PHYSICAL REVIEW A VOLUME 27, NUMBER 3 MARCH 1983

Strand separation in deoxyribonucleic acid (DNA) molecules

P. Jensen, Marko V. Jarić,* and K. H. Bennemann
Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14,
D-1000 Berlin 33, Federal Republic of Germany
(Received 21 May 1982)

A study of degrees of freedom activated by the strand separation in deoxyribonucleic acid (DNA) is initiated. Thermal effects of activated degrees of freedom are analyzed beyond phenomenological treatments. It is shown that these effects may be observed by simultaneously measuring melting profiles and specific-heat curves of DNA.

Strand separation (melting) in deoxyribonucleic acid (DNA) molecules has been a subject of intense experimental and theoretical investigations.[1-4] The theoretical investigations were dominated by attempts to calculate melting profiles $\Theta(T)$ (i.e., fraction of unwound DNA as a function of temperature) from a given DNA base sequence and, conversely, to extract DNA base sequence from an experimentally observed melting curve. These attempts reached an impressive stage.[4]

Notwithstanding their success current theories of DNA melting are phenomenological and there is a need for further understanding of physical, microscopic origins of their phenomenological content. The aim of this paper is twofold: We wish to initialize a theoretical study of microscopic origins of phenomenological parameters and we wish to point out the needed experiments.

In the simplest description of DNA melting, a DNA molecule is considered a one-dimensional sequence of bound nucleotide base pairs. Each base pair is bound by an effective binding (free) energy ΔF (without a loss of generality we will assume that all base pairs are identical). When ΔF changes sign at the melting temperature T_m, half of the DNA is unwound. An intuitive interpretation of this "transition" is that at T_m a loss in enthalpy per base pair is compensated by a loss in entropy per base pair and a pair has equal probability for being bound or unbound. The transition is not sharp, its width depends on temperature dependence of ΔF and on a correlation energy ΔK between adjacent base pairs.

In order to describe ΔF and ΔK in some detail, we will resort to a simple physical picture. Since a base pair is bound by hydrogen bonds there will be a corresponding contribution ΔF_h to the effective binding energy. An additional contribution ΔF_a arises from an activation of degrees of freedom following breaking of the hydrogen bonds. If we assume a total of N activated degrees of freedom (ADF's) per base pair, then

$$\Delta F_a = -\frac{N}{2} kT \ln\frac{T}{\tau} , \tag{1}$$

where k is the Boltzmann's constant; $k\tau = (\prod_{i=1}^{N} k\tau_i)^{1/N}$ is a local average characterization of ADF's, $k\tau_i = \frac{1}{2} D_i \varphi_i^2$ being a typical energy of ith ADF, D_i its "stiffness constant" and φ_i its characteristic "displacement."

Combining the two above-mentioned contributions, we find

$$\Delta F = \Delta F_h - \frac{N}{2} kT \ln\frac{T}{\tau} . \tag{2}$$

Here we can assume with sufficient accuracy that ΔF_h is temperature independent (a temperature dependence coming from entropy is included in the second term).[5] Furthermore, in obtaining Eq. (2) we have assumed that ADF's do not interact. Such an interaction contributes partially to the correlation energy ΔK. It suffices, for our purpose, to assume ΔK to be temperature independent.

In order to contrast Eq. (2) with phenomenological models we observe that an expansion of Eq. (2) to lowest order in $(T-T_m)$ recovers a phenomenological relation $(kT)^{-1}\Delta F = a(T-T_m)$. Consequently, the main departures of the phenomenological models are expected far from the transition region. Indeed, for both $T \rightarrow 0$ and $T \rightarrow \infty$ the lowest-order approximations lead to incorrect qualitative behavior. However, the numerical value of the deviation is insignificant and it is practically unobservable. As a result we must take a closer look at the transition region.

In this region the most important difference between the models can be estimated by expanding Eq. (2) to the next lowest order in $(T-T_m)$. For example, this leads to a shift of the maximum of

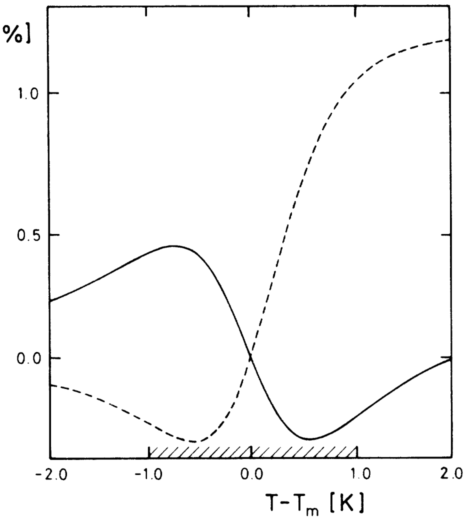

FIG. 1. Function $\delta(T)$ for poly d(A-T)·poly d(A-T) in the present model (full line) and in the phenomenological model (dashed-line) (Ref. 8). Shading indicates the transition region $\sim(\Theta_m')^{-1}$.

$\Theta'(T)\equiv d\Theta/dT$ relative to T_m by only 10^{-2} K [as calculated for poly d(A-T)·poly d(A-T) using the data of Ref. 3]. [We use here the abbreviation poly d(A-T) for poly (deoxyadenine-deoxythymine).] A similar shift occurs in the specific-heat curve. This value puts a limit on precision by which the number of A-T base pairs in a DNA (sub) sequence may be determined to 1 in 10^4.

Apparently, in order to verify our model, physical quantities other than Θ must be considered. For example, temperature dependence of the specific heat or viscosity[6] of DNA solutions may be analyzed. Although the viscosity measurements reveal a reach and intriguing information worth exploring, we will focus on the specific heat which is simpler for theoretical analysis.

We need only discuss a contribution $C(T)$ to the specific heat which arises due to strand separation. Thus, we measure all energies relative to the completely helical DNA. Experimentally, $C(T)$ is obtained by subtracting from total specific heat a contribution extrapolated from below the transition (melting) region.

Few features of $C(T)$ are easily identified. We note that Eq. (2) may be interpreted in terms of an energy spectrum for a single base pair consisting of two states (bound and unbound states) and a continuum (in a phenomenological approach one essentially excludes continuum). Consequently, at low temperatures DNA behaves as a two-state system and we expect a Schottky-type anomaly and a characteristic exponential decay of $C(T)$. On the other hand, at high temperature $C(T)$ corresponds to a system with N noninteracting degrees of freedom (per base pair). Thus, at high T, $C(T)$ should approach $kN/2$. Experimentally, this implies that a subtraction of the specific heat, extrapolated from below melting region, from the high-temperature values should give a constant $kN/2$. Unfortunately, this is unlikely to be observed against a background.

The above conclusions, in conjunction with exact calculations[7] imply that $C(T)$ behaves in a similar fashion as $\Theta'(T)$: They both reach a maximum, C_m and Θ_m', respectively, at approximately T_m and they drop to (approximately) zero at both sides of the transition point.

If curves $C(T)$ and $\Theta'(T)$ are analyzed separately then a difference between a phenomenological and our model (at the most 1–2 % of the peak values) may be absorbed in an adjustment of phenomenological parameters. However, we expect that such adjustments cannot be made *simultaneously* for both $C(T)$ and $\Theta'(T)$. To demonstrate this it is best to consider a function $\delta(T)$ such that

$$\delta(T)=\frac{C(T)}{C_m}-\frac{\Theta'(T)}{\Theta_m'} \, . \qquad (3)$$

In Fig. 1 we present the function $\delta(T)$ calculated for poly d(A-T)·poly d(A-T).[8] It is significant that $\delta(T)$ shows a *qualitatively* different behavior for phenomenological $(KT)^{-1}\Delta F=a(T-T_m)$ and our [cf Eq. (2)] model. For any choice of the parameters $\delta(T)$ qualitatively resembles an odd function of $(T-T_m)$ and its sign is opposite for the two models. An experimental observation of this difference requires a measurement of Θ' and C with a precision better than 1% of Θ_m' and C_m, respectively. This appears feasible.

In conclusion, a few remarks are in order. In this paper we were only concerned with the principal effects of ADF's. For example, we have not treated loop entropy or elasticity effects which may be incluleded in much the same way as in phenomenological theories.[4] Second, we have left for a future work a microscopic identification of ADF's. As a preliminary conclusion, based on energy considerations, we expect ADF's to be associated with rotations around covalent bonds released after base pair breaking. Furthermore, it will be necessary to treat more rigorously an interaction among ADF's. Such an interaction contributes significantly to the correlation energy ΔK.

Apparent success of phenomenological models in describing a relationship between a DNA sequence and its melting curve implies local character of phenomenological parameters. This is in accord with our interpretation of their origin. By giving to phenomenological parameters a more fundamental interpretation we hope to open a door to experimental determinations of microscopic characteristics of DNA molecules. Such characteristics are directly contained in parameter τ and, as we have seen, our parameters may be determined by simultaneous measurement of differential melting and specific-heat curves. We hope that this paper will stimulate necessary microcalorimetric measurements.

We acknowledge partial support from the Deutsche Forschungsgemeinschaft. One of us (M.V.J.) also acknowledges an Alexander von Humboldt research fellowship and the hospitality of Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France, during a part of his work.

---

[*] Present Address: Physics Department, Montana State University, Bozeman, Montana 59717.

[1] See, for example, D. Poland and H. A. Scheraga, *Theory of Helix-Coil Transitions in Biopolymers* (Academic, New York, 1970); A. A. Vedenov, A. M. Dykhne, and M. D. Frank-Kamenetskii, Usp. Fiz. Nauk 105, 479 (1971) [Sov. Phys.—Usp. 14, 715 (1972)].

[2] R. W. Wartell and E. W. Montroll, Adv. Chem. Phys. 22, 129 (1972).

[3] M.L.C. Leung, B. Y. Tong, and F. Y. Wu, Phys. Lett. 54A, 361 (1975).

[4] M. Ya. Azbel, Phys. Rev. A 20, 1671 (1979); Biopolymers 19, 61 (1980), and references therein.

[5] A simplest classical potential which gives Eq. (2) is, for $N=1$, a ΔF_h deep square well of extension φ continued by a harmonic potential characterized by D. Thus, φ is a measure of a difference in the phase (configuration) space occupied by a bound and unbound base pair.

[6] See, for example, J. Eigner and P. Doty, J. Mol. Biol. 12, 549 (1965); R. C. Grant, M. Kodama, and R. D. Wells, Biochemistry 11, 805 (1972).

[7] Fraction of melted DNA $\Theta(T)$, and the specific heat $C(T)$ may be calculated exactly. In the usual manner one establishes a connection with an Ising chain in a (temperature-dependent) external field proportional to ΔF, Eq. (2), and with nearest-neighbor coupling proportional to ΔK.

[8] In concrete calculations we used melting-curve data for poly d(A-T)·poly d(A-T), Ref. 3, Fig. 11. However, for homopolymeric DNA's $\Theta(T)$ is an essentially two-parameter curve (T_m,Θ'_m) and many parameters may fit an experimental melting curve. In this light, our choice $N=8$, $\Delta F_h\approx 2600$ K, $\tau\approx 40$ K and $\Delta K\approx 630$ K, which leads to $a\approx 0.019\ \mathrm{K}^{-1}$, $T_m\approx 319$ K and $U\approx 1.97$ (Ref. 3), should not be taken too seriously. We were lead to this choice by a physical expectation $N\approx 10$, $\Delta F_h\sim 4000$ K, $\tau<\Delta F_h$, and $T_m\sim\Delta K<\Delta F_h$. For further determination of the parameters, a measurement of $C(T)$, i.e., of $\delta(T)$, must be considered.