Conformational Studies of the Smallest Structural Motifs of DNA Detectable via Vibrational Circular Dichroism: Cytidylyl-(3′-5′)-Guanosine and Guanylyl-(3′-5′)-Cytidine

Sheryl S. Birke and Max Diem
Departments of Chemistry and Biochemistry, City University of New York, Hunter College, New York, New York 10021 USA

ABSTRACT The infrared absorption and vibrational circular dichroism (VCD) spectra of buffered aqueous solutions of cytidylyl-(3′-5′)-guanosine (5′(CG)3′) and guanylyl-(3′-5′)-cytidine (5′(GC)3′) are reported. Under low ionic strength conditions, these dinucleotides exhibit VCD features that can be predicted qualitatively from structural data of $(CG)_2$ and $(GC)_2$ sequences of poly(dG-dC) · poly(dG-dC), using the exciton model for infrared VCD intensities.

INTRODUCTION

We report observed and computational infrared (vibrational) CD (VCD) of the dinucleotides 5′(CG)3′ and 5′(GC)3′ in buffered, aqueous solution. The interpretation of the observed VCD spectra, and concentration- and temperature-dependent infrared absorption studies, lead to the conclusion that these dimers exist at low temperature and low ionic strength in a duplex geometry.

VCD is a novel spectroscopic technique that combines the stereochemical and conformational sensitivity of chiroptical techniques such as circular dichroism, with the "fingerprint" sensitivity of vibrational spectroscopy toward molecular structure and conformation. Thus, VCD can reveal subtle aspects of solution conformation of biological molecules (Diem, 1993). In this paper, we shall deal exclusively with circular dichroism in the base carbonyl-stretching vibrations, which occur about 1650 cm^{-1} (6 μm). Our previously published results on model DNA and RNA systems have shown that most of VCD intensity observed in this spectral region is generated by an exciton-type interaction of individual C═O transitions, which are held in a well defined geometric pattern (Gulotta et al., 1989; Zhong et al., 1990). Because the interaction of these dipoles is dependent on the alignment of the dipoles with respect to each other, VCD can be used to monitor the orientation of bases.

We have found excellent agreement between observed VCD intensities and those computed via the exciton approach for model systems such as poly(rC), poly(rG) · poly(rC), poly(dG) · poly(dC), poly(dG-dC) · poly(dG-dC), and others (Zhong et al., 1990; Xiang et al., 1993). In these calculations, canonical structural data were used to define the positions of the carbonyl groups, and empirical values were used for the vibrational frequencies and the transition moments. The agreement between observed and computational results was less satisfactory for small oligonucleotides when canonical structural data were used, indicating that small oligonucleotides have solution conformations that deviate significantly from canonical structures (Birke et al., 1993).

The present study is motivated by our observation, via VCD, of the interaction of an intercalating drug, ethidium bromide (EB) with both poly(dG-dC) · poly(dG-dC) and the tetramer 5′d(CGCG)3′ (Birke et al., 1993b). To interpret the structural changes occurring in the DNA upon drug intercalation, the interaction between EB and the dinucleotides 5′(CG)3′ and 5′(GC)3′ needed to be studied as well. However, the VCD spectra of these simplest nucleic acid fragments, double-stranded, self-complementary 5′(CG)3′ and 5′(GC)3′ have not yet been observed, although the VCD of a small number of dinucleotides had been reported earlier (Annamalai and Keiderling, 1987). The VCD spectra, at 5°C, of these dinucleotides are very different from each other, and may shed light into very subtle details of DNA structure. We found that the simple exciton-type calculations (Tinoco, 1963) referred to above provide an adequate qualitative description of the origin of VCD intensities in nucleic acids.

MATERIALS AND METHODS

The dinucleotides cytidylyl-(3′-5′)-guanosine, also referred to as 5′(CG)3′, and guanylyl-(3′-5′)-cytidine, 5′(GC)3′, were purchased from ICN (Irvine, CA) and Sigma Chemical Co. (St. Louis, MO), respectively, and were used after lyophilization from D_2O to exchange labile protons. They were dissolved in 0.01 M sodium cacodylate/0.1 M NaCl in D_2O to concentrations between 5 and 25 mg/ml.

Temperature-dependent VCD spectra were obtained at 5, 25, and 40°C. Temperature control was attained by circulating thermostated water through the sample holder. The temperatures were determined by placing a thermocouple between the windows of a sample holder filled with water. The reported temperatures are assumed to be accurate within ±3°C.

VCD spectra were collected via an infrared circular dichrograph described elsewhere (Lee and Diem, 1992) at scan speeds of 1 cm^{-1}/s and a spectral bandpass of ~6 cm^{-1}. The VCD spectra shown are coadditions of 30 individual scans each. Infrared absorption spectra were collected simultaneously with the VCD spectra and were verified via FT-IR spectroscopy.

The computed spectra were obtained using the simplest of the exciton-type calculations (the Degenerate Extended Coupled Oscillator, or DECO formalism), which has been described recently (Xiang et al., 1993). Input parameters for these calculations are the carbonyl coordinates from the

Received for publication 1 December 1993 and in final form 22 September 1994.

Address reprint requests to Dr. Max Diem, Department of Chemistry and Biochemistry, CUNY-Hunter College, 695 Park Ave., New York, NY 10021. Tel.: 212-772-5330; 212-772-5332; E-mail: mjnmax@aol.com.

0006-3495/95/03/1045/05 $2.00

appropriate bp sequences $(CG)_2$ and $(GC)_2$ of poly(dG-dC) · poly(dG-dC) and poly(rG-rC) · poly(rG-rC), and the same experimental transition moments and frequencies used before in the computations of polymers. A line width of 15 cm^{-1} was used in the calculations, although this value appears narrow compared with the observed spectra. However, we believe that the observed bandwidth of nearly 60 cm^{-1} is due to conformational variations and the exciton splitting of all coexisting conformers. Coordinates for the canonical polymers were calculated using the program MacroModel (Still, 1989), which uses crystal structural data to determine canonical B-form DNA structures.

The DECO formalism, which is the least accurate of the methods used recently, was used because more sophisticated calculations are not justified at this point due to the very approximate nature of the structural parameters used for the dinucleotides. However, the calculations show qualitative agreement with the experimental data and help define the interactions that give rise to the observed VCD of the polymers.

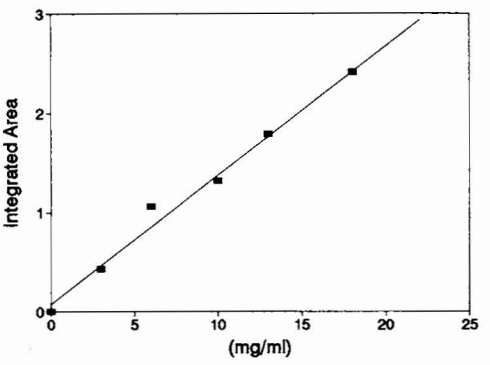

FIGURE 2 Beer-Lambert plot of integrated absorption intensity versus concentration of 5′d(GC)3′.

RESULTS

Fig. 1 shows the observed VCD and infrared absorption spectra of 5′(CG)3′ (*solid heavy lines*) and 5′(GC)3′ (*dashed heavy lines*), at 25 and 15 mg/ml, respectively. Within the concentration range accessible with present instrumentation (5–25 mg/ml), the shape of the VCD spectra for both dinucleotides does not depend on the sample concentration. However, the signal-to-noise ratio of the spectra decreases with decreasing sample concentration to the point where only the most predominant VCD features are observable.

The infrared absorption spectra are very nondescriptive and consist of a very broad, single absorption in both cases with a halfwidth of about 50 cm^{-1}. The absorption spectra are not sensitive to changes in temperature or concentration and exhibit a linear Beer-Lambert plot. Fig. 2 shows this linear behavior of the integrated C=O stretching peak area for 5′(GC)3′, which demonstrates that there is no aggregation of the sample. We conclude that in this concentration range, the dinucleotides are in a duplex form and are not aggregated.

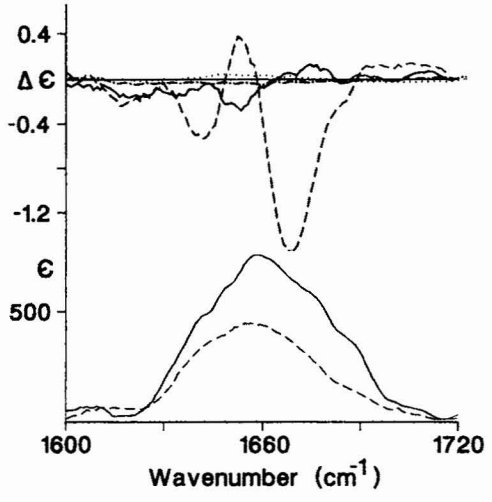

FIGURE 1 (*top*) Observed infrared VCD of 5′(CG)3′ at 25 mg/ml at 5°C (——) and 40°C (- - - -), and 5′(GC)3′ at 15 mg/ml at 5°C (- - - -) and 40°C (· · · · ·) in buffered aqueous (D_2O) solution. (*bottom*) Corresponding infrared absorption spectra of 5′(CG)3′ at 5°C (——) and 5′(GC)3′ at 5°C (- - - -).

The VCD spectra of the two dinucleotides are quite different. 5′(GC)3′ exhibits a strong negative-positive-negative couplet (1637/1658/1672 cm^{-1}), whereas 5′(CG)3′ exhibits a small negative couplet (the negative VCD band at lower wavenumber) at 1678/1655 cm^{-1}.

However, an increase in temperature to 25°C causes a decrease of the VCD amplitudes, and at 40°C there is no longer any VCD observable. This is shown in Fig. 1. We attribute the loss of the VCD signals to be due to the loss of duplex structure of the dinucleotides.

Fig. 3 shows the calculated VCD spectra of $(CG)_2$ (*solid line*) and $(GC)_2$ (*dashed line*) sequences of the canonical B-form poly(dG-dC) · poly(dG-dC). Calculations were also carried out for dimers based on the A-form polymer poly(rG-rC) · poly(rG-rC), vide infra. The overall observed spectral features are reproduced in these calculated data: a negative/positive/negative VCD signal for the $(GC)_2$ sequence, and a negative couplet for the $(CG)_2$ sequence. In addition, there is a shift in the absorption maximum between the two sequences that agrees in direction with the shift observed in the experimental data. The calculated rotational and

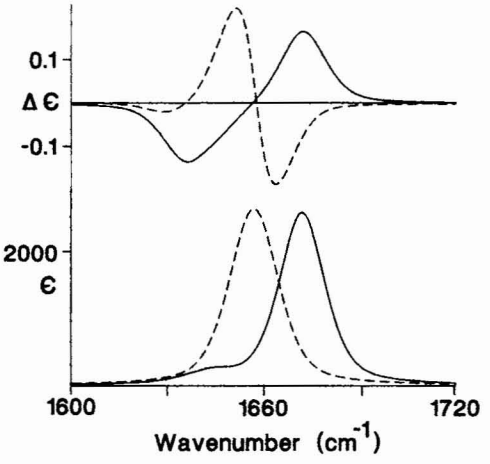

FIGURE 3 Computed infrared VCD (*top*) and absorption (*bottom*) spectra of $d(CG)_2$ (——) and $d(GC)_2$ (- - - -). For details of the computations, see Fig. 4.

dipole strengths for each exciton component, along with the computed frequencies, are summarized in Tables 1 and 2. The former data have been reported previously (Zhong et al., 1990); the results here are qualitatively the same as the original data, but more generally accepted input parameters are used in the present calculations. Also, in the original report the computational results for the double-stranded dinucleotides were mislabeled: what was marked as

C − G
⋮ ⋮
G − C

should have been labeled $(CG)_2$.

DISCUSSION

An inspection of the coordinates of guanosine and cytosine in canonical B-form poly(dG-dC) · poly(dG-dC) reveals that the carbonyl groups in two consecutive bp occur in two distinct geometric patterns. In a single C · · · G bp, the two carbonyl groups are at about 174° with respect to each other; this nearly antiparallel arrangement explains why a single bp (e.g., a 1:1 mixture of guanosine monophosphate and cytosine monophosphate) exhibits no measurable VCD signal in the carbonyl-stretching vibration (Gulotta et al., 1989).

Examining a canonical B-form poly(dG-dC) · poly(dG-dC) sequence from the 5′-end, one finds that in a double-stranded C-G sequence, which we refer to as $(CG)_2$, all four carbonyl groups are within ±5° from being parallel or antiparallel to each other. This is shown in Fig. 4 *a* and has been pointed out previously (Birke et al., 1993). We referred to this case as the "near parallel" geometry. Such an alignment produces good coupling interactions between the dipoles, but minimal induced optical activity. This is manifested by the results summarized in Table 1, which describe a negative couplet with rotational strengths of only $+8/-6 \times 10^{-42}$ [esu · cm]^2 at 1672 and 1635 cm^{-1}, respectively, and a large splitting between the exciton components of 37 cm^{-1}.

The observed data for 5′(CG)3′, shown in Fig. 1 (*solid line*), is very small (near the observation limit) and consists of a negative couplet. This observation is supported by the hypothesis that the geometry of the "near parallel" or "near antiparallel" alignment of the carbonyl groups in the 5′d-(CG)3′ yields low rotational strength and large splitting (cf. Table 1).

A $(GC)_2$ layer, on the other hand, consists of dipoles with a 90° twist between the two layers. This is shown in Fig. 4 *b*. For the perpendicular dipoles, the interaction energy is smaller, resulting in an exciton splitting of only 29 cm^{-1}. However, the rotational strengths are larger ($-31/+36 \times 10^{-42}$ [esu · cm]^2 (cf. Table 1). The observed VCD spectra for the 5′d(GC)3′ (Fig. 1, *dashed line*) are in agreement with the calculated rotational strengths.

The simulated VCD spectra depicted in Fig. 3 do not show the difference in magnitude of the VCD calculated for $(CG)_2$ and $(GC)_2$ as strongly as the observed spectra. This is because the dipolar splitting is less in the "perpendicular" case. Thus, the cancellation of opposite VCD intensity is much more pronounced in $(GC)_2$ because of the smaller splitting of the exciton components. Nevertheless, the calculations predict a negative/positive/negative VCD intensity pattern, which is actually observed for 5′(GC)3′. Again, this observation suggests that in this molecule the carbonyl groups are in a geometry similar to that discussed for $(GC)_2$, namely, the "near

TABLE 1 Calculated absorption and VCD intensity parameters for $d(CG)_2$ and $d(GC)_2$

Frequency (cm^{-1})	Dipole strength (esu^2 cm^2 × 10^{38})	Rotational strength (esu^2 cm^2 × 10^{42})
d(CG(2		
1672.0	82.1	7.7
1647.7	.2	−1.2
1644.8	4.6	−0.95
1635.4	.8	−5.5
$d(GC)_2$		
1659.0	44.8	−30.6
1655.5	36.4	35.8
1654.9	7.9	−3.9
1630.4	.02	−1.3

TABLE 2 Calculated absorption and VCD intensity parameters for $r(CG)_2$ and $r(GC)_2$

Frequency (cm^{-1})	Dipole strength (esu^2 cm^2 × 10^{38})	Rotational strength (esu^2 cm^2 × 10^{42})
$r(CG)_2$		
1684.0	68.1	6.6
1648.7	.7	−3.9
1647.7	17.0	−0.06
1619.7	2.4	−2.7
$r(GC)_2$		
1659.6	26.2	−18.9
1657.6	34.7	34.9
1653.3	26.6	−25.3
1629.4	1.6	9.2

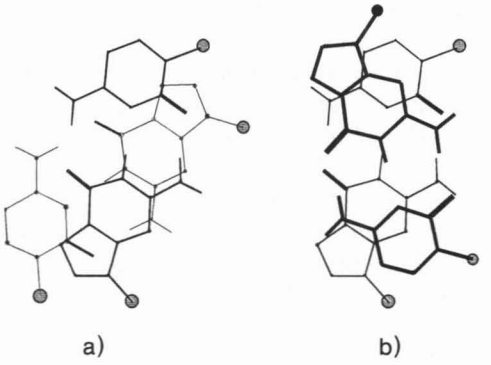

FIGURE 4 Geometries of $d(CG)_2$ (*left*) and $d(GC)_2$ (*right*) sequences of canonical, B-form poly(dG-dC) · poly(dG-dC). The shaded circles at the periphery of the bp indicate the linkage to the ribose sugars. Base pairs drawn with the heaviest lines are closest to the observer. The upper bp in part *a* is the lower bp in *b*. Carbonyl groups are indicated by extra heavy lines.

perpendicular" geometry. Therefore, because the VCD spectrum for the $(CG)_2$ is primarily determined by the "near parallel" alignment and the $(GC)_2$ established by the "near perpendicular alignment, it is not unexpected that the observed and computed VCD spectra are different.

The basic pattern of "near parallel" and "near perpendicular" alignment of dipoles is also found in dimer sequences based on poly(rG-rC) · poly(rG-rC). Consequently, the calculated VCD spectra for dimers based on A-form structures are similar to those for B-form systems. In fact, inspection of Table 2 reveals that the observed negative-positive-negative VCD pattern is reproduced better for the $r(GC)_2$ structure than for $d(GC)_2$. We have found that in the case of the tetramers 5′d(GCGC)3′, 5′d(CGCG)3′, 5′d(CCGG)3′ and 5′d(GGCC)3′, VCD calculations based on A-form coordinates gave slightly better agreement with experimental data than those based on B-form geometries. We interpreted this observation in terms of the more open and unwound structures exhibited by the A-form oligomers (Birke et al., 1993).

The loss of VCD spectral features of both dinucleotides at elevated temperatures demonstrates that these dimers are most likely more open and unwound than the tetramers. In the case of the "near parallel" $(CG)_2$, we find that at 25°C the VCD signal is lost. Therefore, we conclude that $(CG)_2$ exhibits a weak duplex structure due to the unwinding of the helix and end fraying and, consequently, weaker stacking interactions.

The qualitative agreement between observed and calculated data suggests that the geometric pattern of the polymer, namely, the near parallel and near perpendicular alignment of the carbonyl groups, is maintained in the dimers. However, it is expected that significant relaxation of the canonical structure, such as end fraying, base tilting, and unwinding, will occur in small oligonucleotides. We have initiated a program of combined VCD intensity/molecular dynamics (MD) calculations to address this problem. This approach is based on the argument that the enormous width of the observed VCD and infrared absorption spectra is due to the dynamic variations in molecular conformations. These conformational variations are computed at short (1 ps) intervals, and their VCD intensities are coadded. It was found that MD-derived structures reproduce the observed VCD spectra better than those based on canonical poly(dG-dC) · poly(dG-dC) parameters (Kapsis, 1993). These calculations can provide a novel approach to determine average solution conformations of oligonucleotides. Detailed results of these calculations will be reported at a later date.

The more immediate importance of the results reported here is twofold. First, for simple dinucleotides, it appears that observed and calculated spectra agree qualitatively. This result lends further credibility to the simple DECO model, which has been criticized as too simplistic by some theorists. However, in the field of VCD calculations of nucleic acids, in general, this model has performed remarkably well because the interacting dipoles are "weakly coupled" in the terminology of Keiderling (Baur and Keiderling, 1992). This implies that the interaction between the dipoles is mostly a through-space dipolar interaction, whereas in peptides, for example, the corresponding amide C=O stretching modes interact via through-space and through-bond coupling because they are only three bonds apart.

Second, the observed and calculated results here help to define the dominant interactions that determine the VCD of polynucleic acids and explain the differences in the observed VCD spectra of similar species such as 5′d(CGCG)3′ and 5′d(GCGC)3′. In the former, one finds a "near parallel," a "near perpendicular," and another "near parallel" arrangement (Birke et al., 1993). Because the VCD is dominated by short range interactions, and because the "near parallel" arrangements contribute little VCD intensities, the overall VCD of 5′d(CGCG)3′ is mostly due to the center "perpendicular" arrangement. In 5′d(GCGC)3′, on the other hand, two "near perpendicular" patterns are separated by a "parallel" arrangement, and a more complicated VCD pattern is observed (Birke et al., 1993).

It is also interesting to note that in the double-stranded trimer (CGC) · (GCG), the calculated VCD is dominated by the "perpendicular" interaction, and the resulting computed VCD spectra resemble those of poly(dG-dC) · poly(dG-dC) remarkably. Thus, one concludes that the VCD in polynucleic acids is dominated by interactions of bp for which their carbonyl groups are oriented in a right-handed, 90° twist.

CONCLUSIONS

We have shown that VCD of double-stranded dinucleotides sequences are observable in aqueous solutions. The results can be interpreted by the DECO formalism and bp sequences from canonical A- or B-form DNA structures. The qualitative agreement between calculated and observed VCD spectra suggest that the dimers occur in a solution conformation that correspond to $(CG)_2$ and $(GC)_2$ sequences in poly(dG-dC) · poly(dG-dC) or poly(rG-rC) · poly(rG-rC).

The financial support of this research by grant GM 28619 from National Institutes of Health, and several City University of New York Faculty Research Awards are 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.

Baur, P., and T. A. Keiderling. 1992. Computational evaluation of the coupled oscillator model in the VCD of selected small molecules. *J. Am. Chem. Soc.* 114:9100–9105.

Birke, S. S., B. Kagarlovsky, M. Moses, and M. Diem. 1995. Structural changes of DNA upon ethidium bromide intercalation: an infrared vibrational CD study. *J. Am. Chem. Soc.* In press.

Birke, S. S., M. Moses, M. Gulotta, B. Kagarlovsky, D. Jano, and M. Diem. 1993. Infrared CD of deoxy oligonucleotides. Conformational studies of

5′d(GCGC)3′, 5′d(CGCG)3′, 5′d(CCGG)3′ and 5′d(GGCC)3′ in low and high salt aqueous solution. *Biophys. J.* 65:1262–1271.

Diem, M. 1994. Application of infrared CD to the analysis of the solution conformation of biological molecules. *In* Techniques and Instrumentation in Analytical Chemistry. N. Purdie and H. G. Brittain, editors. Elsevier Science Publishers, Amsterdam, The Netherlands. 91–130.

Gulotta, M., D. J. 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.

Kapsis, E. 1993. Molecular dynamics of $d(CGCG)_2$: a conformational analysis using the AMBER 4.0 force field and simulated infrared VCD spectra. Master's thesis. City College of the City University of New York.

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

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

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

Xiang, T., D. J. Goss, and H. Diem. 1993. Strategies for the computation of IR CD and absorption spectra of biological molecules: RNA. *Biophys. J.* 65:1255–1261.

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