Overstretching and force-driven strand separation of double-helix DNA

Simona Cocco,[1] Jie Yan,[2] Jean-Francois Léger,[1] Didier Chatenay,[1] and John F. Marko[2]
[1]*Laboratoire de Dynamiques des Fluides Complexes, CNRS, 3 rue de l'Université, Strasbourg, France*
[2]*Department of Physics, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607-7059, USA*
(Received 8 October 2003; published 20 July 2004)

We analyze whether the "overstretched," or "*S*" form of double-stranded DNA consists of essentially separated, or essentially interacting, polynucleotide strands. Comparison of force-extension data for *S*-DNA and single-stranded DNA shows *S*-DNA to be distinct from both double helix and single-stranded forms. We use a simple thermodynamical model for tension-melted double-stranded DNA, which indicates that the overstretching transition near 65 piconewtons cannot be explained in terms of conversion of double helix to noninteracting polynucleotide strands. However, the single-strand-like response observed in some experiments can be explained in terms of "unpeeling" of large regions of one strand, starting from nicks on the original double helix. We show that *S*-DNA becomes unstable to unpeeling at large forces, and that at low ionic strength, or for weakly base-paired sequences, unpeeling can preempt formation of *S*-DNA. We also analyze the kinetics of unpeeling including the effect of sequence-generated free energy inhomogeneity. We find that strongly base-paired regions generate large barriers that stabilize DNA against unpeeling. For long genomic sequences, these barriers to unpeeling cannot be kinetically crossed until force exceeds approximately 150 piconewtons.

DOI: 10.1103/PhysRevE.70.011910 PACS number(s): 87.14.Gg, 82.39.Pj, 87.15.He, 82.37.Rs

I. INTRODUCTION

Several experimental groups have observed that double-stranded DNA (dsDNA) when placed under tensions exceeding about 65 piconewtons (pN), transforms from its usual *B*-form, to a new form approximately 1.7 times the *B*-form length [1–6]. The transition from *B*-DNA to highly stretched ("overstretched") DNA is remarkably sharp, occurring between roughly 60 and 70 pN, depending on ionic strength [7]. This stretched form of DNA, sometimes called "*S*-DNA," can support tensions of up to 400 pN [6], and is stable at 80 pN for $\approx$10 min [7], suggesting that under some solution conditions it is a distinct, stable state of the double helix.

Rouzina and Bloomfield [9,8] have developed a theory of tension-driven strand separation ("tension-melting") of ds-DNA, and in collaboration with Wenner and Williams [10] applied it to analysis of single molecule experiments. They concluded that overstretched DNA is in fact strand separated [see Fig. 1(a)], i.e., that 65 pN tension depresses the melting temperature of the double helix to room temperature. A series of experiments by Wenner *et al.* showed that changes in solution conditions which favor melting of the double helix all reduce the force needed for overstretching [10]. Wenner *et al.* interpret these results as indicating that *S*-DNA is tension-melted DNA.

In this paper we examine the tension-melting hypothesis of Rouzina *et al.*, in the light of data from a number of experiments. We first consider data for the tension versus extension of single-stranded DNA (ssDNA), which shows that the force response of *S*-DNA does not match either that of one ssDNA, or that of two, parallel ssDNAs (Sec. II). This comparison indicates that *S*-DNA is not simply ssDNA.

Second, we use the approach of Rouzina and Bloomfield [9,8] to construct the free energy of tension-melted dsDNA (Sec. III). We consider two possible tension-melted forms: (a) parallel-separated strands where tension is supported by both ssDNAs [Fig. 1(a)], and (b) "unpeeled" DNA where one strand has fallen off the other, with tension therefore supported by one ssDNA [Fig. 1(b), part *C*]. The parallel-separated form has been suggested by Wenner *et al.* [10], while the unpeeled form has been suggested to occur at low ionic strength by Smith *et al.* [2]. For these two states, we

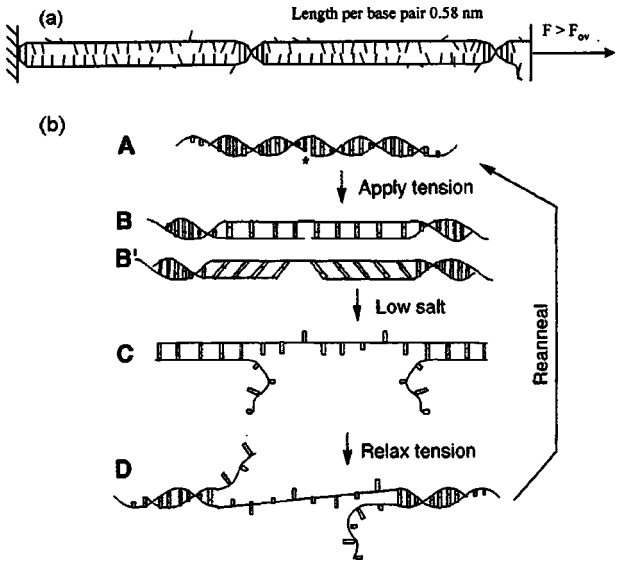

FIG. 1. Proposed structures of overstretched dsDNA. At 150 mM NaCl and *p*H 7.5, the double helix abruptly increases in length by about 70% when a force of about 65 pN is applied. (a) Wenner *et al.* [10] propose that the lengthened DNA is "tension melted" or strand separated, thus with half the total applied tension supported by each ssDNA (reproduced from Ref. [10]). (b) Smith *et al.* [2] propose that at high salt, the lengthened DNA is still base paired (*B*,*B*'); they also propose that at low salt unpeeling from nicks may occur (*C*) which can lead to strong hysteresis when tension is reduced (*D*) (reproduced from Ref. [2]).

use ssDNA elasticity data to estimate their stretching free energies.

The theory of Sec. III uses interstrand binding free energy estimated directly from DNA "unzipping" experiments [4,6,11–13] where dsDNA is mechanically converted to two ssDNA by direct mechanical pulling. These experiments provide for us, in a model-independent way, a value for the free energy difference between dsDNA and separated ssDNA strands at ambient temperature [16–18]. We then show that our results based on DNA unzipping data are consistent with predictions of free energy models based on DNA melting [14]. Therefore, unzipping and melting experiments provide independent and consistent estimates of the free energy cost of separating paired bases.

We consider the two tension-melted forms of DNA to be distinct from the third possibility of a novel stretched and unwound double helix (*S*-DNA), where the two strands are bound together [Fig. 1(b), part *B*]. This model has been suggested by Cluzel *et al.* [1] and Léger *et al.* [5,7,19], based on the observation that the state found at tensions of 60–140 pN is mechanically distinct from ssDNA [7]. Measurements of a well-defined $\approx$35 base pair (bp) helical pitch for overstretched DNA [19,5] also suggest that it has a regular secondary structure, and is not simply disordered ssDNA strands. A further piece of evidence follows from observations of *two* force-plateau transitions, the second higher-force one leading to a ssDNA-like force response [12,4]. This implies that the lower-force ($\approx$65 pN) transition is not strand unpairing.

The notion that overstretched DNA has a secondary structure distinct from that of unpaired stretched ssDNA strands has been theoretically examined by Lebrun and Lavery [20] using molecular modeling, and by Cizeau and Viovy [21] using coarse-grained models. Recently Storm and Nelson [22] have analyzed experimental data to show that *S*-DNA has a bending persistence length $\approx$10 nm, much larger than the 0.7 nm persistence length of ssDNA [2].

We thus use experimental data to estimate the free energy of overstretched DNA, and we compare our results with those for tension-melted forms of DNA. For physiological solution conditions (room temperature, pH 7.5, 150 mM NaCl) we find that near 65 pN *S*-DNA is more stable than either separated or unpeeled ssDNA states, for the average sequence composition of λ-DNA (typical of genomic DNA). However, at higher forces $\approx$120 pN, *S*-DNA becomes unstable to unpeeling of one strand. Thus, for DNA under physiological conditions, the 65 pN transition is from *B* to a distinct *S*-state of DNA. Depending on precisely how many nicks (intermittent breaks along the backbones) are found along a molecule one may observe a second, higher-force transition to unpeeled DNA with ssDNA-like elastic response, as observed in some experiments [4,6,12].

We also find that at physiological conditions, *S* and unpeeled DNA states preempt tension-driven strand separation. However, if the transitions to *S* and unpeeled DNA are suppressed, it may be possible to observe separated strands. This provides a rough microscopic picture for the transition to a strand-separated and highly overwound form of the double helix called "*P*-DNA," observed by Allemand *et al.* [23]. In those experiments on unnicked molecules (forbidding unpeeling) double helix twist was constrained, suppressing formation of *S*-DNA which is known to be unwound relative to *B*-DNA [5].

We have generalized our model to other ionic conditions. At low salt, ssDNA elasticity, the strength of base-pairing interactions, and the free energy of *S*-DNA are all modified. With these effects added, our model predicts a force vs salt-concentration phase diagram. At low salt ($<$25 mM NaCl), unpeeling of *B*-DNA preempts formation of *S*-DNA. At higher salt we find two transitions: a low-force *B* to *S* transition, and a higher force *S* to unpeeling transition.

Effects of base-pair composition are then discussed. The results mentioned above are for a sequence-averaged model with AT/GC composition close to that of λ-DNA (adenine and thymine, or A and T, are less strongly base paired than guanine and cytosine, or G and C). Using AT- or GC-rich DNA will shift the base-pairing energies appreciably and modify our results. We are able to quantitatively understand experiments of Rief *et al.* [6,4] which show that AT-rich DNA goes through a single unpeelinglike transition, while GC-rich DNA is observed to display formation of *S*, and then unpeeling, as force is increased.

We consider the sequence-dependent free energy of the tension-separated parallel-strand state (Sec. IV) in more detail. Using equilibrium statistical mechanics, we show that although some small "bubbles" of locally strand-separated DNA may form below the overstretching transition, these cannot account for the large length change that occurs around 65 pN. We employ a model of base-pair interactions [14] widely used to gauge stability of double-stranded nucleic acids. The only element of the "classical" theory of DNA melting that we do not include is the long-range interaction introduced by the constraint that ssDNA bubbles internal to a long dsDNA have associated with them a loop-closure entropy contribution [15]. Rouzina and Bloomfield have noted that cooperativity generated by loop entropy is greatly reduced by applied force [8], making it a much less important physical effect for tension melting than for purely thermal DNA melting. This, plus the fact that in the tension-melting case the free energy of ssDNA is dominated by its extensive stretching free energy, makes it reasonable to neglect the logarithmic loop interaction.

Finally (Sec. V), we consider the dynamics of unpeeling of *B*-DNA. As in DNA unzipping [18,24–27], unpeeling one strand from a double helix will require passing over a series of sequence-generated free-energy barriers. We first introduce a model for unpeeling which includes sequence dependence. We then analyze the dynamics of unpeeling for forces near 65 pN and show that the sequence-generated barriers cannot be dynamically crossed. This provides further evidence for our major conclusion that *S*-DNA is stable near 65 pN, but now accounting for both kinetic and sequence-inhomogeneity effects. At higher forces near 150 pN, sequence-generated barriers to unpeeling can be crossed. During unpeeling elongation rate-dependent force-extension occurs, similar to that observed experimentally [6,4]. When unpeeling occurs, our dynamical model generates force-extension hysteresis similar to that observed in a few different experiments [2–4,6,7,10].

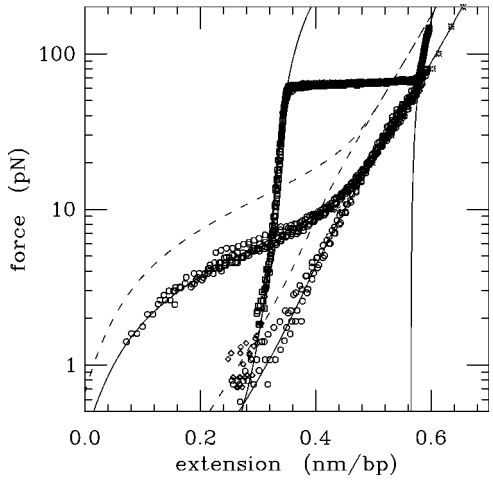

FIG. 2. Force vs extension of ds and ssDNA. Squares show experimental dsDNA data of Léger *et al.* [7,5] for 500 mM NaCl buffer, diamonds show experimental dsDNA data of Smith *et al.* [2] for 1 M NaCl buffer. Data for physiological salinity (150 mM NaCl) are similar, but have a plateau shifted a few pN below the 500 mM result, see Refs. [10,7]. Circles show experimental data of Bustamante *et al.* [12] for ssDNA; stars show high-force ssDNA data of Rief *et al.* [4]. The left, lower-extension curve is for 150 mM NaCl, while the right, higher-extension curve is for 2.5 mM NaCl. The two ssDNA datasets converge at high force, to the behavior $x \approx \ln f$. Solid curves show the theoretical models for *B*-DNA, *S*-DNA, and single ssDNAs used in the text. The dashed curves indicate the force response of two parallel ssDNAs for 150 mM (left) and 2.5 mM (right) NaCl.

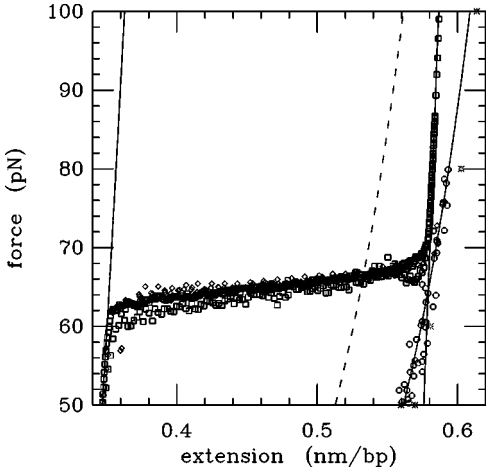

FIG. 3. Force vs extension of ds and ssDNA as in Fig. 2 showing 65 pN force plateau and adjacent states. Above the *B* to *S* transition, the elastic behavior of *S*-DNA is not in accord with either that of one ssDNA (solid curve following circles), or of two ssDNAs (dashed curve).

II. COMPARISON OF *B*-DNA, *S*-DNA, AND ssDNA

A. *B*-DNA elasticity

Figure 2 shows experimental data of Léger *et al.* [7,5], showing the *B* to *S* transition of a single EMBL3 λ -DNA 44.2 kb in length in phosphate buffer of pH 7.4 with 500 mM NaCl (squares). In the same plot data of Smith *et al.* [2] are shown, for overstretching of a Sam7 λ-DNA 48.5 kb in length in Tris buffer with 1 M NaCl (diamonds). Extension is given in nm/bp units; the contour length of *B*-DNA is 0.34 nm/bp. These data agree on the lengths per base pair of the *B* and *S* forms, suggesting that systematic errors in the length measurements are small.

For forces below 65 pN, the molecule is in the *B* form, with force response of a persistent chain polymer with persistence length of roughly 50 nm, with slight contour-length elastic extensibility [28]. Chain bending fluctuations generate the nonlinear low-force elasticity, while elastic stretching of the double helix generates the roughly linear stretching between 20 and 50 pN, with a spring constant per base pair of $f_B/h \approx 1200/0.34$ pN/nm. The force f_B is often quoted as the stretching elastic constant of the double helix. The solid curve following the squares for forces below 40 pN in Fig. 2 is the extension per base pair,

$$x_{\mathrm{ds}}(f) = h\left[1 - \frac{1}{\sqrt{4\beta A f}} + \frac{f}{f_B}\right] \tag{1}$$

for *B*-form persistence length $A = 50$ nm, force constant $f_B = 1230$ pN, and contour length per base pair $h = 0.34$ nm. This formula includes the first correction in inverse powers of f from bending fluctuations, plus linear helix stretching.

Note that the force constant (f_B for *B*-DNA) corresponds to the slope of the force vs extension curves of Fig. 2, multiplied by the *B*-form length per base pair ($h = 0.34$ nm/bp).

B. Transition to *S*-DNA

At 500 mM and 1 M salt concentrations, just above 60 pN, the double helix changes from about 1.05 to 1.7 times *B*-form length (from about 0.36 to 0.58 nm/bp, see Figs. 2 and 3). Beyond this "*B* to *S*" transition, a stiff force response occurs, characterized by a spring constant per base pair of $\approx 1600/0.34$ pN/nm, more than the force constant on the *B* side of the transition. *S*-DNA can be extended to more than 1.8 times the original double helix length (0.62 nm/bp) by forces near 200 pN. At these forces the single-molecule tethers have a lifetime of a few seconds at most, necessitating quick experiments. However, at lower forces ≈80 pN, Léger has reported *S*-DNA to be stable for greater than 10 min in 500 mM NaCl [7].

There is a gradual reduction of the transition plateau force with decreasing salt concentration. Over 1–0.100 M NaCl this drop is about 5 pN per decade of salt concentration [10,7]. In this range of salt concentration, little hysteresis is observed for molecules which do not contain large numbers of nicks [7]. For lower salt concentrations (0.1 M to 1 mM univalent salt), the plateau drops down faster with salt, closer to 6 pN per decade reduction in salt [10]. In this lower-salt regime, appreciable hysteresis is often observed [10].

C. ssDNA

Figure 2 also shows data of Smith *et al.* [12] for the elastic response of a single strand of the 48502-bp Sam7 λ DNA (circles). Results for Tris buffer with NaCl concentrations of 2.5 mM (left curve) and 150 mM (right curve) are

shown. Raw data in microns have been divided by the sequence lengths to obtain nanometers per base pair (nm/bp). Also shown are high-force measurements of Rief *et al.* [6,4] (stars) which have been converted to bases/nm using a factor of 0.34 nm/base. Although we do not include them in this figure, data of Dessinges *et al.* [29] are consistent with the data shown.

A single strand can be extended to about 0.7 nm/base by high forces, but for low forces ssDNA has less extension than a *B*-DNA of the same sequence length. There is no overstretching transition for ssDNA, but instead a gradual extension occurs. The ssDNA extension depends strongly on ionic strength; for low concentration of NaCl (<10 mM) the extension goes roughly as log of force; for high NaCl concentration, a force threshold of ≈ 2 pN must be passed for appreciable extension to occur. This effect is due to a combination of screening of electrostatic self-repulsion, and sticking of the exposed bases to one another [31,30].

The solid curves passing through the ssDNA data are a phenomenological model for the extension per base pair $x_{ss}(f)$. This model is described in the Appendix, and includes the effect of NaCl concentration. It is a smooth and convenient representation of the experimental result over forces up to 200 pN, for NaCl concentrations from 1 mM to 1 M. Note that there is variation in x_{ss} with sequence [30,7], and that the particular form assumed here describes the 48502-bp λ -DNA studied by Bustamante *et al.* [12].

D. Comparison of *S*-DNA to ssDNA

At 65 pN, the length of ssDNA is close to 1.7 times *B*-form length (0.58 bp/nm), close to the length of *S*-form DNA just past the *B* to *S* transition. This coincidence suggests that the *S*-form might be ssDNA, especially when the data are plotted as in Fig. 2. However, this line of reasoning has a few problems.

Figure 3 shows the experimental data of Fig. 2, replotted to focus in on the *S* end of the transition. Again, boxes show dsDNA data of Léger *et al.*, diamonds show dsDNA data of Smith *et al.* [2], and circles show ssDNA data of Bustamante *et al.* [12]. The *B* to *S* transition ends at an extension of about 0.58 nm/bp and a force of 68 pN, again coinciding with the ssDNA data. However, the ssDNA spring constant per base (280/0.34 pN/nm) is less than one-fifth of that of *S*-DNA. Thus *S*-DNA is not easily modeled using the ssDNA force response.

A further problem with this ssDNA interpretation is that the coincidence of extensions at 68 pN occurs for *S*-DNA and *only one* polynucleotide strand. The double helix consists of two strands, and therefore the most reasonable ssDNA model for *S*-DNA should have two parallel and noninteracting strands, with a force response

$$x_{2ss}(f) = x_{ss}(f/2). \tag{2}$$

Figure 3 includes this response (dashed curves); this shows that if two parallel and noninteracting ssDNA's were stretched to extensions of 1.7(0.58 nm/bp), a total force *double* that of one ssDNA would be required, more than 100 pN. Alternately, at 65 pN the extension of two parallel ssDNAs can only be about 1.5 (about 0.51 nm/bp). Finally, the force constant of two parallel ssDNA's is only about one-third that of *S*-DNA. Two parallel noninteracting ssDNAs cannot quantitatively explain the mechanical properties of *S* -DNA.

Without any thermodynamical or statistical-mechanical analysis, examination of experimental data indicates that *S* -DNA has mechanical properties distinct from ssDNA. One might argue [8–10] that *S*-DNA is made up of a mixture of relatively large islands of separated ssDNA, and remnant base-paired *B*-DNA. In this "*B*-ss" scenario, the overstretched state should have a force constant between that of *B* and ssDNA, again less than that observed. So, a simple mixed-phase picture of *S*-DNA is problematic as well.

III. FREE ENERGY OF TENSION-DRIVEN MELTING OF DNA

In this section we discuss a simple free energy model for tension-driven melting of DNA, using the approach of Rouzina and Bloomfield [9,8]. We consider for the moment solution conditions of 150 mM NaCl, and sequence-averaged free energies. From force-extension data the free energy per base pair (per base for ssDNA) at constant force, can be obtained by integration of extension:

$$w(f) = \int_0^f df' x(f'). \tag{3}$$

This can be carried out either for ssDNA or for dsDNA, and gives a free energy relative to zero force. From a statistical-mechanical point of view, Eq. (3) is the log of the partition function of the molecules in the fixed-force ensemble [28]; note that more positive values of this free energy indicate more probable states. Thermodynamically, Eq. (3) is the Legendre transform of the Helmholtz-free-energy-like mechanical work $\int dx f$ naturally defined as a function of extension [8]. Note that the molecules are tethered, so there is no concentration dependence to these free energies. Figure 4 shows the free energy for ssDNA, $w_{ss}(f)$, obtained by numerical integration of $x_{ss}(f)$ for 150 mM NaCl.

A. Relative free energies of *B*-DNA and ssDNA from unzipping

From the definition (3), at zero force we have $w_{ss}(0)=0$ and $w_B(0)=0$. To be treated as relative free energies, we must include the base-pairing free energy of the double helix. We introduce this as a positive free energy per base pair g_0. Thus relative to $w_{ss}(f)$, the free energy of *B*-DNA is $w_B(f) + g_0$.

We can estimate g_0 from experimental data for *unzipping* of DNA, where the two strands are pulled apart from one another [4,11,13]. This type of experiment is unique in providing information about the relative free energies of equithermal ssDNA and *B*-DNA. Previously, estimates of this free energy difference came from studies of thermal DNA melting [32,14], via potentially inaccurate extrapolation of free energies measured near melting points typically >70 C, down to room temperature.

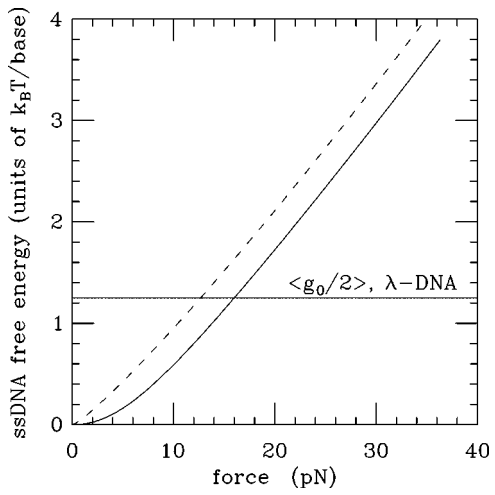

FIG. 4. Free energy per base of ssDNA $w_{ss}(f)$, obtained by numerical integration of extension (see text). Results are shown for 2.5 mM (left,dashed) and 150 mM NaCl (right,solid). Horizontal line at $1.25k_BT$ indicates half the average base-pairing free energy per base for λ-DNA at 150 mM, and intercepts the 150 mM ssDNA curve at the experimentally observed mean unzipping force of 15 pN. Variations in 150 mM NaCl unzipping forces from 10 to 20 pN observed experimentally indicate that the base-opening free energies vary with sequence from 1.2 to $3.4k_BT$ per base pair.

The free energy per pair of unzipped bases (under tension f), relative to that of zipped *B*-DNA (which is not under tension along its length) is just

$$\Delta g_u(f) = 2w_{ss}(f) - g_0. \tag{4}$$

When this quantity is negative, *B*-DNA is stable, the situation at zero force, where $g_0 > 0$ and $w_{ss}(0) = 0$. When Eq. (4) is positive, ssDNA is stable, which is guaranteed to occur for sufficient force since $w_{ss}(f)$ rises monotonically. The force f_u where $g_u(f_u) = 0$, or when $w_{ss}(f_u) = g_0/2$ is the threshold for unzipping to occur. Unzipping experiments on λ-DNA find $f_u = 15$ pN for 150 mM Na^+ buffer. This allows us to determine $g_0 = 2.5k_BT$ (Fig. 4).

1. Sequence effects

The same considerations applied to the sequence-dependent variations of force observed experimentally [4] indicate that the most tightly bound sequences unzip for forces close to 20 pN, giving $g_{0,GC} = 3.4k_BT$. This can be seen by examining Fig. 4, and locating the free energy of ssDNA for 150 mM NaCl at 20 pN; by Eq. (4) this is half the unzipping free energy per base. The most weakly bound sequences have unzipping forces close to 10 pN, giving via Fig. 4 $g_{0,AT} = 1.2k_BT$. In Sec. IV we will show that these g_0 estimates obtained from unzipping experiments agree well with base-pairing free energies obtained from analysis of thermal DNA melting [14].

Although there is almost no dependence of the length per base pair of double helix on sequence, there is evidence of appreciable dependence of the elasticity of ssDNA on its AT fraction. The ssDNA data of Fig. 2 are for λ-DNA which is almost exactly 50%AT; most of the analysis of this paper is applied to this molecule. However, the 70% AT 156GMac DNA studied by Léger *et al.* [7] has an extension versus force which is about 7% longer than that of λ-DNA. If this effect occurs proportionally with AT fraction for a 100%AT DNA, the ssDNA stretching free energy (3) would be almost 20% higher than the λ result shown in Fig. 4. Consequently the estimate of g_0 for pure AT from unzipping experiments would be shifted up to $1.4\ k_BT$.

No data are at present available to test the question of whether highly GC-rich ssDNAs are proportionally shorter. However, if pure GC-ssDNA is 20% shorter than λ-DNA, its stretching free energy would be about 20% less than that shown in Fig. 2. The value of g_0 inferred from unzipping experiments in this case would by shifted down to $2.8\ k_BT$. These shifts in the ssDNA free energy are appreciable, but do not lead to large quantitative changes in our results. We do not consider the sequence dependence of ssDNA length further below.

We emphasize that the estimates of strand-separation free energy (g_0 per base pair) of this section have been computed independently of any particular detailed theory of DNA "melting," usually studied at high temperatures >50 C. Instead we have used only experimental unzipping data at room temperature. In Sec. IV we will show that well-established theories of the detailed sequence dependence of base-pairing free energies are consistent with the results of this section.

B. Tension-driven melting of *B*-DNA

1. Unpeeling of one strand

Using the free energies of ss and *B*-DNA we can calculate the free energy of tension-melted states, following the approach of Bloomfield and Rouzina [8]. First, we consider conversion of a stretched *B*-DNA to one single strand, via unpeeling of the other strand. The free energy of one stretched strand plus the free energy of one relaxed (unpeeled) strand (which has been chosen to be zero), relative to stretched double helix is

$$\Delta g_{\text{unpeel}} = w_{ss}(f) - w_B(f) - g_0, \tag{5}$$

which is plotted for $g_0 = 2.5k_BT$ (Fig. 5, solid).

This free energy difference passes through zero at 62 pN, about where the *B* to *S* transition is observed. Using the base-pairing free energy sequence dependence obtained from analysis of unzipping (g_0 between $1.2k_BT$ and $3.4k_BT$), in the absence of *S*-DNA, unpeeling will occur for a range of forces from 40 pN (in Fig. 4, where the unpeeling free energy of Fig. 5 is about $-1.3k_BT$) to 80 pN (where the unpeeling free energy is about $+0.9k_BT$).

Thus, it is possible to observe tension-driven unpeeling where one strand falls off the other, over a broad range of forces (40–80 pN). This is a likely situation when there are plenty of nicks along a duplex, allowing the melting away of weakly bound regions starting at around 40 pN. However, complete melting away of one strand will require one to sit at a force above 80 pN (note that the free energy of *S*-DNA will be less than that of *B*-DNA). If experiments are done on molecules with no or only a handful of nicks, then only those

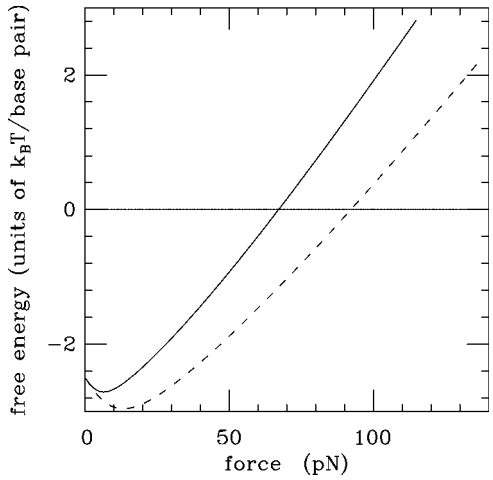

FIG. 5. Free energy per base pair of unpeeled (solid line) and parallel-strand-separated (dashed line) DNA, relative to *B*-form double helix, for 150 mM NaCl. The curves are calculated for *B*-DNA base-pairing energy of $2.5k_BT$, the average value for λ-DNA at 150 mM NaCl. When these free energy curves reach -1.3 and $+0.9k_BT$, *B*-DNA regions with base-pairing energies of 1.2 (AT rich) and $3.4k_BT$ per base pair (GC rich) become unstable. Unpeeling of one strand occurs over a broad force range from 43 pN for AT-rich DNA, to 85 pN for GC-rich DNA. True tension-driven melting, or parallel-strand separation, occurs over a higher force range, from 65 pN (AT rich) to 110 pN (GC rich).

nicks which are in or very near to large patches of AT-rich DNA are going to cause an observable extension signal. In Sec. V we will show that for heterogeneous sequences unpeeling will occur out of equilibrium, as an essentially irreversible process, due to the different unpeeling force associated with AT and GC rich regions.

2. Parallel separated strands

A second possibility is that tension-driven melting occurs, leaving two parallel, isolated strands. Under the assumption that there are no interactions between the two strands, the relevant free energy difference is between two single strands each under tension $f/2$ and the double helix under tension f:

$$\Delta g_{\text{unpair}} = 2w_{\text{ss}}(f/2) - w_{\text{B}}(f) - g_0. \tag{6}$$

This free energy difference (Fig. 5, dashed line, again for $g_0 = 2.5k_BT$) passes through zero at a force of about 93 pN. For weakly and strongly bound sequences the melting points are 60 and 110 pN, corresponding to where Δg_{unpair} crosses -1.3 and $0.9k_BT$.

In the second "unpairing" scenario, tension-driven melting occurs at a force above that of the *B* to *S* transition under physiological conditions. The elasticity of the melted state must be that of two parallel single strands each under tension $f/2$, which has an extension of only 1.5 times *B*-length near 65 pN. This scenario also requires that the highly tensed single strands, which are increasingly forced near one another at high forces, not interact.

C. Free energy of *S*-DNA

Above we have considered tension-driven melting of *B*-DNA. We did not yet consider the *S*-DNA state apparent in Figs. 2 and 3. We compute the free energy of *S*-DNA at 500 mM NaCl by considering the transition to begin from the *B*-DNA state at a force $f_0 = 62$ pN, where the *B*-DNA extension is ≈ 0.95 times the *B*-form contour length (about $x_0 = 0.32$ nm/bp). The transition region for 500 mM NaCl is approximated as being linear, from the *B*-DNA curve at force f_0, to the end of the transition at an extension of about 1.705 times the *B*-form length ($x_1 = 0.58$ nm/bp) and force $f_1 = 68$ pN. Then, we approximate the *S*-DNA force response beyond x_1 as linear, but now with a slope $S = (1600/0.34)$pN/nm (i.e., by $f_S = f_1 + S[x - x_1]$). The stiffness *S* indicates the force needed to extend one base pair of *S*-DNA by a given length.

To compute the *S*-DNA free energy at arbitrary force, we integrate along this force curve to obtain

$$w_{\text{S}}(f) = w_{\text{B}}(f_0) + \tfrac{1}{2}[(x_0 + x_1)(f_1 - f_0) + 2x_1(f - f_1) + (f - f_1)^2/S]. \tag{7}$$

Like w_{B}, w_{S} does not include the base-pairing free energy g_0. Equation (7) is free energy per base pair provided that the extensions are in units of length per base pair, and the *S*-DNA stiffness *S* is in units of force per length. The resultant linear *S*-form force response is included in Figs. 2 and 3 (solid line passing through squares above 68 pN) and is close to the experimental data for forces between 68 and 150 pN.

For NaCl concentrations other than 500 mM, we correct the forces f_0 and f_1 by shifting them down by 5 pN for each decade reduction in NaCl concentration in accord with the effect observed in Refs. [10,5]. As salt concentration is reduced, the free energy of the *S*-DNA state thus is reduced. We do not include the increase in double-helix persistence length that occurs with decreasing salt concentration [10] since most of the *S*-DNA free energy is generated by the transition relative to that coming from the initial extension of the double helix.

The free energy of *S*-DNA at 150 mM NaCl is plotted in Fig. 6 relative to the unpeeling free energy, $[w_{\text{S}}(f) + g_0 - w_{\text{ss}}(f)$, solid curve], and relative to the parallel-strand-separation free energy $[w_{\text{S}}(f) + g_0 - 2w_{\text{ss}}(f/2)$, dashed curve], for 150 mM NaCl, and for the sequence-averaged base-pairing energy $g_0 = 2.5k_BT$. For this case, *S*-DNA is favorable relative to unpeeling in the force range of 20–120 pN. *S*-DNA is even more favored over unpaired strands, having a higher free energy for all forces above 10 pN.

Sequence dependence of the free energy difference between *S* and unpeeled DNA can be roughly gauged from the solid curve in Fig. 6. For example, the stability of *S* relative to unpeeling for GC-rich DNA is determined by when the free energy difference shown is greater than $-0.9k_BT$; thus *S*-DNA is favored over unpeeling for GC-rich DNA for the entire force range shown. For AT-rich DNA, unpeeling is favored over *S*-DNA when the free energy shown in Fig. 6 is less than approximately $+1.3k_BT$, showing that unpeeling is favored over the whole force range. Thus, on a real molecule with inhomogeneous sequence, unpeeling of AT-rich mol-

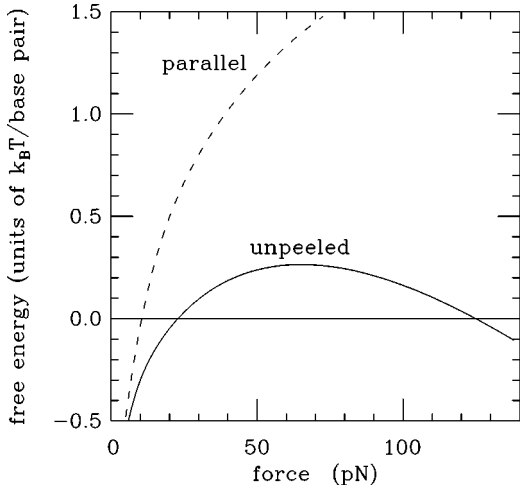

FIG. 6. Free energy per base pair of *S*-DNA, relative to that of unpeeled DNA (solid line), and relative to separated-parallel strands (dashed line) for 150 mM NaCl and base-pairing interaction of $2.5k_BT$. *S*-DNA is stable relative to unpeeled DNA for forces from about 20 to 120 pN. *S*-DNA is also stable relative to base-unpaired DNA for all forces (dashed curve). A shift of the average base-pairing free energy down by $0.3k_BT$/bp is sufficient to make unpeeling favorable relative to *S*-DNA for all forces.

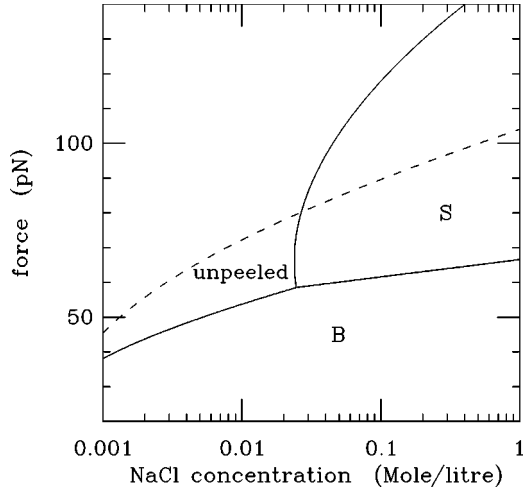

FIG. 7. Salt-force "phase diagram," for sequence-averaged model of DNA (base-pairing interaction $g_0=[2.5+0.2\ \ln(M/0.15)]k_BT$ where M is molar NaCl concentration). For high salt, *B*-DNA transforms first into *S*-DNA, and then unpeels, as force is increased. For low salt, *B*-DNA directly unpeels. Formation of parallel separated strands from *B*-DNA occurs at a higher force (dashed line); parallel separated strands are never favorable relative to unpeeled DNA.

ecule may occur, but GC-rich regions will keep it from spreading to the whole molecule, as long as there are not too many nicks. This will be discussed in more detail in Sec. V.

Similarly, the dashed curve of Fig. 6 shows that parallel-strand separation is not thermodynamically favorable on most of a λ-DNA. Parallel-strand separation is only favorable for the most AT-rich sequences, and even then only for forces below about 60 pN (the point where the dashed curve of Fig. 6 reaches $+1.3k_BT$). Parallel-strand separation can occur at any point in an unnicked double helix; thus, Fig. 6 suggests that a small fraction (the most AT-rich regions) of a λ-DNA might undergo parallel-strand separation in a window of forces near 60 pN. This will be considered in more detail in Sec. IV.

The thermodynamical model developed here allows us to draw a few conclusions. First, for 150 mM NaCl the transition near 65 pN is not force-driven melting, but is a transition to a novel DNA state, i.e., *S*-DNA. However, force-driven unpeeling can be observed on molecules with nicks, to a degree that will depend on the precise sequence and the nick locations. From the free energy model, the likely situation for a small number of nicks on λ-DNA (or other DNAs with similar AT/GC composition) is that small regions of unpeeling may occur near 40 pN, in nicks that happen to be in or very near AT-rich patches. However, GC-rich regions will stop unpeeling from spreading, and the 65 pN transition to the stiff *S*-DNA state will thus occur.

The free energy difference between *S*-DNA and unpeeled DNA is not terribly large. At 150 mM NaCl, *S*-DNA is more favorable by about $0.3k_BT$/bp on average near the 65 pN transition (Fig. 6). Changing the solution conditions so as to favor strand separation (reduction in salt concentration or pH) will increase the degree to which unpeeling will occur. Furthermore, our model predicts that unpeeling becomes more favorable with increasing force (Fig. 6). At high salt we therefore expect a *B*-*S* transition followed by a second S-unpeeling transition at higher force; at low salt we expect only the *B*-unpeeling transition.

By including the salt dependence of *S*-DNA and ssDNA we can compute a force-salt "phase diagram" (see Appendix). To do this we include an ionic strength correction to base-pairing free energy suggested by Santa Lucia *et al.* [14], $\Delta g_0=0.2\ \ln[M/0.150]$, where M is the molar NaCl concentration. Thus higher ionic strength stabilizes the double helix. The resulting phase diagram for the average λ-DNA base-pairing interaction of $2.5\ k_BT$ is shown in Fig. 7. At high salt, we pass from *B* to *S* to unpeeled DNA with increasing force. At low salt, we pass directly from *B* to unpeeled DNA.

The dashed line in Fig. 7 shows the tension-driven strand-separation line (*B*-DNA to two parallel ssDNAs), which is above the unpeeling transition (*B* to one ssDNA under tension, the other relaxed) for all salt concentrations. Separated strands are always less favorable than unpeeled DNA. However, if unpeeling and formation of *S*-DNA are suppressed, tension-driven unpairing might be observable. One way this can be done is by using unnicked molecules and by fixing DNA twist so that the untwisting of the double helix characteristic of the *B* to *S* transition [5,7,19] cannot occur. In fact, this has been done, with the result that a transition to an base-unpaired state called *P*-DNA is observed at forces of about 110 pN. Our model is a starting point for a microscopic model for *P*-DNA. The main additional ingredient required is an estimate of the free energy of the tight right-handed winding characteristic of the unpaired ssDNAs inside *P*-DNA [23].

TABLE I. Base-pairing-stacking free energies of Santa Lucia [14]. Free energies are in k_BT units, and are for 25 C, 150 mM NaCl, pH=7.5. For other salt concentrations the values must be corrected (see text).

Base i and $i+1(5'\rightarrow 3')$	Free energy J_i (150 mM NaCl, pH 7.5, 25 C)
AA	1.68
AT	1.42
AG	2.19
AC	2.42
TA	0.97
TG	2.42
TC	2.12
GG	3.00
GC	3.75
CG	3.68

IV. SEQUENCE-DEPENDENT FREE ENERGY OF TENSION-DRIVEN MELTING

Above we have considered homogeneous sequences, using constant pairing free energy $g_{0,\mathrm{AT}}=1.2k_BT$ for a poly (dA-dT) sequence, $g_{0,\mathrm{GC}}=3.4k_BT$ for a poly (dG-dC) sequence, and the averaged value $g_0=2.5k_BT$ for λ-DNA. In this section we consider sequence effects in detail.

We use a base-pairing free energy based on the nearest-neighbor model of Santa Lucia [14]. Santa Lucia's model can be described in terms of base-pairing variables b_i which we take to be 0 for base-paired *B*-DNA and 1 for unpaired bases. The base-pairing free energy is

$$G_{\mathrm{bp}}=\sum_{i=1}^{N-1}\{J_i b_i b_{i+1}+C[1-\delta_{b_i,b_{i+1}}]\}, \tag{8}$$

where the J_i are the sequence-dependent "stacking" free energies that depend on the base pair at sequence position i and $i+1$ as shown in Table I, and where C is the energy cost associated with the boundaries between B and ssDNA regions. The J_i's are the generalization of the base-pairing free energy g_0, to the sequence-dependent case. For 25 C, pH 7.5 and 150 mM NaCl, *B*-DNA is stable, and the J_i are positive free energy "costs" of breaking base pairs. There is a rather wide variation in the J_i's, from 1.0 to 3.8 k_BT. The boundary energy C is known to be about $3k_BT$, the value we take here.

This model is representative of typical modern theories of the sequence-dependence of base-pairing interactions [34], and describes the free energy contributions from local disruption of hydrogen bonds. The data for such models (Table I) are based on study of the thermal melting of short (usually $<$20 bp) DNA double helices. There is another, longer-ranged, entropic contribution to partially melted DNA under zero tension which is important to purely thermal melting. This is the entropy of internal ssDNA "loops," leading to a free energy cost of bubble creation proportional to the log of the bubble sequence length [36,35]. This is important to thermal melting of large DNAs, since this long-ranged interaction suppresses small bubbles, and leads to cooperative formation of $>$100 bp ssDNA regions [15].

However, for stretched dsDNA this loop entropy contribution is not quantitatively important. We are concerned strictly with force-driven strand separation occuring at forces of at least a few piconewtons. The ssDNA regions will either be free-ended (starting from a nick) and therefore not loops at all, or internal loop regions under tension. In the latter case a few piconewtons of tension will greatly reduce the entropic cost of a loop, as qualitatively discussed by Rouzina and Bloomfield [8]. Furthermore, the dominant contribution to the ssDNA free energy will be the stretching free energy of Sec. III, which is proportional to ssDNA length. Compared to the stretching free energy, the loop entropy is a relatively small correction. Finally, note that the logarithmic term suppresses DNA strand separation, and therefore our model overestimates the likelihood of observation of the parallel-ssDNA state.

Figure 8(a) plots the J_i for the λ-DNA sequence, using the model of Santa Lucia [14] for 150 mM NaCl, pH 7 and 25 C buffer. The average pairing free energy is $\langle J_i\rangle=2.48k_BT$, in accord with the mean value $g_0=2.5k_BT$ taken in Sec. III. However, the local contributions are quite inhomogeneous. The first half of the sequence is richer in GC than the second half, resulting in larger base-pairing free energy density along the first half. For comparison we also show the base-pairing free energy distribution along the 156Gmac plasmid studied by Léger *et al.* [33] [Fig. 8(b)]. This plasmid is about 70% AT, and has a lower average pairing energy of about $\langle J_i\rangle=2.16k_BT$. Again, there are abrupt shifts along the pairing free energy distribution.

Note that the free energies plotted in Fig. 8 have been subjected to a 15 bp-width gaussian smoothing to show the variation in free energy at $>$15 bp scales; without smoothing, one would see just the ten energy levels of Table I.

A. Parallel-strand separation

We now use the Santa Lucia model to study whether inhomogeneous sequence will affect the main conclusion of Sec. III, that the *B* to *S* transition preempts formation of an appreciable amount of tension-melted DNA, in the case where unpeeling cannot occur (e.g., for un-nicked DNA). We treat this using an equilibrium model, since the stretched strands can be expected to be able to fluctuate between paired and unpaired states. Unpeeling, which can be expected to show a more irreversible character, is described in the next subsection using a nonequilibrium model.

We extend the Santa Lucia model by supposing that each base pair i can take one of three states, *B*-DNA ($b_i=0$), ssDNAs ($b_i=1$) and *S*-DNA ($b_i=2$). We treat base-pairing interactions using the model described above, and the mechanics of the three DNA conformations using the free energies w of Sec. III. Thus, the free energy describing a given base-pairing state $\{b_i\}$, under tension f is

$$G=\sum_{i=1}^{N}\{C[1-\delta_{b_i,b_{i+1}}]+J_i\delta_{b_i,1}\delta_{b_{i+1},1}-w_{\mathrm{B}}(f)\delta_{b_i,0} -2w_{\mathrm{ss}}(f)\delta_{b_i,1}-w_{\mathrm{S}}(f)\delta_{b_i,2}\}. \tag{9}$$

The first terms describe the cooperativity and sequence de-

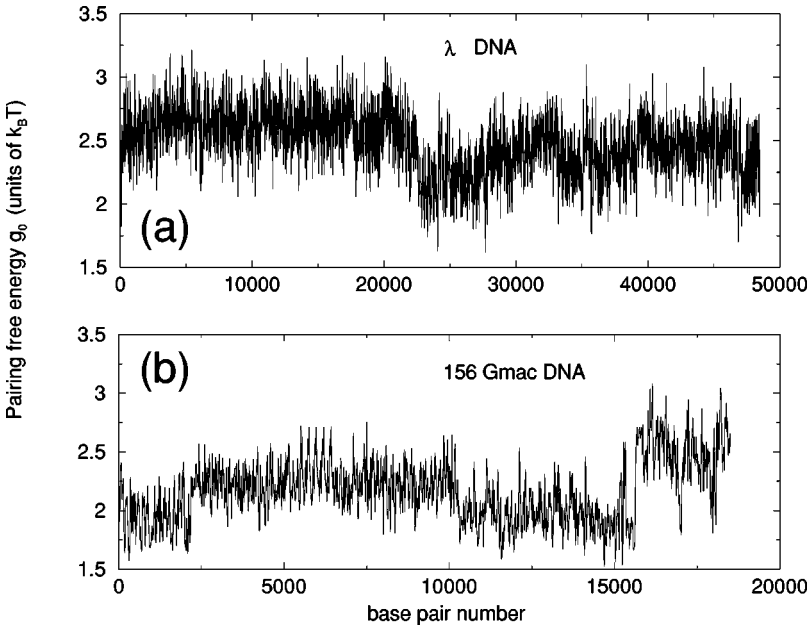

FIG. 8. Base-pairing free energy densities at zero force for (a) λ-DNA (top) and (b) the AT-rich DNA 156Gmac, calculated using the model of Santa Lucia [14] for 150 mM NaCl, *p*H 7.5 and 25 °C (see Table I). Curves shown here are smoothed using a 15 bp width Gaussian to show large-scale variation in free energy density, see text.

pendence of base-unpairing, while the last three terms set the relative free energies of the three conformational states. Note that C also is the energy for *B*-*S* and ss-*S* boundaries. In principle these are independent parameters, but for simplicity we take them all to be the same; estimates of the *B*-*S* boundary energy are in fact close to $C=3k_BT$ [1].

The partition function associated with the $\{b_i\}$ fluctuations uses the Boltzmann factor e^{-G} (we take k_BT as our energy unit), and can be written as

$$Z=\sum_{b_1=0,1,2}\cdots\sum_{b_N=0,1,2}\prod_{i=1}^{N-1}\exp[w_{\mathrm{B}}\delta_{b_i,0}+2w_{\mathrm{ss}}\delta_{b_i,1}+w_{\mathrm{S}}\delta_{b_i,0}-C(1-\delta_{b_i,b_{i+1}})-J_i\delta_{b_i,1}\delta_{b_{i+1},1}]. \quad (10)$$

This can be considered to be a product of N 3×3 matrices, reflecting the one-dimensional, nearest-neighbor-coupled structure of Eq. (9). Salt dependence is taken into account via the salt dependences of the J_i and the w's.

We compute Z numerically for λ-DNA by explicitly carrying out the matrix product of Eq. (10), using the known sequence. The average value of the molecule extension is $k_BT\partial\ln Z/\partial f$, and the averaged number of melted sites is $\partial\ln Z/\partial(2w_{\mathrm{ss}})$ [19].

Figure 9(a) shows DNA force versus extension for λ-DNA. The solid curve is the result for 150 mM NaCl; the dashed curve is 10 mM NaCl; and the dotted curve is 2.5 mM NaCl. As salt concentration is reduced, the transition plateau is reduced, in accord with the results of the previous section. Figure 9(b) shows the fraction of the bases which are strand separated (i.e., neither *B* nor *S* form); for each case, a peak occurs in the ssDNA fraction at roughly the midpoint of the overstretching transition. The amount of strand separation increases with decreasing salt concentration, but even at 2.5 mM at most only about 10% of the molecule strand separates.

This calculation verifies that parallel-strand-separated DNA is not the favorable state at any force, for physiological ionic conditions (150 mM NaCl), in accord with the sequence-averaged model of Sec. III, but now generalizing it to inhomogeneous sequence. If unpeeling is suppressed, e.g., by having a low density of nicks as occurs in vivo, then at forces of 40–50 pN, only the most unstable, AT-rich regions of the molecule start to strand-separate. Then, when *S* becomes favorable relative to *B*, the molecule is transformed to *S* form, including the regions which were starting to strand separate. This is the origin of the well-defined peak in the ssDNA fraction [Fig. 9(b)].

An interesting feature of the ssDNA density is that at zero force and 150 mM NaCl, about 0.5% of the molecule is single stranded [Fig. 9(b)]; this fraction actually is *reduced* by forces of 10–20 pN. This effect is due to *B*-DNA being more easily extended than parallel ssDNAs in this force range, as has been noted by Rouzina and Bloomfield [8]. Increasing force above 20 pN drives increasing single-strand bubble formation up to 50 pN. These "40 pN bubbles" are highly AT-rich regions.

The results of this section are a prediction for experiments on a dsDNA with no nicks and covalent closure of its ends (some viruses carry DNAs which naturally have this structure, e.g., vaccinia [37]). To have zero torque in the molecule during stretching, single-bond connections could be engineered, e.g., to single biotins located at the ends. Alternately, one could test our predictions using molecules with ssDNA end anchors but where a small fraction of the base pairs are chemically crosslinked together, or with terminal GC-rich regions, both of which would block unpeeling from the ends while allowing free rotation of the molecule. Including the effect of the entropy cost of ssDNA "bubbles" will suppress the formation of large unpaired regions, and therefore the present theory at worst slightly overestimates the stability of tension-driven unpairing.

B. Unpeeling free energy at constant force

Recall that for an average base pair free energy $g_0=2.5k_BT$ the unpeeled DNA is thermodynamically stable

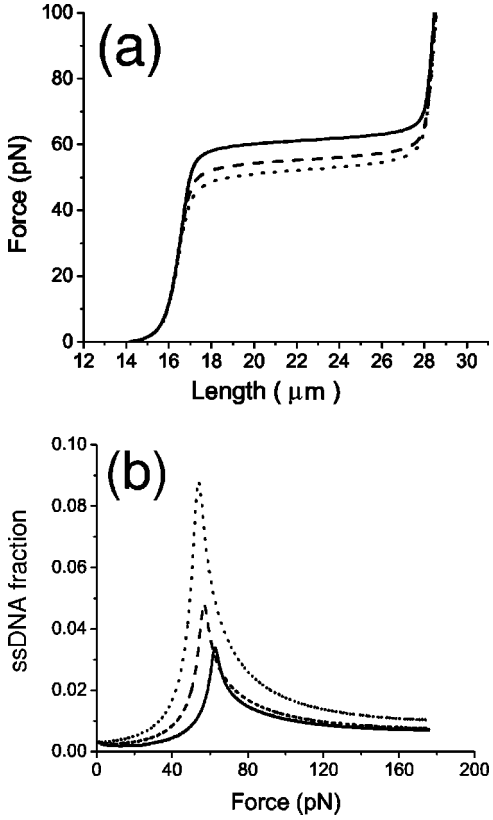

FIG. 9. Inhomogenous-sequence model of equilibrium distribution of *B*, *S*, and parallel-strand-separated states along an unnicked 48.5 kb λ-DNA. (a) Force-extension curve; solid curve shows result for 150 mM NaCl, where the transition is predominantly *B* to *S*. Dashed curve shows result for 10 mM NaCl; plateau is shifted down mainly by the salt-dependence of the *S*-state free energy. Dotted curve shows result for 2.5 mM NaCl; the further shift of the force plateau down to ≈50 pN is partly due to parallel-strand separation. (b) Fraction of bases which are in strand-separated conformation, vs force, with different curves for different salt conditions as in (a). The only case where appreciable strand separation occurs is low salt (dotted, 2.5 mM NaCl). Note that at zero force, the molecule is about ≈0.5% ssDNA.

against *S*-DNA [i.e., $w_{ss}(f)-w_S(f)-g_0>0$] for forces larger than 110 pN, for 150 mM NaCl. Thus for a hypothetical molecule with homogeneous sequence and uniform base pairing energy $g_0=2.5k_BT$, we predict a *B* to *S* transition at 65 pN, and then an *S* to unpeeled transition (assuming the presence of a nick) at 110 pN. However, as shown above, the heterogeneity of genomic sequences (e.g., λ-DNA) leads to a rough base-pairing free energy landscape.

The consequences of this are first, that for a heterogeneous sequence, we can expect unpeeling to depend on the position(s) of nick(s) along the molecule. Recall that for the 3′-strand end anchors routinely used for λ-DNA [38], there are always two locations at the molecule ends at which unpeeling can start. Unpeeling from a particular nick must follow a pathway where a series of high barriers associated with GC-rich regions much be crossed. This means that barrier-crossing dynamics will dominate unpeeling, much as occurs in unzipping of DNA at constant force [24,25,27].

Unpeeling of a heterogeneous sequence should thus be described using a nonequilibrium model of the motion of the dsDNA-ssDNA "fork." This situation is closely related to the dynamics of unzipping of a dsDNA which is done by pulling the two ssDNAs apart by sequence dependent 10–20 pN forces [13,11]. Our model for unpeeling dynamics is closely related to the sequence-dependent dynamical theory of DNA unzipping [17].

We begin by describing the free energy of unpeeling, at first for the simplest case where the only nick on a molecule is at one end (sequence position $i=0$), so that unpeeling can proceed in only one direction (this can be achieved using a single-strand anchor at one end, and a double-strand anchor at the other end, see Ref. [10]). Since end attachments are usually made to the 3′ end, we suppose the strand to unpeel in the $5'\rightarrow 3'$ direction. Unpeeling is thus described in terms of the sequence position of the dsDNA-ssDNA fork. We consider unpeeling of a dsDNA that, depending on the force, is initially in either *B* or *S* form.

For a tethered λ-DNA we imagine unpeeling *n* bases from sequence position 1 (conventionally called the "left end"). Given that usually the 3′ strand ends are tethered, this corresponds to unpeeling the 5′ strand end, in the $5'\rightarrow 3'$ direction. Using Eq. (8), the base-pairing free energy cost is

$$C+\sum_{i=1}^{n-1} J_i. \tag{11}$$

Since the J_i of Santa Lucia are the sequence-dependent analog of the averaged g_0 used in Sec. III, the unpairing free energy for opening the first *n* bases is $\Sigma_{i=1}^{n}\, g_0(0,n)$, where

$$g_0(0,n)=\begin{cases} C & n=1 \\ J_{n-1} & n\geqslant 2 \end{cases} \tag{12}$$

is the incremental free energy cost of unpeeling the *n*th base from the end.

If a nick is made midmolecule, between base pairs n_0 and n_0+1, then unpeeling can proceed in either the $5'\rightarrow 3'$ "forward" direction, or in the opposite $3'\rightarrow 5'$ "reverse" direction. The generalization of Eq. (12) is straightforward:

$$g_0(n_0,n)=\begin{cases} C & n=\pm 1 \\ J_{n_0+n-1} & n\geqslant 2 \\ J_{n_0+n+1} & n\leqslant -2 \end{cases} \tag{13}$$

for unpeeling $|n|$ bases, with the sign of *n* indicating the two directions of strand unpeeling.

Given the sequence dependent $g_0(n_0,i)$, we can write the free energy of a molecule with its first *n* bases after the end-nick unpeeled, in the $5'\rightarrow 3'$ direction:

$$G_f(n)=[w_{ds}(f)-w_{ss}(f)]n+\sum_{i=1}^{n} g_0(0,i), \tag{14}$$

where $w_{ds}(f)$ is the work done by the force during the extension of the double-stranded state, up to the force *f* [the rightmost integral term of Eq. (3)]. The force-extension for dsDNA includes the B elasticity (1) up to the plateau force, then the *B*-*S* transition and *S* elasticity as discussed in Sec.

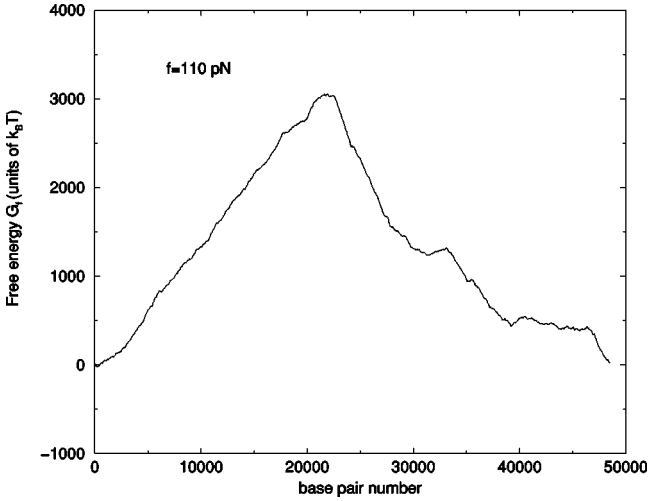

FIG. 10. Free energy of unpeeling from a nick at the end ($n_0=0$) of a λ-DNA, under constant force of 110 pN (150 mM NaCl, *p*H 7.5, 25 C). The two minima for *S*-DNA ($n=0$) and for completely unpeeled ssDNA ($n=48\,502$) are in thermodynamical equilibrium. However, crossing the huge barrier of $3000k_BT$ between the two minima is kinetically forbidden. The free energy has been shifted to zero at $n=0$.

III C. Note that Eq. (14) is zero when $n=0$; i.e., it is free energy relative to the double-stranded form.

Figure 10 shows the free energy for unpeeling the *S*-form of λ-DNA at a force of 110 pN, starting from a nick at the beginning of the sequence, calculated using the Santa Lucia model [14] for 150 mM NaCl, pH 7, 25 C. At this force, the states with $n=0$ and $n=48\,502$ (completely double-stranded and completely unpeeled states, respectively) have the same free energy, and in an equilibrium theory would have equal probability. However, the inhomogeneous sequence generates barriers on many sequence scales. At the largest scale

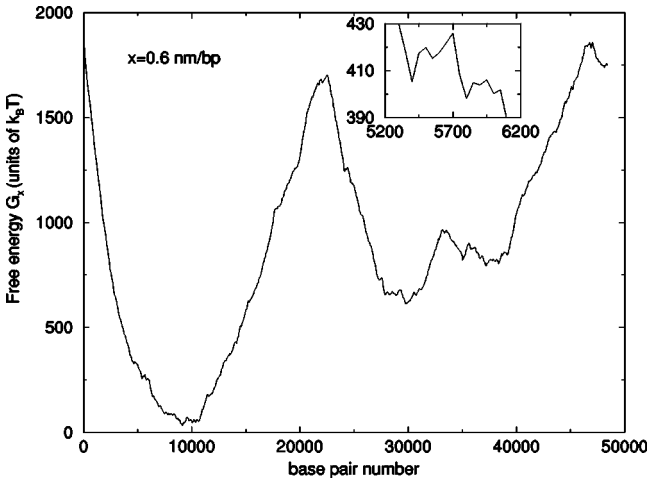

FIG. 11. Free energy of a λ-DNA as a function of the number of base pairs unpeeled from a nick at the end ($n_0=0$), for fixed extension per base pair $x=0.6$ nm/bp (150 mM NaCl, *p*H 7.5, 25 C). The minimum (equilibrium) is at $n\approx 10\,000$; local minima separated by barriers cause "stick-slip" motion when extension x is increased at fixed rate.

there is a giant barrier of about $3000k_BT$. Thus, while an inhomogeneous-sequence equilibrium theory would predict unpeeling from the end at around 110 pN, kinetically unpeeling will still be blocked at this force. Quantitative description of unpeeling of inhomogeneous DNA requires a kinetic theory.

For the general case of an nick at sequence position n_0 far from either end, the ssDNAs either to the left or the right of the nick may unpeel (in opposite $5'\rightarrow 3'$ and $3'\rightarrow 5'$ directions along that strand). Then,

$$G_f(n_+,n_-)=\sum_{i=1}^{n_+} g_0(n_0,i)+\sum_{i=1}^{n_-} g_0(n_0,-i)-nw_{\rm ss}(f) - (N-n)w_{\rm ds}(f), \tag{15}$$

where $n=n_-+n+$, and where the index i on the nicked strand increases $5'\rightarrow 3'$.

C. Unpeeling free energy at fixed extension

Experimentally the *B*-*S*-ssDNA transition has most often been studied by stretching the DNA at a constant rate, using manipulators that control the positions of the ends (e.g., laser traps, cantilevers). In this case the free energy landscape is complex since the amount it is "tilted" will jump up and down as the molecule responds via a series of barrier-crossing events.

The free energy of an *N*-base pair DNA held at extension Nx, with a nick at n_0 and $n=n^+ + n^-$ unpeeled base pairs is

$$G_x(n)=N\,x\,f-nw_{\rm ss}(f)-(N-n)w_{\rm ds}(f) - \sum_{i=1}^{n_+} g_0(n_0,+i)-\sum_{i=1}^{n_-} g_0(n_0,-i), \tag{16}$$

where $n=n_+ + n_-$ and, where $f=f(x,n)$ is the value of the force for which the unpeeled ssDNA and the *S*-DNA are in equilibrium, i.e., the solution of the equation:

$$x=\frac{n}{N}x_{\rm ss}(f)+\frac{N-n}{N}x_{\rm ds}(f). \tag{17}$$

In Fig. 11 we show the free energy landscape obtained for a nick at the origin and extension $x=0.6$ nm/bp (the free energy has been shifted by a constant). The minimum is at $i\approx 10\,000$ unpeeled base pairs; although the pure *S*-DNA and unpeeled states have equal free energy, there is a 1500 k_BT barrier separating these states (plus many other much smaller but still insurmountable barriers). The inset shows local minima separated by free energy barriers. So, when increasing the extension at a constant rate unpeeling will take place by abrupt jumps between minima. This is the unpeeling analog of the "stick-slip" motion seen during unzipping of DNA at fixed extension [13].

V. KINETICS OF STRAND UNPEELING AT FIXED EXTENSION

The sequence-generated barriers (Figs. 10 and 11), plus the necessity of following a sequential pathway over them,

make unpeeling an intrinsically slow process. Laser trapping experiments [10,12] typically use 1–2 min for stretch-release cycles. For λ-DNA (16.3 μm *B*-form length) this corresponds to stretching rates of 300–600 nm/s. In atomic-force-microscope work of Rief *et al.* [6,4] rates of 0.15 to 3 μm/s were studied. Existing experiments show the following evidence for our favored scenario, that a well-defined *B* to *S* transition occurs near 65 pN for physiological ionic strength, and that at higher forces *S*-DNA undergoes unpeeling via barrier-crossing kinetics:

(1) The ≈60 pN *B* to *S* plateau for λ DNA at 150 mM NaCl shows little hysteresis and little change with pulling rate; however, at higher extension and force ($x \approx 0.6$ nm/bp, $f > 100$ pN) the stretching curves show a strong pulling-rate dependence [6,4]

(2) After reaching high forces $f > 100$ pN, relaxation shows force hysteresis that increases at larger rate [4,6,7]

(3) At low ionic strength, and for the AT-rich 156Gmac at 35 C, force hysteresis is observed after stretching to forces <100 pN [7,2].

We now present a simple model for the kinetics of unpeeling. We consider experiments at constant pulling rate v, either stretching ($v > 0$) or relaxation ($v < 0$). For a dsDNA with a nick at sequence position n_0, unpeeling of the nicked strand is described using dissipative dynamics driven by the free energy at fixed extension (16):

$$\frac{1}{r}\frac{dn}{dt} = -\frac{dG_x(n)}{dn} \tag{18}$$

We use the force-balance relation (17), and eliminate time ($dt = dx/v$) to find fork position equations

$$\frac{dn_+}{dx} = \frac{r}{v}\{w_{ss}[f(x,n)] - w_{ds}[f(x,n)] - g_0[n_0, +n_+(x)]\}$$

$$\frac{dn_-}{dx} = \frac{r}{v}\{w_{ss}[f(x,n)] - w_{ds}[f(x,n)] - g_0[n_0, -n_-(x)]\}. \tag{19}$$

Here $f(x,n)$ is obtained by solving the equilibrium condition (17). To describe increasing extension, we integrate (19) with initial condition $n_+(x) = 0$, $n_-(x) = 0$, starting from the value of extension at which $w_{ss} + g_0 = w_{ds}$, and ending at an extension of $x = 0.74$ nm/bp. We use a microscopic base-opening rate of $r = 10^8 s^{-1}$, and we vary v between 1 nm/sec to 1500 nm/sec. For a relaxation run we use a negative rate v in Eq. (19), starting from a fully unpeeled state ($n = N$) at extension $x = 0.74$ nm/bp.

Finally, in order to make integration of Eq. (19) numerically tractable, we smooth the base-pairing potential $g_0(n_0, n)$ using the Gaussian kernel $\exp(-n^2/[2\sigma^2])/(2\pi\delta^2)^{1/2}$. This preserves the barrier structure of the fixed-extension free energy landscape (see Fig. 11) at scales larger than δ, while making Eqs. (19) smooth enough at scales smaller than δ to be easily integrated. We have used a smoothing scale of $\delta = 15$ bp, which eliminates barriers which are less than 5 $k_B T$ high. The dynamics associated with these small barriers will be fast and thermally driven [18]. Our model (19) plus the Gaussian-smoothed g_0 provides a coarse-grained description suitable for studying the relatively slow and driven unpeeling of the molecule as a whole.

A. Unpeeling of λ DNA at 150 mM NaCl

We numerically integrate (19) for the Gaussian-smoothed sequence-dependent pairing free energies for λ DNA in buffer with pH 7, 25 C, and 150 mM NaCl. The nick position n_0 has been varied. Fig. 12(a) shows force extension curves and the number of unpeeled base pairs, for different pulling rate, for λ-DNA with a nick at one end ($n_0 = 0$). At low pulling rates, unpeeling starts when the tension reaches about 180 pN; for further extension, unpeeling progresses [Fig. 12(a), inset]. As pulling rate is increased, unpeeling progresses slowly. These curves compare well with experimental results (see Fig. 4 of Clausen-Schaumann *et al.* [6]), which show a pulling-rate-independent *B* to *S* transition and then a higher-force, less-well-defined, and pulling-rate-dependent "melting" transition.

The unpeeling transition depends on the number and position of the nicks. In the one nick case, if the nick is located in the middle of the sequence the unpeeling is more favorable relative to a nick at one end, since there are two melting directions. Figure 12(b) shows force-extension curves for a few different nick positions, for $v = 300$ nm/s; unpeeling always begins near 150 pN, well separated from the *B* to *S* transition. The sequence that unpeels more rapidly is the one with the nick at position $n_0 = 23\,000$, corresponding roughly to the top of the free energy barrier of Fig. 10.

Figure 13 shows a stretch-relax cycle for a λ-DNA with a nick at $n_0 = 23\,000$, for different pulling rates. Stretching follows the *B*-*S* force-extension curves up to a force of about 150 pN at which unpeeling begins. The force during unpeeling is pulling-rate dependent behavior, as in experiments of Clausen-Schaumann *et al.* (see Fig. 4 of Ref. [6]). During retraction from an unpeeled configuration, the force comes down initially along the ssDNA force-extension curve, showing hysteresis. The molecule starts to reanneal when the force approaches the 60 pN plateau. Reannealing is less complete at higher relaxing rates; at the lowest relaxation rate (100 nm/s) sequence dependent forces are observed.

B. Unpeeling of 156Gmac at 150 mM NaCl

The 156Gmac DNA is a 70% AT-rich DNA with a lower average pairing free energy than λ-DNA (see Fig. 8). Unpeeling-reannealing dynamics for extension followed by retraction at 300 nm/s are shown in Fig. 14(a), for 150 mM NaCl and 35 C. We have used the temperature correction given by Santa Lucia [14]; note that the theory has no adjustable parameters. For this AT-rich molecule and elevated temperature, unpeeling is highly favored in its AT-rich regions, and so it provides an interesting test case for our theory.

This case is also attractive because it has been experimentally studied. Figure 14(b) shows an extension-retraction experiment on a nicked 156Gmac at 35 C by Léger [7], at a rate of about 200 nm/s. Three extension-retraction cycles

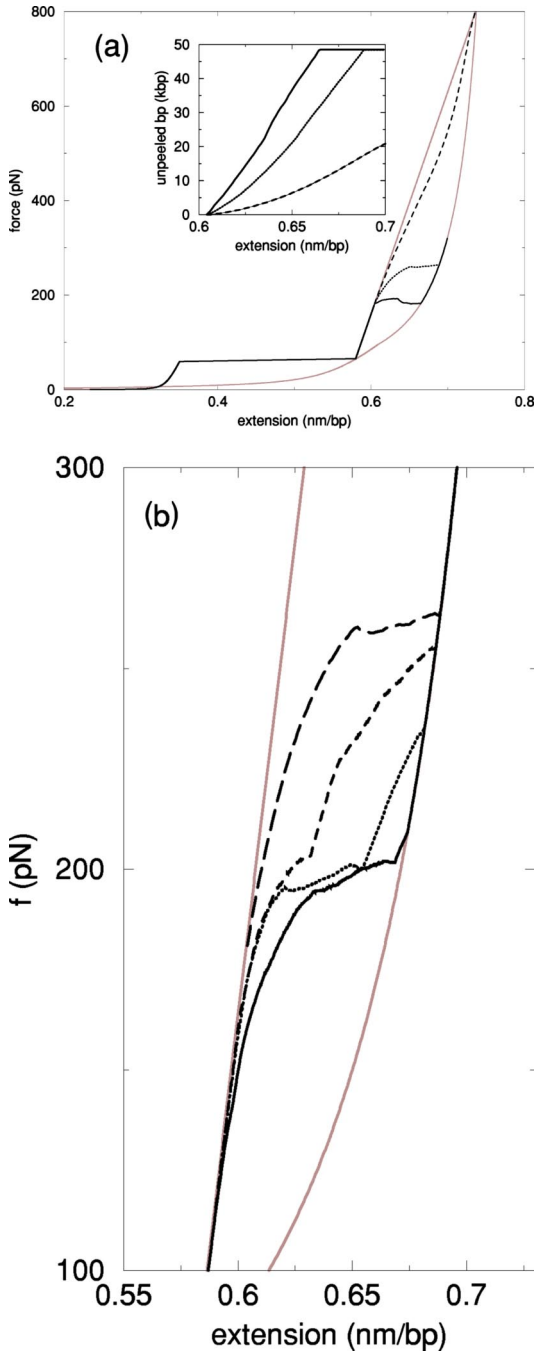

FIG. 12. (a) Stretching λ-DNA with constant extension rate; base-pairing interactions and ssDNA elasticity are for 150 mM NaCl, *p*H 7.5, 25 C. Unpeeling begins from *S*-DNA, from a nick at the beginning of the sequence ($n_0=0$); black curves show extension rates $v=100$ nm/s (solid), 300 nm/s (dotted), 1500 nm/s (dashed). Force is shown vs extension per base pair; for slow extension, unpeeling starts at around 170 pN; rapid extension causes the tension in the molecule during unpeeling to increase. Gray curves show equilibrium *B*, *S*, and ssDNA force response. Inset shows corresponding number of unpeeled base pairs vs relative extension. (b) Calculated unpeeling of λ-DNA extension rate $v=300$ nm/s, for λ DNA, showing results for different nick positions. Force vs the extension relative to *B*-DNA extension for $n_0=23\,000$ (solid), 30 000 (dotted), 10 000 (dashed), 0 (long dashed). Gray curves show equilibrium *B*-DNA, *S*-DNA, and ssDNA force responses.

were done, to three successively higher forces (85 pN, 140 pN, 180 pN). During the first cycle (solid), a *B* to *S* transition is almost reversibly traversed. On the second cycle (dotted) the *B* to *S* transition is followed, but then during return the force falls below the transition, and instead follows a jagged return curve. Finally on the third cycle (dashed line) a transition to a ssDNA response is observed, with a smooth ssDNA-like response curve during return.

Figure 14 shows results for our theory, for a series of three extension-retraction cycles to successively higher peak forces (100 pN, 110 pN, 180 pN, indicated by stars in the figure). On the first extension (solid), unpeeling starts from the nick (located on a weakly bound region) at about 50 pN, before the *B*-*S* transition. However, unpeeling then stops at a more tightly bound region, an example of blockage of unpeeling by GC-rich "insulating" regions. During the first return, both unpeeled strands reanneal at least partially, resulting in a small amount of hysteresis. However, on the second cycle, only one strand reanneals, causing a larger amount of hysteresis. Finally, the third cycle to 180 pN results in full unpeeling and an ssDNA-like return. Thus, our theory is able to describe this rather complex situation where one is gradually driving more of the molecule into the unpeeled state. This example also emphasizes the difference between the reversible *B* to *S* transition which occurs over a rather narrow force range, and the irreversible unpeeling transition which occurs over a broad force window.

C. Low-salt unpeeling kinetics

The above results were for the physiological ionic strength of 150 mM NaCl. Significantly lower ionic strengths reduce the base-pairing free energy, but also the *S* free energy. Based on the sequence-averaged analysis of Sec. III, we expect unpeeling to dominate for ionic strengths below 25 mM. In our kinetic model we account for NaCl concentrations below 150 mM using the experimental results for *S* and ssDNA at lower ionic strength (see Sec. II C), plus the ionic-strength correction for the base-pairing free energy given by Ref. [14]) [$\Delta g_0=0.2\ \ln(M/0.150)$, where M is the molarity of NaCl].

Figure 15(a) shows theoretical results for 10 mM NaCl and 25 C, for a nick in an AT-rich region of λ DNA ($n_0=23\,000$). Two separate extension-retraction cycles are shown, the first to 80 pN (dashed), the second to 150 pN (solid). During extension, partial unpeeling occurs at ≈45 pN, at an extension of about 0.35 bp/nm, but then GC regions stop unpeeling, allowing the remainder of the molecule to undergo a *B* to *S* transition at 60 pN. During the return from 80 pN, a small amount of hysteresis is observed at the low-extension end of the plateau, rather like the 35 C 156Gmac result of the preceding section. However, if extension is made out to 0.70 nm/bp and 150 pN, the molecule fully unpeels, and an ssDNA-like response is observed during retraction.

Figure 15(b) shows an experimental extension-retraction curve for λ-DNA in 10 mM NaCl buffer at 25 C, of Leger [7]. In this run, a peak force of 75 pN was used. A small amount of hysteresis is observed, similar to that observed theoretically [Fig. 15(a)].

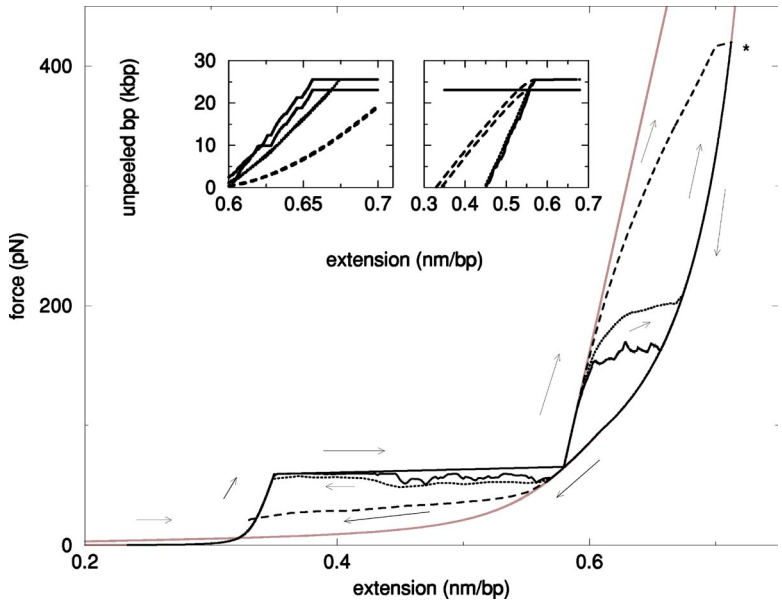

FIG. 13. Force during extension-retraction cycle for λ-DNA, with unpeeling occurring from a nick in the middle of the molecule (n_0 = 23 000) for different rates: v = 100 nm/s (solid), 300 nm/s (dotted), 1500 nm/s (dashed). Gray curve shows equilibrium *B* to *S* and ssDNA force responses. For all rates, extension follows the equilibrium B-S force response up to the force of around 150 pN at which unpeeling starts to occur from the nick. The unpeeling force-extension curves depends on the pulling rate; when unpeeling is complete it reaches the ssDNA curves. During relaxation starting from the position indicated by the star, the reannealing curves follows the ssDNA curve down to roughly the *B*-*S* plateau force; then the two strands of DNA begin to hybridize, and the force-extension curve becomes dependent on the retraction rate. The number of unpeeled base pairs vs relative extension is shown in the left inset during stretching and in the right inset during relaxation.

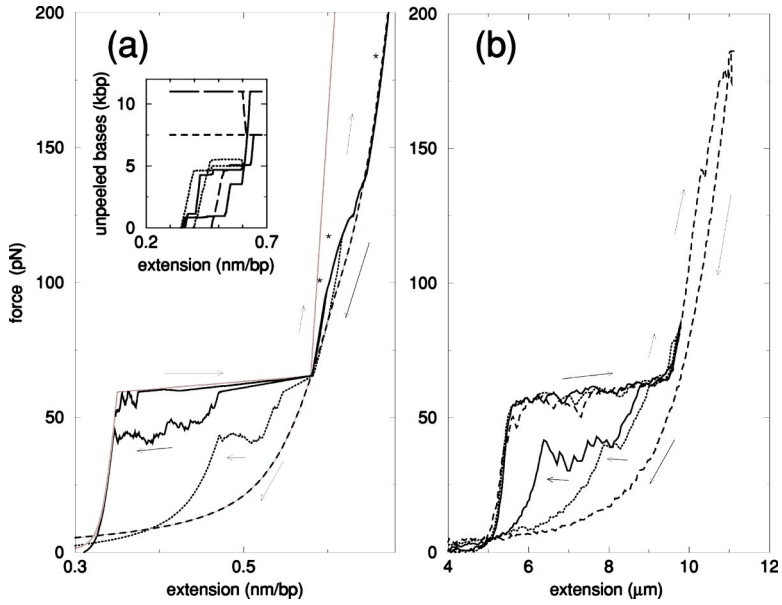

FIG. 14. (a) Theoretical force vs extension during stretch-relax cycle at v = 300 nm/s, for AT-rich DNA 156Gmac, with a nick at n_0 = 11 000, at 35 C. During stretching (solid, arrows leading to higher force and extension) unpeeling starts at 50 pN, before the *B*-*S* transition. Unpeeling proceeds during and after the *B*-*S* transition; the molecule becomes completely unpeeled at 150 pN. Retraction results are shown for three cases where extension is stopped and retraction started (stars) at 100 pN (solid), 110 pN (dotted) and 180 pN (dashed). Hysteresis during retraction is associated with the amount of unpeeling during extension. For the 110 pN case only one of the two unpeeled ssDNAs recombines during retraction; for the 180 pN case the unpeeled DNA remains as ssDNA during retraction. (b) Experimental data [7] on 156Gmac for 35 C, at a rate of approximately v = 200 nm/s. The molecule was subjected to three stretching-relaxation cycles to progressively higher peak forces; forces at which retraction was begun were 85 pN (solid), 140 pN (dotted), and 180 pN (dashed).

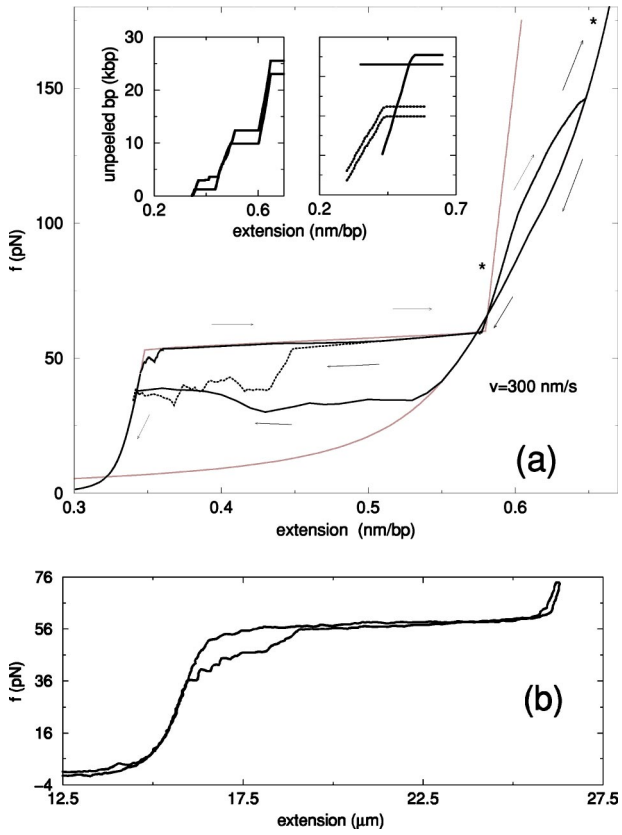

FIG. 15. (a) Theoretical stretching and reannealing dynamics for λ-DNA, for 10 mM NaCl, nick position $n_0=23\,000$, and extension/retraction velocity $v=300$ nm/sec Stars indicate points at which retraction begins. The equilibrium *B*-*S* plateau (gray) is shifted down because of low salt, to ≈55 pN. *B*-DNA starts to unpeel at 40 pN, and is completely unpeeled at 140 pN (solid). Results are shown for retractions beginning at 80 pN (dotted), and 160 pN (solid); for the higher peak force more hysteresis occurs during retraction. Left inset: number of unpeeled base pairs during stretching. Right inset: number of unpeeled bases during relaxation. (b) Experimental data of Leger [7] for λ DNA stretched and relaxed in phosphate buffer with 25 mM NaCl, at 25 C. The hysteresis is similar to that found in the calculation, in the case where relaxation is started at 80 pN.

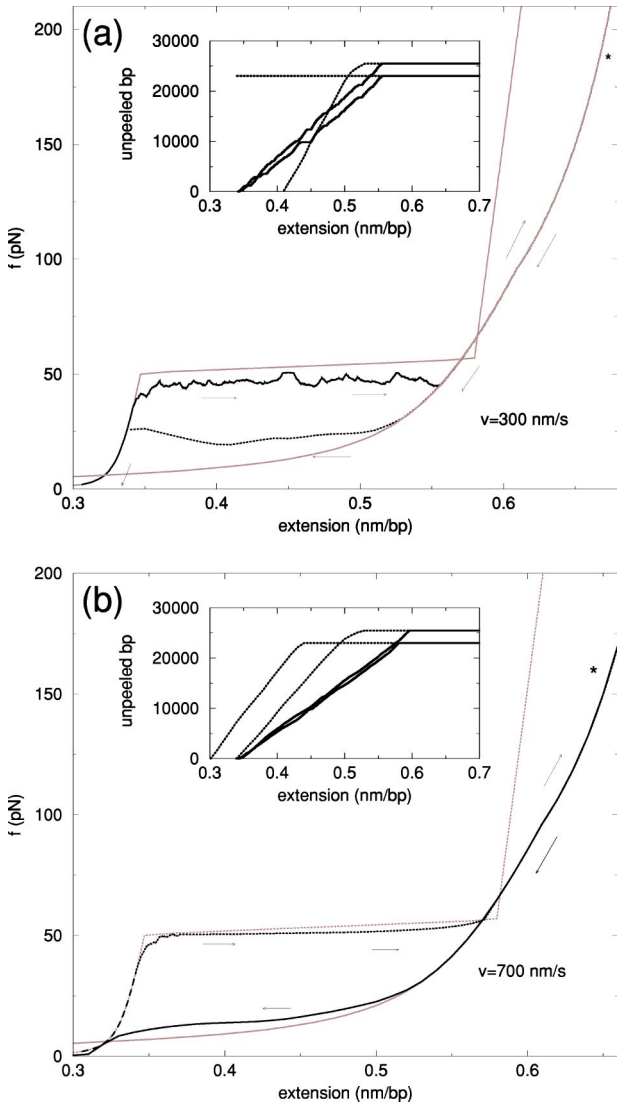

FIG. 16. Theoretical stretching and reannealing dynamics for λ-DNA at 2.5 mM NaCl and 25 C, nick at $n_0=23\,000$. (a) For extension/retraction rate of 300 nm/sec *B*-DNA unpeels at about 40 pN. Retraction leads to large hysteresis. (b) Extension/retraction rate of 700 nm/sec. At this rate unpeeling transition is smoother than the top curve. Reannealing follows the ssDNA force-extension curve.

At an even lower ionic concentration of 2.5 mM NaCl and an extension rate of 300 nm/s the unpeeling transition completely preempts the *B*-*S* transition [Fig. 16(a)]. A rough sequence-dependent unpeeling occurs at forces lower than the *B*-*S* plateau. During retraction hysteresis shows up. For a larger rate of 700 nm/sec [Fig. 16(b)] unpeeling does not occur completely before the *B*-*S* transition and the force-extension curves are smoother.

We do not show results for ionic strengths above 150 mM NaCl. The stabilization of base-paired DNA makes unpeeling progressively unfavorable at higher salt. A sharp *B* to *S* transition followed by a higher-force unpeeling transition occurs in our kinetic calculation in this regime.

VI. CONCLUSION

The main conclusion of this paper is that *S*-DNA as observed by the groups of Smith *et al.* [2], Cluzel *et al.* [1] and others under physiological solution conditions (pH 7.5 buffered aqueous solution with 150 mM Na^+), or at higher NaCl concentrations, cannot be explained in terms of separated DNA strands. The mechanical properties of *S*-DNA are distinct from those of either one (unpeeled) or two parallel single strands. Most simply, *S*-DNA has a larger spring constant than either ssDNA, two parallel ssDNAs, or *B*-DNA at about the same elongation (Sec. II). This is perhaps the most clear experimental signature supporting a picture of the 65 pN 150 mM NaCl transition as being to a dsDNA state with a reasonably robust secondary structure.

To support this conclusion we have analyzed the free energy of tension-driven melting [9,8], first using DNA unzipping data [4,6,13] to fix the relative free energies of ssDNA and *B*-DNA (Sec. III), and then using established base-

pairing-free-energies [14]. to account for sequence effects. We then have compared these strand-separated state free energies to that of *S*-DNA, also computed from experimental data.

For 150 mM NaCl, over the force range of 40 (AT-rich DNA unpeeling from B) to >150 pN (GC-rich DNA unpeeling from *S*) it becomes favorable for one strand of a *B*-DNA to 'unpeel' so that the remaining ssDNA carries all the tension. We emphasize that unpeeling depends on breaks, or "nicks" in the DNA backbone, which allow the unpeeled strand to reach a thermodynamically favorable highly extended state.

We also have analyzed the thermodynamics of a second scenario for tension-driven melting, namely, the generation of parallel, noninteracting strands, where the two strands each carry half the total tension [10,8]. Neglecting the possibility of unpeeling, parallel-strand separation becomes favorable to *B*-DNA for tensions of 60 (AT-rich) to 120 pN (GC-rich), at 150 mM NaCl. This force range is higher than that for unpeeling, reflecting the larger free energy cost to create the parallel-strand-separated state. The origin of the lower free energy cost of unpeeling is simply that the released strand can relax into a random coil state.

Experimental data allows us to estimate the free energy of *S*-DNA and to compare it to these two strand-separated states. At 150 mM NaCl, *S*-DNA becomes thermodynamically favorable relative to both unpeeled and unpaired states for all but the most weakly bound (typically AT-rich) sequences, near 65 pN.

Unpeeling does become more favorable than *S*-DNA for larger forces: thermodynamically it becomes favorable at 120 pN, but barriers generated by GC-rich regions will keep it from propagating on λ-DNA until about 150–200 pN. Our model gives a clear interpretation for the two transitions observed in some experiments [12,6]. The first transition at 65 pN is *B* to *S*-DNA; on molecules with only a few nicks this is a sharp and reversible transition [7]. The second transition at higher forces (≈150 pN for small numbers of nicks) is unpeeling of DNA.

In DNAs of natural origin which have inhomogeneous sequence we have shown that unpeeling will occur inhomogeneously. With increasing force, unpeeling will occur first from nicks that happen to be in weakly bound (AT-rich) regions of a molecule; unpeeling will then spread to strongly bound regions (GC-rich) only at higher forces >150 pN. GC-rich regions thus act to "insulate" against unpeeling. The details of how unpeeling will spread on an inhomogeneous molecule will depend on sequence, and will show a series of plateaus at constant force, or a series of spikes at constant extension rate corresponding to pushing the unpeeling fork through a series of GC-rich regions. The latter unpeeling transition will vary in character from molecule to molecule if the nicks occur randomly. Unpeeling should generally show strong stretching-rate dependence and strong hysteresis on retraction.

Most experiments that show the 65 and 150 pN transitions have been done in "fixed-extension" setups where the force transducer has a rather stiff force constant (AFM, optical trap [12,6]). We expect that if one carries out DNA stretching at *fixed force* as can be done with a magnetic tweezer system [39], one should expect a totally different character of the two transitions. At 65 pN, one will see a reversible force-extension curve, but as unpeeling starts to occur, one should observe a series of extension plateaus (jumps in length) corresponding to thermal activation of the unpeeling fork over GC-rich barrier regions. This is closely related to similar dynamics of DNA unzipping at fixed force [24,25,27,40].

On a molecule with nicks only at its ends, unpeeling might occur near the ends if they are AT rich, for ≈40 to 50 pN, but only until the first GC-rich "barrier" is reached. Given the typical sequence distribution along genomic DNAs, this will typically be only a few hundred base pairs (a nick in a particularly long AT-rich region may cause longer unpeeling at 150 mM NaCl below the *B* to *S* transition, as we have shown for 156Gmac). After unpeeling stalls at the first GC-rich barrier, the remainder of the molecule will convert to *S*-DNA at 62–68 pN. Finally, at higher forces ≈150 pN, unpeeling will become favorable, and all barriers to unpeeling will disappear at >200 pN, causing the second transition to a ssDNA-like force response. We note that we have also found that for 150 mM NaCl, small bubbles of parallel-strand separation occur in the most weakly bound regions of molecule (typically AT-rich regions) near 50 pN.

On molecules with many nicks, many unpeeling initiation sites will reduce the force threshold at which unpeeling effects will be seen, causing a merger of *B* to *S* and unpeeling, and molecular rupture at forces near 60 pN. For solution conditions that favor double-helix melting (e.g., low salt or elevated temperature), either unpeeling, or if no nicks are present, partial parallel-strand separation, can preempt formation of *S*-DNA.

Thus a 5′-3′ ("same strand") tether ought to undergo a gradual unpeeling transition as force is slowly raised above 43 pN; note that the beginning of unpeeling will occur from the *B* state, while much of the unpeeling will occur from the *S* state. If the "free" strand is in fact in the form of many short fragments, they might be observed to leave the stretched strand starting at 38 pN. Nicks introduced at known positions along one strand as would be obtained with single-strand-cutting ("nickase") restriction enzymes could allow precise experiments to test our model. Another controlled experimental situation would be a molecule with no nicks except at one end (i.e., an unnicked molecule with a single-strand connection at one end, and a double-strand anchor at the other); we expect two transitions at 150 mM NaCl: first a nearly reversible *B* to *S* transition at 62 pN, and then unpeeling from the single-strand-anchor end near 150 pN. To eliminate other nicks, such an experiment would be best done on a rather short DNA (e.g., 5 kb).

The *B*-DNA double helix, unpeeled DNA and *S*-DNA all have nearly the same free energy for 25 mM NaCl, near 60 pN (the *B*-*S*-ssDNA "triple point" of Fig. 6). Near this point, small changes in sequence will determine whether a double helix with a nick in it will retain a base-paired structure (as will occur for GC-rich DNA), or alternately whether a short ssDNA region will unpeel (as for AT-rich DNA). Such a mechanism suggests a way for DNA sequence to regulate initiation of general recombination. In the bacterium *E. coli* it is thought that an invading ssDNA becomes coated

with the protein RecA, which facilitates its displacement and replacement of the homologous ssDNA in the recipient chromosome [41]. Intriguingly, RecA has been shown to be able to generate tensions of up to 100 pN on a single DNA molecule [33]. Perhaps part of RecA's function during the early stages of recombination is to apply some tension to the dsDNA being invaded; nicks adjacent to particularly AT rich regions would, according to our analysis, release ssDNA which could facilitate initiation of strand exchange.

A second possible biological relevance of our work concerns the opening of parallel single-stranded regions, which can occur for forces slightly below the *B* to *S* transition for highly AT rich regions. Gene transcription initiation sites just happen to be very AT rich, and have been shown to in general melt at relatively low temperatures [42] and to be the first sequences to open when a molecule is supercoiled (i.e., untwisted) [43]. Our sequence-specific calculations suggest that DNA tension in the 50–60 pN range could play a role in opening AT-rich transcription initiation sites, a necessary step for transcription to begin. RNA polymerase is known to be capable of generating this level of force [44], and so if an RNA polymerase were anchored (e.g., by its transcript as is thought to be the case in E. coli) tension transmitted to the DNA template could open nearby initiation sites. This provides a mechanism for cooperative transcription of adjacent genes.

Note added in proof. We note that Zhou *et al.* have formulated a model of the *B*-DNA to ERRORS-DNA transition, based on failure of stacking interactions, without base unpairing [45]. We thank Z.-C. Ou-Yang for bringing this work to our attention.

ACKNOWLEDGMENTS

This research was supported by the NSF through Grant No. DMR-0203963, and by a grant from the Johnson and Johnson Focused Giving Grant Program. We are grateful to S. B. Smith, M.-N. Dessinges, V. Croquette, and D. Bensimon for sending us their experimental data. J.F.M. and S.C. are grateful to S. Kowalczykowski and R. Monasson for discussions of this problem.

APPENDIX: PHENOMENOLOGICAL MODEL OF ssDNA ELASTICITY

The ssDNA free energy used in this paper is phenomenological, chosen to combine the logarithmic dependence of extension on force seen at ~mM NaCl concentration, with the force-threshold observed for extension seen at >50 mM NaCl concentration. [2,12,30,29]. These effects have been argued to be due to a competition between long-ranged Coulomb repulsion of the charges along the ssDNA backbone and short-ranged self-adhesion of the bases on the basis of computer calculation [31,29]. Similar conclusions follow from analytical calculations based on semiflexible polymers with screened Coulomb interactions [40].

The extension per base pair we use is

$$x_{ss} = h\left(\frac{a_1 \ln f/f_1}{1 + a_3 e^{-f/f_2}} - a_2 - f/f_3\right), \tag{20}$$

where the external force f is in piconewtons, and where $h = 0.34$ nm, $a_1 = 0.21$, $a_2 = 0.34$, $f_1 = 0.0037$ pN, $f_2 = 2.9$ pN, and $f_3 = 8000$ pN. The parameter $a_3 = 2.1 \ln(M/0.0025)/\ln(0.15/0.0025) - 0.1$ depends on NaCl concentration M (in Mol/litre). It is possible for the above extension formula to be less than zero; in this case, we set the extension to be zero. This provides the force threshold needed for extension of ssDNA observed for NaCl concentrations >50 mM [2,12,30,29].

The $\ln f$ provides the characteristic log-force dependence of extension discussed by Zhang *et al.* [31], while the denominator describes the reduction in contour length generated at low force by self-adhesion ("folding") of the chain [40]. The dependence of a_3 on NaCl concentration modulates the strength of this effect. We emphasize that the above model's virtue is to provide a smooth description of experimental data. The free energy per base pair is obtained using the integral (3).

[1] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron, Science **271**, 792 (1996).
[2] S. B. Smith, Y. J. Cui, and C. Bustmante, Science **271**, 795 (1996).
[3] C. G. Baumann, S. B. Smith, V. A. Bloomfield, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **94**, 61850 (1997).
[4] M. Rief, H. Clausen-Schaumann, and H. E. Gaub, Nat. Struct. Biol. **6**, 346 (1999).
[5] J.-F. Leger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko, Phys. Rev. Lett. **83**, 1066 (1999).
[6] H. Clausen-Schaumann, M. Rief, C. Tolksdorf, and H. E. Gaub, Biophys. J. **78**, 1997 (2001).
[7] J.-F. Leger, Ph.D. thesis, l'Université Louis Pasteur Strasbourg I, 2000.
[8] I. Rouzina and V. A. Bloomfield, Biophys. J. **80**, 882 (2001).
[9] I. Rouzina and V. A. Bloomfield, Biophys. J. **80**, 894 (2001).
[10] J. R. Wenner, M. C. Williams, I. Rouzina, and V. A. Bloomfield, Biophys. J. **82**, 3160 (2002).
[11] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, Proc. Natl. Acad. Sci. U.S.A. **94**, 11 935 (1997).
[12] C. Bustamante, D. Smith, and S. Smith, Curr. Opin. Struct. Biol. **10**, 279 (2000).
[13] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, Biophys. J. **82**, 1537 (2002).
[14] J. SantaLucia, Jr., Proc. Natl. Acad. Sci. U.S.A. **95**, 1460 (1998).
[15] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, Woodbury, NY, 1994), pp. 302–312.
[16] S. Cocco, J. F. Marko, and R. Monasson, Proc. Natl. Acad. Sci. U.S.A. **98**, 8608 (2001).
[17] S. Cocco, R. Monasson, and J. F. Marko, Phys. Rev. E **66**,

051914 (2002).
[18] S. Cocco, J. F. Marko, and R. Monasson, Eur. J. Phys. **10**, 153 (2003).
[19] A. Sarkar, J.-F. Leger, D. Chatenay, and J. F. Marko, Phys. Rev. E **63**, 051903 (2001).
[20] A. Lebrun and R. Lavery, Nucleic Acids Res. **24**, 2260 (1996).
[21] P. Cizeau and J.-L. Viovy, Biopolymers **42**, 383 (1997).
[22] C. Storm and P. C. Nelson, Phys. Rev. E **67**, 051906 (2003).
[23] J.-F. Allemand, D. Bensimon, R. Lavery, and V. Croquette, Proc. Natl. Acad. Sci. U.S.A. **95**, 14 152 (1998).
[24] D. Lubensky and D. R. Nelson, Phys. Rev. Lett. **85**, 1572 (2000).
[25] D. Lubensky and D. R. Nelson, Phys. Rev. E **65**, 031917 (2002).
[26] U. Gerland, R. Bundschuh, and T. Hwa, Biophys. J. **81**, 1324 (2001).
[27] M. Prentiss, D. Lubensky, and D. R. Nelson, Proc. Natl. Acad. Sci. U.S.A. **100**, 1694 (2003).
[28] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).
[29] M.-N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette, Phys. Rev. Lett. **89**, 248102 (2002).
[30] B. Maier, D. Bensimon, and V. Croquette, Proc. Natl. Acad. Sci. U.S.A. **97**, 12 002 (2000).
[31] Y. Zhang, H. J. Zhou, and Z. C. Ou-Yang, Biophys. J. **81**, 1133 (2001).
[32] K. J. Breslauer, R. Frank, H. Blocker, and L. A. Marky, Proc. Natl. Acad. Sci. U.S.A. **83**, 3746 (1986).
[33] J.-F. Leger, G. Romano, J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko, Proc. Natl. Acad. Sci. U.S.A. **95**, 12 295 (1998).
[34] D. H. Turner, Curr. Opin. Struct. Biol. **6**, 299 (1996).
[35] V. A. Bloomfield, D. M. Crothers, and I. Tinoco, Jr., *Physical Chemistry of Nucleic Acids* (Harper and Row, New York, 1974).
[36] R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).
[37] J. DeMasi, S. Du, D. Lennon, and P. Traktman, J. Virol. **75**, 10090 (2001).
[38] S. Smith, L. Finzi, and C. Bustamante, Science **258**, 1122 (1992).
[39] T. R. Strick, J.-F. Allemand, D. Bensimon, and V. Croquette, Biophys. J. **74**, 2016 (1998).
[40] S. Cocco, J. F. Marko, R. Monasson, A. Sarkar, and J. Yan, Eur. J. Phys. **10**, 249 (2003).
[41] P. R. Bianco, R. B. Tracy, and S. C. Kowalczykowski, Front. Biosci. **17**, D570 (1998).
[42] E. Yeramian, Gene **255**, 151 (2000).
[43] G. W. Hatfield and C. J. Benham, Annu. Rev. Genet. **36**, 175 (2002).
[44] M. D. Wang, M. D. Schnitzer, H. Yin, R. Landick, J. Gelles, and S. M. Block, Science **282**, 902 (1998).
[45] H. J. Zhou, and Z.-C. Ou-Yang, Phys. Rev. Lett. **82**, 4560 (1999).