

Effects of the eye phase in DNA unzipping

Debaprasad Giri¹ and Sanjay Kumar^{2,*}

¹Physics Section, MMV, Banaras Hindu University, Varanasi 221 005, India

²Department of Physics, Banaras Hindu University, Varanasi 221 005, India

(Received 8 October 2005; revised manuscript received 27 February 2006; published 30 May 2006)

The onset of the “eye phase” (a phase consisting of configurations of eye-type conformations or bubbles in the double-stranded DNA) and its role during the DNA unzipping is studied when a force is applied to the interior of the chain. The directionality of the hydrogen bond introduced here shows oscillations in force-extension curve similar to a “sawtooth” kind of oscillations seen in the protein unfolding experiments. The effects of intermediates (hairpins) and stacking energies on the melting profile have also been discussed.

DOI: 10.1103/PhysRevE.73.050903

PACS number(s): 87.14.Gg, 36.20.Ey, 64.90.+b, 82.35.Jk

Molecular interactions play a key role in living organisms. Recent advances in experimental techniques have allowed nanomanipulation in a single biological molecule and they made it possible to measure these interactions [1]. The aim is to exert a force in the pN range by optical tweezers, atomic force microscopy, etc. and characterize the molecular, elastic, structural, and functional properties of biomolecules [2,3]. In typical experiments of double-stranded DNA (dsDNA) unzipping, a force is applied to the ends of the chain [Fig. 1(a)], and one studies the force-extension curve which shows the elastic properties and the gross features that reflect the local “G-C” vs “A-T” content along the sequence [4–8]. Moreover, the force-temperature diagram shows, below the melting temperature, that the overstretching force [8] decreases nearly linearly with temperature.

Theoretically, DNA unzipping may be studied in different ensembles [9] depending on the experiments. For example, atomic force microscopes (AFM) work in constant distance ensemble (CDE), while a magnetic bead uses the concept of constant force ensemble (CFE). The prediction of unzipping transition based on interacting Gaussian chains [10] raised a lot of interest, and now results are available from the dynamical approach [11], exact solutions of lattice models [12,13], simple models of quenched-averaged DNA [14,15], numerical simulations, and scaling analysis [16,17]. Recently, for a model of interacting polymers where any monomer of one chain can interact with any monomer of the other chain (we call it model A), the role of an intermediate entropy-stabilized phase was recognized, and a force-induced triple point [18] in a force-temperature plane was established.

In most of the models studied for dsDNA [12–16,19], a monomer i of one strand can only interact with the i th monomer of the other strand, which is similar to the models of DNA (we call it model B) proposed earlier by Poland and Scheraga [20]. These models do not take into account the directional nature of the hydrogen bond and underestimate the entropy by restricting the formation of hydrogen bonds. Thus, these models may only give a limited picture of the unzipping transition and do not allow for studying the effect of intermediate states [21].

So far, most of the theories of DNA unzipping have focused on the thermodynamic limit, and therefore, consist of a few parameters typically in the form of an effective base-pair interaction in a simplified way. In all the single molecule experiments, a finite size chain is used, and hence, no “true phase transition” can, in principle, be observed. Still, the phase transition observed in such experiments may be considered as real if the length of the chain exceeds the characteristic correlation lengths. Now, it is becoming possible to go to the other limit of studying shorter segments (~ 10 base pair) at the coarse-grained level [22,23]. The purpose of this paper is to provide the exact results of a semimicroscopic model of short chains by incorporating the directional nature of hydrogen bonds and then propose a method to study the effect of molecular interactions right at the individual base-pair level and their role on melting profile.

In the following, we adopt a more realistic model of DNA, which may be defined in any dimension [24] (herein after we call it model C). A similar model has been used in Ref. [24] in the context of relative stabilities of DNA hairpin structures. The model takes care of important shortcomings of model B and also incorporates some additional features such as the existence of intermediate states, effects of stacking energy, excluded volume properties of nucleotides, and the directional nature of hydrogen bonds.

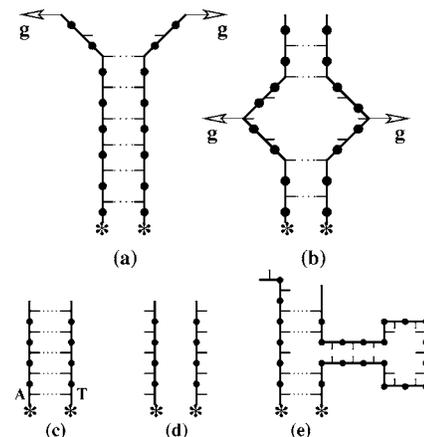


FIG. 1. The schematic of DNA unzipping by force (g) applied at (a) one end of the two strands; (b) at the interior of the strands, which gives rise to the eye-type conformation. (c)–(e) represent the possible conformations of model C. Here, (c) represents the completely zipped state, (d) a nonpairing configuration, and (e) a partial bound state with a hairpin loop.

*On leave at Max-Planck Institute for the Physics of Complex Systems, Noethnitzer, 01187 Dresden, Germany.

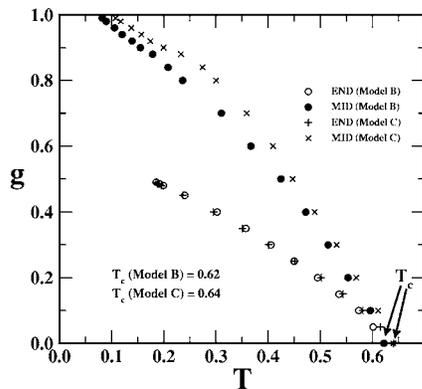


FIG. 2. Variation of critical force [$g_c(T)$] (END and MID cases) as a function of temperature (T) for models B and C.

Although the importance of bubble formation during thermal melting has been recognized, so far no attempt has been made to experimentally explore the phase diagram when a force is applied to the interior of the chain as shown in Fig. 1(b). Such situations occur in many biological processes, for example, during gene expression, RNA forms bubbles or “eye-type” conformations on the DNA. [13] Therefore, we consider two cases: (i) force has been applied at the end of the chain [Fig. 1(a), “END” case or Y case], and (ii) at the middle of the chain [Fig. 1(b), “MID” case or “Eye” case]. The contribution to energy by this force, g , is $-2gx$, where $2x$ is the absolute distance in the x direction between the two strands at the point of application of the force.

We model the two strands (e.g., A in one strand and T in other strand) of a homopolymer DNA by two mutually attracting-self-avoiding walks (MASAWs) on a square lattice as shown in Figs. 1(c)–1(e). The bases are associated with the link between two monomers of a chain as depicted in Fig. 1. In one strand, the bases point toward the right, while on the other they are on the left, as one traverses the chains sequentially. We associate a contact energy $-\epsilon$ (effective base-pair interaction) with each pairing between complementary strands only if the bases are nearest neighbors (short-range nature of the hydrogen bond) and approach each other directly without the strands coming in between [Fig. 1(c)]. Figure 1(e) shows the possibility of the formation of a hairpin (which is not possible in model B) in a single strand of DNA. However, in this case, non-native contribution has been taken into account, but no apparent weight has been assigned to the stem as it is made up of the same nucleotides.

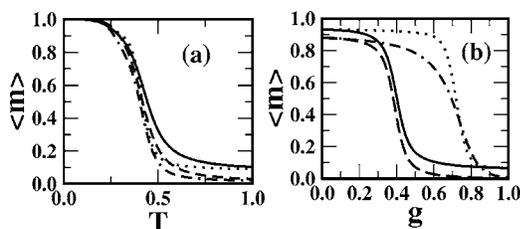


FIG. 3. Variation of scaled $\langle m \rangle$ (a) with temperature (T) at constant force $g=0.25$ (END) and 0.50 (MID); (b) with force (g) at constant $T=0.3$. Here, solid and dashed lines represent the END case, while dot and dotted-dashed line represent the MID case for models B and C, respectively.

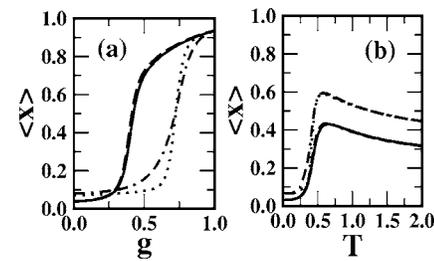


FIG. 4. Plot of scaled $\langle x \rangle$ (a) with force (g) at constant $T=0.3$ and (b) with temperature (T) at constant $g=0.25$ (END) and 0.5 (MID). The lines have the same meaning as in Fig. 3.

The partition function (Z_N) of the system under consideration can be written as a sum over all possible configurations of MASAWs, i.e., $\sum_{m,x} C(m,x) \exp(\beta m \epsilon) \exp(\beta g x)$, where $\beta=1/k_B T$ is the inverse temperature, k_B being the Boltzmann constant. $C(m,x)$ is the number of distinct conformations walks of length $2N$ having m number of intact base pairs whose end (or mid) points are at a distance x apart. We have obtained $C(m,x)$ for $N \leq 16$ and analyzed the partition function using the exact enumeration and a series analysis technique [25,26]. We prefer this technique, because it can predict various phases of the system [27] quite effectively and the scaling corrections can be correctly taken into account by a suitable extrapolation method [25,26]. To achieve the same accuracy in Monte Carlo, a chain of 2 orders of magnitude larger than in the exact enumeration method is required [28]. We set $\epsilon/k_B=1$ and calculate the reduced free energy per base pair from the relation $G(T,g) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln Z(T,g) = \ln \mu(T,g)$ [26]. The limit $N \rightarrow \infty$ is achieved by using the ratio method [25] for extrapolation. The transition point can be obtained from the plot of $G(T,g)$ vs T or from the peak value of $\partial^2 G / \partial (1/T)^2$.

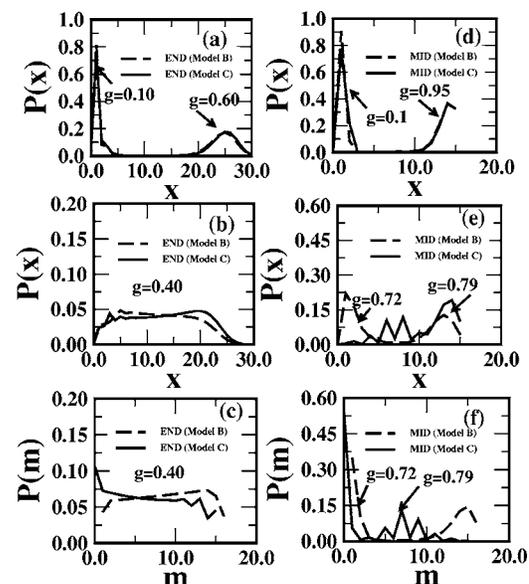


FIG. 5. (a)–(f) represent the $P(x)$ and $P(m)$ of the END case (a)–(c) and the MID case (d)–(f) for different forces at constant $T=0.3$. (e) shows the signature of the eye phase of even widths in the form of oscillation for the MID case in model C but is absent in model B.

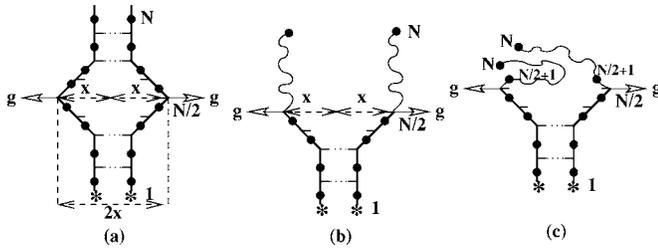


FIG. 6. Schematics showing the different conformations keeping $2x$ ($\leq N^\nu$, where ν is the end-to-end distance exponent) distance constant. For model B, (b) and (c) have the same weight, while for model C, (c) has more weight than (b).

A force-temperature (g - T) phase diagram of models B and C for END and MID case is shown in Fig. 2. The qualitative features of the phase diagram obtained here may be compared with experiments [8]. The phase boundary separates the zipped and the unzipped state. At $T=0$, the critical force can be found from a simple analytic calculation and is equal to 0.5 [18] that is in agreement with the critical force from Fig. 2.

Experimentally, the melting profiles are obtained by monitoring the change in the UV absorbance with temperature that provides the information about the fraction of open base pairs, and the melting temperature is defined when half of the total base pairs get open [29]. Another quantity of experimental interest is to monitor and measure the end separation by varying the force. We calculate these quantities from the expressions $\langle m \rangle = \sum_m C(m, x) \exp(\beta m \epsilon) \exp(\beta g x) / Z_N$ and $\langle x \rangle = \sum_x C(m, x) \exp(\beta m \epsilon) \exp(\beta g x) / Z_N$ and plot their variations with temperature and force in Figs. 3 and 4, respectively. It is evident from Fig. 3 that, with increasing temperature (or force), the number of intact base pairs decreases, and there is a sharp transition from a zipped state to the unzipped state.

Remarkably, to break the same amount of base pairs at low temperature, one requires almost double the force if it is applied in the middle of the chain, rather than the END [Fig. 3(b)] consistent with the exact results [13]. Near the melting temperature, fluctuation dominates and less than double the force is required for unzipping the chain from the middle.

Another interesting observation is the variation of average elongation with force, which shows a monotonically increasing trend [Fig. 4(a)] at constant temperature and approaches unity. Meanwhile, the variation of extension at constant force shows a sharp rise with temperature [Fig. 4(b)] and then a slight decline to approach a value below unity. At constant temperature, there is a transition from the zipped to the unzipped (“rod-like”) state. Keeping force constant, when temperature is varied, there is a transition, and the chain acquires conformations close to the rod-like states. As this temperature is still low, with a further increase in temperature, the entropy of the system increases and the chain acquires a coil-like state at a higher temperature, and thus, the average distance decreases.

We also study the probability distribution curves $P(x)$ with x and $P(m)$ with m for models B and C defined by $P(x) = \sum_m C_N(m, x) \exp(\beta m \epsilon) \exp(\beta g x) / Z_N$ and

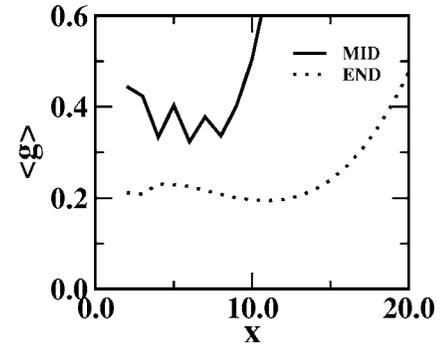


FIG. 7. Plot of $\langle g \rangle$ with distance (x) at constant $T=0.5$.

$P(m) = \sum_x C_N(m, x) \exp(\beta m \epsilon) \exp(\beta g x) / Z_N$, respectively. In Figs. 5(a)–5(f), we have shown $P(x)$ and $P(m)$ for different values of force and a fixed temperature $T=0.3$. The x component of the distribution function gives information about the states of dsDNA. The maxima of $P(x)$ at $x \approx 0$ correspond to the zipped state for a given set with $g=0.1$ and $T=0.3$. For small forces, thermal fluctuations are too weak to unzip the strand, and the DNA remains in the zipped state. This is being reflected in the structureless distribution function with a well-defined peak at the most likely value of the extension [Figs. 5(a) and 5(d)]. For both the models at higher force and at the same temperature ($g=0.6$ for END and 0.95 for MID, respectively, at $T=0.3$), thermal fluctuations have been suppressed by a strong force and strands are found in the segregated rod-like state with more or less an identical distribution peaked at maximum extension. However, striking differences are observed in the probability distribution curves for models B and C at intermediate forces below the transition line (g - T plane), when the force is applied in the middle. No such differences are observed for the END case. For the MID case, the probability distribution curve of model C shows strong oscillations, whereas for model B no such oscillations are observed. This indicates that the model B may be described by two state models. However, model C shows that certain intermediate states (eye shape of even width) are more favorable than others.

It is interesting to note that the g - T phase diagram of models B and C are almost identical without any re-entrance at low temperature. This may be because of the fact that the energy of the ground state and the unbound state for both models are the same. At the center point of the “Y” (end case), the phase boundary is determined by a balance of the net force $-2gx$ and the unzipping potential $-\epsilon m$ with associated entropy. In the thermodynamic limit, directionality of the hydrogen bonds, which appear in the form of entropy of the partial bound states does not play a crucial role in this balance. The absence of re-entrance is understood with the zero entropy of the ground state for both models. As shown recently by Kapri *et al.* [13] for MID case, in the CDE there is a possibility of a coexistence region that is better thought of as an eye phase \equiv two Y joined together. In this case, the separation at the point of application of force is smaller than the fully open case, and thus, such conformations statistically have more weight than the other conformations.

Since the bottom end is kept fixed, the top side of the strand may open due to thermal fluctuations [Fig. 6(b)] and

form a partial bound state as shown in Fig. 6(c), which is more stable than Fig. 6(b). Therefore, in model C, half of the chain undergoes an unzipping transition, while the other half due to the non-native contacts, shows the combined effects of unzipping and slippage (shearing) transition. Thus, in model C, the transition appears more smoother than in model B, which can be seen in Fig. 3(b). The consecutive peaks in $P(x)$ vs x curve [Fig. 5(e)] represent the eye phase of even widths that contribute most to the partition function. Because of native contacts, the contribution of the eye shape conformation is significantly less and small thermal fluctuations are sufficient to unzip the chain in model B.

We substantiate our arguments by extending the calculation also to CDE. The partition function in CDE may be defined as $Z_N(x, T) = \sum_m \exp(\beta m \epsilon)$. The two ensembles are related by $Z_N(T, g) = \sum_x Z_N(x, T) \exp(\beta g x)$ [13,14]. The free energy is given by the relation $F_N(x, T) = -T \ln Z_N(x, T)$. In CFE, the average separation $\langle x \rangle$ fluctuates, while in CDE, one measures the average force to keep the separation constant given by the expression $\langle g \rangle = \partial F_N(x, T) / \partial x$ at a constant temperature [9]. Thus, the force-extension curve obtained is shown in Fig. 7 for model C, which also shows oscillations for the MID case but is constant for the END case. Though a somewhat similar effect has been seen experimentally in molecules like Titin [30], DNA has not been probed so far. The average of force obtained here (0.2 ± 0.05 for the END case and 0.35 ± 0.05 for the MID case at $T=0.5$) also matches with the value shown in Fig. 2.

The stacking energy in the case of homopolymers gets adsorbed in the effective base pair interaction. To see this, we associate an additional energy between two consecutive parallel base pairs (only possible in model C) and found that there is no qualitative change in the phase diagram except the shift in the transition temperature. If stacking energy is negative (attractive interaction), the chain becomes stiffer. However, it does not change the nature of oscillations as observed in probability distribution curves for the MID case [31].

The exact results on short chains of a new semimicroscopic dsDNA that incorporates the directional nature of hydrogen bond show unequivocal signature of an eye phase, without going to the long chain limit. This happens for the case with a force acting in the middle of the dsDNA. The variation of elongation due to the force in different ensembles has a different behavior both qualitatively and quantitatively. We anticipate that refinements in high-precision single molecule experiments will be able to verify these predictions.

We thank Y. Singh, S. M. Bhattacharjee, and Jeff Chen for many fruitful discussions on the subject and University Grants Commission, India for the financial assistance. One of us (S.K.) would like to acknowledge financial support from MPIPES, Dresden, Germany.

-
- [1] M. Rief *et al.*, *Science* **276**, 1112 (1997); C. Bustamante, Z. Bryant, and S. B. Smith, *Nature (London)* **421**, 6921 (2003).
 [2] M. Carrion-Vazquez *et al.*, *Prog. Biophys. Mol. Biol.* **74**, 63 (2000); A. Engel *et al.*, *Curr. Biol.* **9**, R133 (1999).
 [3] U. Bockelmann, *Curr. Opin. Struct. Biol.* **14**, 368 (2004).
 [4] J. C. Sitko *et al.*, *Biophys. J.* **84**, 419 (2003).
 [5] Y. Zeng, A. Montrichok, and G. Zocchi, *Phys. Rev. Lett.* **91**, 148101 (2003).
 [6] P. Cluzel *et al.*, *Science* **271**, 792 (1996).
 [7] S. B. Smith *et al.*, *Science* **271**, 795 (1996).
 [8] I. Rouzina and V. A. Bloomfield, *Biophys. J.* **80**, 894 (2001); *ibid.* **80**, 894 (2001).
 [9] C. Bustamante, J. Liphardt, and F. Ritort, *Phys. Today* **58**, 43 (2005).
 [10] S. M. Bhattacharjee, *J. Phys. A* **33**, L423 (2000).
 [11] K. L. Sebastian, *Phys. Rev. E* **62**, 1128 (2000).
 [12] D. Marenduzzo, A. Trovato, and A. Maritan, *Phys. Rev. E* **64**, 031901 (2001); D. Marenduzzo, S. M. Bhattacharjee, A. Maritan, E. Orlandini, and F. Seno, *Phys. Rev. Lett.* **88**, 028102 (2002).
 [13] R. Kapri, S. M. Bhattacharjee, and F. Seno, *Phys. Rev. Lett.* **93**, 248102 (2004); R. Kapri and S. M. Bhattacharjee, *J. Phys.: Condens. Matter* **18**, 5215 (2006).
 [14] D. K. Lubensky and D. R. Nelson, *Phys. Rev. Lett.* **85**, 1572 (2000); *Phys. Rev. E* **65**, 031917 (2002).
 [15] Pui-Man Lam *et al.*, *Biopolymers* **73**, 293 (2004).
 [16] N. Singh and Y. Singh, *Eur. Phys. J. E* **17**, 7 (2005).
 [17] Jeff Z. Y. Chen, *Phys. Rev. E* **66**, 031912 (2002).
 [18] S. Kumar, D. Giri, and S. M. Bhattacharjee, *Phys. Rev. E* **71**, 051804 (2005).
 [19] M. Baiesi, E. Carlon, and A. L. Stella, *Phys. Rev. E* **66**, 021804 (2002).
 [20] D. Poland and H. A. Scheraga, *J. Chem. Phys.* **45**, 1456 (1966); **45**, 1464 (1966).
 [21] C. Danilowicz *et al.*, *Phys. Rev. Lett.* **93**, 078101 (2004); C. Danilowicz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1694 (2003).
 [22] T. Strunz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11277 (1999).
 [23] M. J. Lang *et al.*, *Nat. Methods* **1**, 1 (2004).
 [24] S. Kumar *et al.*, *Europhys. Lett.* **70**, 15 (2005).
 [25] A. J. Guttmann, *Phase Transition and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic, New York, 1989) Vol. 13; P. K. Mishra, S. Kumar, and Y. Singh, *Physica A* **323**, 453 (2003).
 [26] C. Vanderzande, *Lattice Models of Polymers* (Cambridge University Press, U.K., 1998).
 [27] D. Marenduzzo, A. Maritan, A. Rosa, and F. Seno, *Phys. Rev. Lett.* **90**, 088301 (2003).
 [28] Y. Singh *et al.*, *J. Phys. A* **34**, L67 (2001).
 [29] R. M. Wartel and A. S. Benight, *Phys. Rep.* **126**, 67 (1985).
 [30] M. Rief *et al.*, *Science* **275**, 1295 (1997).
 [31] S. Kumar and D. Giri (unpublished).