Vibrational fluctuations of hydrogen bonds in a DNA double helix with alternating TA type inserts

R. Beger, Y. Feng, and E. W. Prohofsky,
Department of Physics, Purdue University, West Lafayette, Indiana 47907

ABSTRACT The Green's function technique is applied to a study of breathing modes in a DNA double helix which contains a region of different base pairs from the rest of the double helix. The calculation is performed on an alternating poly(dC – dG) · poly(dC – dG) helix in the B conformation with four consecutive base pairs replaced by a model of a biological promoter region with four alternating T-A,A-T base pairs, henceforth referred to as $(TATA)_2$. The average stretch of interbase hydrogen bonds is found to be amplified around the insert. This is likely related to the $(TATA)_2$ insert having a lower stability against hydrogen bond melting than the two semi-infinite poly(dC-dG) · poly(dC-dG) helices. The insert region may be considered to be a site of enhanced tendency to melt in such a helix. The results show that an alternating AT insert of four base pairs has a larger average hydrogen bond stretch inside and outside the insert region than the average hydrogen bond stretch inside and outside an insert of four consecutive A-T base pairs, henceforth referred to as (AAAA) · (TTTT). Calculations are performed which show that the enhancement of the average hydrogen bond stretch around an alternating TA type insert is greatly dependent upon the local modes and not the inband modes. The amount of local mode enhanced average stretch is explored as a function of insert size.

INTRODUCTION

A Green's function technique has been developed to study the vibrational properties of DNA polymers and has been applied to several examples of DNA double helix. These include a G-C helix with four consecutive base pairs replaced by A-T (1), a junction of AT/GC helices (2), a terminus of a DNA homopolymer (3), a junction of DNA double helix and single strands (4), and a defect-mediated melting of DNA homopolymers (5). In these examples, the alterations in interactions between atoms from those of the perfect helix are treated as defects.

The presence of such defects breaks the helical symmetry thereby making the description of the normal modes much more difficult. However, for situations of interest, such defects usually exist in a finite region of the polymer and involve a relatively small number of atoms. The problem is then conveniently solved by the Green's function technique. A variety of realistic DNA molecules can be treated using this technique, including those with structural disorder, such as terminal regions, irregular base pair sequences, and various linking arrangements involving DNA double helices or single strands which occur during DNA replication, recombination, RNA transcription, and other biological processes. Those defects are believed to play an important role in the melting of the DNA double helix.

It has been found that localized modes can exist around defects in DNA polymers. These can enhance the vibrational fluctuations in the hydrogen bonds which link the two complimentary strands of the DNA double helix. This enhanced stretch can in turn initiate melting of the DNA double helix around the defects (1–5). It is also found experimentally that the stability of a portion of DNA helix can be influenced by the nucleotide sequence of adjacent regions (6, 7, 10). Hence, alterations in the nucleotide sequence in regions adjacent to an actual DNA-protein interaction site could influence the recognition of that site, particularly if the frequency of a breathing mode is important for protein binding. It is also believed that an A-T rich region should melt at lower temperatures than a G-C rich region in a double helix with mixed base pairs (8, 9). This could lead to an initiation of melting from that site. Furthermore, poly(dT – dA) · poly(dT – dA) melts at a lower temperature than poly(dA) · poly(dT) (12). Promoter sequence's with $(TATA)_2$ inside the sequence have been seen to be conserved in many DNA promoter regions ~10 base pairs before the genetic initiation codon (11). Because the four alternating T-A,A-T base pairs have the sequence $(TATA)_2$ and have been seen near promoter regions, four alternating T-A,A-T base pairs could offer a potential melting promoter site.

In the present paper, we apply the Green's function technique to study vibrational properties of a DNA double helix which has a finite region containing base pairs different from the rest of the helix. Breathing modes of the hydrogen bonds are of particular interest. The Green's function method is used to study the thermal fluctuation in the average hydrogen bond stretch. The

calculation is carried out for a DNA double helix in the B conformation which has two consecutive alternating T-A,A-T base pairs in the middle and alternating C-G,G-C base pairs for the rest of the helix. This is equivalent to inserting a section of poly(dT – dA) · poly(dT – dA) double helix into a poly(dC – dG) · poly(dC – dG) double helix. Hydrogen bond thermal fluctuations in the $(TATA)_2$ base pairs and in the two semi-infinite poly(dC – dG) · poly(dC – dG) strands a couple of base pairs away from the $(TATA)_2$ insert are examined and compared with those of poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG) in the same conformation. Finally, the Green's function method is used selectively to determine if the size of the insert effects the average hydrogen bond stretch.

FORMALISM

We construct the altered helix system starting with a perfect double helix poly(dC – dG) · poly(dC – dG) and perfect double helix poly(dT – dA) · poly(dT – dA). We cut the perfect helices by setting the force constants linking the atoms across cell boundaries to zero at two planes in each perfect helix two cells, or four base pairs, apart. We then add corresponding forces to connect the finite $(TATA)_2$ section to the two semi-infinite poly(dC – dG) · poly(dC – dG) strands as shown in Fig. 1. We choose to cut the O5-C5 bond so that the number of valence forces involved is a minimum.

The unperturbed helix system, consisting of two infinite double helices is described by the eigenvalue problem

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

where **F** is the force constant matrix, **I** is a unitary matrix, ω and **q** are eigenvalues and eigenvectors in the mass weighted Cartesian (MWC) coordinates. Because there is no interaction before linking between the infinite helix and the finite section, the force constant matrix **F** has a block diagonal form and Eq. 1 can be written as two separate equations,

$$[\mathbf{F}_{\text{poly(dCG)}} - (\omega_{\text{poly(dCG)}})^2\mathbf{I}]\mathbf{q}_{\text{poly(dCG)}} = 0 \tag{2}$$

and

$$[\mathbf{F}_{\text{poly(dTA)}} - (\omega_{\text{poly(dTA)}})^2\mathbf{I}]\mathbf{q}_{\text{poly(dTA)}} = 0. \tag{3}$$

Here Eq. 2 describes the infinite double helix poly(dC – dG) · poly(dC – dG) and is solved in the usual way by utilizing the helical symmetry. Likewise, Eq. 3 describes the infinite double helix poly(dT – dA) · poly(dT – dA) and is solved in the same way. Generalized forces were used in determining the **F** inside the unit cell and for the long range forces (15). The helicies have a C2 symmetry about the x-axis (which is out of the page in Fig. 1). The C2 symmetry can be used to reduce the computation time of Eq. 2 and Eq. 3 (13).

The perturbed helix system is described by a similar eigenvalue equation

$$(\mathbf{F} - \omega^2\mathbf{I} - \mathbf{C})\mathbf{q} = 0, \tag{4}$$

where **F** is the force constant matrix for the unperturbed system. ω and **q** are the new eigenvalues and eigenvectors. The perturbation matrix, **C**, is the change in **F** necessary to bring about the severing of the infinite helix and the joining of the two semi-infinite poly(dC – dG) · poly(dC – dG) strands created, to the finite $(TATA)_2$ section. The perturbation breaks the helical symmetry and normal mode solutions are much more difficult to calculate. However, because **C** is a matrix of relatively low dimensionality, Eq. 4 can be solved in terms of the Green's function of the unperturbed system.

The Green's functions for the unperturbed and the perturbed systems are defined as

$$\mathbf{g} = (\omega^2\mathbf{I} - \mathbf{F})^{-1}$$

and

$$\mathbf{G} = (\omega^2\mathbf{I} - \mathbf{F} + \mathbf{C})^{-1},$$

respectively. They are related by the following Dyson equation (14),

$$\mathbf{G} = \mathbf{g} + \mathbf{gTg} \tag{5}$$

and where

$$\mathbf{T} = \mathbf{C}(1 - \mathbf{gC})^{-1}.$$

In what follows we have to distinguish between two types of vibrations associated with the defect or local changes (our **C** matrix) made in the force constant

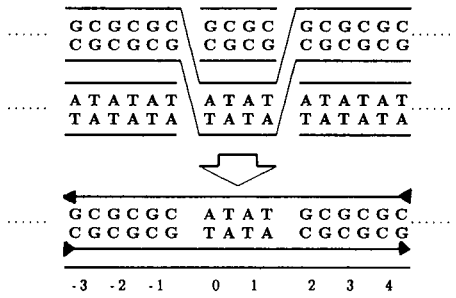

FIGURE 1 The altered double helix is constructed by cutting the perfect helices poly(dC – dG) · poly(dC – dG) and poly(dT – dA) · poly(dT – dA), and joining $(TATA)_2$ base pairs inbetween two semi-infinite poly(dC – dG) · poly(dC – dG) strands. The arrows point in the 5′ to 3′ direction. The numbers at the bottom are unit cell numbers used to identify base pairs in the DNA insert polymer.

matrix. The altered system can have vibrations at new frequencies, i.e., frequencies that were in a band gap in the perfect system with $\mathbf{C} = 0$. These new frequencies are confined to regions near the defect, i.e., base pairs where the changes are made. Their vibrational amplitude falls off exponentially into the unperturbed helix. These modes are called local modes, localized largely to the vicinity of the defect. Other vibrational modes at the defect are at frequencies where vibrational bands exist even in the perfect helices, i.e., with $\mathbf{C} = 0$. These modes are not localized in the same sense. They are called inband modes as when an oscillation is started at the defect, the vibrations will excite vibrations in the rest of the helix. Energy of this frequency can propagate through the systems bands. The inband modes may have larger amplitudes near the defect but are connected to vibrations throughout the system. On the other hand a local mode vibration exists only in the vicinity of the defect.

We are interested in the breathing modes which are relatively low frequency ($\approx 85\ \mathrm{cm}^{-1}$) vibrations characterized by a large average stretch in the hydrogen bonds and presumed to play a significant role in DNA melting. A melting coordinate, s, is defined as the average hydrogen bond stretch amplitude. A thermal mean-square displacement amplitude for the melting coordinate in cell m is then obtained from (5, 15–17)

$$D(m) = \langle s^2 \rangle = \int D^{m}(\omega)\, d\omega + \sum_{\kappa} D_{\kappa}^{m}, \tag{6}$$

where

$$D^{m}(\omega) = \frac{\hbar}{\pi} \operatorname{Im}[G_{ss}^{mm}(\omega^2)] \coth\left(\frac{\hbar\omega}{2kT}\right) \tag{7}$$

and

$$D_{\kappa}^{m} = \frac{\hbar}{2\omega_{\kappa}} s_{\kappa}^{m} s_{\kappa}^{m*} \coth\left(\frac{\hbar\omega_{\kappa}}{2kT}\right) \tag{8}$$

which represent the contribution to the hydrogen bond fluctuation from an inband mode of frequency ω and a localized mode of frequency ω_{κ}, respectively. In Eqs. 7 and 8, T is the temperature. A temperature of 300 K is used in this particular calculation. The normalized average hydrogen bond stretch in cell m associated with the κth localized mode, s_{κ}^{m} is given by

$$s_{\kappa}^{m} = g_{si}^{mn} C_{ij} s_{j\kappa}^{n}, \tag{9}$$

where $s_{j\kappa}^{n}$ is the jth internal coordinate which is directly effected by a perturbation between cell n and cell $n + 1$. Eqs. 6–9 are written in the internal coordinates where the dimension of the perturbation matrix $\mathbf{C}$ is the total number of forces involved by the cutting and joining, which is usually smaller than the total number of Cartesian coordinates involved. In the internal coordinates, $\mathbf{C}$ is a diagonal matrix and its elements are the same in magnitude as the force constants involved in constructing the helix. The negative of the force constant is used to cut a particular bond and a positive force constant, not previously there, is added to create a new connection between particular atoms. Geometrical factors are involved in the actual magnitude of the added and subtracted force constants. The Green's function for the unperturbed system can be written in terms of the solution of the eigenvalue equation (14). In the internal coordinates,

$$g_{ij}^{mn}(\omega^2) = \frac{1}{\pi} \sum_{\lambda} \int_0^{\pi} d\theta \frac{Re[s_i^{\lambda}(\theta) s_j^{\lambda *}(\theta) e^{i(m-n)\theta}]}{\omega^2 - \omega_{\lambda}^2(\theta)}, \tag{10}$$

where $s^{\lambda}(\theta)$, normalized within a unit cell, is the eigenvector belonging to the eigenvalue $\omega_{\lambda}(\theta)$ for a phase angle θ in the internal coordinates. Eq. 10 represents the contributions from all of the vibrational modes of the two semi-infinite poly(dC – dG) · poly(dC – dG) strands and the $(TATA)_2$ insert of the helix. Special care was taken to keep the C2 symmetry about the middle of the insert when the internal coordinates were calculated.

If ω lies within one or more phonon dispersion branches of the perfect helix then the perfect helix Green's function involves an integration over one or more singularities. This difficulty may be circumvented by moving this pole off the real axis by giving ω a small imaginary part i/τ. Physically, this corresponds to an exponential decay of the normal modes caused by, for example, anharmonic interactions causing loss of energy to nearby modes or by dissipative forces such that the molecule would feel in a solution environment. We replace ω^2 by

$$(\omega + i/\tau)^2 = \omega^2 + 2i\omega/\tau - 1/\tau^2$$

$$\approx \omega^2 + i\gamma$$

for small $1/\tau$. The use of ω^2 above in Eq. 10 follows that shown by Feng and Prohofsky (1).

As stated earlier, to reduce the size of the matrix $\mathbf{C}$ and thus the computing time, the long-range Coulomb and Van der Waals interactions are replaced by a set of effective forces (2, 15) which produce equivalent potential energies and a similar spectrum for each DNA homopolymer as those obtained using our more exact model of the long range forces. To emphasize the effects on the breathing modes due to the different base pairs, we used the same set of effective forces for the unperturbed system, (Eq. 1), so that any effect introduced by replacing the Coulomb and Van der Waals forces with effective forces are excluded. In a previous investigation (2), tilting and torsion forces were found to be unimportant in such calculations and they are neglected in the $\mathbf{C}$ matrix to further reduce the computing time.

RESULTS AND DISCUSSION

We carried out the calculation for an infinite DNA double helix in the B conformation, which has a $(TATA)_2$ insert inbetween two semi-infinite poly(dC – dG) · poly(dC – dG) strands. We have found that breathing modes exist around 85 wave numbers in both copolymers which is found accessible by FTIR and Raman techniques (18, 19). We scanned through the entire spectra of both poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG) and found that the average hydrogen bond stretch amplitude is significant only for a few branches of the perfect helix dispersion curves between 60 cm^{-1} and 140 cm^{-1}. Nevertheless, we calculated the average hydrogen bond stretch between 0 and 152 cm^{-1}. The dispersion curves for the frequency range of interest are shown in Fig. 2 and Fig. 3, where Fig. 2 is the phonon dispersion curves of poly(dC – dG) · poly(dC – dG), and Fig. 3 is the phonon dispersion curves of poly(dT – dA) · poly(dT – dA). The dispersion curves are different from those given in reference 2 because the DNA sequences are unique and the torsions for backbone atoms at the cuts and tiltings between bases in adjacent cells are neglected in the present calculation.

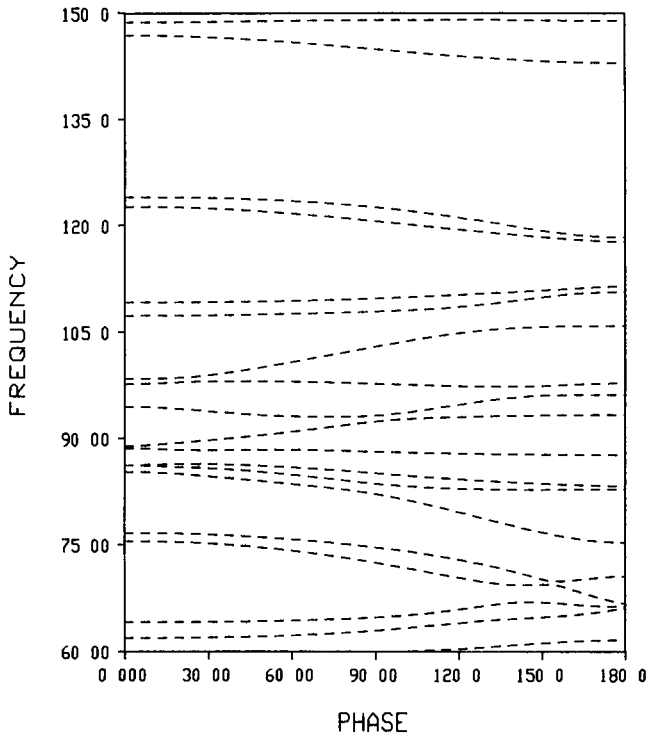

FIGURE 2 Phonon dispersion curves of poly(dC – dG) · poly(dC – dG) from 60 to 150 cm^{-1}. The dashed lines appear to cross but they do not.

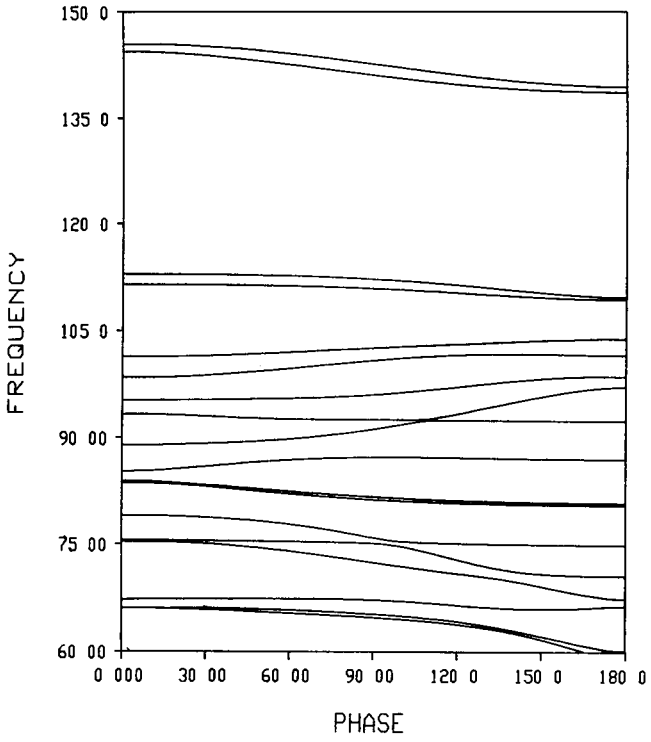

FIGURE 3 Phonon dispersion curves of poly(dT – dA) · poly(dT – dA) from 60 to 150 cm^{-1}.

First, the thermal mean-square amplitude of the average stretch in hydrogen bonds is estimated from Eqs. 6–8 for the $(TATA)_2$ base pairs and 10 base pairs of the poly(dC – dG) · poly(dC – dG) semi-infinite strand closest to the insert section on the right side of the insert. The C2 symmetry about the axis in the middle of the insert reduces the problem by making the both sides of the $(TATA)_2$ insert equivalent. The results are shown in Fig. 4 for the right side of the insert. Here the asterisks represents the total thermal fluctuation, the cross represents the contribution from inband modes, the solid straight lines are the corresponding values for the intact polymers poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG). Significant change in the hydrogen bond fluctuation from that of the perfect helices is observed in the $(TATA)_2$ base pairs and in a few base pairs of the semi-infinite poly(dC – dG) · poly(dC – dG) closest to the $(TATA)_2$ insert. The hydrogen bond fluctuation is increased by a factor of 2.5 around the insert compared with that of poly(dC – dG) · poly(dC – dG). This indicates that the existence of a $(TATA)_2$ insert in two semi-infinite poly(dC – dG) · poly(dC – dG) helices should reduce the stability of the base pairs near the insert. The hydrogen bonds in the $(TATA)_2$ insert or the base pairs of the two semi-infinite poly(dC – dG) · poly(dC – dG) closest to the $(TATA)_2$ section may break

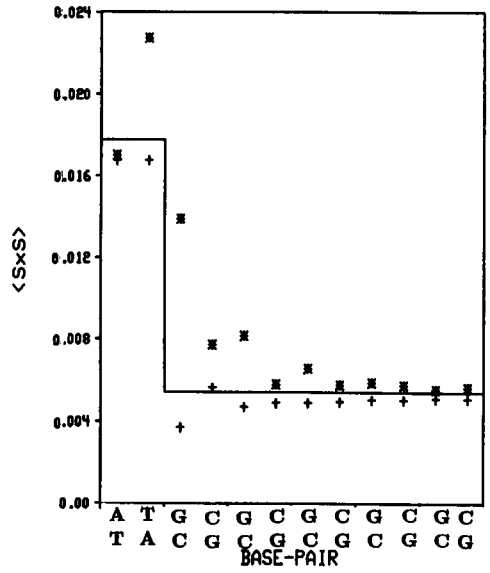

FIGURE 4 Thermal mean-square displacement amplitude D for the melting coordinate is shown with * and without + the local mode contribution. The solid lines are corresponding values for poly(dC – dG) · poly(dC – dG) and poly(dT – dA) · poly(dT – dA). All results for D are in Angstroms squared.

first as temperature increases and thus initiate melting of the double helix. We would expect the T-A hydrogen bonds closest to the edge of the insert to melt first, or all the hydrogen bonds in the insert would melt together before the hydrogen bonds in the two semi-infinite poly(dC – dG) · poly(dC – dG) helices melt.

The dynamics of the base pair separation operates on at least three time scales. Electron motion is the fastest motion and the second fastest motion is associated with the motion of the light hydrogen atom between the heavier atoms making up the hydrogen bond. The slower motions are associated with the motion of the bases relative to each other. The slow motion of the heavy base pairs across the greater hydrogen bond is the breathing motion that can lead to permanent separation of the hydrogen bond. We are studying the actual self consistent melting of $(TATA)_2$ inserts and will be able to tell when the amplitudes of hydrogen bond stretch become important.

One can see that some intensity in s is diminished over that of the perfect poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG) polymers from the inband modes around the insert. This is due to the fact that breathing modes of poly(dT – dA) · poly(dT – dA) and breathing modes of poly(dC – dG) · poly(dC – dG) usually exist in different frequency regions. For example, the average hydrogen bond stretch is largest for dispersion branches in between band 25 at 59.15 cm^{-1} and band 32 at 83.99 cm^{-1} for poly(dT – dA) · poly(dT – dA) whereas the modes with the same characteristics are branches at band 36 at 94.49 cm^{-1} to 105.91 cm^{-1} and in between bands 39 at 117.75 cm^{-1} to band 40 at 124.06 cm^{-1} in poly(dC – dG) · poly(dC – dG). In a region where breathing modes of poly(dC – dG) · poly(dC – dG) exist, the average hydrogen bond stretch is enhanced for the $(TATA)_2$ insert base pairs and reduced in a few base pairs in the two semi-infinite poly(dC – dG) · poly(dC – dG) which are close to the insert. The enhancement in the insert base pairs is usually unimportant because it is still much smaller than the contribution from the regions where breathing modes of poly(dT – dA) · poly(dT – dA) exist. However, the decrease in the base pairs of the semi-infinite poly(dC – dG) · poly(dC – dG) is appreciable because the average hydrogen bond stretch for the base pairs in semi-infinite poly(dC – dG) · poly(dC – dG) is essentially determined by such breathing mode branches. Similarly, in a region where breathing modes of poly(dT – dA) · poly(dT – dA) exist, the average hydrogen bond stretch is appreciably reduced in $(TATA)_2$ base pairs whereas it is slightly increased in a few base pairs in the semi-inifinite poly(dC – dG) · poly(dC – dG) near the $(TATA)_2$ insert. As a result, the $(TATA)_2$ section is stabilized by the two semi-infinite poly(dC – dG) · poly(dC – dG) helices and vice versa. Similar results have been observed in a junction of AT/GC double helix (2) and in a (AAAA) · (TTTT) insert inside two semi-infinite poly(dG) · poly(dC) strands (1).

Comparing this work with the earlier work (1) for a (AAAA) · (TTTT) inside two semi-infinite poly(dG) · poly(dC) strands we see that the average hydrogen bond stretch in a $(TATA)_2$ insert inside two strands is larger inside and outside the insert region than the average hydrogen bond stretch inside and outside the insert region for (AAAA) · (TTTT) inserted into two semi-infinite poly(dG) · poly(dC) helices (1). Because both the inserts and the semi-infinite helices are different it is not possible to assign differences in behavior safely to the change in the insert without further investigations. Fig. 4 shows that the base pair inside the insert region nearest the insert edge has increased its average hydrogen bond stretch by ~25% with the local mode contribution included over the perfect helix average hydrogen bond stretch. The other base pair inside the $(TATA)_2$ insert had an average hydrogen bond stretch that remained about the same as for the perfect poly(dT – dA) · poly(dT – dA) helix.

There were seven local modes found in the first six band gaps that overlapped poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG). The largest local modes were found to be at 114.1 cm^{-1} and at 117.5 cm^{-1}. Fig. 5 shows the local modes contribution to the hydrogen bond stretch for the local modes 114.1 cm^{-1} and 117.5

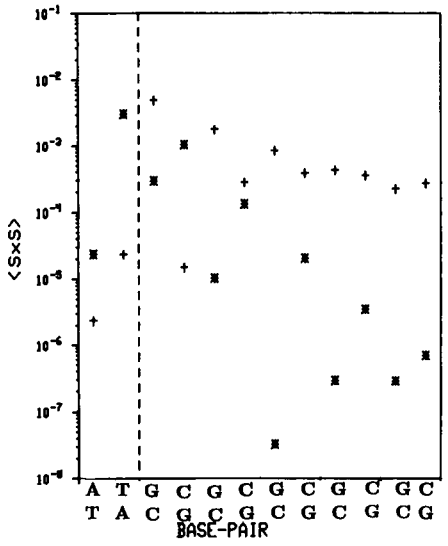

FIGURE 5 The log of the contribution to the hydrogen bond squared amplitude of the local mode at 114.1 cm^{-1} is shown with + and the local mode at 117.5 cm^{-1} is shown with *. The dashed line represents the AT CG junction. The stretch is in units of Angstroms squared.

cm^{-1}. In the (AAAA) · (TTTT) insert, the average hydrogen bond stretch, with the local mode contribution included, showed that three of the four base pairs inside the insert had an average hydrogen bond stretch which was reduced by ~25% when compared with the perfect helix average hydrogen bond stretch. The other base pair inside the insert had an average hydrogen bond stretch that remained about the same as the perfect poly(dA) · poly(dT) helix value. Outside the $(TATA)_2$ insert the average hydrogen bond stretch in the semi-infinite poly(dC – dG) · poly(dC – dG) strands compared with the average semi-infinite poly(dG) · poly(dC) strand hydrogen bond stretch outside of the (AAAA) · (TTTT) insert is about the same. Although the (AAAA) · (TTTT) and the $(TATA)_2$ pair inserts were constructed differently, the results seem to confirm the instability of the alternating poly(dA – dT) · poly(dA – dT) copolymer as seen by Awati and Prohofsky (12). They used MSPA to properly predict that a poly(dA) · poly(dT) polymer melts at a higher temperature than an alternating (dA – dT) · poly(dA – dT) copolymer. It would be nice to see if the (AAAA) · (TTTT) insert also melts at a higher temperature than a $(TATA)_2$ insert, the information thus far uncovered seems to point that way.

We would like to explore the role of the size of the insert region. However, the calculation of the inband contribution in particular requires much machine time. In the four base pair inserts studied so far (1) and the AT/GC junction (2) a pattern has emerged. As described above the inband contribution contributes much the same hydrogen bond stretch as that found for the separated perfect helices. The contribution is somewhat less but is somewhat predictable in its size as well as its dependence on position.

TABLE 1 **The *bp* column lists the base pairs in order going from the insert edge to the middle of the $(TATA)_2$ insert.**

bp	*D*	D^2	D^4	D^6	D^8
1	0.01765	0.02060	0.00798	0.01002	0.00687
2	0.01765	—	0.00122	0.00958	0.01078
3	0.01765	—	—	0.00709	0.00663
4	0.01765	—	—	—	0.00906

D is the perfect helix hydrogen bond stretch and the average hydrogen bond stretches D^2, D^4, D^6, and D^8 are for inserts of two, four, six, and eight base pairs. *D* is in units of Angstroms squared.

We ran a calculation to test whether this would hold for different sized inserts of two, six, and eight alternating TA base pairs as well as the four alternating TA insert discussed above. Table 1 and Table 2 shows the perfect helix hydrogen bond stretch D, and the stretches D^2, D^4, D^6, and D^8, where D^2 stands for the inband average hydrogen bond stretch of the insert polymer with two alternating TA base pairs inside the insert and so forth for D^4, D^6, and D^8 for four, six, and eight base pair inserts. The values inside the insert are in Table 1 and the values outside the insert in the two semi-infinite poly(dC – dG) · poly(dC – dG) strands are in Table 2. The frequencies sampled in this calculation were 10, 23, 30, 40, 50, 60, 66, 70, 80, 90, 95, 100, 111, 120, and 144 inverse cm^{-1}. In Table 1 the base pair column lists the base pairs inside the insert in order from the insert edge until the midpoint of the insert. Table 2 lists the base pairs of the semi-infinite poly(dC – dG) · poly(dC – dG) strand in order from the outside of the insert edge.

TABLE 2 **The *bp* column lists the base pairs in order going away from the insert edge in the semi-infinite poly(dC—dG) · poly(dc—dG) helical strands.**

bp	*D*	D^2	D^4	D^6	D^8
1	0.00647	0.01413	0.00870	0.00641	0.00390
2	0.00647	0.00902	0.00676	0.00639	0.00633
3	0.00647	0.00805	0.00554	0.00530	0.00763
4	0.00647	0.01177	0.00897	0.00921	0.00940
5	0.00647	0.00694	0.00605	0.00599	0.00572
6	0.00647	0.00707	0.00502	0.00526	0.00545
7	0.00647	0.00597	0.00800	0.00760	0.00728
8	0.00647	0.00586	0.00790	0.00755	0.00668
9	0.00647	0.00859	0.00843	0.00832	0.00822
10	0.00647	0.00469	0.00469	0.00490	0.00581

D stands for the perfect helix hydrogen bond stretch and D^2, D^4, D^6, and D^8 are for inserts of two, four, six, and eight base pairs. *D* is in units of Angstroms squared.

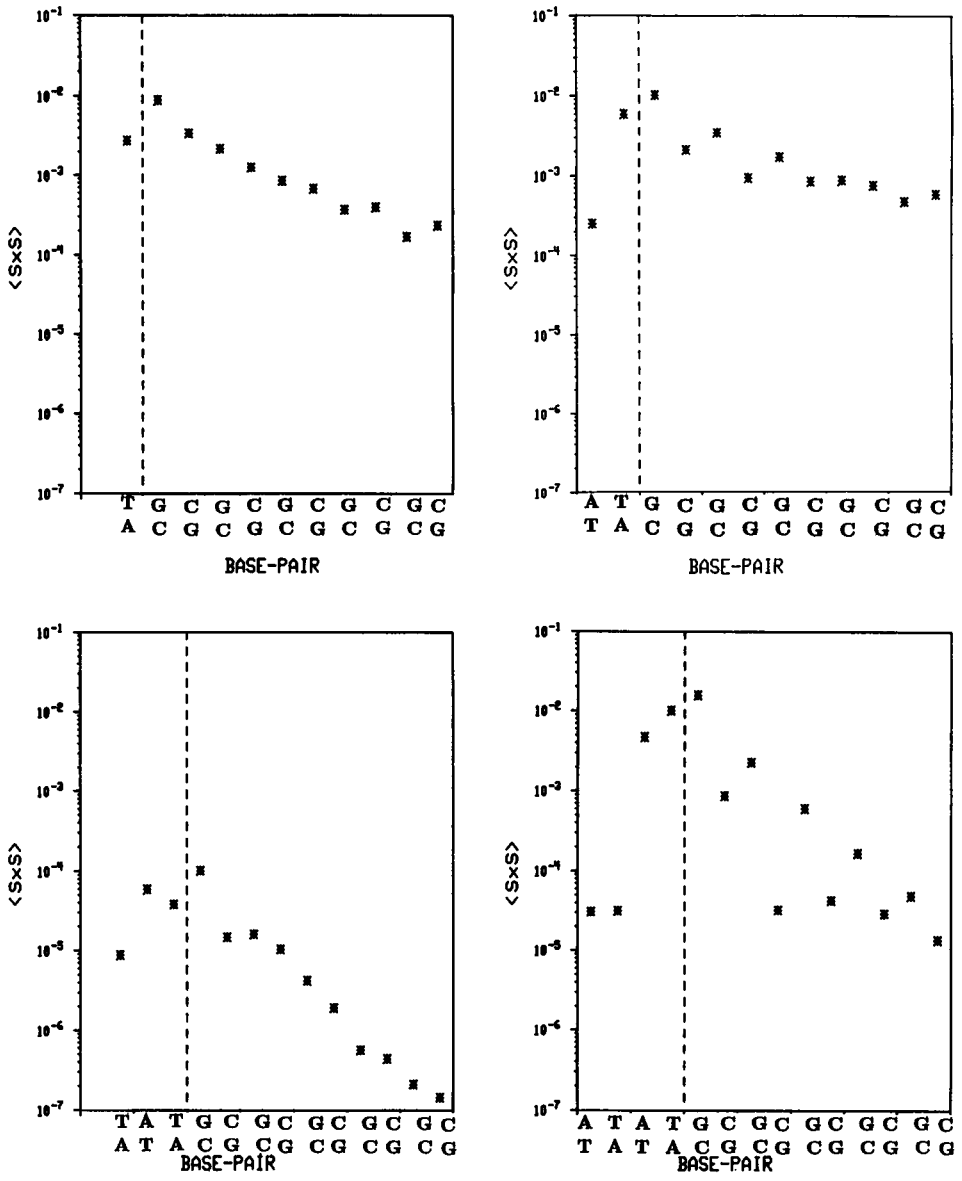

FIGURE 6 (*a*) The log of the contribution to the hydrogen bond squared amplitude D from the local modes for a two base pair insert. We show only the right hand half of the $(TATA)_2$ insert and two semi-infinite poly(dC – dG) · poly(dC – dG) helices. Because the C2 symmetry causes the left hand side to be a mirror image of the right hand side. The AT CG junction is at the dashed line. (*b*) The log of the local mode contribution for the right hand half helix for a four base pair insert. The AT CG junction is at the dashed line. (*c*) The log of the local modes contribution for the right hand half helix for a six base pair insert. The AT CG junction is at the dashed line. (*d*) The log of the local modes contribution for the right hand half helix for an eight base pair insert. The AT CG junction is at the dashed line.

The average hydrogen bond stretch of the base pair just inside the insert for a two base pair insert is large when compared with perfect helix hydrogen bond stretch D. This may be due to the fact that the two base insert has internal coordinates for both sides of the insert in one cell. We believe that the total inband average hydrogen bond stretch for a two base pair insert will still stay near the perfect helix hydrogen bond stretch when local modes are not included (as seen in Fig. 4). The average hydrogen bond stretches found inside and outside of the insert shown in Table 1 and Table 2, respectively, are almost equal to the values shown in Fig. 4. In Table 2, D^4, D^6, and D^8 show a consistent oscillatory behavior after the third base pair. This is to be expected because the inband modes are dependent on the imaginary part of G and G is proportional to $e^{i(m-n)\theta}$. This is also shown in the AT/GC

junction paper (2). In the factor $e^{i(m-n)\theta}$ m and n are cell numbers. This factor makes the hydrogen bond stretches of any individual ω oscillate as the wave travels down the helix. These oscillations will tend to cancel when the integration is over 1,400 ω's. Because D^4, D^6, and D^8 show this expected behavior, we believe it is justifiable to presume that when D^4, D^6, and D^8 are integrated over 1,400 ω's as D^4 was that the total inband average hydrogen bond stretch for the six and eight base pair inserts will be near and lower than the total perfect helix hydrogen band stretch as seen in Fig. 4 for the $(TATA)_2$ insert inside two semi-infinite poly(dC – dG) · poly(dC – dG) helices. It seems that the inband contribution to hydrogen bond stretch is relatively unaffected by details of the inserts except for the smallest insert of two base pairs.

The principal differences in the hydrogen bond stretch from that of the perfect helices is then due to local mode contributions. Because the local modes occur only at a small number of frequencies, the local mode contribution can easily be calculated. It is much simpler to calculate local mode hydrogen bond stretch than to calculate the contribution from the inband modes.

Localized modes tend to enhance the hydrogen bond fluctuation, most significantly at the junctions of the $(TATA)_2$ insert and the two semi-infinite poly(dC – dG) · poly(dC – dG) helices. Figs. 6, *a–d* show the total hydrogen bond stretch due to the local modes for an alternating T-A,A-T insert of two, four, six, and eight base pairs, respectively. The local mode contribution of the six base pair insert is small and there were few local modes as well. The four base pair insert had the most local modes, but the eight base pair insert had the largest total local mode contribution in the first six overlapping band gaps of the perfect poly(dT – dA) · poly(dT – dA) and poly(dC – dG) · poly(dC – dG) dispersions curves. It is reasonable to expect the local modes for the two, six, and eight base pair insert would increase the hydrogen bond stretch around the insert edge as was seen in Fig. 4 for the four base pair insert. One can see that the local mode hydrogen bond fluctuation varies considerably depending on the size of the insert region. The six base pair insert has local mode contributions to hydrogen bond stretch about an order of magnitude less than that for a two, four, or eight base pair insert. If enhanced hydrogen bond instability were the principal factor in effective promoter action, a two, four, or eight base pair insert could be an effective promoter site, a $(TATATA)_2$ would not, as there would be little enhanced hydrogen bond motion leading to enhanced instability.

The average hydrogen bond stretch approaches that of poly(dC – dG) · poly(dC – dG) at about two unit cells or four base pairs from the $(TATA)_2$ insert, indicating that the $(TATA)_2$ insert would not effect hydrogen bond fluctuation in the semi-infinite poly(dC – dG) · poly(dC – dG) helix far away from the insert. However, a number of experimental observations suggest long range effects (21, 22). Several factors which could give rise to long range effects have not been included in this calculation. We have not explored distortions of the basic structure. The junction regions may be displaced from B conformation positions. Another element not included is the synergistic interaction of two unique regions. A third effect not currently included is the synergistic interaction between sequence defects and selfconsistent force constant softening. All of the above should be investigated for long range effects because this calculation predicts no long-range effects by a $(TATA)_2$ insert inbetween two semi-infinite poly(dC – dG) · poly(dC – dG) helices.

As pointed out before, the principal difference between perfect polymer DNA and DNA with special inserts in this calculation is in the appearance of localized modes at particular frequencies. We predict such local modes at 114.1 and 117.5 cm^{-1} for the $(TATA)_2$ insert. These modes should cause an increase in larger hydrogen bond stretch at these frequencies which would cause increased absorption and Raman scattering. It may be possible to test for the occurrence of such localized modes by detecting increased absorption or scattering at these frequencies in systems with inserted $(TATA)_2$ regions.

This work was supported by National Institutes of Health grant GM24443 and Office of Naval Research contract N00014-89-K-0115.

REFERENCES

1. Feng, Y., and E. W. Prohofsky, 1990. Vibrational fluctuations of hydrogen bonds in a DNA double helix with non-uniform base pairs. *Biophys. J.* 57:547–553.
2. Feng, Y., R. D. Beger, X. Hua, and E. W. Prohofsky, 1989. Breathing modes near a junction of DNA double helices. *Phys. Rev. A.* 40:4612–4619.
3. Putnam, B. F., L. L. Van Zandt, E. W. Prohofsky, and W. N. Mei. 1981. Resonant and localized breathing modes in terminal regions of the DNA double helix. *Biophys. J.* 35:271–287.
4. Putnam, B. F., and E. W. Prohofsky. 1983. Localized vibrational modes at a double-helix-single-stand junction. *Biopolymers.* 22: 1759–1767.
5. Kim, Y., and E. W. Prohofsky. 1986. Defect-mediated hydrogen-bond instability of poly(dG)-poly(dC). *Phys. Rev.* B33:5676–5681.
6. Wells, R. D., R. W. Blakesley, S. C. Hardies, G. T. Horn, J. E. Larson, E. Selsing, J. F. Burd, H. W. Chan, J. B. Dodgson, K. F. Jensen, I. F. Nes, and R. M. Wartell. 1977. The role of DNA structure in genetic regulation. *CRC Crit. Rev. Biochem.* 305–340.
7. Burd, J. F., R. M. Wartell, J. B. Dodgson, and R. D. Wells. 1975. Transmission of stability (telestability) in deoxyribonucleic acid. *J. Biol. Chem.* 250:5109–5113.

8. Wada, A., and A. Suyama. 1985. Homogeneous double-helix stability in individual genes. *In* Biomolecular Stereodynamics IV Proceedings of the Fourth Conversation in the Discipline Biomolecular Stereodynamics. R. H. Sarma, editor. Albany, NY. 21–46.

9. Wada, A., and A. Suyama. 1986. Local stability of DNA and RNA secondary structure and its relation to biological functions. *Prog. Biophys. Mol. Biol.* 47:113–157.

10. Burd, J. F., J. E. Larson, and R. D. Wells. 1975. Further studies on telestability in DNA. *J. Biol. Chem.* 250:6002–6007.

11. Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of Escherichia coli promoter sequences. *Nucleic Acids Res.* 15:2343–2255.

12. Awati, K. M., and E. W. Prohofsky, 1989. Anomalous melting of AT-DNA polymers. *Phys. Rev.* In press.

13. Hua, X. M., and E. W. Prohofsky. 1988. Normal-mode calculation for methylated Z-DNA Poly(dG-m^5dC) – Poly(dG-m^5dC). *Biopolymers.* 27:645–655.

14. Bottger, H. 1983. Principles of the theory of lattice dynamics. Physik-Verlag, Weinheim. 87 pp.

15. Putnam, B. F. 1981. Calculation of macromolecular force constants and vibrational properties of a semi-infinite strand of the DNA double helix. Ph.D thesis. Purdue University, West Lafayette, IN. 87–135.

16. Elliott, E. J. 1975. Proceedings of the International School of Physics, Course LV, Lattice Dynamics and Intermolecular Forces. Academic Press, Inc., New York. 342–383.

17. Callaway, J. 1974. Quantum Theory of the Solid State, A and B. Academic Press, Inc., New York. 824 pp.

18. Prohofsky, E. W. 1985. Motional dynamics of the DNA double helix. Biomolecular Stereodynamics IV, Proceedings of the Fourth Conversation in the Discipline Biomolecular Stereodynamics. R. H. Sarma, editor. Albany, NY. 21–46.

19. Powell, J. W., G. S. Edwards, L. Genzel, and A. Wittlin. 1987. Investigation of far-infrared vibrational modes in polynucleotides. *Phys. Rev. A.* 35:3929–3935.

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

21. Sullivan, K. M., and D. M. J. Lilley. 1986. A dominant influence of flanking sequences on a local structural transition in DNA. *Cell.* 47:817–827.

22. Sullivan, K. M., I. H. Murchie, and D. M. J. Lilley. 1988. Long range structural communication between sequences in supercoiled DNA. *J. Biol. Chem.* 263:13074–13082.

23. Patient, R. K., S. C. Hardies, J. E. Larson, R. B. Inman, L. E. Maquat, and R. D. Wells. 1979. Influence of A-T content on the fractionation of DNA restriction fragments by RPC-5 column chromatography. *J. Biol. Chem.* 254:5548–5554.