

Unzipping double-stranded DNA with a force: Numerical results

Jeff Z. Y. Chen

Department of Physics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

(Received 4 October 2001; published 25 September 2002)

A double-stranded DNA molecule pulled with a force acting on the strand terminals exhibits a partially denatured structure or can be completely unzipped when the pulling force goes beyond a critical force. It has been suggested that accompanying the unzipping transition, various power-law properties exist. Through the numerical solution to a model that contains heterogeneous bonding interactions between bases on the two strands, we evaluated the critical forces and the extension-force curves for various degree of sequence disorderliness, and compared the numerical results with predictions from analytical approaches.

DOI: 10.1103/PhysRevE.66.031912

PACS number(s): 87.15.-v, 87.10.+e, 64.70.-p

I. INTRODUCTION

Schematically shown in Fig. 1 is a double-stranded DNA, each strand containing a disordered sequence of bases that are paired with their counterparts on the other strand. The paired segments contain mostly complementary *A-T* and *G-C* pairings. Recent experimental achievements have made it possible to probe the structure of a double-stranded DNA (dsDNA) molecule, typically by applying an external force acting on the terminal bases [1,2]. In recent theoretical studies, it has been suggested that an unzipping critical force exists beyond which the entire dsDNA molecule can be unzipped. Earlier work by Vedenov, Dykhne, and Frank-Kamenetskii concentrated on the theoretical development of the coil-helix transition occurring at the melting transition based on a Ising-model-like treatment [3]. Much recent discussion concerns a simple model in which two strands are modeled by two polymer chains, whose pairing potentials between the different bases have been modeled by a homogeneous short-ranged potential of various types [4–11]. Most work is based on the continuous description of a polymer chain and its mapping to a quantum-mechanical problem. The thermal melting and forced unzipping transition can be represented in a phase diagram with the temperature and external force as parameters, and such a phase diagram has been generated both from continuous and lattice descriptions [10,11]. The heterogeneous nature of the base sequences has been ignored in these studies until recently, when studies of the melting and unzipping transition that contain different pairing potentials have been reported [12,13].

In this paper, we report our numerical solution to the Hamiltonian model (Sec. II) that was proposed by Lubensky and Nelson [13], which contains an explicit simulation of the sequence dependence of the bases. Each generated sequence of dsDNA model is then examined through the numerical solution of a “time” dependent Schrödinger equation, by utilizing the Crank-Nicholson method (Sec. III). We mainly concentrated on discussing the unzipped extension as a function of the external force and the unzipping phase diagram, produced from the numerical simulations (Secs. IV–VI).

II. BASIC MODEL

The basic assumption is to approximate the configuration of each of the two strands in the dsDNA by a Wiener mea-

surement. The spatial position of the s th base, labeled continuously from $s=0$ at one end to $s=N$ at the other, is denoted as $\mathbf{r}_i(s)$, where i ($=1,2$) represents the two strands. Each base unit on the $i=1$ chain interacts with a counterpart unit on the $i=2$ chain. The probability functional for the configuration of the system is then assumed to have the form

$$P[\mathbf{r}_1(s), \mathbf{r}_2(s), \{s\}] \propto \exp \left\{ -\frac{D}{2b^2} \int_0^N \left[\left(\frac{d\mathbf{r}_1}{ds} \right)^2 + \left(\frac{d\mathbf{r}_2}{ds} \right)^2 \right] ds - \beta \int_0^N V[\mathbf{r}_1(s) - \mathbf{r}_2(s), s] ds \right\}, \quad (1)$$

where b is the basic Kuhn length. The pairing interaction, represented by the short-ranged $V(\mathbf{r}, s)$ that explicitly depends on a distribution of the sequence $\{s\}$, should have a force range approximately the same as or smaller than b , and is generally assumed to be different for different base pairs involved. Upon the introduction of two new variables $\mathbf{r}_\pm = \mathbf{r}_1 \pm \mathbf{r}_2$ and the reduction of the probability function to eliminate the \mathbf{r}_+ variable, one arrives at an effective probability function [13]

$$P_0[\mathbf{r}_-(s), \{s\}] \propto \exp \left\{ -\frac{D}{2a^2} \int_0^N \left[\left(\frac{d\mathbf{r}_-}{ds} \right)^2 \right] ds - \beta \int_0^N V[\mathbf{r}_-(s), s] ds \right\}, \quad (2)$$

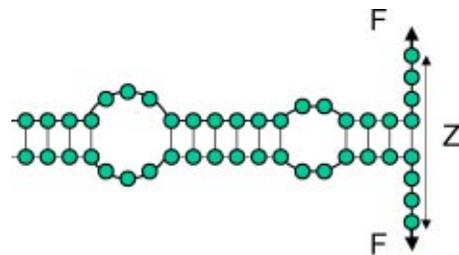


FIG. 1. Schematic representation of the dsDNA model considered. Each pair interacts with $\zeta(s)\epsilon$, where $\zeta(s)$ is a random number, and the sequence is assumed to be heterogeneously disordered.

which contains the relative position vector \mathbf{r}_- only, where $2a^2 = b^2$. The complete probability function with an unzipping force acting on the N th base pairs is then

$$P_{\text{unzip}}[\mathbf{r}_-(s), \{s\}] = P_0[\mathbf{r}_-(s), \{s\}] \exp\{\beta \mathbf{F} \cdot \mathbf{r}_-(N)\}. \quad (3)$$

The main task is to solve the partition function problem of this probability for possible sequences of the pairing distribution, and find system-specific and universal behavior.

It is our interest here to treat the problem numerically. For numerical convenience we further approximate the original probability in a one-dimensional version,

$$P_0[z(s), \{s\}] \propto \exp\left\{-\frac{1}{2a^2} \int_0^N \left[\left(\frac{dz}{ds}\right)^2 - \beta \int_0^N V[z(s), s] ds\right\}. \quad (4)$$

Hence the unzipping probability function can be written as

$$P_{\text{unzip}}[z(s), \{s\}] = P_0[z(s), \{s\}] \exp\{\beta F \cdot z(N)\} \quad (5)$$

with the assumption that the pulling force is directed along the z direction only. One of the most interesting problem in statistical physics is to examine the stability of dsDNA binding using a D -dimensional probability function [Eq. (2)]. A denaturing (or melting) transition might occur without an unzipping force at a high temperature, which invariably leads to the questioning of the very role that dimensionality plays in such a transition [14]. No conclusive evidence has been provided for determining the upper dimensionality of the heterogeneous model. Cule and Hwa [12] have studied the melting transition of a heterogeneous DNA model near the melting transition. They have concluded that the variation of the pairing potentials has a weak effect on the melting behavior of a heterogeneous DNA model.

In principle, to rigorously model real dsDNA systems, we should examine the above probability function in three dimensions. Our main focus here, however, is the unzipping transition under a strong external force at a temperature, significantly away from the melting transition region. Here, we have used the term ‘‘melting transition’’ strictly for the denaturing transition occurring while the temperature varies with no external or a small external unzipping force. Since the governing phenomenon in the strong force region is one dimensional along the direction of the force, examining a one-dimensional version of the original probability function is justified.

The mathematical form of the probability function in Eq. (4) can be related to many other physical phenomena. The closest one is in the study of random copolymer localization under an external field, specifically, at a sharp interface [15]. In the current situation, further simplification can be made if we are only interested in the probability distribution of the N th terminals, not the conformation of the entire chain that would depend on all spatial coordinates of every base pair. The calculation of such a probability function $\psi_0(z, N, \{s\})$,

as a function of the separation distance z , of a DNA model containing a total of N base pairs can be initially made unrelated to any force field. The additional factor of the unzipping effects can be included by considering the exponential factor in Eq. (5).

Going through an analysis of the analogy between the probability function in Eq. (4) and the path integral approach in quantum physics, one can show that $\psi_0(z, N, \{s\})$ in a pairing field V obeys a time dependent Schrödinger-like equation,

$$-\frac{\partial \psi_0(z, N, \{s\})}{\partial N} = \left[-\frac{a^2 d^2}{2 dz^2} + \beta V(z, N) \right] \psi_0(z, N, \{s\}), \quad (6)$$

where the wave function ψ is subject to the initial condition $\psi_0(z, 0, \{s\}) = 1$. The same equation was recently examined by various authors in the context of dsDNA unzipping, when $V(z, N)$ has no sequence dependence [7,9,11]. Zhou also considered the case where the potential is not symmetrically centered at $z=0$ [16]. In our case, since the pairing potential depends on a specific sequence in a dsDNA model, $\psi_0(z, N, \{s\})$ has an explicit sequence dependence.

It thus becomes clear that an effective numeric approach to study the problem is to generate a number of DNA models by specifying statistically independent disordered sequences; through a numerical solution of the Schrödinger equation for each of these models, we can collect statistics both in the conformational ensemble and in the disorder sequence ensemble.

III. NUMERIC PROCEDURE

As we mentioned earlier, the pairing potential is a function of the path coordinate s . Spatially, the potential is short ranged, and has been modeled in other studies by a $-\delta(z)$ function [7,9] or a square-well function [10]. Here, in order to implement a numerical scheme based on the finite-difference scheme, we write

$$V(z, s) = \zeta(s) \epsilon w(z), \quad (7)$$

and use a Gaussian potential well $w(z) = -\exp[-(z/a)^2]/\sqrt{\pi}$. This particular choice of a smooth function will influence the statistics in a small spatial region of size a near $z=0$, but will not cause any serious harm to the large-scale properties in a strong external force that we are targeted to observe. The overall magnitude of the pairing potential is given by ϵ (assumed positive) and the information on the sequence dependence is contained in the dimensionless coefficient $\zeta(s)$.

In this paper, we wish to observe the evolution of different physical properties as the sequence varies from the completely random state to the homogeneous state whose $\zeta(s)$ has no sequence dependence. In particular, we parametrize $\zeta(s)$ in the form

$$\zeta(s) = [\alpha + 4(1 - \alpha)r(s)] / [\alpha + 4(1 - \alpha)], \quad (8)$$

where $r(s)$ is a random number ranging from -1 to 1 . The relative randomness in the system is controlled by the parameter α , such that $\alpha=0$ corresponds to the complete randomness and $\alpha=1$ corresponds to a homopolymer.

It is unknown for real systems to what extent the interaction and the sequence is truly random. Even for the case of $\alpha=0$ where no apparent net pairing force exists, it can be shown that bonding may still occur. Without knowing exactly what value of α we should use in practice, we will discuss the simulation results that span the entire range of $\alpha=[0,1]$. At this point, we wish to stress that using a random number for $\zeta(s)$ is only a theoretical abstraction of the real interaction sequences in these systems. The current numerical procedure can be used unaltered if a more realistic sequence can be used in lieu of $\zeta(s)$ in Eq. (8). Some information on the free energy cost of $A-T$ and $G-C$ pairings, for example, can be found in literature [1,17].

To implement the numerical procedure in practice, we have to confine ourselves to a finite region $z/a=(-L,L)$, which was divided into $2L$ evenly spaced parts. A time step of $\Delta N=1$ and a spatial step of $\Delta z/a=1$ was used in the discrete scheme. The numerical procedure used to invert Eq. (6) follows a standard Crank-Nicholson algorithm, as is further described in [18]. The fact that we are dealing with a tridiagonal matrix to approximate the operator on the right hand side of this equation further reduces the computation load. To effectively collect information on the sequence-dependent features of the problem, in total 10^3 statistically independent sequences have been generated for each given α . The numerical scheme used here, when written in a discrete form, is principally similar to the transfer matrix calculation in [12].

Considering the sequence-independent exponential factor in Eq. (5), we see that the calculation of unzipping probability can be separated into the calculation of the unperturbed probability $\psi_0(z,N,\{s\})$, and then, by multiplying the force dependent factor, performing the final analysis in various values of F . Hence our main task is to collect the statistics of the unperturbed $\psi_0(z,N,\{s\})$.

The numerical procedure and thus the probability function resulting from it depends on L , the dimension of the system considered in the numerical simulation, and N , the total number of pairs in the model. It is important to minimize the finite-size effects due to the limited value of L and N chosen in the actual calculation. Some of the L - and N -dependent effects will be discussed in Sec. VI.

IV. PHASE DIAGRAM AND THE UNZIPPING TRANSITION

In the preceding section we have discussed that the characteristics of the statistical properties of the models can be represented by the probability function of the N th terminal, $\psi_0(z,F,\{s\})$. In this section we calculate the free energy per base pair for the $F=0$ case and relate the calculation to a critical unzipping force in a phase diagram.

To begin with, we discuss the conformational properties of the system under no external force with the system size $L=1600$ and $N=32000$. This set of parameters has the minimal finite-size effect among other systems that have

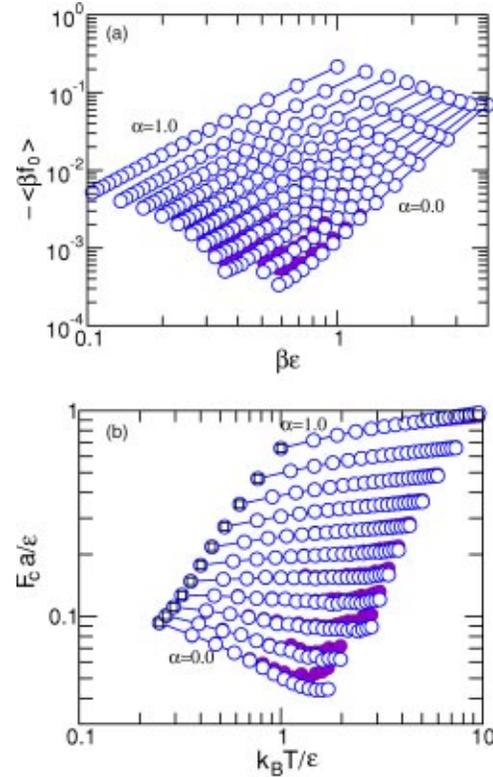


FIG. 2. Free energy per base pair as a function of the inverse temperature (a) and the critical force as a function of the temperature (b). The reduced free energy per base pair, $-\beta\langle f_0 \rangle$, is deduced from the partition function of the considered systems. In plot (a) open symbols correspond to the system size $L=1600$, $N=36000$ and filled symbols to $L=800$, $N=16000$. The difference between the symbols indicates of errors. In plot (b) the free energy in (a) has been used, together with the formula for the critical force ($D=1$) in Eq. (10), to produce the critical lines. The squares in plot (b) represent the points near which the scaling behavior was studied in this work.

been examined, as will be shown later and is used in this section, unless otherwise explained. Knowing the probability function $\psi_0(z,N,\{s\})$ for each generated sequence, we can calculate the free energy per monomer, $f_0(\{s\})$, based on

$$\beta f_0(\{s\}) = -\frac{1}{N} \ln \int_{-L}^L \psi_0(z,N,\{s\}) dz. \quad (9)$$

The sequence-averaged reduced free energy per particle, $\beta\langle f_0 \rangle$, is then calculated from the 10^3 samples that we have generated.

Figure 2(a) shows $\beta\langle f_0 \rangle$ as a function of $\beta\epsilon$ for $\alpha=0.0, 0.1, 0.2, \dots, 1.0$, obtained from our numerical simulations. The open symbols correspond to the calculation result based on $L=1600$, $N=32000$ and the filled symbols correspond to a same calculation based on $L=800$, $N=16000$. The relative displacement of the symbols indicates the magnitude of the numerical error associated with the data.

Earlier, Lubensky and Nelson have determined that the critical pulling force needed to unzip the entire dsDNA is

$$F_c a = \sqrt{2D} \langle f_0 \rangle / k_B T. \quad (10)$$

Displayed in Fig. 2(b) is a phase diagram that was reproduced based on the data in Fig. 2(a) with lines connected between the data points to approximate the critical lines. Curves with ($\alpha > 0.7$) show a different sign of slope from those of smaller α . Marenduzzo and co-workers recently discussed a possible denaturing transition when the temperature is lowered while fixing the unzipping forces in a homopolymer model [10,11]. The critical line for the $\alpha = 1$ case in Fig. 2(b) resembles the critical line produced in their paper, where they have used a square-well potential, rather than the Gaussian potential used here. This unusual behavior is an artifact of the Gaussian attractive well used in the current calculation. We can show by adding a “hard wall” potential at $z=0$ that all the curves display the same slope as the lower α 's. As the critical lines in Fig. 2(b) continuously approach $F_c=0$, we expect a finite melting temperature to occur. The determination of the melting temperature, however, involves much more physical consideration than the current simple polymer model. Due to the repositioning of each base pair to achieve a double helix alignment, one needs to consider not only the bonding energy gain but also the entropic loss associated with the coil-helix transition [3]. Nevertheless, the role that heterogeneity plays near the melting transition in a simple polymer model treatment has been addressed recently in [12].

In this paper, we concentrate on the unzipping transition at a temperature much lower than the melting transition. The squares in Fig. 2(b) specify the location where an analysis of the unzipping behavior was conducted [$\beta\epsilon = \alpha + 4(1 - \alpha)$]. An accurate estimate of the free energy per monomer has a direct consequence on further analysis of the universal properties of the unzipping transition. Listed as the second column in Table I is the averaged free energy with the number in the parentheses showing the digits where $\beta\langle f_0 \rangle$ differs in comparing the calculated results of the next largest system considered ($L=800, N=16000$). The third column in Table I contains the standard deviation of the free energy based on the consideration of the 10^3 sample sequences. The fourth column in Table I contains F_c based on Eq. (10) and column 2, which will be tested below for consistency in a scaling analysis near the unzipping transition.

V. SCALING NEAR THE UNZIPPING TRANSITION

Following Eq. (5), we concluded that $\psi_F(z, N, \{s\})$, the probability function for the N th terminal associated with the presence of an unzipping force F , is obtained from that of a free dsDNA model, $\psi_0(z, N, \{s\})$, by an additional factor $\exp(\beta F z)$:

$$\psi_F[z, N, \{s\}] = \psi_0[z, N, \{s\}] \exp\{\beta F z\}. \quad (11)$$

We next consider the conformational properties when an external force F is applied to the N th terminal. In particular, we have chosen the inverse temperature $\beta\epsilon = \alpha + 4(1 - \alpha)$ for different values of α in the scaling study.

Figure 3(a) shows an example of the averaged and nor-

TABLE I. Statistical properties of the bounded dsDNA model. The second column corresponds to the scaled free energy per base pair, $\beta\langle f_0 \rangle$, for various values of disorderliness represented by α . The number in the parentheses of the second column represents the estimated error due to finite system size, and $\langle \delta f_0 \rangle$ in the third column represents the standard deviation based on 1000 sampled sequence samples. The fourth column is the critical force needed to completely unzip a long dsDNA, as calculated according to Eq. (10) based on the values given in the first column. The fifth, sixth, and seventh columns contain the effective exponent q_1 , q_2 , and q_3 , respectively, as defined in Eqs. (13), (14), and (15).

α	$\beta\langle f_0 \rangle$	$\beta\langle \delta f_0 \rangle$	$\beta F_c a$	q_1	q_2	q_3
0.0	-0.0694(2)	0.000 11	0.3724	1.94	3.99	2.94
0.1	-0.0696(3)	0.000 07	0.3732	1.84	3.83	2.84
0.2	-0.0722(3)	0.000 06	0.3701	1.82	3.78	2.77
0.3	-0.0778(3)	0.000 03	0.3946	1.78	3.65	2.86
0.4	-0.0864(2)	0	0.4159	1.69	3.50	2.83
0.5	-0.0986(1)	0	0.4442	1.67	3.46	2.71
0.6	-0.1145(4)	0	0.4789	1.51	3.07	2.55
0.7	-0.1344(2)	0	0.5186	1.39	2.72	2.41
0.8	-0.1578(2)	0	0.5620	1.27	2.39	2.28
0.9	-0.1848(1)	0	0.6083	1.20	2.18	2.14
1.0	-0.2148(1)	0	0.6555	1.16	2.10	2.10

malized probability $\psi_F(z, N, \{s\})$ as a function of z at three different values of the external force, $\beta F a = 0, 0.32$, and 0.36 when $\alpha = 0$ (complete random sequence). The terminal end can be seen as being highly localized in the vicinity around an average. The fact that a loop segment in the dsDNA prefers to be separated when the unzipping process reaches it makes the terminal density centralize at a finite distance. In sharp contrast, the probability function for the homopolymer case ($\alpha = 1$) shows a more extended, exponential decaying profile [see Fig. 3(b)]. As a matter of fact, one can explicitly show that the large z behavior for a $-\delta(z)$ potential well is indeed an exponential function.

In previous theoretical approaches, the number of unzipped base pairs, $\langle m \rangle$, has been considered in terms of a scaling relationship with $F_c - F$. In order to calculate $\langle m \rangle$ directly, we have to handle the added requirement of storing

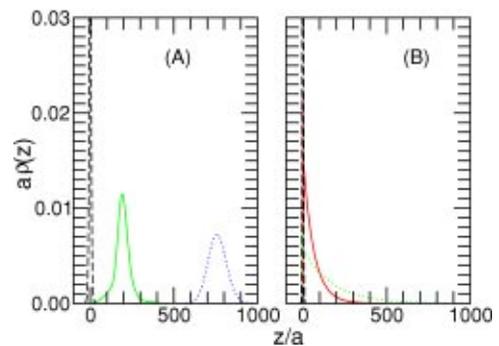


FIG. 3. Probability function of the pulled terminals as a function of the separation distance for $\alpha=0$ (a) and $\alpha=1$ (b). The long dashed lines near $z=0$ represent the density profile at $F=0$. The solid and dotted curves correspond to a finite value of $\beta F a$.

the partition functions of all base pair segments under *explicit* external force F , for $s=1,2,\dots,N$. While such a calculation of $\langle m \rangle$ is not impossible from a numerical approach similar to the current one, the evaluation of the average distance of the N th terminals is much easier. It is straightforward to consider

$$\bar{A}(z, \{s\}) = \frac{\int_{-L}^L A(z) \psi_0(z, N, \{s\}) \exp(\beta F z) dz}{\int_{-L}^L \psi_0(z, N, \{s\}) \exp(\beta F z) dz} \quad (12)$$

for any function of the terminal separation z .

To demonstrate the sequence dependence of the separation distance \bar{z} as a function of the external force F , we have plotted ten different cases in Fig. 4 for $\alpha=0$. Due to the difference in the pairing sequence, some DNA models can be partially denatured with a smaller force, whereas others require a much larger force. The steplike functions are a signature of the disorderliness in the sequence, as have been observed many times experimentally. All curves flatten out near the system size, $z/a=L=1600$.

While individual dsDNA shows its own characteristics, the system averaged quantity over the sequence ensemble demonstrates *universal* conformational properties, as has been recently pointed out [13]. In particular, $\langle \bar{z} \rangle$ can be written in terms of a scaling relationship with respect to the force difference $F_c - F$ near F_c :

$$\langle \bar{z} \rangle / a \propto (F_c - F)^{-q_1}. \quad (13)$$

The top curve in Fig. 5(a) represents $\langle \bar{z} \rangle / a$ for $\alpha=0$ as a function of $\beta(F_c - F)a$ in a double logarithmic scale. We see that within the region of $[0.03, 0.02]$, $\langle \bar{z} \rangle$ indeed follows a power behavior, with a fitted exponent of $q_1=1.94$, which is in good agreement with the theoretical prediction of $q_1=2$ [13]. Exception of the scaling can be seen in the smaller $F_c - F$ region where the saturation of $\langle \bar{z} \rangle$ to a finite value occurs. As a direct consequence of the *compound* finite-size effect shown for various sequences in Fig. 4, the averaged $\langle \bar{z} \rangle$ has a smaller scaling region in a finite system. Further

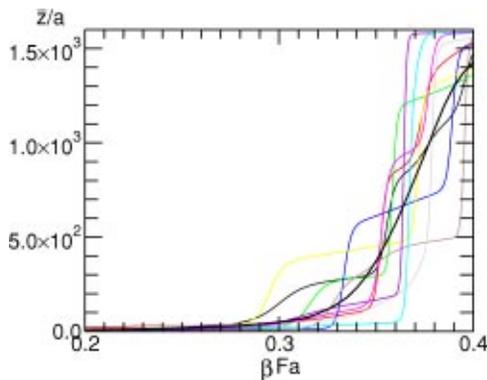


FIG. 4. Typical unzipping force versus extension curves for ten different random sequences ($\alpha=0$). The thick curve represents an average over 10^3 independently generated samples.

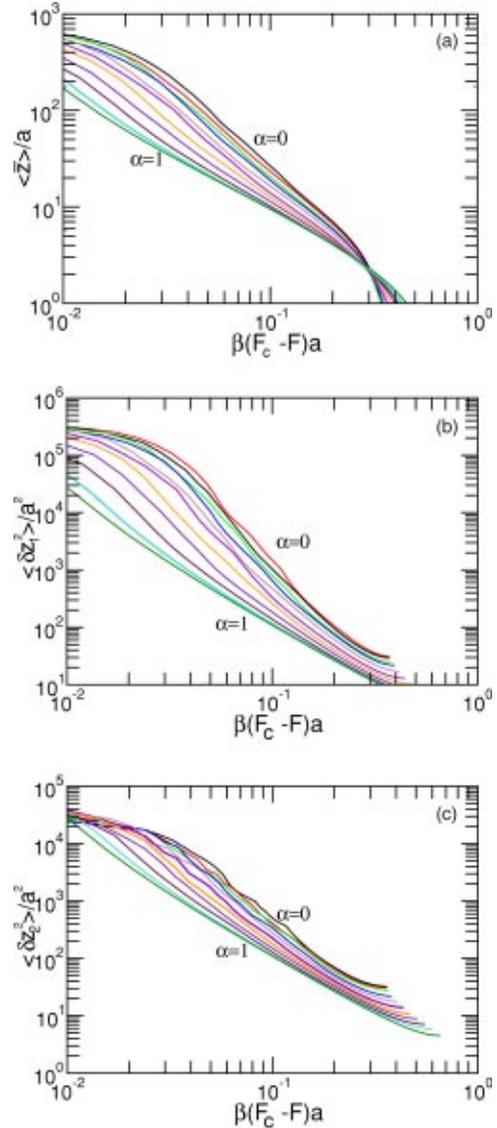


FIG. 5. Scaling behavior of the unzipped extension, $\langle \bar{z} \rangle$ (a), and also δz_1^2 (b) and δz_2^2 (c), as a function of the force difference $F_c - F$. Curves from the top to bottom correspond to $\alpha = 0.0, 0.1, 0.2, \dots, 1.0$. A system size of $L = 1600, N = 32000$ has been used to produce these curves. For comparison, see Eqs. (13)–(15).

evidence that the flattened part is a manifestation of the finite-size effect is provided in the following section.

As a comparison, \bar{z} for the homopolymer case ($\alpha=1$) is also plotted in the same figure (the bottom curve). Considering an inverse δ function pairing potential that produces an exponential profile, one can show analytically that $q_1=1$. The fitted scaling exponent from our numerical results, $q_1=1.16$, is comparable to the theoretical prediction.

The curves between these two extreme cases in Fig. 5 correspond to $\alpha=0.1, 0.2, \dots, 0.9$ from top to bottom. Models with smaller α display an interesting crossover phenomenon: near the transition point the curves tend to approach an asymptotic $q_1=2$ behavior, and in the intermediate region, the curves tend to follow an asymptotic $q_1=1$ behavior. The upward bending trend of some of the curves, in particular,

the $\alpha=0.3$ case, indicates that a crossover point exists between the two power laws. Strictly speaking, a sensible analysis of the data for these α 's should be done by explicitly considering the crossover behavior, which will demonstrate the important aspect of the size of a "critical region" near F_c . Since our inadequate data does not allow a substantial analysis of the crossover, we have simply fitted the data in Fig. 5 to a power law. The *effective* scaling exponent q_1 found this way is listed in the fifth column of Table I.

The different shapes of the distribution function in Fig. 3 can be further characterized by averaged higher moments. In particular, Lubensky and Nelson have discussed two other power laws, which are listed here in terms of averages of z :

$$\delta_1^2 z \propto \langle \bar{z}^2 \rangle - \langle \bar{z} \rangle^2 \propto (F_c - F)^{-q_2}, \quad (14)$$

$$\delta_2^2 z \equiv \langle \bar{z}^2 \rangle - \langle \bar{z} \rangle^2 \propto (F_c - F)^{-q_3}. \quad (15)$$

These scaling laws can also be examined in a similar way. Figures 5(b) and 5(c) demonstrate how the scaling laws are obeyed in double logarithmic plots. Again, the effective exponents q_2 and q_3 are determined, from a least squares linear fit, and are listed in Table I. Clearly, from an inspection of these figures, we conclude that a scaling relationship is well followed. The values of $q_2=3.99$ and $q_3=2.94$ compare favorably with the theoretical prediction of $q_2=4$ and $q_3=3$ for $\alpha=0$. The value of $q_2=q_3=2.10$ also agrees with the expected exponent of $q_2=2$.

VI. FINITE-SIZE EFFECTS

In any numerical analysis one has to treat the finite-size effects carefully. Indeed, there are two factors that need to be considered here. The finite length of model DNA, N , and the finite dimension of the system simulated in the numerical treatment, L .

Taking the $\alpha=0$ case, for example, we have considered various pairs of the parameters N , L in our simulation. For adequately long model DNA (large N), the maximum separation distance between the two pulled terminals is limited by the system dimension L . Taking into account the additional consideration that various sequences can be independently pulled at various distances, the useful information for the averaged quantity, when all the sequences are considered, is limited to a fraction of L .

Figure 6 displays $\langle \bar{z} \rangle$ for various values of N , L . We have chosen the particular value of N for each L in such a way that doubling N would display no major difference in these curves. The approaching of the scaling behavior of large- L systems can be observed in this figure. The most important feature of this series of plots is the indication that the asymptotic power-law behavior is valid only for very large L system, in which the terminal separation has a significant distance such that the corresponding critical region can be reached. The figure clearly demonstrates the need for studying sufficiently large- L systems in order to find the anticipated scaling behavior.

The demand for large L in a numerical calculation, how-

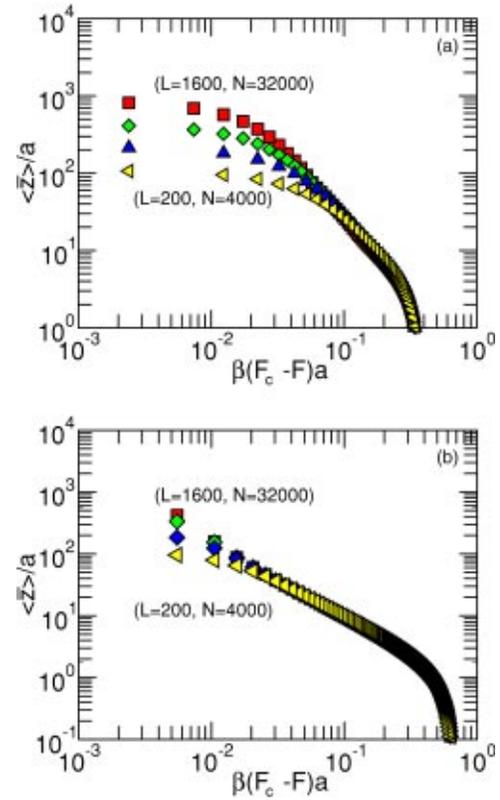


FIG. 6. The finite-size effects on the separation distance of the terminals. Squares, diamonds, up triangles, and left triangles represent system sizes ($L=1600, N=32000$), ($L=800, N=16000$), ($L=400, N=8000$), and ($L=200, N=4000$), respectively. In order to observe the scaling behavior, a large L is required. The two cases, $\alpha=0$ and $\alpha=1$, are displayed separately in (a) and (b).

ever, is restricted by the actual limitation of the precision of a real number. Although most of the calculations in this work were performed in double precision, the subtle balance of the ψ_0 function and the exponential factor in Eq. (11) relies on handling extremely large and small real numbers. We can show that the current system size, $L=1600$, is actually very close to the computational limitation, even with a number of numerical tricks that have been implemented to reduce the severe situations.

The demand for large L in a numerical calculation would also require a longer polymer chain to be simulated. We are dealing with the thermodynamic limit of the unzipping transition where both N and L goes to infinity. So long as N is finite, the system always displays a completely unzipped state as L approaches ∞ , since the energy at $z=\infty$ is lower than any paired or partially paired states. The current model, and indeed our simulations, pertains to another limit, namely, infinitely long DNA's ($N=\infty$) in a large but finite system (finite L). Hence, as L goes up, we need to simulate a much longer dsDNA to avoid a shorter dsDNA being easily pulled apart and starting to accumulate near the boundary of the simulated system at $z=L$.

VII. CONCLUDING REMARKS

In this work, we have carried out a numerical study of a dsDNA unzipping model that contains heterogeneous se-

quences of pairing potential energy. We have examined the unzipping critical force as a function of the temperature and also the scaling behavior near the unzipping transition. We have treated the problem carefully by examining various system sizes to minimize the finite-size effects, which have been shown to play an important role near the transition force. Spanning over the entire range of disorderliness, we have

demonstrated the differences and the similarities between a homopolymer and heteropolymer model for dsDNA.

ACKNOWLEDGMENT

Financial support for this work was provided by the Natural Science and Engineering Research Council of Canada.

-
- [1] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, *Phys. Rev. Lett.* **79**, 4489 (1997); *Phys. Rev. E* **58**, 2386 (1998).
- [2] J. F. Leger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko, *Phys. Rev. Lett.* **83**, 1066 (1999).
- [3] A. A. Vedenov, A. M. Dykhne, and M. D. Frank-Kamenetskii, *Sov. Phys. Usp.* **14**, 715 (1972).
- [4] D. Poland and H. A. Sheraga, *J. Chem. Phys.* **45**, 1456 (1966).
- [5] M. Peyrard, *Europhys. Lett.* **44**, 271 (1998).
- [6] M. Peyrard and A. R. Bishop, *Phys. Rev. Lett.* **62**, 2755 (1989).
- [7] S. M. Bhattacharjee, *J. Phys. A* **33**, L423 (2000); e-print cond-mat/0010132.
- [8] R. E. Thompson and E. D. Siggia, *Europhys. Lett.* **31**, 335 (1995).
- [9] H. Zhou, e-print cond-mat/0007015.
- [10] D. Marenduzzo, A. Trovato, and A. Maritan, *Phys. Rev. E* **64**, 031901 (2001).
- [11] D. Marenduzzo, S. M. Bhattacharjee, A. Maritan, E. Orlandini, and F. Seno, e-print cond-mat/0103142.
- [12] D. Cule and T. Hwa, *Phys. Rev. Lett.* **79**, 2375 (1997).
- [13] D. K. Lubensky and D. R. Nelson, *Phys. Rev. Lett.* **85**, 1572 (2000).
- [14] M. S. Causo, B. Colozzi, and P. Grassberger, *Phys. Rev. E* **62**, 3958 (2000).
- [15] T. Garel, D. A. Huse, S. Leibler, and H. Orland, *Europhys. Lett.* **8**, 9 (1989).
- [16] Note that the combination $\kappa\beta$ should be considered as a temperature independent constant in Ref. [9].
- [17] I. Tinoco, Jr., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla, *Nature New Biol.* **246**, 40 (1973).
- [18] Z. Y. Chen, *J. Chem. Phys.* **111**, 5603 (1999).