# Raman Spectra of Single Crystals of r(GCG)d(CGC) and d(CCCCGGGG) as Models for A DNA, Their Structure Transitions in Aqueous Solution, and Comparison with Double-Helical Poly(dG)·Poly(dC)†

J. M. Benevides,‡ A. H.-J. Wang,§ A. Rich,§ Y. Kyogoku,∥ G. A. van der Marel,⊥ J. H. van Boom,⊥ and G. J. Thomas, Jr.*,‡

*Department of Chemistry, Southeastern Massachusetts University, North Dartmouth, Massachusetts 02747, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan, and Gorlaeus Laboratories, Leiden State University, Leiden, 2300RA The Netherlands*

*Received July 31, 1985*

ABSTRACT: The self-complementary oligonucleotides $[r(CGC)d(CGC)]_2$ and $[d(CCCCGGGG)]_2$ in single-crystal and solution forms have been investigated by Raman spectroscopy. Comparison of the Raman spectra with results of single-crystal X-ray diffraction and with data from polynucleotides permits the identification of a number of Raman frequencies diagnostic of the A-helix structure for GC sequences. The guanine ring frequency characteristic of C3′-endo pucker and anti base orientation is assigned at 668 ± 2 $cm^{-1}$ for both dG and rG residues of the DNA/RNA hybrid $[r(GCG)d(CGC)]_2$. The A-helix backbone of crystalline $[r(GCG)d(CGC)]_2$ is altered slightly in the aqueous structure, consistent with the conversion of at least two residues to the C2′-endo/anti conformation. For crystalline $[d(CCCCGGGG)]_2$, the Raman and X-ray data indicate nucleosides of alternating 2′-endo–3′-endo pucker sandwiched between terminal and penultimate pairs of C3′-endo pucker. The A-A-B-A-B-A-A-A backbone of the crystalline octamer is converted completely to a B-DNA fragment in aqueous solution with Raman markers characteristic of C2′-*endo*/*anti*-G (682 ± 2) and the B backbone (826 ± 2 $cm^{-1}$). In the case of poly(dG)·poly(dC), considerable structural variability is detected. A 4% solution of the duplex is largely A DNA, but a 2% solution is predominately B DNA. On the other hand, an oriented fiber drawn at 75% relative humidity reveals Raman markers characteristic of both A DNA and a modified B DNA, not unlike the $[d(CCCCGGGG)]_2$ crystal. A comparison of Raman and CD spectra of the aqueous $[d(CCCCGGGG)]_2$ and poly(dG)·poly(dC) structures suggests the need for caution in the interpretation of CD data from G clusters in DNA.

Recently, Benevides et al. (1984b) obtained and interpreted laser Raman spectra of the Z-DNA crystals $[d(CGCGCG)]_2$, $[d(m^5CGTAm^5CG)]_2$, and $[d(CGCATGCG)]_2$, the molecular structures of which had been determined by single-crystal X-ray diffraction analysis at atomic resolution. In combining the results of X-ray diffraction and Raman spectroscopy, it was possible to identify a number of Raman lines characteristic of specific nucleoside conformations of Z DNA, including those of both GC and AT base pairs. Since the correlations between Raman frequency and molecular structure apply also to noncrystalline samples, this approach provides a means of identifying nucleoside C3′-endo/syn and C2′-endo/syn conformers in DNA fibers, gels, and solutions by Raman spectroscopy. Raman data from crystalline DNA oligomers and from polydeoxynucleotide fibers and solutions have also been compared with corresponding data from viral nucleic acids and nucleoproteins to identify conformational properties of the native viral genomes (Thomas et al., 1985).

In the present work we report the Raman spectra of two self-complementary oligonucleotides, r(GCG)d(CGC) and d(CCCCGGGG), both of which form miniature double-helical structures. The latter has been investigated by single-crystal X-ray diffraction analysis (Wang et al., 1982a; T. Haran, Z. Shakked, A. H.-J. Wang, and A. Rich, unpublished results), which reveals both C3′-endo/anti and C2′-endo/anti conformers, or minor variants thereof. The double-stranded hybrid hexamer $[r(GCG)d(CGC)]_2$ is similar in sequence to the terminal segments of the larger DNA/RNA hybrid, $[r(GCG)d(TATACGC)]_2$, which forms a double-helical A-type structure containing only C3′-endo/anti nucleoside conformations (Wang et al., 1982b). The A conformation of the decamer is evidently due to the presence of ribonucleotides, since the corresponding deoxy sequence adopts a B-DNA conformation in solution (Mellema et al., 1983). Therefore, we anticipate that the hybrid hexamer $[r(GCG)d(CGC)]_2$ also forms an A-type helix containing C3′-endo/anti conformers.

As far as we know the Raman spectrum of an A-DNA model containing only GC base pairs has not been reported previously. The preferred geometry of a hybrid DNA/RNA structure at physiological conditions is of importance because of the relevance of such structures to transcription, reverse transcription, and DNA replication. The secondary structure of $[r(GCG)d(CGC)]_2$ in aqueous solution is also of interest since the minimum number of hybrid base pairs required to retain an A-helix geometry in water is not known.

The octanucleotide crystal investigated in this work, $[d(CCCCGGGG)]_2$, diffracts to near atomic resolution and is isomorphous with the related crystal, $[d(GGCCGGCC)]_2$, in which all nucleosides except the third and fifth from the 5′-terminus, i.e., C(3) and G(5), assume the A-type geometry

† This research was supported by Grant AI18758 from the National Institutes of Health and Grant INT83-11889 from the National Science Foundation, U.S.–Japan Cooperative Science Program. This is paper 30 in the series Raman Spectral Studies of Nucleic Acids. Paper 29 in this series is Thomas & Benevides (1985).

* Address correspondence to this author.
‡ Southeastern Massachusetts University.
§ Massachusetts Institute of Technology.
∥ Osaka University.
⊥ Leiden State University.

0006-2960/86/0425-0041$01.50/0 

(Wang et al., 1982a; T. Haran, Z. Shakked, A. H.-J. Wang, and A. Rich, unpublished results). The central four base pairs in both of these structures share the same sequence (CCGG), and conformational similarities are therefore anticipated. In the central quartet, the sugar ring pucker alternates between C3′-endo and C2′-endo geometries along one strand (5′ → 3′) and conversely for the complementary strand. The fate of this structure in solution is relevant to a number of recent studies, which suggest that an alternation in ring pucker along one DNA strand (or between two strands) may be a rather common feature of aqueous nucleic acid structure (Arnott et al., 1983; Fodor et al., 1985). The $C_4G_4$ sequence may also serve as an appropriate model for ascertaining the effect of alternating ring pucker upon Raman lines, which serve as markers for A-DNA and B-DNA helices.

Experimental evidence indicates that the transition between B and A forms of DNA depends upon the GC content, with the A structure preferentially stabilized for GC sequences (Drew & Dickerson, 1981; Inui-Nara et al., 1985). In fact, apparently stable A-helix structures have recently been reported for aqueous poly(dG)·poly(dC) (Nishimura et al., 1985) and low-humidity fibers of poly(dA-dT)·poly(dA-dT) (Thomas & Benevides, 1985). We have examined the Raman spectra of fibers and solutions of poly(dG)·poly(dC), as well as both Raman and circular dichroism (CD) spectra of aqueous solutions of $C_4G_4$, over a wide range of solute concentration and compared the spectral data with those obtained from the crystalline GC oligomers. The results indicate conformational similarities between GC sequences of fibrous poly(dG)·poly(dC) and those of the $C_4G_4$ crystal but reveal that the conformations of aqueous GC models are significantly influenced by solvent and salt environment.

## Experimental Procedures

*Materials.* (*A*) *Synthesis and Growth of Single Crystals of Self-Complementary Oligonucleotides.* The oligonucleotides r(GCG)d(CGC) and d(CCCCGGGG) were synthesized by an improved phosphotriester technique in which 1-hydroxybenzotriazol was used as an activating reagent (van der Marel et al., 1981). After the deblocking of the fully protected molecules, both fragments were purified by Sephadex G-50 column chromatography and converted into ammonium salts. The purity of each oligomer was greater than 95%, as judged by HPLC analysis. The crystallization of these molecules has been described in detail elsewhere (Wang et al., 1982a,b; T. Haran, Z. Shakked, A. H.-J. Wang, and A. Rich, unpublished results).

(*B*) *Polynucleotides.* Poly(dG) (lot 427840), poly(dC) (lot 663/82), and poly(dG)·poly(dC) of homogeneous size (lot 709-57, CsCl buoyant density = 1.793 g/cm$^3$) were purchased from Pharmacia P-L Biochemicals. Solutions were prepared either by directly dissolving the polynucleotide duplex at the concentrations specified below in 0.1 M NaCl at pH 7.5 or by first dissociating the duplex at pH 13 and then reannealing it through progressive dialysis against media of gradually decreasing pH down to 7.5. Both methods produced samples of apparently identical secondary structures as judged from Raman results. Fibers of poly(dG)·poly(dC) were drawn as described previously (Prescott et al., 1984) and were sealed in hygrostatic Raman cells maintained at the desired relative humidity (RH).

*Spectroscopy.* (*A*) *Raman Spectroscopy.* Raman spectra of solutions, fibers, and crystals were obtained by methods previously described in detail (Prescott et al., 1984; Benevides et al., 1984b; Thomas & Benevides, 1985) and summarized below.

The octamer or hexamer crystal, authenticated by X-ray diffraction as referenced in the preceding section, was transferred with approximately 10 μL of mother liquor [2-methyl-2,4-pentanediol (40%) + 25 mM sodium cacodylate (pH 7) + 2 mM spermine + 10 mM $MgCl_2$) to a Raman sample cell (Kimax 34507 glass capillary), which was thermostated at 32 °C. The 514.5-nm line from a Coherent Model CR-2 argon laser was focused on the crystal, and the Raman scattering at 90° was collected and analyzed by a Spex Ramalog spectrometer under the control of a North Star Horizon II microcomputer. Spectral data were collected at 1.0-cm$^{-1}$ intervals with an integration time of 1.5 s. A slit width of 8.0 cm$^{-1}$ was employed for each scan of the spectrum from 300 to 1800 cm$^{-1}$. Additional data from some samples were collected at 4.0- and 2.0-cm$^{-1}$ slit widths to search for band splitting as described under Results and Discussion. Spectra of solutions of the oligomers and polymers were obtained similarly. Generally, solution spectra were recorded at or below 32 °C.

Each Raman spectrum displayed in the figures is the average of several scans, of 1.5-cm$^{-1}$ or less repeatability, from which the fluorescent background and scattering by the solvent or mother liquor have been removed by computer-subtraction techniques as described (Benevides et al., 1984b). The spectrum of solvent or mother liquor was always recorded with the same instrument setting employed for solution or crystal. In some of the spectra, the noise was smoothed by a least-squares fit of third-order polynomials to overlapping 15-point regions. This procedure did not measurably alter the Raman line frequencies or intensities. Raman frequencies are believed accurate to within ±2 cm$^{-1}$. The sharp line at 764 cm$^{-1}$, due to mother liquor and measurable to a precision of ±1 cm$^{-1}$ (Benevides et al., 1984b), was employed as an internal frequency standard in spectra of crystals.

(*B*) *Deconvolution of Raman Spectra.* Deconvolution methods for enhancement of resolution in Raman spectra of nucleic acids have been described in detail elsewhere (Thomas & Agard, 1984). Here, deconvolution has been employed only to separate or sharpen partially overlapped bands that cannot be resolved instrumentally. Generally, a Gaussian function of 15-cm$^{-1}$ half-width was employed to deconvolute the spectral envelope in the 600–900-cm$^{-1}$ interval. This desmearing function corresponds closely to the instrument slit function.

(*C*) *Circular Dichroism Spectroscopy.* CD spectra were recorded on a JASCO Model J-500A recording spectropolarimeter, equipped with a Model DP501N data processor (Institute for Protein Research, Osaka University). Melting temperatures of the aqueous oligomers were determined by standard UV methods and yielded the following results. In 0.1 M NaCl, pH 7, the $T_m$ of [r(GCG)d(CGC)]$_2$ was 41 °C, and that of [d(CCCCGGGG)]$_2$ was 47 °C. Both melting transitions were sharp in the sense that 80% of the total UV intensity change over the temperature range 20–65 °C took place within the interval $T_m$ ± 5 °C. Less than 10% of the increase in $A_{260}$ occurred between 20 and 32 °C. These results were confirmed by the Raman spectra, which showed no significant change over the temperature range 20–32 °C. At the latter temperature it is estimated that less than 10% of the helical structure of either oligomer has been melted.

## Results and Discussion

*Molecular Structure and Raman Spectrum of* [*r(GCG)d(CGC)*]$_2$. The molecular hybrid [r(GCG)d(CGC)]$_2$ is expected to form an A-helix structure (Wang et al., 1982b). Figure 1 shows the Raman spectrum of the single crystal of [r(GCG)d(CGC)]$_2$. The close correspondence between the

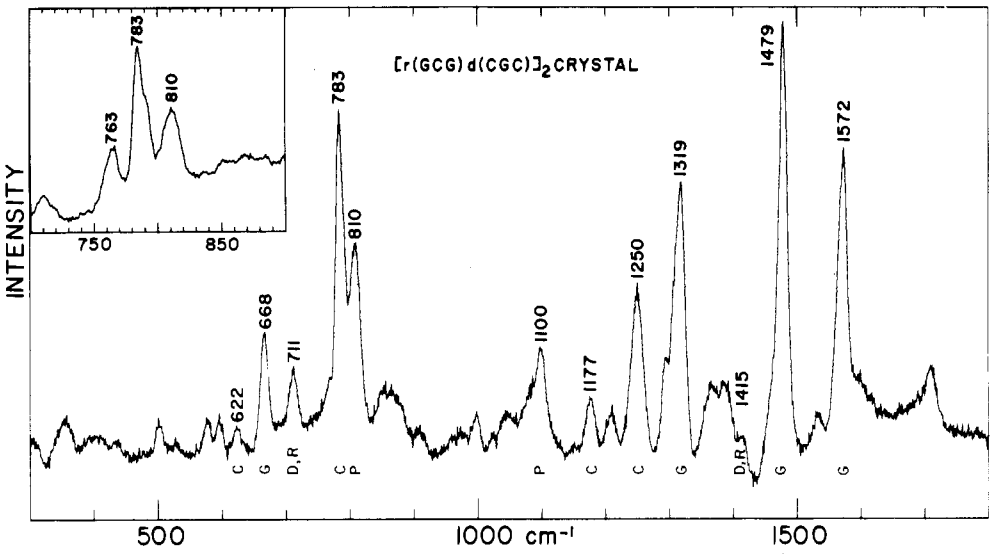



FIGURE 1: Raman spectrum in the region 300–1800 $cm^{-1}$ of the $[r(GCG)d(CGC)]_2$ crystal suspended in mother liquor. The spectrum has been corrected for scattering by the mother liquor as described previously (Benevides et al., 1984b). Raman frequencies of prominent bands are listed in $cm^{-1}$, and assignments are indicated as follows: guanine (G), cytosine (C), phosphate (P), ribose phosphate (R), and deoxyribose phosphate (D). The same notation is used in subsequent figures. The inset shows the region near 800 $cm^{-1}$, recorded at high spectral resolution (2-$cm^{-1}$ slit width), to demonstrate the appearance of a single line at 810 $cm^{-1}$.

spectrum of Figure 1 and Raman spectra of aqueous A-RNA (Lafleur et al., 1972; Thomas & Hartman, 1973) and fibrous A-DNA structures (Erfurth et al., 1975; Prescott et al., 1984) is remarkable and confirms the A-helix structure for the hybrid hexamer in the crystalline state. This result also confirms the transferability of Raman lines characteristic of the A backbone geometry among crystalline, fibrous, and aqueous nucleic acids. The Raman lines or markers that are most useful for distinguishing A-helix geometry of the DNA or RNA backbone are labeled in Figure 1 and several of these are tabulated in Table I. Further discussion of the Raman assignments is given under the next section.

Comparison of Figure 1 with spectra of GC sequences in Z- and B-helical forms shows also that the $[r(GCG)d(CGC)]_2$ crystal lacks Raman lines assigned previously to Z and B backbone geometries, which confirms the earlier Raman marker assignments (Benevides & Thomas, 1983; Benevides et al., 1984b).

*Raman Lines Characteristic of the A-Helix Geometry. (A) The Conformation-Sensitive Guanosine Ring Mode Occurs at 668 $cm^{-1}$.* One of the most useful indicators of nucleoside conformation is a Raman line in the region 620–690 $cm^{-1}$, which arises not from a vibration within the backbone itself but from a vibration of the guanosine ring that is mechanically coupled through the glycosidic bond to the furanose ring (Lord & Thomas, 1967; Thamann et al., 1981; Nishimura et al., 1983; Benevides et al., 1984b). The guanine frequency is therefore influenced by both sugar ring pucker and glycosyl torsion angle. As discussed by Nishimura et al. (1983), C3′-*endo*/*anti*-guanosines should produce the ring frequency at 667 ± 3 $cm^{-1}$. As seen in Figure 1, this frequency occurs as a single sharp band at 668 ± 2 $cm^{-1}$, which defines its position and intensity for the C3′-*endo*/*anti*-G conformation. Since a single, sharp line is observed, both dG and rG nucleosides are confirmed to yield the same C3′-endo/anti marker.

*(B) Backbone Vibrations Correspond Closely to Those of Aqueous RNA and A-DNA Fibers.* The weak Raman line at 711 $cm^{-1}$ is assigned to a main-chain vibration characteristic of C3′-*endo*-ribosyl structures. This line is the analogue of the 710–715-$cm^{-1}$ line observed previously in polyribonucleotides and oligoribonucleotides and assigned to a vibration of the ribose ring involving C–C and C–O stretching motions (Prescott et al., 1974).

Table I: Conformation-Sensitive Raman Frequencies and Assignments from Spectra of $[r(GCG)d(CGC)]_2$ and $[d(CCCCGGGG)]_2$[a]

| $[r(GCG)d(CGC)]_2$ | | $[d\text{-}(CCCCGGGG)]_2$ | | |
|---|---|---|---|---|
| crystal | solution | crystal | solution | assignment[b] |
| 668 | 668 | 663 | | 3′-*endo*/*anti*-G |
| | | 686[c] | 682 | 2′-*endo*/*anti*-G |
| 690 | 690 | *d* | | A helix |
| 711 | 711 | 705 | | A helix |
| | | 792[e] | 784[e] | B helix OPO |
| 810 | 810 | 807 | | A helix OPO |
| | | | 826 | B helix OPO |
| | | 850[c] | | B helix OPO |
| | | | 1092 | B helix $PO_2^-$ |
| 1100 | 1096 | 1099 | | A helix $PO_2^-$ |
| | | 1321[c] | 1333 | 2′-*endo*/*anti*-G |
| 1415 | 1415 | 1415 | | 3′-endo sugar[f] |
| | | | 1420 | 2′-endo sugar[f] |

[a] Frequencies ($cm^{-1}$) are accurate to ±2 $cm^{-1}$. [b] Abbreviations: G, guanosine or deoxyguanosine; OPO, stretching vibration of backbone phosphodiester group; $PO_2^-$, stretching vibration of backbone phosphodioxy group. Other nomenclature is defined in the text. [c] Assigned to C2′-endo/anti or C1′-exo/anti conformer. See text. [d] The region near 690 $cm^{-1}$ is obscured by other bands. [e] Overlapped by a band due to the cytosine ring ca. 782 $cm^{-1}$. [f] See Prescott et al. (1984).

The conformation-sensitive diester OPO stretching vibration gives rise to an intense and sharp line at 810 $cm^{-1}$, while the $PO_2^-$ symmetric stretch occurs at 1100 $cm^{-1}$. These are the frequencies anticipated for a double helix of the A backbone geometry. Further, the observed ratio of Raman intensities ($I_{810}/I_{1100}$ = 1.6 ± 0.1) corresponds to that expected of a pure A helix (Thomas & Hartman, 1973; Benevides & Thomas, 1983).

We note that in RNA fibers the OPO frequency is usually centered near 813 ± 2 $cm^{-1}$ (Thomas & Hartman, 1973), slightly higher than in A-DNA fibers (807 ± 2 $cm^{-1}$) (Prescott et al., 1984). The fact that we observe a single OPO band for the hybrid $[r(GCG)d(CGC)]_2$ at the position intermediate between markers for A RNA and A DNA rather than two bands of equal intensity at the respective positions for A RNA and A DNA, is indicative of a homogeneous backbone geom-

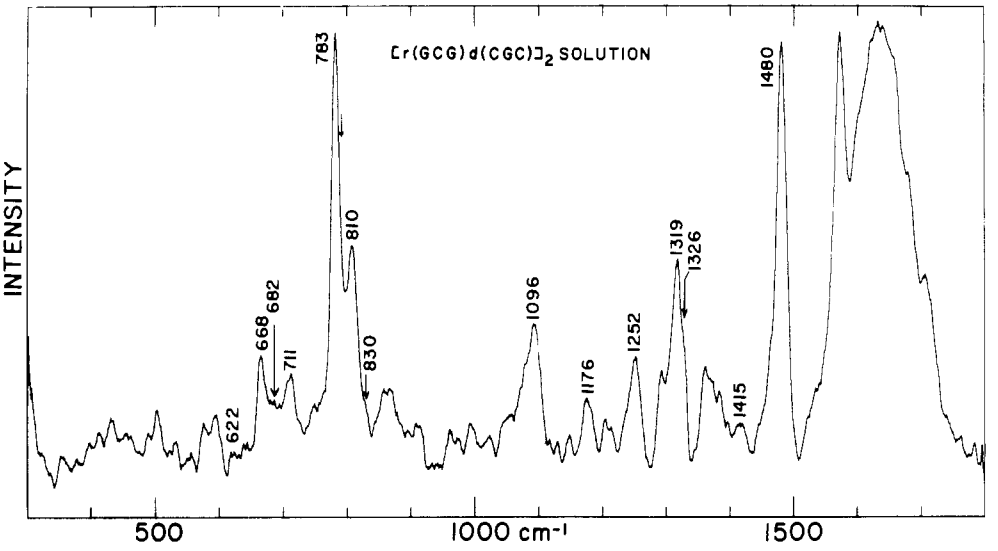



FIGURE 2: Raman spectrum of $[r(GCG)d(CGC)]_2$ dissolved to 20 mg/mL in 0.1 M NaCl solution at pH 7.5. The spectrum has been corrected for scattering by the solvent. Arrows indicate Raman lines associated with nucleoside or backbone conformations of B DNA.

etry in the hybrid, which is distinct from structures of homoribo and homodeoxyribo A helices. It is important to recognize from Figure 1 that the 810-$cm^{-1}$ line displays a half-width as narrow as that of any other line in the Raman spectrum of $[r(GCG)d(CGC)]_2$, indicating that the band is not the composite of two overlapping bands at nominally higher and lower frequencies. We have verified that the 810-$cm^{-1}$ line cannot be resolved into a doublet by recording spectra with an instrumental slit width as low as 2 $cm^{-1}$, as shown in the inset of Figure 1. The experimental finding is also confirmed by Fourier deconvolution of the spectrum, which yields a single deconvoluted peak at 810 $cm^{-1}$ rather than two peaks at higher and lower frequencies (not shown in Figure 1, but see Figure 7). Ribosyl (5′-$CH_2$) and deoxyribosyl (5′-$CH_2$ and 2′-$CH_2$) methylene deformations give rise to a weak, broad band at 1415 $cm^{-1}$, as expected for sugars of the A-helix backbone (Benevides & Thomas, 1983; Prescott et al., 1984).

*(C) The A-Helix Structure of Crystalline [r(GCG)d(CGC)]$_2$ Is Predominantly Unchanged in Aqueous Solution.* The Raman spectrum of aqueous $[r(GCG)d(CGC)]_2$ (Figure 2) indicates that the A-helix structure of the crystal is largely conserved in aqueous solution. However, several Raman lines characteristic of B DNA appear very weakly in the solution spectrum, indicating a slightly less uniform A helix than in the crystal. A quantitative estimate of the loss of A character may be obtained by comparing the value of the Raman intensity quotient ($I_{810}/I_{1100}$) in crystal and solution spectra (Thomas & Hartman, 1973). This quotient falls from 1.6 ± 0.1 in the crystal (Figure 1) to 1.4 ± 0.1 in solution (Figure 2), which corresponds closely to a decrease of 2 of the 12 bases per hexamer duplex.

The partial A → B isomerization, evidenced by the approximate 16% intensity decrease of the 810-$cm^{-1}$ line, is consistent with an "end effect" in which the terminal rG·dC base pairs fall out of the strict A helix register of the crystal and in which rG retains C3′-endo pucker while dC is converted to C2′-endo pucker. Fraying of the double helix would also account for the very weak B backbone marker in the solution spectrum (830-$cm^{-1}$ shoulder to the 810-$cm^{-1}$ line, Figure 1) and would be consistent with the interpretation of NMR results on related structures (Mellema et al., 1983; Sanderson et al., 1983). It is interesting to note that the sharpness of the 668-$cm^{-1}$ line in the crystal (Figure 1) is also attenuated in the solution spectrum (Figure 2), which is consistent with a loss of regularity of the C3′-endo/anti geometry of G residues. This could be explained by partial fraying of the penultimate base pairs, which would result in a reduction of the number of C3′-*endo*-dG residues.

While the conversion of the sugar conformation of the terminal dC residues from C3′-endo to C2′-endo is consistent with both Raman and NMR spectra of aqueous hybrids, we note that the Raman data are also consistent with a conversion of the two dG residues to C2′-endo/anti conformation. This would also explain the intensity loss at 668 $cm^{-1}$, noted above, and apparent intensity gain at 682 $cm^{-1}$ accompanying the transfer of $[r(GCG)d(CGC)]_2$ from crystal to solution (observed as a filling in of the otherwise deep minimum between the 668- and 711-$cm^{-1}$ peaks; cf. Figures 1 and 2). However, such an alternating structure has not been observed previously.

The probable nucleoside conformations for crystal and solution forms of $[r(GCG)d(CGC)]_2$ may be summarized as

crystal

| | | | | |
|---|---|---|---|---|
| 5′ | 3′-endo/anti | rG-dC | 3′-endo/anti | 3′ |
| ↓ | 3′-endo/anti | rC-dG | 3′-endo/anti | ↑ |
| | 3′-endo/anti | rG-dC | 3′-endo/anti | |
| | 3′-endo/anti | dC-rG | 3′-endo/anti | |
| | 3′-endo/anti | dG-rC | 3′-endo/anti | |
| 3′ | 3′-endo/anti | dC-rG | 3′-endo/anti | 5′ |

solution

| | | | | |
|---|---|---|---|---|
| 5′ | 3′-endo/anti | rG-dC | 2′-endo/anti | 3′ |
| ↓ | 3′-endo/anti | rC-dG | 3′-endo/anti | ↑ |
| | 3′-endo/anti | rG-dC | 3′-endo/anti | |
| | 3′-endo/anti | dC-rG | 3′-endo/anti | |
| | 3′-endo/anti | dG-rC | 3′-endo/anti | |
| 3′ | 2′-endo/anti | dC-rG | 3′-endo/anti | 5′ |

Molecular models for $[r(GCG)d(CGC)]_2$ proposed on the basis of the Raman results are shown in Figure 3.

*Raman Spectra and Nucleoside Conformations of [d(CCCCGGGG)]$_2$. (A) [d(CCCCGGGG)]$_2$ Assumes the B-DNA Conformation in Aqueous Solution.* The Raman spectrum of aqueous $[d(CCCCGGGG)]_2$ is compared with the spectrum of an authenticated B-DNA model, low-salt poly(dG-dG)·poly(dG-dC), in Figure 4. The latter spectrum has been discussed and interpreted in detail previously (Benevides & Thomas, 1983). Figure 4 shows that only the B-DNA backbone and only C2′-endo/anti conformers of deoxyguanosine are present in $[d(CCCCGGGG)]_2$. The sharp and prominent bands near 682 (G), 785 (C), and 826 $cm^{-1}$ (OPO) in both spectra confirm that GC pairs, whether alternating

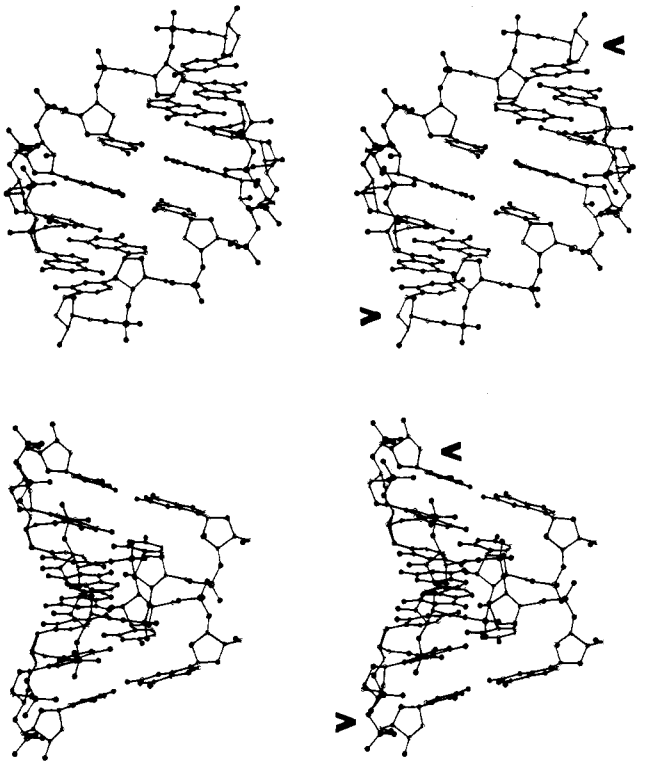


FIGURE 3: Stereo drawing of the model of the double-helical structure of $[r(GCG)d(CGC)]_2$ as determined from the Raman spectrum of the crystal. The model was generated by assuming C3′-endo/anti nucleoside conformations and coordinates of A DNA. Arrows indicate the positions of dC residues of the 3′-terminus, which are believed to assume the C2′-endo/anti conformation in aqueous solution. (Top) View along the 2-fold axis approximately perpendicular to the helix axis; (bottom) view perpendicular to both the 2-fold and helix axes.

or nonalternating, produce the same marker bands in the 600–900-$cm^{-1}$ interval of the Raman spectrum. [The lower intensity of the 826-$cm^{-1}$ line of the octamer in comparison to the corresponding (829 $cm^{-1}$) intensity of the polymer is attributable to end effects in the octamer.]

On the other hand, the 1200–1400-$cm^{-1}$ intervals of the Raman spectra of Figure 4 are clearly different for $[d(CCCCGGGG)]_2$ and poly(dG-dC)·poly(dG-dC). The major difference between the two spectra occurs for the cytosine line near 1240 $cm^{-1}$. In poly(dG-dC)·poly(dG-dC), the cytosine line exhibits the same relative intensity as the corresponding line in calf thymus DNA (Prescott et al., 1984) and other B-DNA polymers (Brown et al., 1972; Thomas et al., 1985). By comparison, a more intense 1240-$cm^{-1}$ line occurs in $[d(CCCCGGGG)]_2$. These results suggest strongly that the higher intensity of the 1240-$cm^{-1}$ line of $[d(CCCCGGGG)]_2$ is due to less hypochromism associated with stacking of cytosine in the $C_4$ sequences.

The inset to Figure 4 shows the CD spectrum of $[d(CCCCGGGG)]_2$ in more dilute aqueous solution (0.1 mg/mL). This CD profile exhibits many similarities to that of B DNA and confirms the conclusion reached from Raman spectra. The $[d(CCCCGGGG)]_2$ solution that yielded the CD spectrum of Figure 4 contained approximately the same ratio of counterion to phosphate as the more concentrated oligonucleotide solutions prepared for Raman spectroscopy. We found, however, that significant changes in salt (and/or buffer) concentration produce substantially altered CD spectra, which could not be easily interpreted in terms of known CD patterns.

*(B) The Molecular Structure of Crystalline [d(CCCCGGGG)]$_2$ Is Not B DNA.* Figure 5 shows the Raman spectrum obtained from the single crystal of $[d(CCCCGGGG)]_2$, which is isomorphous with $[d(GGCCGGCC)]_2$ (Wang et al., 1982a). Comparison with Figure 4 demonstrates that the octamer duplex in the crystal is structurally different from the solution form. Specifically, the secondary structure of the octamer backbone in crystalline $[d(CCCCGGGG)]_2$ is unlike that of B DNA, and the average deoxyguanosine conformation differs slightly from the uniform C2′-endo/anti pattern of the aqueous octamer and the low-salt form of poly(dG-dC)·poly(dG-dC). This is in accord with recent X-ray results indicating that an unusual conformation occurs at the CpG step of the $C_4G_4$ sequence (T. Haran, Z. Shakked, A. H.-J. Wang, and A. Rich, unpublished results).

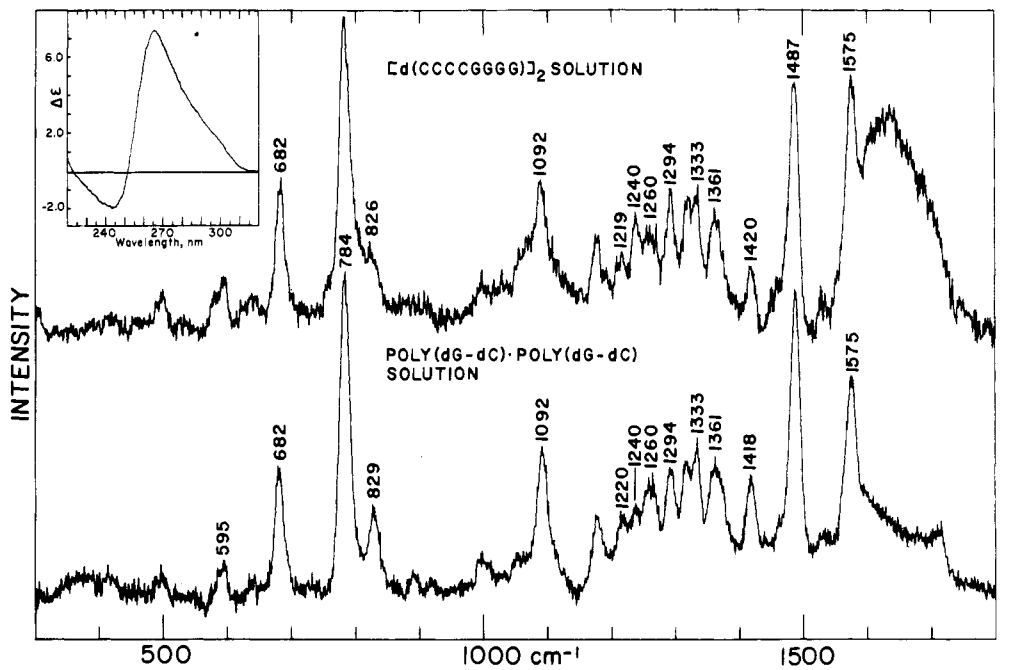



FIGURE 4: (Top) Raman spectrum of $[d(CCCCGGGG)]_2$ dissolved to 20 mg/mL in 0.1 M NaCl solution at pH 7.5. (Bottom) Raman spectrum of poly(dG-dG)·poly(dG-dC) dissolved to 20 mg/mL in $H_2O$ solution at pH 7.5 (no added salt). The inset shows the CD spectrum of aqueous $[d(CCCCGGGG)]_2$.

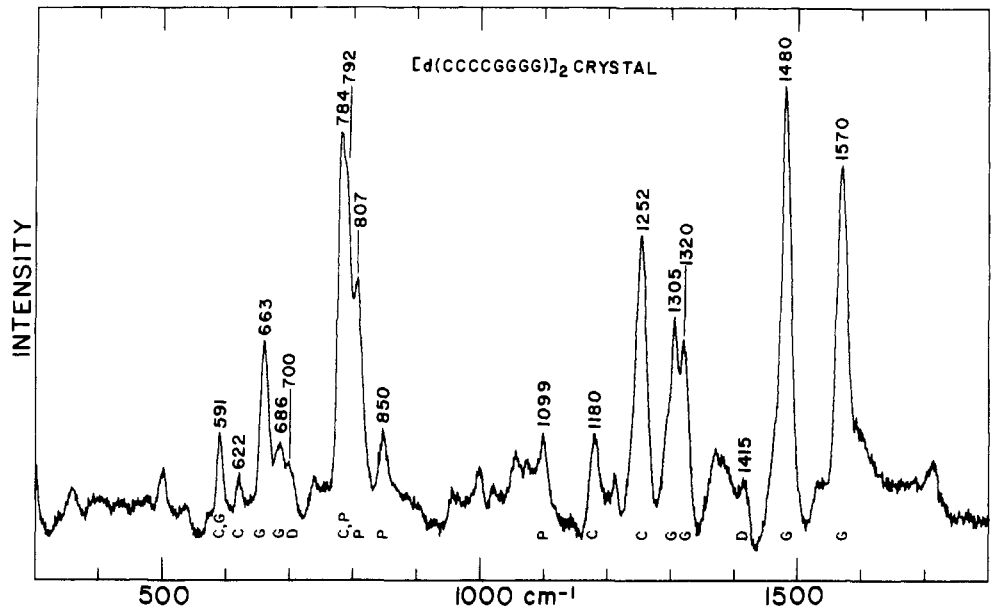



FIGURE 5: Raman spectrum of the $[d(CCCCGGGG)]_2$ crystal suspended in mother liquor (see Figure 1 legend).

We rely upon analogy with the $[d(GGCCGGCC)]_2$ crystal structure, and Raman correlations established previously (Benevides et al., 1984b; Prescott et al., 1984; Benevides & Thomas, 1983), to reach certain conclusions about the molecular structure of crystalline $[d(CCCCGGGG)]_2$.

First, the spectrum of the crystal shows a richness of Raman lines in the 600–900-$cm^{-1}$ interval quite different from a pure B-DNA conformation [cf. Figure 5 with Figure 4 and data of Benevides & Thomas (1983)], indicating a greater diversity of secondary structure than is present in B DNA. Second, the guanosine conformation marker occurs as a doublet (approximate relative peak heights in parentheses): a major peak at 663 $cm^{-1}$ (0.70 ± 0.05) and a minor peak at 686 $cm^{-1}$ (0.30 ± 0.05). This indicates that in the crystal structure roughly two-thirds to three-fourths of the dG residues may adopt the C3′-endo/anti conformation but roughly one-fourth to one-third of the dG residues adopt a conformation that differs from C3′-endo/anti. From previously established correlations, the 686-$cm^{-1}$ line is most likely due to a C2′-*endo*/*anti*-dG conformer, or a minor variant thereof (Benevides et al., 1984b). Note that the additional shoulder near 700–705 $cm^{-1}$ is too high in frequency to be assigned to guanine (Nishimura et al., 1983) but exhibits the position and intensity of the sugar ring mode discussed previously as an A-helix marker (Prescott et al., 1974) (see also Figure 1).

Third, we find evidence for at least two different backbone OPO conformation markers near 800 $cm^{-1}$: 790 (B backbone) and 807 (A backbone). The former occurs as a partially resolved shoulder to the more intense cytosine ring mode at 784 $cm^{-1}$; the latter is reasonably intense. Fourth, the weak Raman line near 850 $cm^{-1}$ (superposed over a broader background) suggests an unusual OPO diester stretch for a small portion of the octamer backbone. We presume the 850-$cm^{-1}$ line to be indicative of backbone geometry within the B family because of its proximity to other B-DNA markers. The diffuse Raman scattering that underlies the 850-$cm^{-1}$ peak is consistent with the occurrence of other B-type secondary structure.

Collectively, these results indicate an approximate 3 to 1 ratio of C3′-endo to C2′-endo (or closely related) G conformers in the octamer crystal, as well as the presence of OPO linkages associated with both A- and B-form DNA backbones. The pucker of the minor G conformer cannot be deduced unequivocally from the Raman data but is reasonably assigned as C2′-endo/anti. Typically, the C2′-*endo*-G marker is expected at 682 ± 2 $cm^{-1}$; here, it occurs at 686 ± 2 $cm^{-1}$. From X-ray studies it is known that both C2′-endo and the closely related C1′-exo conformers can be incorporated into a B-type backbone and both are quite stable in GC sequences of this kind (Wang et al., 1982a). Both would be consistent with the Raman backbone markers noted above. Therefore, we do not rule out the possibility of C1′-exo or another minor variant of the dG sugar ring pucker as the source of the 686-$cm^{-1}$ line. However, for simplicity, in the subsequent discussion we shall refer to the minor dG conformer of crystalline $[d(CCCCGGGG)]_2$ as C2′-*endo*/*anti*-dG.

With reference to the crystal structure of $[d(GGCCGGCC)]_2$ (Wang et al., 1982a), residues G(1), G(2), C(7), and C(8) have sugar conformations close to the C3′-endo pucker of A DNA. The two base pairs on each end therefore comprise miniature A-DNA helices. The sugar puckers of the central four nucleosides, C(3) through G(6), alternate in B-A-B-A fashion. Since the octamers $[d(CCCCGGGG)]_2$ and $[d(GGCCGGCC)]_2$ have the same CCGG sequence in their respective central four bases and since their crystal structures are isomorphous, we anticipate that the molecular structures of both crystals may be very similar to one another. Such an alternation of ring puckers along the central four residues of $[d(CCCCGGGG)]_2$, sandwiched between two miniature A helices, would explain all of the above-noted qualitative features of the Raman spectrum. Furthermore, an alternation in ring pucker for $[d(CCCCGGGG)]_2$, as occurs for $[d(GGCCGGCC)]_2$, would generate guanine ring modes near 662 and 682 $cm^{-1}$ with an approximate 3:1 Raman intensity ratio, in close quantitative agreement with the data of Figure 5. Finally, this interpretation would also be consistent with the occurrence of an irregularity at the CpG step of the $C_4G_4$ sequence (T. Haran, Z. Shakked, A. H.-J. Wang, and A. Rich, unpublished results).

In Figure 6 we show the deconvolution of the $C_4G_4$ crystal spectrum in the region 600–900 $cm^{-1}$. The deconvolution permits better separation of the overlapped Raman lines and more reliable measurement of the relative Raman intensities (band areas). The results show that the ratio of integrated intensities of the 663- and 686-$cm^{-1}$ lines (2.70:1) compares favorably with the expected 3:1 ratio for an A-A-B-A-B-A-A-A geometry like that of $[d(GGCCGGCC)]_2$.

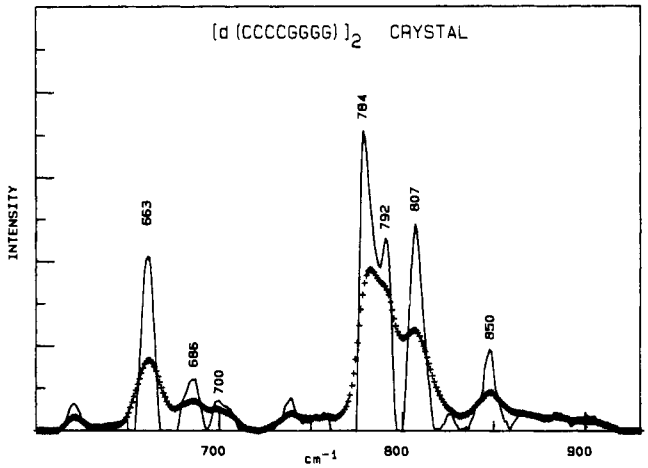
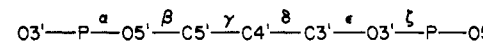



FIGURE 6: Comparison of the observed (+) and deconvoluted (—) spectra of the $[d(CCCCGGGG)]_2$ crystal in the region 600–900 $cm^{-1}$. A Gaussian desmearing function of 15-$cm^{-1}$ half-width was employed in the Fourier deconvolution method of Thomas & Agard (1984).

Table II: Conformation-Sensitive Guanine Ring Frequencies and C3′–C4′ Torsion Angles in Model Structures

| structure | residue | C3′–C4′ torsion angle ($\delta$)[a] | Raman frequency ($cm^{-1}$) |
|---|---|---|---|
| A DNA | G | 83 | 668 |
| d(CCCCGGGG) | G(6) | (85) | 663 |
| B DNA | G | 139 | 682 |
| d(CCCCGGGG) | G(5) | (116) | 686 |

[a] Data from Wang et al. (1984). Values in parentheses are taken from the isomorphous crystal $[d(GGCCGGCC)]_2$. Values for DNA are from model-building analyses. Torsion angles of the backbone are defined as




Finally, we note that C1′-*exo*/*anti*-G remains a plausible candidate for the minor conformer of $[d(CCCCGGGG)]_2$. This would not only account for the somewhat anomalous positions of the guanine marker (686) and OPO marker (850 $cm^{-1}$) but would also be completely consistent with the structure determined for the isomer $[d(GGCCGGCC)]_2$. In any case it seems clear from the Raman data that the proposed B-like base pairs in the $[d(CCCCGGGG)]_2$ crystal, viz., C(3)·G(5) and G(5)·C(3), deviate in a minor way from the so-called "average" B geometry found in the alternating GC sequence of the low-salt form of poly(dG-dC)·poly(dG-dC) (Benevides & Thomas, 1983).

*Comparison of the Guanine Ring Mode and the Torsion Angle of the C3′–C4′ Bond.* In Table II we list the conformation-sensitive guanine ring frequencies from several structures together with the C3′–C4′ torsion angles from model-building and X-ray structure studies (Wang et al., 1982a). Additional studies will be required in order to determine whether a reliable correlation can be established between the sugar torsion angle and the guanine ring mode. Nevertheless, the present data are consistent with a sugar pucker for the internal G(6) residue of $[d(CCCCGGGG)]_2$ close to that found in A DNA and a sugar pucker for the internal G(5) residue more similar to that of B DNA.

*Backbone Structures and Raman Markers in A DNA and B DNA.* One of the most pronounced differences between the solution and crystal spectra of $[d(CCCCGGGG)]_2$ is the intense 807-$cm^{-1}$ line in the latter. As detailed elsewhere (Thomas & Hartman, 1973; Erfurth et al., 1975; Prescott et al., 1984), a line near this frequency is due to a quasi-symmetric OPO stretching vibration and is characteristic of A-DNA backbone geometry. In B DNA, the corresponding backbone vibration gives rise to two Raman lines at ca. 790 (symmetric component) and ca. 835 $cm^{-1}$ (antisymmetric component). Benevides & Thomas (1983) have shown from studies of deuterated DNAs that the B-DNA band observed at 790 $cm^{-1}$ actually comprises two Raman components, a cytosine ring breathing mode centered near 784 $cm^{-1}$ and the OPO symmetric stretching vibration of the B backbone, which occurs closer to 790 $cm^{-1}$. These features are evident in the deconvoluted spectrum of Figure 6.

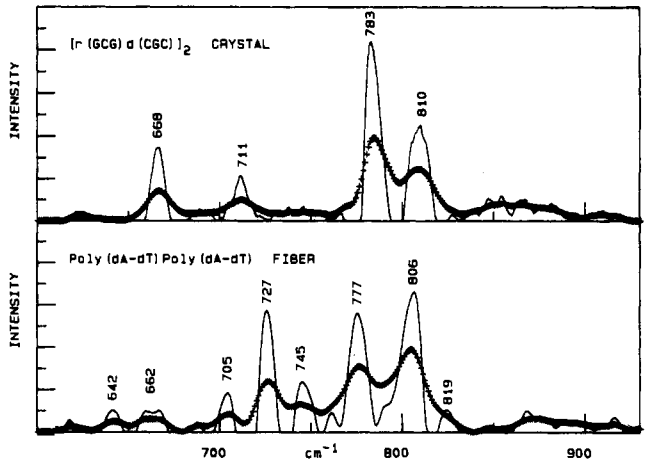



FIGURE 7: (Top) Raman spectrum of $[r(GCG)d(CGC)]_2$ crystal in the 600–900-$cm^{-1}$ region (+) compared with its deconvoluted spectrum (—). (Bottom) Corresponding data for the A-helix fiber of poly(dA-dT)·poly(dA-dT), taken from Thomas & Benevides (1985). Other conditions are as given in Figure 6.

Although it has been suggested elsewhere that A DNA also contains a backbone mode at 790 $cm^{-1}$ (Nishimura et al., 1983, 1985), we find scant evidence for such a Raman line in our spectra of A-DNA models. We have examined the Raman scattering in the 600–900-$cm^{-1}$ region of the DNA/RNA hybrid $[r(GCG)d(CGC)]_2$, as a model for pure A helix containing GC pairs only, and the same spectral region of low RH fibers of poly(dA-dT)·poly(dA-dT), as a model for pure A helix containing AT pairs only. Both the observed and deconvoluted spectra are shown in Figure 7. The results indicate no prominent Raman line centered near 790 $cm^{-1}$ for these A-helix models. The very weak line near 792 $cm^{-1}$ in poly(dA-dT)·poly(dA-dT) can be attributed entirely to a thymidine ring mode not associated with secondary structure and not sensitive to conformation (Thomas et al., 1985). Similarly, the apparent asymmetry in the 783-$cm^{-1}$ line of $[r(GCG)d(CGC)]_2$, discernible only in the high-resolution spectrum of Figure 1 (inset), indicates at best an extremely weak 790-$cm^{-1}$ line in the A helix, of little practical use for structure diagnosis.

Comparison with the $[d(CCCCGGGG)]_2$ crystal (Figure 6) confirms that partial B-DNA character is required to generate a conformation-sensitive 790-$cm^{-1}$ line of appreciable intensity. We conclude that the A helix contains no significantly intense Raman marker near 790 $cm^{-1}$, associated with either GC or AT sequences.

It is interesting to note that the Z-DNA polymer, high-salt poly(dG-dC)·poly(dG-dC), as well as Z-DNA oligomers investigated previously (Benevides & Thomas, 1983; Benevides et al., 1984b), all exhibit a weak Raman line near 790 $cm^{-1}$. In the Z form, C2′-*endo*/*anti*-cytidine is paired with C3′-*endo*/*syn*-guanosine. The 790-$cm^{-1}$ frequency appears to be common to GC structures containing C2′-*endo*/*anti*-cytidine. Further crystal analyses will be conducted to confirm this assignment.

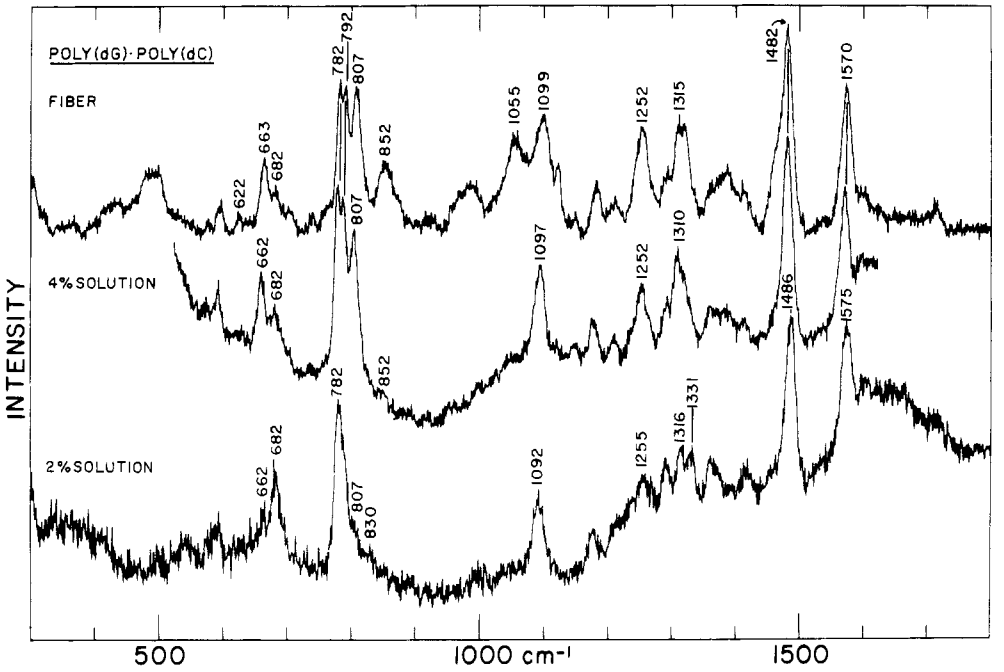



FIGURE 8: Raman spectra of poly(dG)·poly(dC) in the following states: oriented fiber at 75% relative humidity (top); 40 mg/mL in aqueous solution at pH 7.5 (middle); 20 mg/mL in aqueous solution at pH 7.5 (bottom).

*Poly(dG)·Poly(dC) in Fiber and Solution States.* In Figure 8, the Raman spectra of poly(dG)·poly(dC) in fibrous and solution states are shown for comparison with the spectrum of the $[d(CCCCGGGG)]_2$ crystal (Figure 5). The significant features of these data and their interpretations are as follows:

(i) All spectra have a line near 807 $cm^{-1}$, but the line intensity varies greatly from one spectrum to another, indicating different amounts of A-helix structure. The A-helix marker is most intense for the concentrated (40 mg/mL) solution of poly(dG)·poly(dC), followed by the fiber at 75% RH, the $[d(CCCCGGGG)]_2$ crystal, and the more dilute polymer solution (20 mg/mL).

(ii) All spectra have a line between 830 and 852 $cm^{-1}$, the intensity of which varies inversely with that of the 807-$cm^{-1}$ line. Accordingly, as A-helix content diminishes, the alternate structure giving rise to the 830–850-$cm^{-1}$ line(s) increases. Because of the similarity of the 852-$cm^{-1}$ line of fibrous poly(dG)·poly(dC) to the 850-$cm^{-1}$ line of $[d(CCCCGGGG)]_2$, we believe these frequencies have a common structural origin, namely, a "modified" B-type backbone (see Raman Spectra and Nucleoside Conformations of $[d(CCCCGGGG)]_2$). The corresponding Raman line in dilute poly(dG)·poly(dC) occurs at 830 $cm^{-1}$, close to the value expected of a "normal" B DNA.

(iii) Several other Raman markers of poly(dG)·poly(dC) and $[d(CCCCGGGG)]_2$ fit into a pattern consistent with the OPO backbone markers, including the band due to dioxy stretching ca. 1090–1100 $cm^{-1}$. The band has the expected position (1097 $cm^{-1}$) for a predominantly A-DNA structure in the concentrated poly(dG)·poly(dC) solution and the expected position (1092 $cm^{-1}$) for a predominantly B-DNA structure in the dilute solution.

(iv) Regarding guanosine conformation markers, the $[d(CCCCGGGG)]_2$ crystal displays the strongest C3′-endo marker (663 $cm^{-1}$), followed closely by the concentrated poly(dG)·poly(dC) solution and the fiber, whereas the dilute poly(dG)·poly(dC) solution contains the strongest C2′-endo marker (682 $cm^{-1}$).

(v) The different intensity of the cytosine ring mode near 1240–1250 $cm^{-1}$ in all poly(dG)·poly(dC) structures (Figure 8) indicates different base-stacking arrangements in the various structures. The fact that the Raman pattern in the 1200–1400-$cm^{-1}$ region of the $[d(CCCCGGGG)]_2$ crystal resembles more closely the pattern of fibrous poly(dG)·poly(dC) than that of aqueous poly(dG)·poly(dC) is in accord with the previous discussions.

(vi) In summary, although none of the samples represented in Figure 8 appears to contain a homogeneous secondary structure, the spectrum of the concentrated poly(dG)·poly(dC) solution resembles most closely the spectrum of the A-DNA model, $[r(GCG)d(CGC)]_2$ (Figure 1), while the spectrum of the dilute poly(dG)·poly(dC) solution is more like that of B DNA [cf. poly(dG-dC)·poly(dG-dC), Figure 4]. However, it is important to note that the intermediate cases, i.e., the $[d(CCCCGGGG)]_2$ crystal and the fiber of poly(dG)·poly(dC), contain geometries in which the non-A structure departs significantly from that of prototypical B DNA. Under Raman Spectra and Nucleoside Conformation of $[d(CCCCGGGG)]_2$, we have speculated that the modified B-DNA structure of crystalline $[d(CCCCGGGG)]_2$ may involve variants of C2′-*endo*/*anti*-G, including the C1′-exo/anti conformer, as in $[d(GGCCGGCC)]_2$. Clearly, such variants may also be present in the fiber of poly(dG)·poly(dC).

## Conclusions

The DNA/RNA hybrid $[r(GCG)d(CGC)]_2$ adopts a uniform A-helix structure in the crystal examined in this work. The present results validate the assignment of the following Raman markers: C3′-*endo*/*anti*-guanosine at 668 ± 2 $cm^{-1}$ and A backbone at 810 ± 2 $cm^{-1}$. Both deoxyguanosine and deoxyriboguanosine residues of the crystalline hybrid yield the same Raman markers. In aqueous solution, the A-helix structure of $[r(GCG)d(CGC)]_2$ is substantially retained. It appears that only deoxy residues of the hybrid can be converted to the C2′-endo/anti conformation, which may be a consequence of fraying at the ends of the A-type double helix. These results indicate that a hexanucleotide sequence is sufficient to retain a predominantly A-type structure for the hybrid in aqueous solution at physiological conditions.

The $[d(CCCCGGGG)]_2$ crystal contains both A and B backbone characteristics and both C3′-endo and C2′-endo

nucleoside puckers. Raman and X-ray data suggest an overall structure similar to that of $[d(GGCCGGCC)]_2$ (Wang et al., 1982a). Raman markers for the C3′-endo nucleosides and the A backbone of $[d(CCCCGGGG)]_2$ are identical with those of $[r(GCG)d(CGC)]_2$; however, the Raman markers assigned to the B-like residues of the octamer indicate small but significant structural differences from canonical B-form polynucleotide models.

The present results, together with previously published Raman spectra of DNA crystals, fibers, and solutions (Erfurth et al., 1972; Prescott et al., 1984; Benevides et al., 1984b; Benevides & Thomas, 1985; Thomas et al., 1985), indicate substantial variability in the positions of backbone and nucleoside markers associated with B DNA, in contrast to the constancy of markers associated with A DNA and Z DNA. Thus, the Raman marker identified with the B backbone ranges from ca. 826 $cm^{-1}$ in aqueous $[d(CCCCGGGG)]_2$ and ca. 829 $cm^{-1}$ in aqueous low-salt poly(dG-dC)·poly(dG-dC) to values as high as 850 $cm^{-1}$ in crystalline $[d(CCCCGGGG)]_2$ and 852 $cm^{-1}$ in aqueous poly(dG)·poly(dC). A consensus value of 834 ± 5 $cm^{-1}$ is usually cited for DNA of high molecular weight and irregular base sequence. (Corresponding markers of A and Z backbone occur at 807 and 748 $cm^{-1}$, respectively, with little or no observed dependence on base sequence or chain length.) Accordingly, we conclude that the larger the DNA fragment, the greater is the diversity of B-family structures that contribute to the overall secondary structure. It is this "averaged" B structure that gives rise to the band envelope centered at 834 $cm^{-1}$ in the Raman spectra of DNA fibers and solutions. A similar conclusion is reached from X-ray results (Dickerson et al., 1982).

The structure-sensitive guanine ring mode of B DNA (i.e., 680–687 $cm^{-1}$ for C2′-*endo*/*anti*-G conformations) also appears to span a greater range than its counterparts in A-DNA (C3′-endo/anti) and Z-DNA (C3′-endo/syn) structures. This finding is consistent with the conclusion that minor variations in the geometry of B DNA, perhaps associated with base sequence, are reflected also in the Raman markers assigned to nucleoside conformers. Therefore, both the backbone OPO and guanosine conformation markers give indications that high molecular weight B DNA may contain an ensemble of similar but nonidentical conformers, dictated by the local base sequence.

Alternation of sugar ring pucker along each oligonucleotide strand is a conspicuous feature of the structure proposed here for crystalline $[d(CCCCGGGG)]_2$. It is tempting to conclude that such alternation may be a general characteristic of DNA structure. However, alternation of pucker is not found in either crystalline $[r(GCG)d(CGC)]_2$ or aqueous $[d(CCCCGGGG)]_2$. Additional DNA models will have to be investigated in order to determine the relevance of alternating pucker to the A and B forms of DNA.

Poly(dG)·poly(dC) would appear to represent a poor model for characterization of DNA structure. Raman spectra of solutions at different solute concentrations and of fibers are not identical with one another, indicating that effects of intermolecular aggregation may influence or even determine the secondary structure. A similar conclusion has been reached from CD studies of substantially more dilute solutions of poly(dG)·poly(dC) (Y. Kyogoku, J. M. Benevides, and G. J. Thomas, Jr., unpublished results). Apparently, clusters of G residues in oligomer and polymer structures tend to generate anomalous CD patterns, which cannot be interpreted with correlations developed for random or irregular base sequences. The present results confirm the findings of others (Nishimura et al., 1985) and indicate the need for caution in applying empirical CD correlations to DNA fragments containing clusters of G residues.

Finally, we may draw an interesting conclusion from the appearance of a single sharp Raman line at 810 $cm^{-1}$ as the A-helix marker in the $[r(GCG)d(CGC)]_2$ crystal. The occurrence of only one line at 810 $cm^{-1}$ (rather than a doublet at 807 and 813 $cm^{-1}$ due to ribo and deoxyribo residues, respectively) is direct evidence for a uniform structure isomorphous with neither A DNA nor A RNA. Furthermore, the singlet provides indirect evidence that the different A backbone markers observed in A DNA (807 $cm^{-1}$) and A RNA (813 $cm^{-1}$) probably arise from small but significant structural differences in the two kinds of A helix and not from the effects of different mechanical coupling from the two kinds of sugar ring (e.g., coupling with respective C2′ substituents). In other words, the nonidentical Raman markers observed in A RNA and A DNA are more likely due to differences in secondary structure than to differences in primary structure. If this interpretation is correct, we may hope to exploit the Raman OPO frequency to distinguish A-DNA and A-RNA geometries in hybrid structures that manifest such differences or to detect hybrid structures in which RNA conformation dominates DNA, or vice versa, rather than the present situation of an "intermediate" geometry. In future work it will be of interest to investigate the Raman spectra of well-defined RNA oligonucleotide fragments, the molecular structures of which have been determined by X-ray diffraction.

**Registry No.** r(GCG)d(CGC), 99327-09-0; d(CCCCGGGG), 99327-10-3; poly(dG)·poly(dC), 25512-84-9.

## REFERENCES


Arnott, S., Chandrasekaran, R., Hall, I. H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141–4155.

Benevides, J. M., & Thomas, G. J., Jr. (1983) *Nucleic Acids Res. 11*, 5747–5761.

Benevides, J. M., LeMeur, D., & Thomas, G. J., Jr. (1984a) *Biopolymers 23*, 1011–1024.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., Rich, A., & Thomas, G. J., Jr. (1984b) *Nucleic Acids Res. 12*, 5913–5925.

Brown, K. G., Kiser, E. J., & Peticolas, W. L. (1972) *Biopolymers 11*, 1855–1869.

Dickerson, R. E., Drew, H. R., Conner, B. N., Kopka, M. L., & Pjura, P. E. (1983) *47*, 13–24.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.

Erfurth, S. C., Bond, P. J., & Peticolas, W. L. (1975) *Biopolymers 14*, 1245–1257.

Fodor, S. P. A., Starr, P. A., & Spiro, T. G. (1985) *Biopolymers 24*, 1493–1500.

Inui-Nara, H., Akutsu, H., & Kyogoku, Y. (1985) *J. Biochem. (Tokyo) 98*, 629–636.

Lafleur, L., Rice, J., & Thomas, G. J., Jr. (1972) *Biopolymers 11*, 2423–2437.

Lord, R. C., & Thomas, G. J., Jr. (1967) *Spectrochim. Acta, Part A 23A*, 2551–2591.

Mellema, J.-R., Haasnoot, C. A. G., van der Marel, G. A., Willie, G., van Boeckel, C. A. A., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 5717–5738.

Nishimura, Y., Tsuboi, M., Nakano, T., Higuchi, S., Sato, T., Shida, T., Uesugi, S., Ohtsuka, E., & Ikehara, M. (1983) *Nucleic Acids Res. 11*, 1579–1588.

Nishimura, Y., Torigoe, C., & Tsuboi, M. (1985) *Biopolymers* (in press).

Prescott, B., Gamache, R., Livramento, J., & Thomas, G. J.,

Jr. (1974) *Biopolymers 13*, 1821-1845.
Prescott, B., Steinmetz, W., & Thomas, G. J., Jr. (1984) *Biopolymers 23*, 235-256.
Sanderson, M. R., Mellema, J.-R., van der Marel, G. A., Wille, G., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 3333-3346.
Thamann, T., Lord, R. C., Wang, A. H. J., & Rich, A. (1981) *Nucleic Acids Res. 9*, 5443-5457.
Thomas, G. J., Jr., & Hartman, K. A. (1973) *Biochim. Biophys. Acta 312*, 311-322.
Thomas, G. A., & Peticolas, W. L. (1983) *J. Am. Chem. Soc. 105*, 993-996.
Thomas, G. J., Jr., & Agard, D. A. (1984) *Biophys. J. 46*, 763-768.
Thomas, G. J., Jr., & Benevides, J. M. (1985) *Biopolymers 24*, 1101-1105.
Thomas, G. J., Jr., Benevides, J. M., & Prescott, B. (1985) *Biomolecular Stereodynamics III, Proceedings of the Fourth Conversation in Biomolecular Stereodynamics* (Sarma, R. H., & Sarma, M. H., Eds.) Adenine Press, Guilderland, NY (in press).
van der Marel, G. A., van Boeckel, C. A. A., Wille, G., & van Boom, J. H. (1981) *Tetrahedron Lett.*, 3887.
Wang, A. H.-J., Fujii, S., van Boom, J. H., & Rich, A. (1982a) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3968-3972.
Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, C. A. A., & Rich, A. (1982b) *Nature (London) 299*, 601-604.

# Hydrogen-1 Nuclear Magnetic Resonance Investigation of *Clostridium pasteurianum* Rubredoxin: Previously Unobserved Signals†

R. Krishnamoorthi and John L. Markley*
*Department of Chemistry, Purdue University, West Lafayette, Indiana 47907*
Michael A. Cusanovich and C. T. Przysiecki
*Department of Biochemistry, University of Arizona, Tucson, Arizona 85721*
*Received March 19, 1985*

ABSTRACT: Previously unobserved signals were located in the 470-MHz $^1$H NMR spectra of oxidized and reduced rubredoxin (Rd) from *Clostridium pasteurianum*. When the protein was oxidized, some of the resonances broadened beyond detection. Longitudinal relaxation ($T_1$) measurements identified a number of these peaks as arising from residues close to the paramagnetic iron; these resonances exhibited short $T_1$ values attributable to the dominant electron–nuclear dipolar relaxation mechanism. The chemical shifts of these peaks were not strongly dependent on the oxidation state of the protein, although relative ratios of line widths of several peaks in the spectra of oxidized and reduced Rd suggested localized conformational changes of the protein as a result of oxidation. Furthermore, spectra of the oxidized protein collected in the range 8–60 °C revealed no appreciable changes in the chemical shifts of these peaks with temperature. These results seem to point out a negligible dipolar contribution, due to either magnetic anisotropy or zero field splitting, to the observed shifts in the spectrum of oxidized Rd. Resonances were assigned to tyrosine-11 or phenylalanine-49 (but not to either specifically) on the basis of their $T_1$ values and the X-ray diffraction data of the protein molecule [Watenpaugh, K. D., Sieker, L. C., Herriott, J. R., & Jensen, L. H. (1973) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B29*, 943–956; and a further refinement deposited with the Protein Data Bank]. An upfield-shifted peak at about −1.1 ppm in the spectra of both oxidized and reduced Rd was assigned to a methyl group. Broad, rapidly relaxing peaks in the 8–10 ppm spectral region of oxidized Rd and even broader peaks in the 0 to −4.5 ppm spectral region of reduced Rd were identified with the ligated cysteinyl $\alpha$- and $\beta$-hydrogens, respectively, as suggested by an analysis of their chemical shifts, line widths, and longitudinal relaxation properties.

Our interest in high-potential iron–sulfur proteins (HiPIPs)[1] and ferredoxins prompted us to investigate rubredoxin (Rd) as a potential model for assigning the hyperfine-shifted $^1$H NMR peaks in spectra of the more complicated iron–sulfur proteins. For example, in contrast to the HiPIPs, which have 4Fe-4S clusters (Carter, 1977), Rd has no acid-labile sulfur and only one iron atom, which is nearly tetrahedrally coordinated to the sulfur of four cysteinyl ligands (Cammack et al., 1977). Various physical techniques, including X-ray diffraction (Herriott et al., 1970; Watenpaugh et al., 1973, 1979), ESR (Atherton et al., 1966), Mössbauer (Phillips et al., 1970), NMR (Phillips et al., 1970), and laser Raman

† This research was supported in part by grants from the U.S. Department of Agriculture Competitive Research Grants Office, Cooperative State Research Service, Science and Education (82-CRCR-1-1045 to J.L.M.), and the National Institutes of Health (GM 21277 to M.A.C.). The Purdue University Biochemical Magnetic Resonance Laboratory has financial support from Grant RR 01077 from the Biotechnology Resources, National Institutes of Health.
* Address correspondence to this author at the Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin—Madison, Madison, WI 53706.

[1] Abbreviations: HiPIPs, high-potential iron–sulfur proteins; 4Fe-4S, cluster containing four iron atoms and four inorganic sulfur (acid-labile) atoms; Rd, rubredoxin; ESR, electron spin resonance; NMR, nuclear magnetic resonance; ppm, parts per million; DSS, sodium 4,4-dimethyl-4-silapentanesulfonate.

0006-2960/86/0425-0050$01.50/0 