Sample-to-Sample Fluctuations in Heterogeneous DNA

ZH. S. GEVORKIAN,[1,2,3] MAI SUAN LI,[4,5] CHIN-KUN HU[1,6]

[1]Institute of Physics, Academia of Sinica, Nankang, Taipei 11529, Taiwan

[2]Institute of Radiophysics and Electronics, Theory Division, Ashtarak-2,0410, Armenia

[3]Yerevan Physics Institute, Theoretical Department, Alikhanian Br.2,0036, Yerevan, Armenia

[4]Institute of Physics, Polish Academy of Sciences, Al.Lotnikow 32/46,02-668 Warsaw, Poland

[5]Saigon Institute for Computational Science and Technology VNU-ITP, 6 Quarter, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam

[6]Center for Nonlinear and Complex Systems and Department of Physics, Chung Yuan Christian University, Chungli 32023, Taiwan

Received 22 March 2010; revised 16 July 2010; accepted 11 August 2010
DOI: 10.1002/polb.22142
Published online 29 September 2010 in Wiley Online Library (wileyonlinelibrary.com).

ABSTRACT: We studied the structure of heterogenous DNA in the native state. There are two different regimes in the sample-to-sample fluctuations of the free energy in the native state, which can be interpreted via the concept of local free energy of base pairs. In the first low-temperature frozen regime, local free energies are random and there are large sample-to-sample fluctuations for short DNAs. In the high-temperature molten regime, the weakly bounded base pairs are opened and do not give random contribution to the free energy of native DNA. As a result, sample-to-sample fluctuations are suppressed in the molten regime. © 2010 Wiley Periodicals, Inc. J Polym Sci Part B: Polym Phys 48: 2432–2436, 2010

KEYWORDS: calculations; heterogeneous polymers; thermodynamics

INTRODUCTION

The secondary structures of DNA and RNA are determined by many factors: base pairing, base stacking, interaction with solvent, electrostatic interaction, and so forth. One of the most important common phenomena for RNA and DNA is the denaturation or melting of double-stranded helices upon changes in ambient temperatures or solvent conditions.[1–5] In the DNA case, most of the theoretical efforts were addressed to the role of heterogeneity in the melting; for a recent review, see for example ref. 6 and references therein.

Under normal physiological conditions, the secondary structure of RNA consists of helical segments and loops of various sizes. This structure is flexible, because each base of RNA has more than one binding partner. Thus, one expects to see a fine structure in the native state of RNA. As shown by recent studies,[7] there is a phase transition between a molten native state of RNA, where the secondary structure can fluctuate, and a glassy native state, where the secondary structure is essentially frozen.

The native state of DNA is clearly different. Each base on one strand has predominantly one binding partner (complementary base) on the other strand. Therefore, one would expect that the secondary structure of DNA does not fluctuate similar to RNA.

The purpose of this article is to show that the native state of heterogeneous DNA nevertheless has a fine structure related to sample-to-sample fluctuations of the free energy (and other thermodynamical quantities). For higher temperatures, these sample-to-sample fluctuations are suppressed while they are large at lower temperatures implying that the secondary structure of a sufficiently short DNA will have its own, original behavior, different from the averaged (over all samples) behavior. In high-temperature molten regime, only local free energies of strongly coupled base pairs give contribution to the free energy, therefore, free energy is less random and sample-to-sample fluctuations are suppressed. Intermediate states in heterogeneous DNA melting have been also discussed in refs. 8–10.

INITIAL RELATIONS

The Hamiltonian of the system is defined as follows:

$$H=\sum_{n=1}^{N}\frac{K}{2}(y_{n+1}-y_n)^2+V_{n+1}(y_{n+1})\equiv\sum_{n=1}^{N}H_n(y_n), \qquad (1)$$

where y_n is the component of the relative displacement of monomers along the hydrogen bond, and V_n is the pairing interaction between monomers from the first and the second

Correspondence to: ZH. S. Gevorkian (E-mail: gevork@yerphi.am) or C.-K. Hu (E-mail: huck@phys.sinica.edu.tw)

strands. We assume that cyclic boundary conditions $y_{N+1}=y_1$ are fulfilled. This Peyrard–Bishop (PB)[11] type model was used for investigation of denaturation transition of heterogeneous DNA.[12] We will use it for investigation of thermodynamics of DNA in a naturated state.

The Hamiltonian of eq 1 does not consider base-stacking interaction that plays a stabilizing role in DNA, see for example, ref. 13. Note that heterogeneity is mainly associated with the base-pairing interaction. Base-stacking interaction is less heterogeneous, see for example, ref. 14. For investigation of sample-to-sample differences, base-stacking interaction can be included into an effective base-pairing interaction (see below).

Equation 1 implies that the partition function of the system can be written as

$$Z=\int\prod_{n=1}^{N}dy_n\exp[-\beta H(y_n)], \tag{2}$$

where $\beta=1/T$. Let us introduce the partition function of a DNA segment.[12]

$$\Phi_n(y_n)=\int\prod_{k=1}^{n-1}dy_k\exp\left[-\sum_{k=1}^{n-1}\frac{K}{2T}(y_{k+1}-y_k)^2-\frac{V_{k+1}(y_{k+1})}{T}\right]. \tag{3}$$

Then the partition function of the overall system becomes

$$Z=\lim_{N\to\infty}\int dy\Phi_N(y). \tag{4}$$

Using eqs. 2 and 3, one can find a transfer integral equation for $\Phi_n(y)$

$$\int dy'\exp\left[-\frac{K}{2T}(y'-y)^2-\frac{V_{n+1}(y)}{T}\right]\Phi_n(y')=\Phi_{n+1}(y). \tag{5}$$

Here $\Phi_n(y)$ has a physical meaning of "wave function" of the nth monomer pair.[11,12] The recursive eqs. 5 can be represented in the differential form by taking into account that in the integral the main contribution gives the values of y' close to y

$$\Phi_n(y)+\frac{T}{2K}\frac{d^2\Phi_n(y)}{dy^2}=\exp\frac{V_{n+1}(y)}{T}\Phi_{n+1}(y). \tag{6}$$

Transition from integral equation (eq 5) to differential equation (eq 6) is correct provided that wave function is slowly varying on the scale $\sqrt{T/2K}$, which is actually the average distance between strands. It is obvious from the physical reasons that the aforementioned condition will be satisfied for a wide range of temperatures. We will investigate eq 6 in two extreme regimes below.

FROZEN STATE

First, consider the low temperature $T\to 0$ limit. Let us assume that the potential energy of base pairing interaction $V_n(y)$ is a slow varying function of y. Neglecting the derivatives of $V(y)$,[15,16] and in the limit $T\to 0$, we can write eq 6 as

$$\exp\left[\frac{T}{2K}\frac{d^2}{dy^2}-\frac{V_{n+1}}{T}\right]\Phi_n(y)=\Phi_{n+1}(y). \tag{7}$$

Assume that the varying pairing potential has the form $V_n(y)=(\lambda+\eta_n)V(y)$, where $V(y)$ is a given potential well with infinite wall at $y=0$, and λ characterizes the strength of the potential, and η_n are random numbers with zero average. An infinite wall restricts the motion of monomers to real positive values of y.[12] We will search solution of eqs. 7 in the form $\Phi_n(y)=c_n\Phi_0(y)$, where $\Phi_0(y)$ is the homopolymeric ground state wave function, which obeys the following equation:

$$\left[-\frac{T}{2K}\frac{d^2}{dy^2}+\frac{\lambda V(y)}{T}\right]\Phi_0(y)=\frac{\varepsilon_0}{T}\Phi_0(y). \tag{8}$$

Using eqs. 7 and 8, one finds recursive equations for coefficients c_n

$$\exp\left(-\frac{\varepsilon_{n+1}}{T}\right)c_n=c_{n+1}. \tag{9}$$

where $\varepsilon_n=\varepsilon_0+\eta_n V$, and V is some characteristic value of $V(y)$. Equation (9) is easily solved

$$c_n=c_1\exp\left(-\sum_{k=1}^{n}\frac{\varepsilon_k}{T}\right). \tag{10}$$

Using eqs. 10 and 4, and neglecting some additive nonessential terms, we find for the free energy

$$F=-TlnZ=\sum_{n=1}^{N}\varepsilon_n. \tag{11}$$

So, we have found that in the low temperature limit, the free energy of DNA is a sum of local free energies. If we assume that random numbers η_n are uncorrelated, $\langle\eta_n\eta_m\rangle=W\delta_{n,m}$ the same will be correct for local free energies. Taking into account that $\delta\varepsilon_n\sim\eta_n$, and $\langle\eta_n\rangle=0$, one can show that the average free energy equals to its homopolymeric value $\langle F\rangle=N\varepsilon_0(T)$ and sample-to-sample fluctuations equal to $\langle\delta F^2\rangle=WV^2N$. Note that similar to many other disordered systems that contain quenched randomness, one should average free energy rather than partition function, see for example, ref. 17. Remind that V characterizes the typical value of pairing potential, and W characterizes its variance. Relative fluctuations of free energy $\sqrt{\langle\delta F^2\rangle}/F^2=\sqrt{W}V/\varepsilon_0\sqrt{N}$ are small in the thermodynamical limit $N\to\infty$. However for small N they can be important. In DNA frozen state, like many other disordered polymeric systems, those conformations are relevant that ensure minimum energy.[18] The base pairs are locked in the minimums of the pairing potential.

In the next section, we will consider the high-temperature state of the heterogeneous DNA.

MOLTEN STATE

Assuming that $\frac{V_n}{T}\ll 1$, eq 6 can be written in the form

$$\frac{T}{2K}\frac{d^2\Phi_n(y)}{dy^2}=\Phi_{n+1}(y)-\Phi_n(y)+\frac{V_{n+1}(y)}{T}\Phi_{n+1}(y). \tag{12}$$

It is convenient to go to continuous variable t instead of n

$$\Phi_{n+1}(y)\to\Phi(y,t),\quad V_{n+1}(y)\to V(y,t), \tag{13}$$

$$\Phi_{n+1}(y)-\Phi_n(y)\to\frac{\partial\Phi(y,t)}{\partial t}. \tag{14}$$

Substituting eqs 13 and 14 into eq 12 and neglecting the small term $T/2K\frac{d^2}{dy^2}\frac{\partial\Phi}{\partial t}$,[19] one obtains

$$\frac{\partial\Phi(y,t)}{\partial t}-\frac{T}{2K}\frac{d^2\Phi(y,t)}{dy^2}+\frac{V(y,t)}{T}\Phi(y,t)=0. \tag{15}$$

So, formally, we obtain the KPZ (Kardar–Parisi–Zhang) equation[20] that describes surface growth and the directed polymers in random media problems as well.[21,22] The important difference of the DNA case from the general KPZ case is that the randomness of the potential $V(y,t)$ is only on variable t defining the number of a base pair and the average is nonzero and negative. These differences lead to different behavior of free energy fluctuations. As aforementioned, the potential $V(y,t)$ is assumed in the form $V(y,t)=(\lambda+\eta(t))V(y)$, where $V(y)$ is a given short-ranged potential, and $\eta(t)$ is a random noise with zero average.

We will search the solution of eq 15 in the form

$$\Phi(y,t)=\exp\left[-\frac{1}{T}\int_0^t\varepsilon(\tau)d\tau\right]\phi(y,t), \tag{16}$$

where $\phi(y,t)$ weakly depends on t. Substituting eq 16 into eq 15 and neglecting derivative of $\phi(y,t)$ on t, one obtains an eigenvalue problem

$$\left[-\frac{1}{2\beta^2K}\frac{d^2}{dy^2}+(\lambda+\eta(t))V(y)\right]\phi(y,t)=\varepsilon(t)\phi(y,t). \tag{17}$$

Substituting (16) into (4) and neglecting inessential terms, we have

$$F=\int_0^t\varepsilon(\tau)d\tau. \tag{18}$$

Thus, the free energy of DNA is expressed via $\varepsilon(t)$.

Equation 18 suggests that $\varepsilon(\tau)$ plays a role of local free energy. In the homopolymeric case $\eta\equiv0$, it does not depend on τ, and free energy is determined as $F=N\varepsilon_0$. The denaturation transition appears when $\varepsilon_0(T_c)=0$.[11,23]

To investigate the spectrum of Scrödinger equation appearing in eq 18, one should specify the form of attraction potential $V(y)$. For the sake of simplicity, we choose it as the delta-shell with infinite wall at origin,

$$\begin{aligned}V(y)&=-1/y_0\delta(y-y_0),\ y>0\\ V(y)&=\infty,\ y\le0.\end{aligned} \tag{19}$$

Here y_0 plays a role of attraction radius of potential. The contact interaction potential was used in refs. 24 and 25 for surface growth and directed polymers in random media problems. Here, we use it to study the heterogeneous DNA. The quantum-mechanical problem with δ-shell potential can be exactly solved,[16,26] and the energy of bound state is determined as follows

$$\varepsilon(t)=-\frac{T^2}{8Ky_0^2}\ln^2 z_0(t), \tag{20}$$

where $z_0(t)$ is the solution of the following transcendent equation

$$\ln z(t)=2\beta^2K(\lambda+\eta(t))(z(t)-1). \tag{21}$$

We will also need the wave function of the ground state in the delta-shell potential for estimation of the neglected term. It has the following form

$$\begin{aligned}\phi(y,t)&=\frac{1}{2\mathcal{N}^{-1/2}}[e^{-k|y-y_0|}-e^{-k(y+y_0)}];\\ k(t)&=\frac{\sqrt{2K|\varepsilon(t)|}}{T}\end{aligned} \tag{22}$$

with the following normalization constant

$$\mathcal{N}=\frac{1}{4k}[1-e^{-2ky_0}(1+2ky_0)]. \tag{23}$$

Now, we can estimate the neglected term $\partial\phi/\partial t$ and clarify conditions of applicability of adiabatic approximation. It follows from eqs 22 and 23 that $\phi(y_0,t)\sim\sqrt{k(t)}$, therefore we have

$$\frac{1}{\phi(y_0,t)}\frac{\partial\phi(y_0,t)}{\partial t}\sim\frac{1}{k(t)}\frac{\partial k(t)}{\partial t}\sim\frac{1}{\varepsilon(t)}\frac{\partial\varepsilon(t)}{\partial t}. \tag{24}$$

So, the ratio of neglected and hold terms is of order of ratio of energy difference of base pairs because of heterogeneity and energy itself. At high temperatures, main contribution to the local free energy of a base pair gives entropic terms, which are the same for different base pairs; therefore, the local free energy difference between different base pairs is small, and the latter ratio, except the region of small temperatures, elsewhere is much less than unity. Thus, the adiabatic approximation we used is justified in high temperature regime. At very low temperatures, in frozen state, the adiabatic approximation fails because $d\varepsilon(t)/dt\sim d\eta(t)/dt$ and $d\eta(t)/dt$, for random $\eta(t)$, can be large even for small $\eta(t)$.

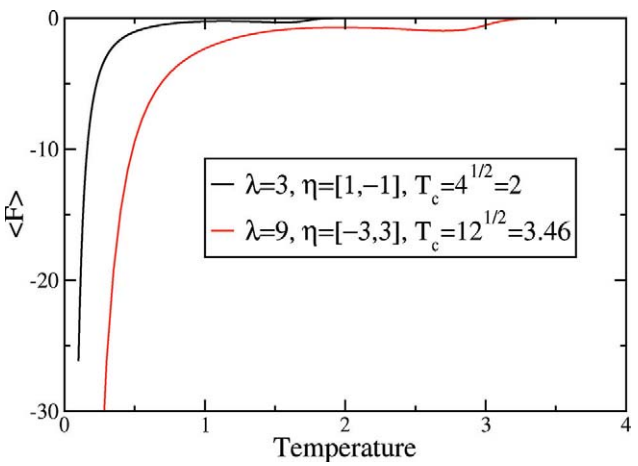

FIGURE 1 Average free energy. We take $2K\equiv4y_0^2\equiv1$, see eqs 19 and 20, for simplicity and $t\equiv1000$. Critical melting temperature of DNA is found from the equation $\langle F(T_c)\rangle=0$. Plots for two different set of parameters ar shown. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

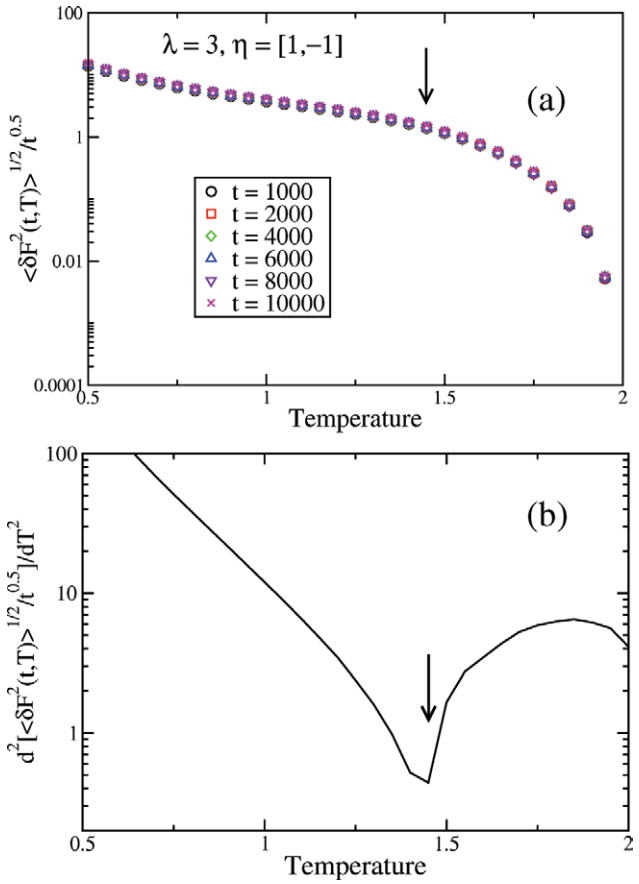

FIGURE 2 (a) Fluctuation of free energy normalized by the number of DNA base pairs for $\lambda = 3$ and $\eta \in [-1, 1]$. (b) Second derivative of free energy sample-to-sample fluctuation has a peculiarity at crossover temperature marked by arrow. Crossover temperature corresponds to the partial opening of base pairs.

A bound state solution of eq 21 exists provided that $2\beta^2 K(\lambda + \eta(t)) > 1$. At temperatures $2\beta^2 K(\lambda + \eta(t)) \leq 1$, a base pair is opened, and local free energy according to eqs 20 and 21 is zero independent of random binding constant $\eta(t)$ and temperature. The critical temperature of homopolymeric ($\eta(t) \equiv 0$) DNA denaturation is determined from this condition $T_c = \sqrt{2K\lambda}$. In the heterogeneous case, within adiabatic approximation, critical temperature is determined by the maximum bound energy $T_c = \max\sqrt{2K(\lambda + \eta(\tau))}$. For example, when $2K \equiv 1, \lambda = 3$ and for uniform distribution $\eta \in [-1, +1]$, $T_c = 2$, correspondingly when $\lambda = 9$ and $\eta \in [-3, +3]$, $T_c = 3.46$ see Figure 1. These are overestimated values. The actual critical temperature is rather determined by average bound energy of base pairs and does not differ from the homopolymeric value significantly.[12] This is consistent with the experiment data, and similar result was obtained in the wetting problem.[27]

Similar to the frozen state, here also DNA free energy is a sum of local free energies eq 18. However, now these energies are no longer completely random. This leads to important peculiarities in free energy. Using eqs 18–21, one can calculate the average free energy and its fluctuation for a given random function $\eta(t)$ numerically. We have carried out such a calculation for uniformly distributed random function $\eta(t) \in [-1, +1]$ and $\eta(t) \in [-3, +3]$. Choosing a continuous range of $\eta(t)$ instead of just two values $\eta = \pm 1$ enables us to take into account effectively different stacking of base pairs. First, we find that unlike the frozen state $\langle F \rangle \neq F_0$, see Figure 1, therefore, the heterogeneity will contribute to the averages of thermodynamical quantities.

More important is the different behavior of fluctuation of free energy

$$\langle \delta F^2(t, T) \rangle = \int_0^t \int_0^t d\tau d\tau' \langle \delta\varepsilon(\tau)\delta\varepsilon(\tau') \rangle, \qquad (25)$$

where $\delta\varepsilon(\tau) = \varepsilon(\tau) - \langle \varepsilon(\tau) \rangle$.

Figures 2(a) and 3(a) show that fluctuations are strongly suppressed in high temperature molten regime. The reason of such behavior is that local free energies of opened weakly coupled base pairs are determined mainly by entropic factors and not by binding constant and are no longer random. In this model, local free energy of a opened base pair is zero independent of binding constant and temperature. In molten regime, fluctuation depends on the number of base pairs as

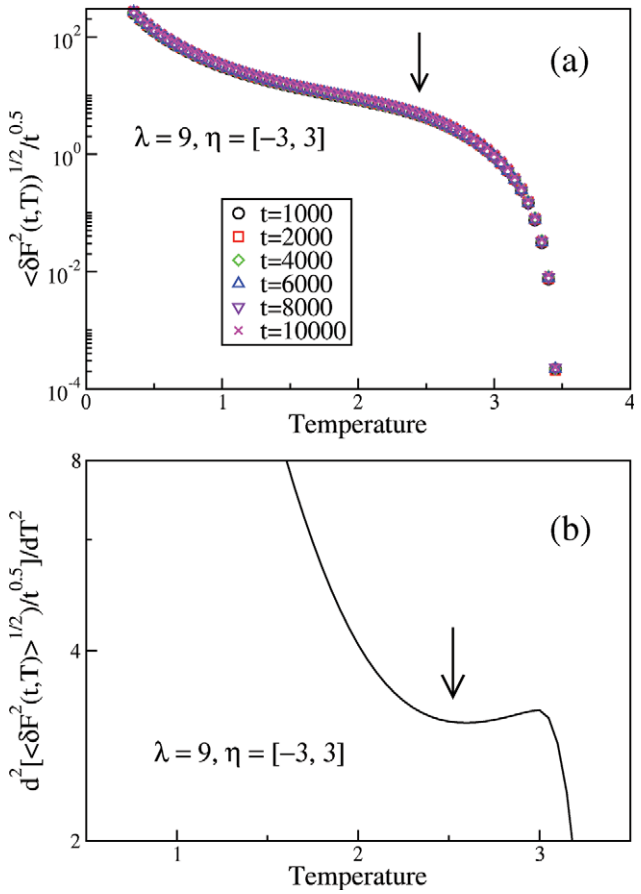

FIGURE 3 (a) Fluctuation of free energy normalized by the number of DNA base pairs for $\lambda = 9$ and $\eta \in [-3, 3]$. (b) Second derivative of free energy sample-to-sample fluctuation has a peculiarity at crossover temperature marked by arrow. Crossover temperature corresponds to the partial opening of base pairs.

$t^{0.5}$ similar to frozen case. However, the prefactor to scaling $\langle \delta F^2(t,T)/t \rangle^{1/2}$ is strongly decreasing function of temperature, see Figures 2 and 3. In frozen state, fluctuations do not depend on the temperature and are relatively large.

At the temperature of partial opening of base pairs, a crossover from one behavior to another occurs. Along with the melting temperature $T_c = \max \sqrt{2K(\lambda + \eta(\tau))}$ DNA heterogeneity introduces a new characteristic temperature $T_{cr} \sim \min \sqrt{2K(\lambda + \eta(\tau))}$. In second plot of Figure 2, $\lambda = 3$ and $\eta = [-1, 1]$, therefore, one obtains the following values for crossover and melting temperatures $T_{cr} \sim 1.4$ and $T_c \sim 2$, respectively. In third plot Figure 3 $\lambda = 9$ and $\eta \in [-3, 3]$, correspondingly $T_{cr} \sim 2.45$, $T_c \sim 3.46$.

DISCUSSION

In conclusion, we have investigated the naturated state of a heterogeneous DNA within a simple theoretical model. Our model is identical to ref. 12 where the influence of heterogeneity on the melting transition was considered. We are considering the sample-to-sample fluctuations in heterogeneous DNA. The heterogeneous DNA problem is similar to the directed polymers in random media. However, in DNA case, potential energy is random only on "time" (number of a base pair) variable, and average potential is a given negative function. For investigation of thermodynamics below the denaturation temperature, an adiabatic approximation was used. One can distinguish two different behaviors within the naturated state. In both regimes, the free energy of the system is represented as sum of local effective free energies of base pairs. However, in frozen state, free energies are random, whereas in molten state, the free energy of open pairs is the same independent of the pair and no longer random. As a result sample-to-sample fluctuations of DNA, free energy are suppressed in the molten regime. Numerically calculating fluctuations for different parameters, we prove that peculiarity is caused by partial opening of base pairs.

This work was supported by the National Science Council of the Republic of China (Taiwan) under Grant Nos. NSC 96-2911-M-001-003-MY3 and NSC 96-2811-M-001-059, National Center of Theoretical Sciences in Taiwan, Academia Sinica (Taiwan) under Grant No. AS-95-TP-A07, and Ministry of Science and Informatics in Poland (grant No 202-204-234). The authors are grateful to A. Allahverdyan and E. Mamasakhlisov for helpful discussions.

REFERENCES AND NOTES

1 Poland, D.; Scheraga, H. A. Theory of Helix-Coil Transitions in Biopolymers; Academic Press: New York, 1970.

2 Wartell, R. M.; Benight, A. S. Phys Rep 1985, 126, 67–107.

3 Morozov, V. F.; Mamasakhlisov, E. S.; Hayryan, S.; Hu, C.-K. Physica A 2000, 281, 51–59.

4 Morozov, V. F.; Mamasakhlisov, E. S.; Grigoryan, A. V.; Badasyan, A. V.; Hayryan, S.; Hu, C.-K. Physica A 2005, 348, 327–338.

5 Mamasakhlisov, E. S.; Hayryan, S.; Morozov, V. F.; Hu, C.-K. Phys Rev E 2007, 75, 061907-1–061907-10.

6 Coluzzi, B.; Yeramian, E. Eur Phys J B 2007, 56, 349–365.

7 Lässig, M.; Wiese, K. J. Phys Rev Lett 2006, 96, 228101-1–228101-4.

8 Badasyan, A. V.; Grigoryan, A. V.; Mamasakhlisov, E. Sh.; Benight, A. S.; Morozov, V. F. J Chem Phys 2005, 123, 194701-1–194701-6.

9 Montrichok, A.; Gruner, G.; Zocchi, G. Euro Phys Lett 2003, 62, 452–458.

10 Cantor, C. R.; Schimmel, P. R. Biophysical Chemistry; Freeman: New York, 1980.

11 Peyrard, M; Bishop, A. Phys Rev Lett 1989, 62, 2755-2758.

12 Cule, D.; Hwa, T. Phys Rev Lett 1997, 79, 2375–2378.

13 Hunter, C. A.; Lu, X.-J. J Mol Biol 1997, 265, 603–619.

14 Frank-Kamenetskii, M. D. Phys Rep 1997, 288, 13–60.

15 Allahverdyan, A. E.; Gevorkian, Zh. S.; Hu, C.-K.; Nieuwenhuizen, Th. M. Phys Rev Lett 2006, 96, 098302-1–098302-4.

16 Allahverdyan, A. E.; Gevorkian, Zh. S.; Hu, C.-K.; Nieuwenhuizen, Th. M. Phys Rev E 2009, 79, 031903-1–031903-15.

17 Pande, V. S.; Grosberg, A. Yu; Tanaka, T. Rev Mod Phys 2000, 72, 259–314.

18 Bryngelson, J. D.; Wolynes, P. G. Proc Natl Acad Sci USA 1987, 84, 7524–7528.

19 Neglected term is of order ε/T. Equation (15) exactly describes homogeneous DNA despite the expansion on ε/T. Therefore, its range of applicability for inhomogeneous DNA is not $T \gg \varepsilon$ but much larger than $T \gg \sqrt{<(\varepsilon(\tau,T) - \varepsilon_0(T))^2>}$.

20 Kardar, M.; Parisi, G.; Zhang, Y. C. Phys Rev Lett 1986, 56, 889–892.

21 Fisher, D. S.; Huse, D. A. Phys Rev B 1991, 43, 10728–10742.

22 Halpin-Healy, T.; Zhang, Y. Phys Rep 1995, 254, 215–414.

23 de Gennes, P. G. Rep Prog Phys 1969, 32, 187.

24 Kallabis, H.; Lässig, M. Phys Rev Lett 1995 75, 1578–1581.

25 Bhattacharjee, S. M.; Mukherji, S. Phys Rev Lett 1993, 70, 49–52.

26 Landau, L. D.; Lifshitz, E. M. Quantum Mechanics; Pergamon Press: Oxford, 1958.

27 Forgacs, G.; Luck, J. M.; Nieuwenhuizen, T. M.; Orland, H. Phys Rev Lett 1986, 57, 2184–2187.