Defect-mediated hydrogen-bond instability of poly(dG)-poly(dC)

Y. Kim and E. W. Prohofsky
Department of Physics, Purdue University, West Lafayette, Indiana 47907
(Received 23 September 1985)

A modified self-consistent phonon theory based on the complex Green's function is developed for a DNA polymer poly(dG)-poly(dC) which contains a defect (dG refers to repeating guanine bases on one strand and dC the repeating cytosine bases on the other strand). The defect simulates a fluctuation in which the hydrogen bonds which bridge the guanine and cytosine in one cell are broken. The theory is then used to analyze the possible instabilities that can arise in neighboring cells as a function of the temperature. A melting or hydrogen-bond instability does occur in the neighboring cells at 350 K. We find a directional effect around the defect when the instability begins to occur. This calculation has no parameters adjusted to fit melting data and is based on the potentials which are adjusted to fit data of the vibrational modes of the DNA homopolymer at room temperature (293 K).

I. INTRODUCTION

Recently, we have developed a modified self-consistent theory[1-3] (MSPA) based on phonon Green's functions.[1] We have applied it to the perfect infinite DNA polymer poly(dG)-poly(dC) to study the melting of the hydrogen bonds. This melting is inferred from an instability in the hydrogen bonds which makes it impossible to find a self-consistent solution. Although the earlier paper predicted a melting temperature[1] reasonably close to the experimental melting temperature,[4-6] it was a mean-field theory which retained helical symmetry. This required, for instance, that every set of hydrogen bonds melt simultaneously. The actual melting can be expected to involve nucleation sites, and these fluctuations are outside the scope of a mean-field approach. In this paper we apply the self-consistent method to the case where a nucleation defect is assumed present in an otherwise perfect helix. In any real polymer there should be a terminus and local defects which could arise for various reasons as well. If there is a defect in the homopolymer, the fluctuations of the hydrogen bonds would be different around this defect from the fluctuations of the perfect helix. It is therefore likely that the physical melting of the polymer is determined by when such defects grow rather than by the conditions which would bring about mean-field melting. Biologically, it is also very important to study the fluctuations around defects, because the presence of the enzyme certainly introduces a defect.[7]

Putnam *et al.*[8] have studied the fluctuation of hydrogen bonds around the terminus interpreting the terminus as a defect. However, they have not included the anharmonic effects of the hydrogen-bond fluctuations, and so they could not study the temperature instability of hydrogen bonds near the terminus and how the hydrogen bonds behave near the melting temperature. In this paper we want to study the hydrogen-bond instability around a simple but physically and biologically meaningful defect as a first attempt to study the role of nucleation site defects. The defect we consider here is as follows. We cut the three hydrogen bonds of the base pair of the cell (0) of the perfect helix. (See Fig. 1.) Mathematically, this means we set the three internal force constants for the hydrogen-bond stretches of the cell (0) to be zero. We want to see how this affects the instability of the hydrogen bonds of the cells (1) and (−1) and what changes in temperature dependence occur.

In a perfect helix, the helical symmetry[9] of the homopolymer factors the vibrational equation of motion into a block diagonal form, thus a relative phase variable θ which is equivalent to Bloch momentum $\mathbf{k}$ in a lattice with translational symmetry is a good quantum number. In applying MSPA (Refs. 1–3) to the melting of the perfect helix we need the Green's function only in the θ coordinate. However, around the defect there is no helical symmetry and thus everything should be done in configurational coordinates. This is the principal reason to do the problem using Green's functions. One can then do calculations in which very localized phenomena are important. This requires integration over bands, and such integrations gives rise to singularities in the Green's functions. To avoid these singularities we must introduce the complex Green's function. We think this may be the first theoretical study of a defect for lattice vibration which include the anharmonic effects through the complex Green's function. We call this method a self-consistent phonon approach based on the complex Green's functions. All the calculations are done using the internal coordinate of the hydrogen-bond stretches,[1] since in our model the melting of hydrogen bonds is physically due to the instability of the hydrogen bonds resulting from the fluctuation of the hydrogen-bond stretch. This coordinate set involves fewer coordinates in the internal coordinate set.

II. PURE LOCALIZED STATES AND COMPLEX GREEN'S FUNCTION

We introduce the defect at a particular cell, cell (0). (See Fig. 1.) For melting to occur this defect must spread throughout the helix. To determine this spread of the

broken bonds we examine the stability of the neighboring cells, cell (-1) and cell (1). To do this we calculate the vibrational fluctuations of the hydrogen bonds of the cells (1) and (-1). The vibration equation of motion of the helix with the defect is

$$(F-\omega^2+C)q=0, \tag{1}$$

where F is the force-constant matrix[10] of the perfect helix in mass-weighted Cartesian (MWC) coordinates, C is the perturbation for the cut of the three hydrogen bonds of the cell (0), and ω^2 is the eigenfrequencies of the system. The Green's function[8] of the perfect helix is written as

$$g_{\alpha\alpha'}(\omega^2;n,n')=[(\omega^2-F)^{-1}]^{nn'}_{\alpha\alpha'}=\frac{1}{\pi}\sum_j\int_0^\pi d\theta\frac{\mathrm{Re}\{q_j^\alpha(\theta)[q_j^{\alpha'}(\theta)]^*\exp[i\theta(n-n')]\}}{\omega^2-\omega_j^2(\theta)}, \tag{2}$$

where $q_j^\alpha(\theta)\exp(in\theta)$ is the eigenfunction of the αth coordinate of the cell (n) of the eigenfrequency $\omega_j(\theta)$ of the band j with relative phase θ. Here the variable θ is a good quantum number of a system with helical symmetry.[9,10]

If there is a defect in the system, new eigenfrequencies which are so-called "pure localized" states may occur which lie within the branch gaps of the perfect helix dispersion curves. Within the branch gaps, the denominator for the perfect helix Green's function (2) involves no singularities and g may be easily calculated. Therefore from (1) the equation of motion can be rewritten as

$$(1-gC)q=0\ . \tag{3}$$

From Eq. (3), the eigenfunctions for the coordinates directly affected by the defect are entirely determined by the solving "small" system of equations

$$(1-g_{aa}C_{aa})q_a=0\ . \tag{4}$$

Eigenfunctions of the pure localized states for the coordinates not directly affected by the defect are given by

$$q_b=g_{ba}C_{aa}q_a\ . \tag{5}$$

In Eqs. (4) and (5), subscripts a and b refer to those coordinates directly and not directly affected by the cut, respectively. The new frequencies for the localized modes are found by searching for zeros of the determinant

$$D(\omega^2)=\det[1-g_{aa}(\omega^2)C_{aa}]\ . \tag{6}$$

Using Eqs. (4) and (6), we first find the localized eigenstate frequencies and then calculate the eigenfunctions of the hydrogen-bond stretches of the cells (1) and (-1) for the localized states.

If ω^2 lies within one or more dispersion bands of perfect helix, then the denominator of Eq. (2) involves one or more singularities. Using the identity,

$$\lim_{\epsilon\to 0}\frac{1}{x+i\epsilon}=\mathrm{P}\left[\frac{1}{x}\right]-i\pi\delta(x), \tag{7}$$

$g(\omega^2)$ within the bands of perfect helix can be rewritten as

$$g(\omega^2+i\epsilon)=\sum_j\mathrm{P}\left[\int_0^\pi\frac{Q_j(\theta)d\theta}{\omega^2-\omega_j^2(\theta)}\right]$$
$$-i\pi\sum_\kappa[d\omega_\kappa^2(\theta_0)/d\theta]^{-1}Q_\kappa(\theta_0), \tag{8}$$

where $Q_j(\theta)$ is the numerator of the integrand of Eq. (2), $\sum_\kappa$ is the sum over the bands in which ω lies, θ_0 satisfies $\omega=\omega_\kappa(\theta_0)$ and P stands for the principal part of the integration. (For the detailed derivation of this equation, see Ref. 8.) Thus the singularities appear as an imaginary part of Green's function. As will be seen in the next section, the imaginary part of the diagonal Green's function is very important for the calculation of the fluctuation of

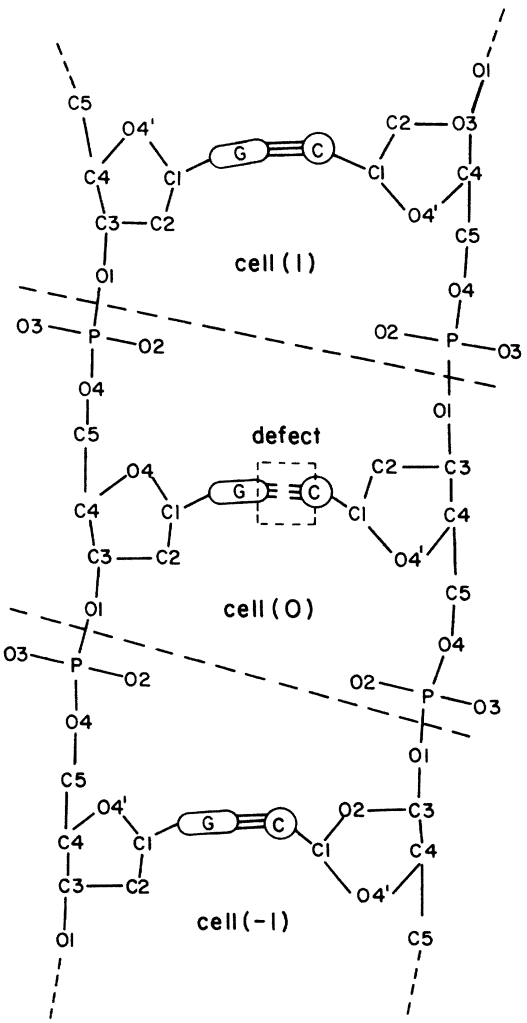

FIG. 1. Portion of three unit cells of poly(dG)-poly(dC) which shows the defect in hydrogen bonds of cell (0).

a bond stretch. From the complex Green's function of the perfect helix, i.e., from Eq. (8), we can easily calculate the defect Green's function $G(\omega^2)$. From Eq. (2) we see that $G(\omega^2)$ satisfies the following equation:

$$G=(\omega^2-F-C)^{-1}=(1-gC)^{-1}g$$
$$=g+gTg\,, \tag{9}$$

where T is the so-called t matrix or scattering matrix which for our defect is written as

$$T=C(1-gC)^{-1}\,. \tag{10}$$

In numerical calculation we first calculate the complex Green's function of the perfect helix through Eq. (8) and then calculate the T matrix using Eq. (10). We then calculate the defect complex Green's function through Eq. (9).

Among all the force constants of DNA we assume that the six force constants for the hydrogen-bond stretches of cells (1) and (−1) change significantly and the other force constants of the cells are assumed not to change significantly. Thus for the self-consistent calculation, to examine the stability of the hydrogen bonds of cells (1) and (−1), we do every calculation in the subspace of the six hydrogen-bond-stretch coordinates of cells (1) and (−1) of the DNA helix internal coordinates. Even though everything is derived in this section using the MWC coordinate, it also can be shown to hold for the internal coordinates.[8] In a previous paper[1] we calculated the eigenfrequencies and the eigenfunctions for the three internal coordinates for the hydrogen-bond stretches of the perfect helix for temperatures below 380 K. We use this information to calculate the eigenfunctions for the six coordinates and the frequencies for the pure localized states from Eqs. (5) and (6) for this case with defect present. For inband modes we calculate the defect complex Green's function through Eqs. (8), (9), and (10). Among the 123 bands of the perfect helix, the "high"-frequency bands numbers 28 and to 123 are essentially dispersionless,[10] thus we assume these can be treated in the Einstein approximation. The other bands whose frequencies are lower than 220 cm^{-1} are treated exactly. This assumption is the same as that for the calculation of the melting of the perfect helix.[1]

III. SELF-CONSISTENT PHONON APPROACH AND COMPLEX GREEN'S FUNCTION

In this section, using the result of the preceding section, we show how the modified self-consistent phonon approach[1-3] (MSPA) is easily formulated through the complex Green's function. In MSPA there are three steps in each iteration to include the anharmonic effects. This is explained in detail in Ref. 1. Here we briefly describe each step so as to include the defect effect.

A. Step 1: Calculation of correlation tensor for the hydrogen-bond stretch

In Ref. 1 the correlation tensor for a hydrogen-bond stretch is shown to be

$$D_i=\langle s^i s^i\rangle=\frac{1}{\pi}\sum_j\int_0^\pi d\theta\frac{1}{2\omega_j(\theta)}\coth[\beta\omega_j(\theta)/2]$$
$$\times s_j^i(\theta)[s_j^i(\theta)]^*\,, \tag{11}$$

where $\langle\ \rangle$ means the thermal average and $s_j^i(\theta)$ is the component of the eigenfunction of the ith hydrogen-bond stretch of band j of the perfect helix. Thus the correlation tensor for a hydrogen-bond stretch is the thermal average of the fluctuation of the hydrogen-bond stretch. Using the relation

$$d\theta\sum_j\frac{1}{2\omega_j(\theta)}s_j^i(s_j^i)^*=d\omega\sum_j\delta(\omega-\omega_j(\theta))$$
$$\times s_j^i(s_j^i)^*[d\omega^2/d\theta]^{-1} \tag{12}$$

and Eqs. (2) and (8), the correlation tensor D_i is reduced to

$$D_i=\frac{1}{\pi}\overline{\int}\,d\omega\,\mathrm{Im}g_i(\omega^2)\coth(\beta\omega/2)\,, \tag{13}$$

where $\mathrm{Im}g_i$ means the imaginary part of the diagonal Green's function of the perfect helix for ith hydrogen bond and $\overline{\int}$ means integration over ω for all the bands of the perfect helix. It is easy to show that this holds for the general case if there are continuous frequency bands. Including the effects of pure localized states in each iteration, the correlation tensor for the ith hydrogen bond with the defect is calculated from the equation

$$D_i=\frac{1}{\pi}\overline{\int}\,d\omega\coth(\beta\omega/2)\mathrm{Im}G_i(\omega^2)$$
$$+\sum_\lambda\frac{1}{2\omega_\lambda}s_\lambda^i(s_\lambda^i)^*\coth(\beta\omega_\lambda/2)\,, \tag{14}$$

where G_i is the defect complex Green's function in each iteration and s_λ^i is the eigenfunction for the λth pure localized states with eigenfrequency ω_λ. First, using the result for the defect complex Green's function for inband modes and the pure localized states of the preceding section, we calculate D_i for the hydrogen-bond stretches of cells (1) and (−1). Then in each iteration we calculate D_i using the result of step 3.

B. Step 2: Force-constant calculation for the hydrogen-bond stretch

Using a Morse potential for each hydrogen-bond stretch from the result of Ref. 1 and D_i of step 1, we calculate the new force constant for the six hydrogen-bond stretches. The Morse potential parameters[1] for the hydrogen bonds are determined from fitting the experimentally observed dynamic behavior, especially for the hydrogen-bond-stretch mode observed around 85 cm^{-1}.[11-15] This step is essentially the same calculation as that in Ref. 1.

C. Step 3: Calculation of new complex Green's function and localized states

Using the given force constants and the calculated results of step 2, we define the force-constants change $\bar{C}$ where $\bar{C}$ is a 6×6 diagonal matrix in the subspace of the hydrogen-bond stretches of cells (1) and (−1) in the internal coordinates of the DNA molecule. The diagonal element of $\bar{C}$ is

$$\bar{C}_i=\bar{\phi}_i-\phi_i \quad (i=1,2,\ldots,6), \tag{15}$$

where $\bar{\phi}_i$ is the calculated force constant from step 2 and ϕ_i is the initially given force constant for the ith hydrogen bond in each iteration. Using $\bar{C}$ instead of C, we can calculate the new complex Green's function and the new localized states for the hydrogen bonds of cells (1) and (−1) by the method described in Sec. II.

IV. RESULTS

In the iteration scheme we define a divergence Δ at each temperature, and using the given information of Sec. II, continue iteration of steps 1, 2, and 3 of the preceding section until

$$\Delta=\text{abs}\left[\frac{\prod_{i=1}^{6}\bar{C}_i}{\prod_{i=1}^{6}\phi_i}\right]<0.0001, \tag{16}$$

where $\bar{C}_i$ is from Eq. (15) and ϕ_i is the force constant for the ith hydrogen bond from the previous iteration. We also compared the current Δ with the Δ_0 of the previous iteration and watched for the instability $\Delta/\Delta_0>1$. At each temperature, using the eigenfrequencies and eigenfunctions of Ref. 1, we first calculate the defect complex Green's function and the pure localized states. Then the iteration of the preceding section is repeated until the self-consistency is established. All the parameters for the Morse potentials for the hydrogen bonds are taken from Ref. 1 and no parameters are altered during the calculations. Starting from 293 K, this procedure is repeated at every 10°. At 350 K no self-consistency is found.

The results of this calculation are displayed in Figs. 2–4. Figure 2 shows the thermal expanded heavy-atom–heavy-atom bond length of the three hydrogen bonds of cells (1) and (−1). It shows that the instability occurs around 350 K. Around the hydrogen-bond defect the instability occurs 30° lower than the case of the perfect helix.[1] In cell (1) as in the case of the perfect helix the hydrogen-bond adjacent to the major groove melts first followed by the other two bonds. [See Fig. 2(a).] However, in cell (−1) the hydrogen bond near the minor groove first shows the instability followed by the other bonds. [See Fig. 2(b).] There is a directional effect around defect. The values for the force constants around the defect are shown in Fig. 3. Here you can also see the directional effect. Figure 4 shows the correlation tensor of the hydrogen-bond stretch as a function of temperature. The correlation tensor which represents the fluctuation of the bond stretch gets very large around 350 K, and it also shows the directional effect.

V. DISCUSSIONS

As we expected, the fluctuation of the hydrogen bonds around a hydrogen-bond defect is much larger than those of the perfect helix at a given temperature. Since fluctuations are expected to occur in the helix the actual observed melting is expected to occur when such fluctuation can grow rather than when the entire helix undergoes mean-field melting. In this calculation we show that growth of a particular fluctuation should occur around 350 K when

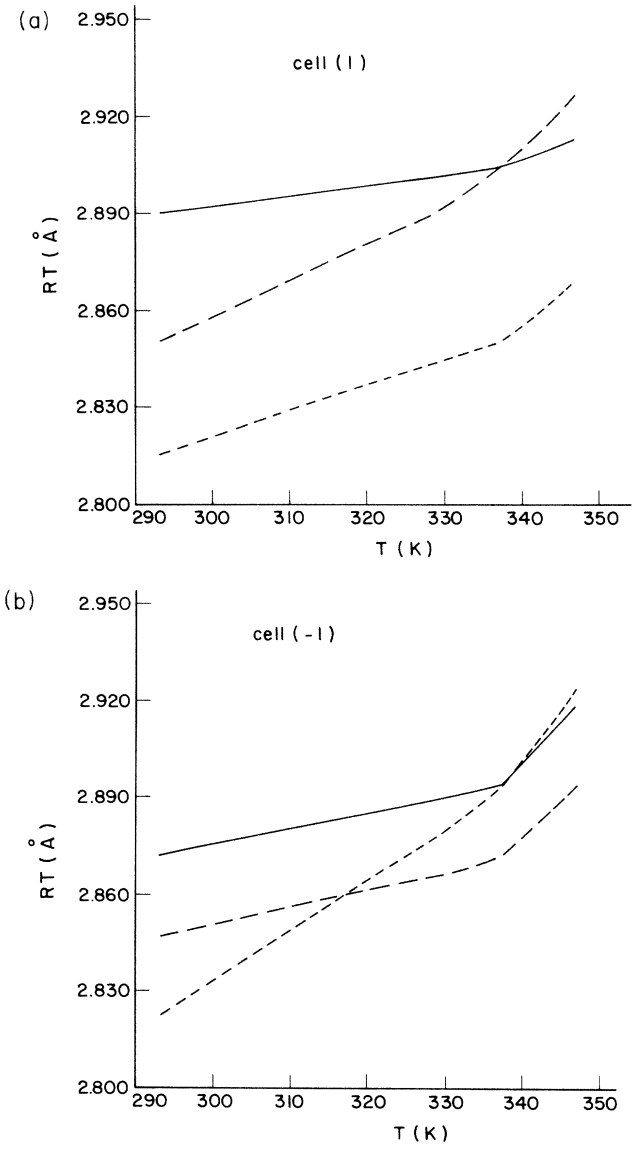

FIG. 2. The mean distance between heavy atoms (thermal expansion) for the hydrogen bonds around the defect as a function of the temperature. The solid line is for the central or N(1)—H—N(3) bond, the long dash line is the O(6)—H—N(4) bond adjacent to the major groove and the short-dashed line is the N(2)—H—O(2) bond adjacent to the minor groove. The curve is plotted as succession of the straight segments between calculated points. The kinks are simply the position of the calculated points. (a) The mean distance for cell (1). (b) The mean distance for cell (−1). By comparing (a) and (b), you can see the directional effect around the defect.

mean-field melting can be expected only above 380 K. The 350-K melting temperature predicted by this calculation is very close to the experimental observation of helix melting although melting temperature does vary with the salt concentration of the DNA sample. The calculations here are based on the eigenfunctions and eigenfrequencies of the perfect helix at low salt concentration.[10,14] As we have explained in Ref. 1, if the salt concentration varies from 10 mM to several hundreds mM, the melting temperature varies approximately 10°. At 19.5 mM Na^+ concentration the reported melting temperature is 360 K.

In our calculation we have ignored the effects of deviations of the fluctuations of the cells from (2) to (∞) and from (-2) to ($-\infty$) from those of the perfect helix. This should be a relatively small effect because of the following reasons. As can be seen from Fig. 4 and the result of Ref. 1,

$$[D_i(\text{defect})-D_i(\text{perfect})]/D_i(\text{perfect})<0.2\,, \qquad (17)$$

where D_i(defect) is a typical correlation tensor among those of the cells from (2) to (∞) or from (-2) to ($-\infty$), and D_i(perfect) is the fluctuation of the perfect helix. This effect on the fluctuation of cells (1) and (-1) is propagated only by the off diagonal Green's function which is a small quantity compared to the diagonal Green's function. Thus it can cause a small correction.

An interesting feature of this calculation is the sensitivity of the melting to phonon amplitude of specific frequencies. The largest contribution to the stretch comes from phonons between 60 and 120 cm^{-1}. It is likely that one can enhance melting by external pumping at these frequencies. We have previously suggested[7] that enzymes may bring about strand separation by excitation of specific phonon modes. Another interesting feature of this cal-

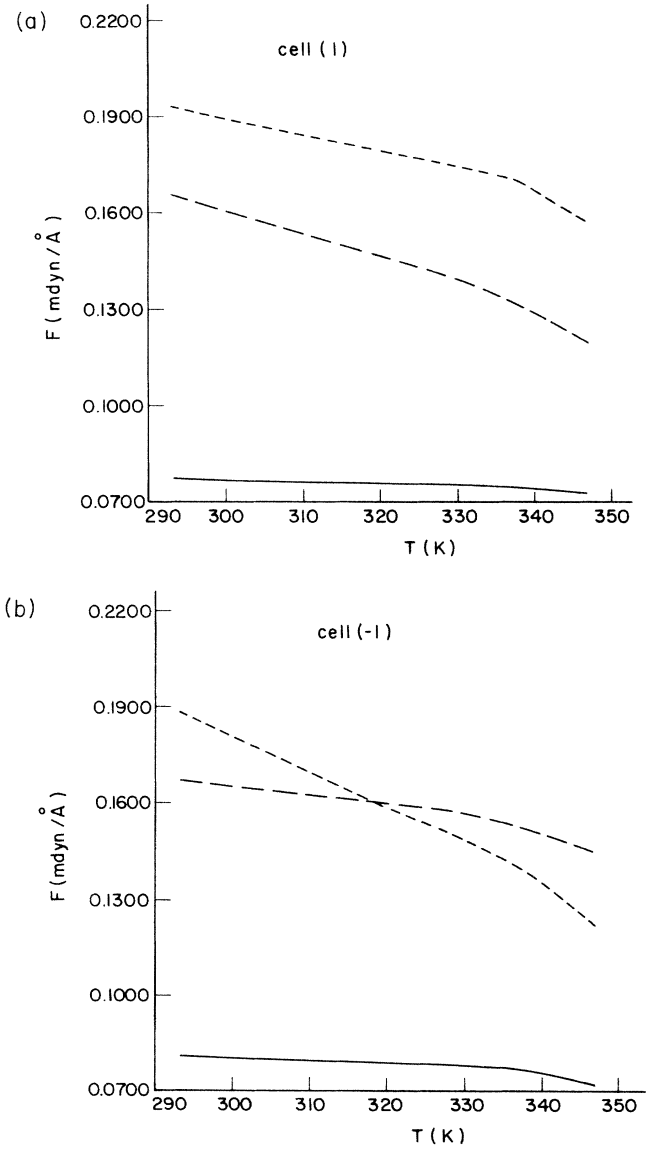

FIG. 3. The effective force constants for the hydrogen bonds of (a) cell (1) and (b) cell (-1) as a function of the temperature. The lines refer to the same bonds as in Fig. 1.

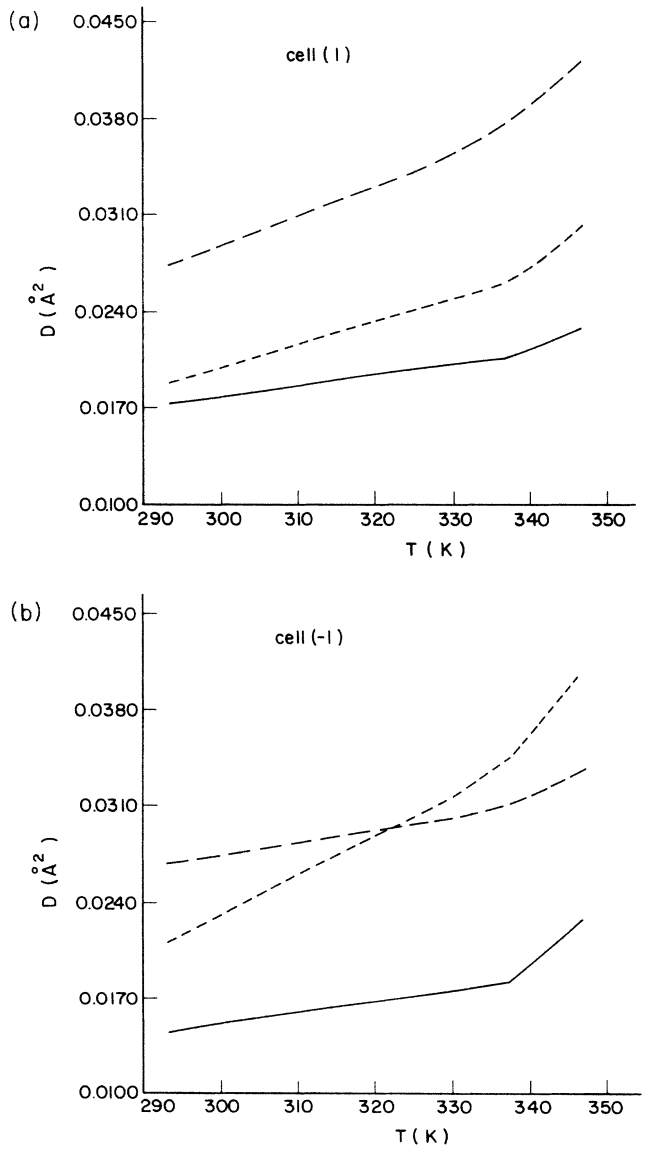

FIG. 4. The correlation tensor D for the hydrogen bonds of (a) cell (1) and (b) cell (-1) as a function of temperature.

culation is the appearance of the dynamical asymmetry in the melting behavior. For the material poly(dG)-poly(dC), melting proceeds in the $+z$ direction ($3' \rightarrow 5'$ in the G backbone) by opening first the hydrogen bond adjacent the major groove. (See Fig. 1.) Melting proceeds in the $-z$ direction ($3' \rightarrow 5'$ in the C backbone) by first opening the hydrogen bond adjacent the minor groove. This asymmetry can only result from the small asymmetry inherent in this sample. It, however, raises the interesting question as to how directional the dynamics can be as a result of specific sequences. It appears that specific base sequences can determine a direction which is easier to melt and which can cause melting in a particular way, i.e., opening to an enzyme interacting in one groove rather than another.

ACKNOWLEDGMENTS

This work has been supported in part by U.S. Naval Air Systems Command Contract No. MDA-850055 and U.S. National Institutes of Health Grant No. GM-24443.

[1] I. Y. Kim, K. V. Devi-Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185 (1985).
[2] Y. Gao and E. W. Prohofsky, J. Chem. Phys. **80**, 2242 (1984); Y. Gao, K. V. Devi-Prasad, and E. W. Prohofsky, J. Chem. Phys. **80**, 6291 (1984).
[3] N. R. Wethramer, Am. J. Phys. **37**, 763 (1969).
[4] O. Gotoh and Y. Tagashira, Biopolymers **20**, 1033 (1981).
[5] C. Schidkraut and S. Lifson, Biopolymers **3**, 195 (1965).
[6] D. R. Monaselidze and G. N. Mgeladze, Biofizika **22**, 950 (1977).
[7] E. W. Prohofsky, Comments Mol. Cell. Biophys. **2**, 65 (1983).
[8] 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 (unpublished).
[9] P. W. Higgs, Proc. R. Soc. London, Ser. A **220**, 472 (1953).
[10] K. V. Devi-Prasad and E. W. Prohofsky, Biopolymers **23**, 1795 (1984); K. V. Devi-Prasad, Ph.D. thesis, Purdue University, 1984 (unpublished).
[11] J. M. Eyster and E. W. Prohofsky, Biopolymers **13**, 2565 (1974); **16**, 965 (1977).
[12] H. Urabe, Y. Tominaga, and K. Kubota, J. Chem. Phys. **78**, 5937 (1983).
[13] S. M. Linsday and J. Powell, Biopolymers **22**, 2045 (1983).
[14] G. S. Edwards, C. C. Davis, M. L. Swicord, and J. D. Saffer, Phys. Rev. Lett. **53**, 1284 (1984).
[15] W. N. Mei, M. Kohli, E. W. Prohofsky, and L. L. Van Zandt, Biopolymers **20**, 833 (1981).