Type II DNA: When the interfacial energy becomes negative

A Letters Journal Exploring the Frontiers of Physics

August 2011

EPL, **95** (2011) 48009
doi: 10.1209/0295-5075/95/48009

www.epljournal.org

Type II DNA: When the interfacial energy becomes negative

Poulomi Sadhukhan[a], Jaya Maji and Somendra M. Bhattacharjee

Institute of Physics - Bhubaneswar 751 005, India

received 25 February 2011; accepted in final form 6 July 2011
published online 5 August 2011

PACS 87.14.gk – DNA
PACS 87.15.Zg – Phase transitions
PACS 74.20.De – Phenomenological theories (two-fluid, Ginzburg-Landau, etc.)

Abstract – An important step in transcription of a DNA base sequence to a protein is the initiation from the exact starting point, called promoter region. We propose a physical mechanism for identification of the promoter region, which relies on a new classification of DNAs into two types, Type I and Type II, like superconductors, depending on the sign of the energy of the interface separating the zipped and the unzipped phases. This is determined by the energies of helical ordering and stretching over two independent length scales. The negative interfacial energy in Type II DNA leads to domains of helically ordered state separated by defect regions, or blobs, enclosed by the interfaces. The defect blobs, pinned by non-coding promoter regions, would be physically distinct from all other types of bubbles. We also show that the order of the melting transition under a force is different for Type I and Type II.

DNA in its double helical form shows a resilience against an external pulling force. The bound state does not allow a force g applied at an end to penetrate up to a critical force $g = g_c$, above which the DNA gets unzipped [1–3]. The transition is first order for temperatures $T < T_c$, where T_c is the denaturation (melting) temperature in the absence of any force [4]. The force-induced unzipping transition of DNA is due to a competition between the bond orientation by force and ordering by base pairing. The formation of a helically ordered dsDNA from denatured strands is a symmetry breaking transition. At a coarse-grained level, the ordered state can be described by an order parameter ψ, with $\psi = 0$ for the denatured state. The external force does not couple directly to this order parameter. Consequently, at a junction of a bound and an unzipped DNA, there is a need to define two length scales: one scale ξ that gives the length over which the DNA ordering is damaged on the bound side of the interface, while the other scale λ gives the distance over which the force penetrates the bound state. The existence of the second scale λ was pointed out by de Gennes in a model involving stretching of the backbone and the hydrogen bonds [5]. Generally one expects interfaces separating phases to be energetically costly (*e.g.* surface tension), but here we show that if $\lambda \gg \xi$, then the interfacial energy, or surface energy, between bound and unzipped DNA can become negative. There can then be a penetration of force in the form of distorted regions or "defect blobs" of length λ enclosing a denatured bubble of size ξ. In analogy to superconductors, when the interfacial energy becomes negative, one gets a mixed phase of DNA and the zipped-to-mixed phase transition becomes continuous. Based on the sign of the zipped-unzipped interfacial energy we classify DNA into two types: Type II has negative interfacial energy whereas Type I is the conventional case with positive interfacial energy. This classification is not related to the existing classification based on DNA conformation.

A Type II DNA has novel features which are of considerable biological and physical implications. To be noted that the defect blobs are different from thermally created bubbles. This is because the bubbles of the latter type would consist of random configurations of denatured strands generated by thermal fluctuations and may have positive interfacial energy. The distinctness of the defect blobs can be a signature for their identification in biological processes. Let us consider the transcription process where the genetic code, determined by the base sequence, is transfered to the amino acid sequence of a protein. For correct transcription, the sequence must be read from the correct starting point on DNA. These starting non-coding regions are called promoter regions and their identification is the first and vital step in transcription [6]. A pulling force or a forced separation in a homogeneous Type II DNA produces a finite density of the defect

[a]E-mail: poulomi@iopb.res.in

blobs [7] (discussed later). The non-coding sequences or the promoter regions may act as inhomogeneities on a DNA and could play the role of pinning centers for the defect blobs. The advantage of physical identification of pinned defect blobs could facilitate recognition of the promoter regions for gene expression (*e.g.*, see [8,9]). So far as physical properties are concerned, Type I and Type II DNA will have different phase diagram and phase transition as discussed later.

Recently, both in experiment [10] and simulation [11], a continuous transition has been observed if the topology is preserved in a stretching experiment by pulling both the strands either at both ends or at one end of an anchored DNA. We also note that a detailed molecular dynamics study [12] of under- or over-wound DNA without writhe, observed the formation of localized sequence-dependent defects which allow the rest of the dsDNA to be in the relaxed normal state. It is known that topoisomerase II may bind anywhere on the DNA but its topology changing activity is restricted to specific sequences (cleavage sites) indicating that geometric distortions get localized around certain sequences [13]. These are consistent with our general predictions, though we like to add that interfacial information in any of these cases are not available.

The thermodynamic description of unzipping of DNA requires three variables, ψ describing the helical ordering (*i.e.*, broken symmetry) and a force-displacement (g, x) conjugate pair, where x is the scaled separation between the two strands at the point of application of force g. On the bound side x can be taken as the response to an internal induced force $\tilde{g}$, so that,

$$x(\tilde{g}) = \chi \tilde{g}, \tag{1}$$

where χ, the stretchability or the response function, is independent of g in the linear response regime. Though we restrict to linear response regime here, the final results can be reproduced for a general force-dependent χ. The variables are chosen such that $\psi = 0$ for the unzipped state, and $\psi \neq 0$ for the ordered state, while $\tilde{g} = 0$ in the bulk of the ordered state. At this point it is to be noted that the order parameter ψ represents helical ordering which is not directly coupled to the external pulling force. As a result we get two independent length scales in the problem. This makes the present treatment different from other existing models.

For a homogeneous state, the Gibbs free energy $G(T, g)$ per unit length at temperature T and a pulling force g is given by

$$G(T, g) = G(T, 0) - W(g), \tag{2}$$

where $W(g) = \int_0^g x(g')\,\mathrm{d}g'$ is the work for stretching. The conditions of phase coexistence at $g = g_c$, *i.e.*

$$G_{\mathrm{z}}(T, g_c) = G_{\mathrm{u}}(T, g_c), \tag{3}$$

and non-penetration of force in the bound state for $g \leqslant g_c$, *i.e.*,

$$G_{\mathrm{z}}(T, g_c) = G_{\mathrm{u}}(T, g_c), \tag{4}$$

(subscripts z and u representing the zipped and the unzipped phases), when combined with eq. (2), give

$$G_{\mathrm{z}}(T, g) = G_{\mathrm{u}}(T, g) + W(g) - W(g_c). \tag{5}$$

Equation (5) agrees with the known exact results of ref. [2] when appropriate $x(g)$ from the exact solution is used. In particular one verifies that $G_{\mathrm{z}} - G_{\mathrm{u}} = \frac{1}{2}\chi(g^2 - g_c^2)$, in the linear response regime (near melting).

Compared to the stretched unzipped state, the zipped phase has to pay a cost $W(g)$ for force expulsion for not following the force-diktat, but gains energy $W(g_c)$ due to binding or ordering. The phase coexistence requires a perfect compensation of one by the other. This compensation may be used to obtain the binding energy of the zipped phase as

$$E_{\mathrm{z}}(T) = W(g_c). \tag{6}$$

This equation may also be used to define g_c from the binding energy.

Let us now consider an inhomogeneous situation of a dsDNA at $T < T_c$ by pulling at one end by a force $g = g_c(T)$ so that there is an interface separating the coexisting zipped and unzipped phases. The interfacial energy is obtained by comparing this mixed state free energy with that of a fully unzipped homogeneous state at $g = g_c$. Needless to say that an interface can be created spontaneously if there is a gain in energy in doing so.

Since far from the interface, the Gibbs free energy density is the same in the two phases, the total free energy $\mathcal{G}$ can be written as

$$\mathcal{G} = \int_{-\infty}^{\infty} G_{\mathrm{u}}(T, g_c)\mathrm{d}z + \sigma, \tag{7}$$

where σ is the "surface tension", and z is a contour length measured along the DNA or the strands, the $z = 0$ point being chosen at the point of interface with $z < 0$ as the unzipped phase.

We start with the free energy functional

$$\mathsf{F}_{\mathrm{tot}} = \int_{-\infty}^{\infty} \mathrm{d}z\, \mathcal{F}\{\psi, x\}, \tag{8}$$

whose minimum gives the equilibrium free energy in a fixed distance ensemble. The functional $\mathcal{F}\{\psi, x\}$ can be taken as

$$\mathcal{F}\{\psi, x\} = F_{\mathrm{u}} + F\{\psi\} + \frac{K_\psi}{2}\left(\frac{\partial \psi}{\partial z}\right)^2 + \frac{K_x}{2}\left(\frac{\partial x}{\partial z}\right)^2 + \int_0^x g(\tilde{x})\,\mathrm{d}\tilde{x}, \tag{9}$$

where $F\{\psi\}$ is the free energy of the homogeneous bulk zipped phase with reference to the unzipped-state free energy F_{u}. In the unzipped state $F\{\psi\} = 0$. K_ψ and K_x are additional "elastic" constants for distortions in ψ and x. The elastic part of the free energy can be extended to torques. The order parameter ψ and force $\tilde{g}$ are not

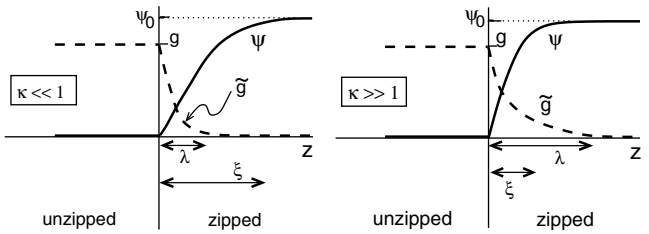

Fig. 1: Schematic diagram of variation of zipping-unzipping order parameter ψ (continuous line) and applied force g (dashed line) inside unzipped and zipped phases. ξ is the length scale of variation of ψ inside the zipped phase and λ is the scale for g. For Type I (left figure), $\kappa = \lambda/\xi \ll 1$ and for Type II (right figure), $\kappa \gg 1$.

coupled in the free energy in the form taken in eq. (9) and consequently, this form is valid only in extreme limits. Further generalizations are not needed for this paper. The Gibbs free energy is obtained from eq. (9) by using the equilibrium values of ψ and x, followed by a Legendre transformation from x to g.

The equilibrium conditions, obtained by minimizing $\mathsf{F}_{\rm tot}$, are

$$\frac{\delta F}{\delta \psi} - K_\psi \frac{\partial^2 \psi}{\partial z^2} = 0, \tag{10}$$

$$-K_x \frac{\partial^2 x}{\partial z^2} + \frac{x}{\chi} = 0, \tag{11}$$

with the condition that

$$\psi = 0, \quad x = x_c = \chi g_c \quad \text{at} \quad z = 0, \tag{12}$$

and

$$\psi = \psi_0, \quad x = 0 \quad \text{at} \quad z \to \infty, \tag{13}$$

ψ_0 being the solution of

$$\frac{\delta F}{\delta \psi} = 0$$

to maximize the interfacial energy. The length scales ξ and λ, giving how fast ψ or $\tilde{g}$ grow or decay inside the zipped phase (see fig. 1), come from eqs. (10) and (11), as

$$\xi^{-2} = \frac{1}{K_\psi} \left(\frac{1}{\psi} \frac{\partial F}{\partial \psi} \right) \Big|_{\psi \to 0}, \text{ and } \lambda^2 = K_x \chi. \tag{14}$$

The equation for λ reduces to the form derived by de Gennes [5] if the elastic constants of his model are used for K_x and χ. The dimensionless ratio $\kappa = \lambda/\xi$ is expected to be different for different sequences of DNA.

For $\kappa \ll 1$, the external force penetrates only a short distance λ into the zipped region. In contrast the order parameter rises to its asymptotic value ψ_0 in a much larger length ξ. One has to pay the energy cost for the damage in ordering over a length scale ξ, and therefore,

$$\sigma \sim E_z \xi \sim \frac{1}{2} \chi g_c^2 \xi, \tag{15}$$

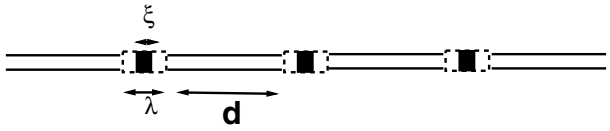

Fig. 2: Schematic diagram of a periodic array of defect blobs. The array has a periodicity d which controls the density of blobs. Each distorted region is of length $\sim \lambda$ with an unzipped core of size $\sim \xi$.

which is positive. This is the conventional scenario of force expulsion of various models on the zipping-unzipping phase transition and this scenario gives the well-known behavior of the unzipping transition.

When $\kappa \gg 1$, the force penetrates a greater distance λ into the sample, so that there is an obvious gain in the stretching energy (*i.e.* reduction of the "positive energy" for force expulsion) over the interval of penetration, over and above the gain by ordering. With x from eq. (11), eq. (7) gives

$$\sigma = -\frac{\chi g_c^2}{2} \lambda, \tag{16}$$

which is negative. Hence, it is possible to lower the free energy of the DNA by creating the interface. The value of κ for transition from Type I to Type II depends on the form of χ which, in turn, depends on the DNA sequence and the secondary structure. It is therefore primarily the sequence but also the secondary structure that determine whether a DNA would behave like Type I or II.

If we now consider the bulk zipped state with a force g, then force penetration may be possible in the form of many isolated distorted regions or blobs. For $\lambda \gg \xi$, with the unzipped core of size ξ costing an energy $E_z \xi$, and the x part of the free energy $\mathcal{F}\{\psi, x\}$ in eq. (9), one finds for a homogeneous chain that a periodic structure of the blobs [7], as in fig. 2, is possible energetically, if $g > g_c/\sqrt{\kappa}$. The initial penetration of force is at $g_{c1} = g_c/\sqrt{\kappa}$ with periodicity $d \to \infty$. The unzipping transition therefore becomes continuous in contrast to the first-order nature for Type I.

The negative interfacial energy is found in Type II superconductor [14] too. Our formulation is similar to that of Type II superconductivity in a one-dimensional geometry. As there is indeed a phase transition in DNA, the Landau theory is justified here. It suffices for a one dimensional case to consider a scalar order parameter.

We may point out a few additional implications of a negative interfacial energy. The penetration of the force is not possible in the conventional polymer models. For any helical or twisted pairs of strings, a pulling force produces over-winding. We expect this over-winding in DNA to be present at the interface, distorting but not vitiating the ordered state. The resulting distortion plays a role in determining the interfacial energy. The penetration of force is via a denatured core of size ξ, surrounded by such a distorted region of size λ. These defect blobs could

be pinned by certain sequences, thereby localizing them in specific regions of the DNA. We speculate that the regions which localize the defect blobs are the non-coding promoter regions. This gives a topological interpretation of the defect blob and it would also be applicable to torque. The existence of the mixed or Type II phase with pinned defect blobs will affect the melting profile under a force, and the force-distance isotherm will show steps originating from the blobs, especially for finite chains. Our analysis shows that the relation between ordering and unzipping is needed to get a negative interfacial energy. The helical ordering is not just base-pairing —it involves stacking and other distant neighbor interactions. Any microscopic model for Type II DNA would have to take these into account. On the experimental front, it is time for a second-generation single molecular experiments that would explore the interfaces on DNA.

To summarize, in this paper we showed that different types of phenomena happen for two regimes of the ratio $\kappa = \frac{\lambda}{\xi}$ of the independent length scales ξ and λ, of DNA order parameter (ψ) and internal force ($\tilde{g}$), respectively. For $\kappa \ll 1$, the interfacial energy is positive, and the unzipping or melting under a force is first order. The external force has no effect inside the ordered or zipped phase, *i.e.*, there is no internal force ($\tilde{g}$) inside as λ is small. This is named Type I. On the contrary, for $\kappa \gg 1$, the interfacial energy becomes negative and the force penetrates the zipped phase in the form of defect blobs. The creation of interfaces are energetically favored, so that interfaces are formed spontaneously. Thus defect blobs are formed inside the ordered phase. Above a force threshold $g > g_{c1}$, there will be a finite density of these defect blobs. The melting under tension of unzipping is second order. This case is named Type II.

REFERENCES

[1] Bhattacharjee S. M., *J. Phys. A*, **33** (2000) L423; cond-mat/9912297.
[2] Marenduzzo D., Bhattacharjee S. M., Maritan A., Orlandini E. and Seno F., *Phys. Rev. Lett.*, **88** (2002) 028102.
[3] Kumar S. and Li M. S., *Phys. Rep.*, **486** (2010) 1.
[4] Gotoh O., *Adv. Biophys.*, **16** (1983) 1.
[5] de Gennes P., *C. R. Acad. Sci. - Ser. IV - Phys.*, **2** (2001) 1505.
[6] Watson J. D. *et al.*, *Molecular Biology of the Gene*, 6th edition (Pearson Education, Singapore) 2008.
[7] Goodman B. B., *Rev. Mod. Phys.*, **36** (1964) 12.
[8] Cao X. Q., Zeng J. and Yan H., *Phys. Rev. E*, **77** (2008) 041908.
[9] Choi C. H. *et al.*, *Biophys. J.*, **95** (2008) 597.
[10] Mameren J. van *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **106** (2009) 18231.
[11] Marenduzzo D., Orlandini E., Seno F. and Trovato A., *Phys. Rev. E*, **81** (2010) 051926.
[12] Randall G. L., Zechiedrich L. and Pettitt B. M., *Nucleic Acids Res.*, **37** (2009) 5568.
[13] Mueller-Planitz F. and Herschlag D., *Nucl. Acids Res.*, **35** (2007) 3764.
[14] Abrikosov A. A., *Sov. Phys. JETP*, **5** (1957) 1174.