Dynamical Scaling of the DNA Unzipping Transition

D. Marenduzzo,[1] Somendra M. Bhattacharjee,[2,3,4,*] A. Maritan,[1,4] E. Orlandini,[3] and F. Seno[3]

[1]*International School for Advanced Studies (SISSA), and INFM, Via Beirut 2-4, 34014 Trieste, Italy*

[2]*Institute of Physics, Bhubaneswar 751 005, India*

[3]*INFM, Dipartimento di Fisica, Università di Padova, Via Marzolo 8, 35131 Padova, Italy*

[4]*The Abdus Salam International Center for Theoretical Physics (ICTP), Strada Costiera 11, 34100 Trieste, Italy*

(Received 6 April 2001; published 28 December 2001)

We report studies of the dynamics of a set of exactly solvable lattice models for the force-induced DNA unzipping transition. Besides yielding the whole equilibrium phase diagram, which reveals a reentrance, these models enable us to characterize the dynamics of the process starting from a nonequilibrium initial condition. The thermal melting of DNA displays a model dependent time evolution. On the contrary, the dynamical mechanism for the unzipping by force is very robust and the scaling behavior is independent of the details of the description and, hence, superuniversal.

DOI: 10.1103/PhysRevLett.88.028102 PACS numbers: 87.15.He, 64.60.Cn, 64.60.Ht, 87.14.Gg

The replication of DNA is a correlated process involving many proteins and other molecules [1] working at different points in space and time. It has recently been shown [2–6] that the force-induced unzipping of DNA is a genuine phase transition different from its thermal melting transition. It was then hypothesized [2] that the initiation of replication at the "origins" along the DNA, e.g., by dnaA for E.Coli [1,7] or by the "origin recognition complex" (ORC) in eukaryotes [8] is like this unzipping near the phase transition point (with dnaA or ORC acting as the force-inducing agent). Recently, techniques like laser tweezers [9] and atomic force microscopes (AFM) [10–12] have been used to study DNA by pulling at one end. The observed hysteresis in AFM unzipping experiments is a signature of a first order transition [11]. Furthermore, the activities of polymerases, topoisomerase, and a few others on a single stranded DNA have also been analyzed in terms of the force exerted by or applied against them [13–15]. Similar in vitro investigation of the full replication process would then require an understanding of the coupling between the opening of the strands and the subsequent events during replication. Such a study involves the dynamics of the unzipping process [3].

The purpose of this Letter is to define a set of models, for which the unzipping transition can be studied exactly. Based on this, a detailed study of the dynamics not relying on any mean-field approximation can be performed. Such studies are so far lacking in literature. Our models are extensions of the Poland and Sheraga model [16], in which the strands are taken as random walkers with base pair interaction between them. Our models are different because there is a force f pulling the open end of the strands (tied at the other end) and the self- and mutual-avoidance (hard-core repulsion) of the strands are incorporated too. This ensures that the zero-force melting temperature T_m takes a finite value also in dimension $d = 2$ unlike in Ref. [16]. Similar models with self-avoidance but with no external force, and not analytically solvable, have been recently proposed to study equilibrium denaturation [17]. Two types of lattice models are considered here: (i) The "Y model" in which the two strands, of length N, are zipped together up to a bifurcation point; i.e., the only allowed configurations are those which have the first $N - m$ monomers bound and the remaining m separated as in a Y. (ii) The "b model" in which configurations with "bubbles" are allowed. In real situations there are single-strand binding (SSB) proteins which bind to the opened strands to prevent rejoining. This justifies the "Y model" where the bubble formation (rejoining) is suppressed. However, in vitro studies need not contain the SSB proteins and bubbles can form anywhere along the strands due to thermal fluctuations. This is realized in the b model (see Fig. 1a). In both cases the equilibrium phase diagram displays a reentrant region at low temperature (T): for a finite range of forces the molecule gets unzipped by *decreasing* T. The dynamics of both the b and the Y models in the various phases and on the phase boundary are then studied, by starting from a nonequilibrium bound state as the initial condition. By using Monte Carlo dynamics, we find that in all regimes above or on the phase boundary, the time evolutions of the order parameters follow dynamical scaling laws. The basic features of both statics and dynamics (but see below) are maintained for any d. This allows us to concentrate on the $d = 2$ case. Generalizations to $d > 2$ are easy and are discussed later.

We model the two strands of DNA by two directed self- and mutually avoiding walks. In $d = 2$ on the square lattice (see Fig. 1a), the two walks are forced to follow the positive direction of the diagonal axis $(1, 1)$ (i.e., the coordinate along this direction always increases). A force f acts along the transverse $(-1, 1)$ direction (the x direction). By measuring the x separation in units of the elementary square diagonal, we say that two complementary monomers are in contact when this separation is 1: a binding energy $-\epsilon$ (we choose $\epsilon = 1$) is gained for each contact. Because of the geometrical properties of the lattice, all these contacts contributing to the energy involve monomers labeled by the same base pair index, as one

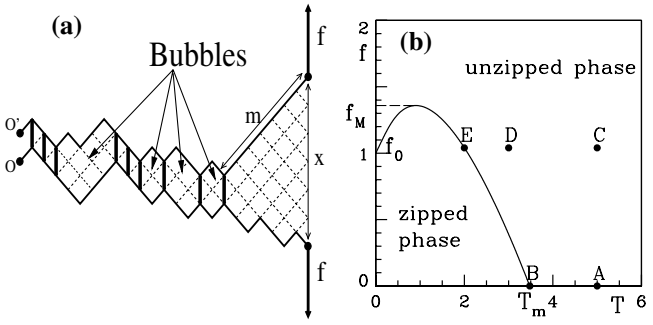

FIG. 1. (a) A typical configuration of the two DNA strands as modeled on a square lattice. Thick vertical lines indicate monomers which are in contact. The quantities m and x are graphically represented. (b) Plot of the f vs T phase diagram for the model. Note the reentrance from $T = 0$ up to $T \simeq 0.260\,193 T_m$.

would require for base pairing in DNA. The quantities of interest are the average open end separation $\langle x \rangle$, the average number of contacts Θ, and the average number of "liberated" monomers (i.e., from the last contact to the end) $\langle m \rangle$ (see Fig. 1a).

We calculate $\langle x \rangle$, for the two N-step chains, one starting at $(0, 1)$ and the other at $(1, 0)$, at $T = \beta^{-1}$ (with the Boltzmann constant $k_B = 1$) and under a force f. A discontinuity or nonanalyticity in $\langle x \rangle$, for $N \to \infty$, would signal a phase transition. We can obtain $\langle x \rangle$ exactly from the partition function $\mathcal{Z}_N(\beta, f)$ as $\langle x \rangle = \frac{\partial \ln \mathcal{Z}_N}{\partial \beta f}$, where $\mathcal{Z}_N(\beta, f) = \sum_{x \geq 1} d_N(x) \exp(\beta f x)$, and $d_N(x)$ represents the fixed distance partition function, i.e., the sum over all interacting pairs of directed chains whose last monomers (open ends) are at distance x. The thermodynamic behavior comes from the singularity closest to the origin of the related grand partition function [6,18,19] $\mathcal{G}(\mathsf{z}, \beta, f) = \sum_{N=0}^{\infty} \mathsf{z}^N \mathcal{Z}_N(\beta, f)$ where z is the step fugacity. From the recursion relation $d_{N+1}(x) = [2d_N(x) + d_N(x+1) + d_N(x-1)][(e^{\beta} - 1)\delta_{x,1} + 1]$, for the b model, together with the boundary conditions $d_N(0) = 0 \;\forall\; N$, and $d_0(x) = \delta_{x,1}$, one finds two competing singularities of $\mathcal{G}(\mathsf{z}, \beta, f)$, namely, $\mathsf{z}_1 = \sqrt{1 - e^{-\beta}} - 1 + e^{-\beta}$ (β dependent), and $\mathsf{z}_2 = \frac{1}{2[1+\cosh(\beta f)]}$ (both β and f dependent). If the smallest singularity is z_2, the molecule is in the unzipped phase; otherwise it is in the zipped one. By equating these two singularities we get the critical line, $f_c(T)$, separating the two phases. For the b model, $f_c(T) = T \cosh^{-1}[p(\beta) - 1]$, with $p(\beta) = 1/(2\mathsf{z}_1)$ (see Fig. 1b). For the Y model, $f_c(T)$ is similar with $p(\beta) = e^{\beta}$.

For both the models, the critical force at $T = 0$ is $f_0 = 1$. There is a maximum threshold f_{M} such that for $f > f_{\mathrm{M}}$ the system is always in the unzipped state. For the b model, $T_m = \frac{1}{\ln 4/3}$ and $f_{\mathrm{M}} = 1.358\,806\ldots$, while $T_m = \frac{1}{\ln 2}$ and $f_{\mathrm{M}} = 1.282\,143\ldots$ for the Y model. Both Θ and $\langle x \rangle$ show discontinuities at $f = f_c(T)$, implying a first order transition, everywhere except for the b model at $f = 0$ in $2 \leq d \leq 5$. For $f_{\mathrm{M}} > f > f_0$ the usual zipping transition is present but if T is still lowered the two strands cleave again through a "cold unzipping." The exact results on reentrance can be understood in a more general way. Let us consider a configuration of DNA which is dominant as $T \to 0$ in our models, in which the first $N - m$ monomers are zipped and the remaining are completely stretched by the force f. The energy and entropy gain with respect to the completely unbounded case is thus $-(N - m) + af(N - m)$ and $(N - m)c$, respectively, where c is the entropy per bounded unit and a is a geometrical factor. Thus $f_c = (1 + cT)/a$ as $T \to 0$ which shows the emergence of reentrance. For our models, $a = 1$ and $c = \log 2$ give the exact low-T phase boundary derived earlier. A lower value of c (which in any case ought to be >0 [20]) may occur if, e.g., the bound region has a larger persistence length. The argument given suggests that reentrance is common to all polymeric models of DNA similar to ours [20]. Even if the realistic parameters of DNA (persistence lengths) might confine this effect to unrealistically low T, with a proper choice of solvent and double stranded polymers reentrance should be observable. In the continuum approximation of Refs. [2,4,5] this effect is not found since these models are not valid in the low T regime as the average bond length vanishes as $T \to 0$.

We now consider the dynamics of the b model and the Y model. In both cases, we start from a nonequilibrium initial condition with the two chains zipped in a zigzag configuration (as at $T = 0$), and let the system evolve at a temperature T and under a force f, with T and f chosen so that the equilibrium state is either on or above the phase boundary. Let us first consider the two-dimensional case. The five regimes considered are marked A–E in Fig. 1b. Numerically, a Monte Carlo dynamics with one-bead local move (discretized Rouse model) is used to monitor the time evolution of m and x (Fig. 1). We accept or reject the moves according to a probability given by the Boltzmann factor as in the standard Metropolis algorithm. We reject configurations where the strands cross. In all cases we find the following dynamical scaling laws [21] to hold good:

$$m(t) = N^{d_1} G_m(t/N^{z_1}) \sim t^{\theta_1}, \tag{1}$$

$$x(t) - x(0) = N^{d_2} G_x(t/N^{z_2}) \sim t^{\theta_2}, \tag{2}$$

with $\theta_i = d_i/z_i$ (omitting the average signs for simplicity of notation). Here N is the length of each chain, $G_{m,x}$ are two scaling functions, and time (t) is measured in units of N Monte Carlo attempts. Equations (1) and (2) also define the exponents $d_{1,2}$ and $z_{1,2}$ for the two variables with $\theta_{1,2}$ describing the early time evolution away from saturations. Note that $d_{1,2}$ can be obtained through equilibrium considerations as one requires $m(t) \sim N^{d_1}$ and $x(t) \sim N^{d_2}$ for $t \to \infty$. The crossover to the equilibrium behavior is described by the "dynamic" exponents $z_{1,2}$. The exponents are obtained from simulations by collapsing (see Fig. 2) the Monte Carlo data according to Eqs. (1) and (2) [22]. The results are summarized in Table I.

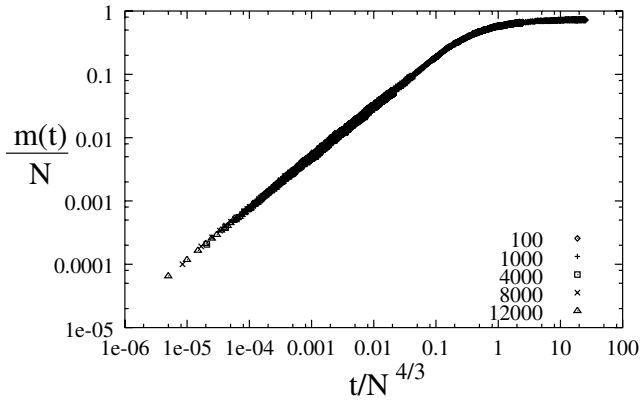

FIG. 2. Plot of $m(t)/N$ vs $t/N^{4/3}$ for various values of N in regime A (see Fig. 1b) for the b model: the collapse of all curves indicates that $d_1 = 1$, $z_1 = 4/3$, and $\theta_1 = 3/4$. Same quality collapses are obtained for the other cases.

The dynamics displays two different time scales mirrored in the difference of the exponents $z_{1,2}$ in regimes A and C. One time scale ($\sim N^{z_1}$) quantifies the time necessary for the unbinding (or unzipping) of the bases while the other ($\sim N^{z_2}$) gives the time needed to open (and to stretch whenever $f \neq 0$) the two chains up to their equilibrium open-end separation. At $T < T_m$ the two processes are virtually the same, because the unbinding (or unzipping) is dragged by the stretching. However, above T_m, the processes decouple and the unbinding gets faster, being controlled by the temperature, yielding $z_1 < z_2$. Furthermore, in the numerical calculations we found large sample-to-sample fluctuations, thus requiring a huge number of runs ($\sim 10^4$) to reduce statistical fluctuations in $m(t)$ and $x(t)$. This is due to the long time correlation present in the system, which keeps samples with different initial histories far apart for any t. We can explain the exponents for the Y model found numerically.

Regime *A:* $f = 0, T > T_m$: Above T_m the dominance of the entropy implies that at every time step one base pair breaks, yielding a linear behavior with $\theta_1 = 1$ and $d_1 = 1$. Also $x(t)$ tends to increase, up to its equilibrium value $\sqrt{N}$ (a fact reflected in the upward derivative at $x = 1$ of the equilibrium probability distribution). This suggests that the dynamics of this quantity is in the same universality class of the $d = 1$ Kardar-Parisi-Zhang equation [21], and so $\theta_2 = 1/3$, $d_2 = 1/2$, and $z_2 = 3/2$.

TABLE I. "Dynamic" and equilibrium exponents for the Y model (Y) and the b model with bubbles (b) as defined in Eqs. (1) and (2). The regimes A, B, C, D, and E are shown in Fig. 1.

Regime	d_1	z_1	θ_1	d_2	z_2	θ_2
A:Y	1	1	1	1/2	3/2	1/3
A:b	1	4/3	3/4	1/2	3/2	1/3
B:Y, b	1	2	1/2	1/2	2	1/4
C:Y	1	1	1	1	2	1/2
C:b	1	4/3	3/4	1	2	1/2
D:Y, b	1	2	1/2	1	2	1/2
E:Y, b	1	3	1/3	1	3	1/3

Regime *B:* $f = 0, T = T_m$: In this regime $\theta_1 = \frac{1}{2} = z_1^{-1}$ because at criticality the probabilities to increase and to decrease m are expected to be equal, so that m performs, roughly speaking, a random walk in time with reflecting boundaries at $m = 0, N$. Also for the open-end separation, steps toward larger or smaller values of x are equally probable, and therefore the time evolution of x is in the universality class of the $d = 1$ Edward-Wilkinson equation [21]. Hence, $\theta_2 = \frac{1}{4}$ and $z_2 = 2$.

Regime *C:* $f > 0, T > T_m$: The strands tend to stretch along the direction of the pulling force. However, once we have pulled the two chains up to an open-end separation x, to increase x further by one unit we first need to move all of the stretched part, which would take a time typically of order x. In other words, one has $x(t + t_0) \sim x(t) + \frac{t_0}{x}$ implying $\theta_2 = \frac{1}{2}$($\theta_1 = 1$ as before and $d_{1,2} = 1$).

Regime *D:* $T < T_m, f > f_c(T)$: Here the only microscopic mechanism for opening the fork is through the applied force: the strands must stretch completely in the vicinity of the bifurcation point and only at this point will the fork liberate one more monomer. Thus, $x \sim m$, and, using arguments as done for regime C, $\theta_{1,2} = \frac{1}{2}$, $d_{1,2} = 1$. In Ref. [3] it was found, in a mean-field approach for a model resembling our Y model, that the time necessary to unzip DNA is $\sim N^2$. This is consistent with our analysis, but works only in this regime.

Regime *E:* $T < T_m, f = f_c(T)$: On the phase boundary, one expects that the cost for unzipping and zipping is the same (the equilibrium probability distribution of having m monomers unzipped or an open-end separation equal to x is flat), so that $x(t + t_0) = x(t) \pm \frac{t_0}{x}$ with equal probability. Therefore, the open-end separation makes a random walk in the rescaled time $\frac{t}{x}$ so that $x \sim (\frac{t}{x})^{1/2}$ implying $\theta_{1,2} = \frac{1}{3}$. Moreover, $d_{1,2} = 1$ since at coexistence there is a finite fraction of liberated monomers. Another way of obtaining $\theta_1 = \frac{1}{3}$ is to demand that a kink liberated at the fork needs to diffuse out of the end before the next one is released. In other words, the rate of change of m is determined by the diffusion of a kink over a distance m. The latter time scale being of order m^{-2}, we expect $dm/dt \sim m^{-2}$ which gives $m \sim t^{1/3}$.

Turning to the b model (bubbles are allowed), the dynamical exponents in regimes B, E, and D with $T \leq T_m$ are the same as in the Y model. This states that at $T < T_m$ not only for statics, as we saw previously, but also for the dynamics, bubbles are not relevant in the scaling properties. At $T \geq T_m$, instead, the opening of bubbles heavily affects the base unpairing process, unlike in the Y model where bubbles are forbidden [23]. The length of the unzipped part now can change by $\pm l(t)$, where $l(t)$ is the typical length of bubbles, and the motion of the fork point can by no means be approximated by a simple random walk (and so θ_1 changes as shown in Table I). The quantity $x(t)$ instead has a dynamics in the b model similar to

the fork case, and indeed θ_2 is the same for both the models in all regimes. We show in Fig. 2 the collapse leading to $\theta_1 = \frac{3}{4}$ in regime A.

An important question is the dependence of our results on dimensionality (let us consider only θ_1 for simplicity). As for the Y model, the arguments we gave above for θ_1 for regimes A–E suggest that there be no d dependence. For the b model, instead, the d independence should be true only in regimes D and E for $T < T_m$, where the Y model gives the exact result; at $T \geq T_m$, on the other hand, bubbles play a dominant role, and so we expect a dependence on d. We confirmed this picture with some calculations on a simpler model which should be in the same universality class of the one under study: that of a single random walk, pinned at the origin by an attractive interaction and subject to a stretching external force. In this system m is defined as the number of monomers from the last visit to the origin to the end of the walk. For $T > T_m$, in regimes A and C, our calculations show that the exponent θ_1 increases as dimension increases, apparently with no upper critical dimension. Just at criticality at zero force, instead, we find that the exponent θ_1 is very close to $\frac{1}{2}$ in any d. The emerging picture of robust results for $T < T_m$ and model-dependent dynamics for $T \geq T_m$ would be preserved even if, in the original models, the directedness constraint is relaxed. Our models can be extended to include sequence heterogeneity though a distinction between a single realization and an ensemble averaged case needs to be made. We, however, expect the reentrance and the existence of dynamic scaling at least in the ensemble-averaged case. As sequence heterogeneity controls the strength of the binding of base pairs, its role becomes less crucial as T increases so we expect that for high enough T also the values of the dynamic exponents of regimes A and C will be recovered.

The observed power-law and scaling behaviors in Eqs. (1) and (2) imply that in our model there is scale invariance. As it is generally accepted [21], this could be due to the presence of strong temperature driven fluctuations or of long range time correlations (necessary to build up a cooperative mechanism) in the unzipping. In a replication process, it is expected that fluctuations in unzipping, after its initiation at the origin, would affect the binding of the next set of proteins. Though the nature of this coupling is not known, we note that the difference in the dynamical scaling at the thermal denaturation point and on the unzipping phase boundary can in principle provide a dynamic selection mechanism for binding of proteins in the Y fork. However, at this point, this is speculative.

In conclusion, based on the exact phase diagrams of the Y model (without bubbles) and the b model (with bubbles) for the force-induced unzipping transition (and denaturation), we investigated the dynamics of unzipping from a nonequilibrium bound state both on and away from the phase boundary. The dynamics shows scaling behaviors in different regimes of the phase diagram. These scalings in most cases could be understood from the plausible mechanisms of unzipping and denaturation as discussed above, except for a quantitative understanding of the scaling in the high T region for the b model. Lastly, the unzipping dynamics on the phase boundary in the presence of a force is distinctly different from the thermal denaturation at zero force. Whether a real biological system takes advantage of these differences to distinguish the unzipped region of DNA from a fluctuation-induced bubble formation remains to be probed.

This work was supported by MURST(COFIN 99).

*Email address: somen@iopb.res.in

[1] A. Kornberg and T. Baker, *DNA Replication* (W. H. Freeman, San Francisco, 1992), 3rd ed.
[2] S. M. Bhattacharjee, J. Phys. A **33**, L423 (2000); **33**, 9003(E) (2000); cond-mat/0010132.
[3] K. L. Sebastian, Phys. Rev. E **62**, 1128 (2000).
[4] D. K. Lubensky and D. R. Nelson, Phys. Rev. Lett. **85**, 1572 (2000).
[5] H. Zhou, cond-mat/0007015.
[6] D. Marenduzzo, A. Trovato, and A. Maritan, Phys. Rev. E **64**, 031901 (2001).
[7] E. Boye, T. Stokke, N. Kleckner, and K. Skarstad, Proc. Natl. Acad. Sci. U.S.A. **93**, 12 206 (1996); W. M. Xia and W. Dowhan, Proc. Natl. Acad. Sci. U.S.A. **92**, 783 (1995).
[8] J. Marx, Science **270**, 1585 (1995); **270**, 1667 (1995); **270**, 1671 (1995); **270**, 1674 (1995).
[9] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot, Proc. Natl. Acad. Sci. U.S.A. **94**, 11 935 (1997).
[10] M. Reif, H. Clausen-Schaumann, and H. E. Gaub, Nature Struc. Biol. **6**, 346 (1999).
[11] D. Anselmetti, J. Fritz, B. Smith, and X. Fernandez-Busquets, Single Mol. **1**, 53 (2000).
[12] T. Strunz, K. Oroszlan, R. Schafer, and H. H. Güntherodt, Proc. Natl. Acad. Sci. U.S.A. **96**, 11 277 (1999).
[13] B. Maier, D. Bensimon, and V. Croquette, Proc. Natl. Acad. Sci. U.S.A. **97**, 12 002 (2000).
[14] G. V. Shivashankar, M. Feingold, O. Krichevsky, and A. Libchaber, Proc. Natl. Acad. Sci. U.S.A. **96**, 7916 (1999).
[15] G. J. L. Wuite, S. B. Smith, M. Young, D. Keller, and C. Bustamante, Nature (London) **404**, 103 (2000).
[16] D. Poland and H. Scheraga, J. Chem. Phys. **45**, 1464 (1966).
[17] M. S. Causo, B. Coluzzi, and P. Grassberger, Phys. Rev. E **62**, 3958 (2000); Y. Kafri, D. Mukamel, and L. Peliti, Phys. Rev. Lett. **85**, 1572 (2000).
[18] R. J. Rubin, J. Chem. Phys. **43**, 2392 (1965).
[19] S. Lifson, J. Chem. Phys. **40**, 3705 (1964).
[20] A rigorous treatment can also be given that supports our claims (details will be presented elsewhere).
[21] See, e.g., A. L. Barabasi and H. E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, U.K., 1995).
[22] S. M. Bhattacharjee and F. Seno, J. Phys. A **34**, 6375 (2001).
[23] Regime B is a borderline case: bubbles heavily affect the evaluation of m but do not change θ_1.