Cent. Eur. J. Phys. • 1-8
Author version

Central European Journal of Physics

Constrained thermal denaturation of DNA under fixed linking number

Research Article

Amir Bar[1,2*], Alkan Kabakçıoğlu[3†], David Mukamel[1‡]

1 Department of Physics of Complex Systems, The Weizmann Institute of Science, Rehovot 76100, Israel

2 Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel

3 Department of Physics, Koç University Sarıyer 34450 İstanbul, Turkey

Abstract: A DNA molecule with freely fluctuating ends undergoes a sharp thermal denaturation transition upon heating. However, in circular DNA chains and some experimental setups that manipulate single DNA molecules, the total number of turns (linking number) is constant at all times. The consequences of this additional topological invariant on the melting behavior are nontrivial. Below, we investigate the melting characteristics of a homogeneous DNA where the linking number along the melting curve is preserved by supercoil formation in duplex portions. We obtain the mass fraction and the number of loops and supercoils below and above the melting temperature. We also argue that a macroscopic loop appears at T_c and calculate its size as a function of temperature.

PACS (2008): 87.15.Zg, 36.20.Ey

Keywords: DNA denaturation • circular DNA • supercoiling

1. Introduction

The two strands of a B-DNA molecule unbind upon heating in water when the entropic free energy gain due to unhybridized single-strand DNA overweighs the enthalpic contribution of hydrogen bonding by duplex formation [1]. Thermal denaturation temperature increases with (and has been used as an experimental probe for) the fraction of GC to AT content [2]. Mechanical denaturation [3–5] has been proposed as a tool for sequencing [6]. Theoretical models of DNA denaturation which employ the entropy-enthalpy interplay as the driving mechanism of the phase transition date back to '60s, to works by Fisher [7], Poland, and Scheraga [8]. It was already observed

* *E-mail: Amir.Bracha@weizmann.ac.il*
† *E-mail: akabakcioglu@ku.edu.tr*
‡ *E-mail: David.Mukamel@weizmann.ac.il*

by then, that the sharpness of the transition depends on the subleading logarithmic correction to the entropy which is dominated by the contribution from flexible, denatured regions (loops). The precise scaling form of the entropy for self-avoiding polymer chains with given connective topology was calculated later by Duplantier [9] and adopted to the DNA melting problem relatively recently [10] by Kafri *et al.*. They observed that the melting process of a homogeneous DNA (i.e., with a binding energy that is uniform beyond some microscopic scale) is a first-order phase transition. In fact, experiments on long DNA molecules display multiple sharp denaturation steps corresponding to the unbinding of regions with different AT/GC content.

These findings apply to DNA molecules with free ends, in which a loop formed by dehybridizing a sequence of base pairs bears no consequence on the rest of the molecule. Things change if the ends are rotationally constrained (e.g., attached to a surface) or covalently bonded to each other, forming a circle. These examples comprise a new physical setup where DNA's thermal denaturation should be investigated. The novel ingredient is the introduction of the winding or the linking number (LK), as a new topological invariant. An immediate consequence of fixing LK is the lack a complete denaturation due to the entanglement of the two strands. However, as we discuss below, this fact does not rule out the presence of a true melting transition at a finite temperature. In fact, the new melting scenario turns out to be richer than the original Poland-Scheraga (PS) model [8]. Finally note that, the problem is not a purely academic exercise. Circular DNAs are ubiquitous not only in nature [2] (all bacteria have them), but also in labs where they are constantly used for DNA replication in PCR machines; while DNA chains with rotationally constrained ends are frequently used in single-molecule experiments [11].

In Section 2, we describe how one can generalize the PS model in order to construct a fixed-LK ensemble. Section 3 investigates the melting behavior with emphasis on the macroscopic loop created above the critical temperature. We conclude with a summary of our results and future prospects.

2. A DNA ensemble with fixed linking number

The Poland-Scheraga model [8] considers a homogeneous DNA, where each unit on a given strand is allowed to bind to a unique complementary unit in the opposite strand, with some binding energy $\epsilon < 0$. Duplex regions are assumed infinitely stiff, therefore have negligible contribution to the polymer's entropy. At nonzero temperatures, a finite fraction of the complementary units is expected to be unbound. A successive sequence of such pairs forms a loop, i.e., a denatured segment. Due to a roughly 10-fold smaller persistence length, these single-stranded DNA loops are entropically favorable. The statistical weight of a loop formed by l successive units is given by

$$\Omega(l) = A\frac{s^l}{l^c}$$

where c is a universal exponent determined by geometric considerations [9] and A and s are phenomenological constants. For a freely fluctuating DNA one finds $c > 2$ which yields a first-order denaturation transition [10].

The model above should be extended when the linking number is fixed. If one assumes that the loops fluctuate freely, the winding number of one DNA single-strand around the other in a loop of length l should be zero on

average, with an rms value $\sim l^{1/2}$ [12]. Therefore we assume that, the turns previously residing on the loops are transfered to the remaining bound portions of the DNA. The Boltzmann factor for a circular DNA chain with total length L and a single loop of length l now has the form

$$\Omega(l)\, Z_b(L-l,l)$$

where the unit length is set to the pitch of the DNA and $Z_b(L-l,l)$ refers to the partition sum for a bound DNA of length $L-l$ with l extra turns.

In order to proceed, we need to introduce a physical mechanism through which the extra turns are accommodated in a DNA duplex. In the limit when the distance between the two strands is negligible relative to the DNA length, a well-known theorem [13, 13]: $LK = Tw + Wr$ relates the linking number to the twisting and bending degrees of freedom. Tw and Wr are called *twist* and *writhe*, respectively, which are defined as follows: Let $\vec{r}(s)$ be the position vector along the centerline parametrised by s, $\vec{t}(s) = d\vec{r}/ds$ be the corresponding tangent vector and $\hat{u}(s)$ be the unit vector pointing from one strand to the other (perpendicular to $\vec{t}$). Then,

$$\begin{aligned} Tw &= \oint \frac{ds}{2\pi}\, \vec{t}(s)\cdot\vec{u}(s)\times d\vec{u}(s)/ds \\ Wr &= \oint \frac{ds\, ds'}{4\pi}\, \vec{t}(s)\times\vec{t}(s')\cdot(\vec{r}(s)-\vec{r}(s'))/|\vec{r}(s)-\vec{r}(s')|^3 \end{aligned}$$

The twist accounts for the helicity of the DNA duplex. For a straight DNA chain (closed at infinity) $Tw = LK_0 \equiv 2\pi L/\lambda$, where λ is the helical pitch. The writhe quantifies the part of the linking number absorbed by the excursions of the backbone. The fraction and the energetic cost of the extra turns transferred from the loops and residing in twist and writhe are determined by the twist and bending moduli, respectively.

The melting behavior in the limit of infinite bending stiffness where the duplex regions remain rigid (as assumed in the PS model) has been investigated in Refs.[14, 15], and for $c > 2$ it was found that the transition becomes continuous [14]. A discussion of this extreme is postponed to the end of the paper, since available numerical evidence points to the plausibility of the opposite extreme. For example, Ref.[16] argues that the longer a circular DNA is, the higher is the fraction of LK that is absorbed by the *writhe*. Therefore, we here choose to relax the assumption of rigid DNA duplexes and allow them to accommodate a finite configurational entropy, while keeping the pitch λ uniform and at its native value (i.e., no overtwist). An analytically trackable model that allows entropy in bound DNA regions was recently proposed [17] and shown by us to display a continuous denaturation transition for $c > 2$, in which the singularity shifts to progressively higher derivatives of the free energy as $c \to 2^+$ [18]. Below we review this model and present further numerical analysis on the statistics of denatured loops and supercoils, with an emphasis on loop condensation which yields a macro-size loop above the melting temperature.

For convenience we assume that, a DNA duplex with an extra linking number forms supercoils with a fixed pitch (e.g., consider tight supercoils where the pitch is determined by matching the grooves of the two facing duplexes). The linking number absorbed by such a supercoil is given by its length upto some proportionality constant which

we set to unity. The partition function for a bound DNA duplex of length L_b and an excess LK of k residing in supercoils can then be expressed as

$$Z_b(L_b,k) \;=\; \sum_{n=1,2,\ldots}\;\sum_{l_1,l_2,\ldots}{}'\;\sum_{k_1,k_2,\ldots}{}'\;\prod_i \omega^{l_i}\prod_j \nu^{k_j} \tag{1}$$

where ω and ν are the Boltzmann factors for a unit length of DNA duplex outside and inside a supercoil, the inner sums with a ($'$) are subject to constraints $\sum_i(l_i+k_i)=L_b$ and $\sum_i k_i = k$, and n counts the pairs of alternating supercoils and unsupercoiled duplex regions. Introducing the loops as well, the canonical partition for the whole DNA is given by

$$Z(L,k) \;=\; \sum_{n=0,1,2,\ldots}\;\sum_{l_0,l_1,\ldots}{}'\;\sum_{\lambda_1,\lambda_2,\ldots}{}'\;\sum_{k_0,k_1,\ldots}{}' Z_b(l_0,k_0)\prod_{i=1}^{n}\Omega(\lambda_i)Z_b(l_i,k_i) \tag{2}$$

where λ_i are the length of denatured loops, n is the number of loops, and ($'$) refers to the restrictions $\sum_i(l_i+\lambda_i)=L$ and $\sum_i(k_i-\lambda_i)=k$. We assume that the end segments (with lengths l_0 and l_n) are bound. It is convenient to switch ensembles and work with the grand sum

$$Q(z,\mu)=\sum_{L,k} Z(L,k)z^L\mu^k \tag{3}$$

$Q(z,\mu)$ can be expressed diagrammatically as below after the DNA configurations over which the sum is performed are grouped according to the number of loops they contain:

$$\begin{aligned} Q(z,\mu) \;&=\; [\dot{—}] \;+\; [\dot{—}\bigcirc\dot{—}] \;+\; [\dot{—}\bigcirc\dot{—}\bigcirc\dot{—}] \;+\; \cdots \\ &=\; \frac{\dot{—}}{1-\bigcirc\dot{—}} \equiv \frac{\tilde{V}(z,\mu)}{1-U(z\mu)\tilde{V}(z,\mu)}\;. \end{aligned} \tag{4}$$

Above, $\tilde{V}(z,\mu)$ ($\dot{—}$) refers to the grand sum for a bound DNA duplex decorated with an arbitrary number of supercoils, and $U(z\mu)$ ($\bigcirc$) is the similar sum for a denatured loop. Using Eq.(1), $\dot{—}$ itself can be arranged into a series involving terms with increasing number of supercoils. Below are the corresponding diagrams where "–" and "|" represent bound segments and supercoils, respectively:

$$\begin{aligned} \tilde{V}(z,\mu)=[\dot{—}] \;&=\; [\_] \;+\; [\perp] \;+\; [\perp\perp] \;+\; [\perp\perp\perp] \;+\; \cdots \\ &=\; \frac{-}{1-\llcorner} \equiv \frac{V(z)}{1-W(z/\mu)V(z)} \end{aligned} \tag{5}$$

The functions $V(x)$, $W(x)$, $U(x)$ can be calculated explicitly as:

$$\begin{aligned} V(z) \;&=\; \frac{\omega z}{1-\omega z} \\ W(x) \;&=\; B\frac{x}{1-\nu x} \\ U(x) \;&=\; \sum_{n=1}^{\infty} A\frac{(sx)^n}{n^c} = A\Phi_c(sx) \end{aligned} \tag{6}$$

where $\Phi_c(x)$ is the polylog function and B is another phenomenological constant associated with the energetic cost of initiating a supercoil. Substituting in Eq.(4), we obtain

$$Q(z,\mu) \;=\; \left[-1+\frac{1}{\omega z}-B\frac{\nu z}{\mu-\nu z}-A\Phi_c(sz\mu)\right]^{-1} \tag{7}$$

The total DNA length is given by $L=\partial\log Q/\partial\log z$, while the (excess) linking number is fixed by the condition $\partial\log Q/\partial\log\mu=k$. For a DNA chain with no extra turns, $k=0$ at all temperatures. Also note that, $\nu=0$ gives the PS model and $\mu=1$ yields an extension with supercoils but no constraint on LK. Below, we present numerical data for the number and mass fraction of loops and supercoil, supporting the analytical results in Ref.[18].

2.1. Melting and loop condensation

Using the standart machinary of equilibrium statistical mechanics, we observe that the thermodynamics limit reached by setting z to the smallest singularity of Q given explicitly by Eq.(4). When $sz\mu<1$ (i.e. for low temperatures) z and μ can therefore be obtained using

$$Q(z,\mu)^{-1} \;=\; 0 \tag{8}$$

$$\partial\log(Q)/\partial\log\mu \;=\; 0 \quad\Rightarrow\quad \partial(Q^{-1})/\partial\mu=0. \tag{9}$$

where Q^{-1} is given by Eq.(7). For $c>2$, a phase transition arises from the nonanalytic behavior observed when the solution to Eqs.(8,9) coincides with the tip of the branch cut for $\Phi_c(sz\mu)$ at

$$sz\mu=1\ . \tag{10}$$

Beyond this point, Eqs.(8,9) together with Eq.(10) overdetermine z and μ. After a detailed analysis, one finds that when $T>T_c$ Eq.(9) acquires a correction due to the appearance of a loop of macroscopic dimensions [18] and is modified as below:

$$\partial\log(Q^{-1})/\partial\log\mu|_{sz\mu=1} \;=\; Aa_0\ . \tag{11}$$

where a_0/ζ_c is precisely the contribution of the macroscopic loop to the average loop size and it is an increasing function of temperature as shown below. In contrast, the average size of microscopic loops increases with temperature only for $T<T_c$ and remains constant at the value ζ_{c-1}/ζ_c for $T>T_c$ (this follows from the expression for $U(x)$ in Eq.(6).)

Below we calculate the fraction of DNA base pairs in denatured loops (m_l) and supercoils (m_s), as well as the number of loops (n_l) and supercoils (n_s), normalized by the length L of the DNA in the limit $L\to\infty$. These

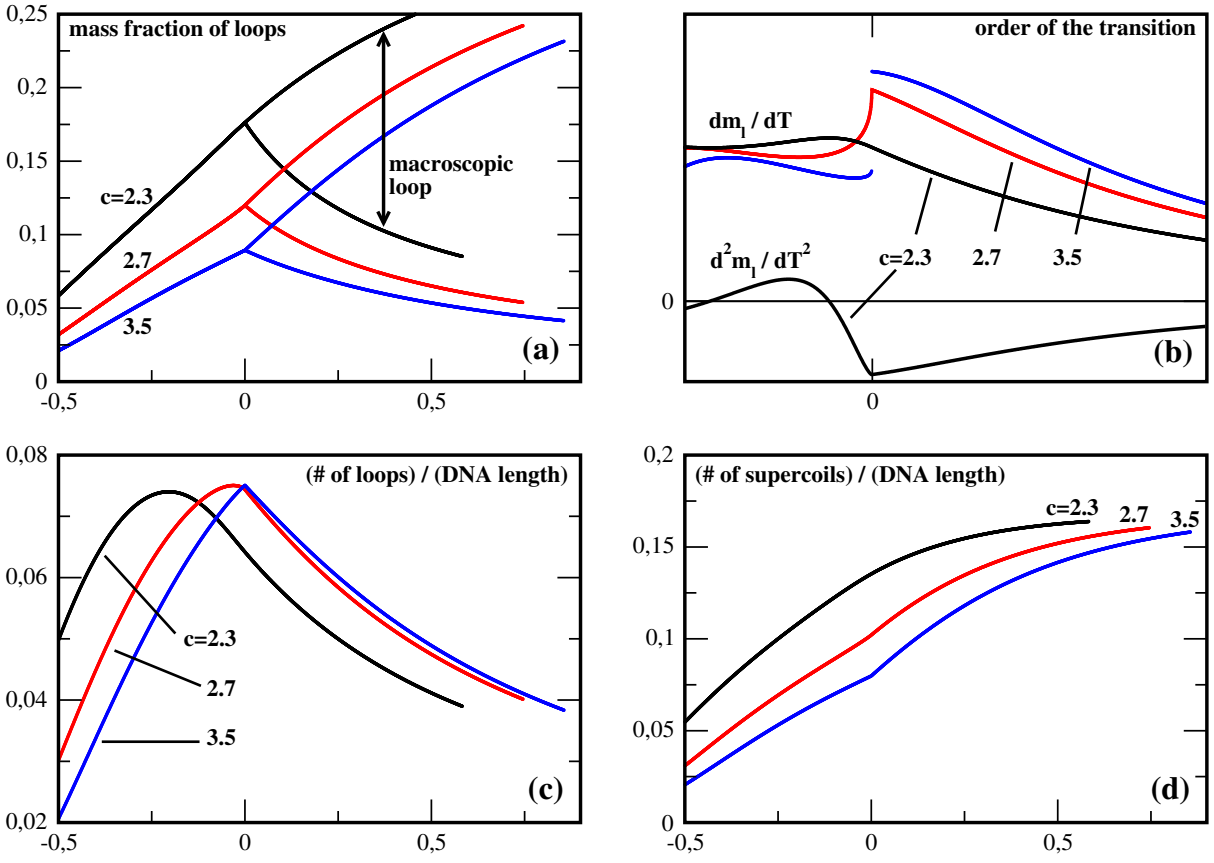

Figure 1. (a) The mass fraction of denatured DNA (the order parameter for melting) *vs* reduced temperature for several values of c. The lower branches are for the microscopic loops only. (b) Temperature derivatives of the order parameter are shown in order to deduce the order of the phase transition for each c value. The vertical axis is scaled arbitrarily for better visibility. (c) The number of denatured loops scaled by the DNA length. The number of loops first increases and then decreases as a function of temperature, with a precursor peak below T_c for $c < 3$. (d) The number of supercoils scaled by the DNA length. A singular behavior at T_c similar to (b) is found in (c) and (d) as well. Horizontal axes in all graphs is the reduced temperature $(T - T_c)/T_c$.

quantities can be expressed as follows:

$$\begin{aligned} m_l &= -\partial \log z/\partial \log s \ , \quad m_s = -\partial \log z/\partial \log \nu \\ n_l &= -\partial \log z/\partial \log A \ , \quad n_s = -\partial \log z/\partial \log B \end{aligned} \tag{12}$$

where z is obtained using Eqs.(8,9) for $T \leq T_c$ and Eqs.(8,10) for $T \geq T_c$. A plot for each of the densities above is given in Fig.1 for several values of c. The loop fraction m_l and its partitioning into contributions from microscopic loops and the macroscopic loop above T_c (with the fraction $a_0(T) : \zeta_{c-1}$) is given in Fig.1a. The temperature derivatives of the order parameter are shown in Fig.1b. It is evident that the transition becomes smoother with decreasing c. The discontinuity at the singular point shifts from the first derivative of m_l ($c = 3.5$) to the second

($c = 2.7$) and the third ($c = 2.3$) as c gets smaller, in agreement with Ref.[18] where the order of the transition was derived analytically. On the other hand, the macroscopic loop grows continuously from zero but with a finite slope at T_c for all $c > 2$. Interestingly, the mass fraction in microscopic loops displays a cusp at T_c and *decreases* above it. The mass fraction residing in supercoils is equal to that in the loops (not shown) by LK conservation. Figs.1c,1d show the *number* of loops and supercoils. While the supercoil number increases monotonously around the melting temperature, the loop number has a more complex dependence on temperature and c, due to the entropic advantage of merging microscopic loops. For $c > 3$, the loop number behaves similar to the mass fraction in microscopic loops. It increases for $T < T_c$ and declines steadily for $T > T_c$, with a cusp at T_c. For $c < 3$, a precursor peak below T_c indicates that loop merger rate catches up with the increase in m_l even before a macroscopic loop appears, so that the number of microscopic loops starts diminishing below T_c. This is understandable considering the form of $\Omega(l)$. The entropic advantage of having few, big loops is larger for smaller c and eventually wins against the positional entropy of having several small loops, as the loop mass density increases with temperature.

3. Conclusion

Thermodynamics of DNA melting with a fixed linking number is qualitatively different than that understood by the standart Poland-Scheraga model. When supercoiling in duplex regions is induced by denatured loops, we derived the mass fraction and the number of denatured loops and supercoils in the thermodynamic limit. We showed that a macroscopic loops appears at the melting temperature and grows steadily with temperature above it. Our preliminary calculations (to be reported elsewhere) suggest that this feature is not specific to the supercoiling mechanism, and also applies to the case when one considers overtwisting of the DNA duplexes as the means for LK conservation. Implications of a fixed LK on the melting dynamics and whether a macroscopic loop (through merger of microscopic loops) is dynamically accessible are interesting questions to be investigated in the future.

This work has been supported by the Scientific and Technological Research Council of Turkey (TUBITAK) through Grant TBAG-110T618, and by the Israel Science Foundation (ISF).

References

[1] R. Wartell and A. Benight, Physics Reports **126**, 67 (1985).

[2] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, and A. Campbell, *Essential cell biology* (Garland Science New York, 2004).

[3] I. Rouzina and V. Bloomfield, Biophysical journal **80**, 882 (2001).

[4] S. Cocco, R. Monasson, and J. Marko, Proceedings of the National Academy of Sciences **98**, 8608 (2001).

[5] D. Marenduzzo, A. Trovato, and A. Maritan, Phys. Rev. E **64**, 031901 (2001).

[6] A. Marziali and M. Akeson, Annual review of biomedical engineering **3**, 195 (2001).

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

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

[9] B. Duplantier, Journal of Statistical Physics **54**, 581 (1989).

[10] Y. Kafri, D. Mukamel, and L. Peliti, Phys. Rev. Lett. **85**, 4988 (2000).

[11] Z. Bryant, M. Stone, J. Gore, S. Smith, N. Cozzarelli, and C. Bustamante, Nature **424**, 338 (2003).

[12] E. Rensburg, E. Orlandini, D. Sumners, M. Tesi, and S. Whittington, Journal of Physics A: Mathematical and General **26**, L981 (1993).

[13] J. White, American Journal of Mathematics **91**, 693 (1969).

[14] J. Rudnick and R. Bruinsma, Phys. Rev. E. **65**, 030902(R) (2002).

[15] T. Garel, H. Orland, and E. Yeramian, Arxiv preprint q-bio/0407036 (2004).

[16] M. Sayar, B. Avşaroğlu, and A. Kabakçıoğlu, Phys. Rev. E **81**, 041916 (2010).

[17] A. Kabakçıoğlu, E. Orlandini, and D. Mukamel, Phys Rev E. **80**, 010903(R) (2009).

[18] A. Bar, A. Kabakcioglu, and D. Mukamel, Phys. Rev. E **84**, 041935 (2011).