An Approximate Approach to DNA Denaturation

L. VALKO and P. GERŠI

Department of Physical Chemistry, Faculty of Chemical Technology, Slovak Technical University, Radlinského 9, 81237 Bratislava, Slovakia

Abstract. A quantum-statistical-mechanical theory of equilibrium states was developed to present a simple realistic model of DNA denaturation. Applying the Feynman free energy minimum principle approach to the free energy of an anharmonic oscillator representing an interchain H-bond between DNA strands a simple formula for its melting temperature was derived. Experimental Raman and IR vibrational data were used, and if not available the vibrational data for individual H-bonds in guanine-cytosine and adenine-thymine base pairs were assigned using the Prohofsky self-consistent phonon approach; moreover, the Morse potential parameters were taken over from the same source. On the basis of them the H-bond force constants and then H-bond melting temperatures were calculated for individual H-bonds in the base pairs. The calculated average melting temperatures were compared with the experimental ones for various samples of DNA. Very reasonable results were obtained.

Key words: Hydrogen bond – Morse potential – Feynman free energy minimum principle – DNA denaturation – Melting temperature

Introduction

The dynamics of DNA transcription is among the most fascinating problems of modern biophysics because it is at the basis of life. However, it also is a very difficult problem due to the complex role played by RNA polymerases in the process. It is now well established (Freifelder 1987) that local denaturation of DNA is involved so that it is interesting to investigate the thermal denaturation of the double helix as a preliminary step to the understanding of the transcription. Moreover, studies of molecular dynamics and structure of nucleic acids are of general importance for the understanding of their functions. As such they have been subject of several theoretical and experimental investigations. Because of the importance of the functions played by DNA denaturation in replication, transcription and recombination processes, many theoretical and experimental techniques have been developed to

study melting or strand separation of DNA. One successful theoretical analysis has been the modified self-consistent phonon approach (MSPA), which provides a fundamental understanding of the melting temperatures for a number of DNA homopolymers and copolymers (Gao et al. 1984; Kim et al. 1985; Prabhu et al. 1989).

Let us briefly review the development of experimental work that has been done in recent years. The experimental background comes mainly from Raman and IR spectroscopic results that are an alternative approach (Urabe and Tominaga 1981), more closely connected to theoretical lattice dynamics data (Eyster and Prohofsky 1977). The evaluation of the experimental and theoretical work allowed a conclusion to be made, namely that vibrational normal-mode analysis of infrared and Raman experiments suggested that local melting could be achieved through breathing modes (Prohofsky et al. 1979) as described by MPSA technique which provides an adequate understanding of the melting process based on microscopic motion of atoms.

Thermodynamic characteristics which are insensitive to details of the nucleotide molecular structure but are highly sensitive to the weak connections, or interchain H-bonds between DNA strands, determining the DNA melting temperature, seems to be most attractive in this respect. The breaking of these weak connections depends essentially upon the type of the nucleotide sequence. By mutual comparison, thermodynamic methods are integral methods and are usually not sensitive to the microscopic structure of the thermodynamic system investigated. However, this makes it difficult to reconstruct the microscopic structure from thermodynamic parameters. Fortunately, this general statement is not applicable to DNA. In other words, by a simple change of variables it is possible to obtain the probability of the given composition of a DNA section of a given length from the dependence of DNA integral heat of absorption or optical density on the temperature, and ligand concentration in the solution (Azbel 1973). The extensive thermodynamic instability in the framework of the Ising model including the formation of bubbles in DNA has been studied in theory (Valko 1983).

As can be seen from what has been published so far in the field, the non-linear lattice dynamics model of DNA has recently been the subject of many investigations. It is clear that a realistic model of DNA denaturation and its melting will exhibit non-linear effects owing to the very large amplitude motions known to exist. Exactly this is issue we want to investigate in this paper. However, our approach uses a very simple description of the DNA molecule restricted to minimum possible degrees of freedom.

Formulation of the Working Formula

If the melting of hydrogen bonds is physically due to the instability of H-bonds

resulting from the fluctuation of H-bond stretch at a certain temperature, the important co-ordinate for the melting is the stretch of the H-bond. In this paper we treat the hydrogen-bond stretches in fundamental internal coordinations of DNA responsible for its melting. Given the above considerations, we describe DNA by a Hamiltonian which is similar to the standard form used in molecular mechanics and molecular dynamics studies. Our aim is to explore denaturation using the method of statistical mechanics. It is an elegant way to avoid the problems that occur in a straightforward statistical approach. A system which can be approached this way must have bounded interatomic interaction potentials. The interaction energy must approach zero at large separation. The H-stretch not only is a dynamic element in the melting calculation but it is also used to monitor the melting. We follow Feynman's (1972) free energy minimum principle approach. As a result of the large amplitude of the motions occurring during the transcription or denaturation of DNA, a dynamical model must intrinsically include non-linearities. In our simple model the H-bonds are treated as an asymmetric potential well. Here we want to go further but still keep the model as simple as possible in an attempt to determine the fundamental mechanism of the melting. We use the Morse potential to represent the transverse interaction of the base pairs through H-bonds. However, our discussion below can readily be adapted to any pair potential, the Morse potential however is especially convenient from analytical aspect. The Morse potential has the form (Schroeder and Lippincot 1957)

$$V(x) = V_0 (1 - \mathrm{e}^{-\alpha x})^2 - V_0 \tag{1}$$

Where $x = r - r_0$, α, r_0, V_0, are the Morse constants. V_0 is the depth of the potential well which is determined from the dissociation energy of the H-bond (Prabhu et al. 1990). Expanding $V(\mathrm{x})$ about $x = 0$ in a Maclaurin series we obtain

$$V(x) = \frac{\mu\omega^2}{2} x^2 + \beta x^3 \tag{2}$$

where μ is the reduced mass of H-bond model X-H$\cdots$Y; ω is the angular frequency defined as $\omega = (k_{\mathrm{eff}}/\mu)^{1/2} = (2\alpha^2 V_0/\mu)^{1/2}$ (where k_{eff} is the effective force constant); and β is the first anharmonic correction defined by the relation

$$\beta = \frac{1}{3!} \left(\frac{\partial^3 V}{\partial x^3} \right)_{x=0} = -\alpha^3 V_0 \tag{3}$$

Because of the anharmonicity, when temperature increases and denaturation of DNA is probable the mean position of oscillation H-bonds moves out, as shown in Fig. 1. Bearing this in mind we treat the problem of DNA melting using the Feynman minimum principle written as

$$F \leq F_0 + \langle H - H_0 \rangle_0 \tag{4}$$

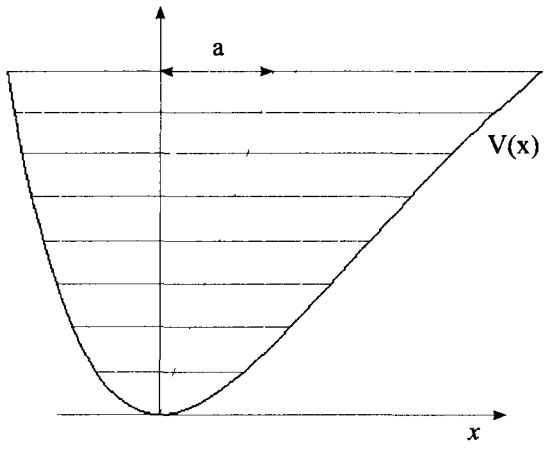

Figure 1. Mean position of oscillation is displaced.

Here, F is the true free energy and H is the corresponding true Hamiltonian. In our dynamic model of denaturation of DNA we take

$$H_0 = \frac{p^2}{2\mu} + \left(\frac{\mu\omega^2}{2}\right)(x-a)^2 \tag{5}$$

which describes a harmonic oscillator with its centre displaced by an amount a. F_0 is then free energy of a system of H_0, for which we can write

$$F_0 = k_B T \left[2\sinh\left(\frac{\hbar\omega}{2k_B T}\right)\right] \tag{6}$$

Our task now is to find

$$\langle H - H_0\rangle_0 = \left\langle \frac{\mu\omega^2}{2}x^2 + \beta x^3 - \frac{\mu\omega^2}{2}(x-a)^2 \right\rangle_0 \tag{7}$$

If the substitution $y = x - a$ is made in Eq. 6, then taking into account that

$$H_0 = \frac{p^2}{2\mu} + \left(\frac{\mu\omega^2}{2}\right)y^2 \tag{8}$$

we can easily calculate the expectation values of the various powers of y. By symmetry,

$$\langle y\rangle_0 = \langle y^3\rangle_0 = 0 \tag{9}$$

and

$$\langle y^2\rangle_0 = \left(\frac{\hbar}{2\mu\omega}\right)\coth\left(\frac{\hbar\omega}{2k_B T}\right) \tag{10}$$

Taking all into account, Eq. 4 yields

$$F \leq F_0 + \langle H - H_0 \rangle_0 = F_0 + \left(\frac{\mu\omega^2}{2}\right) a^2 + \beta a^3 + 3\beta a \langle y^2 \rangle_0 \tag{11}$$

The best estimation of F from Eq. 10 is obtained by minimising the right-hand side. Then, differentiation of Eq. 11 with respect to a leads to

$$3\beta a^2 + \mu\omega^2 a + 3\beta \langle y^2 \rangle_0 = 0 \tag{12}$$

Anharmonic effects of the hydrogen-bond fluctuation are responsible for temperature H-bonds instability near the melting temperature. Since a is a complex number quantity at the DNA melting point, of the two solutions of the initial equation (12) the one that satisfies this requirement is

$$\left(\frac{\mu\omega^2}{6\beta}\right)^2 \leq \langle y^2 \rangle_0 \tag{13}$$

or with respect to Eq. (10)

$$\left(\frac{\mu\omega^2}{6\beta}\right)^2 \leq \frac{\hbar}{2\mu\omega} \coth\left(\frac{\hbar\omega}{2k_B T}\right) \tag{13'}$$

At this type of H-bond stretching motion the system – DNA is in unstable equilibrium and therefore DNA denaturates. At the temperature at which this occurs there is a breakdown of DNA stability, as reflected in a sudden divergence of oscillator centre displacement or sudden reduction in the interbase H-bond force constants this is taken as the melting temperature of a given H-bond in the DNA. The melting temperature T_m which it occurs is

$$T_m = \frac{\hbar\omega}{k_B \ln[(\Gamma + 1)/(\Gamma - 1)]} = \frac{\hbar\omega}{2k_B \operatorname{arcth} \Gamma} \tag{14}$$

where

$$\Gamma = \frac{2\mu\omega}{\hbar} \left(\frac{\mu\omega^2}{6\beta}\right)^2 \tag{15}$$

In DNA, the most significant structural changes occur in the interbase H-bonds. This is because these bonds are the weakest in the system. As a result, the vibrational normal modes as well as the dynamic stability of these interbase H-bonds are changed. Here we present how this effect can be explicitly incorporated into our approach.

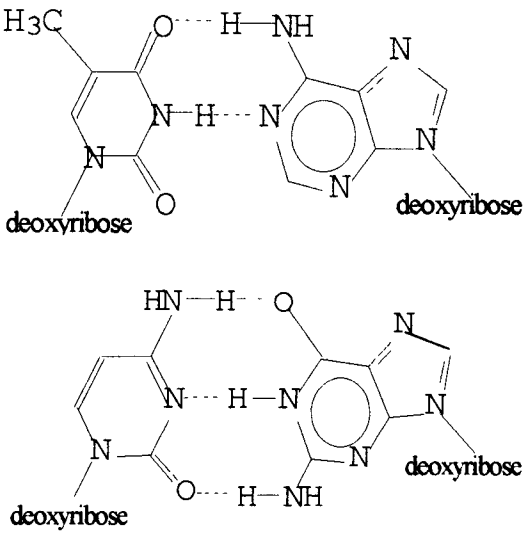

Figure 2. The structures of an adenine-thymine and a guanine-cytosine base pair. The sugar-phosphate backbones are not shown.

Analysis

The DNA molecule is a double helix of two strands of nucleotides (bases and backbone sugar-phosphate group). These two strands are linked together through H-bonds between the complementary bases (Fig. 2). One is an A-T base pair and the other is a G-C base pair. For a single DNA molecule, melting temperature mainly depends on G-C -to- A-T ratio. Poly(dA).poly(dT) has two H-bonds. They are referred to as adjacent to the minor groove N(1)-H···N(3), and adjacent to the major groove N(6)-H···O(4). The three H-bonds in poly(dG).poly(dC) are referred to as the central H-bond N(1)-H···N(3), the H-bond adjacent to the major groove of the molecule O(6)-H···N(4), and the H-bond adjacent to the minor groove N(2)-H···O(2).

Our presented theoretical approach, which only is exact in the limit of strong stacking interactions, shows that the mean stretching of the H-bonds is a complex number quantity when the temperature approaches the melting temperature. This described situation is consistent with thermal denaturation of DNA and it shows that, in spite of its simplicity, the proposed model can respect this basic feature of the DNA molecule. However, the presented statistical thermodynamic approach does not tell how melting occurs: does it start locally as observed experimentally or is it a homogeneous process? As follows from the working formula (14), evaluation of the melting temperature should be done using the vibrational data of the indi-

vidual H-bonds in the A-T and G-C base pairs. However, the melting temperature is extremely sensitive to the vibrational frequency of the H-bonds. To be successful in our estimation of the melting temperature the appropriate choice of the Morse potential parameters V_0, α, r_0 for such a simplified model of DNA is crucial. The calculated and the experimentally obtained parameters used in our simple DNA melting model have to be chosen so as to give realistic properties for the proposed DNA melting model in terms of vibrational frequencies and depths of the potential wells.

The experimental background comes mainly from Raman and IR spectroscopic results. As already mentioned the DNA double helix undergoes strand separation melting in which the interbase H-bonds have to be broken. The motion of the helix involved in this bond disruption are low frequency collective motion of the entire macromolecule. A detailed theory or description of melting of such a system is quite different from that of a small molecule with a much simpler dynamics. The major motion involving the H-bonds is the so-called 'breathing' motion in which the two bases oscillate in opposition, stretching and compressing the H-bonds. The starting point of our approach is to assign experimental or MSPA-calculated vibrational frequencies of Morse potential to the individual hydrogen-bonds in the base pairs. Unfortunately, it is not yet clear what vibrations of the whole DNA system belong to normal modes associated with the bonding hydrogen atoms. Having this fact in mind we rely on the only fact that the frequencies of H-bonds X-H$\cdots$Y depend sensitively on the X-H$\cdots$Y distances involved, and it is clear that these frequencies will be strongly temperature dependent, as actually observed (Urabe et al. 1985; Peticolas et al. 1987). It follows from the low frequency Raman spectra of DNA that there is a broad response at 85 cm^{-1} in DNA solution, originating from the characteristic motion of the double helical DNA which disappears when DNA is thermally denaturated (Urabe and Tominaga 1981). This peak was assigned to minor groove N(2)-H$\cdots$O(2) in G-C base pair. It is known that the Raman modes soften and therefore it would be very difficult to experimentally determine the situation of the Raman modes branching. However, we have MSPA vibrational frequency modes in which softening and branching were taken into account. In this respect, the vibrational frequency of 129 cm^{-1} was assigned to the minor groove of the DNA molecule O(6)-H$\cdots$N(4) (Powell et al. 1987). From MSPA force constant data vibration mode of 67 cm^{-1} was assigned to the central H-bond N(1)-H$\cdots$N(3) in G-C base pair (Kim et al. 1985). That adjacent to the minor groove N(1)-H$\cdots$N(3) was assigned the vibration mode of 110 cm^{-1}, and the H-bond adjacent to the major groove N(6)-H$\cdots$O(4) was assigned 144 cm^{-1} in A-T base pair (Powell et al. 1987). From the known vibration modes and the reduced mass effective force constants were calculated using the known relation $k_{\text{eff}} = \omega^2 \mu$. Within the simplified quasidiatomic H-bond model X-H$\cdots$Y, μ represents the effective reduce mass of the two interacting atoms defined as

Table 1. Parameters of the Morse potential function and calculated melting temperature T_m

Base pair	H-bond	μ 10^{-26} kg	$\tilde{\nu}$ cm^{-1}	V_0 10^{-20} J	α 10^{10} m^{-1}	$-\beta$ 10^{20} J.m^{-3}	T_m K
AT	N(1)-H···N(3)	1.2033	110[a]	2.12	1.10	2.85	335
AT	N(6)-H···O(4)	1.2862	144[a]	2.21	1.46	6.92	345
GC	N(1)-H···N(3)	1.2033	67[b]	2.15	0.67	0.64	343
GC	N(4)-H···O(2)	1.2862	129[a]	2.56	1.22	4.62	405
GC	N(2)-H···O(2)	1.2862	85[ac]	2.50	0.81	1.34	399

a) Powell et al. 1987, b) Kim et al. 1985, c) Urabe and Tominaga 1981

$\mu = [(m(X)+m(H))\cdot m(Y)]/[m(X)+m(Y)+m(H)]$. Therefore, the force constants used here are effective force constants in the sense that they should be interpreted as a superposition of the stretch, bend torsion type short-range force, also including some contribution of interhelical interactions (Kim et al. 1985). To complete our evaluation of the DNA melting temperature definite the all parameters of the Morse curve representing H-bond have to determined (Fig. 1). Therefore, at least the first anharmonic correction β defined by relation (3) has to be estimated. To do this we need to know the Morse parameter V_0 and the scaling factor α. As known the localisation of energy in the H-bonds depends on frequency and temperature. Some frequencies contribute more to the H-bond stretches than do others. In this connection one may say that a major disagreement in this regard is the depth of potential V_0 and the dynamic behavior around stable minimum (Techera et al. 1990) on which the values of the Morse potential parameters depend. Because of this the selection of the values of the Morse potential parameters is most important. From the known effective force constants the length scaling factors of H-bonds were determined which are also shown in Table 1. From the given values of V_0 and α the first anharmonic correction was calculated for individual H-bonds in both base pairs (Table 1). The MSPA theory renders a considerable variety of the Morse potential parameter sets for different sets of charges and different fittings for the various other nonbonding interactions. Because of this, we used the slightly modified Prohofsky's self-consistent phonon approach to Morse potential parameters, e.g., the well depths, especially of N(1)-H···N(3), N(6)-H···O(4) in A-T base pair and N(1)-H···N(3) in G-C base pair. The temperature at which H-bond instability becomes manifested was taken as the melting temperature of the H-bonds. Column 6 in Table 1 lists the melting temperatures of the individual H-bonds, calculated using Eq. 14.

The most reasonable explanation of the calculated temperatures for individual H-bonds in both base pairs lay in assuming that the two H-bonded base pairs A-T and G-C contribute unequally to the stability of the helix, thereby causing T_m to

Table 2. Comparison between the calculated and the experimental melting temperatures for various DNA samples.

Bacteria	T_m	$T^a_{\exp}$
Streptococcus pneum.	361	360
Escherichia coli	365	364
Serratia marcescens	369	367
Mycobacterium phlei	371	370
poly(dA).poly(dC)	340	342
poly(dG).poly(dC)	382	383

a) Marmur and Doty 1959

depend on the mean composition of either the whole DNA or of large segments thereof. Marmur and Doty (1959) prepared a number of protein-free samples of DNA of different base compositions from various sources (Table 2). In reality, as has been shown by them the melting temperatures of homopolymeric A-T and G-C base pairs differ by 50 K, so that the H-bond energies of A-T and G-C must be different as already documented. This is not surprising since three H-bonds can form between a G-C base pair and only two between an A-T base pair.

We see that the melting temperatures calculated for individual H-bond in G-C and A-T base pairs differ from the experimental ones (Table 1). There has been a number of works in which theoretically calculated melting profiles of sequenced DNA were compared with experimentally observed ones. However, these theoretical profiles differ significantly from observed profiles for a variety of DNA's (Gotoh and Tagashira 1981).

Since there are various individual H-bonds in DNA the melting point of which differ we can define with little ambiguity (statistical independence) the number averaged melting temperature

$$\langle T_m \rangle = \sum_i \frac{N_i T_{mi}}{N} = \sum_i x_i T_{mi} \tag{16}$$

Here, $N = \sum_i N_i$, and $x_i = N_i/N$ is the fraction of H-bonds which melt at temperature T_{mi}. If H-bonds having any particular melting point are averaged, the number averaged melting temperature is obtained which is essentially the first moment of the number distribution. The fact that Eq. (16) defines the number averaged melting temperatures can be introduced. If $|\Gamma| \gg 1$, then $\ln[(\Gamma+1)/(\Gamma-1)] \approx 2/\Gamma$ and Eq. (14) can be approximated as

$$T_m \cong \frac{k^3_{\text{eff}}}{36 k_B^2 \beta^2} \tag{17}$$

or if we consider $\beta = -\alpha^3 V_0$, then

$$T_m \cong \frac{2}{9k_B} V_0 \tag{18}$$

With a little manipulation, this can be shown to be

$$\langle T_m \rangle \cong \frac{2}{9k_B} \sum_i x_i V_{0i} \tag{19}$$

here, V_{0i} is the depth of the i-th H-bond. Having this result in mind, the linear dependence of the melting temperature on the fraction GC base pairs x_0 can be now obtained directly

$$\langle T_m \rangle = T_{A-T} + (T_{G-C} - T_{A-T}) x_0 \tag{20}$$

Our considerations did not account for the formation of loops. According to observations the loops lead to a decrease in the transition temperature (Crothers 1968). In order to utilise this we calculated the dependence of the denaturation temperature, $\langle T_m \rangle$, on the guanine-cytosine content for various DNA samples. However, our approach permits an estimation of this linear $\langle T_m \rangle$ dependence of DNA for just of one base pair at a time; for A-T this is calculated to be 340 K, and for G-C base pair the value was 379 K. The calculated values of are listed in Tab. 2. ranges from 360 K for DNA from *Streptococcus pneumoniae* to 370 K for that from *Mycobacterium phlei*, in a manner directly proportional to the G-C content. The calculated $\langle T_m \rangle$ values are reproducible to within 1 K for different carefully prepared samples (Marmur and Doty 1959) and hence they may be considered as intrinsic characteristics of total DNA from a given species.

Conclusion

The model presented in this paper can be considered as a direct application of the Feynman minimum principle approach to DNA melting; the presented model describes the dynamics of the H-bonds between base pairs responsible for DNA melting. The predicted melting temperatures of DNA are close to observed values and they rank the melting behavior in the proper order. Based on the obtained results it can be said that the proposed model provides a satisfactory qualitative description of the DNA dynamics. More importantly, our simple approach would serve as a further test to the validity of the MSPA theory. As the temperature increases the H-bonds are expected to become more distended and softer. This is expected to be reflected in certain modes showing dominant H-bond motion frequency falling. We hope that in agreement with experimental results the present

model will enable to describe the fluctuational openings and the formation of bubbles that grow and combine with each other to lead to complete denaturation of the DNA molecule. In a forthcoming paper we shall consider the H-bond vibration dependence on the temperature in the framework of the DNA model presented herein.

Acknowledgements. This work was supported by grant No. 1/1396/94 to L. Valko from Slovak Agency for Science.

References

Azbel M. Ya. (1973): The inverse problem for DNA. Biopolymers **12,** 1591—1609

Crothers A. D. (1968): Calculation of melting curves for DNA. Biopolymers **6,** 1391—1404

Eyster J. M., Prohofsky E. W. (1977): Soft modes and the structure of the DNA double helix. Phys. Rev. Lett. **38,** 371—373

Feynman R. F. (1972): Statistical Mechanics. W. A. Benjamin Inc., New York

Freifelder D. (1987): Molecular Biology. Jones and Bartlett, Boston

Gao Y., Devi-Prasad K. V., Prohofsky E. W. (1984): A self-consistent microscopic theory of hydrogen bond melting with application to poly(dG).poly(dC). J. Chem. Phys. **80,** 6291—6298

Gotoh O., Tagashira Y. (1981): Stabilities of the nearest-neighbour doublets in double-helical DNA determined by the fitting calculated melting profiles to observed profiles. Biopolymers **20,** 1033—1042

Kim Y., Davi-Prasad K. V., Prohofsky E. W. (1985): Self-consistent phonon theory of mean-field hydrogen-bond melting of poly(dG)-poly(dC). Phys. Rev. **32,** 5185—5189

Marmur J., Doty P. (1959): 1. Dependance on composition of the configurational stability of deoxyribonucleic acids. Nature **183,** 1427—1429

Peticolas W. L., Kubasek W. L., Thomas G. A., Tsuboi M. (1987): Biological Applications of Raman Spectroscopy. Wiley & Sons, New York

Powell J. W., Edwards G. S., Genzel L., Kremer F., Wittlin A., Kubasek W., Peticolas W. (1987): Investigation of far-infrared vibrational modes in polynucleotides. Phys. Rev. A **35,** 3929—3939

Prabhu V. V., Young L., Prohofsky E. W. (1989): Hydrogen-bond melting in B-DNA copolymers in a mean-field self-consistent phonon approach. Phys. Rev. B **39,** 5436—5443

Prabhu V. V., Young L., Awati K. W., Zhuang W., Prohofsky E. W. (1990): Defect-mediated hydrogen-bond melting in B-DNA polymers. Phys. Rev. B **41,** 7839—7845

Prohofsky E. W., Lu K. C., Van Zandt L. L., Putman B. E. (1979): Breathing modes and induced resonant melting of the double helix. Phys. Lett. A **70,** 492—494

Schroeder R., Lippincot E. R. (1957): Potential function model of hydrogen bonds II. J. Phys. Chem. **61,** 921—928

Techera M. A., Daemen L. L., Prohofsky E. W. (1990): Comment on DNA solitons with realistic parameters. Phys. Rev. A **42,** 5033—5034

Urabe H., Tominaga Y. (1981): Low frequency Raman spectra of DNA. J. Phys. Soc. JPN. **50,** 3543—3544

Urabe H., Hayashi H., Tomonaga Y., Nishimura Y., Kubota K., Tsuboi M. (1985): Collective vibrational modes in assembly of DNA and its application to biological systems. Low frequency Raman spectroscopy. J. Chem. Phys. **82,** 531—535
Valko L. (1983): Thermodynamic instability in DNA. Gen. Physiol. Biophys. **2,** 499—518

Final version accepted December 12, 1995