assistance in adapting his program, DNRP53. We further appreciate numerous intriguing discussions with D. Klehr and the expert technical assistance of K. Maass.

REFERENCES

Adachi, Y., Käs, E., & Laemmli, U. K. (1989) *EMBO J, 8*, 3997–4006.
Anderson, J. N. (1986) *Nucleic Acids Res. 21*, 8513–8533.
Blasquez, V. C., Xu, M., Moses, S. C., & Garrard, W. T. (1989) *J. Biol. Chem. 264*, 21183–21189.
Bode, J., & Maass, K. (1988) *Biochemistry 27*, 4706–4711.
Bode, J., Pucher, H. J., & Maass, K. (1986) *Eur. J. Biochem. 158*, 393–401.
Chomczynski, P., & Sacchi, N. (1987) *Anal. Biochem. 162*, 156–159.
Cockerill, P. N., & Garrard, W. T. (1986) *Cell 44*, 273–282.
Cockerill, P. N., Yuen, M. H., & Garrard, W. T. (1987) *J. Biol. Chem. 262*, 5394–5397.
Darby, M. K., Herrera, R. E., Vosberg, H. P., & Nordheim, A. (1986) *EMBO J, 5*, 2557–65.
Dinter, H., & Hauser, H. (1987) *Eur. J. Biochem. 166*, 103–109.
Dirks, W., Mittnacht, S., Rentrop, M., & Hauser, H. (1989) *J. Interferon Res. 9*, 125–133.
Duggleby, R. G. (1984) *Comput. Biol. Med. 14*, 447–455.
Duggleby, R. G. (1988) *J. Theoret. Biol. 130*, 123–124.
Forrester, W. C., Novak, U., Gelinas, R., & Groudine, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86* (14), 5439–5443.
Gasser, S. M., & Laemmli, U. K. (1986) *EMBO J. 5*, 511–518.
Gasser, S., & Laemmli, U. K. (1987) *Trends Genet. 3*, 16–22.
Gross, D. S., & Garrard, W. T. (1987) *Trends Biochem. Sci. 12*, 293–297.
Hofmann, J. F.-X., Laroche, T., Branas, A. H., & Gasser, S. M. (1989) *Cell 57*, 725–737.
Homberger, H. P. (1989) *Chromosoma 98*, 99–104.
Izaurralde, E., Mirkovitch, J., & Laemmli, U. K. (1988) *J. Mol. Biol. 200*, 111–126.
Izaurralde, E., Käs, E., & Laemmli, U. K. (1989) *J. Mol. Biol. 210*, 573–585.
Jarman, A. P., & Higgs, D. R. (1988) *EMBO J. 7*, 3337–3344.
Käs, E., & Chasin, L. A. (1987) *J. Mol. Biol. 198*, 677–692.
Käs, E., Izaurralde, E., & Laemmli, U. K. (1989) *J. Mol. Biol. 210*, 587–599.
Klehr, D., & Bode, J. (1988) *Mol. Gen. (Life Sci. Adv.) 7*, 47–52.
Lutfalla, G., Blanc, H., & Bertolotti, R. (1985) *Somatic Cell Mol. Gen. 113*, 223–238.
Mirkovitch, J., Mirault, M. E., & Laemmli, U. K. (1984) *Cell 39*, 223–232.
Phi-Van, L., & Strätling, W. H. (1988) *EMBO J. 7*, 655–664.
Probst, H., & Herzog, R. (1985) *Eur. J. Biochem. 146*, 167–171.
Scheuermann, R. H., & Chen, U. (1989) *Genes Dev. 3*, 1255–1266.
Sperry, A. O., Blasquez, V. C., & Garrard, W. T. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5497–5501.
Stief, A., Winter, D. M., Strätling, W. H., & Sippel, A. E. (1989) *Nature (London) 341*, 343–345.
Stockhaus, J., Eckes, P., Blau, A., Schell, J., & Willmitzer, L. (1987) *Nucleic Acids Res. 158*, 3479–3491.
Tsutsui, K., Tsutsui, K., & Muller, M. T. (1988) *J. Biol. Chem. 263*, 7235–7241.
van Assendelft, G. B., Hanscombe, O., Grosveld, F., & Greaves, D. R. (1989) *Cell 56*, 969–977.
Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C., & Axel, R. (1977) *Cell 11*, 223–232.
Wirth, M., Bode, J., Zettlmeissl, G., & Hauser, H. (1988) *Gene 73*, 419–426.
Xu, M., Hammer, R. E., Blasquez, V. C., Jones, S. L., & Garrard, W. T. (1989) *J. Biol. Chem. 264*, 21190–21195.

DNA Solution Conformation via Infrared Circular Dichroism: Experimental and Theoretical Results for B-Family Polymers†

Wenxin Zhong, Miriam Gulotta, Dixie J. Goss, and Max Diem*
Department of Chemistry, City University of New York, Hunter College, 695 Park Avenue, New York, New York 10021
Received November 13, 1989; Revised Manuscript Received March 28, 1990

ABSTRACT: Infrared (vibrational) circular dichroism (VCD) has been observed for the DNA models $d(CG)_5$, poly(dG-dC)·poly(dG-dC), poly(dG)·poly(dC), poly(dA-dT)·poly(dA-dT), and poly(dA)·poly(dT) in the B-conformation in buffered, aqueous solution. The observed results are quantitatively interpreted in terms of the exciton model for coupled carbonyl stretching vibrational states.

Circular dichroism in vibrational transitions, observed in the infrared spectral region, has recently been applied to the elucidation of nucleic acid solution conformation (Annamalai & Keiderling, 1987; Gulotta et al., 1989). For nucleic acids, both electronic CD and VCD signals arise from the dipolar coupling of (electronic or vibrational) transitions localized within the bases. These transitions interact, or couple, and produce absorption and circular dichroism features that are governed by the exciton theory (Tinoco, 1963).

However, VCD offers the advantage over conventional electronic CD, observed in the visible/UV spectral region, of more and narrower spectral features. In addition, for certain transitions, such as carbonyl stretching vibrations, the vibrational modes giving rise to the spectral features are well localized, and the dipole change occurs in a direction that is nearly parallel to the bond connecting the carbon and oxygen

†This research was supported by the National Science Foundation (CHE 86-07070 to D.J.G.), an American Heart Association–NYC Established Investigatorship (D.J.G.), the National Institutes of Health (GM 28619 to M.D.), and several City University of New York Faculty Research Awards. The VCD instrument utilized was constructed with funds derived from a previous NSF grant to M.D. (CHE 86-07934). DNA structural computations were carried out on the departmental molecular modeling facility, purchased under RCMI Grant RR 03037.

0006-2960/90/0429-7485$02.50/0 © 1990 American Chemical Society

atoms. Thus, the infrared VCD observed in the carbonyl stretching vibrations can be interpreted in terms of the relative positions of the carbonyl groups of the purine and pyrimidine bases, and consequently, information on the base alignment can be obtained.

RNA models were studied via VCD in the carbonyl stretching region, 1550–1750 cm^{-1} (Annamalai & Keiderling, 1987), and the dependence of the observed signals on base composition was established. We reported previously the first observation of VCD in DNA models containing cytosine and guanine residues (C-G polymers) in the same spectral region (Gulotta et al., 1989), along with the first observation of the changes in the VCD features due to a salt-induced transition from the B- to the Z-form in poly(dG-dC)·poly(dG-dC). In addition, the framework for the interpretation of the observed VCD spectra, the "extended coupled oscillator" (ECO) model, was developed in that paper. In the ECO model, the coupling of interacting, degenerate vibrational transitions creates VCD intensities that may be computed easily (vide infra). The ECO is a straightforward extension of the classical coupled oscillator (CO) model (Tinoco, 1963), which has its roots in the exciton theory of optical activity. Our original ECO calculations suggested that the VCD observed in DNA between 1550 and 1750 cm^{-1} is entirely due to coupled carbonyl stretching states, since these calculations reproduced the sign and magnitude of the observed signals. We found that the ECO formalism is a better model for the description of the VCD of a large polymer than the simple two-group CO model, although the signs of the RNA VCD signals were predicted properly with the CO model (Annamalai & Keiderling, 1987).

In this paper, we report the dependence of VCD and infrared absorption spectra of base composition and sequence in model DNA samples, which exist in aqueous (D_2O) solution in a B-type conformation. We find that C-G polymers give quite different spectra than adenine/thymine- (A-T) containing polymers. In addition, the VCD spectra show sensitivity toward base sequence, since the spectra of poly(dG)·poly(dC) are different from those of poly(dG-dC)·poly(dG-dC).

To facilitate the interpretation of the spectra and spectral differences, computations are carried out to reproduce observed spectra from structural data alone. We find, in general, excellent agreement between observed and theoretical spectra. Thus, it appears that VCD can be used as a novel spectroscopic technique to sample local effects in DNA solution structure in an extremely sensitive way.

Experimental Procedures

All infrared absorption and VCD spectra were recorded on the Hunter College dispersive VCD instrument, the design and performance of which have been reported before in detail (Diem et al., 1988; Diem, 1990). Commercial samples (Pharmacia, Inc.) were lyophilized from D_2O and dissolved in cacodylate buffer in D_2O to a concentration of between 20 and 40 mg/mL. Samples were contained between two 19 mm diameter CaF_2 disks held apart by a 25-μm Teflon spacer. Total sample volume required is 8 μL. The temperature of the samples during spectral data acquisition was ca. 30 °C.

All VCD spectra shown are the averages of between 10 and 20 individual scans, each taking about 10 min. Base lines obtained for the buffer have been subtracted from the sample spectra.

Theory and Computational Procedures

One of the dominant mechanisms that give rise to VCD intensities is the coupling of polar, vibrational transitions such as the purine or pyrimidine C=O stretching vibrations, arranged in a helical geometry. If one vibrational quantum is absorbed by these degenerate oscillators, the resulting vibrationally excited state 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.

In the ECO description of the optical activity of n interacting dipoles, the rotational strength R, and hence the VCD intensities, is given by

$$R_k = -(\pi\nu_0/c)\sum_{i=1}^{n}\sum_{j>i}^{n}\mathbf{c}_{ik}\mathbf{c}_{jk}[\mathbf{T}_{ij}\cdot\mu_i \times \mu_j] \quad (1)$$

where the $\mathbf{c}_{ij}$ are the eigenvector components of the (dipole–dipole) interaction matrix given by

$$\mathbf{V}_{ij} = \frac{\mu_i\cdot\mu_j}{|\mathbf{T}_{ij}|^3} - \frac{3(\mu_i\cdot\mathbf{T}_{ij})(\mu_j\cdot\mathbf{T}_{ij})}{|\mathbf{T}_{ij}|^5} \quad (2)$$

$\mathbf{T}_{ij}$ is the distance vector between dipole μ_i and μ_j, ν_0 is the center frequency (in wavenumbers) of the unperturbed transition, c is the velocity of light, and the subscript k refers to the kth exciton component ($1 < k < n$). The infrared absorption intensities can be obtained from the dipole strengths D, defined by

$$D_k = \sum_{i=1}^{n}\mathbf{c}_{ik}^2\mu_i^2 + \sum_{i=1}^{n}\sum_{j>i}^{n}\mathbf{c}_{ik}\mathbf{c}_{jk}(\mu_i\cdot\mu_j) \quad (3)$$

The computations of dipole and rotational strengths for the polymers are carried out by use of Cartesian coordinates of the C and O atoms of the various carbonyl groups on the purine and pyrimidine bases, the transition frequency ν_0 of an unperturbed carbonyl transition, and its monomeric dipole strength D, which is proportional to μ^2. For the conversion between rotational (or dipole) strengths (in $esu^2\ cm^2$) and the extinction coefficients (in L mol^{-1} cm^{-1}), the approximations

$$D \approx (9.2 \times 10^{-39})\pi\epsilon_{max}w/\nu_0 \quad (4)$$

and

$$R \approx (2.3 \times 10^{-39})\pi\Delta\epsilon_{max}w/\nu_0 \quad (5)$$

were used. Here, w denotes the width of the observed VCD or absorption band. Equations 4 and 5 hold for Lorenzian band shapes; for Gaussian bands, the factor π needs to be replaced by $\sqrt{\pi}$. All calculations reported here were carried out for double-stranded models containing 10 residues in each strand, such as $d(CG)_5$·$d(CG)_5$ or $d(G)_{10}$·$d(C)_{10}$. In the case of $d(CG)_5$·$d(CG)_5$, we compare the observed and calculated spectra directly, since experimental VCD data have been measured for the decamer (Gulotta et al., 1989). For all other samples, the comparison between computed and observed spectra was carried out by normalizing the observed polymer and computed decamer spectra to one base pair absorption and VCD. This procedure is necessary since decamers of the desired composition are not readily available, whereas the polymers are. Comparing decamers and polymers is legitimate, since it was found previously that the spectra of a double-stranded decamer exhibit VCD and absorption features very similar to the ones observed for a polymer of well over 100 base residues.

For the computations reported below, it was assumed that the VCD intensities observed between 1550 and 1750 cm^{-1} are mostly due to coupled C=O stretching states. This is, of course, a somewhat crude assumption, since C=C and C=N

Table I: Vibrational Frequencies (cm^{-1}) of the Deoxynucleotide Monomers[a] in D_2O

	freq[a]	ϵ_{max}[b]	assignment[c]
d(A)	1580	300	C=C, C=N stretch
	1630	1500	C=C, C=C stretch
d(T)	1630	980	
	1665	1150	C=O, C=C
	1690 sh	700	C=O, C=N
d(C)	1615	600	
	1645	1250	C=O, C=N
d(G)	1580	750	C=C, C=N
	1665	780	C=O, C=C
d(G)-d(C)	1580	770	
	1615	410	
	1650 br	950	

[a] All frequencies are believed accurate to ±3 cm^{-1}. Abbreviations: sh, shoulder; br, broad. [b] ϵ values are approximate, due to difficulties in determining the accurate cell path length and sample concentrations (due to hydration). [c] Assignments from Tsuboi et al. (1973). The assignments give the dominant terms in the potential energy distribution. These calculations were performed for residues U, C, A, and G.

stretching motions of the aromatic ring systems occur in this frequency range as well. However, our own infrared absorption data on the monomers and 1:1 molar mixtures of monomers and previous vibrational assignments and normal coordinate calculations (Tsuboi et al., 1973) suggest that the strong features above 1650 cm^{-1} in the absorption and VCD spectra are due to C=O stretching vibrations and that all C=C and C=N stretching vibrations occur below 1640 cm^{-1}. The monomer vibrational frequencies are listed in Table I.

There are two additional arguments in favor of the assumption that the observed VCD is due mostly to C=O coupling. First, using the correct geometries of the C=O oscillators in poly(dG-dC)·poly(dG-dC), we show that the C=O stretching exciton components cover a range of over 60 cm^{-1}. Considering a half-width of each components of about 20 cm^{-1}, it follows that any vibrational or VCD intensity between 1600 and 1680 cm^{-1} is, possibly, due to C=O exciton components.

Second, we note that the large VCD feature observed by Keiderling in poly(A) (Annamalai & Keiderling, 1987), which is devoid of carbonyl oscillators, has its negative VCD maximum at just above 1635 cm^{-1}. In poly(dA)·poly(dT) positive VCD intensity is observed at this wavenumber, and it is unlikely that this signal is due to the C=C and C=N vibrations which cause the signal in poly(A). Thus, we believe that the dominant features of the observed VCD spectra in the four polymers reported here are indeed due to C=O exciton states and that C=C and C=N stretching vibrations may be ignored to a first approximation. However, it is conceivable that in certain favorable geometries, e.g., in poly(A), they may interact to produce effects as well.

All calculations were carried out with a FORTRAN 77 program, written in our laboratory, which calculates the dipole interaction matrix and rotational and dipole strengths for all exciton components from atomic coordinates of the dipoles and the carbonyl stretching transition moment available from the monomer. The computations are very fast and require only a few seconds of execution time on a fast personal computer for 20 interacting dipoles. Atomic coordinates of the carbonyl groups were derived from a generalized canonical B-type DNA structure, the basis of fiber diffraction data from the Arnott data base (Arnott et al., 1976).

Results and Discussion

(*a*) *Oligomers.* Figure 1 shows observed infrared absorption and VCD data for the double-stranded decamer $d(CG)_5$·

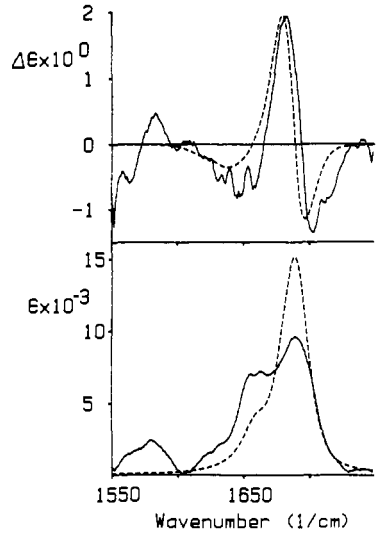


FIGURE 1: Calculated (dashed traces) and observed (solid traces) infrared VCD (top) and absorption (bottom) spectra of $d(CG)_5$. See text for experimental and computational details.

$d(CG)_5$, henceforth referred to simply as $d(CG)_5$. As mentioned above, this is the only small polymer reported in this study, and all the following spectra were obtained on polymers with a molecular weight of well over 100 000. Also depicted in Figure 1 are the computed absorption and VCD spectra for $d(CG)_5$. The ordinate units, ϵ and $\Delta\epsilon$, are not normalized with respect to the number of base pairs; i.e., they are expressed in L cm^{-1} (mol of polymer)$^{-1}$, vide infra. For all GC polymers, the carbonyl transition frequency and intensity were taken from a 1:1 CMP/GMP mixture in D_2O (ν_0 = 1650 cm^{-1}; ϵ_{max} = 950 L mol^{-1} cm^{-1}), since the individual C=O stretching peaks observed for both CMP and GMP coalesce into one broad peak in the mixture.

For computed spectra, band shapes were employed that are 50:50 mixtures of Lorenzian and Gaussian contours. It was found that such a mixed composition improves the fit between observed and computed spectra, although pure Gaussian or pure Lorenzian band shapes give satisfactory results as well. The band composition, in fact, is the only adjustable parameter in the calculations; all other input data, such as coordinates or dipole moments, were treated as constants. In light of this fact, the agreement between observed and computed data is striking. The magnitude and positions of the calculated VCD spectra are within experimental errors. The agreement between calculated and observed absorption data is not quite as impressive, although peak positions and the total integrated absorption strength are well reproduced.

Figure 2 depicts the individual exciton components, as well as their sum, which is, of course, the computed spectrum of Figure 1. The largest individual intensity is nearly an order of magnitude larger than the resulting sum of all components. The cancellation of positive and negative VCD intensities to reproduce the final spectrum, as shown in Figure 2, depends on the accurate prediction of both the intensities and positions of each component. In view of this fact, the agreement achieved for $d(CG)_5$ is even more remarkable.

In order to further investigate how the cancellation of VCD intensities determines the overall computed band shapes, we have computed VCD spectra for even smaller (G-C) polymers. The simplest model system used was a single GC base pair. Since the two carbonyl groups in a GC base pair are virtually coplanar, there is very little chirality induced by their coupling. Furthermore, the coupling energy is relatively small, since the distance of the centers of masses of the two oscillators is large

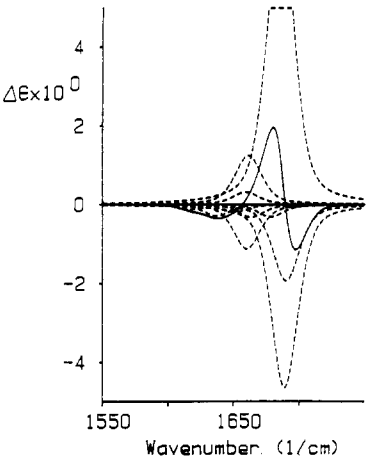


FIGURE 2: Individual VCD components (dashed) and their sum (solid) used in the computation of the $d(CG)_5$ model.

Table II: Frequencies, Infrared Absorption, and VCD Intensities for d(CG)·d(CG) and d(GC)·d(GC)

		IR intensity		VCD intensity	
	freq (cm^{-1})	esu^2 cm^2 × 10E38	L mol^{-1} cm^{-1}	esu^2 cm^2 × 10E42	L mol^{-1} cm^{-1}
G—C	1667	98.5	2276	−71.2	−0.66
¦ ¦	1664	77.4	1782	75.5	0.70
C—G	1658	0.8	18	−1.6	−0.01
	1608	0.1	1	−2.8	−0.02
C—G	1702	166.1	3912	15.8	0.15
¦ ¦	1645	0.2	6	−1.8	−0.02
G—C	1631	5.3	119	−1.4	−0.01
	1622	1.8	41	−12.6	−0.11

(about 5 Å). Thus, a single GC base pair will not produce any observable VCD.

Next, the interaction of four carbonyl groups in the following two double-stranded dimers was considered:

```
dG - dC          dC - dG
 ¦    ¦           ¦    ¦
dC - dG          dG - dC

 d(CG)            d(GC)
```

In these two segments, the orientation of the four carbonyl groups differs drastically, and consequently, entirely different VCD patterns are expected. These results are summarized in Table II. In double-stranded d(GC), the four carbonyl groups are virtually all parallel, albeit in different planes. Thus, the induced chirality is low, but the coupling is large, as manifested by the large splitting in the IR absorption spectra (cf. Table II). In double-stranded d(CG), on the other hand, the second set of carbonyl groups are twisted nearly 90° with respect to the first (lower) set of dipoles. This arrangement results in large induced chirality (Table II) with a sign pattern similar to the one observed in poly(dG-dC)·poly(dG-dC), poly(dG)·poly(dC), and $d(CG)_5$.

The arrangement pattern of dipoles found in d(CG) and d(GC) is prototypical for B-type double helixes consisting of CG bases: there will be, between consecutive base pairs, either the "near-parallel" alignment of dipoles or the "near 90° twist" between them. Thus, it is of interest to examine the two segments

```
dC - dG - dC          dG - dC - dG
 ¦    ¦    ¦           ¦    ¦    ¦
dG - dC - dG          dC - dG - dC

d(GCG)·d(CGC)         d(CGC)·d(GCG)
```

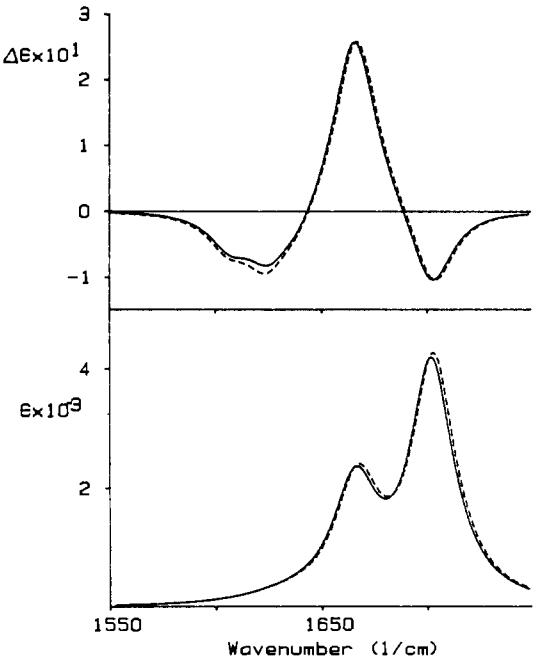


FIGURE 3: Calculated VCD (top) and absorption (bottom) spectra of d(GCG)·d(CGC) (solid) and d(GCG)·d(GCD) (dashed).

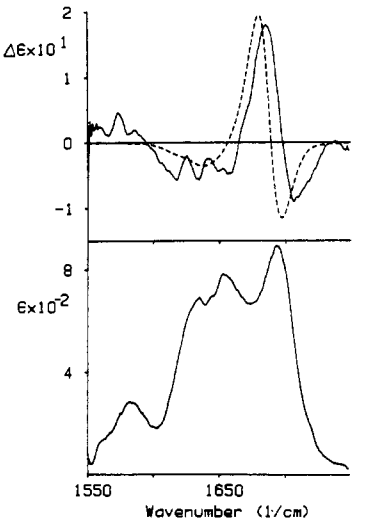


FIGURE 4: Calculated (dashed trace) and observed (solid trace) infrared VCD (top) and absorption (bottom) spectra of poly(dG-dC)·poly(dG-dC). Data are normalized to one base pair extinction coefficients.

The calculated VCD and infrared traces are shown in Figure 3. From the virtually identical spectra it is clear that the order, in which the dipole pattern occurs, is immaterial: both models contain one set of near-parallel and one set of near 90° twist in different order and exhibit virtually identical computed IR VCD and absorption spectra. In addition, these species exhibit overall VCD spectral features similar to the ones observed for $d(CG)_5$. Thus, it appears that even short segments (trimer and larger) will show spectra similar to the ones observed for oligomers and polymers and that the agreement between observed and calculated data displayed in Figure 1 is not coincidental.

(*b*) *Polymers.* We now turn to the discussion of the polymers. Observed infrared absorption and VCD spectra for poly(dG-dC)·poly(dG-dC), poly(dG)·poly(dC), poly(dA-dT)·poly(dA-dT), and poly(dA)·poly(dT) are shown in Figures 4–7. Also shown are VCD spectral simulations for the structurally analogous decamers. Here, both experimental and

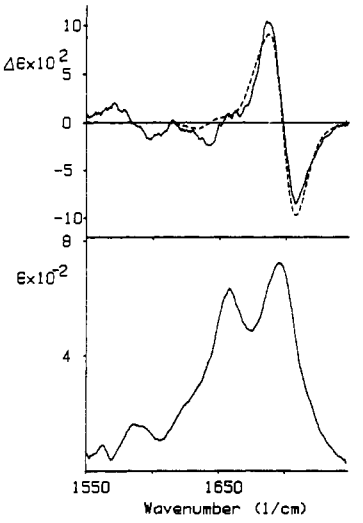


FIGURE 5: Calculated (dashed trace) and observed (solid trace) infrared VCD (top) and absorption (bottom) spectra of poly(dG)·poly(dC). Data are normalized to one base pair extinction coefficients.

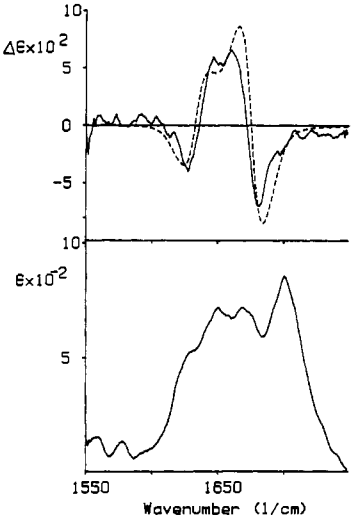


FIGURE 6: Calculated (dashed trace) and observed (solid trace) infrared VCD (top) and absorption (bottom) spectra of poly(dA)·poly(dT). Data are normalized to one base pair extinction coefficients.

theoretical data are expressed in units of L cm^{-1} (mol of base pairs)$^{-1}$, i.e., the spectra are normalized to a one base pair chain length.

The most significant observation about the four polymer VCD spectra is the fact that they all differ significantly. Further inspection of Figures 4–7 reveals that the two G-C polymers are similar to each other, but very different from those of the A-T polymers. Thus, it is clear that VCD distinguishes the base composition and the composition of a strand in the simple DNA models discussed here. Electronic CD, on the other hand, is less sensitive toward base composition, since B-type polymers give very similar CD spectra. The remainder of this paper will deal with the implications of the spectral differences observed for the different polymers and an attempt to interpret the observed spectral differences.

Both G-C polymers, poly(dG-dC)·poly(dG-dC) and poly(dG)·poly(dC), show VCD spectra that are dominated by a large negative–positive couplet (1700 and 1670 cm^{-1}, respectively; cf. Figures 4 and 5), with a broader, slightly negative signal between 1580 and 1650 cm^{-1}. The zero crossing of the VCD trace in the main couplet occurs nearly exactly under the maximum of the absorption spectrum. The absorption feature is very different than that of a CMP/GMP

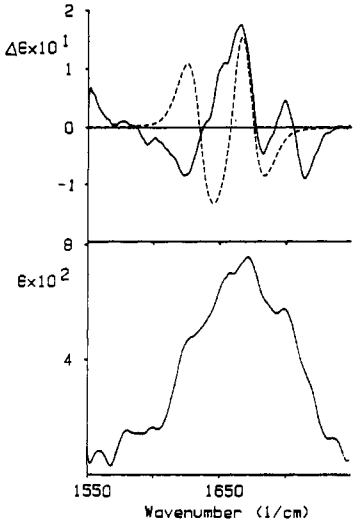


FIGURE 7: Calculated (dashed trace) and observed (solid trace) infrared VCD (top) and absorption (bottom) spectra of poly(dA-dT)·poly(dA-dT). Data are normalized to one base pair extinction coefficients.

mixture, indicating the large coupling which occurs in the C=O stretching modes. Furthermore, the C=C and C=N stretching vibrations, which are strong in GMP, are barely observed in the polymer.

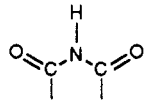
In the VCD spectrum of poly(dG-dC)·poly(dG-dC), the positive portion of the spectrum is larger than the negative part by about a ratio 2:1, whereas in the VCD spectrum of poly(dG)·poly(dC) the positive and negative intensities are nearly identical. Over the entire C=O stretching region, however, the integrated VCD intensities nearly equal zero in both cases, as expected from the exciton theory.

The calculated VCD intensities in Figures 4 and 5 are not drawn to scale but differ from the observed intensities by a factor of 2 in poly(dG-dC)·poly(dG-dC) and a factor of 4 in poly(dG)·poly(dC). In both cases, the calculated VCD intensities are too large, although all data are normalized to one base pair. However, it should be pointed out that the absolute determination of experimental VCD intensities is still somewhat ambiguous, mostly due to the uncertainty in the absorption cell path length and in the exact composition of the sample (for example, the content of sodium ions, inert salts, or water). It is gratifying to see, however, that the computed VCD spectra reproduce the observed intensity patterns extremely well as indicated by the well-balanced couplet in poly(dG)·poly(dC) (Figure 5) and the more positive couplet with an additional low-frequency negative signal in poly(dG-dG)·poly(dG-dC), Figure 4. Finally, we would like to point to a difference in the VCD spectra of double-stranded $d(CG)_5$ (Figure 1) and poly(dG-dC)·poly(dG-dC), Figure 5. The VCD couplet in the polymer is shifted to higher wavenumber, as can be seen by comparison with the computed spectra. Such a difference could be due to changes in the exciton levels between the decamer and the polymer or even to slight structural changes between the samples.

We now turn to the discussion of the A-T polymers. Figures 6 and 7 show the IR and VCD spectra of poly(dA-dT)·poly(dA-dT) and poly(dA)·poly(dT). As pointed out before, these VCD spectra are quite different from the traces displayed for the G-C polymers in Figures 4 and 5. In poly(dA)·poly(dT), the highest frequency absorption peak exhibits virtually no VCD intensity, and a large negative–positive couplet (1675/1650 cm^{-1}) is followed by a positive–negative couplet (1640/1625 cm^{-1}). In poly(dA-dT)·poly(dA-dT), there is an

additional negative–positive couplet centered at about 1700 cm^{-1} in the absorption peak which shows no signal in poly(dA)·poly(dT). The absorption spectra of poly(dA-dT)·poly(dA-dT) appear to be very poorly resolved.

The computation of the spectral features of A-T polymers is much more complicated than that of G-C polymers. This is due to the fact that the adenosine residue bears no carbonyl group at all but the thymine residue has two carbonyl groups, which are in close proximity and in ideal geometry for coupling:



In principle, there exist several different approaches to the simulation of A-T VCD spectra. The first approach assumes that all carbonyl groups (i.e., the C2 and C4 carbonyl groups at all thymine residues) have an unperturbed absorption frequency of ca. 1650 cm^{-1} as in GMP and the computational procedure discussed for the G-C polymers may be used unaltered. In this case, the two carbonyl groups located on one thymine residue couple so heavily that their coupling energy exceeds 100 cm^{-1}. Consequently, the computed VCD spectra for both $d(A)_{10}$·$d(T)_{10}$ and double-stranded $d(AT)_5$ are dominated by two large couplets, one located at about 1730 cm^{-1} and the other at about 1580 cm^{-1}. This clearly disagrees with the observed data, and the method is not useful.

The second approach employs reduced coupling between certain C=O groups, on the basis of the following additional information. The infrared absorption spectrum of thymine monophosphate (TMP) in D_2O shows four distinct peaks between 1550 and 1750 cm^{-1}, at 1598 (shoulder), 1634, 1668, and 1694 (shoulder) cm^{-1}. These data on TMP agree well with the Raman data in aqueous and organic media reported for the uracil residue (Bandekar & Zundel, 1983a,b). On the basis of relative intensities, one could assign the two dominant bands as the symmetric and antisymmetric combination states of two coupled C=O stretching vibrations, as expected from a coupled oscillator model.

If, indeed, the two frequencies observed for the TMP monomer (1634 and 1668 cm^{-1}) are due to coupled states, one needs to introduce a parameter that reduces the coupling energy from the computed value of over 100 cm^{-1} to the observed value of 34 cm^{-1}. Physically, the reduction of the coupling can be visualized as a "local" dielectric constant, since there is an N–H group, and not vacuum, separating the two carbonyl oscillators, or it could be visualized as a vibrational interaction force constant with the effect of reducing the coupling. This approach did not produce computed VCD spectra that were at all similar to the observed ones. Furthermore, previous studies strongly suggested that the different frequencies of the C2 and C4 carbonyl groups are not due to dipole–dipole interactions but to chemical differences and different involvement in interstrand hydrogen bonding.

Thus, a third alternative for the computation of the VCD spectra was followed. For this approach, we assume that the carbonyl group at atom C2, which is involved in hydrogen bonding to the adenosine base, has an unperturbed vibrational frequency of 1634 cm^{-1} and the C=O group at C4 has a frequency of 1668 cm^{-1}. This assignment is based on previous vibrational calculations of uracil by Bandekar and Zundel.

Since these frequencies are relatively far apart, we assume that the C2 and C4 carbonyl groups do not couple with each other. Thus, we model the spectra of poly(dA)·poly(dT) and poly(dA-dT)-poly(dA-dT) by independent strands of C2 and C4 carbonyl groups which do not interact.

By use of this procedure and coordinates of the thymine carbonyls in $d(A)_{10}$·$d(T)_{10}$, simulated spectra were obtained that are in excellent agreement with the ones observed for poly(dA)·poly(dT), Figure 6. The observed and computed VCD spectra are dominated by a negative–positive couplet at 1668 cm^{-1} and a positive–negative couplet at 1634 cm^{-1}. Their sum reproduced the relative negative and positive intensities and the frequencies associated with these peaks very well. However, the overall magnitude of the (observed) polymer VCD spectrum is lower, by a factor of 2, than the scaled computed data. Similar behavior was found earlier for the G-C polymers, vide supra.

Figure 7 shows observed and simulated VCD data for the remaining polymer, poly(dA-dT)·poly(dA-dT). The observed VCD and absorption spectra differ significantly from those of the polymers discussed so far. Furthermore, the computed VCD spectra do not agree well with the observed ones.

We attribute these differences to the possibility that poly(dA-dT)·poly(dA-dT), at the concentrations utilized, could partially or totally exist in a triple-stranded conformation, involving Hoogstein base pairs. Triple helixes are not likely to occur in any of the other polymers at the sample conditions used.

Triple helixes were observed previously by Keiderling and co-workers in poly(A)·poly(U)·poly(U) samples in RNA model studies (Keiderling, private communication). These triple helixes are, generally, characterized by much more complex VCD patterns, although the absorption spectra are well defined. Since the absorption spectrum in Figure 7 is broad and very ill defined, we conclude that the sample consisted of a mixture of double- and triple-stranded polymers. In that case, the differences between observed and computed VCD spectra may be explained.

Experimental proof for this hypothesis has to wait until VCD spectra can be collected from poly(dA-dT)·poly(dA-dT) samples at much lower concentrations and longer sample path lengths. This presents difficulties, since the D_2O background absorption at 6 μm exceeds 60% at path lengths of 25 μm. Instrumentation capable of measurements under even more adverse conditions, i.e., more light loss due to the solvent, is presently being developed in our laboratory.

Conclusions

We have shown that VCD data can be collected for a number of polynucleic acids from buffered aqueous solution. VCD can distinguish not only base composition (e.g., G-C- from A-T-containing DNA models) but also can discriminate structural features due to the composition of the strands. An interpretation of the VCD signals is possible in terms of a simple computational procedure that simulates observed spectral data.

Acknowledgments

We thank Dr. Mihaly Mezei for carrying out the coordinate calculations for all the model polymers studied. We also thank Prof. Tim Keiderling for the prepublication data on triple-helical RNA.

Registry No. $d(CG)_5$, 114742-74-4; d(CG), 15178-66-2; d(GC), 23405-83-6; d(GCG)·d(CGC), 90376-18-4; poly(dG-dC), 36786-90-0; poly(dG)·poly(dC), 25512-84-9; poly(dA)·poly(dT), 26966-61-0; poly(dA-dT)·poly(dA-dT), 24939-09-1.

References

Annamalai, A., & Keiderling, T. A. (1987) *J. Am. Chem. Soc. 109*, 3125.

Arnott, S., Campbell Smith, P. J., & Chandrasekara, R. (1976) in *Handbook of Biochemistry and Molecular Biochemistry*, 3rd ed. (Fasman, G. D., Ed.) Vol. II, p 412.
Bandekar, J., & Zundel, G. (1983a) *Spectrochim. Acta 39A*, 337.
Bandekar, J., & Zundel, G. (1983b) *Spectrochim. Acta 39A*, 343.
Diem, M. (1990) in *Vibrational Spectra and Structure* (Durig, J., Ed.) Vol. 19, Elsevier Science Publishers, Amsterdam, The Netherlands (in press).
Diem, M., Roberts, G. M., Lee, O., & Barlow, A. (1988) *Appl. Spectrosc. 42*, 20.
Faulkner, T. F. (1976) Ph.D. Dissertation, University of Minnesota.
Gulotta, M., Goss, D. J., & Diem, M. (1989) *Biopolymers 28*, 2047.
Tinoco, I. (1963) *Radiat. Res. 20*, 133.
Tsuboi, M., Takahashi, S., & Harada, I. (1973) in *Physicochemical Properties of Nucleic Acids* (Duchesne, J., Ed.) Vol. 2, p 91, Academic Press, New York.

Studies of the Mechanism of the Δ^5-3-Ketosteroid Isomerase Reaction by Substrate, Solvent, and Combined Kinetic Deuterium Isotope Effects on Wild-Type and Mutant Enzymes†

Liang Xue, Paul Talalay, and Albert S. Mildvan*
Department of Biological Chemistry and Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205
Received February 23, 1990; Revised Manuscript Received May 3, 1990

ABSTRACT: Δ^5-3-Ketosteroid isomerase (EC 5.3.3.1) catalyzes the isomerization of Δ^5-3-ketosteroids to Δ^4-3-ketosteroids by a conservative tautomeric transfer of the 4β-proton to the 6β-position using Tyr-14 as a general acid and Asp-38 as a general base [Kuliopulos, A., Mildvan, A. S., Shortle, D., & Talalay, P. (1989) *Biochemistry 28*, 149]. On deuteration of the 4β-position (97.0%) of the substrate, $k_{cat}(H)/k_{cat}(4\beta\text{-D})$ is 6.1 in H_2O and 6.3 in D_2O. The solvent isotope effect, $k_{cat}(H_2O)/k_{cat}(D_2O)$, is 1.6 for both the 4$\beta$-H and 4$\beta$-D substrates. Mutation of Tyr-55 to Phe lowers k_{cat} 4.3-fold; $k_{cat}(H)/k_{cat}(4\beta\text{-D})$ is 5.3 in H_2O and 5.9 in D_2O, and $k_{cat}(H_2O)/k_{cat}(D_2O)$ with the 4β-H and 4β-D substrates is 1.5 and 1.7, respectively, indicating concerted general acid–base catalysis in either the enolization or the ketonization step of both the wild-type and the Tyr-55 → Phe (Y55F) mutant enzymes. An additional slow step occurs with the Y55F mutant. Smaller isotope effects on K_m are used to estimate individual rate constants in the kinetic schemes of both enzymes. On deuteration of the 4α-position (88.6%) of the substrate, the secondary isotope effect on k_{cat}/K_m corrected for composition is 1.11 ± 0.02 with the wild-type enzyme and 1.12 ± 0.02 with the Y55F mutant. These effects decrease to 1.06 ± 0.01 and 1.07 ± 0.01, respectively, when the 4β-position is also deuterated, thereby establishing these to be kinetic (rather than equilibrium) secondary isotope effects and to involve a proton-tunneling contribution. Deuteration of the 6-position of the substrate (92.0%) produces no kinetic isotope effects on k_{cat}/K_m with either the wild-type (1.00 ± 0.01) or the Y55F mutant (1.01 ± 0.01) enzyme. Since a change in hybridization from sp^3 to sp^2 occurs at C-4 only during enolization of the substrate and a change in hybridization at C-6 from sp^2 to sp^3 occurs only during reketonization of the dienol intermediate, enolization of the substrate constitutes the concerted rate-limiting step. Concerted enolization is consistent with the right angle or antarafacial orientations of Tyr-14 and Asp-38 with respect to the enzyme-bound substrate and with the additive effects on k_{cat} of mutation of these catalytic residues [Kuliopulos, A., Talalay, P., & Mildvan, A. S. (1990) *Biophys. J. 57*, 39a].

The enzyme Δ^5-3-ketosteroid isomerase (EC 5.3.3.1) of *Pseudomonas testosteroni* catalyzes the isomerization of Δ^5-3-ketosteroids to the conjugated Δ^4-3-ketosteroids, accelerating the rate of this process by a factor of $10^{9.5}$ (Kuliopulos et al., 1987) (Figure 1). The enzyme-catalyzed reaction is a stereospecific, intramolecular, cis, diaxial conservative transfer of the 4β-proton to the 6β-position at a rate that approaches the diffusion limit. As previously pointed out (Kuliopulos et al., 1989), the overall reaction is unlikely to be a concerted process on theoretical grounds since concerted, suprafacial, 1,3-sigmatropic shifts are forbidden by orbital symmetry rules (Woodward & Hoffmann, 1970; Alder et al., 1971). Moreover, spectroscopic, isotopic, and kinetic studies have provided evidence for the existence of an enolic intermediate (Wang et al., 1963; Malhotra & Ringold, 1965; Bantia & Pollack, 1986; Eames et al., 1990). The enolization of the substrate would be facilitated by a proton donor that protonates the carbonyl group of the steroid and a proton acceptor that removes the 4β-proton and transfers it to the 6β-position.

NMR docking of the spin-labeled substrate analogue spiro[doxyl-2,3′-5α-androstane]-17′β-ol into the 2.5-Å resolution X-ray structure of isomerase revealed that either Tyr-55 or Tyr-14 might function as the proton donor and that Asp-38 was appropriately positioned to function as the proton acceptor (Kuliopulos et al., 1987b). Cloning and expression of the gene

† This work was supported by NIH Grants DK 28616 (to A.S.M.) and DK 07422 (to P.T.). The interactive Graphics Facility of the Department of Biophysics, The Johns Hopkins University School of Medicine, was established and is maintained by NIH and NSF grants and by a gift from the Richard-King Mellon Foundation.

0006-2960/90/0429-7491$02.50/0 © 1990 American Chemical Society