# Hydrogen-bond melting in B-DNA copolymers in a mean-field self-consistent phonon approach

V. V. Prabhu, L. Young, and E. W. Prohofsky
*Department of Physics, Purdue University, West Lafayette, Indiana 47907*
(Received 9 September 1988)

We explore hydrogen-bond melting in B-DNA for the copolymers poly(*d*GC)-poly(*d*GC), poly(*d*AC)-poly(*d*GT), poly(*d*AG)-poly(*d*CT), and poly(*d*AT)-poly(*d*AT). Here *d*GC refers to a repeating sequence of a guanine base followed by a cytosine base on one strand. The melting is explored through a mean-field self-consistent phonon theory based on the Green's-function method. By our calculations, the melting temperatures are 385, 366, 357, and 325 K for the four helices, respectively. The onset of melting, in helices with both A-T and G-C base pairs, is either in the A-T pair or in the G-C pair, depending on the effective near-neighbor interaction.

## INTRODUCTION

A copolymer of DNA is a one-dimensional infinite lattice which has for a unit cell two base pairs of DNA. As is well known, the base pairs could be either adenine-thymine (A-T) or guanine-cytosine (G-C). We have four possible copolymers since the unit cell of two base pairs could be composed of a G-C and a C-G pair, an A-T and a C-G pair, an A-T and a G-C pair, or an A-T and a T-A pair. We shall hereafter refer to these helices as GCGC, ACGT, AGCT, and ATAT, respectively. The bases are held together by hydrogen bonds. We explore melting of these hydrogen bonds using a mean-field self-consistent phonon approach (MSPA). Kim *et al.*[1,2] have applied MSPA to the hydrogen-bond melting of a DNA homopolymer having the base pairs G-C for a unit cell. The procedure is in five essential steps.

(1) We find eigenvectors and eigenfrequencies of the lattice with all force constants in the harmonic approximation. (2) From the eigenvectors and frequencies, we find the mean-square stretch amplitudes of the hydrogen bonds. (3) Each hydrogen bond is encased in a Morse potential well, and an effective force constant averaged over the stretches is calculated. (4) The altered hydrogen-bond force constants change the lattice dynamics. New eigenfrequencies are found by the Green's-function method. From perturbation approximations we find the new eigenvectors. (5) We iterate through steps 2–4 at each temperature until the reduction in the hydrogen-bond force constants converges. The melting temperature is taken as that at which the change in hydrogen-bond force constants diverges after a few iterations. We now explain each step briefly.

### Step 1: finding the dispersion relations for a dna lattice

We model the copolymer as a one-dimensional infinite lattice where a unit cell is of two base pairs. On account of helical symmetry the eigenvalues and eigenvectors are obtained from the secular equation[3]

$$|A-\omega^2 I|=0 , \tag{1}$$

$$A(\theta)\equiv B^{+}(\theta)F(\theta)B(\theta) , \tag{2}$$

where $B$ is the transformation matrix between mass-weighted Cartesian coordinates (MWC) and internal coordinates. $F$ is the matrix of force constants of internal coordinates, and $\theta$ is degrees of phase shift from one unit cell of two base pairs to the next. $\theta$ extends over the first Brillouin zone. The dimensionality of the secular equation is equal to the number of degrees of freedom in a unit cell. For a unit cell of two base pairs it is 246.

$$F=F_V+(F_0+F_N)+F_{\mathrm{H}} , \tag{3}$$

where $F_V$ is the matrix of valence force constants within a unit cell, $F_0$ is the matrix of nonbonded interaction force constants between the bases of a unit cell, and $F_N$ is the matrix of long-range interaction force constants between a unit cell and ten neighbors to either side. $F_{\mathrm{H}}$ is the diagonal matrix of hydrogen-bond force constants which are listed in Table I. The valence force constants are refined by us based on spectral data[3] above 400 $\mathrm{cm}^{-1}$. The nonbonded interaction between nearest-neighbor base atoms has an electrostatic term and a Van der Waals term.[4] The interaction is of the form

$$f_{ij}=\left|\frac{2\eta e_i e_j}{r_{ij}^3}-\frac{42A}{r_{ij}^8}\right| , \tag{4}$$

where $r_{ij}$ is in Å and $e_i$ is the net charge in units of electron charge on atom $i$. When $f_{ij}$ is in mdyn/Å, $\eta$ is 2.31. $A$ is 1.85. The atomic positions are from Arnott *et al.*[5]

TABLE I. Hydrogen-bond lengths and force constants at 293 K.

| | Bond | Length (Å) | $k$ $\left[\frac{\text{mdyn}}{\text{Å}}\right]$ |
|---|---|---|---|
| G-C | N(1)—H—N(3) | 2.8491 | 0.210 |
| G-C | O(6)—H—N(4) | 2.8328 | 0.185 |
| G-C | N(2)—H—O(2) | 2.8457 | 0.167 |
| A-T | N(1)—H—N(3) | 2.8807 | 0.142 |
| A-T | N(6)—H—O(4) | 2.9506 | 0.053 |

 

The charges are based on calculations by Miller[6] but reduced by 2.3 based on comparison of these to a few charges determined experimentally by fitting to x-ray scattering data.[4] The nonbonded interactions between adjacent unit cells are taken as force constants

$$f_{ij}=\frac{2\eta e_i e_j}{\sqrt{\epsilon_i}\sqrt{\epsilon_j}r_{ij}^3} . \tag{5}$$

Two dielectric constants are needed as some atoms see each other through DNA while others see each other through water of hydration. Both dielectric parameters are set to 1.0 within a unit cell and fixed at the long-range value beyond 10 Å. The long-range value for DNA was set at 6.0 and for water at 9.0.[7]

The choice of 9.0 for water was guided by several facts. Lee *et al.*[7] indicate that the water dipole orientation time near DNA varies between values corresponding to frequencies of 4–80 GHz. The dielectric constant is thus frequency dependent and only approaches its static value for frequencies well below 4 GHz. We find 9.0 appropriate to calculations in the sub-THz range. This is essential to fit the compressional acoustic mode in the dispersion relations to that obtained in inelastic neutron scattering data on DNA.[8]

The force constants for the hydrogen bonds were found from the Lippincott-Schroeder potential.[9] All these terms are included in the $F$ matrix of Eq. (1) and diagonalized for various values of $\theta$. The above calculation is for ACGT, and it fits experimentally observed neutron scattering data.[10,8] ACGT has the same base content as native DNA with alternating purine-pyrimidine bases on a single strand. The same parameters are used for all four copolymers.

For AGCT, having all purines on one strand and all pyrimidines on another, the Van der Waals term is not included. This was also found necessary in the homopolymer poly($d$G)-poly($d$C) by Kim *et al.*[1] The inclusion of the Van der Waals term causes an anomalous melting temperature discussed in three paragraphs prior to the concluding paragraph of this paper.

## MSPA

We now employ the mean-field self-consistent phonon approach (MSPA), which requires iteration through steps 2, 3, and 4. The procedures in each iteration are as follows.

### Step 2: calculation of mean-square stretching amplitude of the hydrogen bond

The mean-square stretching amplitude for the hydrogen bonds $D_i$ is given by

$$D_i=\langle s_i\cdot s_i\rangle ,$$

where $s_i$ are the internal stretch coordinates of the hydrogen bonds. The calculation proceeds as follows. The mean energy of a simple harmonic oscillator is

$$\langle E_n\rangle=\langle (n+\tfrac{1}{2})\hbar\omega\rangle . \tag{6}$$

The mean-square amplitude of the oscillator, at a given temperature $T$, can be shown to be

$$\langle a_n^2\rangle=\frac{\hbar}{\omega}\coth\frac{\hbar\omega}{2kT} . \tag{7}$$

Our calculations yield eigenvectors whose components give the relative amplitudes of motion. We convert the relative amplitudes to the actual amplitudes. Let $[A^{\lambda}(\theta)]$ be an eigenvector whose components are the actual amplitudes of the mass-weighted Cartesian coordinates in a unit cell. There are $3N$ coordinates, and we have to consider waves traveling in both directions. For a mode at $\theta$ and in band $\lambda$, if $\xi^{\lambda}(\theta)$ is the relative eigenvector which is normalized,

$$[A^{\lambda *}(\theta)][A^{\lambda}(\theta)]=3N\frac{1}{2}\frac{\hbar}{\omega_{\lambda}(\theta)}\coth\frac{\hbar\omega_{\lambda}(\theta)}{2kT}\times[\xi^{\lambda *}(\theta)][\xi^{\lambda}(\theta)] . \tag{8}$$

The stretch internal coordinates can be obtained from the transformation matrix $B$,

$$s_i=\sum_k \xi_k B_{ki} . \tag{9}$$

Using $B$ on both sides of Eq. (8)

$$[s_i^{\lambda *}(\theta)s_i^{\lambda}(\theta)]_A=3N\frac{1}{2}\frac{\hbar}{\omega_{\lambda}(\theta)}\coth\frac{\hbar\omega_{\lambda}(\theta)}{2kT}\times[s_i^{\lambda *}(\theta)s_i^{\lambda}(\theta)]_R . \tag{10}$$

The subscripts $A$ and $R$ are for absolute and relative, respectively. Averaging over all the bands at a given $\theta$ and over all values of $\theta$ for a given band we obtain

$$D_i=\frac{1}{\pi}\sum_{\lambda}\int_0^{\pi}d\theta\frac{\hbar}{2\omega_{\lambda}(\theta)}\coth[\hbar\omega_{\lambda}(\theta)/2kT]\times s_i^{\lambda}(\theta)s_i^{\lambda *}(\theta) . \tag{11}$$

Here, the 13 lowest frequency bands are treated exactly and the other bands are treated in the Einstein approximation, i.e., assumed dispersionless. We have verified that treating more than 13 bands exactly does not in any way alter the results.

### Step 3: force constant calculation for the hydrogen-bond stretching

While all force constants in the lattice are assumed to remain confined to the symmetric regions of their potential wells and hence have an unchanging harmonic force constant until melting temperature, we selectively introduce anharmonicity in the force constants of the hydrogen bonds. The hydrogen bonds are encased in asymmetric Morse potentials (Fig. 1) where they are allowed to increase in length with temperature. An effective force constant averaged over all stretches in the Morse well is calculated. At 293 K this effective force constant is equated to the harmonic force constant of the hydrogen bond found from the Lippincott-Schroeder[9] calculations and used in step 1. It is this effective force constant that breaks down at the melting temperature. To calculate the effective force constant consider the true Hamiltonian

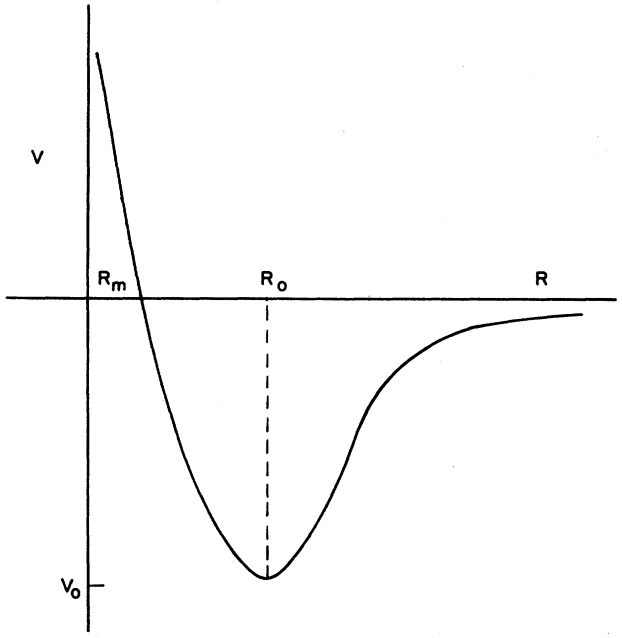


FIG. 1. The Morse potential well. The parameter $a$ is inversely proportional to the square root of the "width."

$$H_T = E_k + V(r_1, \ldots, r_n) , \tag{12}$$

where $E_k$ is the kinetic energy part and $V(r_1, \ldots, r_n)$ is the true many-body potential. This potential has harmonic and anharmonic contributions. Since the final solution is phononlike, one can write an effective Hamiltonian

$$H_E = E_k + \tfrac{1}{2}\phi u^2 , \tag{13}$$

where $\phi$ is an effective force constant. $\phi$ is determined by

$$\frac{d}{d\phi}(F_T - F_E) = 0 , \tag{14}$$

where $F_T$ and $F_E$ are the true and effective free energies, respectively. The free energy $F$ is given by the expression

$$F = -kT \ln \mathrm{Tr}[\exp(-H/kT)] , \tag{15}$$

$$\phi = \frac{\int du\, e^{-u^2/2D} \dfrac{d^2V}{du^2}}{\int du\, e^{-u^2/2D}} . \tag{16}$$

$d^2V/du^2$ is the local force constant for various amounts of displacement. The effective force constant is the weighted average of the force constant, weighted over the statistical range of displacements for a given temperature. At the commencement of the iterations this is set equal to the harmonic force constant of the hydrogen bond at 293 K. In MSPA we set

$$V[d_i(T) + \mu(T)] = V[d_i(T) - \mu(T)] , \tag{17}$$

$d_i(T)$ is the mean equilibrium bond length of the $i$th hydrogen bond. $\mu(T)$ is the amplitude of oscillation of the thermal phonons. Here $\mu(T)$ is set equal to the full width at half maximum of $e^{-u^2/2D_i}$, which is $2\sqrt{\ln 4 D_i}$. For the $i$th hydrogen bond,

$$\phi_i = \frac{\int_{R_m}^{\infty} du\, e^{-u^2/2D_i} \dfrac{d^2V[d_i(T)+u]}{du^2}}{\int_{R_m}^{\infty} du\, e^{-u^2/2D_i}} . \tag{18}$$

The lower limit of the integral $R_m$ is the point where the hard core is struck raising the energy to that at infinity (Fig. 1). The integration can be done in a closed form,[2]

$$\phi = 2a^2 V_0 \frac{A - B}{C} , \tag{19}$$

$$A = 2\{\exp[2a^2 D - 2a(R_T - R_0)]\} \times \mathrm{erfc}\left[\frac{2aD - R_T + R_m}{\sqrt{2}D}\right] , \tag{20}$$

$$B = \exp\left[\frac{a^2 D}{2} - a(R_T - R_0)\right] \times \mathrm{erfc}\left[\frac{aD - R_T + R_m}{\sqrt{2}D}\right] , \tag{21}$$

and

$$C = \mathrm{erfc}\left[\frac{R_m - R_T}{\sqrt{2}D}\right] , \tag{22}$$

where erfc stands for the complementary error function

$$R_m = R_0 - (1/a)\ln 2 , \tag{23}$$

where $R_0$ is determined in the succeeding paragraph.

We approximate the potential for the hydrogen bond as a Morse potential.[2] This is depicted in Fig. 1. The Morse potential has the form[11,9]

$$V(R) = V_0(1 - e^{-a(R - R_0)})^2 - V_0 . \tag{24}$$

The Morse constants $V_0$, $a$, $R_0$ are determined as follows.[12] Using the initial solution of the vibrational modes at 293 K, we can calculate the mean-square stretching amplitude for each hydrogen bond using Eq. (11). From x-ray data we know the length of each hydrogen bond at 293 K. $R_0$ is obtained using Eq. (24) in Eq. (17). If $d_i$ is the equilibrium bond length

$$d_i(293) = R_0 - \frac{1}{a}\ln\frac{2\sinh(a\mu_{293})}{\sinh(2a\mu_{293})} , \tag{25}$$

$$\mu(T) = 2(2\ln 2)^{1/2} D^{1/2} . \tag{26}$$

Here $V_0$ is determined from the dissociation energy of the hydrogen bond.[12,9] Now $a$ remains to be calculated. In step 1 the force constants for the hydrogen bonds were chosen by the use of the Lippincott-Schroeder potential. These are listed in Table I. Using $V_0$ and $R_0$ in Eq. (18) we determine $a$'s that reproduce the force constant.

The parameters $V_0$, $a$, $R_0$ thus calculated at room temperature are kept the same throughout the calculation. The Morse constants are listed in Table II.

TABLE II. Morse potential constants for hydrogen bonds.

| | Bond | $V_0$ $\left[\frac{\text{mdyn}}{\text{Å}}\right]$ | $a$ $\left[\frac{1}{\text{Å}}\right]$ | $R_0$ (Å) |
|---|---|---|---|---|
| GCGC | | 0.025 44 | 2.210 | 2.804 |
| ACGT | N(1)—H—N(3) | 0.011 36 | 4.860 | 2.741 |
| AGCT | G—C | 0.028 13 | 2.184 | 2.783 |
| GCGC | | 0.031 77 | 1.975 | 2.746 |
| ACGT | O(6)—H—N(4) | 0.026 45 | 2.480 | 2.717 |
| AGCT | G—C | 0.033 44 | 1.963 | 2.734 |
| GCGC | | 0.028 17 | 1.950 | 2.771 |
| ACGT | N(2)—H—O(2) | 0.021 86 | 2.468 | 2.736 |
| AGCT | G—C | 0.030 96 | 1.922 | 2.750 |
| ACGT | | 0.012 58 | 3.720 | 2.745 |
| AGCT | N(1)—H—N(3) | 0.019 14 | 2.520 | 2.758 |
| ATAT | A—T | 0.011 15 | 3.275 | 2.788 |
| ACGT | | 0.016 38 | 1.580 | 2.792 |
| AGCT | N(6)—H—O(4) | 0.023 70 | 1.240 | 2.800 |
| ATAT | A—T | 0.017 08 | 1.465 | 2.817 |

**Step 4: calculation of new frequencies and eigenfunctions based on the Green's function**

Using the given force constants and initial solution, new force constants for the hydrogen bond are calculated through steps 2 and 3. Let the force constant change be $C$, where $C$ is an $n\times n$ diagonal matrix. Here $n$ is the number of hydrogen bonds per unit cell; $n$ is 4 for ATAT, 5 for ACGT and AGCT, and 6 for GCGC. The diagonal element of $C$ is

$$C_i=\bar{\phi}_i-\phi_i (i=1,n) , \tag{27}$$

where $\bar{\phi}_i$ is the calculated force constant, and $\phi_i$ is the initially given force constant for the $i$th hydrogen bond. The perturbation $C$, in the force constant matrix $A$ of Eq. (1), which does not break the helical symmetry, modifies the secular equation to

$$[A(\theta)-\omega^2+C]\xi(\theta)=0 . \tag{28}$$

Instead of solving this equation we use the Green's function method.[13,14] If we define the Green's function $g(\omega^2,\theta)$ as

$$g(\omega^2,\theta)=[\omega^2-A(\theta)]^{-1} , \tag{29}$$

then Eq. (28) is equivalent to

$$[1-g(\omega^2,\theta)C]\xi(\theta)=0 . \tag{30}$$

The new eigenfrequencies satisfy

$$\det[1-g(\bar{\omega}_\lambda^2,\theta)C]=0 , \tag{31}$$

where det stands for determinant and $\bar{\omega}_\lambda$ is the new eigenfrequency of band $\lambda$ at $\theta$. In internal coordinates Eq. (30) is reduced to the determinant of an $n\times n$ matrix, where

$$g_{ij}(\omega^2,\theta)=\sum_\alpha \frac{s_i^\alpha(\theta)s_j^{\alpha*}(\theta)}{\omega^2-\omega_\alpha^2(\theta)} . \tag{32}$$

Here $\omega_\alpha$ is the given eigenfrequency of band $\alpha$ at $\theta$, and $s_i^\alpha(\theta)$ is the eigenfrequency of the $i$th hydrogen-bond stretch. The numerical process of finding the new frequencies is as follows.

As the temperature increases the force constants decrease. The decrease in the force constants ensures that the new frequencies are lower than the corresponding older frequencies. That the new frequencies are indeed lower has been meticulously checked at each iteration. The new frequencies $\bar{\omega}_\lambda$'s are then determined by the following method. First, we divide the interval between $\omega_\alpha^2$ and $\omega_{\alpha-1}^2$ into subintervals. $\omega_\alpha>\omega_{\alpha-1}$. We calculate the determinant of Eq. (30) at the boundary of each subinterval. We then multiply the values of the determinant at successive boundaries and pick out the subinterval across which the product is negative. A negative product indicates that the determinant has changed sign within the given subinterval. Then by linear interpolation of the determinant at the subinterval, we find the new frequency $\bar{\omega}_\alpha$.

The new eigenfunctions $\bar{s}_i^\lambda(\theta)$'s are calculated by the perturbation method. In the MWC coordinate, the new eigenfunction $\bar{\xi}_i^\lambda$ is in the perturbation approximation

$$\bar{\xi}_i^\alpha=\xi_i^\alpha+\sum_{\lambda(\neq\alpha)}\sum_{l,k}\frac{\xi_l^{\lambda*}C_{lk}\xi_k^\alpha}{\omega_\alpha^2-\omega_\lambda^2}\xi_i^\lambda . \tag{33}$$

In internal coordinates, this can be shown to be

$$\bar{s}_i^\alpha=s_i^\alpha+\sum_{\lambda(\neq\alpha)}\sum_j\frac{s_j^{\lambda*}C_j s_j^\alpha}{\omega_\alpha^2-\omega_\lambda^2}s_i^\lambda , \tag{34}$$

where $\bar{s}_i^\alpha$ is the new eigenfunction.

**Step 5: iteration**

We define a divergence $\delta$ and continue iteration of steps 2, 3, and 4 at each temperature until

$$\delta=\text{abs}\left[\frac{\prod_{i=1}^n C_i}{\prod_{i=1}^n \phi_i}\right]<(0.1)^n , \tag{35}$$

where $C_i$ is from Eq. (27), $\phi_i$ is the force constant for the $i$th hydrogen bond from the previous iteration, and $n$ is the number of hydrogen bonds per unit cell. We watch for $\delta/\delta_0>1$, where $\delta_0$ is from the previous step and $\delta$ from the current step. At $T=293$ K, the first calculation is carried out to check the self-consistency of the given solution. At each temperature after 293 K, iterations through steps 2, 3, and 4 were performed until self-consistency was established. At the temperatures indicated in Table III, no self-consistent solutions were found.

## RESULTS

The results of this calculation are displayed in Figs. 2–7. We have drawn plots of the mean-square stretching amplitude of the hydrogen bonds as a function of temperature for GCGC and ACGT (Figs. 2 and 3). These are followed by plots of force constants of the hydrogen bonds as a function of temperature (Figs. 4 and 5). The plots of ATAT and AGCT are similar to those of GCGC and ACGT, respectively, and are found in Ref. 12. It

TABLE III. Calculated and observed melting temperatures in K.

| Mol. | Calc. | Expt. |
|---|---|---|
| GCGC | 385 | 372 |
| ACGT | 366 | 350 |
| AGCT | 370 | 344 |
| | 357[a] | |
| ATAT | 325 | 318 |
| Z-GCGC | 425 | ??? |

[a]The melting temperature of 357 K is without the inclusion of the Van der Waals term and all data in tables and figures of AGCT refer to this case. When included, the melting temperature is 370 K.

must be noted that in each plot the two halves are over the same range of temperature. In each figure we use, a solid line for the—N(1)—H—N(3) bond in G-C, a solid line for the—N(1)—H—N(3) bond in A-T, a long dashed line for the—O(6)—H—N(4) bond in G-C, a long dashed line for the—N(6)—H—O(4) bond in A-T, a short dashed line for the—N(2)—H—O(2) bond in G-C.

The behavior of the hydrogen bonds of the two base pairs of a unit cell is depicted in the first and second half of each plot. Thus, the first half of each plot in ACGT depicts the behavior of the hydrogen bonds A-T and the second half of C-G. The first half of each plot in GCGC depicts the hydrogen bond behavior in G-C and the second half, of the C-G pair. The plots of different copolymers are drawn to the same scale to facilitate comparison. In Figs. 2 and 3 we see that the displacement is small for all temperatures below the melting temperature. This indicates that the bond behaves as a restorable spring, undergoing gradual elongation until the melting temperature. At that temperature it ceases to be a reasonable bond. The loss of the Hookian response is indicated by the sudden increase in the values of the mean-square stretch amplitudes and the sudden drop in the values of the force constants.

As the temperature increases, the vibrational mean displacement increases. In the asymmetric potential, the point of return is pushed to regions of weaker restoring forces. The effective force constant is the weighted average over force constants over the entire potential well. As the turn around point is pushed further out, the feeble contributions to the force constant reduce the effective force constant. The mean-thermal displacement in-

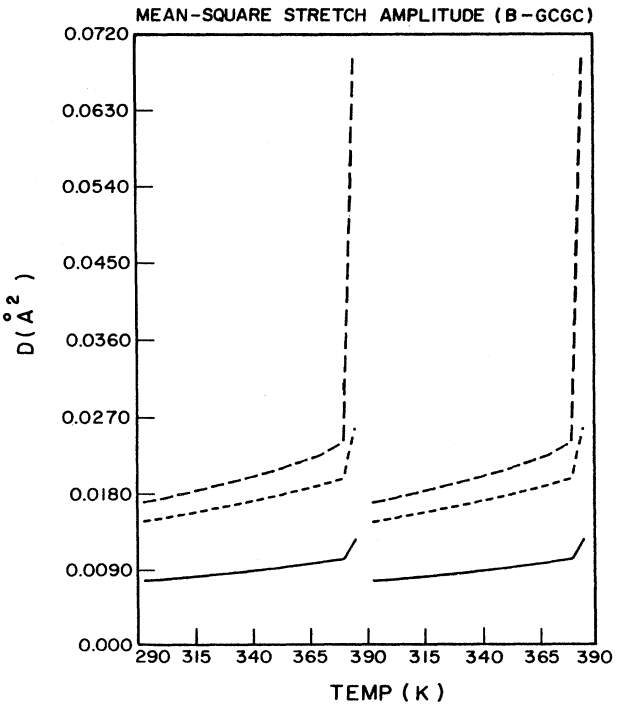


FIG. 2. The mean-square stretch amplitudes of the hydrogen bonds in GCGC as a function of temperature. The three bonds to the left are of the base pair G-C, while the three bonds to the right are of the base pair C-G succeeding G-C in GCGC. The solid line is for the central bond N(1)—H—N(3). The long-dashed line is the O(6)—H—N(4) bond adjacent to the major groove. The short-dashed line is the N(2)—H—O(2) bond adjacent to the minor groove. The first and second half of the temperature axis denote the same range of temperature.

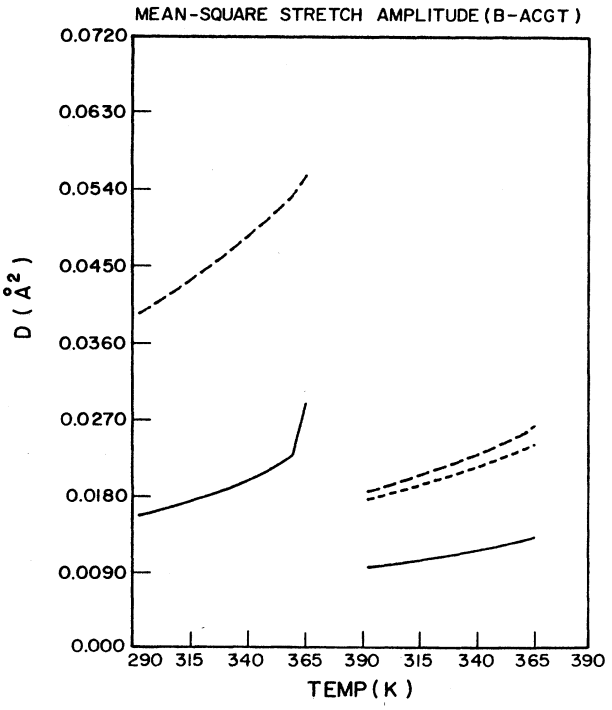


FIG. 3. The mean-square stretching amplitudes for the five hydrogen bonds in ACGT as a function of temperature. The two bonds to the left are of the base pair A-T, while the three bonds to the right are of the base pair C-G succeeding A-T in ACGT. The solid line is the N(1)—H—N(3) bond on both sides. The long-dashed line is the N(6)—H—O(4) bond on the A-T side and O(6)—H—N(4) on the C-G side. The short-dashed line on the C-G side is the N(2)—H—O(2) bond. The first and second half of the temperature axis denote the same range of temperature.

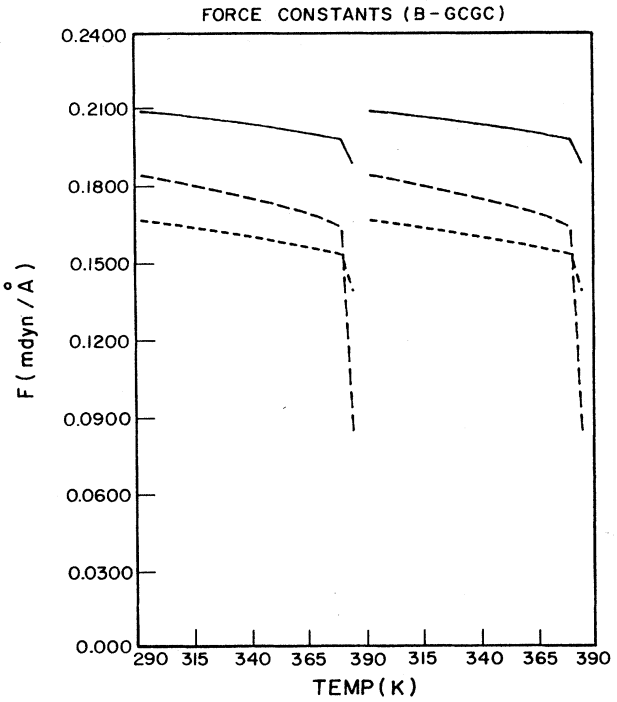


FIG. 4. The effective force constants for the six hydrogen bonds in GCGC as a function of the temperature. The lines refer to the same bonds as in Fig. 1.

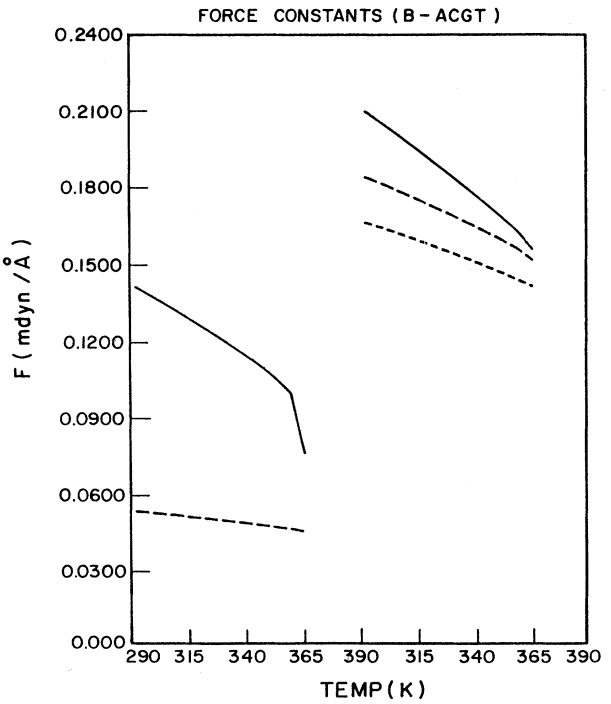


FIG. 5. The effective force constants for the five hydrogen bonds in ACGT as a function of the temperature. The lines refer to the same bonds as in Fig. 2.

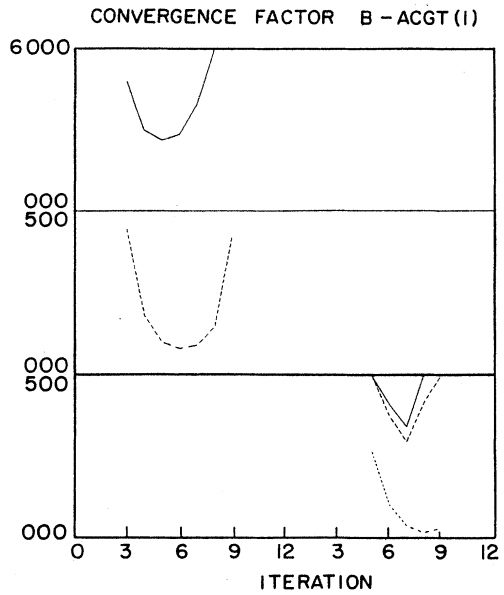


FIG. 6. In ACGT(1) the melting commences in the bonds of the A-T region and spreads to the C-G region. The right half of the figure belongs to the A-T pair and the left half to the C-G pair. The first box at the top shows the first bond to melt. The instability then spreads through the bonds of the lower boxes. Both halves of the iteration axis are over the same range.

creases even further. The return force is very weak and the vibration is anharmonic. This coupled behavior and the resulting feedback causes the instability at high enough temperature.

The breakdown of self-consistency indicates the onset of melting. This reflects itself in the shift from conver-

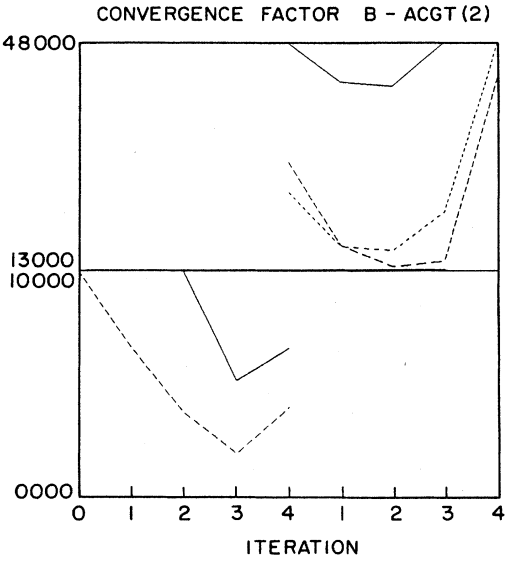


FIG. 7. In ACGT(2) the melting commences in the bonds of the C-G region and spreads to the A-T region. The left half of the figure depicts the C-G pair and the right half, the A-T pair. The box at the top shows the first bonds to melt. The instability then spreads through the bonds of the lower boxes. Both halves of the iteration axis are over the same range.

gence to divergence of $\delta$ in Eq. (35). To explore patterns in the melting of hydrogen bonds and examine the actual bond at which the instability sets in, we plot a convergence factor for each bond. It is the difference in the values of the force constant across an iteration amplified by $10^6$ at the melting temperature. The melting commences at the bond(s) whose convergence factor first diverges after a few iterations. Within 1 K the instability rips through all the bonds. After the onset of instability, $D$ the mean-square stretch amplitude becomes large, the force constant drops, and all approximations involved in the formalism become suspect. The calculations are indicators of the onset of melting only.

In GCGC the melting commences in the bonds adjacent to the major groove.[12] In ATAT both bonds in each base pair melt simultaneously. The actual bond where the melting commences in ACGT and AGCT depends on the form of the near-neighbor interactions in Eq. (4).[12]

The convergence factors of ACGT as a function of iteration at the melting temperature are shown in Figs. 6 and 7. The left half of the plot is the A-T pair and the right half, the C-G pair. The two halves of the axis are over the same range of iteration. The uppermost box shows the bonds that melt first. The melting then spreads through the bonds in the lower boxes.

When the dielectric parameters between nearest neighbors are set equal to the long-range values of Eq. (5), the melting commences in the A-T pairs and spreads to the C-G pair. ACGT, with this inherent interaction, is denoted as ACGT(1) and depicted in Fig. 6. For a distance-dependent dielectric constant in the Coulombic interaction, which is 1.0 within a unit cell and increases to the long-range values as in Eq. (5), melting commences in the C-G regions and spreads to the A-T regions. This is ACGT(2) described in Fig. 7. The diagrams for AGCT are similar and are included in Ref. 12.

When both electrostatic and Van der Waals nonbonded interactions are included we calculate for AGCT an anomalous melting temperature of 370 K. This is much higher compared to its experimentally observed melting temperature, than is the case for the other three copolymers. In fact this temperature is higher than that of ACGT whereas the observed melting temperature of AGCT is lower than that of ACGT. Of the four copolymers studied here AGCT is the only one that has all purine bases on one strand and all pyrimidine bases on the other. The purines are larger than the pyrimidines. But when all purines are on the same strand there is no interleaving across the hydrogen bonds of stacked bases. There is very strong base overlap in a given strand. We believe that the high melting temperature is due to too large a stacking energy which in our formulation occurs via the nonbonded force constants between neighboring bases. To explore this idea we reduced these nonbonded interactions in the calculationally simplest manner by eliminating the Van der Waals interaction but retaining the stronger, more effective, electrostatic interactions as described by Eqs. (4) and (5). In this modified calculation the melting temperature was found to be 357 K which would put it in line with the other calculated melting temperatures as shown in Table III. Clearly our models of nonbonded interactions are still somewhat crude and in need of further refining.

The difference in melting temperature caused by altering the Van der Waals interaction illustrates a point that could be easily misunderstood. Because of our emphasis on hydrogen-bond behavior one may overlook the fact that stacking energy via nonbonded interaction plays a substantial role in the workings of this theoretical approach. The importance of stacking energies has been recognized by workers in DNA for a long time. There is hope for better refinement of the nonbonded interactions as they also affect the frequency of vibrational modes which can be observed independently of melting temperature observations.

Table III lists the calculated and observed melting temperatures. We note that the melting temperatures from the MSPA are about 15 K higher than that reported[15] at 19.5-m$M$ $Na^+$ concentration. The force constants we have used at 293 K represent a low salt concentration. At high concentrations of sodium ions (1–6 mol), the melting temperature decreases[16] with increase in salt concentration. The melting temperatures derived from a mean-field theory are expected to be higher than the experimentally observed temperatures. In our calculation, we assume the DNA is infinitely long, but the terminus and local defects of a real sample cause fluctuations that induce transient base pair openings. These effects make the melting temperature lower than the mean-field melting temperature we have calculated. Calculations by Kim *et al.*[17] indicate that the temperatures calculated from defect mediated melting calculations are lower than the mean-field calculations. Then, they favorably compare with the experimental results. We find MSPA as a successful approach to hydrogen-bond melting in the DNA copolymers.

## ACKNOWLEDGMENTS

This work is supported in part by National Institute of Health, Grant No. GM24443, and Office of Naval Research Nos. N00014-86-K-0252 and N00014-87-K-0162.

[1]Y. Kim, K. V. Devi-Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185 (1985).

[2]E. W. Prohofsky, in *Biomolecular Stereodynamics IV, Proceedings of the Fourth Conversation in the Discipline Biomolecular Stereodynamics,* State University of New York, 1985, edited by R. H. Sarma and M. H. Sarma (Adenine, New York, 1985), pp. 21–46.

[3]K. V. Devi-Prasad and E. W. Prohofsky, Biopolymers **23**, 1795

(1984); K. V. Devi-Prasad, Ph.D. thesis, Purdue University, 1984.

[4]S. M. Lindsay, J. Powell, E. W. Prohofsky, and K. V. Devi-Prasad, in *Structure and Motion: Membranes, Nucleic Acids and Proteins,* edited by E. Clementi, G. Corongiu, M. H. Sarma, and R. H. Sarma (Adenine, New York, 1985), pp. 531–551.

[5]S. Arnott and R. Chandrasekaran (private communication).

[6]K. J. Miller, Biopolymers **18**, 959 (1979).

[7]S. A. Lee, S. M. Lindsay, J. W. Powell, T. Weidlich, N. J. Tao, G. D. Lewen, and A. Rupprecht, Biopolymers **26**, 1637 (1987).

[8]V. V. Prabhu, W. K. Schroll, L. L. Van Zandt, and E. W. Prohofsky, Phys. Rev. Lett. **60**, 1587 (1988).

[9]R. Schroeder and Lippincott, J. Phys. Chem. **61**, 921 (1957).

[10]H. Grimm, H. Stiller, C. F. Majkrzak, A. Rupprecht, and U. Dahlborg, Phys. Rev. Lett. **59**, 1780 (1987).

[11]N. C. Baird, Int. J. Quantum Chem. Quantum Biol. Symp. No. 1, **49**, (1974).

[12]V. V. Prabhu, Ph.D. Thesis, Purdue University, 1988.

[13]B. F. Putnum, L. L. Van Zandt, E. W. Prohofsky, and W. N. Mei, Biophys. J. **35**, 271 (1981); B. F. Putnum, Ph.D. Thesis, Purdue University, 1981.

[14]J. M. Eyster and E. W. Prohofsky, Biopolymers **13**, 2505 (1974); **16**, 965 (1977).

[15]O. Gotoh and Y. Tagashira, Biopolymers **20**, 1033 (1981).

[16]D. R. Monaselidze and G. N. Mgeladze, Biofizika **22**, 950 (1977).

[17]Y. Kim and E. W. Prohofsky, Phys. Rev. B **33**, 5676 (1986).