Calculation of far-infrared absorption in polymer DNA

L. Young, V. V. Prabhu, and E. W. Prohofsky
Department of Physics, Purdue University, West Lafayette, Indiana 47907
(Received 3 March 1988; revised manuscript received 23 September 1988)

We calculate the far-infrared absorption of a specific form of DNA polymer—the *B* conformation, poly(dA)·poly(dT)—using eigenvectors from a lattice dynamical model refined with inelastic neutron scattering data. Our various calculations include the isolated helix, the addition of counterions, and the addition of a "spine of hydration" in the minor groove. We find agreement with experimental infrared data except for one mode. We then attempt to model a coupling between two helices and show some evidence that this mode is interhelical.

I. INTRODUCTION

DNA or deoxyribonucleic acid is a long-chain molecule comprising two strands joined by hydrogen bonds. The two strands wind about a common axis, each forming a helix. For this reason DNA is referred to as a double helix. The strands themselves are made of up of polymeric units called nucleotides. A sugar ring, a phosphate group, and a base form a nucleotide. The nucleotides of DNA have four possible bases: adenine (A), guanine (G), thymine (T), and cytosine (C). The genetic code is contained in the sequence of these bases along the strand.[1]

DNA is found in several different forms or conformations labeled by letters; for example, *A*, *B*, and *Z* are three well-known conformations as shown in Saenger.[2] We will examine the *B* conformation of polymer DNA—a right-handed double helical form. Of specific interest is the homopolymer poly(dA)·poly(dT), for in samples of these polymers the most sharply defined infrared lines are seen.[3] The notation poly(dA)·poly(dT) means that one strand contains only adenine (A) bases and the other, only thymine (T) bases. When the same base is found along the entire strand, the polymer is called a homopolymer. The letter d in poly(dA)·poly(dT) denotes deoxyribonucleic acid, the other nucleic acid being simply ribonucleic acid or RNA.

Our approach to the calculation of far-infrared absorption uses lattice dynamics to calculate the vibrational modes of poly(dA)·poly(dT). The resulting frequencies and eigenvectors are then used to calculate the absorption. Since our initial prediction of vibrational modes in polymer DNA in the 1–100 cm^{-1} frequency range,[4,5] several experimental groups have used Brillouin,[6] low-frequency Raman,[7] and infrared absorption[3] measurements to study this region of the spectrum. Resonant rather than dissipative modes have been observed and in one case the lifetime of such a mode has been determined.[6] The results have been in good agreement with our initial predictions of vibrational modes. Also, the relatively sharp lines available from the far-infrared absorption measurement, which was carried out at both room temperature and at low temperature, have indicated a temperature dependence in agreement with our calculations.[8] Although we have predicted vibrational modes in the far-infrared region, we have not previously calculated the extent of absorption of each mode. Nor have we previously checked the effects of including ions and water molecules explicitly in the calculation instead of accounting for those effect solely with a dielectric constant. Such a check on the isolated helix calculation would expose any accidental cancellation of effects. The latest challenge to our theory is then: Can one achieve good fits and assignments to observed lines in this region, not only for the frequencies but for the relative absorption as well, in a more complex calculation taking into account the water, counterions, and other polymer DNA helices that surround the helix of interest?

In addition to this new Raman and infrared data, inelastic neutron scattering data are also now available.[9] These neutron data are of particular interest to us as they have allowed refinement of our model of the long-range nonbonded interactions in polymer DNA.[10] These nonbonded interactions are the predominant interaction at lowest frequencies; that is, $\nu < 10$ cm^{-1}. The nonbonded interactions are unimportant at higher frequencies ($\nu > 200$ cm^{-1}) where the bonded interactions predominate; yet, they are still important for determining frequencies in the range of interest ($10 < \nu < 200$ cm^{-1}) in this paper. This paper will then use the nonbonded model, based on fitted inelastic neutron scattering data, along with a bonded force constant model developed by refining higher-frequency Raman and infrared data,[11] to calculate the frequencies and the absorption of lines in the far-infrared in poly(dA)·poly(dT) to compare with the observed far-infrared absorption.

As demonstrated by Powell *et al.*[3] relatively sharp infrared lines can be found in good samples of polymer DNA, especially poly(dA)·poly(dT). In these best samples, the polymer DNA is more ordered than in the poorer samples which give rise to much broader absorption peaks. In the more ordered polymer DNA samples, one would expect strong interhelical interactions. Such strong interhelical interactions bring into question the calculational approximation of a single isolated helix. Previous studies have centered on the intrahelical modes, associated with motions mostly in a single helix. Now, in addition, one has strong likelihood of interhelical

modes—modes in which the interactions between helices are dominant.

In this paper we present a simple set of calculations which are aimed at separating out the intrahelical and interhelical modes. We consider how well a simple homopolymer lattice-dynamics calculation can predict the frequency and relative absorption of the intrahelical modes. We do this by comparing the isolated helix vibrational modes to those from several calculations in which the helix interacts with its surroundings in increasing detail. These calculations are for interactions with counterions, interactions with water molecules, and then the combination of the two. We find the method works fairly well for those modes we believe to be intrahelical. Furthermore, we develop a crude calculation to predict the frequencies of the helix interacting with another helix in its surroundings—a calculation of interhelical modes. We find that the principal feature in the measured infrared spectrum that does not correspond to our intrahelical calculations corresponds roughly to a frequency predicted by our crude interhelical calculation. Although the far-infrared frequencies were predicted[4,12] long before sharp lines were seen, the absorption in this range, using an infinite-helix approximation and a nonbonded interaction refined to fit inelastic neutron scattering data, has not been calculated. Also, vibrational modes have not been calculated for the problem in which counterions and water molecules are treated explicitly instead of as a continuum, and the problem of interhelical coupling has not been explored previously.

II. REVIEW

A. Calculations of vibrational modes

We begin with the application of theories of solid-state physics to polymer DNA and model the DNA homopolymer as a one-dimensional infinite lattice (or periodic arrangement of atoms) whose periodic units (or unit cells) are nucleotide pairs or base pairs. The entire double helical homopolymer is generated by taking one base pair and repeating it along the helix axis in the manner of translations and rotations about this axis as if the unit cell were on a screw. In fact, such a lattice is said to have a screw symmetry with a translation operation of p, the pitch of the helix, and a rotation operation of ψ, the pitch angle. Each application of this screw operation yields the next unit cell. Then, according to Bloch's theorem, for an infinite lattice, the motion of the atoms in one repeat unit is the same as that in another except that the two may be of different phase. Therefore, the problem of studying the entire lattice is reduced to the study of one unit cell with all possible phase relationships between succeeding unit cells.

Given the atomic coordinates, masses, charges, and force constants, we find the kinetic and potential energy of the lattice, the equations of motion, and thus, the secular equation as in Eyster and Prohofsky.[13,14] The secular equation is derived in terms of the mass-weighted Cartesian displacement coordinates (MWC) q_i,

$$q_i=\sqrt{m_i}x_i \ , \tag{1}$$

m_i being the mass of, and x_i the displacement of the ith atom. Then the secular equation is given as

$$|\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)-\omega^2\mathbf{I}|=0 \ . \tag{2}$$

$\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)$ should actually be written as $[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_v+[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_{nb}$, where the first term is for valence or bonded force constants, and the second term is for nonbonded interactions.

In $[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta]_v$, $\mathbf{F}$ is the diagonal matrix of force constants of the internal coordinates S_t. $\mathbf{B}(\theta)$ serves to transform $\mathbf{F}$ to MWC coordinates and also account for the phase difference θ between unit cells n, $\mathbf{B}(\theta)=\sum_n \mathbf{B}^n\mathbf{R}^n e^{in\theta}$, where $\mathbf{R}$ is a rotation matrix. The elements of $\mathbf{B}$ are calculated as in Chap. 4 of Wilson *et al.*[15] The valence force constants are refined by Lu,[11] based on spectral data above 400 cm^{-1}. Thus the rows of $\mathbf{B}$ correspond to the numbering of internal coordinates and the columns to the numbering of the x,y,z components of the MWC coordinates.

For $[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_{nb}$, the same concepts hold for $\mathbf{F}$ and $\mathbf{B}$, although in practice we store the matrix elements somewhat differently. We treat the nonbonded interactions between the two atoms of different unit cells as a stretch internal coordinate. Since all of these interactions are stretches, we do not number the internal coordinates but instead label them according to the atoms involved.

As an example, we compare $\mathbf{B}$ for the bonded term with $\mathbf{B}$ for the nonbonded term. For the bonded term, $B_{3,5}$ corresponds to the third internal coordinate, S_3 and the y component of the MWC coordinate of the second atom (atom 1 components are 1,2,3, atom 2 components are 4,5,6,. . .). If S_3 is a stretch between atoms 2 and 6, $B_{3,4}$, $B_{3,5}$, $B_{3,6}$ $B_{3,16}$, $B_{3,17}$, $B_{3,18}$ will be the only nonzero elements on row three. But in the nonbonded term $\mathbf{B}$ is a three-dimensional matrix, $B_{2,6,2}$ corresponding to the y (or second) component of atom 2 (since the third dimension counts x,y,z for the atom listed in the first dimension) involved in the stretch with atom 6. Since the nonbonded problem involves around $(40)^2$ stretches and no other type of internal coordinate, the three-dimensional $\mathbf{B}$ is more compact for computing purposes.

Upon adding these parts together, we can solve the secular equation for the eigenfrequencies and eigenvectors; that is, the resonant frequencies and the relative amplitudes of the atoms of the lattice. Thus we have an adequate expression of its dynamics.

B. Modification of nonbonded potential energy

The potential energy comprises bonded and nonbonded terms which leads to $\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)=[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_v+[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_{nb}$. We follow the method of Eyster and Prohofsky[14] in calculating the nonbonded force constants, modifying the functional form,

$$F_{ij}=F_1+F_2 \ , \tag{3}$$

$$F_1=2\eta\frac{e_ie_j}{\sqrt{\epsilon_i\epsilon_j}r_{ij}^3} \ , \tag{4}$$

$$F_2 = 42\frac{A}{(r_{ij})^n} , \tag{5}$$

where i and j count atoms in different unit cells. F_1 is for long-range or Coulombic interactions, the ith atom having a charge e_i, associated with a dielectric screening ϵ_i and separated from atom j by a distance r_{ij}. The charges e_i are the absolute values of charges taken from calculations by Miller,[16] the hydrogen atoms' charge (as well as mass) being included in the atoms to which they are attached.

We take the absolute value of the charges for the following reason. For a molecule in a stable configuration, the effective force constant on each atom of the lattice must be positive, just as a spring attached to a mass has a positive force constant. Because we are not calculating the force constants from an *ab initio* calculated potential energy, we face the possibility of a negative force constant due to the negative charges in the Coulombic function. Thus the force constant is Coulombic only in the sense that it depends on the magnitudes and separation of the charges. Such a model not only agrees with measurements of acoustic velocity,[5] but also fits neutron scattering data.[10]

The Miller charges are then reduced by a factor of 2.31 based on comparison to a few charges determined experimentally by fitting to x-ray scattering data.[17] $\eta = 0.43$, which includes a conversion factor and the scaling of 2.31 for each charge. The conversion factor is needed because the units of the force constants are mdyn/Å, but the charges are given in units of electronic charge and the coordinates in angstroms. Two distinct dielectric screenings ϵ_i, ϵ_j are used, as some atoms (mostly in the backbone) see each other through water, whereas others (particularly in the bases) see each other through DNA. This difference explains the stronger base-base interactions. We let the dielectric vary linearly according to the distance between the atoms up to an effective separation r_{eff}.[18] For distances greater than r_{eff} the dielectric is treated as a constant. If the atom separation is less than the contact distance r_{con} (the minimum separation allowed between the centers of two atoms), the dielectric constant is taken to be unity. We use an effective distance of 10 Å and a contact distance of 2 Å.

The formula for the dielectric function is

$$\epsilon_{ij} = \epsilon_c + \frac{(\epsilon_0 - \epsilon_c)}{(r_{\text{eff}} - r_{\text{con}})}(r_{ij} - r_{\text{con}}) . \tag{6}$$

ϵ_c is the dielectric constant at the contact distance r_{con}, ϵ_0 is the dielectric constant at the effective separation r_{eff}, and r_{ij} is the separation of atoms i and j. For the interaction between the base atoms, $\epsilon_0 = 6.0$. We use $\epsilon_0 = 9.0$ for the backbone atoms' interaction.

The choice of a value of 9 for the long-range dielectric value of water near DNA is due to the fact that we are interested in vibrational modes in the THz (30–300 cm^{-1}) region. Our earlier model of these nonbonded interactions was fitted to Brillouin scattering observations of the acoustic velocity in polymer DNA fibers at 5–20 GHz in frequency. This is not an entirely satisfactory frequency range for fitting as it is low enough for the effective dielectric constant to be dominated by that of the water near the polynucleotide double helix.

The low-frequency dielectric constant of water is very large but the higher-frequency value is considerably less. Lee *et al.*[19] have used acoustic attenuation data to estimate the rotational relaxation time of water molecules near the polynucleotide double helix. They indicate that the nearest layer of water can rotate only up to frequencies ≈ 4 GHz and that bulk water can rotate up to frequencies ≈ 80 GHz, the latter in agreement with other estimates of the rotational relaxation time for bulk water.

The 50–20 GHz acoustic data are still in the frequency range where the water dipoles can rotate and effectively shield electric fields. The inelastic neutron data is at frequencies of hundreds of GHz and, therefore, above the region of large water dielectric response. We were able to use the neutron data to fit a non-bonded interaction model that should be appropriate for even higher frequencies, where the principal element in that model is the electrostatic interaction shielded by a model of dielectric response.[20] Once above the water rotational relaxation frequency the dielectronic constant of water should be a relatively slowly varying parameter.

F_2 describes the van der Waals interaction when $n = 8$. A is the phenomenological parameter adjusted to a value of 0.12 to fit the infrared absorption data. The long-range interaction is for the unit cell and 19 neighbors on either side. The van der Waals interaction is much shorter in range, involving only one neighbor on each side of the unit cell. The new set of eigenvectors obtained from this refinement is used to calculate the absorption.

C. Polymer DNA absorption calculation

We apply the expressions developed by Kohli *et al.*[21] to calculate the absorption of radiation by poly(dA)·poly(dT) which is in the B conformation. For the relative intensity of the peaks, we calculate the Einstein absorption coefficients,

$$B_L = \frac{8\pi^3}{\hbar^2}|\langle \mu_{l\theta}^L \rangle|^2 , \tag{7}$$

$$B_T = \frac{8\pi^3}{\hbar^2}|\langle \mu_{l\theta}^T \rangle|^2 , \tag{8}$$

where L and T refer to the longitudinal (along the helix axis) and transverse (perpendicular to the helix axis) electromagnetic fields and $\langle \mu_{l\theta} \rangle$ is the transition matrix element of the electric dipole moment operator for the l,θ vibrational mode, l labeling the band and θ, the phase shift from cell to cell.

Assuming the absorption of radiation precipitates a transition between the two lowest vibrational states, these matrix elements are proportional to

$$\langle \mu_{l\theta}^L \rangle = \frac{1}{\sqrt{2}} \sum_i \frac{e_i}{\sqrt{m_i}} \left[\frac{\hbar N}{2\omega_{l\theta}} \right]^{1/2} \delta_\theta S_{il}^0 , \tag{9}$$

$$\langle \mu_{l\theta}^{T} \rangle = \frac{1}{\sqrt{2}} \sum_i \frac{e_i}{\sqrt{m_i}} \left[\frac{\hbar N}{2\omega_{l\theta}} \right]^{1/2} (S_{il}^{+} \delta_{\theta-\psi} + S_{il}^{-} \delta_{\theta+\psi}) , \tag{10}$$

the sum being over the atoms of the repeat unit with charge e_i and mass m_i, with N, the number of repeat units or base pairs. Also,

$$\delta_\alpha \equiv \frac{\sin\left[N\frac{\alpha}{2}\right]}{N\sin\left[\frac{\alpha}{2}\right]} \tag{11}$$

and

$$S_{il}^{\pm} = \frac{q_{ix}^{l} \pm i q_{iy}^{l}}{2} e^{\mp i\phi}, \quad S_{il}^{0} = q_{ix}^{l} , \tag{12}$$

where $q_{ix}^{l}, q_{iy}^{l}, q_{iz}^{l}$ are the components of the lth vibrational mode eigenvector and ϕ is the angle the transverse component of the electromagnetic field makes with the $\hat{\mathbf{e}}_x$ direction.

For our case of an infinite chain, the absorption occurs only for $\theta = 0.0$ in Eq. (9) and $\theta = \pm\psi$ in Eq. (10), according to Kohli *et al.*[21] Then for each base pair ($N = 1$),

$$\langle \mu_{l\theta=0}^{L} \rangle = \frac{1}{\sqrt{2}} \sum_i \frac{e_i}{\sqrt{m_i}} \left[\frac{\hbar}{2\omega_{l\theta=0}} \right]^{1/2} q_{iz}^{l} , \tag{13}$$

$$\langle \mu_{l\theta=\psi}^{T} \rangle = \frac{1}{\sqrt{2}} \sum_i \frac{e_i}{\sqrt{m_i}} \left[\frac{\hbar}{2\omega_{l\theta=\psi}} \right]^{1/2} \frac{q_{ix}^{l} + i q_{iy}^{l}}{2} e^{-i\phi} , \tag{14}$$

and

$$B_L = \frac{8\pi^3}{\hbar^2} \frac{\hbar}{4\omega_{l,\theta=0}} \left| \sum_i \frac{e_i}{\sqrt{m_i}} q_{iz}^{l} \right|^2 , \tag{15}$$

$$B_T = \frac{8\pi^3}{\hbar^2} \frac{\hbar}{4\omega_{l,\theta=\psi}} \left| \sum_i \frac{e_i}{\sqrt{m_i}} \frac{(q_{ix}^{l} + i q_{iy}^{l})}{2} \right|^2 . \tag{16}$$

Noting that $\omega_{l,\theta=+\psi} = \omega_{l,\theta=-\psi}$ (due to time-reversal symmetry), and

$$\left| \sum_i \frac{e_i}{\sqrt{m_i}} (q_{ix}^{l} + i q_{iy}^{l}) \right|^2 = \left| \sum_i \frac{e_i}{\sqrt{m_i}} (q_{ix}^{l} - i q_{iy}^{l}) \right|^2$$

then the total transverse Einstein absorption coefficient is

$$B_T = (2) \frac{8\pi^3}{\hbar^2} \frac{\hbar}{4\omega_{l,\theta=\psi}} \left| \sum_i \frac{e_i}{\sqrt{m_i}} \frac{(q_{ix}^{l} + i q_{iy}^{l})}{2} \right|^2 . \tag{17}$$

In describing absorption data, experimentalists often use the extinction coefficient k which is the complex part of the index of refraction $N(\nu), N(\nu) = n + ik$. The absorption coefficient κ_ν is then $\kappa_\nu = 2\omega k$, where ω is in wave numbers. To relate B_L and B_T to experimental data, we note from Kohli *et al.*[21] that the absorption coefficient is proportional to the Einstein absorption coefficient

$$\kappa_\nu = N_0 \frac{h\nu}{c} B S(\nu) , \tag{18}$$

where κ_ν is the absorption coefficient at frequency ν, N_0 is the number of absorbing centers per cubic centimeter, ν is the frequency of absorption, and $S(\nu)$ is the lineshape function. Choosing a Lorentzian line shape with half-width δ_L, the absorption coefficient evaluated when absorption is maximum becomes

$$\kappa_\nu^{\max} = N_0 \frac{1}{\pi} \frac{h}{c} \frac{\nu_{\max}}{\delta_L} B . \tag{19}$$

The extinction coefficient evaluated at this maximum is then

$$k^{\max} = N_0 \frac{h}{4\pi^2 c \delta_L} B , \tag{20}$$

where we have used $\omega_{\max} = 2\pi\nu_{\max}$. Using Eqs. (7) or (8), we may write $k^{\max}$ in the form

$$k^{\max} = N_0 \frac{4\pi^2}{\hbar \delta_L c} |\langle \mu_{l,\theta} \rangle|^2 . \tag{21}$$

Because the best estimates of δ_L will come from the experimental graph of the extinction coefficient, we will first calculate only $|\langle \mu_{l,\theta} \rangle|^2$. Having made all the calculations, we will then take the best results and calculate the extinction coefficients.

III. RESULTS

A. The isolated helix

For the isolated helix calculation in which the ions and water molecules neighboring polymer DNA atoms are treated as a continuum, the results are given in the first column of Table I. The absorption peaks found by Powell *et al.*,[3] in irradiating Na–poly(dA)·poly(dT), are at 63, 83, 86, 100, and 110 wave numbers (cm^{-1}) at a temperature of 7 K as shown in Fig. 1. Our calculation

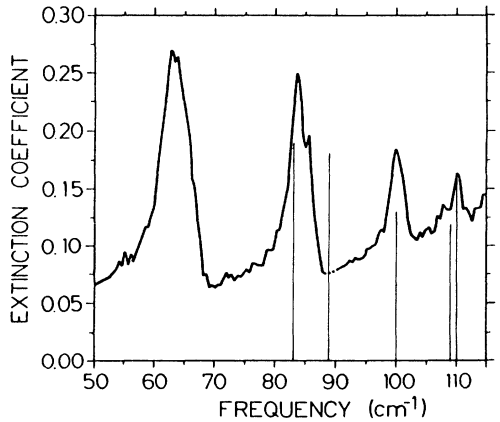

FIG. 1. The locations of the peaks in the measured absorption spectrum of Powell *et al.* (Ref. 3). The vertical lines give the calculated value of the extinction coefficient at its maximum.

TABLE I. Comparison of frequencies and square dipole moments of the isolated helix calculation with those of the calculations in which details of the polymer DNA environment are explicitly added.

Isolated		Counterions		Spine		Ions + Spine		Interhelical	
Mode (cm^{-1})	$\|\mu\|^2\times10^{-19}$ (e cm)^2	Mode (cm^{-1})	$\|\mu\|^2\times10^{-19}$ (e cm)^2	Mode (cm^{-1})	$\|\mu\|^2\times10^{-19}$ (e cm)^2	Mode (cm^{-1})	$\|\mu\|^2\times10^{-19}$ (e cm)^2	Mode (cm^{-1})	$\|\mu\|^2\times10^{-19}$ (e cm)^2
34.0	0.0176	34.5	0.4860	34.0	0.0223	34.6	0.4828	33.0	1.5048
				38.0	0.1727	38.0	0.0051		
39.0	0.6347	40.0	0.4669	39.0	0.4830	40.0	0.4871	40.0	0.5307
45.5	0.4326			45.0	0.4139			45.0	0.3889
		47.0	0.4395	49.0	0.0571	47.0	0.3825		
51.0	0.0268			51.0	0.0657	50.0	0.1294	50.0	0.1464
		52.0	0.0032			52.0	0.2525	52.0	0.4404
53.0	0.0612	54.0	0.3130						
								60.0	0.1603
								61.0	0.1089
				67.0	0.0254	66.5	0.0029		
74.0	0.0570	74.0	0.0246					76.0	0.0117
				82.0	0.1072	81.0	0.2558		
83.0	0.1800	83.0	0.1986					84.0	0.2527
				86.0	0.1264	86.0	0.2028		
89.0	0.1280	88.0	0.5064	90.0	0.0738	90.0	0.2950	89.0	0.6959
		98.0	0.6127	98.0	0.1271	97.0	0.2338	98.0	0.4793
		99.0	0.2116			99.0	0.4481		
100.0	0.1532	101.0	0.4712	102.0	0.0563	101.0	0.3673	100.0	0.2074
						103.0	0.2235	104.0	0.3584
109.0	0.1081			109.6	0.0671				
110.0	0.1417			110.0	0.1622				
		114.0	0.0912			114.0	0.0949		

in the region of the experimental data yields sizable peaks at 39, 46, 83, 89, 100, 109, and 110 cm^{-1} for $\theta=\psi$ (Table I) and 100 cm^{-1} for $\theta=0.0$ (with a value of 0.89×10^{-19} for $|\langle\mu\rangle|^2$). The polarization selections are different for the $\theta=0.0$ and $\theta=\psi$ modes. For absorption at $\theta=0.0$, the incident electric field must be parallel to the helix axis. For the incident field, perpendicular to the helix axis in the ordered film, the absorption is at $\theta=\psi$. The relative contributions to each depend on experimental details.

We see no peak at 63 cm^{-1} and do not have a mode at that frequency. However, outside the range of possible observations in the infrared absorption experiment, we have two highly absorbing modes at 39 and 46 cm^{-1}. Using the eigenvectors of the modes, we find the largest amplitude motion of the free phosphate oxygens occurring in the 39 and 46 cm^{-1} modes (the motion being in the adenine backbone for 39 cm^{-1} and in the thymine backbone for 46 cm^{-1}). These modes are particularly likely to be affected by interhelical interactions as the phosphate groups are on the outside of the helix and highly charged. In packing two right-handed helices parallel to each other, regions of close contact between helix backbones necessarily occur as they must cross each other or distort from helical symmetry. It is likely that such modes are pushed up in frequency by interhelical interactions. We will return to this point when we discuss our calculation with interhelical interactions.

B. Counterions

A solution of DNA requires counterions for stabilization, but we have not included the presence of these ions explicitly in our earlier models. Simulations by Clementi and Corongiu[22] indicate that sodium counterions, strongly attracted to the free phosphate oxygens, follow helical patterns along the backbones. Such interactions with the phosphate groups could alter the frequency of phosphate vibrational modes.

The most difficult step in adding counterions to our model is determining their positions relative to our present set of coordinates. We placed sodium ions an average distance of 2.7 Å away from the free phosphate oxygens, one counterion for each phosphate group. The distance is in approximate agreement with calculations by Clementi and Corongiu.[22] Once the positions are found, the force constants can be calculated. It is likely that the main interaction of the ions and helix is via electrostatic forces. Addition of counterions requires an increase in the size of $[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_{nb}$, the nonbonded term in the secular equation. The new elements for $\mathbf{B}$ and $\mathbf{F}$ are calculated as before in the non-bonded case, the ions interacting with the polymer DNA atoms as well as other ions over four turns of the helix. We do not alter the dielectric used in calculating $\mathbf{F}$ because the counterions are added along the perimeter of the double helix. They are not interspersed between the polymer DNA atoms involved in the nonbonded interaction. They affect the frequencies because of their charge.

Solving this secular equation at $\theta=\psi$ yields eigenfrequencies of 35, 40, 47, and 53 cm^{-1} in the range of interest, 34–65 cm^{-1}, as compared with 39, 46, 51, and 53 cm^{-1} for the isolated helix. The addition of counterions yields no mode near 63 cm^{-1}. Calculation of the absorp-

tion with these eigenvectors does show a decrease in the intensity of the peaks around 40 and 52 cm^{-1} (Table I), while that of the modes in the 53–101 cm^{-1} range show increases in absorption. The exception is the decrease in absorption of the mode at 74 cm^{-1}. In particular, the absorption of the mode at 88 cm^{-1} has more than tripled, while the mode at 83 cm^{-1} shows only a slight increase in absorption. Instead of one peak at 100 cm^{-1}, we find three—98, 99, and 101 cm^{-1}—and none at 110 cm^{-1}. The values of $|\langle\mu\rangle|^2$ are given in Table I.

C. Spine of hydration

The water molecules of the solution which surround the atoms of DNA form hydrogen bonds with some of the nitrogen and oxygen atoms, the bond lifetime being about 10^{-11} sec.[6] Thus, for frequencies greater than a few wave numbers, this layer of water molecules is part of the polymer DNA structure. Kopka *et al.*[23] describe the most distinct pattern of water molecules in *B* poly(dA)·poly(dT) as a "spine of hydration." In her data the spine is formed by water molecules which zigzag from base pair to base pair along the helix axis in the minor groove. Between two base pairs, a water molecule forms a hydrogen bond with a thymine O2 atom in one base pair and an adenine N3 atom in the neighboring base pair. All along the helix, water molecules are found joining neighboring *A*-*T* base pairs. Then to complete the spine, these water molecules are connected to each other by forming hydrogen bonds with another water molecule located between those which bond to the bases (Fig. 7 of Kopka *et al.*[23]).

Aside from this distinct pattern, a total of about 20 water molecules per base pair is found. We limit ourselves to including only the "spine of hydration" in our model, the difficulty being the placing of the molecules in our unit cell configuration which is not exactly the same as the crystal configuration of Kopka *et al.*[23] Also, by adding only the two water molecules per unit cell required for the spine, we do not affect the dielectric used in the nonbonded interaction. The dielectric, which depends on the distance between atoms, is unity for near neighbors, so the water molecules have no effect on the local nonbonded interaction. For larger distances, the addition of two water molecules per unit cell is unlikely to affect the nonbonded interaction enough to alter the dielectric. Nor do the valence force constants require alteration. They were refined based on spectral data above 400 cm^{-1}, a range not affected by the addition of these water molecules.

The hydrogen bond lengths for our calculation vary from 2.64 to 2.89 Å. We use the Lippincott-Schroeder model[24] to calculate the hydrogen-bond stretch force constants which vary from 0.66 (for the 2.64-Å bond) to 0.08 mdyn/Å (2.89-Å bond). We then approximate the angle bend force constants by taking the average of the stretch force constants of the bonds involved in the angle bend and then dividing by seven. The choice of $\frac{1}{7}$ of the magnitude of the stretch force constants is based on our comparison of the valence angle bend and stretch force constants. These approximated angle bend force constants vary in strength from 0.05 to 0.09 mdyn Å/rad^2. Because of the hydrogen bonding of the water molecules to certain poly(dA)·poly(dT) atoms, we increase the size of the bonded term in the secular equation, $[\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)]_v$. The hydrogen bond force constants are placed along the diagonal of **F** and the elements of **B** are calculated as before. Including the mass of the hydrogen atoms of the water with its oxygen atom, we carry out the normal mode calculation. Using these new eigenvectors, we then calculate the absorption. Although we now see a mode in the 60-cm^{-1} range (at 67 cm^{-1}), it shows relatively little absorption (Table I). In the range 70–110 cm^{-1} we have significant peaks at 82, 86, 98, and 110. In general, the absorption has decreased relative to the isolated helix calculation, although the 86-cm^{-1} mode is as absorbing as the 89-cm^{-1} mode of the isolated helix and the mode at 110 cm^{-1} shows an increased absorption.

D. Ions plus spine of hydration

We next calculate the modes for a system of poly(dA)·poly(dT), ions, and water molecules. Each time an atom is added to the model, the size of the dynamical matrix, $\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)$ increases. Its rows and columns correspond to the numbering of the MWC coordinate components of the atoms. Adding ions increased $\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)$ through the nonbonded term, while adding water molecules increased $\mathbf{B}^{\dagger}(\theta)\mathbf{F}\mathbf{B}(\theta)$ through the bonded term. Thus, when we find the secular equation for this system by adding the two terms, we must make sure the ion and water matrix elements do not overlap. The addition of two ions and two water molecules per unit cell increases the size of the final matrix for the unit cell from 123×123 to 135×135. The eigenvectors from this calculation yield a further reduction in absorption of the 67-cm^{-1} mode (Table I). As would be expected, in general, the absorption has decreased relative to the calculation in which ions alone are added and increased relative to the calculation in which water molecules alone are added. Specifically , in the 70–110-cm^{-1} region, the modes with significant peaks are at 81, 86, 90, 97, 99, 101, and 103 cm^{-1}, as seen in Table I.

Although the addition of ions does seem to affect the modes with large amplitude motion of the free phosphate oxygens, and the addition of water molecules does give a mode near 63 cm^{-1}, such results are still not in good agreement with the experimental observation of a highly absorbing peak at 63 cm^{-1}. More likely, some other mechanism will move the 39 and 45 cm^{-1} modes to higher frequency—namely, interhelical coupling.

E. Interhelical coupling

Another mechanism for stiffening of the 39 and 45 cm^{-1} modes is an electrostatic interaction between the phosphate groups of different helices with the sodium counterions included. One can make a crude attempt to model such a coupling by using transverse periodic-boundary conditions. We essentially generate an infinite array of base pairs (including a sodium ion at each phosphate group) along $\hat{\mathbf{e}}_y$ in one base pair's (or unit cell's)

frame of reference, where $\hat{e}_y$ is approximately parallel to the interbase hydrogen bonds. Then we take this array and repeat it along the original base pair's helix axis by the translations and rotations that give polymer DNA its helical symmetry. Of course, we have not succeeded in generating separate helices, but this model should still be valid to look at the effect of the interaction of phosphate groups of nucleotides of different helices on the motion of the free phosphate oxygens. Lindsay[25] provides the distance between helix axes from fiber diffraction experiments. We use 28 Å as this distance; that is, in generating the transverse array of base pairs, the distance between the helix axis of each base pair is 28 Å.

We use the same electrostatic formalism as for the calculation of the addition of counterions to the model. This time the interaction is between the ± 1 transverse repeat unit and the initial repeat unit (cell 0), where ± 1 refers to a translation of ± 28 Å $\hat{e}_y$. Another modification is due to the amount of phosphate clashing between helices. Our model includes a clash for each phosphate group, but such clashes should only occur approximately every five vertical unit cells for neighboring helices. For this reason, we decrease the force constants by $\frac{1}{5}$ to improve the frequencies, realizing that details of our calculation, such as amplitudes of specific atomic motion will be less meaningful.

In the range of interest, interhelical coupling yields modes at 33, 40, 45, 50, 52, 60, and 61 cm^{-1} (Table I). These modes show less motion for the free phosphate oxygen atoms. The modes at 40 and 45 cm^{-1} are now less absorbing than the 39 and 45 cm^{-1} modes of the isolated helix. Also, we now have modes with higher absorption in the 50–65 cm^{-1} range (Table I), whereas before, with the isolated helix, we had no greatly absorbing modes in that region. With the addition of ions, only one significantly absorbing mode appeared in that range. In fact, the absorption of the 52-cm^{-1} mode is near the magnitude of that of the 39 and 45 cm^{-1} modes. In the 70–100 cm^{-1} range, the effect on the peaks relative to the isolated helix is similar to that of adding counterions alone, except that the modes are shifted slightly higher: 84, 89, 98, 100, and 103 cm^{-1}.

F. Extinction coefficient calculation

Although we have seen the need for a calculation which includes details beyond that of the isolated double helix, the latter calculation still provides the best match to the experimental data, except for the 63-cm^{-1} mode. The magnitude of the square dipole moment is in general increased by the addition of counterions, but the relative magnitudes among the peaks given by experiment are better matched by the isolated helix calculation. Adding the water molecules tends to decrease the magnitude of the square dipole moment but does provide a mode at 67 cm^{-1}, not seen in the isolated or counterion calculations. However, the mode shows little promise for great absorption. The calculation with the ions plus spine is a good candidate for calculating the extinction coefficient to compare with experiment except that the results show no mode at 110 cm^{-1}. The success of the interhelical calculation is two modes in the 60-cm^{-1} range with magnitudes of square dipole moments larger than that provided by any other calculation in the same range. However, overall, the relative magnitudes of the experimental data are best matched by the isolated helix calculation.

For this reason, we use the isolated helix data to calculate the extinction coefficients [Eq. (21)]. The estimates of the half-width, δ_L are obtained from the experimental curve of Fig. 1. Lee *et al.*[19] provide an estimate of the density of a sample similar to that used in the experiment of Powell. The densities are given as a function of the relative humidity and range from 1.0 to 1.4 g/cm^3. This density corresponds to N_0 on the order of 10^{21} where N_0 is the number of absorbing centers per cubic centimeter. We have used $N_0 = 1.0 \times 10^{20}$ for the purpose of plotting the results of our calculation on the graph of the experimental data on the same scale for comparison. Thus the value of the extinction coefficient at each maximum, given by our calculation, is represented by the vertical lines in Fig. 1.

IV. CONCLUSIONS

We do find that our single helix calculations fit fairly well the observed modes believed to be intrahelical. This result is likely as the modes calculated in this frequency regime are principally stretch modes of the interbase hydrogen bonds. Earlier calculations on poly(dA)·poly(dT) showed modes at 66, 80, 85, 103, and 110 cm^{-1}, but no absorption was calculated. Since then the nonbonded interactions have been fitted to experimental data covering a wider range of frequencies and phase differences. The current model also lays the foundation for the explicit treatment of water molecules and counterions. The fact that even the addition of counterions and water molecules shifts these modes by only a few wavenumbers reinforces their assignment as hydrogen-bond stretches. Also, these modes are not strongly connected to interhelical interactions.

We now consider the most important insight gained through this calculation. When interhelical coupling is added some, but not all, modes are shifted in frequency. As a result of this calculation, modes appear at 60 and 61 cm^{-1}. Interhelical coupling is expected to be particularly important in those cases where the helices are relatively well ordered. Most importantly, these well ordered cases are the ones that lead to the best resolved spectra. In fact, the mode at 63 cm^{-1} is seen experimentally only in the samples in which the helices are well ordered as is the case with the experimental data examined in this paper. The evidence is that one needs to take into account the interhelical interactions to explain the best spectral data. We see that shifts in frequency of the order needed to obtain better agreement with experiment can be obtained with crude approximations of the interhelical interactions.

This crude calculation of interhelical effects has several flaws which make the calculation suspect and in particular the calculation of the absorption is inaccurate. Adding the transverse array of nucleotide pairs to approximate interhelical coupling actually creates a new mole-

cule altogether. Thus, the symmetry is not the symmetry of the original system. The eigenvectors in particular will be incorrect, although the frequencies may be close to those of a system with interhelical effects. The eigenvectors, however, are what are important in calculating absorption strength.

For properly analyzing a system having both intrahelical and interhelical modes, one must use an entire crystalline unit cell. According to Lindsay,[25] such a cell would have at a minimum, two to three ten base-pair sections of helix, plus all the counterions and water of hydration. The secular matrix would be roughly 30 times larger in each dimension than our homopolymer matrices. For example, the secular matrix for a homopolymer is 123×123; for a copolymer it is two times larger in each dimension or 246×246; and for a three ten base-pair section it is 30 times larger in each dimension or 3690×3690. Although methods to approximate such calculations are beyond the scope of this paper, a more sophisticated calculation of interhelical coupling is being developed.

ACKNOWLEDGMENTS

We wish to thank Professor K. S. Girirajan for the use of his variable dielectric models. We also wish to thank Professor John Powell for permission to use Fig. 1, Powell *et al.*, on the observed far-infrared absorption. This work was supported by the National Institutes of Health (NIH) Grant No. GM24443 and the U. S. Office of Naval Research (ONR) Grant No. N00014-86-K-0252.

[1]G. Felsenfeld, Sci. Am. **253**, 58 (October, 1985).
[2]W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, Berlin, 1984), Chap. 11.
[3]J. W. Powell, G. S. Edwards, L. Genzel, F. Kremer, A. Wittlin, W. Kubasek, and W. Peticolas, Phys. Rev. A **35**, 3929 (1987).
[4]E. W. Prohofsky, K. C. Lu, L. L. Van Zandt, and B. F. Putnam, Phys. Lett. **70A**, 492 (1979).
[5]K. V. Devi Prasad and E. W. Prohofsky, Biopolymers **23**, 1795 (1984).
[6]N. J. Tao, S. M. Lindsay, and A. Rupprecht, Biopolymers **26**, 171 (1987).
[7]Y. Tominaga, M. Shida, K. Kubota, H. Urabe, Y. Nishimura, and M. Tsuboi, J. Chem. Phys. **83**, 5972 (1985).
[8]Y. Kim, K. V. Devi Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185 (1985).
[9]H. Grimm, H. Stiller, C. F. Majkrzak, A. Rupprecht, and U. Dahlborg, Phys. Rev. Lett. **59**, 1780 (1987).
[10]V. V. Prabhu, W. K. Schroll, L. L. Van Zandt, and E. W. Prohofsky, Phys. Rev. Lett. **60**, 1587 (1988).
[11]K. C. Lu, E. W. Prohofsky, and L. L. Van Zandt, Biopolymers **16**, 2491 (1977).
[12]Y. Gao, K. V. Devi-Prasad, and E. W. Prohofsky, J. Chem. Phys. **80**, 6291 (1984).
[13]J. M. Eyster and E. W. Prohofsky, Biopolymers **13**, 2505 (1974).
[14]J. M. Eyster and E. W. Prohofsky, Biopolymers **16**, 965 (1977).
[15]E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations* (Dover, New York, 1980), Chap. 4.
[16]K. J. Miller, Biopolymers **18**, 959 (1979).
[17]D. A. Pearlman and S. H. Kim, Biopolymers **24**, 327 (1985).
[18]A. J. Hopfinger, *Conformational Properties of Macromolecules* (Academic, New York, 1973), Chap. 2, Sec. VI.
[19]S. A. Lee, S. M. Lindsay, J. W. Powell, T. Weidlich, N. J. Tao, G. D. Lewen, and A. Rupprecht, Biopolymers **26**, 1637 (1987).
[20]V. V. Prabhu and E. W. Prohofsky (private communication).
[21]M. Kohli, W. N. Mei, E. W. Prohofsky, and L. L. Van Zandt, Biopolymers **20**, 853 (1981).
[22]E. Clementi, G. Corongiu, IBM DPPG Research Report No. POK-1, 1981 (unpublished).
[23]M. L. Kopka, A. V. Fratini, H. R. Drew, and R. E. Dickerson, J. Mol. Biol. **163**, 129 (1983).
[24]R. Schroeder, and E. R. Lippincott, J. Phys. Chem. **61**, 921 (1957).
[25]S. M. Lindsay, in Proceedings of the International Symposium on Computer Analysis for Life Science, Ohmsha, Tokyo, 1986, edited by C. Kawabata and A. R. Bishop (unpublished).