Sequence Dependence of Conformations of Self-Complementary Duplex Tetradeoxynucleotides Containing Cytosine and Guanine†

Gerald A. Thomas and Warner L. Peticolas*

ABSTRACT: The four self-complementary tetradeoxynucleotides which contain only cytosine and guanine are 5′-d-(CpGpCpG)-3′, 5′-d(CpCpGpG)-3′, 5′-d(GpCpGpC)-3′, and 5′-d(GpGpCpC)-3′. The Raman spectra of aqueous solutions (about 0.05 M in monomer) of these tetranucleotides at pH 7 and 2 °C show clearly that these self-complementary tetranucleotides form double-stranded duplex structures of the canonical B type when the NaCl concentration is 0.5 M NaCl. If the temperature is raised to 50 °C, the Raman spectra show that in each case the double-helical B form melts in a noncooperative way to a disordered single-chain form. On the other hand, if the salt concentration is raised to saturation, the Raman spectrum of only one of these four tetranucleotide solutions at 2 °C is changed in any substantial way. The Raman spectrum of the tetranucleotide 5′-d(CpGpCpG)-3′ at 2.2 °C and at 4 M or higher salt concentration strongly resembles that of double-helical Z-form poly(dC-dG) taken under similar conditions. We conclude that the tetramer 5′-d(CpGpCpG)-3′ is the only self-complementary double-helical tetranucleotide containing only cytosine and guanine in which the B–Z transition can be induced by increasing the salt concentration. This tetramer has several types of stacking interactions which differ markedly from stacking interactions in the other tetramers and may account for the enhanced stabilization of its Z conformation.

Recently various oligonucleotides have been synthesized and crystallized, and their conformation in the crystalline state has been determined by X-ray diffraction techniques (Wing et al., 1980; Wang et al., 1979, 1981; Drew et al., 1980; Drew & Dickerson, 1981; Viswamitra et al., 1978). This work has led to the discovery of new conformations of oligodeoxyribonucleotides. The conformation of a number of d(CG)$_n$ oligomers has been determined, and in each case in the crystalline state the conformation has been found to exhibit a left-handed helix with a zigzag form leading to the name Z-DNA. Recently, Thaman et al. (1981) have shown that the Raman spectra of crystals of d(CG)$_3$ oligomers in the Z conformation exhibit Raman bands characteristic of high-salt form of double-helical poly[d(G-C)] as well as the high-salt form of the duplex hexamer, d(CG)$_3$. In addition, Raman studies by Nishimura et al. (1983) have extended this work with specific conformational assignments for these Raman marker bands. Consequently, the existence of specific Z-form marker bands in the Raman spectra of oligonucleotides containing C and G has been well established. Earlier work in this laboratory and elsewhere has established the existence of marker bands for the A, B, and C forms of DNA by direct comparison of X-ray diffraction patterns and Raman spectral lines (Erfurth et al., 1972, 1975; Brown & Peticolas, 1975; Brown et al., 1972; Thomas & Hartman, 1973; Goodwin & Brahms, 1978). Furthermore, very recently it has been shown possible to characterize mixed forms of DNA in which both C3′ and C2′ ring puckers are seen (Thomas & Peticolas, 1983a,b).

The biological significance of Z-DNA is the subject of much discussion (Wang et al., 1979, 1981; Drew et al., 1980; Behe & Felsenfeld, 1981; Klysic et al., 1981). In general, it is presumed that the left-handed Z configuration does not support biological functions directly but serves as a structural control element for processes involved in genetic expression. If the Z-DNA is used as a control device, the ease with which a segment of DNA can be induced into the Z form must be dependent upon the base sequence. All of the oligomers so far found in the Z conformation seem to be oligomers of alternating CG sequence. However, it is not clear just exactly what base sequences are necessary to sustain the Z conformation. In order to study the effect of base sequence on the B to Z transition, we have prepared double-stranded complexes of self-complementary tetramers containing only dG and dC and examined their conformation by classical laser Raman spectroscopy. The classical technique permits the obtaining of Raman spectra of solutions at very high molar concentrations. Since the formation of double-stranded duplex structures of the tetranucleotides is governed by the law of mass action, it seems that Raman spectroscopy is well suited to such an investigation since duplex formation can be obtained at high concentrations and low temperature. Earlier studies using circular dichroism apparently failed to show evidence of duplex structures in tetramers. Presumably this is due to the fact that circular dichroism measurements can only be made easily on solutions with an absorption of less than 2. Such low concentrations of tetramers probably prevent them from forming double-stranded duplex structures (Quadrafolio et al., 1981).

Materials and Methods

The following oligodeoxynucleotides were purchased from P-L Biochemicals: 5′-d(CpGpCpG)-3′, lot 670-67; 5′-d-(CpGpG)-3′, lot 670-59; 5′-d(GpGpCpC)-3′, lot 670-61; 5′-d(GpCpGpC)-3′, lot 734-16. Sequence conformation was performed by the supplier utilizing a modified procedure of the chemical method of Maxam & Gilbert (1979). Each oligonucleotide was supplied as the corresponding sodium salt and was employed without further purification. All other materials were reagent or spectral grade in quality and were obtained from conventional chemical sources.

Raman samples were prepared in the following manner. Oligodeoxynucleotides were dissolved at a concentration of 0.25 OD unit/μL corresponding to an approximate concentration of 2 × 10^{-2} M in nucleotide base. Samples were prepared

---

† From the Department of Chemistry and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403. *Received December 13, 1983.* This work was supported by Grant 76-82222 from the National Science Foundation and Grant GM-15547 from the National Institutes of Health.

* Address correspondence to this author at the Department of Chemistry.

0006-2960/84/0423-3202$01.50/0

under conditions of low sodium chloride concentration by dissolving the oligonucleotide in either a 1 or 0.5 M sodium chloride solution. Samples exposed to elevated sodium chloride concentrations were prepared either by dissolving the oligonucleotide directly in a solution containing 4 M NaCl or by adding solid sodium chloride to a previously prepared oligonucleotide solution. Oligonucleotide solutions were prepared both in the presence and in the absence of 0.02 M sodium perchlorate. The sodium perchlorate was initially added as an internal standard. Samples were also prepared in the absence of sodium perchlorate to eliminate the possibility of spectral interferences. At the concentrations employed no detectable differences in the behavior of the oligodeoxynucleotides in the presence and absence of added sodium perchlorate were observed. Samples were also prepared in the presence of magnesium by the addition of solid magnesium chloride or by the addition of an aliquot of concentrated magnesium chloride solution to samples dissolved previously in 1 or 0.5 M NaCl. After equilibration a 5-μL aliquot of each sample was placed in a capillary tube which was inserted in a thermostated copper block, and its Raman spectrum was taken with the use of an argon laser.

Raman data for each tetramer were collected under various conditions of salt concentration and temperature in order to determine if each oligonucleotide forms a double-stranded structure and if so to which conformational genus this duplex structure belongs. The basic components of the apparatus include a Spectra Physics Model 165 argon ion laser, a Spex 1301 spectrometer interfaced to a Varian 620/i computer enabling computer-controlled data acquisition and data reduction capabilities, and a temperature-regulated sample compartment. The 514.5-nm line of the argon ion laser was used for the Raman spectra. Spectra were recorded by using approximately 200-mW of laser power measured at the sample. Spectra were recorded with a spectral resolution of 4.6 $\mathrm{cm^{-1}}$.

Results and Discussion

The present investigation is concerned with the elucidation of the conformational state of the self-complementary tetraoligonucleotides d(CpGpCpG), d(GpCpGpC), d(GpGpCpC), and d(CpCpGpG) in dilute and concentrated salt solutions. In particular, we are interested in the ability of these oligomers to form duplex or double-helical structures and the tendency of these duplex structures to go from a helical structure belonging to the B genus to the Z genus as the salt concentration is increased. Reference spectra of the B and Z form of poly[d(G-C)] in solutions of high and low salt concentrations were obtained in order to have Raman spectra which could be used for direct comparison with the Raman spectra obtained on the oligomers. These reference spectra are shown in Figure 1. From the unusually high quality of these spectra a large number of distinct differences in the Raman spectra of the two forms can be observed. In the transformation from B to Z form a drastic decrease occurs in the intensity of the band at 682 $\mathrm{cm^{-1}}$ due to guanine with the generation of a new band at 627 $\mathrm{cm^{-1}}$. A reduction also occurs in the intensity of the marker band for the C2′-endo furanose ring conformation at 829 $\mathrm{cm^{-1}}$. Other changes in going from the B to the Z form include intensity increases at 1265 and 1317 $\mathrm{cm^{-1}}$, an intensity reduction at 1420 $\mathrm{cm^{-1}}$, and a shift in the band at 1362 to 1355 $\mathrm{cm^{-1}}$. The intensity increase in the band at 1317 $\mathrm{cm^{-1}}$ in the Z form obscures the visibility of a band at 1335 $\mathrm{cm^{-1}}$ which results in the apparent collapse of the two bands at 1317 and 1335 $\mathrm{cm^{-1}}$ in the B genus into a single peak at 1317 $\mathrm{cm^{-1}}$ in the Z form. Two bands, one at 1418 $\mathrm{cm^{-1}}$ and one at 1426 $\mathrm{cm^{-1}}$, are visible in the spectrum of the Z form while only one band at 1420 $\mathrm{cm^{-1}}$ is visible in the Raman spectrum of the B genus. Other less prominent and ill-defined spectral alterations are also observable. The more significant band changes which we take as defining the B to Z transition to its fullest extent are listed in Table I. These data together with Figure 1 will be used to analyze the spectra of the tetramers.

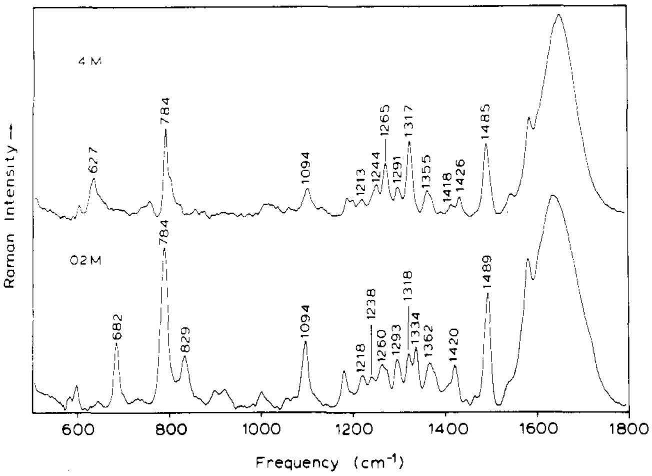

FIGURE 1: Raman spectra of poly[d(G-C)] (0.02 M in nucleotide) in neutral salt solution: 0.02 M (bottom spectrum) and 4 M (top spectrum). The spectra were taken with a 514-nm line from an argon laser.

Table I: Significant Characteristic Raman Spectral Alterations Which Accompany the Transformation in Poly[d(G-C)] from B Form to Z Form[a]

Raman band B form ($\mathrm{cm^{-1}}$)	change	Raman band Z form ($\mathrm{cm^{-1}}$)
	inc	627
682	dec	
829	dec	
1260	inc	1265
1318	inc	1317
1334		NR
1362	shift	1355
1420	dec	1418
NR		1421

[a] inc, intensity increase; dec, intensity reduction; NR, not resolvable.

The tetranucleotide d(CpGpCpG) has been previously studied by circular dichroism (Quadrafolio et al., 1981). In this previous investigation no change in the CD spectrum of the tetramer in dilute solution was observed in response to increasing salt concentration. Due to the lack of observable salt-dependent changes in the CD spectra of the tetramer, the previous investigators concluded that under the conditions employed the tetramer does not undergo a transformation to a left-handed form. The behavior of the tetramer observed in the CD investigation is probably due to the lack of formation of a duplex structure by the tetramer at the concentration of tetramer employed, namely, less than 1 mM. Evidence to support this contention is provided by the lack of temperature-dependent hyperchromism noted by these authors. In contrast, the concentrations we have used are about 100 times higher, and thus, duplex formation is observed below 15 °C. The Raman spectrum of d(CpGpCpG) in 0.5 M NaCl recorded from a sample maintained at 2.2 °C is shown in the middle spectrum of Figure 2. This spectrum is quite similar to the spectrum observed for poly(dG-dC) at an NaCl concentration of 0.02 M. The band of moderate strength at 830 $\mathrm{cm^{-1}}$ is readily identifiable as being due to the C2′-endo furanose ring conformation (Thomas & Peticolas, 1983a,b). Furthermore, the progression of bands in the base vibrational

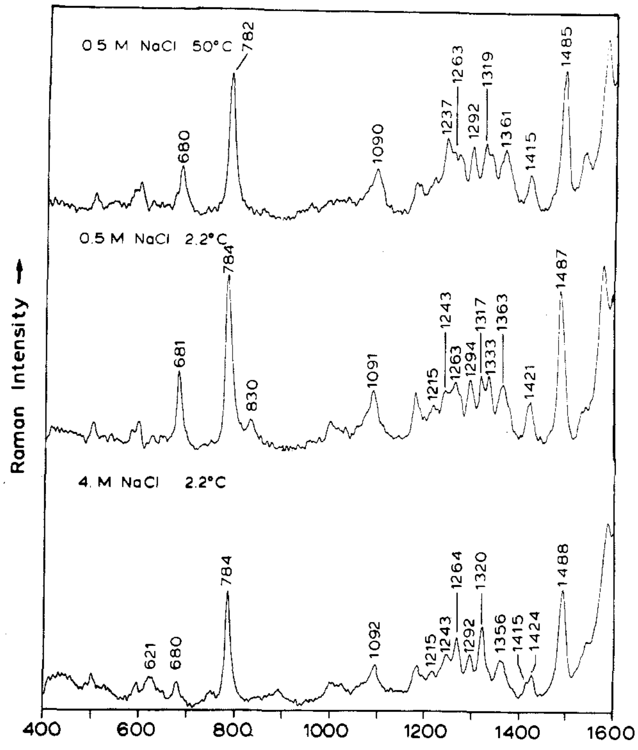

FIGURE 2: Raman spectra of d(CpGpCpG) in 0.5 M NaCl solution taken at 2.2 °C (middle spectrum) and at 50 °C (top spectrum) and in 4 M NaCl solution at 2.2 °C (bottom spectrum). The center spectrum is the one belonging to the B genus with which the other two should be compared. The top spectrum is the melted or disordered form while the bottom spectrum is of the high-salt form belonging to the Z genus.

region from 1200 to 1600 cm^{-1} duplicates the spectral properties of the right-handed B form of poly(dG-dC) except for minor intensity variations. Further evidence that the tetramer has formed a double-stranded duplex structure is obtained from the melting behavior which is observed in the Raman spectrum as the temperature is slowly raised from 2.2 to 50 °C. The top spectrum in Figure 2 shows the spectrum of the melted duplex structure. As expected, elevation of the temperature to 50 °C causes marked alterations in the spectrum of the tetramer due to disruption of base-stacking and base-pairing interactions. The band at 830 cm^{-1} is no longer visible due to the flexible nature of the DNA backbone (Thomas & Peticolas, 1983a,b). An intensity reduction at 681 cm^{-1} is also noted as has been seen in the previous work of Erfurth & Peticolas (1975) on the melting of ordinary DNA. This is due to a hypochromic reduction in the intensity of a guanine ring vibration. The intensity of the band at 1237 cm^{-1} increases markedly upon melting. The increase of the band at 1237 cm^{-1} is due to the unstacking of cytosine (Chou & Thomas, 1977; Strommen & Peticolas, 1982). The bands at 1317 and 1333 cm^{-1} also change relative intensity in comparison to one another. At 2.2 °C the 1317- and 1333-cm^{-1} bands appear as two resolvable bands of almost equal intensity. At 50 °C the band at 1333 cm^{-1} is no longer resolvable and appears as a shoulder on the band at 1319 cm^{-1}. Both bands are presumably due to guanine vibrations. The transformation in these two bands is at present not assigned. Thermal disruption of the secondary structure of this tetramer as well as the other tetramers to be discussed occurs progressively over a wide range of temperatures between 15 and 50 °C. The melting profile is fairly broad. This indicates that the cooperativity of thermal denaturation is extremely low which is what one may expect based upon the length of the oligomer.

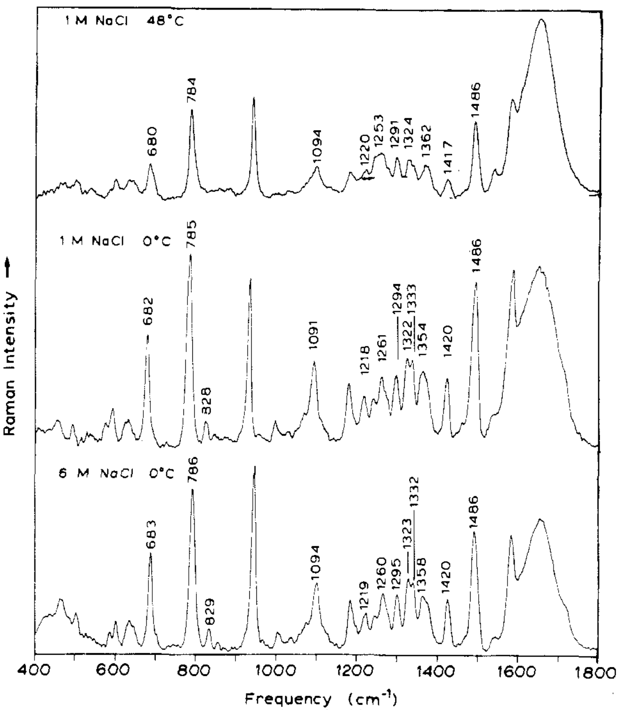

FIGURE 3: Raman spectrum of the tetradeoxynucleotide d(GpCpGpC). The center spectrum is that of the B form at 0 °C, the top spectrum is that of the melted B form at 48 °C, and the bottom spectrum is of the tetramer in greater than 4 M aqueous NaCl.

At high salt concentration this tetramer, d(CpGpCpG), exhibits a Raman spectrum characteristic of the Z form. The bottom spectrum in Figure 2 is the spectrum of this tetramer in 4 M NaCl recorded at 2.2 °C. As in the case of double-helical poly[d(G-C)] at high salt, we find a reduction in the intensity of the Raman bands at 680, 830, and 1415 cm^{-1}. A strong new band at 621 cm^{-1} appears. Intensity increases are observed in the Raman bands at 1264 and 1319 cm^{-1}. This latter change results from the replacement of the two bands at 1317 and 1333 cm^{-1} which are observed in the low-salt form with a single band at 1319 cm^{-1}. The band at 1363 cm^{-1} in the low-salt tetramer spectrum shifts to 1356 cm^{-1} in the high-salt form. In total, these changes are essentially the same as those found in the spectrum of poly[d(G-C)] in going from the low-salt B form to the high-salt Z form. Since the transition has also been found in the hexamer (Thamman et al., 1981), we may conclude that oligomers of CG form double-stranded duplex structures which can be induced into the Z form at high salt concentrations. It is also of interest to note that this tetranucleotide has been shown to adopt the Z form in the crystalline state (Drew et al., 1980).

The tetramer d(GpCpGpC) was also studied by using the Raman effect in the exact same manner as described above for d(CpGpCpG). The high- and low-temperature spectra of the low salt form as well as the high-salt, low-temperature form are shown in Figure 3. Again as in all of the tetramer Raman spectra figures, the middle spectrum is that of the B genus at low-temperature and moderate salt concentration while the top and the bottom spectra are the melted low-salt form and the cold high-salt form. The low-temperature form of the d(GpCpGpC) tetramer was taken in 1 M NaCl at 0 °C and closely resembles the low-temperature, low-salt form of d-(CpGpCpG). The band at 933 cm^{-1} is due to perchlorate as well as the weak structure at 630 cm^{-1}. Melting of the tetramer shows changes similar to that discussed for the d-(CpGpCpG) tetramer. For example, the 828-cm^{-1} band is

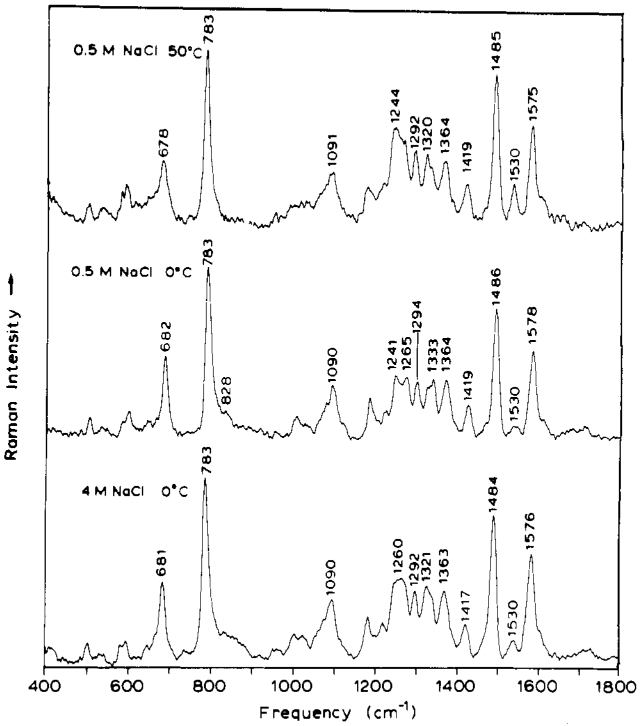

FIGURE 4: Raman spectra of the tetramer d(CpCpGpG) in 0.5 M NaCl, pH 7, at 0 °C (middle spectrum) and at 50 °C (top spectrum) and in 4 M to saturated salt solution (bottom spectrum).

absent in the melted form. Similarly a reduction in the band at 680 cm^{-1} is also noted. Variation in the relative intensity of the band which appears at approximately 1243 cm^{-1} at 0 °C is also observed. As the temperature is raised to 50 °C, a group of bands centered around 1240–1260 wavenumbers merge which results in the formation of a broad structure centered at 1253 cm^{-1}. The collective intensity of these bands at 1253 cm^{-1} can be estimated from a comparison of the total intensity in this region to that of the neighboring bands at 1220 and 1291 cm^{-1} as well as by an examination of the relative intensity of this band in comparison to the perchlorate peak at 933 cm^{-1}. Because of the extreme overlap of the bands in the 1253 wavenumber region, the individual bands are very difficult to resolve even by curve-fitting techniques. However, qualitatively the changes are obvious and may be used in the characterization of the disordered state of this tetramer.

The high-salt form of the d(GpCpGpC) tetramer does not resemble the high-salt form of either the poly[d(G-C)] or the d(CpGpCpG) tetramer. Attempts to induce a transition in the conformation of d(GpCpGpC) by increasing the NaCl concentration to saturation or by adding magnesium chloride to saturation did nothing to change the spectrum significantly. Thus, it must be concluded that the tendency of sections of guanine- and cytosine-containing DNA to form the Z conformation is dependent upon the specific starting sequence of the alternating guanine and cytosine residues.

The Raman spectra of the tetradeoxynucleotide d(CpCpGpG) is of particular interest since Conner et al. (1982) have shown that the molecular structure of d(ICpCpGpG) in the crystalline form is a fragment of the right-handed double helical A-DNA. Since Raman spectroscopy differentiates clearly between the A and B conformations, it is of interest to see what the conformation of d(CpCpGpG) is in solution under varying conditions of salt and temperature. Figure 4 shows the Raman spectra of the tetramer d(CpCpGpG) under varying conditions of temperature and salt. However, with this tetramer the Raman marker bands for the backbone conformation were so weak in the normal spectra that we had to subtract the water spectrum in order to assess more clearly the conformation of the tetramer from the bands in the 800–840-cm^{-1} region of the spectrum. Again in Figure 4 the middle spectrum is the spectrum of the tetramer in low (0.5 M) NaCl at 0 °C. Notice that in this spectrum the 828-cm^{-1} band is present but it is rather broad, which we take to indicate that though the duplex appears to be to a large extent in the B form, it is more flexible than ordinary DNA [see, for example, Thomas & Peticolas (1983)]. The weakness of the band at 829 cm^{-1} suggests that the duplex structure is not a rigid B form but may contain some conformational variability. Even though the duplex appears to be structurally flexible, there is no evidence in the Raman spectra that suggests that the duplex adopts an A-type helical form in solution under the conditions employed. The presence of an A-type helical form would be indicated by the presence of a strong band between 805 and 814 cm^{-1}. In the high-temperature spectrum (top) the 828-cm^{-1} band is completely gone and the 1244-cm^{-1} band has increased as has the 1530-cm^{-1} band. These changes indicate a disordering of the duplex structure of the tetramer. The changes which occur in the base vibrational region are more substantial than the extent of the decrease of the band at 829 cm^{-1}. This suggests that at low temperatures the duplex structure of the tetramer retains a significant degree of order even though the structure appears to be quite flexible and contains a reduced proportion of B form in comparison to the other tetramers discussed. In the high-salt, low-temperature Raman spectrum at the bottom of Figure 4 we find that the backbone marker band region is very broad and diffuse while the fingerprint region (1200–1400 cm^{-1}) somewhat resembles the high-temperature form. In particular, the 1321-cm^{-1} band in the high-salt form is higher than the 1333-cm^{-1} band which is also the case for the high-temperature disordered form. Thus, one may conclude that the tetranucleotide d(CpCpGpG) appears to be more easily disordered or more flexible than the other tetradeoxynucleotides.

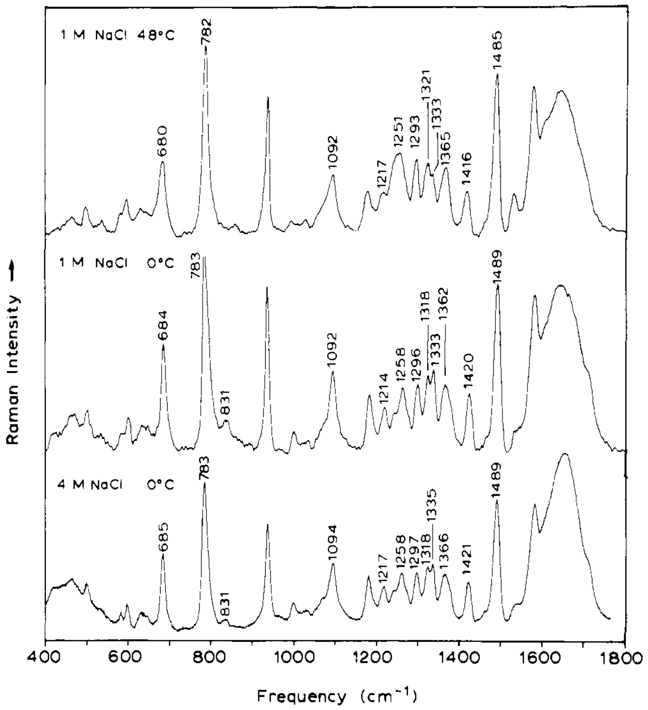

FIGURE 5: Raman spectra of the tetramer d(GpGpCpC) in 1 M NaCl, pH 7, at 0 °C (middle spectrum) and at 48 °C (top spectrum) and in 4 M NaCl at 0 °C (bottom spectrum).

The Raman spectra of the tetramer d(GpGpCpC) under similar conditions of salt and temperature are shown in Figure

5. These spectra resemble those of the tetramer d(GpCpGpC) in Figure 3. The middle spectrum shows that of the usual B-genus Raman spectrum while the top spectrum is of the disordered form. Again we see the change in the intensity of the two components of the 1320–1334-cm^{-1} doublet upon melting. There remains a slight feature at 850 cm^{-1} of unknown origin, and the strong band at 934 cm^{-1} is due to perchlorate. In this tetramer the high-salt form also appears as a somewhat flexible B-genus structure as the 831-cm^{-1} band is weak and broadened. The similarity to the melted low-salt form is apparent. Thus, it appears that high salt has a destabilizing tendency on these tetramers even if they do not go into the Z-genus family under these conditions.

Conclusions

The four self-complementary tetradeoxynucleotides that contain only cytosine and guanine exhibit different conformations under varying conditions of salt and temperature. The tetramer d(CpGpCpG) is the only one of the four that will go into a Z form in saturated salt solution. It is of special interest to note that the only other alternating pyrimidine–purine tetramer studied, d(GpCpGpC), does not adopt a form belonging to the Z genus in solutions of high or saturated salt. The retention of the B form by the tetramer d(GpCpGpC) in solutions of high salt concentration is probably due in part to both a sequence-dependent enhanced stabilization of the B form for this tetramer and a sequence-dependent destabilization of the corresponding left-handed Z form. It is of interest to learn what molecular interactions might account for the conformational differences in these two tetramers with alternating cytosine and guanine residues. Some clues may be obtaining from a careful study of the models showing the structure of Z-DNA.

In high molecular weight Z-DNA composed of alternating cytosine and guanine residues the cytosine ring at the 5′ position stacks closely to the corresponding cytosine ring of the opposite strand in the 3′ direction. This means that even-numbered oligomers with sequences of alternating cytosine and guanine residues which start from the 5′ end with cytosine will have one more such interstrand stacking interaction than those starting from the 5′ end with guanine. Thus, d(CpGpCpG) contains two such interstrand cytosine–cytosine stacking interactions while d(GpCpGpC) contains only one such stacking interaction if both tetramers are constrained to the Z-type conformation. Second, there is a strong interaction in the Z-type helix involving the oxygen of the furanose ring associated with the cytosine in the 5′ direction and the corresponding guanine ring in the 3′ direction. It is reasonable to suppose that it is this furanose oxygen–guanine ring interaction that is responsible for the remarkable changes in the frequency of the ring vibrations of guanine in Z-DNA. Examination of models of the Z-DNA structure shows that this interaction would occur with all four guanine residues in the tetramer d(CpGpCpG) but only with two or three guanine residues in the d(GpCpGpC) sequence. Finally, in the Z-type DNA there is an interaction along the same chain: proceeding from the 5′ to the 3′ direction the guanine ring on each chain stacks with the succeeding cytosine ring on the same strand. However, there are only two of these interactions (one in each strand) in the Z form of the tetramer d(CpGpCpG) while there are four such interactions in a hypothetical Z form of d(GpCpGpC). Thus, it may be that the third interaction tends to destabilize the Z form.

As a result of these three distinct interactions one can distinguish between short alternating purine–pyrimidine and pyrimidine–purine sequences containing guanine and cytosine on the basis of the differing number of each of the above three types of interactions when each sequence is constrained to adopt a Z form. It is at this time difficult to evaluate quantitatively the relative energetic importance of these types of interactions, but it seems plausible that the difference in the number of these interactions which can be made in each of the tetramer sequences exerts either a stabilizing or a destabilizing influence upon the Z form for the purine–pyrimidine sequence. Enhanced stabilization of the B form in combination with destabilization of the left-handed Z form may be responsible for the failure of three of the tetramers studied in this work to adopt a Z-like structure. It seems reasonable to expect that these factors will be more prominent in oligomers of short length or in short segments of defined alternating sequence along a normal DNA chain. As a consequence one would expect the differences between alternating pyrimidine–purine and purine–pyrimidine sequences to diminish in sequences of longer length.

The tetramer d(CpCpGpG) appears to be the most flexible or the least ordered under all conditions. Thus, we see that there is a difference in the stability and conformation for each of these four tetramers. It is hoped that these experimental data will be of help in providing useful information to those workers involved in developing force fields for nucleic acids.

Acknowledgments

We acknowledge the helpful discussions with Lawrence McIntosh of the Institute of Molecular Biology, University of Oregon.

References

Behe, M., & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1619–1623.

Brown, E. B., & Peticolas, W. L. (1975) *Biopolymers 14*, 1259–1271.

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

Chou, C. H., & Thomas, G. J., Jr. (1977) *Biopolymers 16*, 765–789.

Conner, B. N., Takano, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1982) *Nature* (*London*) *295*, 294–299.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 152*, 723–736.

Drew, H., Tankano, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature* (*London*) *286*, 567–573.

Erfurth, S. C., & Peticolas, W. L. (1975) *Biopolymers 14*, 247–264.

Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938–941.

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

Goodwin, D. D., & Brahms, J. (1978) *Nucleic Acids Res. 5*, 835–850.

Klysik, J., Stirdivant, S. M., Larson, J. E., Hart, P. A., & Wells, R. D. (1981) *Nature* (*London*) *290*, 672–677.

Maxam, H. M., & Gilbert, W. (1979) *Methods Enzymol. 65*, 449–559.

Nischimura, Y., Tsuboi, M., Nakano, T., Higuchi, S., Sato, T., Shida, T., Uesugi, S., & Ikehara, M. (1982) *Nucleic Acids Symp. Ser. 11*, 169–172.

Quadrifoglio, F., Manzini, G., Vasser, M., Dinkespield, K., & Crea, R. (1981) *Nucleic Acids Res. 9*, 2195–2206.

Strommen, D. P., & Peticolas, W. L. (1982) *Biopolymers 21*, 969–978.

Thamann, T. J., Lord, R. C., Wang, A. H. T., & Rich, A. (1981) *Nucleic Acids Res. 9*, 5443–5457.

Biochemistry **1984**, *23*, 3207–3217

Thomas, G. A., & Peticolas, W. L. (1983a) *J. Am. Chem. Soc. 105*, 986–992.

Thomas, G. A., & Peticolas, W. L. (1983b) *J. Am. Chem. Soc. 105*, 993–996.

Thomas, G. J., Jr., & Hartman, K. A. (1973) *Biochim. Biophys. Acta 312*, 311.

Viswamitra, M. A., Kennard, O., Shakked, Z., Jones, P. G., Sheldrick, B. M., Salisburg, S., & Pauello, L. (1978) *Nature* (*London*) *273*, 687–690.

Wang, A. H. J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., & Rich, A. (1979) *Nature* (*London*) *282*, 680–686.

Wang, A. H. J., Quigley, G. J., Kolpak, F. J., van der Marel, G., van Boom, J. H., & Rich, A. (1981) *Science* (*Washington, D.C.*) *211*, 171–176.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature* (*London*) *287*, 755–758.

Deoxyguanosine–Deoxyadenosine Pairing in the d(C-G-A-G-A-A-T-T-C-G-C-G) Duplex: Conformation and Dynamics at and Adjacent to the dG·dA Mismatch Site[†]

Dinshaw J. Patel,* Sharon A. Kozlowski, Satoshi Ikuta, and Keiichi Itakura

ABSTRACT: Nuclear magnetic resonance (NMR) has been used to monitor the conformation and dynamics of the d(C_1-G_2-A_3-G_4-A_5-A_6-T_6-T_5-C_4-G_3-C_2-G_1) self-complementary dodecanucleotide (henceforth called 12-mer GA) that contains a dG·dA purine–purine mismatch at position 3 in the sequence. These results are compared with the corresponding d(C-G-C-G-A-A-T-T-C-G-C-G) dodecamer duplex (henceforth called 12-mer) containing standard Watson–Crick base pairs at position 3 [Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 428–436]. The dG·dA interaction at position 3 was monitored at the guanosine exchangeable H-1 and nonexchangeable H-8 protons and the nonexchangeable adenosine H-2 proton. We demonstrate base-pair formation between anti orientations of the guanosine and adenosine rings on the basis of nuclear Overhauser effects (NOE) observed between the H-2 proton of adenosine 3 and the imino protons of guanosine 3 (intra base pair) and guanosines 2 and 4 (inter base pair). The dG(anti)·dA(anti) pairing should result in hydrogen-bond formation between the guanosine imino H-1 and carbonyl O-6 groups and the adenosine N-1 and NH_2-6 groups, respectively. The base pairing on either side of the dG·dA pair remains intact at low temperature, but these dG·dC pairs at positions 2 and 4 are kinetically destabilized in the 12-mer GA compared to the 12-mer duplex. We have estimated the hydrogen exchange kinetics at positions 4–6 from saturation-recovery measurements on the imino protons of the 12-mer GA duplex between 5 and 40 °C. The measured activation energies for imino proton exchange in the 12-mer GA are larger by a factor of ~2 compared to the corresponding values in the 12-mer duplex. This implies that hydrogen exchange in the 12-mer GA duplex results from a cooperative transition involving exchange of several base pairs as was previously reported for the 12-mer containing a G·T wobble pair at position 3 [Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I., Jr. (1982) *Biochemistry 21*, 6567–6574]. We have assigned the nonexchangeable base protons by intra and inter base pair NOE experiments and monitored these assigned markers through the 12-mer GA duplex to strand transition. We demonstrate that replacing two dG·dC by two dG·dA base pairs reduces the melting temperature of the dodecanucleotide by ~17 °C. The phosphorus spectrum of the 12-mer GA differs significantly from that of the 12-mer, indicative of changes in the phosphodiester backbone in order to accommodate the dG(anti)·dA(anti) base-pair formation in the interior of a DNA duplex.

There has been a great deal of recent experimental (Neidle & Berman, 1983; Patel et al., 1982a) and theoretical (Olson, 1982; Tilton et al., 1983; Olson et al., 1983; Levitt, 1983; Tidor et al., 1983) activity aimed at elucidating the conformational features of DNA in the crystalline (Dickerson et al., 1983; Viswamitra, 1983; Wang et al., 1983), fiber (Arnott et al., 1983; Zimmerman & Pfeiffer, 1983), and solution (Pardi et al., 1981; Patel et al., 1983a; Feigon et al., 1983) states. Much less is known about the role of non-Watson–Crick pairing and the consequences base-pair mismatches at and adjacent to the modification site. (Lomant & Fresco, 1975; Topal & Fresco, 1976a,b; Rein et al., 1983; Haasnoot et al., 1979, 1980; Wallace et al., 1979).

The most definitive investigations have focused on the G·U interaction in transfer RNA (Schimmel & Redfield, 1980; Hare & Reid, 1982) and the dG·dT interaction in DNA duplexes (Early et al., 1978; Patel et al., 1982c). This purine–pyrimidine mismatch involves two imino proton–carbonyl group hydrogen bonds on the basis of a nuclear Overhauser effect (NOE) between the guanosine and thymidine protons in the wobble pair.

There are few reports on purine–purine mismatches in nucleic acid duplexes. One example has been observed for yeast phenylalanine transfer RNA crystals, which contain a

---

[†] From AT&T Bell Laboratories, Inc., Murray Hill, New Jersey 07974 (D.J.P. and S.A.K.), and the Molecular Genetics Department, City of Hope National Medical Center, Duarte, California 91010 (S.I. and K.I.). *Received November 7, 1983*. The high-field NMR experiments were performed at the NMR facility for biomedial research located at the Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology. The NMR facility is supported by Grant RR0095 from the Division of Research Resources (National Institutes of Health) and by the National Science Foundation under Contract C-670. The contributions of S.I. and K.I. were funded by NIH Grant GM 28651.

0006-2960/84/0423-3207$01.50/0