Strategies for the Computation of Infrared CD and Absorption Spectra of Biological Molecules: Ribonucleic Acids

Ting Xiang, Dixie J. Goss, and Max Diem
Departments of Chemistry and Biochemistry City University of New York, Hunter College, New York, New York 10021 USA

ABSTRACT We report observed and computed infrared (vibrational) circular dichroism spectra of a number of polyribonucleic acids in aqueous solutions in the 1600–1750 cm^{-1} spectral region, in which C=O and some nucleotide base ring stretching vibrations occur. The experimental data are compared with results calculated using different levels of sophistication within the exciton approach. We find that observed band shapes are generally well reproduced by these models, particularly if care is taken to determine the direction of the vibrational dipole transition moments accurately.

INTRODUCTION

Vibrational circular dichroism (VCD) has been reported for polydeoxy- and polyribonucleic acids, in the stretching vibrations of the carbonyl groups of the bases, and a few other vibrations (Annamalai and Keiderling, 1987; Gulotta et al., 1989; Zhong, et al., 1990). Empirically, we and others were able to correlate these observed VCD features with known solution structures, such as B, Z, and triple-stranded helical forms. In addition, we were able to develop a simple method to reproduce observed VCD spectra computationally, using an approach based on the exciton formalism. The reason of adopting this approach for the original model calculations is described below.

Ab initio quantum mechanical calculations to determine molecular vibrational frequencies and the appropriate electric and magnetic dipole derivatives necessary to compute VCD and IR absorption intensities are totally out of the question for the systems reported here due to the enormous size of the molecules being studied. Semi-empirical calculations are not sufficiently accurate to calculate infrared absorption intensities, let alone VCD features. Thus, we feel that at this time, the best method for VCD calculations for biological macromolecules are empirical calculations, where the observed vibrational frequency and the infrared intensity of a monomeric unit are used as input parameters, and VCD and absorption spectra of the polymer are calculated based on different levels of approximation within the exciton approach. We referred to these original calculations as the "extended coupled oscillator" model (Gulotta et al., 1989). These calculations assume that dipole-dipole coupling between the monomeric subunits is the only source of the induced optical activity. Baur and Keiderling (1993) have shown recently that the vibrational exciton approach is a valid approximation in the case of base vibrations in nucleic acids, which they call "weakly coupled transitions." In strongly coupled vibrations, such as the amide I vibrations in peptides, the interactions of the peptide linkages are partially due to dipolar coupling, and partially due to through bond ("kinetic energy") coupling.

The extended coupled oscillator approach successfully for guanine (G) and cytosine (C) containing poly- and oligonucleic acids in both A and B forms. In the case of adenine (A) and thymine (T) (or uracil, U) containing nucleotides, the original calculations did not work out well, due to the fact that the interaction of the two carbonyl groups on T (or U) appear too strongly in the calculations. Thus, one motivation for this work was to develop a formalism which works for all model nucleotides, regardless of their base composition. Furthermore, we found it necessary to include vibrations other than the C=O stretching motion in the model calculations, since A does not contain any C=O groups, yet shows strong VCD in the 1600–1650 cm^{-1} region. The other bases, C, G, and T (or U) exhibit weak VCD couplets around 1630 cm^{-1}, in addition to the main VCD signals between 1650 and 1700 cm^{-1}. These lower frequency transitions have been ignored so far in our model calculations. In order to include these other vibrational modes, we need to consider the fact that their dipole transition moments could couple with those of the C=O vibration. Thus, we developed the formalism for VCD calculations of weakly interacting sets of nondegenerate oscillators.

This paper is organized as follows: First, the equations for the degenerate extended coupled oscillator (DECO) will be reiterated, followed by an introduction of the equations for the nondegenerate case (NECO). An alternative computational approach for the NECO case, named the "pseudo-single strand" (PSS) approach from CD theory, will be discussed.

Next, this paper will present model calculations for various single and double stranded polymers, using the DECO, NECO, and PSS approach. The effect of including near degenerate vibrations and geometric considerations will be discussed. In the case of nucleotides containing T (or U), the single most important improvement of the computational results was obtained when the direction of the dipole derivative was adjusted with respect to the C=O bond direction. This

Received for publication 23 March 1993 and in final form 16 June 1993.
Address reprint requests to Max Diem at Departments of Chemistry and Biochemistry, City University of New York, Hunter College, 695 Park Avenue, New York, NY 10021.

0006-3495/93/09/1255/07 $2.00

correction is very important, particularly if the carbonyl stretching vibration couples with the ring vibrations of the base. However, these corrections require semi-empirical vibrational calculations to determine the amount of mixing of vibrational coordinates. These methods to estimate the direction of the dipole derivative will be presented, and results obtained for poly(rA)·poly(rU) will be shown.

THEORY

In this section, we wish to present the background and the equations necessary for the understanding of the model calculations presented here. The underlying theory utilized is the exciton model proposed originally to explain the splitting of vibronic energy levels of formally degenerate states of solid molecules in a unit cell (Davydov, 1962). Later, this model was adapted (Moffit, 1956; Tinoco, 1963; Bayley et al., 1969) to explain the optical activity of helical polymers. In this model, one assumes that the excited state of a polymer (or unit cell) composed of two or more monomers is delocalized over all the monomers. Thus, if one photon is absorbed into the excited state, it is impossible to determine which of the monomer units is actually excited. Rather, the excited state wave function is best described by a sum over all possible one-quantum excitations. This implies that the excitation is no longer localized on one of the oscillators, but is delocalized over the entire array of identical oscillators. This delocalized excitation is referred to as an "exciton." The dipolar coupling between the transitions lifts their degeneracy; consequently, one observes as many discrete exciton energy levels as there are interacting dipoles.

The equations necessary to define the excited state wavefunction in terms of the dipole-dipole interaction, which acts as the perturbation to mix the eigenstates, have been derived by Moffit (1956). Subsequently, Tinoco (1963, 1964) applied these equations to derive expressions for the dipole and rotational strengths for the cases of two and infinitely many interacting dipoles. This work presents the starting point of our efforts to compute VCD and infrared absorption intensities of biopolymers.

DECO model

We first turn to the previously published model which we refer to henceforth as the "degenerate extended coupled oscillator (DECO)" model. Earlier, we referred to this approach as the extended coupled oscillator, since it merely is an extension of the principles derived by Tinoco (1963, 1964) to an N-mer of interaction dipoles. However, to distinguish it from the models to follow, we wish to designate it by the new name mentioned above. We have pointed out previously that these calculations are carried out at an empirical level, that is, the molecular vibrational frequency of the monomeric subunit, and the magnitude of the dipole derivative, are taken from experimental data.

In the DECO description of the optical activity of N interacting dipoles, the rotational strength R, and hence the VCD intensities, for the kth exciton component ($1 \leq k \leq N$) is given by:

$$R_k = -(\pi\tilde{\nu}_o/c) \sum_{i=1}^{N} \sum_{j>i}^{N} c_{ik}c_{jk}[\mathbf{T}_{ij} \cdot \boldsymbol{\mu}_i \times \boldsymbol{\mu}_j] \quad (1)$$

where c is the velocity of light, and the c_{ij} are the eigenvector components of the (dipole-dipole) interaction matrix:

$$V_{ij} = \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{|\mathbf{T}_{ij}|^3} - \frac{3(\boldsymbol{\mu}_i \cdot \mathbf{T}_{ij})(\boldsymbol{\mu}_j \cdot \mathbf{T}_{ij})}{|\mathbf{T}_{ij}|^5} \quad (2)$$

$\mathbf{T}_{ij}$ is the distance vector between dipoles $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$,

$$\mathbf{T}_{ij} = X_j - X_i \quad (3)$$

and X_i is the coordinate of the center of mass of an oscillator, and $\tilde{\nu}_o$ is the frequency (in wavenumbers) of the unperturbed transition. V_{ij} is a symmetric matrix with zeros along the diagonal and the interaction of all dipoles in the molecules at their appropriate places. For typical vibrational transition moments, the coupling energies vary from about 50 cm^{-1} for parallel or antiparallel dipoles in close proximity to fractions of a wavenumber.

The infrared absorption intensities can be obtained from the dipole strengths D, defined by

$$D_k = \sum_{i=1}^{N} \sum_{j=i}^{N} c_{ik}c_{jk}(\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) \quad (4)$$

The perturbed energies (vibrational frequencies) for each of the k exciton levels are given by

$$\tilde{\nu}_k = \tilde{\nu}_o + \Delta\tilde{\nu}_k \quad (5)$$

where $\Delta\tilde{\nu}_k$ are the eigenvalues of the interaction matrix V_{ij}. Both rotational and dipole strengths are converted to simulated spectra by overlaying (Gaussian, Lorenzian, or mixed) band profiles with suitable parameters.

In the dimeric case, the perturbation matrix V is given by

$$V = \begin{bmatrix} 0 & V_{12} \\ V_{21} & 0 \end{bmatrix}$$

(with $V_{12} = V_{21}$), which has eigenvectors with elements $\pm \frac{1}{2}\sqrt{2}$. When these values are substituted into Eq. 1, the well known case of the dimeric coupled oscillator (Eq. 6) is obtained which predicts the rotational strengths R for the symmetric $|+\rangle$ and antisymmetric $|-\rangle$ combination states of two interacting vibrations according to

$$R^{\pm} = \mp(\pi\tilde{\nu}_o/2c)\mathbf{T}_{12} \cdot \boldsymbol{\mu}_1 \times \boldsymbol{\mu}_2 \quad (6)$$

This equation was derived independently by Holzwarth and Chabay (1972) and by Tinoco (1963).

NECO model

It can be shown that for the nondegenerate case, equations very similar to Eqs. 1, 2, and 4 can be derived. However, the

interaction energy matrix V is defined as

$$V = \begin{bmatrix} \tilde{\nu}_1 & V_{12} & \cdots \\ V_{21} & \tilde{\nu}_2 & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \quad (7)$$

where $\tilde{\nu}_i$ are the nondegenerate frequencies observed, and the terms in V_{ij} are defined as before (Eq. 2). The eigenvalues of Eq. 6 give the vibrational frequencies of the perturbed nondegenerate oscillators, and the eigenvectors are used to calculate the rotational and dipole strengths as follows:

$$R_k = -(\pi/c) \sum_{i=1}^{N} \sum_{j>i}^{N} c_{ik} c_{jk} [(\tilde{\nu}_j X_j - \tilde{\nu}_i X_i) \cdot \boldsymbol{\mu}_i \times \boldsymbol{\mu}_j] \quad (8)$$

$$D_k = \sum_{i=1}^{N} \sum_{j=1}^{N} c_{ik} c_{jk} (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) \quad (9)$$

Eq. 8 is a simplified expression of an equation which was derived originally by Tinoco (1962), for the case of more than one photon being absorbed into the excited state of the polymer (the case of exciton-exciton interactions). The NECO expression for the rotational strength (Eq. 8) differs from that in the DECO case (Eq. 1) by weighting the position vectors of the dipoles by a frequency factor, and Eq. 8 simplifies to Eq. 1 if $\tilde{\nu}_j = \tilde{\nu}_i$. The equation for the dipole strengths (Eq. 9) is unchanged between the case of the degenerate and nondegenerate extended coupled oscillator.

PSS calculations

The "pseudo single-strand" formalism is a simplification of the NECO approach, which lets one treat much larger sections of polymers. It was first developed by Rhodes (1961) who treated double stranded helices as a single strand of degenerate dimers. In our approach, we consider T degenerate and nondegenerate transitions in each subunit, which might be a set of base pairs. The total number of oscillators in the polymer is N, as before. Thus, the number of subunits S, is given by

$$S = N/T \quad (10)$$

For each subunit, one obtains an interaction matrix given by Eq. 7, with eigenvalues E_p and eigenvectors c_{ip}, and $1 \leq p \leq T$. Next, one writes the interaction matrix for the entire polymer:

$$V^p = \begin{bmatrix} 0 & V^p_{nm} \\ V^p_{mn} & 0 \end{bmatrix} \quad (11)$$

where the dimension of the matrix V^p is $S \times S$, and the V^p_{nm} are the interaction between two subunits n and m, given by

$$V^p_{mn} = \sum_{i=1}^{T} \sum_{j=1}^{T} c_{ip} c_{jp} V_{mi,nj} \quad (12)$$

where the c_{ip} were obtained by the diagonalization of the subunit matrices. $V_{mi,nj}$ are the interaction energies of oscillators i and j, located on groups m and n. Diagonalization of Eq. 11 yields eigenvalues ΔE^p_k and eigenvectors ΔA^p_{mk}, with $1 \leq k \leq S$. The complete set of eigenvalues can be obtained as follows:

$$E^p_k = E_p + \Delta E^p_k \quad (13)$$

The rotational and dipole strengths are given by

$$R^p_k = (\pi/c) \sum_{i=1}^{T} \sum_{j=i}^{T} \sum_{m=1}^{S} \sum_{n=1}^{S} c_{ip} c_{jp} A^p_{mk} A^p_{nk} \times [(\tilde{\nu}_j X_{nj} - \tilde{\nu}_i X_{mi}) \cdot \boldsymbol{\mu}_{nj} \times \boldsymbol{\mu}_{mi}] \quad (14)$$

$$D^p_k = \sum_{i=1}^{T} \sum_{j=i}^{T} \sum_{m=1}^{S} \sum_{n=1}^{S} c_{ip} c_{jp} A^p_{mk} A^p_{nk} (\boldsymbol{\mu}_{mi} \cdot \boldsymbol{\mu}_{nj}) \quad (15)$$

The advantage of this method is that the largest matrix which needs to be diagonalized is an $S \times S$ matrix, whereas in the NECO and DECO calculations, the matrix size is $N \times N$.

For all computational approaches, the resulting dipole and rotational strength (in esu^2 cm^2), were converted to extinction coefficient ϵ or $\Delta\epsilon$ (in liters mol^{-1} cm^{-1}) via the approximations

$$D \approx 9.2 \times 10^{-39} \pi \epsilon_{max} w/\tilde{\nu}_o \quad (16)$$

$$R \approx 2.3 \times 10^{-39} \pi \Delta\epsilon_{max} w/\tilde{\nu}_o \quad (17)$$

Here, w denotes the width of the observed VCD or absorption band. Eqs. 16 and 17 hold for Lorenzian band shapes; for Gaussian bands, the factor π needs to be replaced by $\sqrt{\pi}$.

EXPERIMENTAL CONSIDERATIONS

All infrared VCD and absorption spectra reported here were collected on a dispersive VCD spectrometer at Hunter College, which has been described in detail (Lee and Diem, 1992). This instrument is optimized at about 6 μm, but has a sharp low wavenumber cut-off of 1600 cm^{-1}.

All polyribonucleic acids were obtained commercially (Sigma). They were lyophilized to dryness from D_2O to remove exchangeable protons, and dissolved to a concentration of between 10 and 40 mg/ml in D_2O/sodium cacodylate buffer solution. Samples are contained between 19-mm diameter CaF_2 plates held apart by Teflon spacers. The exact concentration and path length conditions employed for each of the polymers is given in the figure legends. All spectra were obtained at 20°C except those of poly(rG) at low pH, which were collected at 10°C (cf. caption to Fig. 4). The spectra are averages of 30 scans, obtained at a scan speed of 1 cm^{-1}/s and a 1-s time constant. Baselines obtained for the sample cells filled with the solvent/buffer were subtracted from the RNA spectra, to account for small instrument baseline artifacts. All spectra reported here agree with those observed earlier by Annamalai and Keiderling (1987).

Coordinates for the nucleic acid bases for canonical structures where obtained from the program MacroModel (Still,

1989), operating on an Evans & Sutherland graphics workstation, or were derived from literature data in the case of poly(rC). For small oligonucleotides, canonical geometries are not particularly accurate in describing the nuclear coordinates, but in the polymers reported here, these geometries are adequate.

Normal coordinate calculations of the nucleic acid bases were carried out using the AMPAC or similar semi-empirical quantum mechanical methods to obtain molecular energies and their second derivatives. These derivatives permit molecular vibrations to be calculated, but like most unscaled quantum mechanical calculations, the calculated vibrational frequencies are up to 200 cm^{-1} higher than the observed ones. However, we could establish that the character of the vibration, in terms of the atomic displacement vectors, are reasonably well reproduced by these calculations. Thus, we used them to establish the direction of some of the vibrational displacements in the bases. The AMPAC calculations were carried out on an IBM 6000 RISC station.

VCD and infrared absorption calculations, using the DECO, NECO, and PSS approaches, were carried out via FORTRAN programs written in our laboratory, and executing on fast personal computers. The dipole and rotational strength were converted to spectral features by overlaying band shapes as indicated above, and storing the spectra in the same format as the observed data to allow for direct comparison.

Transition dipole moments were derived from the integrated absorption intensities of the corresponding monomers, assuming no interaction between transitions. The values agree well with those reported in the literature (Annamalai and Keiderling, 1987; Zhong, et al., 1990).

RESULTS AND DISCUSSION

In the following section, we shall present observed VCD and infrared absorption spectra of various single stranded, double stranded, and more highly aggregated polymers, and shall compare these observed spectral features with computed data, using the models introduced earlier. In all cases, we attempted to utilize the simplest model (DECO), if at all possible, and resorted to the more sophisticated models if the DECO approach failed.

Poly(rC)

We shall start with the experimental and computational results on poly(rC), which exists in buffered aqueous solution in a single stranded, right-handed helical conformation (Arnott et al., 1976). The experimental VCD and infrared absorption spectra are shown in Fig. 1. The absorption shows two distinct peaks, at 1617 and 1656 cm^{-1}. Two distinct VCD couplets are observed under the absorption peaks, with the VCD zero crossing occurring at the peak maxima. Both are positive, conservative couplets, which is typical for a right-handed helix. Conservative in this context implies that the negative and positive intensities in a couplet are equal; furthermore, we define a positive couplet to be one where the positive lobe is at lower energy (wavenumber) than the negative one.

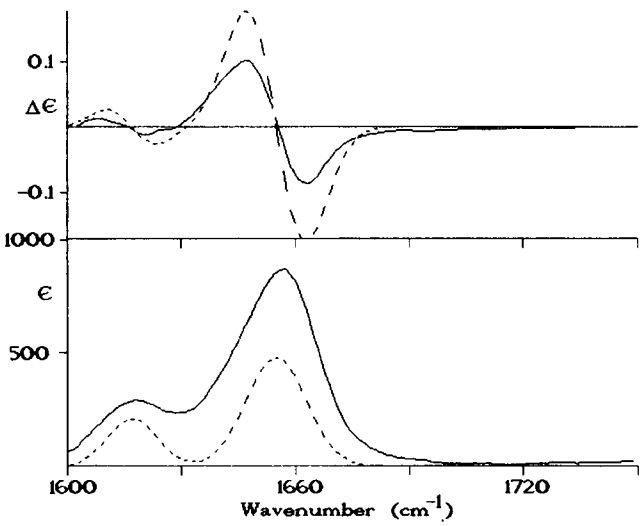

FIGURE 1 Observed and calculated infrared VCD (*top*) and absorption (*bottom*) spectra of poly(rC) in D_2O/sodium cacodylate buffer. Experimental conditions (*solid traces*): [poly(rC)] = 40 mg/ml, T = 20°C, path length = 50 μm, data expressed in molar extinction units per base. Computational conditions (*dashed traces*): ϵ_{max}(1650 cm^{-1}) = 980 liters/(mol cm), w = 18 cm^{-1}. ϵ_{max}(1612 cm^{-1}) = 450 liters/(mol cm), w = 15 cm^{-1}. Two independent DECO computations for 30 bases, results normalized to one base.

The transitions have been assigned (Tsuboi et al., 1973) as follows. The strong peak at 1656 cm^{-1} is typical for the C=O stretching vibration, whereas the 1617 cm^{-1} peak was assigned to a mixture of the C_5=C_6 and the C_4—C_5 stretching coordinates (cf. Fig. 2 *a*). AMPAC calculations for this latter vibration yielded a potential energy distribution of about 95% C_5=C_6 and 4% C_4—C_5. Thus, we assumed that the transition moment of this vibration lies along the double bond. Furthermore, we assume that the carbonyl stretching moment lies along the bond connecting the carbon and oxygen atoms. These dipole moments are indicated in Fig. 2 *a*.

The computed VCD spectra are shown as dashed traces in Fig. 1. Calculations were carried out for a single strand of 30 bases, for which the coordinates were constructed from the data presented by (Arnott et al., 1976). The computed spectra are plotted in units of molar extinction coefficients per base to permit a direct comparison with the experimental data.

The spectra were calculated under the assumption that the two transitions at 1617 and 1656 cm^{-1} do not interact. Thus, two separate DECO VCD and infrared absorption spectra were calculated according to Eqs. 1 and 4 for the two transitions, and subsequently added. Since both observed VCD spectra are nearly conservative, this assumption appears justified, and the agreement between observed and calculated absorption and VCD band shapes and frequencies confirms the validity of this approach. Since the two exciton manifolds were calculated separately, the relative directions of the dipole moment vectors is immaterial.

Thus, the VCD of poly(rC) in the 1600 to 1750 cm^{-1} spectral region is an example of a system which can be repro-

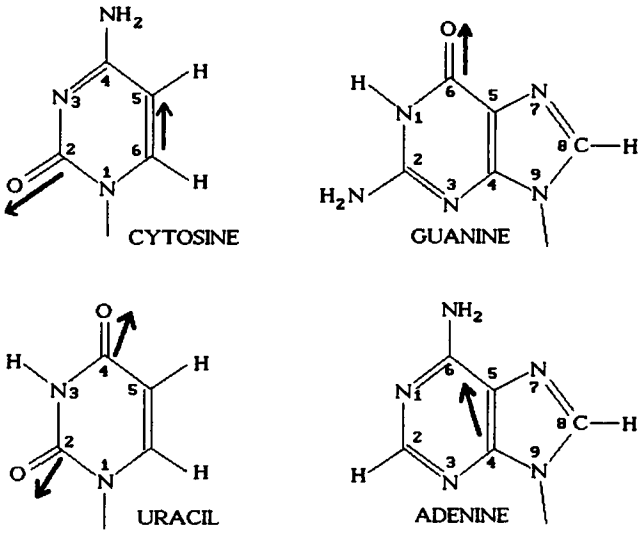

FIGURE 2 Structures of the nucleic acid bases. The directions of the appropriate transition moments are based on semi-empirical calculations, and are drawn approximately.

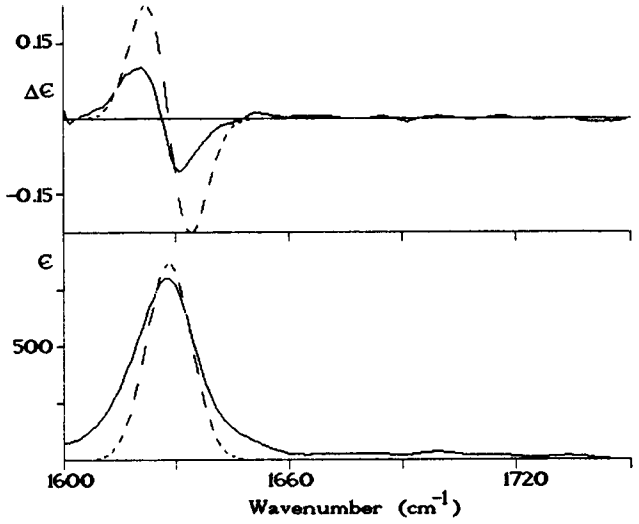

FIGURE 3 Observed and calculated infrared VCD (*top*) and absorption (*bottom*) spectra of poly(rA) in D_2O/sodium cacodylate buffer. Experimental conditions (*solid traces*): [poly(rA)] = 30 mg/ml, T = 20°C, path length = 50 μm, data expressed in molar extinction units per base. Computational conditions (*dashed traces*): $\epsilon_{max}(1625\ cm^{-1})$ = 1100 liters/(mol cm), w = 16 cm^{-1}. DECO computation for 30 bases, results normalized to one base.

duced computationally by considering two noninteracting transitions and calculating their spectral properties independently.

Poly(rA)

Adenine, shown in Fig. 2 *b*, does not contain a C=O functional group, and therefore, does not exhibit infrared absorptions between 1650 and 1700 cm^{-1}, cf. Fig. 3. The peak observed at about 1627 cm^{-1} has been assigned to a coupled C_4=C_5/C_5—C_6 stretching motion (Tsuboi et al., 1973). Since this transition is the only vibration considered in the calculations for poly(rA), its orientation within the purine ring system is immaterial, and only the relative orientations of the ring systems in subsequent bases matter. However, for the calculations on poly(rA)·poly(rU), to be discussed later, the orientation of the transition moment with respect to the transitions in the U base needed to be established. Thus, we determined and used the proper direction of this transition in the calculations on poly(rA) as well, cf. Fig. 2 *b*. Structural features of poly(rA) were derived from program MacroModel, using single stranded RNA parameters. VCD calculations were carried out for a 30-base single strand, using a ϵ_{max} value given by Tsuboi et al. (1973).

The VCD spectrum of this vibration is conservative as well, and the VCD zero crossing occurs at the absorption maximum. Computational results based on the DECO approach reproduce this behavior well, and are shown in Fig. 3. Thus, the calculated and observed results on poly(rA) and poly(rC) indicate that a simple exciton type mechanism is responsible for the vibrational optical activity of these polymers in both C=O and C=C stretching vibrations. The computational model overestimates the VCD intensities in both cases by a factor of two; yet, in both cases there is good qualitative agreement between theoretical and experimental results, in both the C=O stretching modes as well as the less polar C=C—C stretching mode.

Poly(rG)

At neutral pH, the observed and calculated spectra of poly(rG) do not agree at all. However, it is well established that poly(rG) exists at room temperature and pH of 7 in an associated form which can be described as a quadruple helix. When the pH is lowered to 3 and the temperature raised, partial melting to a single stranded form occurs. Ideally, one would like to maintain the temperature above 90°C for some time to insure a complete phase transition. However, the infrared cells used will not allow a temperature this high to be reached without significant sample leakage; thus, the samples were heated quickly to above 70°C, and quenched back to 10°C. The phase transition was followed by UV-CD: the single stranded form shows a negative couplet (257/290 nm), whereas the multiple strand form has positive CD at 290 nm.

The multistranded poly(rG) shows an infrared absorption at 1688 cm^{-1} with a low frequency shoulder at about 1660 cm^{-1}. The VCD of this form is small, and consists of a broad negative intensity at 1615 cm^{-1}, positive VCD at 1675 cm^{-1}, and slightly negative VCD at 1700 cm^{-1} (cf. Fig. 4). The high frequency zero crossing of the VCD trace occurs at the absorption maximum. Using published coordinates (Arnott et al., 1974) for the quadruple strand, the low frequency infrared absorption maximum was predicted properly by DECO calculations, yet the main VCD couplet was predicted with the wrong sign pattern.

After heating poly(rG) at low pH to above 70°C, and quickly cooling it to 10°C, the spectra shown in Fig. 4 (*dashed traces*) are obtained. CD data indicate that under these conditions, the polymer is mostly single stranded, with a small fraction of quadruple-strands. The infrared absorption is shifted to 1698 cm^{-1}, and the positive VCD band is

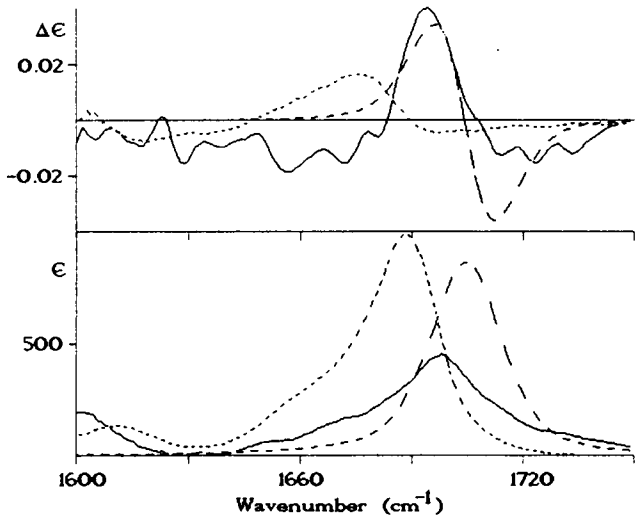

FIGURE 4 Observed and calculated infrared VCD (*top*) and absorption (*bottom*) spectra of poly(rG) in D_2O. Experimental conditions: solid line: [poly(rG)] = 10 mg/ml, pH = 3.5, T = 10°C (after heating to 70°C), path length = 100 μm, data expressed in molar extinction units per base. Long dashes: [poly(rG)] = 30 mg/ml, pH = 7.0, T = 20°C, pathlength = 50 μm, data expressed in molar extinction units per base. Computational conditions (*short dashes*): ϵ_{max}(1655 cm^{-1}) = 980 liters/(mol cm), w = 16 cm^{-1}. DECO computation for 30 bases, results normalized to one base.

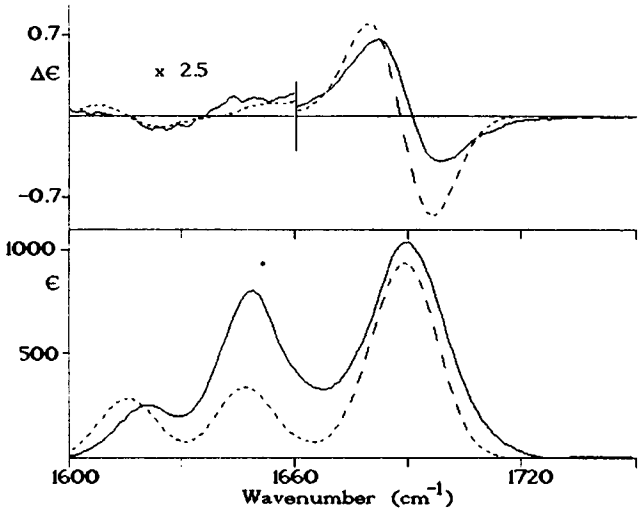

FIGURE 5 Observed and calculated infrared VCD (*top*) and absorption (*bottom*) spectra of poly(rG)·poly(rC) in D_2O/sodium cacodylate buffer. Experimental conditions (*solid traces*): [poly(rC)·poly(rG)] = 30 mg/ml, T = 20°C, path length = 50 μm, data expressed in molar extinction units per base pair. Computational conditions (*dashed traces*): Cytosine: ϵ_{max}(1650 cm^{-1}) = 980 liters/(mol cm), w = 18 cm^{-1}. Cytosine: ϵ_{max}(1612 cm^{-1}) = 450 liters/(mol cm), w = 15 cm^{-1}. Guanosine: ϵ_{max}(1655 cm^{-1}) = 980 liters/(mol cm), w = 18 cm^{-1}. NECO computation for C=O groups, independent DECO calculation for C=C—C vibration, 20 base pairs, results normalized to one base pair.

narrower and more intense. This VCD spectrum can be reproduced reasonably well with simple DECO calculation (not shown), which also predicts properly the shift of the absorption maximum to much higher frequency. Of all observed VCD spectra reported here, the absolute magnitude of the signal is least reliable for poly(rG), because heating the sample above 70°C causes some solvent leakage from the cell, and consequently, not very accurately known concentrations. Nevertheless, the agreement between observed and calculated spectral data is acceptable.

Poly(rG)·poly(rC)

Poly(rC)·poly(rG) exists in buffered, low ionic strengths (0.05 M NaCl) aqueous solution in a double stranded, right-handed conformation belonging to an A-family of conformations. The observed and calculated VCD and absorption spectra are shown in Fig. 5. The observed spectra are somewhat different from those observed for B form poly(dG)·poly(dC); thus, it appears that A and B form polymers can be distinguished *via* VCD (Wang and Keiderling, 1992). The main features of poly(rC)·poly(rG) are three absorptions at 1621, 1649, and 1690 cm^{-1}, whereas in poly(dG)·poly(dC), the two high frequency peaks overlap, and the low frequency peak is observed only as a shoulder. The VCD spectrum of the ribonucleic acid consists of three couplets: a weak positive one with a zero crossing at 1617 cm^{-1}, a weak negative one with a zero crossing at 1636 cm^{-1}, and the main positive-negative signal with a zero crossing at 1691 cm^{-1}. This signal is positively biased. The corresponding VCD signal of poly(dG)· poly(dC) is much more conservative.

It is interesting to note that guanosine and cytosine monomers have their absorption peaks below 1660 cm^{-1}, at 1655 and 1650 cm^{-1}, respectively. In poly(rG)·poly(rC), however, the highest wavenumber peak is observed at 1690 cm^{-1}. Thus, this peak must be due to the interaction of the C=O dipoles, which shifts some exciton components to higher and some others to lower frequencies. This behavior is reproduced well by the calculated results, depicted in Fig. 5. It is also noteworthy that for guanine and adenine, although they have relatively similar structures, the vibrations in the 1600 to 1650 cm^{-1} region are quite different. Adenine, devoid of a C=O group, has a strong C=C—C stretching vibration at 1625 cm^{-1}, with a dipole moment comparable of that of a C=O stretching vibration. Guanine, on the other hand, only has the C=O stretching vibration in the 1600 to 1700 cm^{-1} region, and the corresponding C=C—C vibration occurs below 1600 cm^{-1} (Tsuboi et al., 1973).

The calculated VCD spectra were obtained using the NECO equations (Eqs. 7–9), and C=O stretching vibrations for the guanine (1655 cm^{-1}) and cytosine (1650 cm^{-1}) groups, and an independent C=C—C stretching mode on cytosine (1612 cm^{-1}), as discussed above. These results are shown in Fig. 5. DECO calculations reproduced the VCD intensity not quite as well, and predicted the main infrared absorption and the zero crossing of the main couplet at significantly lower wavenumber. PSS calculations were also carried out, using the three vibrations enumerated above as the transitions in one subunit. The PSS calculations (not shown) reproduced the low frequency VCD slightly better

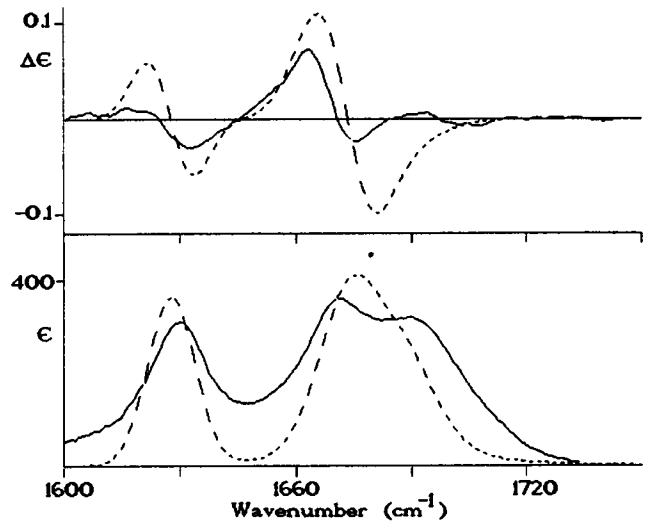

FIGURE 6 Observed and calculated infrared VCD (*top*) and absorption (*bottom*) spectra of poly(rA)·poly(rU) in D_2O/sodium cacodylate buffer. Experimental conditions (*solid traces*): [Poly(rA)·poly(rU)] = 30 mg/ml, T = 20°C, path length = 50 μm, data expressed in molar extinction units per base pair. Computational conditions (*dashed traces*): Uracil: ϵ_{max}(1687 cm^{-1}) = 1100 liters/(mol cm), w = 18 cm^{-1}. Uracil: ϵ_{max}(1654 cm^{-1}) = 650 liters/(mol cm), w = 18 cm^{-1}. Adenosine: ϵ_{max}(1623 cm^{-1}) = 1100 liters/(mol cm), w = 14 cm^{-1}. Three independent DECO calculation for C=C—C, O=C—C, and O=C—N vibrations, 20 base pairs, results normalized to one base pair.

than the NECO calculations, but reproduced the infrared absorption intensities less accurately.

Poly(rA)·poly(rU)

Calculations on uracil-containing polymers are more difficult to carry out, since the two C=O groups on the U residue (or T residue in DNA) are in such close proximity that their coupling is exaggerated in standard DECO calculations. Our previous attempt to calculate poly(dA-dT)·poly(dA-dT) spectra failed, and the poly(dA)·poly(dT) spectra suffered from a misassignment of the two C=O frequencies (Zhong et al., 1990). Here, we wish to report a corrected calculation for poly(dA)·poly(dU). The only vibration of the A residues is taken to be identical to the one described above for poly(rA). However, the two carbonyl stretching vibrations in U require special attention, since the appear to mix more with base stretching modes than C=O modes on other nucleotides. The AMPAC calculations give the following contributions of ring vibrations to the modes observed at the wavenumbers listed (cf. Fig. 2 *c*)as follows. $O{=}C_2$: 1687 cm^{-1}, 86% $O{=}C_2$, 14% C_2-N_1 (uracil); $O{=}C_1$: 1654 cm^{-1}, 79% $O{=}C_4$, 21% C_4—C_5 (uracil); $C_4{=}C_5$—C_6: 1625 cm^{-1}, 73% $C_4{=}C_5$, 27% C_5—C_6 (adenine).The approximate directions of the dipole moments are given in Fig. 2. Since these three vibrations are relatively far apart, interactions between them were expected to be relatively small, and therefore, independent DECO calculations were carried out for all three vibrations. The results for poly(rA)·poly(rU) are shown in Fig. 6, and the agreement between observed and calculated spectra is very good.

CONCLUSIONS

For the single stranded nucleic acids poly(rC), poly(rA), and poly(rG), a simple degenerate extended coupled oscillator formalism can be used to reproduce the VCD features observed. The DECO approach also is valid for poly(rA)·poly(rU), but here, the directions of the transition moments need to be adjusted carefully. In C and G containing nucleic acid polymers, it is best to treat the carbonyl groups as nondegenerate coupled oscillators.

Support of this research by the National Institute of General Medical Sciences (GM 28619) and by several City University of New York faculty research awards is gratefully acknowledged.

REFERENCES

Annamalai, A., and T. A. Keiderling. 1987. VCD of poly(ribonucleic acids). A comparative study in aqueous solution. *J. Am. Chem. Soc.* 109:3125–3132.

Arnott, S., R. Chandarasekaran, and C. M. Martilla. 1974. Structures of polyinosinic acid and polyguanylic acid. *Biochem. J.* 141:537–543.

Arnott, S., R. Chandarasekaran, and A. G. W. Leslie. 1976. Structures of the single stranded polyribonucleotide polycytidylic acid. *J. Mol. Biol.* 106: 735–748.

Baur, P., and T. A. Keiderling. 1993. Ab initio simulations of the VCD of coupled peptides. *J. Am. Chem. Soc.* In press.

Bayley, P. M., E. B. Nielsen, and J. A. Schellman. 1969. The rotatory properties of molecules containing two peptide groups: theory. *J. Phys. Chem.* 73:228–243.

Davydov, A. S. 1962. Theory of Molecular Excitations. McGraw Hill, New York.

Gulotta, M., J. D. Goss, and M. Diem. 1989. IR VCD in model deoxyoligonucleotides: observation of the B→Z phase transition and extended coupled oscillator calculations. *Biopolymers.* 28:2047–2058.

Holzwarth, G., and I. Chabay. 1972. Optical activity of vibrational transitions: a coupled oscillator model. *J. Chem. Phys.* 57:1632–1635.

Lee, O., and M. Diem. 1992. Infrared CD in the 6 μm spectral region: design of a dispersive spectrograph. *Anal. Instrum.* 20:23–43.

Moffit, W. 1956. Optical rotatory dispersion of helical polymers. *J. Chem. Phys.* 25:467–478.

Rhodes, W., 1961. Hypochromism and other spectral properties of helical polynucleotides. *J. Am. Chem. Soc.* 83:3609–3617.

Still, C. 1989. MacroModel software. Version 2.5. Columbia University, New York.

Tinoco, I. 1962. Theoretical aspects of optical activity: polymers. *In* Advances in Chemical Physics. I. Prigogine, editor. Interscience Publishers, New York. 4:113–160.

Tinoco, I. 1963. The exciton contribution to the optical rotation of polymers. *Radiat. Res.* 20:133–139.

Tinoco, I. 1964. Circular dichroism and rotatory dispersion curves for helices. *J. Am. Chem. Soc.* 86:297–298.

Tinoco, I., and R. W. Woody. 1963. Absorption and rotation of light by helical polymers: the effect of chain length. *J. Chem. Phys.* 38:1317–1325.

Tsuboi, M., S. Takahashi, and I. Harada. 1973. Physicochemical Properties of Nucleic Acids. Vol. 2. J. Duchesne, editor. Academic Press, New York. 91.

Wang, L., and T. A. Keiderling. 1992. VCD Studies of the A to B conformational transition in DNA. *Biochemistry.* 31:10265–10271.

Zhong, W., M. Gulotta, D. J. Goss, and M. Diem. 1990. DNA solution conformation via IR CD: experimental & theoretical results for B-family polymer. *Biochemistry.* 29:7485–7491.