Why is the DNA Denaturation Transition First Order?

Yariv Kafri, David Mukamel and Luca Peliti*

Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot 76100, Israel
(November 8, 2000)

We study a model for the denaturation transition of DNA in which the molecules are considered as composed of a sequence of alternating bound segments and denatured loops. We take into account the excluded-volume interactions between denatured loops and the rest of the chain by exploiting recent results on scaling properties of polymer networks of arbitrary topology. The phase transition is found to be *first order* in $d = 2$ dimensions and above, in agreement with experiments and at variance with previous theoretical results, in which only excluded-volume interactions within denatured loops were taken into account. Our results agree with recent numerical simulations.

PACS numbers: 87.14.Gg, 05.70.Fh, 64.10+h,63.70.+h

Thermal denaturation or melting of double-stranded DNA is the process by which the two strands unbind upon heating. The nature of this transition has been investigated for almost four decades [1–7]. Experimentally, a sample containing molecules of a specific length and sequence is prepared. Then the fraction of bound base-pairs as a function of temperature, referred to as the melting curve, is measured through light absorption, typically at about 260 nm. For heterogeneous DNA, where the sequence contains both AT and GC pairs, the melting curve exhibits a multistep behavior consisting of plateaus with different sizes separated by sharp jumps. These jumps have been attributed to the unbinding of domains characterized by different frequencies of AT and GC pairs. The sharpness of the jumps suggests that the transition from bound to unbound is first order.

The early theoretical models [2,3], which we refer to as Poland-Scheraga (PS) type models, consider the DNA molecule as composed of an alternating sequence of bound and denatured states (see, e.g., Fig. 1). Typically a bound state is energetically favored over an unbound one, while a denatured segment (loop) is entropically favored over a bound one. Within the PS type models the segments which compose the chain are assumed to be non-interacting with one another. This assumption considerably simplifies the theoretical treatment and enables one to calculate the resulting free energy. In the past the entropy of the denatured loops has been evaluated by modelling them either as ideal random walks [8] or as self-avoiding walks [9]. It has been found that within this approach the denaturation transition of DNA is continuous both in two and three dimensions and it becomes first order only above four dimensions. It was suggested [9] that taking into account the interaction between loops would sharpen the transition.

A version of the PS model which includes the excluded-volume interactions between the various segments of the chain has been recently introduced [10]. This extension makes the model analytically intractable, but numerical studies in three dimensions indicate that these interactions indeed drive the transition first order.

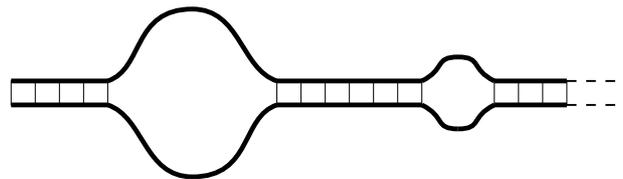In this Letter we consider analytically the effects of

FIG. 1. Schematic representation of the Poland-Scheraga model.

excluded-volume interaction between the various segments of the chain. Although we treat this interaction only in an approximate way, we are able to give some insight into the unbinding mechanism and on the nature of the transition. Our approach makes use of recent important results on the entropy of self-avoiding polymer networks [11,12]. We find that this interaction drives the transition *first order* in $d = 2$, 3 and $4 - \varepsilon$ dimensions. To proceed we introduce and discuss the PS model. A scaling argument which takes into account the interaction between the segments of the DNA chain is then presented and used to study the transition. Finally, we comment on disorder and compare our results to those obtained from more recent models of DNA denaturation [4–7].

The PS model considers two strands, made of monomers, each representing one persistence length of a single strand (ordinarily ~ 40 Å [13]). For simplicity we take boundary conditions where the monomers at the ends of the molecule are always bound. All other monomers on the chain can be either bound or unbound to a specific matching monomer on the second chain. The binding energy $E_0 < 0$ is taken to be the same for all matching monomers. A typical DNA configuration is shown in Fig. 1. It is made of sequences of bound monomers separated by denatured loops. The statistical weight of a bound sequence of length k is $\omega^k = \exp(-kE_0/T)$, where T is the temperature [14]. On the other hand the statistical weight of a denatured sequence of length k is given by the change in entropy due to the added configurations arising from a loop of

length $2k$. For large k this has the general form As^k/k^c, where s is a non-universal constant and the exponent c is determined by the properties of the loop configurations. For simplicity, we set $A = 1$. 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, $\mathcal{Z}$, is given by

$$\mathcal{Z} = \sum_{M=0}^{\infty} G(M) z^M = \frac{V_0(z) U_L(z)}{1 - U(z)V(z)}, \qquad (1)$$

where $G(M)$ is the canonical partition function of a chain of length M, z is the fugacity and

$$U(z) = \sum_{k=1}^{\infty} \frac{s^k}{k^c} z^k, \qquad V(z) = \sum_{k=1}^{\infty} \omega^k z^k, \qquad (2)$$

$V_0(z) = 1 + V(z)$ and $U_L(z) = 1 + U(z)$. Equation (1) can be verified by expanding the partition function as a series in $U(z)V(z)$. The factors $V_0(z)$ and $U_L(z)$ properly account for the boundaries. The average chain length, $\langle L \rangle$, is set by choosing a fugacity such that $\langle L \rangle = \partial \ln \mathcal{Z}/\partial \ln z$. We wish to evaluate the order parameter θ, i.e., the fraction of bounded pairs, as a function of temperature, in the limit $\langle L \rangle \to \infty$. The average number of bounded pairs in a chain is given by $\langle m \rangle = \partial \ln \mathcal{Z}/\partial \ln \omega$, so that

$$\theta = \lim_{L \to \infty} \frac{\langle m \rangle}{\langle L \rangle} = \frac{\partial \ln z^*}{\partial \ln \omega}. \qquad (3)$$

Here z^* is the value of the fugacity in the limit $\langle L \rangle \to \infty$. This is the lowest value of the fugacity for which the partition function (1) diverges, i.e., for which $U(z^*)V(z^*) = 1$. Using $V(z) = \omega z/(1 - \omega z)$ one has

$$U(z^*) = 1/(\omega z^*) - 1. \qquad (4)$$

It is clear that the nature of the denaturation transition is determined by the dependence of z^* on ω. The function $U(z)$ is independent of ω. It is finite for $z < 1/s$ and diverges when $z > 1/s$. On the other hand the function $1/V(z)$ increases continuously as ω decreases (corresponding to an increase in T). Thus, as the temperature increases from $T = 0$, the fraction of attached monomers decreases and z^* increases. However if z^* reaches the value $z_c = 1/s$ (so that $1/V(1/s) \geq U(1/s)$) any further increase of the temperature leaves z^* unchanged so that $\theta = 0$. Therefore the transition takes place at $z^* = z_c = 1/s$. Its nature is determined by the behaviour of $U(z)$ in the vicinity of z_c. This is controlled in turn by the value of the exponent c. There are three regimes:

1. For $c \leq 1$, $U(z_c)$ diverges, so that z^* is an analytic function of ω and no phase transition takes place.

2. For $1 < c \leq 2$, $U(z_c)$ converges but its derivative $U'(z)$ diverges at $z^* = z_c$. Thus $\theta \propto \partial z^*/\partial \omega$ vanishes continuously ($\theta \sim |T - T_c|^{(2-c)/(c-1)}$) at the transition and the transition is *continuous*.

3. For $c > 2$, both $U(z)$ and its derivative converge at $z^* = z_c$ and the transition is *first order*. Here, in contrast to the continuous case, the average size of a denatured loop is finite at the transition.

The value of the exponent c can be obtained by enumerating random walks which return to the origin, so that $c = d\nu$. For ideal random walks this yields $c = d/2$: thus there is no transition at $d \leq 2$, a continuous transition for $2 < d \leq 4$ and a first order transition only for $d > 4$ [8]. On the other hand, for self-avoiding random walks the excluded volume interaction modifies the exponent to $c = 3/2$ for $d = 2$ and $c \approx 1.764$ for $d = 3$. The transition is thus sharper, but still continuous, in three dimensions [9].

In the above treatment the entropy of a loop of length $2l$ was taken to be of the form s^l/l^c. In determining the exponent c the interaction of the loop with the rest of the chain has been ignored. In the following we argue that, when the interaction of the loop with the rest of the chain is taken into account, the above form for the loop entropy is still valid, but the exponent c is modified. The excluded volume interaction of the loop with the rest of the chain is expected to lower the entropy of the loop yielding a larger value of the exponent c.

We evaluate the exponent c in this situation by considering the number of configurations of a loop of length $2l$ embedded in a chain of length $2L$ (see Fig. 2). In reality the chain itself is composed of an alternating sequence of loops and bound segments of various sizes. However, in the present analysis we ignore the detailed structure of the rest of the chain and consider it as being in a bound state. The fact that the rest of the chain is made of both bound and unbound segments increases the excluded volume interaction and is thus expected to increase further the value of c. This point will be discussed later.

In order to evaluate the exponent c for our configuration, we use the results obtained by Duplantier [11,12] for the entropy of general polymer networks: Consider a polymer network composed of segments and vertices with an arbitrary given topology. The results state that for a network composed of segments with an average length L_A the number of configurations is proportional to $\propto s^{L_A} L_A^{\gamma_G - 1}$ where γ_G is an exponent which depends *only* on the topology of the network. It is given by

$$\gamma_G = 1 - d\nu \mathcal{L} + \sum_{N \geq 1} n_N \sigma_N, \qquad (5)$$

where $\mathcal{L}$ is the number of independent loops, n_N is the number of vertices with N outgoing legs, and σ_N is an exponent related to such a vertex. The exponents σ_N are exactly known in $d = 2$ and in $d = 4 - \varepsilon$ to $O(\varepsilon^2)$.

Consider now the relevant topology to the problem discussed in this work (Fig. 2). We are interested in enumerating the configurations of the network in the limit $l/L \ll 1$, when the loop size is much smaller than the rest of the chain. To do this we note that when only two

FIG. 2. The topology of the loop embedded in a chain. The distance along the chain from a vertex of type $V1$ to the nearest vertex of type $V3$ is given by L. The distance between the vertices of type $V3$ is equal to l.
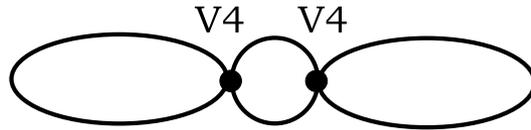

FIG. 3. An extreme topology where the loop of length $2l$ is between two denaturated loops of size $2L$ each. The vertices of order 4 are denoted by $V4$.

length scales determine the geometry of the network, as in the case depicted in Fig. 2, the number of configurations can be written as [12]

$$\Gamma = s^{L+l}(L+l)^{\gamma_{\text{loop}}-1}g(l/L), \qquad (6)$$

for large L and l. Here $g(x)$ is a scaling function and γ_{loop} can be evaluated using eq. (5). For the topology considered above of a loop embedded in two segments (Fig. 2) we have: one loop, $\mathcal{L} = 1$; two vertices of order one, $n_1 = 2$, corresponding to the two free ends of the chain (denoted by $V1$ in the figure); two vertices of order three, $n_3 = 2$ (denoted in the figure by $V3$). Using eq. (5) we obtain

$$\gamma_{\text{loop}} = 1 - d\nu + 2\sigma_1 + 2\sigma_3. \qquad (7)$$

We are interested in a loop size much smaller than the length of the chain, $l/L \ll 1$. Clearly, in the limit $l/L \to 0$, 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 $x \ll 1$,

$$g(x) \sim x^{\gamma_{\text{loop}}-\gamma}. \qquad (8)$$

Thus the l-dependence of Γ, which gives the change in the number of configurations available to the loop, is

$$\Gamma \propto s^l l^{\gamma_{\text{loop}}-\gamma} \cdot s^L L^{\gamma-1}. \qquad (9)$$

It is therefore evident that for large l and L and in the limit $l/L \ll 1$ the partition sum is decomposed into a product of the partition sums of the loop and that of the rest of the chain [15]. The excluded volume effect between these two parts reflects itself in the value of the effective exponent c associated with the loop entropy. This result is very helpful since it enables one to extend the PS approach to the case of interacting loops. From Eq. 9 one sees that the appropriate effective exponent c is

$$c = \gamma - \gamma_{\text{loop}} = d\nu - 2\sigma_3. \qquad (10a)$$

In $d = 2$ $\sigma_3 = -29/64$ [11] and $\nu = 3/4$ yielding

$$c = 2 + 13/32. \qquad (10b)$$

In $d = 4 - \varepsilon$ to $O(\varepsilon^2)$ one has $\sigma_3 = -3\varepsilon/16 + 9\varepsilon^2/512$ and $\nu = 1/2(1 + \varepsilon/8 + 15/4(\varepsilon/8)^2)$ yielding

$$c = 2 + \varepsilon/8 + 5\varepsilon^2/256. \qquad (10c)$$

In $d = 3$ one may use Padé and Padé-Borel approximations to obtain $\sigma_3 \approx 0.175$ [12] which with the value $\nu \approx 0.588$ [12] yielding $c \approx 2.115$.

Equations (10) are our main result. The fact that the exponent c is found to be larger than 2 in $d = 2, 3$ and $4 - \varepsilon$ indicates that the transition is first order in $d \geq 2$.

As noted above this approach does not take into account all the excluded volume interactions in the system since the rest of the chain is assumed to be in a bound state. In practice however the rest of the chain is by itself composed of loops and bound segments of various sizes. The existence of loops is expected to enhance the excluded volume interaction. Taking into account all these configurations in the partition sum is a formidable task. However, we estimate below the effect of these interactions on the partition sum and show that it does not modify the main conclusion of this analysis, namely that the transition is first order.

For a given loop-bound configuration of the rest of the chain it is straightforward to check that the l dependence of the partition sum (9) in the scaling limit is valid as long as the two bound segments which are attached to the loop under consideration are long. On the other hand when at least one of these segments is short and the adjacent open loop is large the effective exponent c is modified. In order to estimate this exponent we consider the extreme case where the rest of the chain is composed of two loops each of size $2L$ (see Fig. 3). A similar analysis to that presented above yields for the value of c,

$$c = d\nu - \sigma_4, \qquad (11a)$$
$$= 2 + 11/16, \qquad \text{in } d = 2, \qquad (11b)$$
$$= 2 + \varepsilon/4 - 15\varepsilon^2/128, \qquad \text{in } d = 4 - \varepsilon, \qquad (11c)$$

where the values $\sigma_4 = -19/16$ in $d = 2$ and $\sigma_4 = -\varepsilon/2 + 11(\varepsilon^2/8)^2$ in $d = 4 - \varepsilon$ dimensions [11] are used along with those of ν. Using $\sigma_4 \approx 0.46$ obtained by Padé and Padé-Borel approximations gives in $d = 3$ the value $c \approx 2.22$. It is easy to verify that this scaling behavior is unchanged as long as one of the bound states connected to the loop under consideration is small (so there is at least one vertex of order four). Evidently c for this configuration is larger than that of the configuration in Fig. 2. One can show that for arbitrary loop-segment

configuration of the chain the effective c is given by either Eqs. (10) or Eqs. (11) depending on the length of the bound segments attached to the loop. Since in both cases the exponent c is found to be larger than 2 in $d = 2$, $d = 4 - \varepsilon$ and in $d = 3$, this analysis strongly suggests that the transition is first order for $d \geq 2$. The analysis presented above is valid as long as the length l of an individual loop is much smaller than the total length of the chain. The fact that the transition is found to be first order, so that the average loop size remains finite at the transition, makes the analysis self consistent.

The model treated above assumes, for simplicity, that the monomers at the ends of the two chains are bounded. In reality the ends need not be attached and one should consider open boundary conditions. Similar analysis for the case where one of the ends of the chain is open yields the same result. Note that one end of the chains must remain attached for a bound state to exist. In the above approach entangled configurations with non equivalent knot topologies, which are possible in $d = 3$, are counted [11] as they should when one end of the chains is open.

Finally, we compare the model to recently studied models for the denaturation transition. These have focused on the stacking energy generating a varying stiffness in the bound and denatured state [4,5,7] or on the helical structure of the DNA molecule [6]. All these studies are within a directed polymer approach and thus ignore the effects studied in this Letter.

In summary, we have shown that the reduction in the number of available configurations of denatured loops due to the excluded-volume interactions between a denatured loop and the rest of the chain is sufficient to drive the denaturation transition first order. This is done by extending the PS type models to include these interactions. Our results have a direct implication on recent numerical simulations of the model [10]. The model studied here treats the binding energy between matching monomers is taken to be equal, although each base can bind only to its corresponding base on the other strand. This last assumption is reasonable when the heterogeneity of the chain is taken into account. For short segments to bind the detailed sequences of the two segments have to match and this is less likely to take place in a heterogeneous system. It is therefore important to consider the effect of the heterogeneity of the binding energy between the matching base pairs on the order of the transition. Also, another effect which has recently been emphasized is the varying stiffness of the phosphate backbone. The model considered above does not treat these effects.

It is instructive to note that although the average loop size is finite at the transition, the loop size distribution is rather broad and decreases as $1/l^c$ for large l. For $c < 3$ this yields a diverging variance of the loop size. It would be of great interest to test this prediction experimentally.

We are indebt to M. J. E. Richardson and H. Orland for many inspiring discussions and their involvement earlier stages of our study of the DNA denaturation transition.

* Michael Visiting Professor. On leave of absence from Dipartimento di Scienze Fisiche and Unità INFM, Università "Federico II", Complesso Monte S. Angelo, I–80126 Napoli (Italy).

[1] For a review see R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985); O. Gotoh, Adv. Biophys. **16**, 1 (1983).

[2] For a review see D. Poland and H. A. Scheraga (eds.), *Theory of Helix-Coil Transitions in Biopolymers* (Academic, New York, 1970); F. W. Wiegel, in C. Domb and J. L. Lebowitz (eds.): *Phase Transitions and Critical Phenomena Vol. 7*, Pg. 101 (Academic, New York, 1983).

[3] T. L. Hill, J. Chem. Phys. **30**, 383 (1959); B. H. Zimm, J. Chem. Phys. **33**, 1349 (1960); S. Lifson, J. Chem. Phys. **40**, 3705 (1964); M. Ya. Azbel, Phys. Rev. A **20**, 1671 (1979).

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

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

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

[7] N. Theodorakopoulos, T. Dauxois ans M. Peyrard, cond-mat/0004487.

[8] D. Poland and H. A. Scheraga, J. Chem. Phys. **45**, 1456 (1966); J. Chem. Phys. **45**, 1464 (1966).

[9] M. E. Fisher, J. Chem. Phys. **45**, 1469 (1966).

[10] M. S. Causo, B. Coluzzi and P. Grassberger, Phys. Rev. E **62**, 3958 (2000).

[11] B. Duplantier, Phys. Rev. Lett. **57**, 941 (1986); J. Stat. Phys. **54**, 581 (1989).

[12] L. Schäfer, C. von Ferber, U. Lehr and B. Duplantier, Nucl. Phys. B **374**, 473 (1992).

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

[14] For simplicity, the configurational entropy related to the embedding of the chain of bound monomers in ambient space is neglected. One may check that it will not modify our main results.

[15] This scaling analysis is also valid for chain configurations considered by J. des Cloizeaux (J. Physique **41**, 223 (1980)), where the probability of contact between two points along a self-avoiding chain is calculated. These configurations were also considered in [11]. It should be noted that the simple approach described above holds as long as, in the limit $l \to 0$, no higher-order vertices (in the sense discussed above) are generated.