ELSEVIER

Physica A 306 (2002) 39–50

PHYSICA A

www.elsevier.com/locate/physa

Denaturation and unzipping of DNA: statistical mechanics of interacting loops

Y. Kafri[a], D. Mukamel[a,*], L. Peliti[b]

[a]*Department of Physics of Complex Systems, The Weizmann Institute of Science, Rehovot 76100, Israel*

[b]*Dipartimento di Scienze Fisiche and Unità INFM, Università "Federico II", Complesso Monte S. Angelo, I–80126 Napoli, Italy*

Abstract

When DNA molecules are heated they undergo a denaturation transition by which the two strands of the molecule are separated and become unbound. Experimental studies strongly indicate that the denaturation transition is first order. The main theoretical approach to study this transition, introduced in the early 1960s, considers microscopic configurations of a DNA molecule as given by an alternating sequence of non-interacting bound segments and denaturated loops. Studies of this model usually neglect the repulsive, self-avoiding, interaction between different loops and segments and have invariably yielded continuous denaturation transitions. It is shown that the excluded volume interaction between denaturated loops and bound segments may be taken into account using recent results on the scaling properties of polymer networks of arbitrary topology. These interactions are found to drive the transition first order, compatible with experimental observations. The unzipping transition of DNA which takes place when the two strands are pulled apart by an external force acting on one end may also be considered within this approach, again yielding a first-order transition. Although the denaturation and unzipping transitions are thermodynamically first order, they do exhibit critical fluctuations in some of their properties. This appears, for example, in the algebraic decay of the loop size distribution at the thermal denaturation and in the divergence of the length of the end segment as the transition is approached in both thermal- and force-induced transitions. © 2002 Published by Elsevier Science B.V.

PACS: 87.14.Gg; 05.70.Fh; 64.10+h; 63.70.+h

Keywords: DNA denaturation; Self-avoiding walks; Polymer networks; Phase transitions

---

[*] Corresponding author.
E-mail address: fmukamel@wicc.weizmann.ac.il (D. Mukamel).

PII: S0378-4371(02)00483-1

1. Introduction

Melting, or thermal denaturation of DNA has been a subject of extensive experimental and theoretical studies since the early 1960s [1]. In this process, the two complementary strands of the DNA molecule which are bound to each other by hydrogen bonds detach upon increasing the temperature until they fully separate. Below the melting temperature the DNA is composed of a series of open loops separated by bound segments. The typical size of open loops increases with temperature, as the melting point is approached. The primary experimental tool for studying this process is the measurement of UV light absorption at a wavelength of about 270 nm. Light at this wavelength is preferentially absorbed by the single strands and it thus provides a measure for the fraction of the bounded pairs, $\theta(T)$, at any given temperature T. This is known as the optical melting curve of the DNA. One expects θ to decrease with temperature, and to vanish at the melting temperature T_{M}. Typically, the melting curve exhibits a series of plateaus and sharp steps whose size and temperature position depend on the sequence of the particular molecule under study. This is a result of the fact that the two types of base pairs composing a DNA molecule, AT and CG have different binding energies, with the AT bond being somewhat weaker than the CG one. The steps thus correspond to the melting of various segments of the DNA which are characterized by different local sequences with a different average binding energy. Thus, a segment containing relatively high concentration of AT bonds melts at a lower temperature than a segment of a similar length but with high concentration of CG bonds. The regions in the DNA chain which correspond to steps in the melting curve are usually referred to as cooperatively melting regions (CMR). The melting curve thus provides statistical information on the DNA sequence. It exhibits strong hysteresis effects. This is interpreted as an indication that denaturation (upon heating) and renaturation (upon cooling) of the CMRs take place at different temperatures (for a review see [1]). Direct observation of the DNA denaturated loop structure is provided by electron microscopy studies [2]. Experimentally, the steps in the melting curve (which correspond to the melting of a finite region) are found to be rather sharp with some rounding, usually attributed to finite size effects. This indicates that the melting transition is first order.

More recently, the introduction of new techniques such as optical tweezers and atomic force microscopy [3,4] has allowed the manipulation of single biological molecules. This made it possible to study a wider variety of physical properties of the DNA molecule. For example, optical tweezers have been used to apply a force and pull apart the two strands at one end of the molecule. It is found that a phase transition takes place at a critical force where the molecule is unzipped and the two strands are separated [5].

Thermal denaturation has been studied theoretically since the early 1960s. The early models, which we refer to as Poland–Scheraga (PS) type, or non-interacting loops models [6,7] consider the molecule as being composed of an alternating sequence of independent bound and denaturated segments. A bound segment is energetically favored over an unbound one, while a denaturated segment (a loop) is entropically favored. Within this approach the interaction between the different segments of the molecule is not taken into account. This assumption simplifies the analysis considerably. The order of the transition is found to be determined by a parameter c which characterizes the statistical weight of a loop. The number of configurations of a loop of length ℓ behaves as s^{ℓ}/ℓ^{c} for large ℓ. Here s is a non-universal constant. It has been shown [8] that the phase transition is first order if $c > 2$ and second order if $1 < c \leqslant 2$. For $c < 1$ no transition takes place and the strands are always bound. Using either a random walk [8] or self-avoiding random walk [9,10] [1] models for the loops has yielded a continuous transition in both $d = 2$ and 3 dimensions. It was suggested [9] that self-avoiding interactions *between* the various parts of the chain (and not just within loops) would further sharpen the transition even possibly making it first order. However, theoretical tools for carrying out this analysis had not been available at the time.

Numerical simulations of finite chains where excluded volume interactions are fully taken into account strongly suggest that the transition is indeed first order [11]. Recently, an approach has been introduced within which self-avoiding interactions both within a loop and between a loop and the rest of the chain may be accounted for [12,13]. To carry out the analysis of this model one has to enumerate the configurations of a loop embedded in a chain with self-avoiding interactions. This has been done by taking advantage of recent results obtained by Duplantier et al. [14,15] for the number of configuration of a general polymer network. It is found that the statistical weight of a loop embedded in a chain has the same general form as before, namely s^{ℓ}/ℓ^{c}. However, the parameter c is now modified and becomes larger than two in $d \geqslant 2$ dimensions. Thus, self-avoiding interactions make the transition first order in two dimensions and above. However, the transition is found to be accompanied by critical fluctuations in some properties of the chain. For example, the loop size distribution is found to decay algebraically at the transition. Indeed, the probability distribution for loops of length ℓ, $P(\ell)$, behaves as $P(\ell) \sim 1/\ell^{c}$ at the transition. This behavior was recently confirmed in numerical simulations of the model where the excluded volume interactions have been fully taken into account. The value of the measured exponent c agrees well with the theoretical predictions. It is also found that when the boundary conditions are such that the chain is open at one end the length of the end segment diverges as $1/|T - T_{\mathrm{M}}|$ when the melting temperature T_{M} is approached.

The model may be extended to consider the unzipping transition which takes place when a force of magnitude f is applied to separate the two strands. The model yields a first-order unzipping transition with the end segment length diverging as $1/|f - f_{\mathrm{U}}|$ near the unzipping critical force f_{U}. These results has previously been found in non-interacting loops models [17–21]. In calculating the critical force near the melting transition it is found that $f_{\mathrm{U}} \sim |T - T_{\mathrm{M}}|^{\nu}$ at variance with predictions of non-interacting loops models which yield $f_{\mathrm{U}} \sim |T - T_{\mathrm{M}}|^{1/2}$ [20]. Here ν is the correlation length exponent of a self-avoiding walk.

---

[1] This estimate is for all configurations, including knotted ones. However, it has been shown that in $d = 3$ the number of *unknotted* configurations is given by a formula of the same form but with a slightly smaller s and with an exponent c which is practically unmodified.

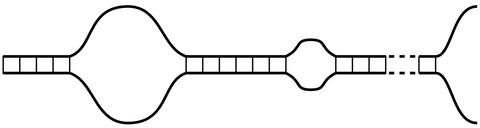

Fig. 1. Schematic representation of a microscopic configuration of the DNA molecule.

In a different class of models of denaturation [22–25] and unzipping [17–21], the two strands of the DNA are modeled by directed polymers. In this approach excluded volume interaction is not accounted for and the denaturation transition is invariably found to be continuous. A recent numerical study suggests that by incorporating the bending energies of the single and double strands into the model the denaturation transition becomes rather sharp, possibly first order [25] . However, it has been argued that as long as the attractive binding interaction between the two strands is short range, the transition is in fact continuous and the sharpness observed is a crossover effect [13]. Recently, studies in which the self-avoiding interaction is partially taken into account have been carried out. In one study it has been argued that taking into account excluded-volume interactions between the two strands while neglecting them within each strand leads to an effective long-range potential between the strands making the transition first order [26]. A different approach which takes into account the excluded volume interactions within each strand while neglecting the inter-strand interactions yields a continuous transition [16].

In this paper, the PS approach to thermal denaturation is reviewed and extended to include self-avoiding interactions. This approach is then extended to consider the unzipping transition.

2. Thermal denaturation of DNA

2.1. Non-interacting loops model

The PS model, introduced in the early 1960s [6,7], considers two strands, each composed of a sequence of monomers. Each monomer represents one persistence length of a single unbound strand. Typically, this is about ~ 40 Å [27], or roughly 8 bases. The persistence length of double-stranded DNA is at least an order of magnitude larger [28]. Monomers on one chain can be either unbound or bound to a specific matching monomer on the other chain. The interactions between a monomer and other monomers on the second strand or on the same strand are ignored. The monomers at one end of the molecule are assumed to be always bound. The binding energy $E_0 < 0$ between matching monomers is taken to be the same for all monomer pairs.

A typical DNA configuration is shown in Fig. 1. It is made of an alternating sequence of bound segments and denaturated loops. The configuration ends with two denaturated strands. The statistical weight of a bound sequence of length ℓ is given by $w^\ell = \exp(-\ell E_0/T)$, where T is the temperature and the Boltzmann constant k_B is set to 1. On the other hand, a denaturated loop does not carry an energy and its statistical weight is derived from its degeneracy. In this model, it is assumed that the loop is fully flexible, and thus it is described by a random walk which returns to the origin after 2ℓ steps. Considering all possible such walks the statistical weight for large ℓ has the form $\Omega(2\ell)=s^\ell/\ell^c$, where s is a non-universal constant and the exponent c is determined by the properties of the loop configurations.[2] Finally, the statistical weight of the end segment, which consists of two denaturated strands each of length ℓ, takes the form $\Lambda(2\ell)=s^\ell/\ell^{\bar{c}}$ for large ℓ, where $\bar{c}$ is in general not equal to c.[2] The values of the exponents c and $\bar{c}$ depend on the specific model used for the loops and the end-segments, respectively (see below).

Using the weights assigned to each segment of the chain the total weight of any given configuration may be calculated. For example the weight of a chain which consists of a bound segment of length ℓ_1, a denaturated loop of length ℓ_2, a bound segment of length ℓ_3, and a pair of denaturated strands of length ℓ_4, is given by $w^{\ell_1}\Omega(2\ell_2)w^{\ell_3}\Lambda(2\ell_4)$.

The model is most easily studied within the grand canonical ensemble where the total chain length L is allowed to fluctuate. The grand canonical partition function, $\mathscr{Z}$, is given by

$$\mathscr{Z}=\sum_{L=0}^{\infty}Z(L)z^L=\frac{V_0(z)Q(z)}{1-U(z)V(z)}\,, \tag{1}$$

where $Z(L)$ is the canonical partition function of a chain of length L, z is the fugacity, and the functions $U(z)$, $V(z)$ and $Q(z)$ are defined by

$$U(z)=\sum_{\ell=1}^{\infty}\Omega(2\ell)z^\ell=\sum_{\ell=1}^{\infty}\frac{s^\ell}{\ell^c}z^\ell\,, \tag{2}$$

$$V(z)=\sum_{\ell=1}^{\infty}w^\ell z^\ell\,, \tag{3}$$

$$Q(z)=1+\sum_{\ell=1}^{\infty}\Lambda(2\ell)z^\ell=1+\sum_{\ell=1}^{\infty}\frac{s^\ell}{\ell^{\bar{c}}}z^\ell \tag{4}$$

with $V_0(z)=1+V(z)$. Eq. (1) can be verified by expanding the partition function as a series in $U(z)V(z)$. The factors $V_0(z)$ and $Q(z)$ properly account for the boundaries. To set the average chain length, L, one has to choose a fugacity such that $L=\partial\ln\mathscr{Z}/\partial\ln z$. This implies that the thermodynamic limit $L\to\infty$ is obtained by letting z approach the lowest fugacity z^* for which the partition function (1) diverges. This can arise either from the divergence of the numerator or from the vanishing of the denominator. The relevant situation, at low temperature, is the second one, which corresponds to z^* satisfying $U(z^*)V(z^*)=1$. Above the transition, namely in the denaturated phase, the numerator diverges. Moreover, when one considers the problem of DNA unzipping by applying an external force on the strands, a divergence arising from a boundary factor will play an important role.

---

[2] In fact the number of configurations is only proportional to the expression given. For simplicity the proportionality constant taken to be 1.

The fraction of bound monomer pairs θ is the experimentally measured quantity and the order parameter of the transition. Its temperature dependence in the thermodynamic limit $L \to \infty$ can be calculated from the behavior of $z^*(w)$. The average number of bound pairs in a chain is given by $\langle m \rangle = \partial \ln \mathscr{Z} / \partial \ln w$, so that

$$\theta = \lim_{L\to\infty} \frac{\langle m \rangle}{L} = \frac{\partial \ln z^*}{\partial \ln w} \; . \tag{5}$$

Detailed analysis of $z^*(w)$ shows that the nature of the transition is determined by the exponent c as follows: (a) $c \leqslant 1$—no phase transition; (b) $1 < c \leqslant 2$—continuous phase transition; (c) $c > 2$—first-order phase transition. The nature of the phase transition is thus directly related to the number of configurations of long denaturated loops within the chain. In the case of a continuous transition, the order parameter θ behaves near the melting point as $\theta \sim |T - T_{\mathrm{M}}|^{(2-c)/(c-1)}$.

In the early studies of this problem the exponent c was evaluated by enumerating all random walks of a given length which return to the origin [8]. It is easy to show that in d dimensions the model yields $c{=}d/2$. This implies that there is no transition for $d \leqslant 2$, a continuous transition for $2 < d \leqslant 4$ and a first-order phase transition for $d > 4$. The model was subsequently extended to include the repulsive short range interaction which exists between the strands constituting a loop. In this approach the loop is modeled as a self-avoiding walk [9]. This yields $c = d\nu$, where ν is the exponent associated with the radius of gyration R_{G} of a self-avoiding walk. For a walk of length L one has $R_{\mathrm{G}} \sim L^{\nu}$, with $\nu = \frac{3}{4}$ in $d = 2$ and $\nu \approx 0.588$ in $d = 3$. This yields $c = \frac{3}{2}$ in $d = 2$ and $c = 1.766$ in $d = 3$. Thus the transition is continuous in both cases, although it is sharper than when the repulsive interaction is neglected altogether.

The two estimates of the exponent c described above treat the loop as an isolated object and thus neglect its interaction with the rest of the chain. This simplification is essential for carrying out the analysis. In the next section, an approach is described within which the repulsive interaction between a loop and the rest of the chain may be accounted for. Although these interactions are treated only in an approximate way, this approach yields insight into the unbinding mechanism and on the nature of the transition.

2.2. Interacting loops approach

To account for the excluded volume interactions between a loop and the rest of the chain one notes that a microscopic configuration of the DNA molecule is composed of many bound and unbound segments of various length. In evaluating the number of available configurations of a loop the interactions with all these bound and unbound segments have to be taken into account. This problem is simplified by neglecting the internal structure of the rest of the chain. One thus considers a loop embedded in a flexible chain (see Fig. 2) and studies the number of configurations of a chain endowed with this topology, assuming that it is self avoiding [35]. One can show that in the limit where the loop length, 2ℓ, is much smaller than the length of the rest of the chain, $2L$, the statistical weight of this topology can be written as a product of the statistical weight of the loop with that of the chain. The weight of the loop is found to be of the same form as that of a free loop but with a different exponent c. This exponent is found

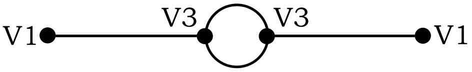

Fig. 2. The topology of the loop embedded in a chain. The length of the chain from a vertex of type V1 to the nearest vertex of type V3 is L. The length of each of the two strands connected to the V3 vertices is ℓ.

to be larger than two in dimensions 2 and above, yielding a first-order denaturation transition.

To carry out this analysis one uses results obtained by Duplantier et al. [14,15] for the number of configurations of polymer networks of arbitrary topology. This represents an extension of the well known results for the number of configurations of a simple self-avoiding random walk [29]. In that case it is known that the number of configurations scales as $\Gamma_{\text{linear}} \sim s^L L^{\gamma-1}$, where L is the length of the polymer, s is a non-universal geometrical constant and γ is a universal exponent. The exponent γ is known exactly in $d=2$, numerically in $d=3$ and via an ε expansion in $d=4-\varepsilon$. Above $d=4$ self-avoiding interactions becomes irrelevant and thus the number of configurations of self-avoiding random walks scales as that of ordinary random walks, yielding $\gamma=1$. The generalization of this result to an arbitrary polymer network goes as follows [14,15]: Consider a branched self-avoiding polymer $\mathscr{G}$ of arbitrary topology. The polymer is made of N chains of lengths $\ell_1, \ell_2, \ldots, \ell_N$. These are tied together at vertices with different number of legs. A vertex with k legs is said to be of order k ($k \geqslant 1$). The number of vertices of order k is denoted by n_k. The number of configurations of the network, $\Gamma_{\mathscr{G}}$, is then given by

$$\Gamma_{\mathscr{G}} \sim s^L L^{\gamma_{\mathscr{G}}-1} g\left(\frac{\ell_1}{L}, \frac{\ell_2}{L}, \ldots, \frac{\ell_N}{L}\right), \tag{6}$$

where $L = \sum_i \ell_i$ is the total length of the network and g is a scaling function. This expression is valid asymptotically as $L \to \infty$. The function g is smooth when its arguments are finite. However, it may be singular when at least one of its arguments approaches zero (which amounts to a crossover to a different topology of the network). Thus, when the thermodynamic limit is taken such that all lengths ℓ_i scale in the same way, the number of configurations is simply given by $\Gamma_G \sim s^L L^{\gamma_G - 1}$. Note that this relation is valid even when the persistence length of each of the chains composing the graph is different. The exponent γ_G depends only on the topology of the network and is given by

$$\gamma_G = 1 - \nu d \mathscr{L} + \sum_{k \geqslant 1} n_k \sigma_k . \tag{7}$$

Here $\mathscr{L}$ is the number of independent loops in the network, d the spatial dimension and ν is the exponent related to the radius of gyration of a self-avoiding random walk. The scaling dimensions σ_k, defined for $k \geqslant 1$, are known exactly in $d=2$ from conformal invariance

$$\sigma_k = (2-k)(9k+2)/64 , \tag{8}$$

and to order ε^2 in $d = 4 - \varepsilon$.

$$\sigma_k = (\varepsilon/8)(2-k)k/2 + (\varepsilon/8)^2 k(k-2)(8k-21)/8 + O(\varepsilon^3) . \tag{9}$$

Also, estimates for the values of the exponents in $d=3$ are available through Padé and Padé–Borel approximants. Clearly, $\sigma_2=0$ as one would expect and above $d=4$, where the self-avoiding interaction is irrelevant, all the exponents σ_k are zero.

Consider now the topology depicted in Fig. 2. The quantity of interest is the number of configurations of the network in the limit $\ell \ll L$, when the loop size is much smaller than the length of the rest of the chain. Using the results by Duplantier (see Eq. (6)), the number of configurations can be written as

$$\Gamma \sim s^{L+\ell}(L+\ell)^{\gamma_{\rm loop}-1} g(\ell/L) \tag{10}$$

for large L and ℓ. Here, $g(x)$ is a scaling function and $\gamma_{\rm loop}$ can be evaluated using Eq. (7). For the topology considered above of a loop embedded in two segments (Fig. 2) one has

$$\gamma_{\rm loop} = 1 - d\nu + 2\sigma_1 + 2\sigma_3 . \tag{11}$$

The limit of interest is that of a loop size much smaller than the length of the chain, $\ell/L \ll 1$. Clearly, in the limit $\ell/L \to 0$, the number of configurations should reduce to that of a single self-avoiding open chain, which, to leading order in L, is given by $s^L L^{\gamma-1}$, where $\gamma = 1 + 2\sigma_1$. This implies that in the limit $x \ll 1$

$$g(x) \sim x^{\gamma_{\rm loop}-\gamma} . \tag{12}$$

Thus the number of configurations is given by

$$\Gamma \sim s^{\ell}\ell^{\gamma_{\rm loop}-\gamma} s^L L^{\gamma-1} . \tag{13}$$

It is therefore evident that, for large ℓ and L and in the limit $\ell/L \ll 1$, the partition sum is decomposed into a product of the partition sums of the loop and that of the rest of the chain. The excluded volume interaction between the loop and the rest of the chain is reflected in the value of the effective exponent c. This result is very helpful since it enables one to extend the Poland–Scheraga approach described in the previous section to the case of interacting loops. From Eq. (13) one sees that the appropriate effective exponent c is given by

$$c = \gamma - \gamma_{\rm loop} = d\nu - 2\sigma_3 . \tag{14}$$

In $d=2$, $\sigma_3 = -\frac{29}{64}$ [14] and $\nu = \frac{3}{4}$, yielding $c = 2 + \frac{13}{32}$. In $d = 4 - \varepsilon$ to $O(\varepsilon^2)$, one has $\sigma_3 = -3\varepsilon/16 + 9\varepsilon^2/512$ and $\nu = \frac{1}{2}(1 + \varepsilon/8 + 15/4(\varepsilon/8)^2)$, yielding $c = 2 + \varepsilon/8 + 5\varepsilon^2/256$. In $d=3$, one may use Padé and Padé–Borel approximations to obtain $\sigma_3 \approx -0.175$ [15] which with the value $\nu \approx 0.588$ [15] yields $c \approx 2.115$. The value of the exponent c is unaffected by the different persistence length of a bound and unbound DNA segment. This is since as stated above Eq. (6) is valid also when the persistence length of different polymers composing the network are different.

The fact that the exponent c is found to be larger than 2 in $d = 2$, $4-\varepsilon$ and 3 strongly suggests that the transition is *first order* for any $d \geqslant 2$. The analysis assumes that the size of the loops in the system is much smaller than the total chain length. The fact that the transition is first order implies that the loops size remain finite as the transition is approached from below. This makes the analysis self-consistent. Note that the loop size distribution, $P(\ell)$, is rather broad at the transition and behaves, for large ℓ, as $P(\ell) \sim 1/\ell^c$. Thus, high enough moments of the loop size distribution always diverge. Thus, although the transition is first order for $c > 2$ it exhibits some critical properties. Since c is found to satisfy $2 < c < 3$ the variance of the loops size is predicted to diverge. A recent numerical study of the loop size distribution, where the excluded volume interactions have been fully taken into account, has verified this prediction with a measured value of $c = 2.10 \pm 0.02$ [16].

In the approach described above, the rest of the chain is approximated by a bound segment. In reality, it is composed of an alternating sequence of loops and bound segments. To test the validity if this approximation one may consider the other extreme whereby the rest of the chain is taken to be a macroscopic loop. The exponent c for this case is found to be larger but very close to the one obtained above [12,13].

To evaluate the length of the end-segment near the transition one needs both exponents c and $\bar{c}$. Using the methods described above one finds $\bar{c} = -(\sigma_1 + \sigma_3)$. This yields $\bar{c} = \frac{9}{32}$ in $d = 2$, $\bar{c} \simeq 0.092$ in $d = 3$, and $\bar{c} = \varepsilon/8 + O(\varepsilon^2)$ in $d = 4 - \varepsilon$, while above four dimensions clearly $\bar{c} = 0$. This suggests that the exponent $\bar{c}$ is smaller than one for any $d \geqslant 2$. Using these results one can show [13] that near the melting transition the end segment length, ξ, diverges like

$$\xi \sim \frac{1}{|T - T_{\mathrm{M}}|} . \tag{15}$$

3. The unzipping transition

Inspired by the development of experimental techniques in which a single molecule can be accurately manipulated (see e.g. [30]), the DNA unzipping transition has become a subject of extensive theoretical studies [17–21,31–34]. Many of these studies use a directed polymer approach to model the DNA, where self-avoiding interactions are not accounted for. The effect of chain heterogeneity has also been considered in some of these studies [18,19]. In this section, the analysis of the PS model is extended to consider the unzipping of homopolymers with self-avoiding interactions. It is shown that the unzipping transition is first order. The dependence of the critical unzipping force on the temperature at low forces, namely near the melting temperature, is also calculated.

Consider a configuration where the corresponding monomers at one end of the chain are bound together, while a force $\vec{f}$ is applied on the two monomers at the other end of the chain, pulling the two strands apart. In this setup, the grand canonical partition function takes the form

$$Z = \frac{V_0(z)O(z)}{1 - U(z)V(z)} , \tag{16}$$

where the factor $O(z)$ is the grand partition function of the open tail under force. One has

$$O(z) = 1 + \sum_{\ell=1}^{\infty} Z_{\mathrm{end}}(\ell) z^{\ell} , \tag{17}$$

where $Z_{\text{end}}(\ell)$ is the canonical partition function of an open end composed of two strands, each of length ℓ.

To evaluate $Z_{\text{end}}(\ell)$ note that when no force is applied the partition sum takes the form $Z_{\text{end}}(\ell) = \Lambda(2\ell) \sim s^{\ell}/\ell^{\bar{c}}$, where, as discussed above, $\bar{c}$ can be evaluated and is smaller than one. When a force $\vec{f}$ is applied, one has

$$Z_{\text{end}}(\ell) = \Lambda(2\ell) \int \mathrm{d}\vec{r}\, p_{\ell}(r) \exp(\vec{f} \cdot \vec{r}/T) , \tag{18}$$

where $p_{\ell}(r)$ is the probability distribution of the end-to-end distance in the absence of a force. It is assumed that $p_{\ell}(r)$ has the same scaling form as that of linear polymers $p_{\ell}(r) = R^{-d} \hat{p}(r/R)$. Here, R is a scaling length related to ℓ by $R \simeq R_0 \ell^{\nu}$, where ν is the correlation length exponent of a linear polymer. For $x \gg 1$ $\hat{p}(x)$ takes the form [34]

$$\hat{p}(x) = P x^{\mu} \exp(-D x^{\lambda}) , \tag{19}$$

where P and D are constants, $\lambda = 1/(1-\nu)$ and the exponent μ is given by

$$\mu = (d/2 + \nu d - \bar{c})/(1-\nu) . \tag{20}$$

A saddle point evaluation of Eq. (18) yields that for $fR/T \gg 1$

$$O(z) \simeq 1 + \sum_{\ell=1}^{\infty} \left[zs \exp\left(A (f R_0/T)^{1/\nu} \right) \right]^{\ell} . \tag{21}$$

According to Eq. (21) at temperatures below the melting temperature T_{M} the end segment partition sum $O(z)$ diverges at a critical, unzipping force f_{U}, given by $\exp(-A(f_{\text{U}} R_0/T)^{1/\nu}) = sz^*(w)$, for forces which are not too small. Here, $z^*(w)$ is the solution of $U(z^*)V(z^*) = 1$ (corresponding to an infinitely long polymer). At this point the average length of a loop in the bulk is finite. Hence, the unzipping transition is first order.

Near the transition, the length ξ of the end segment diverges like $|z^* - z_{\text{U}}|^{-1}$, where $z_{\text{U}} = \exp(-\kappa (f_{\text{U}}/T)^{1/\nu})/s$. Since z^* is regular in f, one has $\xi \sim |f - f_{\text{U}}|^{-1}$ or $\xi \sim |T - T_{\text{U}}(f)|^{-1}$. Thus, the two strands separate gradually from the end as the critical force is approached. Nonetheless, the unzipping transition is first order. The reason is that the transition takes place at a temperature below the denaturation melting temperature T_{M} where the loop size distribution in the interior of the chain decays exponentially with the loop size. Thus, at this point the average loop size in the interior of the chain is finite. On the other hand the length of the end segment is finite as long as $f < f_{\text{U}}$ and its contribution to the order parameter θ and to the entropy is negligible. Therefore, both the order parameter and the entropy exhibit a discontinuity in their values in the unzipped state at the transition. Note that an analysis of Eq. (21) shows that the critical force f_{U} behaves like

$$f_{\text{U}} \sim |T - T_{\text{M}}|^{\nu} \tag{22}$$

as $T \to T_{\text{M}}$, at least as long as the forces are not too small (i.e., $fR/T \gg 1$). This result is independent of the value of c.

4. Summary

In this paper, the Poland–Scheraga type (i.e., non-interacting loops) models of DNA denaturation are extended to fully (though approximately) account for the self-avoiding interaction within the molecule. This analysis is made possible by the scaling theory developed over the last 15 years for polymer networks with arbitrary topology. It is shown that the self-avoiding interaction drives the denaturation transition first order. It is also found that the model exhibits critical behavior in some of its properties, such as its loop size distribution and the length of the end segment. Experimental study of these properties would be of great interest. The model is also extended to study the unzipping transition induced by an external force.

References

[1] R.M. Wartell, A.S. Benight, Phys. Rep. 126 (1985) 67;
O. Gotoh, Adv. Biophys. 16 (1983) 1.
[2] V.M. Pavlov, J.L. Lyubchenko, A.S. Borovik, Y. Lazurkin, Nucl. Acids. Res. 4 (1977) 4052;
A.S. Borovik, Y.A. Kalambet, Y.L. Lyubchenko, V.T. Shitov, E. Golovanov, Nucl. Acids. Res. 8 (1980) 4165.
[3] K. Svoboda, S.M. Block, Ann. Rev. Biophys. Biomol. Structure 23 (1994) 247;
A. Ashkin, Proc. Natl. Acad. Sci. USA 94 (1997) 4853.
[4] H.G. Hansma, J. Vac. Sci. Technol. B 14 (1995) 1390.
[5] U. Bockelmann, B. Essevaz-Roulet, F. Heslot, Phys. Rev. Lett. 79 (1997) 4489;
B. Essevaz-Roulet, U. Bockelmann, F. Heslot, Proc. Natl. Acad. Sci. USA 94 (1997) 11 935.
[6] D. Poland, H.A. Scheraga (Eds.), Theory of Helix-Coil Transitions in Biopolymers, Academic Press, New York, 1970;
F.W. Wiegel, in: C. Domb, J.L. Lebowitz (Eds.), Phase Transitions and Critical Phenomena, Vol. 7, Academic Press, New York, 1983, p. 101.
[7] T.L. Hill, J. Chem. Phys. 30 (1959) 383;
B.H. Zimm, J. Chem. Phys. 33 (1960) 1349;
S. Lifson, J. Chem. Phys. 40 (1964) 3705;
M.Ya. Azbel, Phys. Rev. A 20 (1979) 1671.
[8] D. Poland, H.A. Scheraga, J. Chem. Phys. 45 (1966) 1456;
D. Poland, H.A. Scheraga, J. Chem. Phys. 45 (1966) 1464.
[9] M.E. Fisher, J. Chem. Phys. 45 (1966) 1469.
[10] A.L. Kholodenko, T.A. Vilgis, Phys. Rep. 298 (1998) 251.
[11] M.S. Causo, B. Coluzzi, P. Grassberger, Phys. Rev. E 62 (2000) 3958.
[12] Y. Kafri, D. Mukamel, L. Peliti, Phys. Rev. Lett. 85 (2000) 4988.
[13] Y. Kafri, D. Mukamel, L. Peliti, cond-mat/0108323.
[14] B. Duplantier, Phys. Rev. Lett. 57 (1986) 941;
B. Duplantier, J. Stat. Phys. 54 (1989) 581.
[15] L. Schäfer, C. von Ferber, U. Lehr, B. Duplantier, Nucl. Phys. B 374 (1992) 473.
[16] E. Carlon, E. Orlandini, L. Stella, cond-mat/0108308.
[17] S.M. Bhattacharjee, J. Phys. A 33 (2000) L423;
S.M. Bhattacharjee, Erratum: 33 (2000) 9003.
[18] D.K. Lubensky, D.R. Nelson, Phys. Rev. Lett. 85 (2000) 1572;
D.K. Lubensky, D.R. Nelson, cond-mat/0107423.
[19] N. Hatano, D.R. Nelson, Phys. Rev. Lett. 77 (1996) 570;
N. Hatano, D.R. Nelson, Phys. Rev. B 56 (1997) 8651.
[20] D. Marenduzzo, A. Trovato, A. Maritan, Phys. Rev. E 64 (2001) 031901.

[21] S. Cocco, R. Monasson, J.F. Marko, Proc. Natl. Acad. Sci. USA 98 (2001) 8608.

[22] M. Peyrard, A.R. Bishop, Phys. Rev. Lett. 62 (1989) 2755.

[23] D. Cule, T. Hwa, Phys. Rev. Lett. 79 (1997) 2375.

[24] S. Cocco, R. Monasson, Phys. Rev. Lett. 84 (1999) 5178.

[25] N. Theodorakopoulos, T. Dauxois, M. Peyrard, Phys. Rev. Lett. 85 (2000) 6.

[26] T. Garel, C. Monthus, H. Orland, Europhys. Lett. 55 (2001) 132.

[27] M.T. Record, S.J. Mazur, P. Melanon, J.H. Roe, S.L. Shaner, L. Unger, Ann. Rev. Biochem. 50 (1981) 997.

[28] V.A. Bloomfield, D.M. Crothers, I. Tinoco, Nucleic Acids - Structure, Properties, and Function, University, Sausalito, 2000.

[29] P.-G. de Gennes, Scaling Concepts in Polymer Physics, Cornell, Ithaca, 1979.

[30] B. Smith, L. Finzi, C. Bustamante, Science 258 (1992) 1122;
T.R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, V. Croquette, Science 271 (1996) 1835;
J.-F. Léger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, J.F. Marko, Phys. Rev. Lett. 83 (1999) 1066.

[31] S.M. Bhattacharjee, D. Marenduzzo, cond-mat/0106110.

[32] D. Marenduzzo, S.M. Bhattacharjee, A. Maritan, E. Orlandini, F. Seno, cond-mat/0103142.

[33] E.A. Mukamel, E.I. Shakhnovich, cond-mat/0108447.

[34] J. des Cloiseaux, G. Jannink, Les polymères en solution: leur modélisation et leur structure, Les Éditions de Physique, Les Ulis, 1987.

[35] Y. Kafri, D. Mukamel, L. Peliti, cond-mat/0112179.