Krauhs, E., Little, M., Kempf, T., Hofer-Warbinek, R., Ade, W., & Ponstingl, H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4156–4160.
Lambeir, A., & Engelborghs, Y. (1981) *J. Biol. Chem. 256*, 3279–3282.
Lin, C. M., Singh, S. B., Chu, P. S., Dempcy, R. O., Schmidt, J. M., Pettit, G. R., & Hamel, E. (1988) *Mol. Pharmacol. 34*, 200–208.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
Luduena, R. F., & Roach, M. C. (1981) *Biochemistry 20*, 4444–4450.
McClure, W. O., & Paulson, J. C. (1977) *Mol. Pharmacol. 13*, 560–575.
Penefsky, H. S. (1977) *J. Biol. Chem. 252*, 2891–2899.
Pettit, G. R., & Singh, S. B. (1987) *Can. J. Chem. 65*, 2390–2396.
Pettit, G. R., Cragg, G. M., Herald, D. B., Schmidt, J. M., & Lohavanijaya, P. (1982) *Can. J. Chem. 60*, 1374–1376.
Pettit, G. R., Singh, S. B., Niven, M. L., Hamel, E., & Schmidt, J. M. (1987) *J. Nat. Prod. 50*, 119–131.
Pettit, G. R., Singh, S. B., Schmidt, J. M., Niven, M. L., Hamel, E., & Lin, C. M. (1988) *J. Nat. Prod. 51*, 517–527.
Pettit, G. R., Singh, S. B., Hamel, E., Lin, C. M., Alberts, D. S., & Garcia-Kendall, D. (1989) *Experientia 45*, 209–211.
Ponstingl, H., Krauhs, E., Little, M., & Kempf, T. (1981) *Proc. Natl. Acad. Sci. U.S.A. 28*, 2757–2761.
Ray, K., Bhattacharyya, B., & Biswas, B. B. (1981) *J. Biol. Chem. 256*, 6241–6244.
Roach, M. C., & Luduena, R. F. (1984) *J. Biol. Chem. 259*, 12063–12071.
Roach, M. C., Bane, S., & Luduena, R. F. (1985) *J. Biol. Chem. 260*, 3015–3023.
Roach, M. C., Trcka, P. P., Jurd, L., & Luduena, R. F. (1987) *Mol. Pharmacol. 32*, 432–436.
Rosner, M., Hsu, F.-L., & Brossi, A. (1981) *J. Org. Chem. 46*, 3686–3688.
Schmitt, H., & Atlas, D. (1976) *J. Mol. Biol. 102*, 743–758.
Serrano, L., Avila, J., & Maccioni, R. B. (1984) *J. Biol. Chem. 259*, 6607–6611.
Sheir-Neiss, G., Lai, M. H., & Morris, N. R. (1978) *Cell 15*, 639–647.
Sherline, P., Leung, J. T., & Kipnis, D. M. (1975) *J. Biol. Chem. 250*, 5481–5486.
Taafrout, M., Rouessac, F., & Robin, J.-P. (1983) *Tetrahedron Lett. 24*, 197–200.
Williams, R. F., Mumford, C. L., Williams, G. A., Floyd, L. J., Aivaliotis, M. J., Martinez, R. A., Robinson, A. K., & Barnes, L. D. (1985) *J. Biol. Chem. 260*, 13794–13802.
Yeh, H. J. C., Chrzanowska, M., & Brossi, A. (1988) *FEBS Lett. 229*, 82–86.

The Length of a Junction between the B and Z Conformations in DNA Is Three Base Pairs or Less†

Zhongyi Dai, Gerald A. Thomas, Elisabeth Evertsz, and Warner L. Peticolas*
Department of Chemistry and Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403
Received February 21, 1989; Revised Manuscript Received May 4, 1989

ABSTRACT: Recently it has been suggested that double-helical complexes formed between the DNA sequences $(CG)_n(A)_m$ and their conjugates, $(T)_m(CG)_n$, would be candidates for the formation of a B–Z junction in aqueous solution at high salt concentrations [Peticolas et al. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2579–2583]. The junction was predicted to occur between a B-type helix in the $d(A)_m \cdot d(T)_m$ section and a Z-type helix in the self-complementary $(CG)_n \cdot (CG)_n$ sequence. In this paper we report Raman experiments on the deoxyoligonucleotides d(CGCGCGCGCGAAAAA) and d(CGCGCGAAAAA) and their complements. It is found the latter compound cannot be induced into the Z form in saturated salt solution but that the former sequence goes into a B–Z junction at 5.5 M salt. From a comparison of the relative intensity of the Raman conformational marker bands for B and Z DNA for both the A–T and C–G base pairs, it is shown that in 5.5 M NaCl solution none of the A–T base pairs are in the Z form, but nine of the C–G base pairs are in the Z form. The remaining three C–G base pairs are either in the junction or in the B form. Thus, the junction is formed from three or less C–G base pairs. If the solution is made 95 μM with $NiCl_2$, then the entire duplex goes into the Z form and the Raman bands of the adenine are completely changed into those of the Z form. A similar three base pair B–Z junction has recently been reported by Sheardy and Winkle [(1989) *Biochemistry 28*, 720–725], who used different base sequences for both the B and Z tracts. This indicates that the short B–Z junction length may be independent of the base sequence.

DNA–protein interaction appears to be governed by the sequence of hydrogen-donating and -accepting groups that occur along the major and minor grooves of the DNA (Seeman et al., 1976; Berg & von Hippel, 1987). When DNA is in its canonical B form, the distances between the hydrogen-bond-donating and -accepting portions of the helix can be obtained from the coordinates of the atoms in the base pairs along the chain. It is becoming increasingly evident that DNA exhibits polymorphism and can possibly go into forms other than the B form. Since it is unlikely that much cellular DNA is in a form different from the canonical B form, it is necessary that unusual conformations such as the Z form begin and end with a junction. Even the B form may be subject to change de-

† This work was supported by grants from the National Institutes of Health (GM15547 and GM33825).

0006-2960/89/0428-6991$01.50/0

pending on environmental conditions (Thomas et al., 1987, 1989; Benevides et al., 1988). As we have pointed out previously (Peticolas et al., 1988), even if the DNA does not go completely into a Z or an A form, the ability of a sequence to go easily into a non-B form means that interaction with a DNA binding protein could cause this sequence to be deformed along the B to A or B to Z conformational coordinate pathway. Such a deformation could lead to a better fit between the hydrogen-bond-donating and -accepting parts of the DNA and the protein. Recent Raman and X-ray crystallographic evidence exists for induced structural changes in the DNA upon binding of restrictions enzymes (McClarin et al., 1986; Thomas et al., 1989). For this reason it is of interest to learn what sequences can be induced into which conformations. Recently we have suggested that certain repeated base code words permit the DNA to go into an A or a Z form under mildly dehydrating conditions such as concentrated salt or salt/alcohol solutions (Peticolas et al., 1988). It was also suggested that a B–Z junction would most easily form if sequences were made up of a series of Z directing code words followed by a sequence of B directing code words or vice versa (Peticolas et al., 1988). In an effort to test these suggestions and to determine on upper limit for the minimum distance over which a B–Z junction could form, we have synthesized two oligomers that should in principle be capable of giving a Z–B junction. We have found that one of these sequences does give a junction in aqueous solution at salt concentrations of 3 M or higher. We estimate that the junction occurs over three or less Watson–Crick base pairs. This conclusion of a three-base-pair B–Z junction was recently reached in a completely different manner by Sheardy and Winkle (1989), who studied the conformation of the duplex formed by the hexadecanucleotide $d(C^*G)_4$-$(GACT)_2$ (C^* is 5-methyldeoxycytidine) and its conjugate. These authors used NMR and CD techniques. Our work provides an independent confirmation of their estimate by a different technique on a different sequence. It may be significant that the two different sequences give the same junction length.

It has been reported that a combination of 5.5 M NaCl and 95 μM $NiCl_2$ can induce alternating AT sequences into the Z form (Ridoux et al., 1988). It was of interest to see if $NiCl_2$ could change the conformation of the B–Z junction to a completely Z form. When $NiCl_2$ was added to the concentrated salt solution of the 17-mer, a striking change in the Raman spectra was obtained. The new bands indicate that essentially all of the $d(A)_5(T)_5$ portion of the duplex has undergone a B to Z transformation. It appears that $NiCl_2$ salt is able to change a string of A's and T's into the Z form. The Z form for this sequence has previously been reported in plasmids by McLean et al. (1986). It appears that rather stringent environment conditions are required to induce sequences of A and T into the non-B form.

Materials and Methods

The four oligonucleotides $d[(CG)_3(A)_5]$, $d[(T)_5(CG)_3]$, $d[(CG)_6(A)_5]$, and $d[(T)_5(CG)_6]$ were synthesized by the University of Oregon Biotechnology Laboratory utilizing an Applied Biosystems Model 380 DNA synthesizer that uses phosphoramidite chemistry. The oligodeoxynucleotides following trityl group removal were purified by reverse-phase HPLC on a Vydac C4 10 mm × 250 mm column. The oligomers were eluted with an acetonitrile and 0.1 M triethylamine acetate (pH 6.5) gradient. Purified fractions were concentrated to dryness by rotary evaporation. The samples were dissolved in double-distilled water and lyophilized. The solution samples of each of the oligomers were prepared by dissolving a portion of the sample in 0.5 M NaCl solutions at an approximate DNA concentration of 0.4 OD/μL, which corresponds to about 25 μg/μL. The sample were sometimes very slow in dissolving. To speed up the solvation, the pH was raised to 10 by adding 0.1 NaOH. After solvation was complete, 0.1 HCl was added to bring the pH back to 7. To obtain a fully complementary duplex structure and avoid the presence of single strands, the solutions of the complementary strands at pH 7 and 0.5 M NaCl were mixed in exactly equal concentration. These concentrations were determined from the height of the 680-cm^1 band of guanine and the 1090-cm^{-1} phosphate band relative to the height of the water band at 1640 cm^{-1}. The 680-cm^{-1} band intensity is directly proportional to the concentration of the guanine bases, and the height of the 1090-cm^{-1} band is directly proportional to the concentration of the phosphate groups. Thus, these two Raman intensities form a redundant set of concentration measurements. The two solutions—one of each component—were mixed in a volume ratio that gave an equal concentration for each single chain as determined by the concentration of both the guanine and phosphate residues. This procedure gives an exact 1:1 mixture of the two complementary strands within the ability of the Raman effect to see the difference. This may be seen in Raman spectra of the mixtures of the complementary 11-mers in 5.5 M salt. If either of the two components were in excess, then it would form an immediate Z conformation with the CG portions of the chain. That would lead to the appearance of a band at 625 cm^{-1} in the high-salt spectrum of the 11-mer duplex. In Figure 5 no such band at 625 cm^{-1} is present. Consequently, one can see how well this method of assuring an exact 1:1 paring is. To facilitate the formation of the complete duplex structure, the mixture was heated in warm (60 °C) water for several minutes and allowed to cool slowly. This was necessary to melt out the C–G pairs formed in the solutions of the individual components. Raman spectra were taken as salt concentrations between 0.5 M and saturation. To convert the entire 17-mer to the Z form, the solution of the junction in 5.5 M NaCl was made 95 μM in $NiCl_2$ following the procedure of Ridoux et al. (1988). The Raman equipment has been described elsewhere (Wang et al., 1987a,b). The CD spectra were taken with a Jasco 600 spectrometer that is in the laboratory of Professor John Schellman.

Results and Discussion

We started our studies with the oligomer d-(CGCGCGAAAAA), which we will call the Watson chain, and its Crick complement d(TTTTTCGCGCG). The sequence of five T's (or A's) was chosen because five bases are required for half the turn of a B helix, while the three pairs of C–G are half the turn of a Z helix. It was hoped that the six C–G Watson–Crick base pairs on one end would go into the Z form at high salt but that the subsequent five A–T Watson–Crick pairs would remain in the B form on the other end. When it was found that the fully complementary double helix $d[(CG)_3(A)_5]$·$d[(T)_5(CG)_3]$ failed to leave the B form under saturated salt conditions, a pair of oligomers with a longer CG sequence were synthesized, $d[(CG)_6A_5]$ and d-$[T_5(CG)_6]$. In this case the $(CG)_6$ is capable of forming an entire turn of a Z helix. Again we call $d[(CG)_6(A)_5]$ the Watson strand and $d[(T)_5(CG)_6]$ the Crick strand.

Before studying the heteroduplexes formed by each pair of the fully complementary chains, we considered it prudent to examine the behavior of each of these chains individually in aqueous solution. Figure 1 shows the Raman spectra of a pH 7 aqueous solution of the Watson chain of the 11-mer under

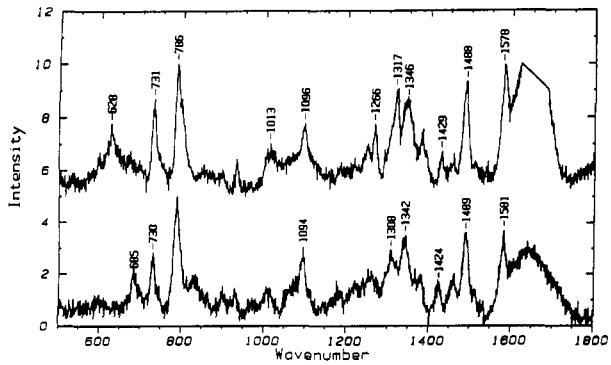

FIGURE 1: Raman spectra of a pH 7 aqueous solution of the Watson 11-mer, d(CGCGCGAAAAA), at 0.5 (bottom) and 5.5 M (top) salt solution.

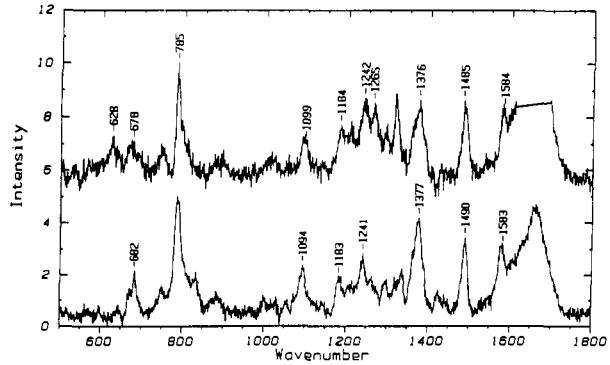

FIGURE 2: Raman spectra of the Crick 11-mer, d(TTTTTCGCGCG), at 0.5 (bottom) and 5.5 M (top) salt solution.

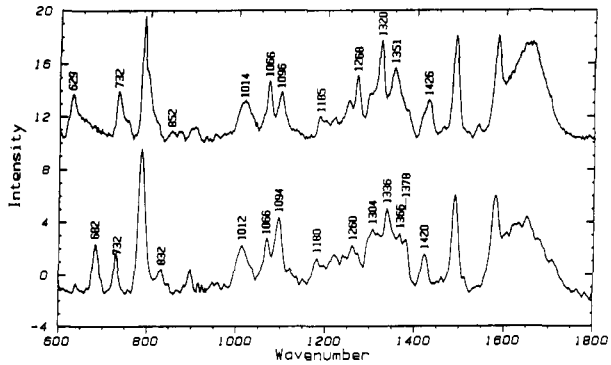

FIGURE 3: Raman spectra of the Watson 17-mer, d(CGCGCGCGCGCGAAAAA), at 0.5 (bottom) and 5.5 M (top) salt solution.

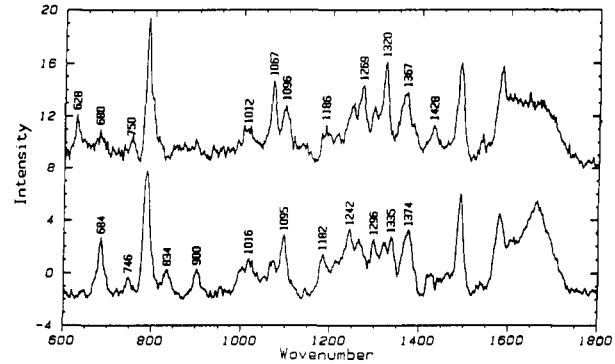

FIGURE 4: Raman spectra of the Crick 17-mer, d(TTTTTCGCGCGCGCGCG), at 0.5 (bottom) and 5.5 M (top) salt solution.

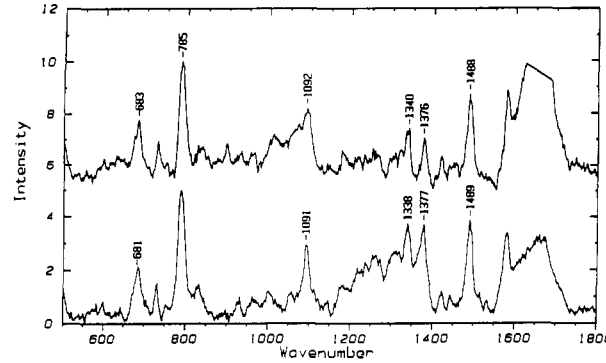

FIGURE 5: Raman spectra of the Watson–Crick duplex formed from a one to one mixture of the 11-mers whose spectra are shown in Figures 1 and 2 at 0.5 (bottom) and 5.5 M (top) salt solution.

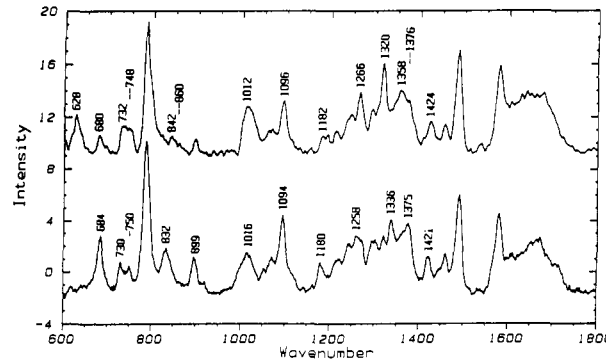

FIGURE 6: Raman spectra of the Watson–Crick duplex formed from a one to one mixture of the 17-mers whose spectra are shown in Figures 3 and 4 at 0.5 (bottom) and 5.5 M (top) salt solution.

conditions of low (bottom spectrum) and high (top spectrum) salt. Similarly, Figure 2 shows Raman spectra of solutions of the Crick chain under similar conditions. Figures 3 and 4 show the Raman spectra of the Watson and Crick chains separately for the 17-mers under identical conditions. In all four of the bottom spectra of Figures 1–4 the presence of a band at 835 cm^{-1} shows that some B-form secondary structure is present (Erfurth et al., 1972, 1975). We ascribe this secondary structure to dimerization of each of these oligomers in aqueous solution leading to a double-helical section containing the $(CG)_n$ portions of the individual chains. Each of the four bottom spectra in Figures 1–4 appears to be of junctions between the B-form double-helical $(CG)_n$ sequences and the single strands of either A or T that remain at either end. When the salt concentration is increased to saturation, all four of the resulting top spectra in Figures 1–4 show evidence of Z-form double-helix formation. This is shown by the appearance of the characteristic bands at 625 and 1320 cm^{-1} that have been shown to be marker bands for the Z form (Thamann et al., 1981). This change into the Z form is to be expected since the only secondary structure appears to be the $d[(CG)_n]\cdot d[(CG)_n]$ double helix formed between the self-complementary sections of each strand. Both the B and the Z form must involve a junction between a double-helical and a single-stranded form that looks something like

```
3'-TTTTTCGCGCG-5'                    3'- CGCGCGCGCGCGAAAAA-5'
   5'-GCGCGCTTTTT-3'          5'- AAAAAGCGCGCGCGCGC-3'
```

where, to illustrate these structures, we have shown the self-association of the Crick strand of the 11-mer and the Watson strand of the 17-mer.

To prepare a heteroduplex double-helical B form, solutions of the Watson and Crick chains for each of these oligomers were mixed under low-salt conditions. As described in the above section, the resulting mixture was heated to 60 °C to melt out the self-associated dimers. When the mixture was recooled to room temperature, a fully formed heteroduplex B structure was obtained. Salt was added to saturation in an

attempt to induce a partial or total Z conformation. The resulting Raman spectra for the 11-mer and the 17-mer are shown in Figures 5 and 6. Again, the bottom spectra are those obtained in low salt, and the top spectra are those obtained in high salt.

The spectra presented in Figure 5 show that when the Watson and Crick chains of the 11-mer are mixed one to one in low-salt solution, the homodimers formed from the individual strands are no longer present, but a fully heteroduplex Watson–Crick double helix is present. There are two ways that one can be certain that the self-association of the Watson and Crick chains has been disrupted and a fully heteroduplex double helix is obtained. The first is the change in the Raman spectra at low salt that shows an increase in the intensity of the B marker band at 835 cm^{-1}. The second is seen in the top spectrum in Figure 5 that shows the Raman spectrum obtained when the salt concentration is raised to 5.5 M NaCl. This Raman spectrum does not show any marker bands characteristic of the Z form. Since solutions containing either the Watson or Crick chains alone show the characteristic Z marker bands when salt is added to saturation, it is apparent that the dimers of the constituent Watson and Crick chains have been broken up and a fully heteroduplex Watson–Crick double helix is formed. The absence of the 625-cm^{-1} band shows that neither the Watson nor the Crick chain is present in any appreciable excess since the presence of an excess of either chain would result in the appearance of the Z marker band in the self-association of the excess single chain.

We interpret the failure of the heteroduplex double helix to show any Z conformation to mean that when the sequence of five A's in the Watson strand are hydrogen bonded to the five T's in the Crick strand, the resulting B helix is so energetically stable that the six C–G Watson–Crick base pairs can not go into a Z form under saturated salt conditions. (We did not try more rigorous methods of inducing a Z form such as adding alcohol or heavy metals.) It seems that the creation of a junction is not energetically favorable for this fully duplexed 11-mer. Consequently, it is of interest to see if a similar double helix with a longer sequence of alternating CG would help in the formation of a B–Z junction.

Figure 6 shows the Raman spectra of a solution of the combined Watson and Crick chains of the 17-mer that have been mixed at low salt and annealed as described above. These spectra show the presence of a B-form double helix at low salt (bottom spectrum) and the presence of considerable Z conformation in the CG portion of the double helix at high-salt solution (top spectrum). This may be seen from the bands at 625 and 1320 cm^{-1} that are present in the top spectrum. The longer CG sequence in the 17-mer appears to be responsible for its going into the Z form. Figure 7 shows a Raman spectrum of the 17-mer in high salt in the region 600–950 cm^{-1} that has been resolved into a minimum number of Lorenztian lines by using the computer program described elsewhere (Patapoff et al., 1988). This computer program automatically gives the area in each of the bands. In this way the relative intensity of the bands may be obtained exactly. In a careful study of the intensity of the Raman bands in going from the B to the Z form, Benevides and Thomas (1983) showed that the total intensity of the B marker band at 680 wavenumbers and of the Z marker band at 625 wavenumbers remained constant during the B to Z transition. This result has been repeatedly confirmed in our own laboratory in our studies on the B to Z transition in oligonucleotides (Wang et al., 1987a,b). As can be seen from Figure 7 the 680-cm^{-1} band is much weaker than the 625-cm^{-1} band. The exact intensities

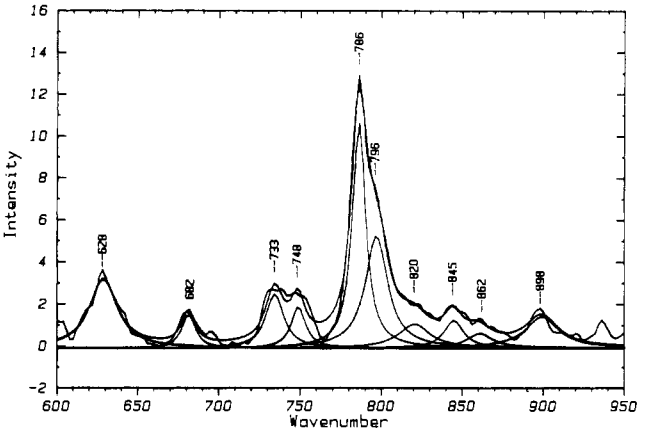

FIGURE 7: Raman spectrum of the Watson–Crick duplex shown in the top of Figure 6 after curve resolution into the minimum number of Lorentzians.

of the two bands normalized to 1.0 is given by the computer as 0.75 ± 0.03 for the 625-cm^{-1} band and 0.25 ± 0.03 for the 680-cm^{-1} band. From these measurements we can obtain the fact that three-fourths or nine of the C–G base pairs are in the Z form. The most reasonable assumption is that the remaining three C–G base pairs are in the junction. The guanine in these base pairs give rise to the band at 680 cm^{-1}. Although this band is a marker band for the B form when double-helical DNA is present, this band is also observed in single disordered DNA chains containing guanine. It is possible that the 680-cm^{-1} band contains a contribution due to a weak unresolved band of thymine that occurs in the region 665–670 cm^{-1}. To be completely accurate, we should subtract the intensity of the contribution of the thymine band from the 680-cm^{-1} band before taking the ratio. Unfortunately, we do not know what this intensity is. This subtraction would result in a reduction in the intensity of the 680-cm^{-1} band. This in turn would decrease the calculated number of base pairs in the junction to below three base pairs. Any error we may have made in ignoring the thymine band would cause us to overestimate the length of the junction. It seems unlikely that the thymine band is of any importance in this measurement, but it could result in our having overestimated the length of the B–Z junction. Consequently, it is concluded that three base pairs is the maximum number that are in the junction.

Although it is extremely unlikely, it is possible that there is an equilibrium between a set of B helices and a set of Z helices that would show the same Raman marker bands. In other words, how can we be certain that all of the A and T tracts are in the B form? All of the marker bands for the Z form of DNA that have been cited so far are due to Thamann et al. (1981) and only correspond to vibrations associated with either the G or the C bases. Recently it has been found possible to put poly[d(A-C)]·poly[d(G-T)] and poly[d(A-T)] into the Z form in aqueous solution through the use of high NaCl concentration with the addition of heavy metal ions (Ridoux et al., 1988). Also, it has been found that certain oligomers containing all four bases crystallize in the Z form (Benevides et al., 1984; Wang et al., 1987a,b). This has permitted the obtaining of marker bands that are due to adenine and thymine. For example, Figure 1 of Ridoux et al. (1988) shows that three bands at 727, 748, and 790 cm^{-1} are due to adenine, thymine, and thymine–phosphate, respectively, in the B form. These bands change dramatically in the Z form with the presence of a single strong band at 746 cm^{-1} that is due to a mixture of phosphate and thymine vibrations and a backbone band at 815 cm^{-1}. A comparison of the region

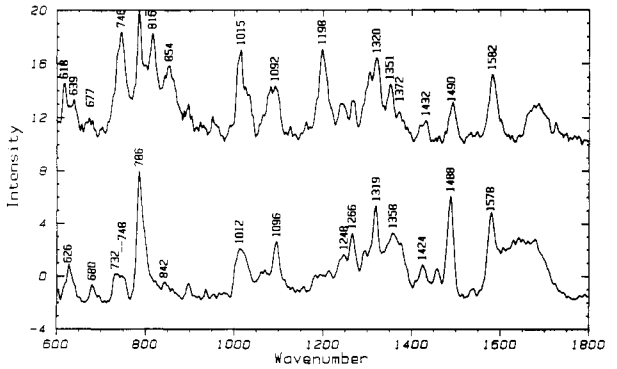

FIGURE 8: Comparison of the Raman spectra of the 17-mer Watson–Crick duplex in 5.5 M NaCl (bottom) and in 5.5 M NaCl and 95 μM $NiCl_2$ (top).

800–720 cm^{-1} of the top Raman spectrum of Figure 6 with Figure 1 of Ridoux et al. (1988) shows plainly that the three bands seen at 727, 748, and 790 cm^{-1} are representative of adenine, thymine, and thymine–phosphate in the B form. From these results we must conclude that all of the A and T tracts are in the B form. Thus, there cannot be an equilibrium between oligomers in the B form and those in the Z form but a true junction between the Z form in the CG tract and the B form in the A and T tracts with the junction occurring in the C–G base pairs adjacent to the A and T tracts.

This conclusion can be even further strengthened by causing the whole 17-mer to go into the Z form. To induce the A and T tracts of the 17-mer into the Z form, we employed the method used by Ridoux et al. (1988) to induce poly[d(A-T)] into the Z form. $NiCl_2$ was added to a solution of the B–Z junction of the 17-mer in 5.5 M NaCl so that the resulting solution was 95 μM in $NiCl_2$. This caused the A and T tracts to go into the Z form so that the whole 17-mer became a left-handed double helix. This may be seen by the rather dramatic changes in the Raman spectrum of the 17-mer upon the addition of $NiCl_2$ that is shown in the top spectrum of Figure 8. The intensity of the 746- and 816-cm^{-1} bands is increased by a substantial amount, which indicates the formation of a Z form in the AAAAA·TTTTT tract. These bands, due to adenine, appear to be almost identical with those reported by Ridoux et al. (1988) for poly[d(A-T)] that was induced into the Z form by using the same concentration of $NiCl_2$. The bottom spectrum of Figure 8 is that of the 17-mer in 5.5 M NaCl before the addition of $NiCl_2$ and shows a comparison between the A and T bands in the junction and in Z DNA. Although McLean et al. (1986) have previously reported that tracts of A and T in plasmid DNA will go into the Z form, there appears to be no previous report on the conversion of continuous tracts of A and T going into the Z form in a relaxed DNA.

Although the Raman evidence is definitive, it seemed worthwhile to look at the CD spectra of the 17-mer at low and high salt to obtain direct evidence of the presence of a left-handed double helix. Figure 9 shows the CD spectra of the 17-mer at high and low salt concentrations. It is evident that considerable change in the handedness of the helix has occurred in going from low to high salt. Comparison with reference spectra of the B and Z forms shows that the low-salt spectrum is characteristic of the B form and the high-salt spectrum is characteristic of a mixture of right-handed and left-handed forms. The CD spectra present evidence that is consistent with the presence of B–Z junctions in this oligomeric DNA duplex at high salt concentrations. We have not tried to quantitate the CD spectrum as this has already been done by Sheardy and Winckle (1989) for a similar sequence.

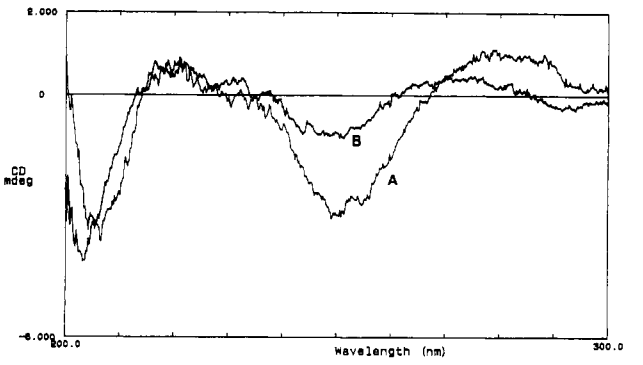

FIGURE 9: Circular dichroic spectrum of the double-helical oligomer d(CGCGCGCGCGCGAAAAA)·d(TTTTTCGCGCGCGCGCG) in aqueous low- and high-salt solutions. Spectrum A is in low salt, and spectrum B is in high salt.

In discussing the distance over which a B–Z junction could be formed, we must keep in mind that sequence may play a very important role in the junction length. It has been shown that almost any sequence can be forced into a Z form in a negatively supercoiled plasmid to relieve the tension in the chain induced by the supercoiling (Singleton et al., 1982; Johnston & Rich, 1985; McLean et al., 1986). Under these conditions it may be that the length of the B–Z junction is very long. This is a completely different situation from the oligomers that we have made that are specifically designed, by virtue of their base sequence, to go into a junction at high salt concentrations over a very short number of base pairs. The occurrence of such junctions in relaxed DNA oligonucleotides has been considered by Sheardy (1988), Peticolas et al. (1988), and Sheardy and Winkle (1989). The presence in a DNA of sections of several alternating pyrimidine–purine pairs followed by a sequence of A_n or T_n will have a conformational potential energy surface that is unique because of the ease with which the former will go into the Z form and the difficulty with which the latter will leave the B form. Such sequences could form virtual junctions that would remain in the B form but undergo deformations such as bending or twisting that could play a role in the protein–nucleic acid interactions. Finally, it is of great interest that the junction length of the B–Z junction obtained by Sheardy and Winkle (1989) using CD and NMR measurements is exactly the same (three base pairs) as our result obtained entirely from the relative intensities of the B and Z Raman marker bands of guanine. It is also of interest that they were able to obtain a junction in a 16-mer of which only 8 residues were the CG. We would interpret this as being due to the fact that A and T tracts have a stronger tendency to stay in the B form than do the d(ACTGACTG) tract used by Sheardy and Winkle for their B tract. The fact that such short junctions can in principle appear in DNA means that very short alternating forms of B and Z DNA could occur in native DNA if the environmental effects were such as to stabilize these conformational changes.

REFERENCES

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

Benevides, J. M., Want, A. H.-J., van der Marel, G. A., van Boom, J. H., Rich, A., & Thomas, G. J., Jr. (1984) *Nucleic Acids Res. 14*, 5913–5925.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1988) *Biochemistry 27*, 931.

Berg, O. G., & von Hippel, P. H. (1987) *J. Mol. Biol. 193*, 723–750.

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.
Johnston, P. D., & Rich, A. (1985) *Cell (Cambridge, Mass.) 42*, 713-724.
McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P. J., Boyer, H. W., Grable, H. W., & Rosenberg, J. M. (1986) *Science 243*, 1526.
McLean, M. J., Blaho, J. A., Kilpatrick, M. W., & Wells, R. D. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5884-5888.
Patapoff, T. W., Thomas, G. A., Wang, Y., & Peticolas, W. L. (1988) *Biopolymers 27*, 493-507.
Peticolas, W. L., Yang, W., & Thomas, G. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2579-2583.
Ridoux, J. P., Liquier, J., & Taillandier, E. (1988) *Biochemistry 27*, 3874-3878.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804-808.
Sheardy, R. D. (1988) *Nucleic Acids Res. 16*, 1153-1167.
Sheardy, R. D., & Winkle, S. A. (1989) *Biochemistry 28*, 720-725.
Singleton, C. K., Klysik, J., Stirdivant, S. M., & Wells, R. D. (1982) *J. Biol. Chem. 257*, 10159-10165.
Thamann, T. J., Lord, R. C., Wang, A. H.-J., & Rich, A. (1981) *Nucleic Acids Res. 9*, 5443-5457.
Thomas, G. A., Patapoff, T. W., Wang, Y., & Peticolas, W. L. (1987) in *Book of Abstracts, Fifth Conversation in Biomolecular Stereodynamics* (Sarma, R. H., Ed.) State University of New York at Albany, Albany, NY.
Thomas, G. A., Kubasek, W. L., Peticolas, W. L., Greene, P., Grable, J., & Rosenberg, J. M. (1989) *Biochemistry 28*, 2001-2009.
Wang, Y., Thomas, G. A., & Peticolas, W. L. (1987a) *Biochemistry 26*, 5178-5186.
Wang, Y., Thomas, G. A., & Peticolas, W. L. (1987b) *J. Biomol. Struct. Dyn. 5*, 249-274.

Three-Dimensional Solution Structure of a DNA Duplex Containing the *Bcl*I Restriction Sequence: Two-Dimensional NMR Studies, Distance Geometry Calculations, and Refinement by Back-Calculation of the NOESY Spectrum†

Kevin M. Banks,‡ Dennis R. Hare,§ and Brian R. Reid*,‖

Departments of Chemistry and of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195, and Hare Research, Woodinville, Washington 98072

Received August 1, 1988; Revised Manuscript Received May 10, 1989

ABSTRACT: A three-dimensional solution structure for the self-complementary dodecanucleotide [d-(GCCTGATCAGGC)]$_2$ has been determined by distance geometry with further refinements being performed after back-calculation of the NOESY spectrum. This DNA dodecamer contains the hexamer [d(TGATCA)]$_2$ recognized and cut by the restriction endonuclease *Bcl*I, and its structure was determined in hopes of obtaining a better understanding of the sequence-specific interactions which occur between proteins and DNA. Preliminary examination of the structure indicates the structure is underwound with respect to idealized B-form DNA though some of the local structural parameters (glycosyl torsion angle and pseudorotation angle) suggest a B-family type of structure is present. This research demonstrates the requirements (resonance assignments, interproton distance measurements, distance geometry calculations, and NOESY spectra back-calculation) to generate experimentally self-consistent solution structures for short DNA sequences.

The recognition of short DNA sequences by proteins (operator sequences by repressors, promoter sequences by RNA polymerase, and restriction sequences by endonucleases) is remarkably specific (Ohlendorf & Matthews, 1983). The affinities between these DNA sequences and their specific proteins are known to be regulated by a number of highly sensitive intermolecular forces, which include electrostatic interactions between the positively charged amino acids and the negatively charged phosphate backbone, sequence-specific intermolecular hydrogen bonding, and van der Waals interactions (Berg & Blomberg, 1978; Berg et al., 1981; Takeda et al., 1986; von Hippple & McGhee, 1972; Ohlendorf & Matthews, 1983). These facts suggest that the three-dimensional structure of the DNA, as well as that of the protein, must play a tremendously important role in the protein/DNA recognition phenomenon (Lomonossoff et al., 1981; Rhodes, 1982; Drew & Travers, 1984, 1985). Within the recent literature, various DNA sequences have been studied with many different spectroscopic techniques with the aim of determining their three-dimensional structures: the goal of these studies was to obtain a better understanding of the mechanisms that regulate DNA-protein interactions (Dickerson & Drew, 1981; Shakked et al., 1983). Additionally, the application of two-dimensional ^{1}H nuclear magnetic resonance spectroscopy (NMR) methods to biopolymers has proven to be a powerful technique, yielding an abundance of structural information when used carefully. In particular, the observation and interpretation of the nuclear Overhauser effect (NOE), a dipolar,

†Supported in part by Program Project Grant P01GM 32681 to B.R.R. and SBIR Grant GM 35620 to D.R.H.
‡Department of Chemistry, University of Washington.
§Hare Research.
‖Department of Chemistry and Biochemistry, University of Washington.

0006-2960/89/0428-6996$01.50/0