Dynamics of DNA melting

2009 J. Phys.: Condens. Matter 21 034110

(http://iopscience.iop.org/0953-8984/21/3/034110)

J. Phys.: Condens. Matter **21** (2009) 034110 (9pp) doi:10.1088/0953-8984/21/3/034110

Dynamics of DNA melting

A Bar[1], Y Kafri[2] and D Mukamel[1]

[1] Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel
[2] Department of Physics, Technion, Haifa 32000, Israel

Received 2 July 2008, in final form 23 November 2008
Published 17 December 2008
Online at stacks.iop.org/JPhysCM/21/034110

Abstract
The dynamics of loops at the DNA denaturation transition is studied. A scaling argument is used to evaluate the asymptotic behavior of the autocorrelation function of the state of complementary bases (either open or closed). The long-time asymptotic behavior of the autocorrelation function is expressed in terms of the entropy exponent, c, of a loop. The validity of the scaling argument is tested using a microscopic model of an isolated loop and a toy model of interacting loops. This suggests a method for measuring the entropy exponent using single-molecule experiments such as fluorescence correlation spectroscopy.

1. Introduction

Melting, or thermal denaturation of DNA, is the process by which the two strands of the DNA molecule become fully separated upon an increase in temperature [1, 2]. At low temperatures the strands are partially unbound by forming fluctuating loops where the two strands are locally separated. As the melting temperature T_M is approached the average loop size increases, yielding full denaturation at T_M. Melting of DNA has been extensively studied over the years both theoretically and experimentally. The natural order parameter of the denaturation transition is the fraction of bound base-pairs. This was measured using specific-heat and UV absorption experiments. Simple models have yielded theoretical expressions for thermodynamic properties. Two main approaches have been developed. One, known as the Peyrard–Bishop model, considers the two strands as directed polymers interacting via a short-ranged potential [3]. One then focuses on the distance between complementary pairs as the melting transition is approached. The other, known as the Poland–Scheraga (PS) model, represents the DNA molecule as an alternating sequence of bound segments and open loops and focuses on the fraction of bound base-pairs [4]. Within the PS approach self-avoiding interactions, which are inherently long-range, may be taken into account. As has been shown these interactions affect the loop entropy, which controls the nature of the melting transition [5, 6].

Recently, single-molecule techniques such as optical tweezers [7, 8], magnetic traps [9, 10] and fluorescence correlation spectroscopy (FCS) [11–13] have been used to probe properties of the melting process. Other techniques, such as quenching, have also been applied [14]. Some of the experiments utilize an external force to induce unzipping of the two strands and study their dynamics. In others, the distance between two complementary base-pairs is probed by FCS without applying an external force. These experimental methods enable one to study not only bulk properties but rather microscopically fluctuating quantities. Inspired by these experiments, theoretical treatments of dynamical properties of DNA have been developed. Numerous studies have focused on the effective dynamics of the ends of an isolated loop, ignoring the detailed dynamics of the two single strands which compose the loop [12, 13, 15–17]. Recently, a model which incorporates the dynamics of the single strands has been suggested and the survival probability of an isolated loop has been calculated [18]. A toy model for the dynamics of interacting loops has also been introduced and analyzed [19].

It has recently been shown that studying the loop dynamics may yield information on the loop entropy [18]. Within the PS approach the dependence of the entropy of a loop on its length plays a dominant role in determining the thermodynamic behavior near the transition. On general grounds one can argue that the entropy of a loop of length n takes the form $S = k_B \log(\Omega(n))$, where $\Omega(n) \sim s^n/n^c$ is the number of loop configurations. Here s is a model-dependent constant and c is a universal exponent. The numerical value of c has been debated over the years. It was found to be modified when the excluded-volume interactions, which are long-ranged in nature, are taken into account [4–6]. When interactions between loops are neglected, and excluded-volume interactions are taken into account only within each loop an exponent $c \simeq 1.76$ was found [5]. On the other hand, when excluded-volume interactions both within a loop and between the loop and the rest of the chain are taken into account, the entropy exponent was found to increase to $c \simeq 2.12$ [2, 6]. This latter result, which predicts a first-order denaturation transition for

homopolymers, has been verified numerically [21]. Note that, when mismatches between the two strands are allowed to take place, the loop exponent c becomes $c-1$ and the order of the transition may change accordingly. While numerical studies of the homopolymer model with excluded-volume interactions yield a clear first-order transition [22], a direct experimental measurement of c is rather difficult and has not been carried out so far. In [18] it was shown that at the melting transition the time dependence of the base-pair autocorrelation function depends on the parameter c. The base-pair autocorrelation function is defined as $C_i(t) = \langle u_i(t+\tau)u_i(\tau)\rangle$, where $u_i(t) = 1, 0$ is a variable which indicates if base-pair i is open (1) or closed (0) at time t and $\langle\cdot\rangle$ denotes an average over τ. The behavior of the autocorrelation function was studied theoretically away from the melting transition [15, 16] and may, in principle, be obtained experimentally by FCS studies. In these experiments the states of a specific base-pair is monitored. So far, FCS experiments have been restricted to short molecules [11]. Measuring the exponent c requires extending these studies to longer molecules.

In the present paper we elaborate on and extend the analysis presented in [18] for the dynamical behavior of homopolymers at the melting transition. The dynamics of a single loop is studied using a simple model, whose validity is then verified in detail using numerical simulations. Within this model the entropy exponent c is introduced as a free parameter which may be chosen at will. While the studies in [18] were tested numerically only for the case $c = 3/2$, here we test the robustness of the results for models with arbitrary values of c. We then consider a toy model, similar to the one considered in [19], which indicated that the results still hold when the interaction between loops is taken into account.

This paper is organized as follows: in section 2 we study the single-loop model using both the scaling argument and microscopic models. In section 3 results for the many-loops model are presented. Finally, we end with a brief summary.

2. Single-loop dynamics

We start by considering the dynamics of an isolated loop. In this approach one ignores processes like merging of loops and the splitting of a large loop into two or more smaller ones. This may be justified by the fact that the cooperativity parameter, which controls the statistical weight of opening a new loop, is estimated to be rather small, $\sigma_0 \approx 10^{-4}$ [20]. Thus splitting a loop into two is unfavorable. Also, the average distance between loops, which within the PS model is proportional to $1/\sigma_0$, is large, making the independent loop approximation plausible. In section 3 we introduce a simple model to effectively take into account the interactions between loops and show that these interactions do not modify the results obtained asymptotically within the single-loop approach.

Within the single-loop dynamics, we assume that a loop may change its length by closing or opening of base-pairs at its two ends. It survives as long as its two ends do not meet. Let $G(n_0, t)$ be the survival probability of a loop of initial length n_0 at time t. As discussed above, the quantity of interest is the equilibrium autocorrelation function

$$C(t) \approx \frac{\sum_{n_0=1}^{\infty} P_{\rm eq}(n_0) n_0 G(n_0, t)}{\sum_{n_0=1}^{\infty} P_{\rm eq}(n_0) n_0}, \tag{1}$$

where for simplicity of notation we have dropped the site index i. Here $P_{\rm eq}(n_0)$ is the probability of having a loop of length n_0 in equilibrium. Hence, $n_0 P_{\rm eq}(n_0)$ is the probability of a particular site to belong to a loop of length n_0. Note that we assume that site i remains open as long as the loop survives. This approximation does not affect the behavior of the autocorrelation function in the scaling limit.

We proceed by first presenting a scaling analysis demonstrating that, in the case of a homopolymer and at criticality, the autocorrelation function decays at large t as $C(t) \sim t^{1-c/2}$ for $c > 2$, while it remains finite, $C(t) = 1$, for $c < 2$. These results are then tested and verified using numerical simulations for various values of c.

2.1. Scaling analysis

In the case of a homopolymer and at criticality it has been shown that the equilibrium loop size distribution is $P_{\rm eq}(n) \sim 1/n^c$. To estimate the survival probability of a loop of length n_0, we consider dynamics under which the loops are non-interacting and do not split into a number of smaller loops. Similar to [15, 16] we further assume that the loop is in a local thermal equilibrium at any given time during its evolution. The validity of this assumption will be discussed in detail below. The loop free energy is thus given by $f \propto c \ln n$. Within the framework of the Fokker–Planck equation, the probability distribution of finding a loop of size n at time t, $P(n, t)$, is given by

$$\frac{{\rm d}P(n,t)}{{\rm d}t} = D\frac{\partial}{\partial n}\left[\frac{c}{n} + \frac{\partial}{\partial n}\right] P(n,t), \tag{2}$$

where D is the diffusion constant. Here we have taken the continuum limit and assumed the dynamics to be over-damped. This equation has to be solved with the boundary condition $P(0, t) = 0$ and initial condition $P(n, 0) = \delta(n - n_0)$. The survival probability of the loop is then given by $G(n_0, t) = \int_0^{\infty} {\rm d}n P(n, t)$.

Within the scaling approach the survival probability is written in the form

$$G(n_0, t) = g(Dt/n_0^z), \tag{3}$$

with $z = 2$. In the appendix we show that the asymptotic behavior of the scaling function for small and large values of the argument is

$$g(x) \sim 1 \qquad \text{for } x \ll 1 \tag{4}$$

$$g(x) \sim x^{-\frac{1+c}{2}} \qquad \text{for } x \gg 1. \tag{5}$$

The autocorrelation function (equation (1)) may thus be written as

$$C(t) \approx \frac{\int_1^N n_0^{1-c} g(Dt/n_0^2)\,{\rm d}n_0}{\int_1^N n_0^{1-c}\,{\rm d}n_0}, \tag{6}$$

where the system size N is taken to infinity in the thermodynamic limit. We first consider the long-time behavior for $c \leqslant 2$. In this case the integrals are controlled by the upper limit N, where $g(Dt/n_0^2) \sim 1$. Both numerator and denominator diverge as N^{2-c} so that $C(t) \sim 1$ for $t \gg 1$. On the other hand, for $c > 2$ both integrals are independent of the upper limit. Changing variables to $y = n_0/\sqrt{Dt}$ yields

$$C(t) \approx (Dt)^{1-c/2} \frac{1}{\langle n_0 \rangle} \int_{1/\sqrt{Dt}}^{\infty} y^{1-c} g(y^{-2})\, \mathrm{d}y, \qquad (7)$$

where $\langle n_0 \rangle$ is the average loop size. The asymptotic behavior of $g(y^{-2})$ at small y (equation (5)) implies that the integral converges for $t \to \infty$, yielding $C(t) \sim t^{1-c/2}$. Hence

$$C(t) \sim \begin{cases} 1 & \text{for } c \leqslant 2 \\ t^{1-c/2} & \text{for } c > 2. \end{cases} \qquad (8)$$

This expression suggests that measuring $C(t)$ at criticality may be used to determine the entropy exponent c. In particular it can be used to distinguish between the case of a continuous transition ($c \leqslant 2$), where $C(t) = 1$, and a first-order phase transition ($c > 2$), where $C(t)$ decays to zero at long times.

In the above analysis it is assumed that the loop is at local equilibrium at any given time. For this assumption to be valid, the survival time of large loops has to be much longer than its equilibration time. A typical survival time of a loop of length n scales as n^2. On the other hand, simple models for the dynamics of microscopic loop configurations, which are usually based on diffusion processes, yield relaxation times which also scale as n^2. Thus the two typical time scale in the same way with the loop size and it is not *a priori* clear that, during the evolution of the loop, it is at local equilibrium. Note that off-criticality the loop size changes linearly in time and therefore the assumption of local equilibrium is clearly not valid.

In the following we introduce and study a model for the loop dynamics. This is done in two steps: first, we consider the simpler case of $c = 3/2$ discussed in [18]. We then generalize this approach to arbitrary values of c. We find strong evidence that the local equilibrium assumption holds asymptotically for the model. It is thus argued that within the model the local equilibrium assumption is valid.

2.2. Microscopic dynamical model for $c = 3/2$

In this section we introduce and analyze a simple model of loop dynamics corresponding to $c = 3/2$. Within the model, the loop is described by a fluctuating interface (or a string), interacting with an attractive substrate in $d = 1+1$ dimensions. Here the interface height variable corresponds to the distance between complementary bases. The interface configurations are those of a restricted solid-on-solid (RSOS) model defined as follows (see figure 1): let $h_i = 0, 1, 2 \ldots$ be the height of the interface at site i. The heights satisfy $|h_i - h_{i+1}| = \pm 1$. Consider a loop between sites 0 and n (where n is even) as shown in figure 1. Outside the loop the interface is bound to the substrate so that $h_{-2k} = h_{n+2k} = 0$ and $h_{-2k-1} = h_{n+2k+1} = 1$ for $k = 0, 1, \ldots$. Inside the loop the heights $h_1 \cdots h_{n-1}$ can take any non-negative value which is consistent with the RSOS conditions. For simplicity we allow only one end of the loop to fluctuate while the other is held fixed. This should not modify any of our results, since the dynamics of the two ends of long loops are uncorrelated with each other. We consider a random sequential dynamics in which the loop configuration and its length are free to fluctuate. Thus the dynamical moves are as follows:

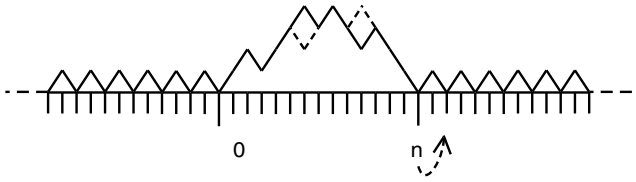

Figure 1. A typical microscopic configuration of the loop in the RSOS model. Dashed lines indicate possible dynamical moves of the interface.

$$h_i \to h_i \pm 2 \qquad \text{with rate 1 for sites } 1 \leqslant i \leqslant n-1, \qquad (9)$$

as long as the resulting heights are non-negative and the RSOS condition is satisfied. For $i = n$ the loop length is changed according to the rules

$$\begin{aligned} n &\to n+2 \qquad \text{with rate } \overline{\alpha}/4 \\ n &\to n-2 \qquad \text{with rate } \overline{\alpha}, \end{aligned} \qquad (10)$$

where n can decrease only if $h_{n-2} = 0$. At the other end the height is fixed, $h_0 = 0$. It is straightforward to verify that the number of configurations of a loop of size n is given by $2^n/n^c$ with $c = 3/2$ for large n. This is a result of the fact that the number of walks of length n in $d = 1 + 1$ dimensions is 2^n and the probability of first return is $n^{-3/2}$. The ratio, $1/4$, between the two length-changing processes in equation (10) is chosen such that in the large n limit the loop is not biased to either grow or shrink. This corresponds to the model being at the denaturation transition point, which is determined by equating the free energies of the pinned segment and that of the open loop. Combining this with detailed balance yields the ratio between the rates. The parameter $\overline{\alpha}$ determines the rate of the length changing processes: $\overline{\alpha} = 0$ corresponds to the dynamics of a loop of a fixed length. As $\overline{\alpha}$ is increased the length-changing processes become faster. In section 2.3 this model is generalized to include a power-law potential between the interface and the substrate. This will allow us to study other values of c.

In a realization of this dynamics one of the $n + 1$ attempts defined above, equations (9) and (10), is chosen at any given time. Of these, $n - 1$ are attempts to update the height at sites $1, 2, \ldots, n - 1$. The other two are attempts to update the position of the edge by a move either to the right or to the left. One attempted move of the edge defines a Monte Carlo sweep.

In order to test the validity of equation (2) we compare its predictions with results obtained from numerical simulations of the model above. In the numerical simulation we find good data collapse, when plotted against t/n_0^z with $z \gtrsim 2$,

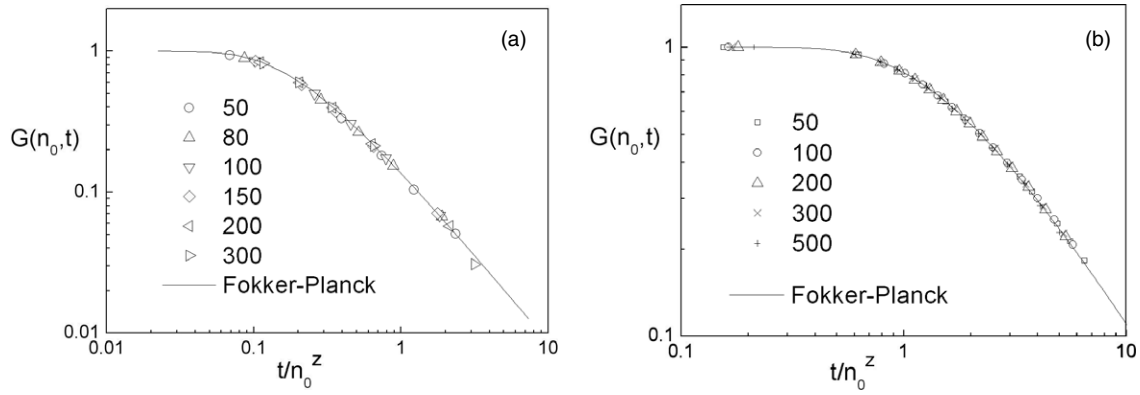

Figure 2. Data collapse of the survival probability (averaged over 4×10^4 realizations) for some values of n_0 with (a) $\overline{\alpha} = 1$ and $z = 2.2$, and (b) $\overline{\alpha} = 0.1$ and $z = 2.07$. The line corresponds to a numerical solution of equation (2).

depending on the value of $\overline{\alpha}$, rather than the expected Fokker–Planck value $z = 2$. With these modified z exponents the survival probability agrees well with the results obtained from the discrete version of the Fokker–Planck equation. The results are summarized in figure 2 where the survival probability is plotted as a function of the scaling variable $t/n_0^{2.2}$ and $t/n_0^{2.07}$ for $\overline{\alpha} = 1$ and $\overline{\alpha} = 0.1$, respectively, for several values of the loop size n_0. The question is whether the discrepancy in the value of z is a result of a finite size effect or does it persist in the large n_0 limit. For the Fokker–Planck equation to properly describe the system it is essential to show that z approaches 2 in the large n_0 and t limit.

In the following we argue that, in fact, the value $z = 2.2$ in the case of $\overline{\alpha} = 1$ (and $z = 2.07$ for $\overline{\alpha} = 0.1$) is a result of finite size effects. For large systems the value $z = 2$ is expected to be recovered. To check this point we calculate numerically the variance of the loop size:

$$w^2(t) = \langle (n(t) - \langle n(t) \rangle)^2 \rangle, \tag{11}$$

where $\langle \cdot \rangle$ denotes an average over realizations of the dynamics. In order to evaluate the temporal growth of $w^2(t)$ we define a variable $\sigma_+(t)$ which takes the value 1 if the length of the loop increases at time t and 0 otherwise. Similarly, we define $\sigma_-(t)$ and $\sigma_0(t)$ for steps which decrease the loop size and steps in which the loop size does not change, respectively. Clearly $\sigma_+(t) + \sigma_-(t) + \sigma_0(t) = 1$. The dynamics of the chain, equation (10), implies that in the limit of large n_0 one has

$$\langle \sigma_+(t) \rangle = \langle \sigma_-(t) \rangle = \alpha/8; \qquad \langle \sigma_0(t) \rangle = 1 - \alpha/4, \tag{12}$$

where $\alpha = \overline{\alpha}/\max\{1, \overline{\alpha}\}$ in accordance with the random sequential dynamics. Denoting $U(t) \equiv \sigma_+(t) - \sigma_-(t)$, it is easy to see that

$$\frac{\Delta w^2(t)}{\Delta t} \equiv w^2(t) - w^2(t-1)$$
$$= 4\langle U(t)^2 \rangle + 8 \sum_{\tau=1}^{t-1} \langle U(\tau) U(t) \rangle, \tag{13}$$

where

$$\langle U(\tau)U(t) \rangle = \langle \sigma_+(\tau)\sigma_+(t) \rangle + \langle \sigma_-(\tau)\sigma_-(t) \rangle$$
$$- \langle \sigma_-(\tau)\sigma_+(t) \rangle - \langle \sigma_+(\tau)\sigma_-(t) \rangle. \tag{14}$$

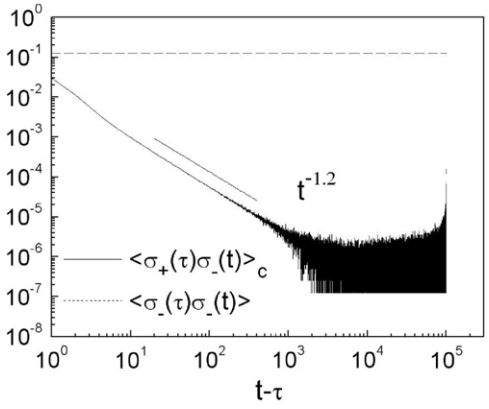

Figure 3. Correlation functions of the σ variables as obtained by averaging over 1.9×10^5 realizations, for $n_0 = 4000$.

It is evident that a loop-increasing step at time t, ($\sigma_+(t) = 1$), is uncorrelated with steps which took place at time $\tau < t$. Thus $\langle \sigma_+(\tau)\sigma_+(t) \rangle = \langle \sigma_-(\tau)\sigma_+(t) \rangle = \alpha^2/64$. Numerically we find $\langle \sigma_-(\tau)\sigma_-(t) \rangle = \alpha^2/64$ (see figure 3). Using these results we finally obtain

$$\frac{\Delta w^2(t)}{\Delta t} = \alpha - 8 \sum_{\tau=1}^{t-1} \left[\langle \sigma_+(\tau)\sigma_-(t) \rangle_c \right], \tag{15}$$

with $\langle \sigma_+(\tau)\sigma_-(t) \rangle_c \equiv \langle \sigma_+(\tau)\sigma_-(t) \rangle - \alpha^2/64$. Numerical simulations of the dynamics show strong correlation between $\sigma_+(\tau)$ and $\sigma_-(t)$ with an algebraic decay in $t - \tau$ (see figure 3). It is interesting to note that the dynamics of the chain induces such long-range temporal correlations between steps of the edge mediated by the loop dynamics.

By extrapolating the sum on the right-hand side of equation (15) using the asymptotic form $B(t - \tau)^{-\gamma}$ with $B \approx 0.015$ and $\gamma \approx 1.2$, deduced from figure 3, we find that the sum converges to a non-zero value. This is demonstrated in figure 4 for $\overline{\alpha} = 1$ and $\overline{\alpha} = 0.1$. For example, in the case $\overline{\alpha} = 1$ the sum converges to $\approx 0.84 < \alpha = 1$, indicating that $w^2(t) \approx 0.16t$ at large t, which in turn yields $z = 2$. The slow power-law convergence towards the asymptotic value implies that it may require large systems to observe the long-time behavior of equation (8).

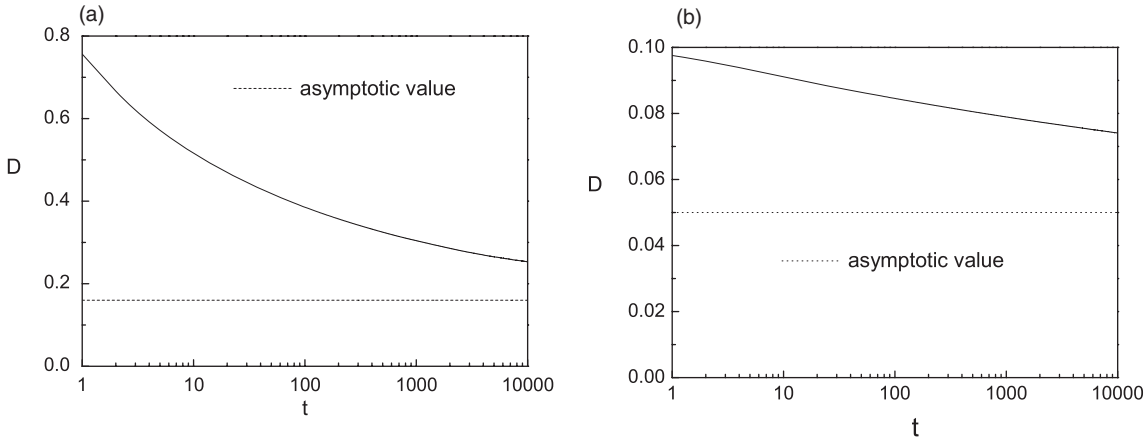

Figure 4. The diffusion coefficient $D \equiv \Delta w^2(t)/\Delta t$, as calculated by equation (15), averaged over 270 000 runs with $n_0 = 2000$ for (a) $\overline{\alpha} = 1$ and (b) $\overline{\alpha} = 0.1$. The slow decay of D can be easily observed.

2.3. Microscopic dynamical model for arbitrary c

In this section we generalize the model of section 2.2 to consider the case of arbitrary c. This can be done within a $(d = 1 + 1)$-dimensional model by introducing a repulsive interaction between the substrate and the interface. Taking an interaction of the form A/h^2, where h is the distance between the interface and the substrate and A is a constant, results in an equilibrium weight of a loop of the form $1/n^c$. The exponent c is related to the interaction strength A [23, 24].

In order to derive the relation between A and c one notes that at the critical point the distribution of the distance between the interface and the substrate decays algebraically at large distances, $Q(h) \sim 1/h^{\kappa}$. It has been shown that, for an interface model for which self-avoiding interactions play no role, the exponent κ is related to the loop exponent c by [25]

$$c = (\kappa + 3)/2. \tag{16}$$

We proceed by introducing a specific model and evaluate κ in terms of the interaction parameter A. One then obtains the loop exponent c from equation (16). We consider an RSOS interface model with the Hamiltonian

$$H(h_1, h_2, \ldots, h_n) = \sum_i \left[-\varepsilon \delta_{h_i,0} + \frac{A}{h_i^2}(1 - \delta_{h_i,0}) \right], \tag{17}$$

where, as before, $h_i = 0, 1, 2 \ldots$ and $h_i - h_{i+1} = \pm 1$. In this Hamiltonian $\varepsilon > 0$ represents the binding energy between the substrate and the interface. To evaluate $Q(h)$ we write down the eigenvalue equation of the transfer matrix corresponding to the Hamiltonian equation (17). For $h > 1$ the equation is

$$\mathrm{e}^{-\beta A/h^2}\Psi_{h-1} + \mathrm{e}^{-\beta A/h^2}\Psi_{h+1} = \lambda \Psi_h. \tag{18}$$

Here λ is the eigenvalue and Ψ_h are the components of the eigenvector. The distance distribution is given by $Q(h) \propto \Psi_h^2$. At criticality the eigenvector component, at large h, has a form $\Psi(h) = 1/h^{\kappa/2}$. By using this form in equation (18) we find the relation

$$\beta A = \tfrac{1}{8}\kappa(\kappa + 2). \tag{19}$$

Combining this with equation (16) yields

$$\beta A = \tfrac{1}{8}(2c - 3)(2c - 1). \tag{20}$$

We now use the model, equation (17), to study numerically the dynamics of a loop with $c \neq 3/2$. The dynamics of the model is similar to that introduced in section 2.2, but with the transition rates of equation (9) modified according to the Hamiltonian equation (17). Namely, the updating rates are given by

$$\begin{aligned} h_i &\to h_i + 2 \qquad \text{with rate } 1 \\ h_i &\to h_i - 2 \qquad \text{with rate } \mathrm{e}^{-\beta A\left((h_i-2)^{-2} - h_i^{-2}\right)}, \end{aligned} \tag{21}$$

as long as the resulting heights are non-negative and the RSOS condition is satisfied. The presence of the long-range interactions also changes the ratio between the rates by which the loop grows ($R(n \to n + 2)$) and shrinks ($R(n + 2 \to n)$). This ratio is given by

$$\frac{R(n \to n + 2)}{R(n + 2 \to n)} = \mathrm{e}^{-\beta\varepsilon}. \tag{22}$$

The critical temperature is found by equating the free energy of the loop with that of the bound segment. This yields

$$4 = \mathrm{e}^{-\beta(A-\varepsilon)}, \tag{23}$$

where $A - \varepsilon$ is the energy of a pair of sites in the bound segment and 4 is the statistical weight of a pair of sites in the open loop. Combining this with equation (22) gives

$$\begin{aligned} n &\to n + 2 \qquad \text{with rate } \overline{\alpha}\mathrm{e}^{-\beta A}/4 \\ n &\to n - 2 \qquad \text{with rate } \overline{\alpha}. \end{aligned} \tag{24}$$

We have simulated the dynamics of equations (21) and (24) for $A = 0.15$ and 0.5. These values of A correspond to $c \approx 1.74(<2)$ and $c \approx 2.12(>2)$, respectively. In figure 5 we present the loop size distribution for these two values of the parameter A. The resulting c values fit well with the predictions.

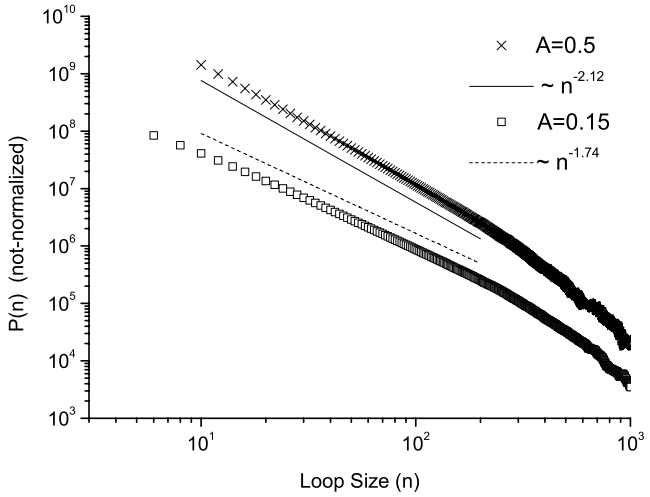

Figure 5. The loop size distribution for $A = 0.15$ and $A = 0.5$ as measured in numerical simulations, and theoretical resulting exponents $c \approx 1.74$ and $c \approx 2.12$, respectively. The theoretical curves show a good fit with the measured data.

In studying the survival probability of a loop we follow the same approach which was applied in section 2.2 for $c = 3/2$. Similar results were obtained for the case of $c > 3/2$. In figure 6 we present the survival probability as obtained from numerical simulations of the model. We find good data collapse, but again with a modified exponent $z = 2.2$ for $\overline{\alpha} = 1$. The scaling function fits well with that obtained from a numerical integration of a discrete version of equation (2).

We have also calculated the step–step autocorrelation function as for the case $c = 3/2$ and found similar results. In particular, we find that the exponent γ seems to have a weak dependence on A, with $\gamma \sim 1.4$ for both $A = 0.15$ and $A = 0.5$ (figures not shown).

3. Many-loops model

In section 2 we analyzed the dynamics of a single loop. We found that it is well described by the Fokker–Planck equation (2) for asymptotically large loops. In the present section we extend this model to consider interaction between loops. This is done by considering a chain composed of an alternating series of loops and bound segments. Each loop and bound segment is characterized only by their respective lengths. In contrast to the study of the dynamics of a single loop here no internal degrees of freedom are associated with a loop. Within this model loops evolve by growing, shrinking, splitting and merging, together with creation and annihilation processes. The rates of the various processes are chosen so that the system evolves to the equilibrium loop length distribution at large times. While the choice of rates is not unique they are taken to be compatible with the single-loop dynamics whenever applicable. A similar approach has recently been applied to study dynamical features such as the approach to equilibrium near the denaturation transition [19]. From this analysis we extract the behavior of the autocorrelation function of a base-pair inside a dsDNA where many interacting loops coexist.

3.1. Definition of the model

The DNA configurations can be represented by an alternating sequence of bound base-pairs and loops. We denote by $[k]$ a bound segment with length k and (l) a loop of length l, with $k, l > 0$. A given configuration of the DNA is thus represented by $[k_1](l_1)[k_2](l_2)\ldots$. In terms of these variables the dynamics of the model is defined by the following rates:

- Motion of a loop edge. This corresponds to the same processes which were considered in the dynamics of an isolated loop in section 2:

$$\begin{aligned} [k](l) &\to [k-1](l+1) && \text{with rate } \left(\frac{l}{l+1}\right)^c \\ [k-1](l+1) &\to [k](l) && \text{with rate 1} \\ (l)[k] &\to (l+1)[k-1] && \text{with rate } \left(\frac{l}{l+1}\right)^c \\ (l+1)[k-1] &\to (l)[k] && \text{with rate 1.} \end{aligned} \tag{25}$$

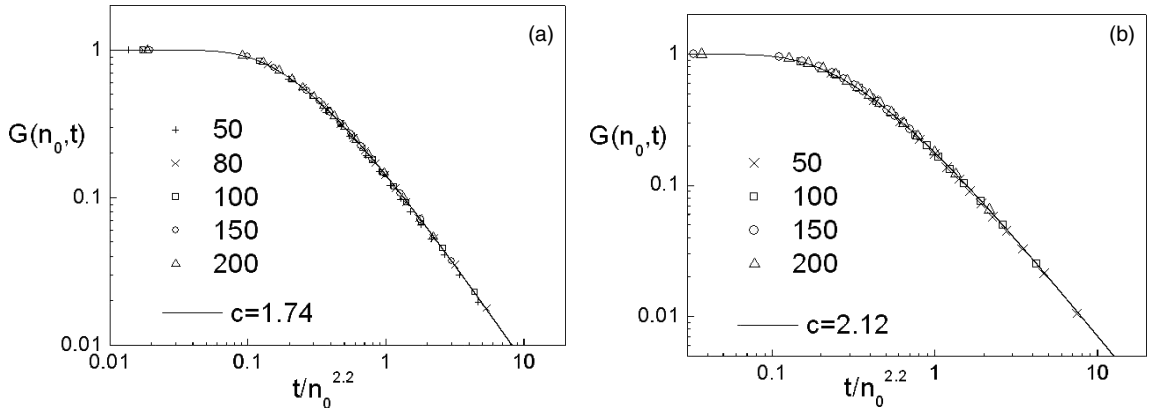

Figure 6. Data collapse of the survival probability (averaged over 4×10^5 realizations) for some values of n_0 with $z = 2.2$, with (a) $c = 1.74$ ($A = 0.15$) and (b) $c = 2.12$ ($A = 0.5$). The line corresponds to a numerical solution of a discrete version of equation (2) with corresponding values of c.

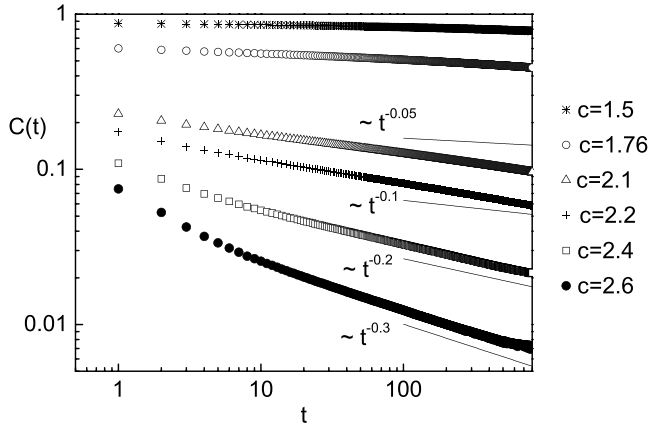

Figure 7. Normalized autocorrelation functions as measured in the simulation of the many-loops model, for $\sigma_0 = 0.0001$, $L = 100\,000$ and 50 000 repetitions. The thin lines indicate the expected behavior of the autocorrelations for the appropriate values of c.

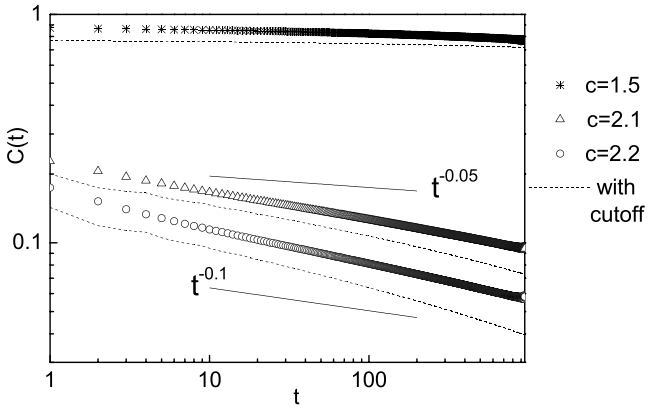

Figure 8. The same as figure 7, with the numerical calculations of the sum with a cutoff.

These processes are executed as long as the lengths of the resulting loops and bound segments are non-zero.

- Splitting and merging of loops:

$$\begin{aligned}(l_1 + l_2 + 1) &\to (l_1)[1](l_2)\\ &\text{with rate } \frac{\sigma_0}{\zeta(c)}\left(\frac{l_1 + l_2 + 1}{l_1 l_2}\right)^c \qquad (26)\\ (l_1)[1](l_2) &\to (l_1 + l_2 + 1) \qquad \text{with rate 1.}\end{aligned}$$

In addition we consider creation and annihilation of loops.

- Creation and annihilation of loops:

$$\begin{aligned}[k_1 + k_2 + 1] &\to [k_1](1)[k_2] \qquad \text{with rate } \frac{\sigma_0}{\zeta(c)(1-\sigma_0)}\\ [k_1](1)[k_2] &\to [k_1 + k_2 + 1] \qquad \text{with rate 1.}\end{aligned} \qquad (27)$$

Here σ_0 is the cooperativity parameter and $\zeta(c) = \sum_{n=1}^{\infty} n^{-c}$. It is straightforward to verify that the choice of rates satisfies detailed balance with respect to the equilibrium weight for the loop sizes at criticality $P(n) = \sigma_0 \frac{n^{-c}}{\zeta(c)}$.

3.2. Numerical simulation

To check that indeed interactions between loops do not modify the asymptotic behavior of the autocorrelation function $C(t)$ we simulate the model equations (25)–(27). We use the experimental relevant value $\sigma_0 = 10^{-4}$ [20] and consider a DNA length of 100 000 base-pairs. The autocorrelation is evaluated by monitoring the state of 1000 base-pairs uniformly distributed within the DNA. Figure 7 shows the results for various values of c along with the theoretically expected slopes.

While the results for large values of c agree well with the theory, there is a systematic deviation from the predicted slopes for smaller values of c close to 2. These deviations could be attributed to the finite length of the simulated system. For example, it is clear that, for $c < 2$, the autocorrelation function of a finite system decays to zero at long times rather than remaining constant. This is due to the fact that there is an upper cutoff on the loop size available. Only for an infinite system is $C(t)$ expected to remain constant (=1) at long times. In order to check this point we introduce an upper cutoff $N_{\max}$ to the loop size in the equation for the autocorrelation function:

$$C(t) \approx \frac{\sum_{n_0=1}^{N_{\max}} P_{\rm eq}(n_0) n_0 G(n_0, t)}{\sum_{n_0=1}^{N_{\max}} P_{\rm eq}(n_0) n_0}. \qquad (28)$$

The loop size $N_{\max}$ is chosen so that it appears roughly one time during a run. For runs which are not too long this can be estimated using $\sigma_0 L R P(N_{\max}) = 1$, where L is the system size and R is the number of Monte Carlo repetitions performed. In figure 8 the results of the simulations are compared with the theoretical expression, equation (28), which is summed numerically.

In summary, we find that our scaling predictions are generally confirmed by the numerical simulations of the many-loops model. However, for values of c close to 2, deviations are found. These seems to be related to finite size effects.

4. Conclusions

In this paper the dynamics of loops at the denaturation transition was studied both within a single-loop model and a many-loops approach. In particular, special care was given to the applicability of the Fokker–Planck equation. It was shown that the long-time decay of the autocorrelation function of the state of complementary bases (closed or open) is sensitive to the value of the loop exponent. In particular, for $c < 2$ it remains finite while for $c > 2$ it decays as $t^{1-c/2}$.

Throughout this paper we have considered homopolymers where the binding energy between different base-pairs is constant. In typical DNA molecules the binding energy is not homogeneous. While a preliminary treatment of the effects of disorder was given in [18], it remains an important and interesting question.

Acknowledgments

The support of the Israeli Science Foundation (ISF) and the Albert Einstein Minerva Center for Theoretical Physics is

gratefully acknowledged. YK also acknowledges support by the US–Israel Binational Science Foundation (BSF).

Appendix. Asymptotic behavior of the return probability

In this appendix we derive the asymptotic behavior of the survival probability corresponding to the Fokker–Planck equation:

$$\frac{\mathrm{d}P(n,t)}{\mathrm{d}t} = D\frac{\partial}{\partial n}\left[\frac{c}{n} + \frac{\partial}{\partial n}\right]P(n,t) \tag{29}$$

with the boundary conditions

$$P(0,t) = 0; \qquad P(\infty,t) = 0; \qquad P(n,0) = \delta(n-n_0). \tag{30}$$

To do so, we first perform a Laplace transform:

$$\overline{P}(n,s) = \int_0^\infty \mathrm{e}^{-st} P(n,t)\,\mathrm{d}t \tag{31}$$

to obtain

$$s\overline{P}(n,s) - \delta(n-n_0) = D\frac{\partial}{\partial n}\left[\frac{c}{n} + \frac{\partial}{\partial n}\right]\overline{P}(n,s). \tag{32}$$

Integrating over a small interval around n_0 yields

$$\partial_n \overline{P}_<(n)|_{n=n_0} - \partial_n \overline{P}_>(n)|_{n=n_0} = \frac{1}{D} \tag{33}$$

where $\overline{P}_<(n)$ and $\overline{P}_>(n)$ are the solutions of equation (32) for $n < n_0$ and $n > n_0$, respectively. By defining $x = \sqrt{s/D}n$ and $P(n,s) = (Ds)^{-\frac{1}{2}} f(\sqrt{s/D}n)$ equation (32) becomes

$$f''(x) + \frac{c}{x}f'(x) - \left(1 + \frac{c}{x^2}\right)f(x) = 0 \tag{34}$$

which has the solution

$$f(x) = Ax^{\frac{1-c}{2}} I_{\frac{1+c}{2}}(x) + Bx^{\frac{1-c}{2}} K_{\frac{1+c}{2}}(x). \tag{35}$$

Here I_ν and K_ν are modified Bessel functions of the first and second kind [27]. Using their asymptotic behavior and the boundary conditions (30) we find $B = 0$ for $x < x_0 = \sqrt{s/D}n_0$ and $A = 0$ for $x > x_0$. Denoting $x_< = \min(x, x_0)$ and $x_> = \max(x, x_0)$ and using equation (33) the Laplace transform of the loop size distribution is given by

$$\overline{P}(s,x) = \frac{\left(\frac{x}{x_0}\right)^{(1-c)/2} I_{(1+c)/2}(x_<) K_{(1+c)/2}(x_>)}{\sqrt{Ds}\left(I'_{(1+c)/2}(x_0)K_{(1+c)/2}(x_0) - I_{(1+c)/2}(x_0)K'_{(1+c)/2}(x_0)\right)}. \tag{36}$$

Using standard methods [26] we integrate this expression to find the Laplace transform of the survival probability

$$\overline{G}(s,x_0) = \int_0^\infty \overline{P}(s,x)\sqrt{D/s}\,\mathrm{d}x = \frac{K_{\frac{1+c}{2}}(x_0)\left(I_{\frac{c-1}{2}}(x_0) - \frac{(\frac{1}{2}x_0)^{\frac{c-1}{2}}}{\Gamma((1+c)/2)}\right) + I_{\frac{1+c}{2}}(x_0)K_{\frac{1-c}{2}}(x_0)}{s\left(I'_{\frac{1+c}{2}}(x_0)K_{\frac{1+c}{2}}(x_0) - I_{\frac{1+c}{2}}(x_0)K'_{\frac{1+c}{2}}(x_0)\right)}. \tag{37}$$

The asymptotic behavior of $\overline{G}(s, n_0 = \sqrt{D/s}x_0)$ for long and short times can be extracted from the behavior of the Bessel functions. For small s equation (37) turns into

$$\overline{G}(s,n_0) \approx \begin{cases} \Phi(c)s^{\frac{c-1}{2}} & c \leqslant 1 \\ \dfrac{n_0^2}{2D(c-1)} + \Phi(c)s^{\frac{c-1}{2}} & c > 1 \end{cases} \tag{38}$$

where $\Phi(c)$ is a constant which depends on c. From this we can extract the asymptotic form of the survival probability for long times: $G(n_0, t \gg n_0^2/D) = g(\xi = Dt/n_0^2 \gg 1) \sim \xi^{-\frac{1+c}{2}}$. The behavior for short times can be obtained in a similar fashion, yielding $g(\xi \ll 1) \approx 1$. In sum, we find that the survival probability for a loop of initial size n_0 has the scaling form

$$G(n_0,t) = g\left(\frac{Dt}{n_0^2}\right), \tag{39}$$

with the asymptotic behavior

$$\begin{aligned} g(\xi \gg 1) &\sim \xi^{-\frac{1+c}{2}} \\ g(\xi \ll 1) &\sim 1. \end{aligned} \tag{40}$$

References

[1] For a review see Wartell R M and Benight A S 1985 *Phys. Rep.* **126** 67
Gotoh O 1983 *Adv. Biophys.* **16** 1
[2] Kafri Y, Mukamel D and Peliti L 2002 *Eur. Phys. J.* B **27** 135
[3] Peyrard M and Bishop A R 1989 *Phys. Rev. Lett.* **62** 2755
[4] Poland D and Scheraga H A 1966 *J. Chem. Phys.* **45** 1456
Poland D and Scheraga H A 1966 *J. Chem. Phys.* **45** 1464
[5] Fisher M E 1966 *J. Chem. Phys.* **45** 1469
[6] Kafri Y, Mukamel D and Peliti L 2000 *Phys. Rev. Lett.* **85** 4988
[7] Bockelmann U, Thomen P, Essevaz-Roulet B, Viasnoff V and Heslot F 2002 *Biophys. J.* **82** 1537
[8] Essevaz-Roulet B, Bockelmann U and Heslot F 1997 *Proc. Natl Acad. Sci. USA* **94** 11935
[9] Danilowicz C, Kafri Y, Conroy R S, Coljee V W, Weeks J and Prentiss M 2004 *Phys. Rev. Lett.* **93** 078101
[10] Danilowicz C, Coljee V W, Bouzigues C, Lubensky D K, Nelson D R and Prentiss M 2003 *Proc. Natl Acad. Sci. USA* **100** 1694
[11] Altan-Bonnet G, Libchaber A and Krichevsky O 2003 *Phys. Rev. Lett.* **90** 138101
[12] Ambjörnsson T, Banik S K, Krichevsky O and Metzler R 2006 *Phys. Rev. Lett.* **97** 128105
[13] Ambjörnsson T, Banik S K, Krichevsky O and Metzler R 2007 *Biophys. J.* **92** 2674
[14] Zeng Y, Montrichok A and Zocchi G 2003 *Phys. Rev. Lett.* **91** 148101
[15] Hanke A and Metzler R 2003 *J. Phys. A: Math. Gen.* **36** 473
[16] Bicout D J and Kats E 2004 *Phys. Rev.* E **70** 010902(R)
[17] Fogedby H and Metzler R 2007 *Phys. Rev.* E **76** 061915
[18] Bar A, Kafri Y and Mukamel D 2007 *Phys. Rev. Lett.* **98** 038103
[19] Kunz H, Livi R and Suto A 2007 *J. Stat. Mech.* P06004
[20] Blossey R and Carlon E 2003 *Phys. Rev.* E **68** 061911
[21] Carlon E, Orlandini E and Stella A L 2002 *Phys. Rev. Lett.* **88** 198101
[22] Causo M S, Coluzzi B and Grassberger P 2000 *Phys. Rev.* E **62** 3958

[23] Zia R K P, Lipowsky R and Kroll D M 1988 *Am. J. Phys.* **56** 160
[24] Lipowsky R 1991 *Europhys. Lett.* **15** 703
[25] Baiesi M, Carlon E, Kafri Y, Mukamel D, Orlandini E and Stella A L 2003 *Phys. Rev.* E **67** 021911
[26] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
[27] Abramowitz M and Stegun I A (ed) 1972 *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* 10th edn (New York: Dover)