-\int_{r_{ds}}^{r}dr'\,p(r')/t\right) \quad (r>r_{ds}), \tag{38}$$

where $p(r)=\sqrt{2m(r)[V(r)-g_{ds}]}$ [26]. The coefficient A can be calculated by connecting the two expressions of $\psi_{ds}(r)$, in such a way that both the wave function and its derivative are continuous functions of r.[1]

The logarithm of the end base-pair radius distribution

$$W(r)=-k_BT_0t\ln\psi_{ds}(r), \tag{39}$$

is shown in the upper curve of Fig. 4. $W(r)$ is the free energy associated with the separation $2r$. i.e., the work done in pulling the two ends apart.

Two regimes can be seen as r grows. The work done initially grows quickly with r, up to a value $W^*\equiv W(r^*)\simeq 36k_BT/b$, see Appendix C. This initial regime corresponds to the separated strands still being within the large energy barrier. For $r>r^*$, one enters a second regime where the portion of the molecule in the barrier region is no longer changing, and, therefore, where the additional work done approaches its thermodynamically expected value, $W(r)\approx r\sqrt{2C(-g_0)}=(2r)f_u$, independent of b [23]. The derivative of $W(r)$ with respect to $2r$ gives the average force $f_{ds}(r)$ that must be applied to the molecule ends to keep their half-separation at r (Fig. 4).

This calculation gives a prediction for the small-r behavior that is not possible from the thermodynamical perspective of Sec. IV. At very short distances we predict a surprisingly large force barrier. We do not know of any direct measurement of this force barrier, although we will show below that experiments on the spontaneous unpairing of short double helicies [8] provide indirect but quantitative evidence for it.

The physical origin of the large force barrier is in the potential well due to hydrogen bonding plus the additional barrier associated with the reduction in DNA strand rigidity as one passes from dsDNA to ssDNA. The hydrogen-bonding interactions U_H alone have a well depth of $\approx 2k_BT$, and vary over a range of 0.5 Å (Inset of Fig. 9). Thus, initiation of the first broken base pair requires an initial force barrier of roughly $40k_BT/\text{nm}\approx 160$ pN to be crossed. The additional $\approx 2k_BT$ barrier in the potential V resulting from the change in strand rigidity with r boosts the total force barrier to the ≈ 300 pN of Fig. 4. The potential barrier height is roughly $(k_BT/2)\ln(E/C)$ and corresponds to the change in partition function per segment associated with this rigidity change.

[1]The connection condition requires that $\psi(r)$ and $\psi'(r)$ be continuous at some point r_A. At zero torque, we find that $r_{ds}\simeq 10.11$ Å and $r_A\simeq 10.19$ Å. The connecting point r_A lies thus in the validity regions of the Morse wave function, $a(r-R)\leqslant 1$, and of the WKB approximation, $m(r)V'(r)/p(r)^{3/2}\ll 1$, which holds as soon as r is slightly above the classical turning point r_{ds} [26].

B. Free energy functional and shape of opening fork

For a partially opened dsDNA, the average shape of the opening ''fork'' can be calculated by finding a suitable stationary point of the energy functional (26). This average configuration $r^*(n)$ obeys the equation

$$m(r^*)\frac{d^2r^*}{dn^2} = -\frac{1}{2}m'(r^*)\left(\frac{dr^*}{dn}\right)^2 + V'(r^*) \qquad (40)$$

where $(\,)'$ denotes derivative with respect to r. Equation (40) is just the ''classical'' equation of motion for a particle with coordinate r^* where the role of ''time'' is played by n, and where the potential energy is $-V(r)$. Again, we have the unusual feature of a position-dependent mass $m(r)$.

This equation of motion, combined with the equilibrium initial condition $r^*(n=0)=r_{ds}$, $dr^*/dn=0$, can be solved numerically. Inverting the solution $r^*(n)$, we obtain the shape of the opening fork, that is the optimal value of opened bases N_{ss} as a function of the end-opening radius r (Fig. 2). The behavior of $N_{ss}(r)$ is plotted in Fig. 5.

As the strands are progressively pulled apart, the number of bases unzipped initially grows as the square root growth of distance between the strand ends: $N_{ss}(r) \simeq 2\sqrt{E(r-r_{ds})/V'(r_{ds})}$. For large distances, the number of unzipped bases tends to an asymptotically linear dependence on r, $N_{ss}(r)\propto r\sqrt{C/2/(-g_0)}$ in accord with the thermodynamic result of Sec. IV. These predicted features of initial unzipping would be interesting to test experimentally.

VIII. KINETIC THEORY OF UNZIPPING

We have so far studied the equilibrium unzipping of the chain, and now we turn to the kinetics of denaturation. Figure 11 shows a tridimensional representation of the potential V (Fig. 9) that the molecule has to overcome to denature. Denaturation is an activated process that we now study within Langer's nucleation theory [27,28].

We consider the Langevin dynamics of unzipping at room temperature, using the energy (26),

$$\zeta\frac{\partial r(n,t)}{\partial t} = -\frac{\delta\mathcal{G}}{\delta r(n,t)} + \epsilon(n,t). \qquad (41)$$

The random forces $\epsilon(n,t)$ are uncorrelated Gaussian variables with zero mean and variance $2k_BT\zeta$. The friction coefficient is on the order of $\zeta=6\pi\eta R$ with $R=10$ Å, where the viscosity of the water is $\eta=10^{-3}$ Pa sec. Opening of a dsDNA is in general an activated process, requiring thermal fluctuation over the energy barrier of the last section.

A. Nucleation transition-state theory

1. Phase space, energy hypersurface, and saddle point

Consider the phase space in which the unzipping dynamics take place. A molecular configuration $\mathcal{R}$ is represented by a point in space, and is assigned a (free) energy $\mathcal{G}[\mathcal{R}]$ (26). The ds and ssDNA states are separated by high-energy configurations that are only briefly visited during an unzipping transition. Depending on force, torque, and temperature, initial dsDNA or ssDNA configurations are metastable and will eventually decay to the thermodynamically stable state. In general, this requires a free energy barrier to be crossed, and the most likely barrier crossing point will be the saddle-point $\mathcal{R}^*$ of the free energy hypersurface $\mathcal{G}[\mathcal{R}]$. $\mathcal{R}^*$ is a particular configuration that is a deformation of the initial metastable state that includes a critical bubble of the stable state. This ''bubble'' will then grow, as the configuration $\mathcal{R}$ moves away from $\mathcal{R}^*$ down in free energy toward the stable state.

The kinetic rate for leaving a metastable state in this picture is given by an Arrhenius-like formula [27,28],

$$\nu \simeq \nu_0 e^{-\beta G^*}. \qquad (42)$$

Here G^* is the free energy of the saddle point relative to the metastable phase. Calculating $\mathcal{R}^*$ and G^* requires the saddle points of the functional (26), i.e., solution of Eq. (40) with boundary conditions of the departure phase.

2. Use of the free energy

In his original formulation of nucleation theory, Langer considered escape from metastable states at temperatures that were low relative to the purely energetic barrier height E^*. In the limit of zero temperature, the waiting time for a fluctuation that can stimulate a transition from the metastable to stable states goes to infinity. As the temperature is made finite, but small with respect to the energetic barrier height, activated processes allow the equilibrium state to be reached, with a relaxation time of the order of $\exp(-E^*/(k_BT))$.

As we have seen in the preceding sections, the situation we face here is formally slightly different. The absolute temperature is not small, and in fact the barrier between ssDNA and dsDNA is of mainly entropic origin. However, application of nucleation theory is possible even when the effective Hamiltonian $\mathcal{G}$ (26), (36) includes a temperature-dependent (entropic) potential, since the starting point of the theory is the Langevin equation, which may contain temperature-dependent potential terms.

3. Identifying the metastable and stable states

When a force f, larger than the critical unzipping force f_u, is applied to the molecule, the dsDNA phase is metastable with respect to the open ssDNA state (Fig. 3). Nucleation will, therefore, take place around the saddle point, which includes an open fork in the dsDNA [Fig. 10(a)]. The dissociation rate ν_- is in this case directly given by formula (42).

At forces $f<f_u$ below the equilibrium unzipping threshold, dsDNA is thermodynamically stable, but in a finite molecule may reach the metastable ssDNA phase by thermal fluctuation. In this case, we can infer the unzipping rate from the rate for the *reverse* reaction, which is of the form of a decay of a metastable (ssDNA) state.

Let ν_- be the kinetic rate for dissociation and ν_+ the kinetic rate for recombination when the two strands are already in close contact. Both ν_-, ν_+ depend on the length N of the sequence and on applied force f. The forward and reverse rates are related by detailed balance,

$$\nu_-(N,f) = \nu_+(N,f)e^{-\beta N\Delta g(f)}, \qquad (43)$$

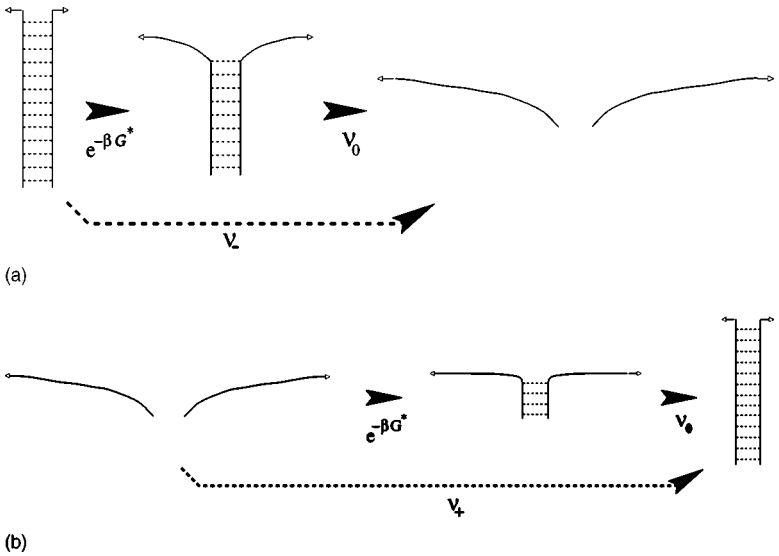

FIG. 10. Sequences of states used to calculate kinetics of barrier crossing during unzipping. (a) When $f>f_u$, the unzipped double helix is thermodynamically favored $[\Delta g(f)<0]$ and the unzipping rate ν_- can be calculated directly. (b) When $f<f_u$, the *B*-DNA double helix is thermodynamically favored $[\Delta g(f)>0]$ and the unzipping rate ν_- must be calculated using the closing rate ν_+, where the *B*-DNA double helix is the final state. The nucleation bubble in this process is 4 bp long. Then, the detailed balance condition $\nu_-=\nu_+ e^{-N\beta\Delta g(f)}$ can be used to determine the opening rate ν_-.

where $\Delta g(f)$ is the free energy excess per bp of ssDNA relative to dsDNA. Knowing ν_+ and $\Delta g(f)$, therefore, gives the dissociation rate ν_-. The recombination rate ν_+ measures the escape rate for the metastable ssDNA phase. The saddle-point configuration for this transition is a bubble of a few paired bases, at the end of two nearly completely unzipped ssDNAs [Fig. 10(b)].

We next compute the probability to nucleate a bubble of stable phase inside the metastable phase, as a function of the applied force, i.e., $\nu_-(f)$ for $f>f_u$ and $\nu_+(f)$ for $f<f_u$. We discuss how to calculate the activation barrier free energy G^* using an instanton technique. In Appendix D we show how the WKB approximation may be used to obtain the same results.

B. Nucleation bubble

1. Extremization of the free energy functional

The activated configuration in the phase space, and its activation free energy are (Sec. VIII A 1)

$$G^*=\mathcal{G}[\mathcal{R}^*]-\mathcal{G}[\mathcal{R}_M], \tag{44}$$

where $\mathcal{R}_M$ is the average configuration of the chain in the metastable state M (note $M=$ds or ss). As discussed in Sec. VIII A 1, the activated configuration can be thought of as made of two pieces. All but a finite number of monomers (numbered by n running from $-N+N^*$ to 0) are in the metastable, initial phase, with a small nucleation bubble ($0\leqslant n\leqslant N^*$) of the stable phase located at the extremity of the molecule (Fig. 11).

From Eq. (44), the activation free energy is just the free energy excess of the nucleation bubble with respect to the metastable phase. We therefore extremize

$$G^*=\int_0^{N^*} dn\left\{\frac{1}{2}m(r^*(n))\left(\frac{dr^*}{dn}\right)^2+V(r^*(n))-g_M\right\}$$
$$-2\,f(r^*(0)-r^*(N^*)), \tag{45}$$

to find the shape of the bubble, $r^*(n)$. Note that G^* is not simply equal to the potential barrier height as it would be for a single particle moving in the potential V, but also takes into account the rigidity of the strands.

For $0<n<N^*$, $r^*(n)$ satisfies Eq. (40), expressing the balance of forces at each point along the chain. The analogy

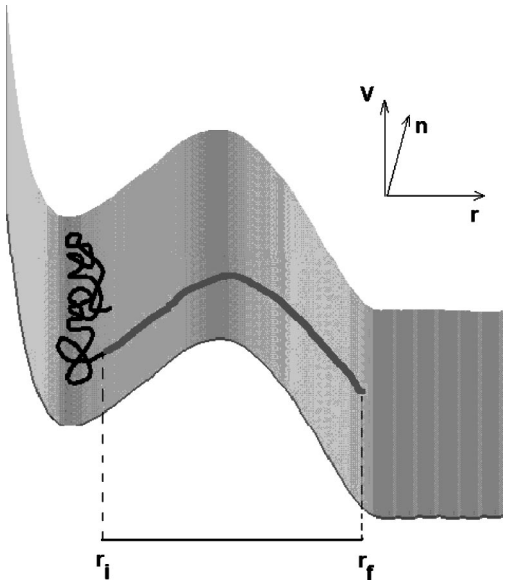

FIG. 11. Schematic representation of the saddle-point polymer conformation $r^*(n)$, which crosses the barrier of $V(r)$. Here, r_i and r_f are the radii of the extremities of the nucleation bubble.

with the equation of motion of a unidimensional particle with coordinate r evolving with "time" n is useful. At the extremities of the bubble, the functional differentiation of Eq. (45) gives back the equation of motion (40) with an additional contribution to the rhs of Eq. (40) equal to $-2f(\delta(n)-\delta(n-N^*))$. Integrating out this equation of motion over the vanishingly small interval $-\epsilon \leqslant n \leqslant +\epsilon$ $(N^*-\epsilon \leqslant n \leqslant N^*+\epsilon)$, with $\epsilon \to 0^+$, shows that the velocity dr^*/dn is discontinuous at the extremities of the bubble, and leads to the following boundary conditions:

$$\left.\frac{dr^*}{dn}\right|_{n=0^+} = \frac{2f}{m(r^*(0))}, \tag{46}$$

$$\left.\frac{dr^*}{dn}\right|_{n=(N^*)^-} = \frac{2f}{m(r^*(N^*))}. \tag{47}$$

These two boundary conditions are not sufficient to solve Eq. (40) since N^* is not determined yet. We need one more condition, which expresses the continuity between the metastable bulk of the chain $(n<0)$ and the adjacent extremity of the nucleus $(n>0)$. From Eq. (45), the free energy per base pair at position n in the nucleus is equal to

$$g^*(n) = \frac{1}{2}m(r^*(n))\left(\frac{dr^*}{dn}\right)^2 + V(r^*(n)) - 2f\frac{dr^*}{dn}, \tag{48}$$

and must coincide with the metastable value g_M in $n=0$. Using the boundary value of the velocity (46) and definition (35), we obtain

$$\hat{V}(r^*(0)) = g_M. \tag{49}$$

Equation (40) can now be integrated by determining $r^*(0)$ from Eq. (49) and using the boundary condition (46) for the velocity. Once the bubble trajectory $r^*(n)$ is known, the activation free energy can be computed from Eq. (45).

2. Nucleation bubble free energy

The transition state free energy G^* can be expressed in terms of a first integral of Eq. (40),

$$Q = -\frac{1}{2}m(r^*(n))\left(\frac{dr^*}{dn}\right)^2 + V(r^*(n)). \tag{50}$$

From Eqs. (46) and (49), $Q=g_M$. Along the bubble trajectory, the velocity is a simple function of the radius,

$$\frac{dr^*}{dn} = \sqrt{\frac{2[V(r^*(n))-g_M]}{m(r^*(n))}}. \tag{51}$$

Using Eq. (47) and (51), we find that the radius $r^*(N^*)$ at the extremity of the bubble is given by $\hat{V}(r^*(N^*))=g_M$. A similar relation holds at the other end of the bubble, see Eq. (49). In addition, Eq. (51) can be used to change the variable in Eq. (45) from n to r.

We thus obtain the following expression for the nucleation bubble free energy:

$$G^* = \int_{r_i}^{r_f} dr\sqrt{2m(r)[V(r)-g_M]} - 2f(r_f-r_i), \tag{52}$$

where $r_i[=r^*(0)]$ and $r_f[=r^*(N^*)>r_i]$ are the radii at the two endpoints, fixed using

$$\hat{V}(r_i) = \hat{V}(r_f) = g_M, \tag{53}$$

with $\hat{V}(r)$ defined in Eq. (35). Note that when ssDNA is the metastable phase, the integration limits of Eq. (52) are r_i, with $\hat{V}(r_i)=g_{ss}=-2f^2/C$, and $r_f=\infty$. The activation free energy G^*, therefore, is finite.

Appendix D shows how these results can be alternately obtained from the Schrödinger equations (23), (34) using the WKB method.

3. Nucleation bubble shape

Once r_i and r_f are determined by the boundary condition (53), the activation free energy G^* follows by numerical integration of Eq. (52). The shape of the bubble is obtained by numerical solution of Eq. (40). For $f>f_u$, both r_i and r_f are finite and the number N^* of base pairs in the bubble may be obtained from integration of Eq. (51),

$$N^*(f) = \int_{r_i}^{r_f} dr\ \sqrt{\frac{m(r)}{2[V(r)-g_M]}}. \tag{54}$$

For forces larger than f_u, we find that $N^*(f)$ decreases slowly with f, with, e.g., $N^*(f_u=12\text{ pN})=4$, $N^*(f=40\text{ pN})=3$.

For forces $f<f_u$, r_f goes to infinity because the potential is flat at infinite radius, and formula (54) must be used with care. The number of base pairs in the bubble can then be defined from the change of slope of the corresponding $r*(n)$ curve. Beyond some radius $r_t=r^*(N^*)$, the average half distance $r^*=n(2f)/C$ between extremities increases linearly with n, and thus corresponds to ssDNA [Sec. IV B 2 and Fig. 10(b)]. We precisely define N^* as the bp index at which the second derivative d^2r^*/dn^2 is maximal. We find $N^*(f)\simeq 4$ for all the forces $f<f_u$. The opening forks for $f=0$, 5, and >12 pN are shown in Fig. 12.[2]

A simple approximate calculation of N^* at the critical force $f=f_u$ shows that $N^*\simeq\sqrt{E}/b$ for ratios E/C ranging from 100 to 1000 ($E/C\simeq 490$ in our model), see Appendix C. The linear dependence of N^* upon $1/b$ is reasonable, since the wider the barrier, the longer the nucleus to overcome it. The square root dependence on the dsDNA stiffness E follows from the form of Eq. (54).

[2]No confusion should be made between the opening forks at zero force (Fig. 12) and at fixed distances between extremities (Fig. 5). The latter corresponds to fluctuations around dsDNA equilibrium represented by ssDNA bubbles as in Fig. 10(a), while the former corresponds to the case of metastable ssDNA, with an activated bubble sketched in Fig. 10(b).

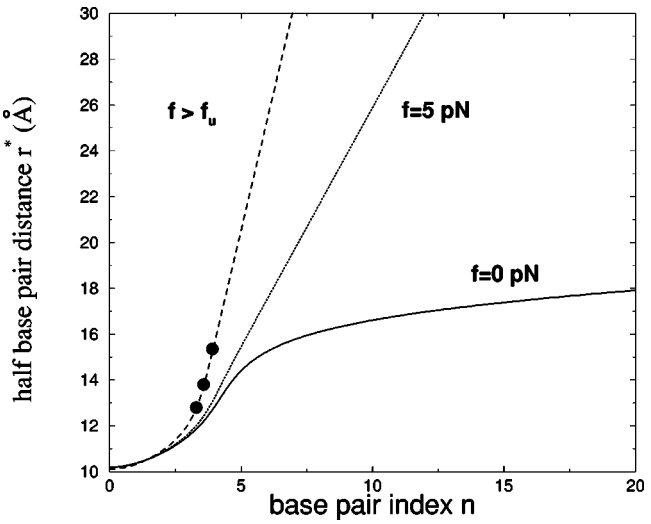

FIG. 12. Saddle-point trajectories $r^*(n)$ connecting dsDNA and ssDNA states, calculated with the instanton technique for different values of the unzipping force ($f=0,5,>12$ pN). For $f<f_u$, the trajectory goes to infinity [see Figure 10 (b)], and the size of the nucleation bubble is defined from the base-pair index n at which the slope changes, giving $N^*\simeq 3$. For $f>f_u$, all trajectories lie roughly on the same curve (dashed line), starting from $r_i\simeq r_{ds}\simeq 10$ Å and halting at some force dependent radius r_f at a finite bp index N^*, see Figure 10(a). The locations of $r_f(f)$ are shown for forces $f=12$, 15, 30 pN (top to bottom, heavy dots).

4. Saddle-point values of angular variables

In the previous calculation, we considered the Langevin dynamics using an effective Hamiltonian for $r(n)$ once angular degrees of freedom were integrated out. To be more complete, one might consider the coupled dynamics of twist angles $\theta(n)$ and radii $r(n)$. In doing so, the saddle-point configuration for the radius $r^*(n)$ turns out to be identical to the motion equation (40). The optimal solution $\theta^*(n)$ for the twist angle reads

$$\theta^*(n)=2\ \arcsin\left(\frac{\sqrt{L^2-H^2}}{2r^*(n)}\right), \tag{55}$$

where L, H are defined in Appendix A. Here, $\theta^*(n)$ decreases quickly from its dsDNA value $\theta_0\simeq 2\pi/10.4$ to zero as n runs from 0 to N^*.

C. The kinetic rate ν_0

So far we have discussed the most likely transition state for escape from a metastable state. To ascertain the corresponding time of escape, we need to know the rate ν_0; see Eq. (42). ν_0 is a rate describing the growth of the activated bubble along the unique descending path from the saddle configuration. Following Refs. [27], [28], we have

$$\nu_0=\frac{|\lambda^*|}{\zeta}, \tag{56}$$

where λ^* is the only negative eigenvalue of the Hessian matrix $\mathcal{M}(n,n')$ of $\mathcal{G}$ around the saddle point $\mathcal{R}^*$. The leading contribution to $\mathcal{M}$ is given by the second derivative of the Morse potential. $U_H''(r^*(n))\delta(n-n')$. The minimum of U_H'', reached in $r^*=R+\ln 4/a$ and equal to $-Da^2/4$, is our estimate for λ^*. We therefore obtain the rate $\nu_0=10^{12}\ \mathrm{s}^{-1}$.

Here we have focused only on the initial nucleation time, and have not discussed the time required for the unzipping ''fork'' to move down the molecule. For long repeated molecules, an additional contribution, of the order of N ($f>f_u$) or N^2 ($f=f_u$), accounting for the propagation of the bubble along the chain up to complete opening should be included in the net time [30].

IX. RESULTS FOR DYNAMICS OF UNZIPPING

We now discuss results of the theory discussed in the preceding section. We treat first the case where the force is held constant. We then discuss the case of constant loading rate, where the force is increased as a linear function of time.

A. Constant force

The average dissociation time $t_-(f)=1/\nu_-(f)$, is plotted in Fig. 6 for molecules of length $N=10$, 20, and 30 bp, for fixed unzipping force. Below the threshold force for unzipping ($f<f_u$), this ''lifetime'' is exponentially dependent on molecule length, which is to be expected from the growth of the total denaturation free energy as $\sim N$ (43).

At zero force, there is, in addition to the length-dependent part of the denaturation free energy, an activation free energy $G_0^*=13k_BT$. This leads to the unzipping time $t_-(N)=10^{-12}\ e^{13+1.4N}=10^{-6.3+0.6N}$ s (Fig. 6, inset). This formula corresponds to the time associated with spontaneous unzipping of an N-nucleotide double helix in free solution, and for $N=5$ bp is the relatively short timescale $t_-(5)=0.3$ ms. For a 30 bp DNA, $t_-(30)\approx 40$ years. Thus dsDNAs beyond ≈ 25 bp are essentially stable in free solution.

For $f>f_u$, $t(f)$ is, within the theory of the preceding section, independent of molecule length, and decreases exponentially with f up to $f=f_b=230$ pN. In this regime, once the initial force barrier is crossed, the applied force is large enough to continue unzipping the DNA. For forces beyond the force barrier $f>f_b$ dissociation is immediate (i.e., $t_-=1/\nu_0$).

The size of the nucleation bubble $N^*\simeq 4$ bp depends weakly on the force, and is essentially fixed by the shape of the barrier.

1. Linearization of the dissociation rate near f_u

The logarithm of the dissociation rate is generally not a linear function of force f (Fig. 6), as it would be for a single degree of freedom [29], but, rather, shows a strong quadratic dependence on f at small forces. This nonlinearity comes from Eq. (43) and the functional dependence of the polymeric chain free energy upon force, e.g., Eq. (10). However, for forces near the critical force f_u, an approximate linear relationship of this sort holds (Fig. 6).

The derivative of the dissociation rate with force is, in part, determined by the derivative of G^* (52). Below the unzipping force ($f<f_u$) we have

$$\frac{dG^*}{df} = \lim_{r_f \to \infty} 2\left(\frac{2f}{C}N^*(f,r_f) - [r_f - r_i(f)]\right)$$
$$= 2\left(\frac{2f}{C}N^*(f) - [r_t(f) - r_i(f)]\right). \quad (57)$$

To obtain this expression we have used expression (54). Here r_i depends on the force while r_t corresponds to the inflection point of $r^*(n)$ (Sec. VIII B 3).

For forces larger than the critical force ($f > f_u$), we use the boundary condition (53) to obtain

$$\frac{dG^*}{df} = -2[r_f(f) - r_i(f)]. \quad (58)$$

For $f \leqslant f_u$, see Eq. (43), we also need the derivative of the free energy difference $\Delta g(f)$ with force

$$N\frac{d}{df}\Delta g(f) = 2N\frac{f}{C}. \quad (59)$$

In the vicinity of f_u we therefore obtain

$$\ln t_-(f,N) = \begin{cases} -\ln \nu_0 + \beta G_u^* - 2\beta x_<(f - f_u) & (f \leqslant f_u) \\ -\ln \nu_0 + \beta G_u^* - 2\beta x_>(f - f_u) & (f \geqslant f_u), \end{cases} \quad (60)$$

where

$$x_< = d_u(N - N^*) + r_t(f_u) - r_i(f_u), \quad (61)$$

$$x_> = r_t(f_u) - r_i(f_u). \quad (62)$$

The energies and lengths entering into the linearized expression (60) have a straightforward physical interpretation. In the vicinity of the critical force ($f \approx f_u$), the ds and ssDNA states have nearly the same free energy. The transition state bubble separating them has free energy $G_u^* = 16.5 k_B T$; this is the free energy barrier involved in dissociation of a dsDNA (or for the reverse recombination reaction of two ssDNAs). The length $d_u = 5$ Å is the projection of a DNA monomer in the force direction, $N^* = 4$ bp is the number of base pairs in the nucleation fork, and $r_t(f_u) - r_i(f_u) = 4$ Å is the difference in radii between the first and last base of the fork. These lengths define the position of the transition state along the reaction coordinate. The total distance between the two DNA extremities in this transition state is $2[R + r_t(f_u) - r_i(f_u)] = 28$ Å.

B. Constant loading rate

1. Rupture force distribution

We now consider the situation where the unzipping force is a linear function of time. The time at which the molecule opens, or ruptures, is stochastic in nature, thanks to thermal fluctuations. The distribution of the rupture force as a function of the loading rate λ (measured in pN/sec) is given by [29]

$$\ln P(f) = -\frac{1}{\lambda}\int_0^f df'\, \nu_-(N,f') + \ln \nu_-(N,f), \quad (63)$$

where $\nu_-(N,f)$ is the dissociation rate at fixed force f and length N computed in the previous subsection. We recall that for $f < f_u$, ν_- has a strong N dependence, while for $f > f_u$ it is independent of N (43). The rupture force distribution is shown for $N = 10$ and for a few loading rates in Fig. 13.

2. Most probable rupture force

The most probable rupture force $f^*(N,\lambda)$ is given by the location of the maximum of the function (63) shown in Fig. 13, and is plotted in Fig. 7 for a few molecule lengths N. The curves for $N = 10$ can be understood in terms of the rupture force distributions of Fig. 13. For small loading rates (e.g., $\lambda = 0.5$ pN/s) $P(f)$ has a single maximum at $f = 0$, so the most probable rupture force f^* (Fig. 7) is zero. As loading rate is increased, a second local maximum appears for $\lambda \approx 10^{0.75}$, but still the most probable rupture force is zero. In this regime the loading rate is so slow that thermal fluctuations open the molecule before force can become significantly different from zero.

However, at $\lambda_1 = 10^{0.78} = 6$ pN/s, the second maximum at Fig. 13 exceeds the maximum at $f = 0$, and the most probable rupture force f^* (Fig. 7) jumps to $f^* = 2.5$ pN. As loading rate is further increased, f^* grows as the peak of Fig. 13 moves to higher force values. Near $\lambda = 4$ pN0/s the rupture force distribution (63) develops a cusp at $f = f_u$, due to the discontinuous derivative of $\nu_-(N,f)$ (see Fig. 6).

For $\lambda < \lambda_2 = 10^{4.6}$ (e.g., $\lambda = 4.4$ pN/s in Fig. 13) the maximum of the rupture force distribution is located on the part of the curve determined by the dissociation rates $\nu_-(N, f < f_u)$, for forces smaller than the critical force. In this range unzipping occurs while dsDNA is thermodynamically stable, so f^* shows a strong dependence on molecule length N.

For $\lambda_2 < \lambda < \lambda_3 = 10^{5.5}$ pN/s the rupture force distribution maximum is at the cusp of the curve in $f = f_u$, generating a plateau in the most probable rupture force of Fig. 7. For $\lambda > \lambda_3$ the maximum of the rupture force distribution is located at a force above f_u, where the dissociation rate increases with force, but is N independent. Thus f^* increases for $\lambda > \lambda_3$ until the loading rate approaches $\lambda_4 \simeq 10^{14}$ pN/sec. Beyond this point, the rupture force remains constant and equal to $f^* = f_b = 230$ pN. For $\lambda > \lambda_3$, unzipping occurs out of equilibrium, with the force increasing too rapidly for the molecule to respond. The loading rate λ_3 thus separates equilibrium and nonequilibrium timescales for constant loading-rate unzipping experiments.

Both λ_1 and λ_2 depend on N and diminish as N increases. For N sufficiently large, e.g. $N = 100$, f^* does not substantially differ from f_u for all loading rates $\lambda < \lambda_3$. The most probable rupture force is thus f_u for $\lambda < \lambda_3$, but becomes loading-rate dependent for $\lambda > \lambda_3$.

3. Linearization near critical force

The linearized result for the log of the dissociation rate (60) can be used to compute the dependence of the most probable rupture force on loading rate near the critical force f_u,

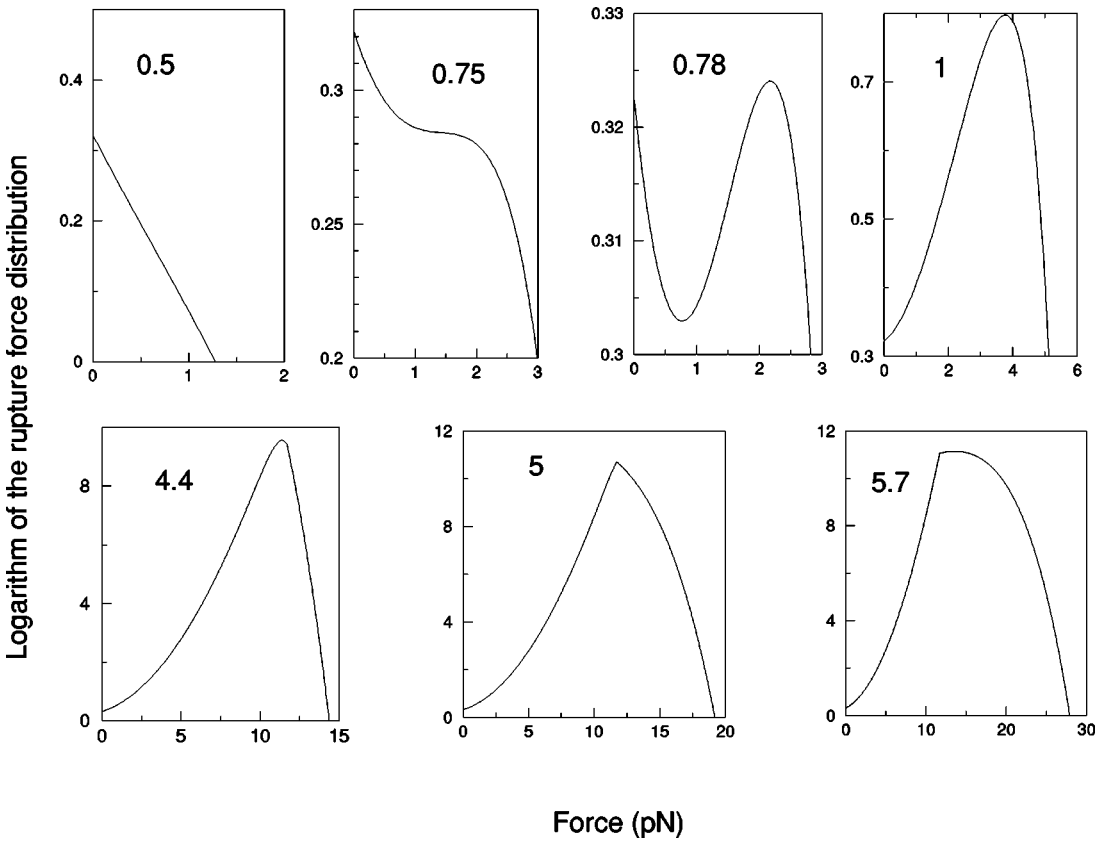

FIG. 13. Logarithm of rupture force distribution $P(f)$ for $N=10$ and different loading rates $\lambda=10^{0.5}$, $10^{0.75}$, $10^{0.78}$, 10^{1}, $10^{4.4}$, 10^{5}, and $10^{5.7}$ pN/s. For small loading rates, the maximum of the rupture force distribution is located at zero force, and jumps discontinuously to a finite force when $\lambda>\lambda_1=10^{0.77}$ pN/s. For $\lambda_2=10^{4.6}<\lambda<\lambda_3=10^{5.5}$ pN/s, $P(f)$ displays a cusp in $f=f_u$, which is thus the most probable rupture force. Above λ_3, a new maximum develops and the typical rupture force starts growing again as a function of λ. At very high loading rate $\lambda>\lambda^4=10^{12}$ pN/s, the rupture force reaches its highest value of $\approx$250 pN (not shown).

$$f^*(N,kv)=f_u+\begin{cases}\dfrac{k_BT}{2x_<}\ln|\lambda/\lambda_2|, & \lambda\leqslant\lambda_2=k_BT\nu_u/2x\\ 0, & \lambda_2\leqslant\lambda\leqslant\lambda_3=k_BT\nu_u/2x\\ \dfrac{k_BT}{2x_>}\ln[\lambda/\lambda_3], & \lambda\geqslant\lambda_3,\end{cases} \tag{64}$$

with $\nu_u\equiv\nu_-(f_u,N)$, and $x_<$ and $x_>$ as in Eqs. (61), (62). From Eq. (64), we see that λ_3 is inversely proportional to the time to cross the barrier separating dsDNA and ssDNA at coexistence, and to the difference in opening between the first and the last base pair of the transition state bubble.

X. COMPARISON WITH EXPERIMENTS

Many of the theoretical conclusions of the preceding sections are in accord with published experimental data. Our theoretical result for the critical unzipping force f_u is in accord with results of Bockelmann, Essevaz-Roulet, and Heslot [2] and Rief, Clausen-Schaumann, and Gaub [11]; our dissociation rate at zero force as a function of the number of base pairs N of Fig. 6 is in agreement with the experiments of Pörschke [8] and Bonnet, Krichevsky, and Libchaber [9], and the behavior of the rupture force f^* as a function of N and of the loading rate λ shown in Fig. 7 is in rough accord with data from AFM experiments on DNA unpairing by force by Strunz *et al.* [10].

A. Equilibrium and macroscopic unzipping

The experiments of Bockelmann, Essevaz-Roulet, and Heslot and of Rief, Clausen-Schaumann, and Gaub examine the opening of a large number of base pairs and according to Fig. 7 are performed with loading rate $\lambda<\lambda_2$. Therefore, as for the $N=100$ curve, the rupture force f^* at which the two strands separate is approximately equal to the critical force f_u, as is observed experimentally.

The macroscopic description of Eq. (9) within the Gaussian and FJCL approximations allows us to relate the critical force to the free energy of denaturation. From our Gaussian approximation we obtain $f_u^{\rm gaus}\simeq 12$ pN, and from the more microscopic model of Sec. V we obtain $f_u^{\rm FJCL}\simeq 11$ pN. These force values correspond to a total DNA denaturation free energy $\Delta g=1.4k_BT$. This critical force is in accord with the unzipping forces between 10 and 15 pN observed by Essevaz-Roulet, Bockelmann, and Heslot for λ-DNA [2]. Our critical force is also in the midrange of results of Rief, Clausen-Schaumann, and Gaub for homogeneous-sequence

DNA, $f_u^{AT}=9$ pN for poly AT poly TA, and $f_u^{GC}=20$ pN for poly CG poly GC.

Equation (9) with the f_u^{gaus} (10), f_u^{FJCl} (12) (see Fig. 8) leads to an estimate of the free energy of denaturation for different sequences, directly from the critical force values. The Bockelmann-Heslot range $10<f_u<15$ pN corresponds to values for the FJCL model in the range $1.3k_BT<\Delta g^{FJCL}<2.4k_BT$, and for the Gaussian model $1.0k_BT<\Delta g^{gaus}<2.3k_BT$. The λ-DNA used has a roughly random sequence, with GC content ranging between 30% and 60%. Using the FJCL expression for the single strand free energy Bockelmann and Heslot [2] have fit the experimental variations of the critical force along the λ DNA.

The critical forces found by Rief *et al.*—$f_u^{AT}=9$ pN and $f_u^{GC}=20$ pN—via Fig. 8 give $\Delta g_{gaus}^{AT}=0.8k_BT$, $\Delta g_{FJCL}^{AT}=1.1k_BT$, $\Delta g_{gaus}^{GC}=4.2k_BT$, and $\Delta g_{FJCL}^{GC}=3.5k_BT$, respectively. These values for the free energy of denaturation are compatible with the ones listed by Breslaner [3]: $\Delta g^{AT}\approx 2k_BT$, $\Delta g^{GC}\approx 5.6k_BT$ for a ionic concentration of 1 M NaCl buffer, which is greater than the 150 mM used in the experiments of Bockelmann and Heslot and of Rief, Clausen-Schaumann, and Gaub

The torque-force phase diagram of Fig. 3 gives a prediction for the critical unzipping force for an experiment that combines the unzipping experiment done in the group of Essevaz-Roulet, Bockelmann, and Heslot [2] and the twisting experiment done by Strick, Bensimon, and Croquette [15] and by Leger *et al.* [31].

B. Unzipping kinetics

Figure 6 shows the dissociation time t_- at zero force as a function of the molecule length. The dissociation time varies as $10^{-\alpha+\delta N}$, with

$$\alpha=\log_{10}\nu_0-\frac{G_0^*}{\ln 10}=6.3 \tag{65}$$

and

$$\delta=\frac{\Delta g}{\ln 10}=0.6. \tag{66}$$

These values are close to results obtained by Pörschke *et al.* for poly *A*-poly *U* DNAs, $\alpha=8$ and $\delta=0.5$. The value of δ is just $\Delta g=1.4k_BT$ only, and the result found by Pörschke gives $\Delta g=1.2k_BT$. The inverse width of the barrier b has been set to $b=1.75$ to obtain the opening time for $N=5$ bp $t_-\simeq 0.3\times 10^{-3}$ s found by Bonnet, Krichevsky, and Libchaber, $t_-\simeq 0.5\times 10^{-3}$ s.[3] Our calculations give the size of the nucleation bubble to be equal to $N^*\simeq 4$ bp, in agreement with the estimate of Pörschke. This value depends only weakly on applied force.

[3]For such a short sequence, no extra contribution to the time coming from the propagation of the opening fork along the chain is expected.

C. Dissociation under fixed loading rate

1. Opposing 5′-5′ strand unbinding experiment of Strunz et al.

In the experiment of Strunz *et al.* [10], the stretching force is applied to opposite 5′-ends. In our model, the stress is applied on the adjacent 5′ and 3′ ends of opposite strands [Fig. 1(c)]. This basic difference makes a precise, quantitative comparison of our results for the rupture force f^* as a function of the molecule length N and loading rate λ impossible.

However, the general features of the dependence of f^* upon N and λ shown in Fig. 7 should be quite robust, and applicable at least qualitatively to the experiments of Strunz *et al.* [10]. Experimentally the dependence of f^* upon N (Fig. 3 of Ref. [10]) was observed meaning that the loading rate is in the range $\lambda_1<\lambda<\lambda_2$. The experimental data show a linear dependence of the log of the dissociation rate on force close to f_u, as in our theory. From a linear fit to the data it is possible to deduce the characteristic length coefficients, activation free energy and critical force (see results of Sec. IX A 1).

For $f<f_u$, we write, as in Ref. [10] and in Eq. (60),

$$\nu(f)=\tilde{\nu}e^{\beta x f}, \tag{67}$$

where $x=2x_<$ from formula (60). However, we note that (in contrast to the interpretation of Ref. [10])

$$\tilde{\nu}=\nu_0 e^{-G_u^*+\beta x f_u} \tag{68}$$

is not the thermal dissociation rate at zero force, but instead is related to the activation energy of the transition state, and the length coefficients at the critical force. We write (again formally as in Ref. [10])

$$\tilde{\nu}=10^{a_1-a_2N}. \tag{69}$$

Again, we do not identify the fitted parameters a_1, a_2 as the zero-force dissociation parameters α (65) and δ (66) of the preceding section. Instead, we have

$$a_1=\log_{10}\nu_0-(G_u^*+\beta x_0 f_u)/\ln 10 \tag{70}$$

and

$$a_2=\beta x_1 f_u/\ln 10. \tag{71}$$

The length coefficients entering Eqs. (68) and (69) are $x=x_0+x_1N$ where $x_0=2[r_t(f_u)-r_i(f_u)-d_uN^*]$ and $x_1=2d_u$ from formula (60). The coefficients x_0 and x_1 have been also fitted by Strunz *et al.* from the slope of the linear fit of the loading rate. In the interpretation given by our kinetic theory, x_1 is the projection of the monomer length difference between the final ssDNA and the initial dsDNA states in the force direction; x_0 is the difference between the length $2(r_t-r_i)$ of the transition state along the force direction and the length d_uN^* of a ssDNA with the same number of bases of the activated bubble.

From the experimental data of Strunz *et al.* we find $x_1=0.7\pm 0.3$ Å and $x_0=7\pm 3$ Å. These values are different from the values we found in the case of unzipping, x_1

$=2d_u=10$ Å and $x_0=2[r_l(f_u)-r_i(f_u)-d_uN^*]=-32$ Å because they refer to a different geometry and a different kind of transition. Indeed, in the experiment of Strunz *et al.*, the force is applied along the molecular axis; as pointed out by the experimenters [10], the value $d_u \simeq 1$ Å is roughly the difference in axial length per base pair between overstretched *S*-DNA (which they consider as the initial state) and completely stretched and denatured ssDNA.

Finally, from the data of Strunz *et al.* we can deduce the stretching force at which the rupture force shows a plateau, $f_u=a_2/x_1 \ln 10=0.5k_BT/0.7$ Å $\simeq 70$ pN. The fact that f_u is larger than the value for transverse unzipping of ≈ 12 pN is due to the lower displacement and hence energy gain associated with longitudinal stretching, relative to that occuring during transverse unzipping [10].

2. Experiments of Rief, Clausen-Schaumann, and Gaub

Rief, Clausen-Schaumann, and Gaub observed a *B*-DNA to *S*-DNA transition, followed by a *S*-DNA to ssDNA transition, when DNA was stretched by a force acting on the opposite 5′-3′ ends of the same polynucleotide strand. In experiments on λ-DNA, the *B*-*S* transition force f^* showed no dependence on loading rate, so in that case f^* is the critical force f_u, and the experiments were, therefore, done in the regime $\lambda < \lambda_3$. In addition, the ds-*S* transition was found to have a highly sequence dependent f_u.

By contrast, the *S*-ss transition f^* depends on loading rate, indicating that the loading rate kv is larger than λ_3, indicating that $\lambda_3 > 6500$ pN/s. Thus f^* gives only an upper bound for f_u. These different behaviors for the *B*-*S* and *S*-ss transitions, obtained for the same loading rates suggest [see Eq. (64)], that the energy barrier is higher for the *S*-ss transition than for the *B*-*S* transition.

XI. CONCLUSION

In this paper we have analyzed the equilibrium and nonequilibrium aspects of unzipping of double-stranded DNA by applied force. This subject is of intrinsic theoretical importance, but is also highly relevant to the unpairing of DNA strands that occurs *in vivo*, and also directly addresses single-molecule micromanipulation unzipping experiments done under totally controlled conditions. We have mainly focused on a semimicroscopic model that accounts for the finite range of base-pairing interactions. This model has enough detail to analyze mechanical behavior at the few-Å scale relevant to molecular-biological events, but is still simple enough to allow a detailed mathematical analysis without having to resort to numerical simulations.

Our semimicroscopic model is constrained so as to have its large-scale equilibrium behavior in accord with ''thermodynamic'' descriptions of unzipping. In particular, this means that the hybridization free energy per base must be $\approx 1.4k_BT$ and that the elastic behavior assumed for the unpaired single-stranded regions must be consistent with results of direct micromanipulation experiments.

Given these thermodynamic constraints, our model naturally leads to an important conclusion, namely, that the semimicroscopic potential well that holds adjacent bases together in the double helix includes an appreciable *barrier*. The origin of this barrier is in the free energy difference due to nucleic acid backbone conformational fluctuations between dsDNA and ssDNA forms. This free energy barrier (about $2k_BT$ per base) adds to the base-pairing potential (about another $2k_BT$ per base in depth) to produce a $\approx 4k_BT$ total barrier that must be traversed over a few Å distance to separate two bases. This leads to a total force barrier of >250 pN.

This remarkably high force barrier is at first worrisome, since the forces observed during essentially ''macroscopic'' experiments are close to the ≈ 12 pN critical force for unzipping. However, this high barrier is difficult to directly observe since most experiments are carried out under conditions where distance between extremities cannot be controlled on the Å scale.

However, we have shown that the barrier to initiation of unzipping has been observed indirectly, in the initiation free energy for conversion of dsDNA to ssDNA. Our theory relates this well-known feature of DNA hybridization free energy to the base-pairing interactions and backbone elasticity. It would be of great interest to carry out AFM fixed-extension single-molecule experiments to directly measure the forces encountered in equilibrium as unzipping is initiated.

We have shown that a second way to gain insight into the unzipping initiation barrier is to study the kinetics of unzipping. We have treated this problem as a Kramers-type activated barrier-crossing problem, although one with many degrees of freedom. Our result is a rich kinetic theory, with results in accord with available experimental data [8,10]. However, many of our results await detailed study experimentally. Our kinetic theory underscores two important features. First, a transition state of ≈ 4 bp extent, which plays a critical role in the unzipping dynamics, has been identified. This ''fork'' structure is the analog of a transition state in a chemical reaction. One could imagine that this state could be a target for enzymes that accelerate DNA unzipping *in vivo*, since a fork of ≈ 4 bp is about the size of ''typical'' DNA-binding proteins. Second, a rich behavior of the most probable rupture force as a function of the loading rate and the sequence length has been obtained (Fig. 7), and has allowed us to interpret unbinding AFM experiments [10]. In this respect, our work extends Evans' theory [29] to the unbinding of polymeric objects, and complements recent works in the field [33].

The clearest limitation of our current analysis is that we do not treat the inhomogeneity in hybridization free energy due to inhomogeneous sequence along real DNAs. In the present paper, our aim was study of the initiation barrier, which is rather large compared to the effect that sequence can have at short scales. However, at large length scales, sequence inhomogeneity will play a crucial role in unzipping kinetics. Associated with this limitation is our neglect of the dynamics of fork propagation after unzipping initiation. Our Kramers-type treatment with its focus on the transition state ''fork,'' is appropriate for understanding only the initial stages of unzipping. This makes our theory adequate for describing the kinetics of denaturation of short sequence mol-

ecules. Rather distinct theoretical approaches closer to those used to study domain-growth kinetics are in order to study the motion of a ds-ss fork down a long (>100 bp) DNA.

Our results may be relevant to other problems, including the unzipping of helix-loop RNA and DNA structures, which have recently been studied using single-molecule micromanipulation techniques [34,35]. Also, our results are a starting point for consideration of more complex problems involving dsDNA–ssDNA–protein interactions. Both the role of proteins bound near an fork on unzipping, and the effect of unzipping on bound proteins are biologically relevant, and open to detailed study using single-molecule micromanipulation techniques.

ACKNOWLEDGMENTS

We thank C. Bouchiat for communication of his results prior to publication, and for useful discussions. Work at UIC was supported by NSF Grant DMR-9734178, by the Petroleum Research Foundation of the American Chemical Society, and by the Research Corporation. S.C. was partly funded by an A. della Riccia grant. R.M. was supported in part by the MRSEC Program of the NSF under Award number DMR-9808595.

APPENDIX A: ENERGY OF THE DISCRETE MODEL

In this appendix, we show how the discrete model of Refs. [5], [6] can be converted to Eq. (17), once the angular variables are integrated out, and a continuum limit is taken.

A. Radial transfer matrix

The discrete model transfer matrix has the form $T_0(r,r')=X(r,r')\ Y_0(r,r')$, with X and Y_0 given by formulas (10) and (12) of Ref. [6]. The radial portion X is

$$X(r,r')=\sqrt{rr'}\,\exp\{-\beta[V_s(r,r')+\tfrac{1}{2}U_H(r)+\tfrac{1}{2}U_H(r')]\}, \tag{A1}$$

where U_H is the Morse potential (denoted V_m in Ref. [6]). The backbone interaction is $V_s(r,r')=E(r,r')(r-r')^2$ with $E(r,r')\equiv E\exp[-b(r+r'-2R)]$ as in Eq. (5) of Ref. [6].

B. Simplifying form for the angular transfer matrix

Consider the angular part of the transfer matrix, and define the on-site potential $y(r)$ via

$$Y_0(r,r)=\frac{1}{r}\exp\{-\beta y(r)\}, \tag{A2}$$

$y(r)$ can be computed for large radii ($r\gg R$) from the definition of $Y_0(r,r)$,

$$Y_0(r,r)=\int_0^\pi d\theta\exp\beta[V_b(r,r,\theta)+\Gamma\theta] \tag{A3}$$

$$\simeq\int_0^L\frac{dx}{r}\exp[-\beta K(\sqrt{L^2-x^2}-H)^2]\left(1+\frac{\beta\Gamma x}{r}\right)$$

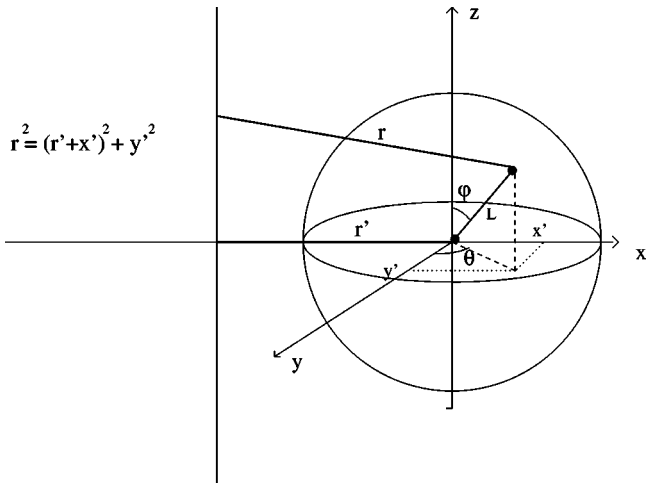

FIG. 14. Given the distance r' from one end of the rod to the left axis (helical axis of the molecule of dsDNA), the other end lies with a uniform probability on the sphere of radius d (denoted by L on the figure as in Appendix A) centered around the first end. We call r the distance of the second end to the axis. Let us call x', y' the coordinates of the projection of the second extremity onto the equatorial plane of the sphere, perpendicular to the left axis: $x'=d\sin\varphi\sin\theta$, $y'=d\sin\varphi\cos\theta$. The equation giving r as a function of r', x', y' is shown on left side of the figure. See text for further details.

$$(\theta=x/r), \tag{A4}$$

from which we obtain

$$y(r)\simeq-\Gamma\frac{R_1}{r},$$

$$R_1\equiv\frac{\int_0^L dx\,x\exp[-\beta K(\sqrt{L^2-x^2}-H)^2]}{\int_0^L dx\exp[-\beta K(\sqrt{L^2-x^2}-H)^2]}\simeq 6\ \text{Å}. \tag{A5}$$

A numerical check shows that Eq. (A5) is a good approximation to $-k_BT\ln Y_0(r,r)$, up to an r-independent additive constant, even when $r\simeq R$. Since small values of r are disallowed from contributing by the strong repulsive part of the potential, Eq. (A5) can be taken as a global approximation.

Now define

$$\hat{\rho}(r,r')=\frac{Y_0(r,r')}{Y_0(r,r)}. \tag{A6}$$

At large distance r, r', with $r-r'$ smaller or equal to the "rigid-rod" length L [5,6], $\hat{p}$ may be interpreted, up to an r, r'-independent multiplicative constant, as the joint probability ρ that the extremities of a rigid rod of length L be at distances r',r from a fixed reference axis (Fig. 14):

$$\rho(r,r')=\frac{1}{4\pi}\int_0^{2\pi}d\theta\int_0^\pi d\varphi\,\sin\varphi\,\delta(r-g(\theta,\varphi,r')), \tag{A7}$$

where

$$g(\theta,\varphi,r') = \sqrt{r'^2 + L^2\sin^2\varphi + 2Lr'\sin\theta\sin\varphi}. \quad \text{(A8)}$$

Integrating the angle θ out and defining $x=\cos\varphi$, we find

$$\rho(r,r') = \frac{2}{\pi}\int_0^1 \frac{dx}{\sqrt{h(x)}}\chi(h(x)), \quad \text{(A9)}$$

where

$$h(x) = \frac{4r'^2L^2}{r^2}(1-x^2) - \left(r - \frac{r'^2}{r} - \frac{L^2}{r}(1-x^2)\right)^2, \quad \text{(A10)}$$

and where $\chi(z)=1$ if $z\geqslant 0$, and 0 otherwise. In the large r, r' limit with $\bar{r}=r-r'$, we obtain

$$\lim_{r'\to\infty}\rho(r'+\bar{r},r') \equiv \rho(\bar{r}) = \frac{1}{\pi L}\int_0^{\sqrt{1-(\bar{r}/L)^2}} \frac{dx}{\sqrt{1-x^2-(\bar{r}/L)^2}}$$

$$= \begin{cases} \dfrac{1}{2L} & \text{if } -L<\bar{r}<L, \\ 0 & \text{otherwise.} \end{cases} \quad \text{(A11)}$$

The above calculation follows from the fact that single strands are represented as freely-jointed chains in our model.

For small wave numbers ($k<1/L$) relevant to a continuum representation of our model, the Fourier transform of ρ is approximately

$$\bar{\rho}(k) = \int_{-\infty}^{\infty} d\bar{r}\,\rho(\bar{r})e^{ik\bar{r}}$$

$$= \frac{\sin(kL)}{kL}$$

$$= 1 - \frac{L^2}{6}k^2 + O(k^4)$$

$$\simeq \exp\left(-\frac{L^2}{6}k^2\right), \quad \text{(A12)}$$

which is the Fourier transform of

$$\rho_{\text{gaus}}(r,r') = \frac{1}{\sqrt{2\pi\beta C_{\text{ss}}}}\exp\left(-\beta\frac{C_{\text{ss}}}{2}(r-r')^2\right), \quad \text{(A13)}$$

with

$$C_{\text{ss}} = \frac{3k_BT}{L^2}. \quad \text{(A14)}$$

This result simply indicates that the freely jointed chain has the long-wavelength behavior of a Gaussian polymer, ρ_{gaus}. Here C_{ss} is the three-dimensional Gaussian entropic stiffness of the single strand.

We determine the multiplicative constant $\hat{\rho}(r,r')/\rho_{\text{gaus}}(r,r')$ by requiring $\rho_{\text{gaus}}(r,r'=r)=1$ from Eq. (A6), and find

$$\hat{\rho}(r,r') = \exp\left(-\beta\frac{C_{\text{ss}}}{2}(r-r')^2\right). \quad \text{(A15)}$$

We finally find the following approximate expression for the transfer matrix; see Eqs. (A1), (A2), (A5), (A6), (A15):

$$T_0(r,r') = \exp\left\{-\frac{\beta}{2}\{(E\exp[-b(r+r'-2R)]+tC_{\text{ss}}) \times (r-r')^2 + U(r) + U(r')\}\right\}, \quad \text{(A16)}$$

where U is defined in Eq. (19).

C. Hamiltonian for radial variables

In the discrete version of the model of Ref. [5], [6] the single strands were assumed to be symmetrically distributed around the helical axis of the molecule (one strand was free to fluctuate while the other was just its axial reflection). Here, we resolve fluctuations of both strands, doubling the total length of independent segments, and leading to the effective entropic stiffness $C=2C_{\text{ss}}$.

Taking these contributions into account gives the transfer matrix

$$T_0(r,r') = \exp\left\{-\frac{\beta}{2}[m(r,r')(r-r')^2 + U(r) + U(r')]\right\}. \quad \text{(A17)}$$

The stiffness $m(r,r')$ (18) is a sum of dsDNA stacking interactions and intrinsic ssDNA entropic stiffness. There is a slight change of notation from Ref. [6], in the stacking function $E(r,r')$: both stacking constant E and inverse length b have been multiplied by a factor of two in the present paper. The effective Hamiltonian (17) for a configuration of radial variables $\mathcal{R}=\{r(0),r(1),\ldots,r(n)\}$ comes immediately from Eq. (A17).

APPENDIX B: CONTINUUM FORMULATION

A. Extension of *n* to a continuous variable

An alternative, nonsymmetric choice for the transfer matrix with the same eigenvalues as T_0 (A17) is

$$T_0^{\text{ass}}(r,r') = \exp\{-\beta[\tfrac{1}{2}m(r,r')(r-r')^2 + U(r)]\}. \quad \text{(B1)}$$

In a continuous version of this model, we look for a similar formula when the base-pair index is increased from n to $n+\epsilon$ with $\epsilon\ll 1$. In this regime, r and r' do not differ by more than $\sqrt{\epsilon}$, allowing some simplifications to be made. We now present two replacements T_1^{ass} and T_1 for T_0^{ass}, T_0 respectively, obtained from such approximations. T_1^{ass} and T_1 reduce to exactly T_0^{ass}, T_0 respectively, when the stiffness is r-independent.

Nonsymmetric model of this paper. Our continuum scheme uses an approximation to Eq. (B1). Assuming $r \simeq r'$, we transform the transfer matrix $T_0^{\rm ass}$ into

$$T_1^{\rm ass}(r,r') = \exp\{-\beta[\tfrac{1}{2}m(r)(r-r')^2 + U(r)]\}, \quad \text{(B2)}$$

where the stiffness is defined in Eq. (24).

Bouchiat's symmetric model. An alternative scheme was proposed by Bouchiat [32]. Let us define the function $\chi(r) = \int_R^r dx \sqrt{m(x)}$. Bouchiat considered the following symmetric transfer matrix:

$$T_1(r,r') = \exp\left\{-\frac{\beta}{2}\{[\chi(r)-\chi(r')]^2 + U(r) + U(r')\}\right\}. \quad \text{(B3)}$$

B. Schrödinger operator for the nonsymmetric model with zero force

The eigenvector equation for the transfer matrix $T_1^{\rm ass}(r,r')$ is

$$\int dr' T_1^{\rm ass}(r,r')\psi(r') = \exp(-\beta g)\psi(r). \quad \text{(B4)}$$

In the continuum limit of the axial distance h going to zero we introduce in Eq. (B4) a parameter ϵ which goes to zero as h,

$$\exp\left(-\beta\frac{\epsilon}{h}U(r)\right)\int dr' \exp\left(-\beta\frac{m(r)h}{\epsilon}(r-r')^2\right)\psi(r')dr' = \exp\left(-\beta\frac{\epsilon}{H}g\right)\psi(r). \quad \text{(B5)}$$

Since the Gaussian term concentrates the integral contribution of r' near r, we can expand the eigenfunction $\psi(r')$ near r. Introducing the integration variable $z = (r-r')/\sqrt{\epsilon}$, neglecting additive constants and expanding we obtain obtain

$$\left[1-\beta\frac{\epsilon}{h}\left(U(r)+\frac{1}{2\beta}\ln\frac{h\beta m(r)}{\epsilon\pi}\right)\right]\sqrt{\frac{2\beta h m(r)}{2\pi}} \times \int dz \exp(-\beta h m(r) z^2) \times \left[\psi(r) + \frac{\epsilon z^2}{2}\frac{\partial^2}{\partial r^2}\psi(r) + O(\epsilon^2)\right] = \left(1-\beta\frac{\epsilon}{h}g\right)\psi(r). \quad \text{(B6)}$$

Integrating the Gaussian terms, and putting back $\epsilon = h$ we obtain Eq. (23)

C. Schrödinger operator for Bouchiat's symmetric model

Bouchiat has shown that Eq. (B3) allows a simple change of variables from the radii r to the new variables χ. Then, the kinetic term in the χ variables has a constant stiffness and the corresponding Schrödinger operator can be written down. After returning to the r variables this is

$$\mathcal{O}_1 = -\left(\frac{k_B T}{\sqrt{2m(r)}}\frac{\partial}{\partial r}\right)^2 + U(r) + \frac{k_B T}{2}\ln m(r), \quad \text{(B7)}$$

This is symmetric, since the scalar product between functions φ_1, φ_2 reads $\langle\varphi_1|\varphi_2\rangle = \int d\chi\, \varphi_1(\chi)\varphi_2(\chi) = \int dr \sqrt{m(r)}\varphi_1(r)\varphi_2(r)$.

D. Relation between symmetric and nonsymmetric models

Defining $\overline{\psi}_1(r)$ to be the ground state wave function of Eq. (B7). Bouchiat has shown that $\psi_1(r) = [m(r)]^{1/4}\overline{\psi}_1(r)$ obeys Eq. (23) with an additional contribution to the potential V,

$$\delta V_1(r) = t^2\frac{5m'(r)^2 - 4m(r)m''(r)}{32m(r)^3}. \quad \text{(B8)}$$

The extra contribution is a small correction both inside and beyond the Morse well, and does not lead to a significant change to the values of $g_{\rm ds}$ and $g_{\rm ss}$. In addition, Bouchiat has shown that the presence of a kernel in the symmetric scalar product induces a small additional contribution $\delta f_1(r) = tm'(r)/[4m(r)]$ to the force $f_{\rm ds}(r)$ (Sec. VII A) needed to keep the two strands at fixed distance $2r$. As stated in Sec. V B, the transfer matrix $T_1^{\rm ass}$ we chose can be seen as a Gaussian approximation to the exact transfer matrix T_1 proposed by Bouchiat.

E. Schrödinger equation including force

The transfer matrix including force f is

$$T_f(r,r') = \exp\{-\beta[\tfrac{1}{2}m(r)(r-r')^2 + U(r) - 2f(r-r')]\}, \quad \text{(B9)}$$

and

$$\int dr'\, T_f(r,r')\psi(r') \simeq \exp\left\{-\beta\frac{\epsilon}{h}\left(U(r) - \frac{2f^2}{m(r)} + \frac{1}{2\beta}\ln\frac{h\beta m(r)}{\epsilon\pi}\right)\right\}\sqrt{\frac{2\beta h m(r)}{2\pi}} \times \int dz \exp\left\{-\beta h\left(\sqrt{\frac{m(r)}{2\sqrt{\epsilon}}}z - \sqrt{2}f/\sqrt{m(r)}\right)^2\right\}\left[\psi(r) + \sqrt{\epsilon}z\frac{\partial}{\partial r}\psi(r) + \frac{\epsilon z^2}{2}\frac{\partial^2}{\partial r^2}\psi(r) + O(\epsilon^2)\right], \quad \text{(B10)}$$

from which we obtain Eq. (34).

APPENDIX C: NUCLEATION BUBBLE LENGTH AND FREE ENERGY

We now derive simple expressions for the free energy G^* and length N^* of the nucleation bubble. Our aim is to obtain the dependence of these quantities on the inverse length b; this parameter is not critical to the description of macroscopic unzipping, but is central to the kinetics of unzipping initiation. In this appendix we calculate results at external force equal to the critical value f_u.

To simplify our results we approximate the potential V in the barrier region by the following piecewise linear form (Fig. 9):

$$V_{\rm app}=\begin{cases}\ln(E/C)(R+R_b-r)/R_b & \text{if } R<r<R+R_b,\\ 0 & \text{if } r>R+R_b,\end{cases} \tag{C1}$$

where

$$R_b=\frac{1}{b}\ln(E/C), \tag{C2}$$

is an approximate value for the width of the barrier. The nucleation bubble free energy G_u^* in the presence of an external force $f=f_u$ is

$$G_u^*\simeq\int_R^{R+R_b}dr\left(\sqrt{2m(r)[V_{\rm app}(r)-g_0]}-2f_u\right)$$
$$\simeq\frac{\sqrt{2C}}{b}\mathcal{G}(E/C,-g_0), \tag{C3}$$

where

$$\mathcal{G}(y,z)\equiv\int_1^y\frac{dx}{d}\left(\sqrt{(x+1)(\ln x+2z)}-\sqrt{2z}\right). \tag{C4}$$

The length of the activated bubble N_u^* can be computed similarly. An accurate approximation for the value of the radius at which the slope of $r^*(n)$ changes [i.e., where the third derivative of $r^*(n)$ vanishes] is $r_t=R+R_b$. Using Eqs. (40) and (51) we obtain

$$N_u^*\simeq\int_R^{R+R_b}dr\sqrt{\frac{m(r)}{2[V_{\rm app}(r)-g_0]}}\simeq\frac{\sqrt{C}}{b}\mathcal{N}(E/C,-g_0), \tag{C5}$$

where

$$\mathcal{N}(y,z)\equiv\int_1^y\frac{dx}{x}\sqrt{\frac{x+1}{\ln x+2z}}. \tag{C6}$$

Within the above approximation, we find $G_u^*\simeq(36/b)k_BT$ and $N_u^*\simeq(6/b)$ bp.

APPENDIX D: WKB CALCULATION FOR NUCLEATION BUBBLE FREE ENERGY

The free energy g_M and wave function $\psi_M(r)$ of the metastable phase obey Schrödinger equation (34). The WKB method can be used to calculate the tunneling amplitude of the wave function through the barrier. We note that we must introduce a new parameter to control the WKB calculation, since we are unable to consider arbitrary absolute temperature (the ssDNA phase free energy is dominated by thermal fluctuation, as is the barrier between dsDNA and ssDNA). We therefore introduce a parameter η, modifying Eq. (34) to be

$$\left[-\frac{(k_BT)^2}{2m(r)}\eta^2\frac{\partial^2}{\partial r^2}+\frac{2f(k_BT)}{m(r)}\eta\frac{\partial}{\partial r}+\hat{V}(r)\right]\psi_M(r)$$
$$=g_M\psi_M(r). \tag{D1}$$

Notice that Eq. (34) is recovered when $\eta=1$, and that a WKB approximation will be accurate for $\eta\to0$. The activation energy $G^*(\eta)$ will be computed for small η, and then η will be extrapolated to provide an estimate for $G^*(\eta=1)$.

The WKB approximation gives the wave function in the barrier range, $r_i<r<r_f$ defined through Eq. (53) under the semiclassical form

$$\psi_M(r)=\exp\left[-\frac{W(r)}{(k_BT)\eta}\right]\{\alpha_0(r)+\eta\alpha_1(r)+\eta^2\alpha_2(r)+\cdots\}. \tag{D2}$$

This result can then be put into into Eq. (D1), giving at lowest order in η the following quadratic equation for $W'(r)=dW/dr$:

$$[W'(r)]^2+4fW'(r)-2m(r)[V(r)-g_M]=0. \tag{D3}$$

Solving Eq. (D3) we obtain the activation energy $W(r)$ for any radius r inside the barrier, plus the corresponding wave function from Eq. (D2). The activation free energy G^* for leaving the metastable state equals $W(r_f)$ and is given by Eq. (52). The radii r_i and r_f are the classical turning points at which the ''momentum'' $W'(r)$ vanishes. We have checked the validity of the WKB approximation in the range $r_i<r<r_f$, see Sec. VII A.

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell*, 3rd ed. (Garland Press, New York 1994).

[2] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot, Proc. Natl. Acad. Sci. U.S.A. **94**, 11 935 (1997); U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, Phys. Rev. E **58**, 2386 (1998).

[3] K. Breslauer, Proc. Natl. Acad. Sci. U.S.A. **83**, 3746 (1986).

[4] A previous account of our main results was presented in S.

Cocco, R. Monasson, and J. F. Marko, Proc. Natl. Acad. Sci. U.S.A. **98**, 8608 (2001).
[5] S. Cocco and R. Monasson, Phys. Rev. Lett. **83**, 5178 (1999).
[6] S. Cocco and R. Monasson, J. Chem. Phys. **112**, 10 017 (2000).
[7] S. Cocco, Ph.D. thesis, University of Rome ''La Sapienza,'' 2000, available at http://ludfc39.u-strasbg.fr/cocco/
[8] D. Pörschke, J. Mol. Biol. **62**, 361 (1971).
[9] G. Bonnet, O. Krichevsky, and A. Libchaber, Proc. Natl. Acad. Sci. U.S.A. **95**, 8602 (1998).
[10] T. Strunz, K. Oroszlan, R. Schafer, and H. J. Guntherodt, Proc. Natl. Acad. Sci. U.S.A. **96**, 11 277 (1999).
[11] M. Rief, H. Clausen-Schaumann, and H. E. Gaub, Nat. Struct. Biol. **6**, 346 (1999).
[12] P. Cluzel *et al.* Science **271**, 792 (1996).
[13] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[14] J. SantaLucia, Jr., Proc. Natl. Acad. Sci. U.S.A. **95**, 1460 (1998).
[15] T. R. Strick, D. Bensimon, and V. Croquette, Genetica **106**, 57 (1999).
[16] P. G. De Gennes *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1985).
[17] R. E. Thompson and E. D. Siggia, Europhys. Lett. **31**, 335 (1995).
[18] H. Zhou, Y. Zhang, and Z. C. Ou-Yang, Phys. Rev. E **62**, 1045 (2000).
[19] T. Dauxois and M. Peyrard, Phys. Rev. E **51**, 4027 (1995).
[20] E. Prohofsky, *Statistical Mechanics and Stability of Macromolecules* (Cambridge University Press, Cambridge, 1995).
[21] H. Urabe and Y. Tominaga, J. Phys. Soc. Jpn. **50**, 3543 (1981).
[22] D. Cule and T. Hwa, Phys. Rev. Lett. **79**, 2375 (1997).
[23] D. K. Lubensky and D. R. Nelson, Phys. Rev. Lett. **85**, 1572 (2000); Phys. Rev. E **65**, 031917 (2002).
[24] S. M. Bhattacharjee, J. Phys. A **33**, L423 (2000).
[25] P. M. Morse, Phys. Rev. **34**, 57 (1929).
[26] D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951).
[27] J. S. Langer, Ann. Phys. (Leipzig) **34**, 258 (1967).
[28] J. S. Langer, Ann. Phys. (Leipzig) **41**, 108 (1969).
[29] E. Evans and K. Ritchie, Biophys. J. **72**, 1541 (1997).
[30] K. L. Sebastian, Phys. Rev. E **62**, 1128 (2000).
[31] J. F. Léger, G. Romano, A. Sarkar, J. F. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko, Phys. Rev. Lett. **83**, 1066 (1999).
[32] C. Bouchiat (private communication).
[33] D. Bartolo, I. Derenyi, and A. Ajdari, e-print cond-mat/0106543.
[34] C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith, Curr. Opin. Struct. Biol. **10**, 279 (2000).
[35] J. Liphardt, B. Onoa, S. B. Smith, I. T. Tinoco, and C. Bustamante, Science **292**, 733 (2001).