Non-contiguous regions of Z-DNA in a DNA dodecamer

Nhi Hua, Gijs A. van der Marel¹, Jacques H. van Boom¹ and Juli Feigon*

Department of Chemistry & Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90024-1569, USA and ¹Gorlaeus Laboratories, University of Leiden, 2300 RA Leiden, The Netherlands

Received April 25, 1989; Revised and Accepted August 28, 1989

ABSTRACT

The conformation of the self-complimentary DNA dodecamer d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) has been investigated in a variety of salt and solvent conditions by one and two-dimensional ¹H NMR. In low salt aqueous solutions, the molecule forms a regular B-DNA structure similar to the unmodified dodecamer. However, in aqueous solution containing high salt concentration and methanol, the dodecamer adopts a structure in which the br⁵CGbr⁵CG ends of the molecule are in a Z-DNA like conformation and the AATT region is neither standard B-DNA nor Z-DNA. The implications of these results for the structure of junctions between B and Z-DNA and the sequence specificity of Z-DNA are discussed.

INTRODUCTION

Z-DNA was first characterized as a structure adopted by DNA polymers composed of alternating dC and dG in high salt solutions, as assayed by spectroscopic changes in the CD spectrum¹. The crystal structures of d(CG)₃² and d(CG)₂³ showed that these spectroscopic changes did indeed correlate with a non-B-DNA structure, termed Z-DNA by Rich and co-workers², which was a left handed double helix with alternating *syn* G and *anti* C Watson-Crick base pairs. Since purines adopt the *syn* conformation more readily than pyrimidines, strict alternation of purines and pyrimidines was originally thought to be a requirement for Z-DNA formation. Z-DNA formation has since been demonstrated in some sequences which contain A·T as well as G·C base pairs and also in some sequences which are not strictly alternating purine and pyrimidine (reviewed in 4,5). However, Z-DNA does not seem to form in stretches of d(A-T)_n longer than a few base pairs⁶. Attempts have been made to predict the Z-DNA forming potential of various sequences⁷, and while the rules governing Z-DNA formation are not fully understood, it is clear that both A·T and out of purine-pyrimidine alternation A·T and G·C base pairs increase the energy required for the B-Z transition in a given DNA sequence. If two or more A·T base pairs, either in purine-pyrimidine alternation or out of alternation, are inserted between runs of alternating d(CG)_n, the entire sequence may adopt a Z-DNA structure or only the CG bases may be Z-DNA, depending on the conditions and the

sequence^{6,8,9,10}. Such structures have been analyzed by means of two-dimensional gel electrophoresis and nuclease digestion studies^{6,8}. If Z-DNA does not form through the entire sequence, then there is effectively a junction region between the Z-DNA and non-Z-DNA regions. However, to date there have been no X-ray crystallographic or NMR structural studies of Z-DNA oligonucleotides containing more than two consecutive A·T base pairs.

In this work, we have used one and two-dimensional ¹H NMR to obtain qualitative information on the structure of the modified DNA dodecamer d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) under conditions where all or part of the molecule might adopt the Z-DNA conformation. The cytosines were brominated at the 5 position in order to make it easier to form Z-DNA, since this modification has been shown to strongly favor Z-DNA formation², and it is difficult to effect the B-Z transition for short oligonucleotides in solution. Previous NMR studies on d(br⁵CGbr⁵CGATbr⁵CGbr⁵CG)¹¹ and crystallographic studies on d(br⁵CGATbr⁵CG)¹² have shown that these sequences containing two out-of-alternation A·T base pairs will adopt a regular Z-DNA structure under appropriate conditions, with the thymine bases in the *syn* conformation. The sequence studied here contains four A·T base pairs, the first and fourth of which are out of purine-pyrimidine alternation with the rest of the molecule. The numbering scheme used for the dodecamer and possible

Scheme 1

structures which this sequence might adopt are illustrated below (Scheme I). We have found that under appropriate conditions (high salt, addition of methanol) the dodecamer adopts a conformation in which the $\text{br}^5\text{CGbr}^5\text{CG}$ ends of the molecule are in a Z-DNA like conformation and the AATT middle is neither B nor Z-DNA (Scheme I b). Thus, the sequence AATT disrupts the formation of Z-DNA for this molecule and a junction is formed between the Z-DNA ends and the middle of the molecule. This gives us information both on the sequence requirements for Z-DNA and on the structure of junctions between B and Z-DNA.

MATERIALS AND METHODS

Sample preparation. The DNA dodecamer $\text{d}(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ was synthesized by an improved phosphotriester method and purified as previously described¹³. The sodium form of the dodecamer was obtained by passage of the DNA over a Bio-Rad AGW50-X4 cation exchange column. Low salt samples were prepared by dissolving the DNA in water and adding appropriate amounts of NaCl for the desired final concentration. The pH was adjusted by addition of HCl or NaOH. The sample was lyophilized and redissolved in D_2O twice and transferred to a 5mm NMR tube. The sample was redried in the NMR tube under a stream of $\text{N}_2(\text{g})$ and redissolved in 99.996% D_2O to a final volume of 400 μl and 2 mM DNA duplex. Samples in H_2O were prepared by redrying and redissolving the sample in 90% $\text{H}_2\text{O}/10\%$ D_2O . Higher salt/methanol samples were prepared by addition of appropriate amounts of NaCl and MgCl_2 , drying the sample, and redissolving in the desired final amount of D_2O or H_2O followed by addition of methanol- d_4 (99.5%, Aldrich).

NMR spectroscopy. All of the NMR experiments were carried out at 500 MHz on a General Electric GN500 spectrometer. Chemical shifts for the aqueous samples were determined from the chemical shift of H_2O which had been previously calibrated as a function of temperature and salt relative to DSS. Chemical shifts for the samples containing methanol were determined using the methanol methyl peak as a chemical shift reference at 3.30 ppm. One-dimensional spectra in H_2O were obtained using a 11 spin echo pulse sequence $(90_x - \tau - 90_x - \Delta - 90_x - 2\tau - 90_x - \Delta)^{14}$. The delay τ was adjusted so that the region of maximum excitation was centered between the aromatic and the imino resonances ($\tau = 50 \mu\text{sec}$).

NOESY (nuclear Overhauser effect spectroscopy) spectra in D_2O were obtained in the pure absorption mode following the method of States, et al.¹⁵ using the standard $90^\circ - t_1 - 90^\circ - \tau_m - 90^\circ$ -acquire pulse sequence¹⁶ with presaturation of the residual HDO peak during the 2 sec recycle delay. Spectra were acquired with 2048 complex points in t_2 and 228-456 t_1 values. The sweep width was 5000 Hz in both dimensions and 64, 128, or 192 scans were acquired per t_1 value. COSY (correlated spectroscopy) spectra were obtained with the standard $90^\circ - t_1 - 90^\circ$ -acquire pulse sequence¹⁷ except the second pulse was about 65° , and for the high salt/methanol sample the residual HOD and methanol- d_4 peaks were alternately irradiated for 500 msec during the 2 second recycle delay. Other acquisition

Figure 1. ^1H NMR spectra of $\text{d}(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in D_2O . (A) Sample in 50 mM NaCl, pH 6.3, 2 mM duplex in 400 μl at 37°C. (B) Same sample in 1 M NaCl, 11 mM MgCl_2 , 56% v/v methanol- $\text{d}_4/\text{D}_2\text{O}$ in 700 μl .

parameters were the same as for the NOESY. DQF-COSY (double-quantum filtered correlated spectroscopy) spectra were acquired as above with standard pulse sequences¹⁸ and phase cycling for hypercomplex spectra¹⁵ (spectra not shown).

One-dimensional spectra were processed with the GE NMR software (GEM16). Two-dimensional spectra were transferred via magnetic tape or ethernet to a VAX 8800 or microvax II computer and processed with the NMR processing program FTNMR (Hare Research). Processing parameters are indicated in the figure legends.

RESULTS

Spectra as a function of salt and solvent conditions. One dimensional spectra of $\text{d}(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in D_2O are shown in Figure 1. Figure 1a shows the spectrum observed for the sample in low salt buffer (low salt form). Under these conditions the molecule adopts a B-DNA conformation similar to that observed for the unmodified dodecamer $\text{d}(\text{CGCGAATTCGCG})$ ¹⁹ (discussed below). Dramatic changes are observed in aqueous solutions containing high salt concentrations and methanol, and these changes are enhanced by the addition of MgCl_2 (high salt/methanol form). Figure 1b shows the spectrum obtained when the salt and solvent are changed to 1 M NaCl, 11 mM MgCl_2 , 50% v/v methanol- $\text{d}_4/\text{D}_2\text{O}$. Under these conditions the molecule adopts a non-B-DNA conformation which will be discussed in detail below. The aromatic region of the spectra is

Figure 2. Aromatic region of the ^1H NMR spectra of $\text{d}(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in D_2O . (A) Expanded region from Figure 1a. (B) Same sample in .5 M NaCl, 50% v/v methanol- $\text{d}_4/\text{D}_2\text{O}$. (C) Expanded region from Figure 1b. Resonance assignments indicated on the figure are discussed in the text.

shown in Figure 2. Figure 2a is an expanded region of Figure 1a. Assignments of the aromatic resonances are indicated and were obtained by standard sequential assignment techniques as discussed below. The sample in Figure 2b contains .55 M NaCl and 50% methanol- d_4 . Under these conditions the sample contains a mixture of two different conformations, corresponding to the low salt B-DNA form and the high salt/methanol form, in slow exchange on the NMR time scale. Figure 2c is the aromatic region of the sample in Figure 1b. Assignments of the aromatic resonances for this high salt/methanol form are indicated. These were obtained primarily by exchange spectroscopy²⁰ using the sample in Figure 2b, as discussed below.

NOESY spectra of the low salt form. A NOESY spectrum of the dodecamer in low salt solution is given in Figure 3. The sample gives a pattern of crosspeaks typical of that normally observed for B-DNA^{19,21,22} and similar to that observed for the related unmodified dodecamer $\text{d}(\text{CGCGAATTCGCG})$ ¹⁹. In particular, strong intra and interbase H6,H8-H2', H2'' (box A) crosspeaks and weak H6,H8-H1' (box B) crosspeaks are observed, indicative of all bases being in

Figure 3. NOESY spectrum of the low salt form of d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) in D₂O at 37°C and $\tau_m = 300$ msec. Sample is the same as Figure 1a. Region of crosspeaks between the (a) aromatic and H2',H2'', (b) aromatic and H1' and (c) aromatic and thymine methyl are boxed. 441 t_1 values of 64 scans each were collected. The spectrum was zero-filled in t_1 and apodized with a squared sine-bell phase shifted by 90° in both dimensions.

Table 1: Chemical Shifts of the Non-Exchangeable Resonances of the Low Salt Form of d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG)[†]

BASE	H6, H8	AH2, CH ₃	H1'	H2'	H2''	H3'	H4'
C1	7.98		5.69	1.92	2.40	4.66	4.03
G2	8.00		6.00	2.72	2.87	4.97	4.34
C3	7.57		5.57	1.82	2.27	4.75	4.10
G4	7.84		5.49	2.65	2.74	4.98	4.30
A5	8.07	7.23	5.95	2.66	2.89	5.03	4.41
A6	8.08	7.59	6.13	2.55	2.89	4.98	4.43
T7	7.07	1.27	5.87	1.96	2.54	4.80	4.18
T8	7.41	1.55	6.10	2.22	2.60	4.90	4.19
C9	7.82		5.60	2.10	2.41	4.85	4.12
G10	7.93		5.94	2.68	2.81	4.99	4.35
C11	7.63		5.74	1.89	2.31	4.76	4.13
G12	7.90		6.13	2.59	2.37	4.64	4.13

[†]Sample conditions same as Figure 3.

the *anti* conformation expected for B-DNA. The observed interbase crosspeaks from the methyl resonances (T₈Me-T₇Me, T₇Me-A₆H₈) are indicative of a right-handed DNA conformation (box C). Sequential assignments were made via the base-H1' crosspeaks and confirmed by base-H2', H2'' connectivities in NOESY and H1'-H2', H2'' connectivities in NOESY and COSY (Figure 6a) spectra. Assignments of the AH2 resonances were made by imino-AH2 NOES observed in NOESY spectra in water (spectra not shown). These assignments, along with H3' and H4' assignments obtained from COSY connectivities (not shown), are summarized in Table 1.

NOESY spectra of the high salt/methanol form. A NOESY spectrum of the dodecamer in high salt/methanol solution is given in Figure 4. Under these conditions, a portion of the sample aggregates and precipitates. In addition, the probe sensitivity is significantly decreased due to the high salt. Therefore the spectra obtained have worse signal-to-noise than the spectra obtained on the sample in low salt. However, several additional NOESY spectra were obtained at the same and other mixing times, as well as under different conditions, so we are confident that our assignments of the weak crosspeak are correct. The pattern of crosspeaks is significantly different from that observed for B-DNA. In particular, some very strong crosspeaks are observed in the base-H1' region, indicative that some of the bases are in the *syn* conformation. In addition, the sequential interbase methyl crosspeaks are no longer observed. Since the structure of the molecule under these conditions was not known, and since sequential base-H1' connectivities are not observed for this conformation, it is not possible to assign the resonances using standard procedures. Assignments were therefore made by exchange spectroscopy from the previously assigned low salt B-DNA form²³.

Exchange spectroscopy for assignment of the high salt/methanol form. A portion of a NOESY spectrum showing the aromatic resonances and their crosspeaks for the sample containing a mixture of the low salt and high salt/methanol conformation of the

Figure 4. NOESY spectrum of the high salt/methanol form of $d(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in D_2O at 40°C and $\tau_m = 350$ msec. Sample is the same as Figure 1b. 228 t_1 values of 192 scans each were collected. The spectrum was zero-filled in t_1 and apodized with a squared sine-bell phase shifted by 70° in both dimensions.

Figure 5. Aromatic region of a NOESY spectrum of $d(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in D_2O at 40°C and $\tau_m = 250$ msec. Sample is the same as Figure 2b and contains a mixture of two conformations. Exchange crosspeaks connecting resonances from the two conformations of the dodecamer are indicated by solid lines. The dotted line indicates an NOE crosspeak. Assignments of the aromatic resonances are indicated in the spectrum at the top; boxed assignments correspond to the resonances from the high salt/methanol conformation. 321 t_1 values of 64 scans each were collected. The spectrum was zero-filled in t_1 to 1K points and apodized with a squared sine-bell phase shifted by 50° in both dimensions.

dodecamer is shown in Figure 5. Most of the crosspeaks observed in this region of the spectrum under these conditions arise from exchange between the two conformations during the 250 msec mixing time. Assignments of the resonances arising from the low salt (B-DNA) conformation in this sample were obtained by extrapolation from the assignments obtained on the sample in low salt (Figure 3). Some chemical shift differences are observed for the B-DNA conformation of the sample in low salt and the sample under these higher salt methanol conditions. This may be due to solvent effects, a slight change in conformation, and/or that the exchange rate is not quite slow (on the NMR time scale) under these conditions. In spite of these changes in chemical shift, we were still able to obtain the B-DNA assignments in this sample unambiguously, since sequential base-H1' NOEs are also observed in this spectrum (not shown). Once the assignments for the B-DNA conformation were confirmed in this sample, the exchange crosspeaks among the aromatic protons were used to assign the aromatic resonances of the high salt/methanol form of the dodecamer. The exchange crosspeaks between the two conformations are connected by solid lines in Figure 5. The methyl resonances were assigned based on their crosspeaks with their TH6s, and these assignments were confirmed by observation of exchange crosspeaks with the B-form. The H1' resonances in the high salt/methanol form were assigned as above, by observation of exchange crosspeaks with the previously assigned B-DNA form in the H1' region (not shown). Once the H1' resonances were assigned, the H2' and H2'' resonances were assigned based on NOE crosspeaks to the H1' which are also observed in this spectrum (not shown). The H2' and H2'' resonances were somewhat arbitrarily distinguished from each other on the basis of chemical shift, with the lower field resonance assigned

Table 2: Chemical Shifts of the Non-Exchangeable Resonances of the High Salt/Methanol Form of d(br⁵CGBr⁵CGAATTbr⁵CGBr⁵CG)[†]

BASE	H6,H8	AH2, CH ₃	H1'	H2' ^a	H2'' ^a
br ⁵ C1	7.827		5.832	1.798	2.718
G2	7.793		6.208	2.727	2.828
br ⁵ C3	7.576		5.872	1.798	2.718
G4	7.754		6.177	2.046	2.506
A5	8.216	8.07	5.819	1.683	2.621
A6	8.216	8.06	6.217	2.550	2.704
T7	7.371	1.57	6.018	2.078	2.249
T8	7.475	1.72	6.039	2.161	2.291
br ⁵ C9	7.708		5.824	1.606	2.432
G10	7.810		6.219	2.465	2.745
br ⁵ C11	7.611		6.031	2.267	2.441
G12	7.840		6.263	2.406	3.187

[†]Assignments of H6, H8, H2, TCH₃, and H1' resonances identified by exchange spectroscopy with low salt form (Figure 5).

^aAssignments of H2' and H2'' resonances identified from NOESY crosspeaks with H1' proton resonances (Figure 4). See text.

to the H2'' resonance for all but the terminal G. This decision was based on observed chemical shifts for H2' and H2'' proton resonances in B-DNA¹⁹ and on published assignments for the Z-DNA form of d(br⁵CG)₃²⁴. The assignments obtained for the high salt/methanol structure are summarized in Table 2.

COSY spectra of the two conformations. Portions of the COSY spectra of the dodecamer in low salt and high salt/methanol showing the region of crosspeaks between the H1' and H2',H2'' resonances are shown in Figure 6. In low salt, crosspeaks are observed between each H1' and its scalar coupled H2' and H2''; these crosspeaks are used to confirm sugar assignments obtained from NOESY spectra. The crosspeaks in this region for the sample in high salt/methanol are less well resolved, and several crosspeaks are either missing or weak. In particular, all of the CH1'-CH2'' crosspeaks are weak or unobserved and all of the

Figure 6. Portion of COSY spectra of d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) in D₂O at 40°C showing the region of crosspeaks between the H1' and H2',H2'' resonances. (A) Low salt sample (same as Figure 1a). (B) High salt/methanol sample (same as Figure 1b). Assignments are indicated on the sides of the spectra. For (A) 456 t₁ values were collected and for (B) 300 t₁ values were collected. The spectra were zero-filled in t₁ to 1K points and apodized with a squared sine-bell in both dimensions.

GH1'-GH2' crosspeaks with the possible exception of G₂ are unobserved or weak. In contrast, strong NOESY crosspeaks between these protons are observed for all but G₄. Thus the low intensity of these crosspeaks indicates a small coupling constant for the relevant protons. These differences in the H1'-H2' and H2'-H2" coupling patterns for the G and C sugars indicate that their sugar conformations are not the same (see discussion).

Base-H1' NOESY stacked plots. Stacked plots of NOESY spectra of the base-H1' crosspeak region of the dodecamer under two different solvent conditions are shown in Figure 7. Figure 7a shows the regular pattern of weak inter and intrabase H8,H6-H1' crosspeaks normally observed for the *anti* conformation of bases in B-DNA. This region also normally contains strong CH5-CH6 crosspeaks, which are absent here due to bromination of the Cs at the 5 position. In marked contrast to the crosspeak pattern observed for B-DNA are the crosspeaks observed for the sample in high salt/methanol (Figure 7b). The sample is in somewhat different conditions (.5 M NaCl, 40 mM MgCl₂, and 40% v/v methanol-d₄/D₂O) from the sample in Figure 1b, 2c, and 4, and under these conditions the guanine H8 resonances are somewhat better resolved. Four strong intrabase GH8-H1' crosspeaks are observed, indicating that all of the guanines are in the *syn* conformation observed in Z-DNA^{25,26}. The cytosines, adenines, and thymines all show only weak intranucleotide base-H1' crosspeaks in this spectrum. The T7H6-H1' crosspeak is particularly weak, and cannot be observed in this spectrum, but was observed in NOESY spectra taken with longer mixing times. The low intensity of these crosspeaks may be partially a result of conformational flexibility in the molecule under these conditions.

Imino proton spectra. Spectra of the imino proton resonances of the dodecamer in low salt and high salt/methanol at 40°C are shown in Figure 8. At this temperature, for the B-DNA conformation (Figure 8a), resonances are observed for all but the terminal G·C imino protons, with the three G·C iminos resonating between 12 and 12.8 ppm and the two A·T iminos between 13.5 and 13.8 ppm. Under high salt/methanol conditions (Figure 8b), the G·C imino resonances shift mostly to lower field. The spectrum of the G·C imino resonances is very similar to that observed for the related decamer d(br⁵CGbr⁵CGATbr⁵CGbr⁵CG) under conditions where the entire molecule adopts the Z-DNA conformation¹¹ (inset to Figure 8). Some residual B-DNA at this temperature contributes to the intensity of

Figure 7. Stacked plots of region of NOESY spectra containing the aromatic-H1' crosspeaks of d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) in D₂O. (A) Low salt sample (same spectrum as Figure 3). (B) High salt/methanol sample containing 0.5 M NaCl, 40 mM MgCl₂, 40% v/v methanol-d₄/D₂O, pH 6.6, in 400 µl at 40°C, and τ_m = 100 msec. The four strong GH8-H1' crosspeaks are indicated, as well as the weak AH8 and T8H8-H1' crosspeaks. 300 t₁ values of 128 scans each were collected. The spectrum was zero-filled in t₁ and apodized with a squared sine-bell phase shifted by 90° in t₁ and line-broadened by 1 Hz (exponential multiplication) in t₂.

Figure 8. Imino proton spectra of $d(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ in H_2O at 30°C . (A) Low salt sample (same sample as Figure 1a in H_2O). (B) High salt/methanol sample (same sample as Figure 1b in H_2O). Because the region of maximum excitation is centered between the aromatic and the imino resonances the resonance intensity of the A·T iminos appears smaller than it actually is. Spectra are line-broadened by 3 Hz. Inset: Imino proton spectrum of the Z-DNA form of $d(\text{br}^5\text{CGbr}^5\text{CGATbr}^5\text{CGbr}^5\text{CG})$ at 35°C in high salt/methanol. Sample conditions are given in reference 11.

the imino resonances. There is also some intensity from the terminal imino resonances, which exchange much more slowly with water in Z-DNA than B-DNA³⁹. Imino proton resonances from the thymine bases in the sample under these conditions appear as two partially resolved resonances at around 13.7 ppm. One dimensional NOEs were obtained on the A·T imino resonances in the high salt/methanol form (not shown). At this temperature, due to exchange with water, only weak NOEs are observed. (At lower temperatures substantially more B-DNA is present and the resolution is much poorer.) The A·T imino resonance at lower field showed an NOE to the AH2 resonance assigned to A₅ while the A·T imino resonance at higher field showed NOEs to A₅ and A₆ as well as to the AH2s from the residual B-form DNA.

³¹P NMR spectra. ³¹P NMR spectra of the dodecamer in low salt

Figure 9. 202 MHz proton-decoupled ^{31}P NMR spectra of $\text{d}(\text{br}^5\text{CGbr}^5\text{CGAATTbr}^5\text{CGbr}^5\text{CG})$ at 30°C . (A) Low salt sample (same as Figure 1a). (B) High salt/methanol sample (2 M NaCl, 10 mM MgCl_2 , 1 mM duplex, pH 7.2 in 56% v/v methanol- $\text{d}_4/\text{D}_2\text{O}$). Some residual B-DNA in this sample is indicated by the dotted lines. For (A), 256 transients of 8K points and spectral width 1500 Hz and for (B) 1500 transients of 4k points with a spectral width of 4000 Hz were collected with a recycle delay of 2 s. Spectra are line-broadened by 2 Hz. Spectra are referenced relative to external TMP (1% in 0.1M NaCl in D_2O).

and high salt/methanol at 30°C are shown in Figure 9. At low salt, the ^{31}P resonances of the dodecamer are clustered in a region around -4.4 ppm. The spectrum in high salt/methanol is markedly different. While most of the resonance intensity is centered around -4.2 ppm, resonances also appear downfield at -2.3 and -3.2 ppm as well as upfield at -5 ppm. Resonance intensity from some B-DNA in this sample is indicated by the dotted lines. Downfield shifted resonances, usually around 2 ppm from the rest of the phosphate resonances, are characteristic of the GpC steps in Z-DNA^{11,25}. For this dodecamer, if the ends are Z and the middle is not, one would expect shifted phosphates from the two GpC steps to give an integrated intensity of 2:9 relative to the rest of the resonances. After subtracting out the estimated contribution from B-DNA resonance intensity, we get a relative integrated intensity of the peak at -2.3 ppm to the resonances between -3.6 and -5.2

Table 3: Crosspeak Intensities for H1'-H2' and H1'-H2" COSY Crosspeaks for d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) in High Salt/Methanol[†]

BASE	H1'-H2'	H1'-H2"
C ₁	strong	no
G ₂	--- ^a	no
C ₃	strong	weak
G ₄	no	medium
A ₅	strong	weak
A ₆	strong	strong
T ₇	strong	medium
T ₈	strong	medium
C ₉	strong	weak
G ₁₀	no	strong
C ₁₁	medium	no
G ₁₂	medium	strong

[†]Qualitative relative intensities from Figure 6b.

^aCan't tell due to overlapping resonances

ppm of approximately 1:8. Since we are unable at this time to assign these resonances, it is difficult to explain this ratio or the resonance at -3.2 ppm. We did observe that the intensity of the resonance(s) at -2.3 ppm was temperature dependent, with the relative intensity actually increasing (relative to the resonances arising from the high salt/methanol form) as the temperature decreased. This may indicate a temperature dependent mobility of the CG ends of the dodecamer under these conditions.

DISCUSSION

Conformation of the dodecamer in low salt. Analysis of NOESY spectra of d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) (Figure 1) indicates that in low salt aqueous solution the dodecamer adopts a regular B-DNA structure. The chemical shifts of the unmodified bases and the pattern of NOESY crosspeaks are similar to those observed for the related unmodified dodecamer d(CGCGAATTCGCG)¹⁹. Since our primary purpose here was to investigate the structure of this dodecamer under conditions where part or all of the molecule might adopt a Z-DNA conformation, we have not attempted here to get a more detailed structure of the low salt, B-DNA conformation.

Conformational transition between the two forms. Addition of a combination of high salt and methanol to the dodecamer solution results in dramatic spectral changes which indicate that the molecule adopts an alternative conformation under these conditions. Addition of methanol alone (up to 50% v/v) to the solution in low salt does not significantly change the conformation of the DNA, although some spectral changes are observed. Addition of high concentrations of salt with no added methanol also does not effect the conformational change (spectra not shown). At intermediate salt concentrations in methanolic solution, such as in Figure 2b, both a B-DNA conformation and the

alternative conformation are observed in equilibrium with each other in slow exchange on the NMR time scale. In 1M NaCl and 50% methanol there is about 60% of the alternative conformation (spectra not shown). Addition of 11 mM MgCl₂ to that sample resulted in almost complete conversion to the alternative conformation (Figure 1b, 2c). Thus, addition of divalent cations significantly stabilizes the high salt/methanol conformation. Some residual B-DNA, in intermediate exchange with the high salt/methanol conformation at this temperature, appears as a broad background in the spectra. We were also able to effect the transition at lower NaCl concentrations by using higher concentrations of MgCl₂ (up to 40 mM) (Figure 7b), but this resulted in increased sample aggregation and precipitation.

The transition from B-form DNA to the high salt/methanol form was also temperature dependent, with the amount of the alternate conformation increasing with increasing temperature. For this reason, most of our analysis was done on samples at 40 °C.

The transition to the high salt/methanol form is accompanied by large chemical shift changes for many of the proton resonances in the molecule. The GH8 resonances all show large upfield shifts of .06-.21 ppm compared to the low salt form. The CH6 resonances, with the exception of C₃, also all shift upfield. Upfield shifts of these resonances were also observed in the Z-DNA conformations of d(br⁵CGbr⁵CGAThr⁵CGbr⁵CG)¹¹ and d(br⁵CG)₃²⁴. The AH8, AH2, TH6, and TMe resonances all shift downfield, indicative of a decrease in stacking interactions for these bases.

Sugar and base conformation in the high salt/methanol form. Crosspeak intensities in NOESY spectra of the high salt/methanol form of the dodecamer indicate that all four of the guanines are in the *syn* conformation, while the rest of the bases are *anti* (Figure 7b). The alternation of *anti* cytosines and *syn* guanines is consistent with the br⁵CGbr⁵CG ends of the molecule adopting the Z-DNA conformation¹¹. The *anti* conformation for all of the AATT bases in the center indicates that this part of the molecule is not Z-DNA. Further evidence for a Z-DNA conformation for the ends of the dodecamer and a non-Z-DNA conformation for the middle is obtained from the sugar H1'-H2' and H1'-H2'' couplings (Figure 6b). In Z-DNA, the sugars are generally in the C2' endo conformation for the bases that are *anti* (generally cytosines) and in the C3' endo conformation for the bases that are *syn* (generally guanines)². The C2' endo conformation is also the average sugar pucker for B-DNA. Sugar coupling constants for these two sugar puckers are J_{1'2'} = 9.5 and J_{1'2''} = 5.8 for C2' endo and J_{1'2'} = 1.5 and J_{1'2''} = 7.7 for C3' endo²⁷. In the Z-DNA conformation of d(br⁵CG)₃, reported coupling constants for the internal guanines are about J_{1'2'} = 2.0 and J_{1'2''} = 8.0 and those for the cytosines are about J_{1'2'} = 9.0 and J_{1'2''} = 5.7²⁸. The observed COSY crosspeaks and their qualitative intensities for the high salt/methanol form are summarized in Table 3. The observed pattern of H1'-H2' and H1'-H2'' crosspeaks is consistent with a predominantly C3'-endo sugar pucker for all of the guanines except the terminal G₁₂ and a predominantly C2'-endo sugar pucker for all of the cytosines. This provides further evidence for a Z-DNA conformation for the

br⁵CGbr⁵CG ends of the dodecamer. The sugar conformations for the internal AATT residues appears to vary somewhat but have an average around C2'-endo. Although the proton NMR data are consistent with a Z-DNA conformation for the CG ends of the dodecamer in high salt/methanol, the ³¹P spectra indicate that there may be some deviation from a regular Z-DNA structure in these ends. Further work will need to be done to explain the phosphorus NMR results.

Base pairing in the high salt/methanol form. The above analysis of the base and sugar conformation establishes that in the br⁵CGbr⁵CG ends of the molecule the Gs are *syn* with a predominantly C3'-endo sugar pucker and the Cs are *anti* with a predominantly C2'-endo sugar pucker, consistent with a Z-DNA conformation for the ends of the molecule. In addition, it is clear that the AATT center of the dodecamer is not Z-DNA, since neither A₇ nor T₉ are in the *syn* conformation or have a C3'-endo sugar pucker, which would be expected if the Z-DNA conformation was propagated through the AATT tract. However, the structure of these bases is not clear. One possible conformation for the dodecamer which would be consistent with the results discussed in the preceding section would be a hairpin structure with a four base Z-DNA stem and a four base AATT loop (Scheme 1d). Imino proton spectra of the dodecamer in high salt/methanol show two A·T imino resonances (Figure 8b). These resonances appear at chemical shifts normally observed for hydrogen bonded A·T imino resonances²⁹, and therefore indicate that the AATT tract is base paired. Observation of NOEs between the imino and their AH2 resonances indicates that these are Watson-Crick base pairs. If the molecule was in a Z-DNA (or B-DNA) hairpin, we would expect to observe the thymine imino resonances at around 10-11 ppm. Based on steric considerations, we would not expect the adenines and thymine in the loop to be base paired or at most only A₅-T₈. Therefore, we conclude that in high salt/methanol the dodecamer adopts a conformation in which the br⁵CGbr⁵CG ends of the molecule have the base and sugar conformation observed in Z-DNA and the AATT center is not Z-DNA.

Implications for the structure of B-Z junctions. Relatively little is known about the structure of junctions between Z and B-DNA. The structure of potential B-Z junctions in supercoiled plasmids has been probed by nuclease digestion, two dimensional gel electrophoresis, and chemical footprinting studies^{9,10,30-36}. These studies and cited references have been variously interpreted to indicate that B-Z junctions are conformationally flexible, have different structures depending on sequence, may be up to several bases in length, and may contain anywhere from zero to several unpaired bases. Kang and Wells concluded, based on reactions of bromoacetaldehyde with recombinant plasmids containing segments of Z-DNA, that B-Z junctions contain few, if any, unpaired bases¹⁰. However, Wells and other workers have more recently concluded that the junction does contain non-paired bases^{32,34,35}.

The results obtained here have some implications for the structure of B-Z junctions. First, the fact that the br⁵CGbr⁵CG ends of the molecule appear to form a Z-DNA helix implies that the junction region does not extend significantly into the Z-DNA

forming part of the molecule. Second, the junction region does not contain any unpaired bases. Imino proton resonances are observed between 12 and 14 ppm for all of the thymines and guanines in the molecule, indicating that all of the bases are base paired. Third, the presence of a junction region appears to result in a higher flexibility or disorder than for B-DNA.

The length of the structural perturbation we would call a B-Z junction cannot be obtained from this molecule. The entire AATT region appears to be non-B-DNA, but since there is Z-DNA on both sides of the AATT tract and therefore two junctions, this places a lower limit of two base pairs for a B-Z junction. It is not clear to what extent the junction perturbs the Z-DNA ends of the molecule. Almost all of the base resonances show downfield shifts in high salt/methanol from their resonance positions in the low salt, B-DNA, form. In addition, the sequential A₆H₈-T₇Me and T₇Me-T₈Me connectivities observed in NOESY spectra of the B-DNA form are absent in the high salt/methanol sample. Sequential base-H1'-base connectivities are also not seen, indicating some departure from standard B-DNA conformation. These chemical shift and NOE changes in the AATT tract imply considerable unstacking or unwinding of the DNA in this region. Interestingly, chemical probing of B-Z junctions with osmium tetroxide led to the suggestion by Johnston and Rich that there was partial or transient unstacking of T and C bases in junction regions³¹. Some unwinding of the DNA in the AATT region (see reference to 6 below) would also be consistent with the results obtained here.

Sheardy and Winkle³⁸ have used CD and one-dimensional NMR spectra of the imino proton resonances and phosphorus resonances of a DNA hexadecanucleotide containing one potential B-Z junction in order to obtain information on B-Z junctions. They conclude that only three base pairs are involved in the junction and one of them is dramatically distorted. The base pairs in the junction are suggested by the authors to be the last two alternating m⁵C·G base pairs (7 and 8) in the Z-DNA end of the molecule and an adjacent A·T base pair (9). Since assignments of the imino protons at high salt (B-Z junction) were made on the basis of the low salt (B-DNA) assignments and the ³¹P spectra are not assigned at all, and the arguments used for their conclusions are not clear to us, it is difficult to evaluate their data relative to our own.

Sequence specificity of the B-Z transition. The results obtained here indicate that a regular Z-DNA structure is not propagated through four A·T base pairs in the sequence AATT even though they are flanked by Z-DNA. Wells and coworkers⁸ have looked at runs of A·T base pairs inserted between (CG)_n in supercoiled plasmids and assayed for the formation of Z-DNA using two-dimensional gel electrophoresis and nuclease digestion studies. They concluded that consecutive A·T base pairs, whether alternating (TA)₁₋₃ or (TT)₁₋₂, could adopt a Z-DNA structure when flanked by alternating (CG)₃₋₅. In contrast, our sequence (AATT) will not adopt the Z-DNA conformation when flanked by (br⁵CG)₂. This difference may be due to the fact that we are using different techniques both to induce and detect the Z-DNA conformation. However, it is more likely that this is a real difference in the

sequence specificity of Z-DNA formation. A notable difference in the sequences of A·T base pairs discussed above is that our sequence is the only one that begins and ends with an A·T base pair out of purine-pyrimidine alternation with the CG tract. Ellison, et al.⁶ have looked at the ability of (TA)₄ to adopt the Z-DNA conformation when flanked by G·C base pairs. They found that in negatively supercoiled plasmids the two d(CG)₆ portions of the insert which flanked the (TA)₄ center adopted the Z-DNA conformation, while the (TA)₄ formed an underwound structure with a helical repeat that appears to be intermediate between B and Z-DNA. Thus, they effectively also have two Z-DNA junctions which appear to alter the conformation of the (TA)₄ center of the insert. This result is consistent with what we observe in our dodecamer for the AATT center under conditions where the ends of the molecule are Z-DNA.

CONCLUSIONS

In this work we have shown that under appropriate salt and solvent conditions the DNA dodecamer d(br⁵CGbr⁵CGAATTbr⁵CGbr⁵CG) adopts a conformation in which the br⁵CGbr⁵CG ends of the molecule are Z-DNA and the AATT center is neither B or Z-DNA. Although the conditions used to effect this transition from the B-DNA structure observed at low salt are not physiologically relevant, the qualitative results we have obtained on this structure are relevant in terms of the conformation of B-Z junctions and the sequence specificity of Z-DNA. We conclude that the run of four A·T base pairs in which the first and last are out of purine-pyrimidine alternation with the br⁵CG part of the molecule are sufficient to disrupt the formation of Z-DNA, and instead adopt a conformation which is neither Z-DNA nor standard B-DNA. Although the AATT region has an altered structure from standard B-DNA, it contains no unpaired bases.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health Grant R01 GM 37254-01 and Office of Naval Research Contract No. N00014-88-K-0180 to JF. We thank Karl M. Koshlap and Dara E. Gilbert for careful reading of this manuscript and help in obtaining some of the spectra.

*To whom correspondence should be addressed

REFERENCES

1. Pohl, F.M. and Jovin, T.M. (1972) J. Mol. Biol. **67**, 375-396.
2. 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 **282**, 680-686.
3. Drew, H., Takano, T., Tanaka, S., Itakura, K., Dickerson, R.E. (1980) Nature **286**, 567-573.

4. Rich, A., Nordheim, A. and Wang, A.H.-J. (1984) *Ann. Rev. Biochem.* **53**, 791-846.
5. Wells, R.D., Amirhaeri, S., Blaho, J.A., Collier, D.A., Hanvey, J.C., Hsieh, W.-T., Jaworski, A., Klysik, J., Larson, J.E., McLean, M.J., Wohlrab, F. and Zacharias, W. (1988) In Wells, R.D. and Harvey, S.C. (eds.), *Unusual DNA Structures*. Springer-Verlag, New York, pp. 1-21.
6. Ellison, M.J., Feigon, J., Kelleher, R.J. III, Wang, A.H.-J., Habener, J.F. and Rich, A. (1986) *Biochem.* **25**, 3648-3655.
7. Ellison, M.J., Kelleher, R.J. III, Wang, A.H.-J., Habener, J.F. and Rich, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8320-8324.
8. McLean, M.J., Blaho, J.A., Kilpatrick, M.W. and Wells, R.D. (1986) *Proc. Natl. Acad. Sci.* **83**, 5884-5888.
9. Singleton, C.K., Klysik, J. and Wells, R.D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2447-2451.
10. Kang, D.S. and Wells, R.D. (1985) *J. Biol. Chem.* **260**, 7783-7790.
11. Feigon, J., Wang, A.H.-J., van der Marel, G.A., van Boom, J.H. and Rich, A. (1985) *Science* **230**, 82-84.
12. Wang, A.H.-J., Gessner, R., van der Marel, G.A., van Boom, J.H. and Rich, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3611-3615.
13. van der Marel, G. A., van Boeckel, C. A. A., Wille, G. and van Boom, J. H. (1981) *Tetrahedron Lett.* **22**, 3887-3890.
14. Sklenar, V., Brooks, B. R., Zon, G. and Bax, A. (1987) *FEBS Lett.* **216**, 249-252.
15. States, D. J., Haberkorn, R. A. and Ruben, D. J. (1982) *J. Magn. Reson.* **48**, 286-292.
16. Kumar, A., Ernst, R. R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Comm.* **95**, 1-6.
17. Aue, W.P., Bartholdi, E., & Ernst, R.R. (1976) *J. Chem. Phys.* **64**, 2229-2246.
18. Rance, M., Sørensen, Bodenhausen, G., Wagner, G., Ernst, R. R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117**, 479-485.
19. Hare, D. R. Wemmer, D. E., Chou, S. H., Drobny, G., Reid, B. R. (1983) *J. Mol. Biol.* **171**, 319-336.
20. Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.* **71**, 4546-4553.
21. Scheek, R. M., Russo, N., Boelens, R., Kaptein, R. (1983) *J. Am. Chem. Soc.* **105**, 2914-2916.
22. Feigon, J., Wright, J. M., Leupin, W., Denny, W. A. and Kearns, D. R. (1982) *J. Am. Chem. Soc.* **104**, 5540-5541.
23. Feigon, J., Wang, A.H.-J., van der Marel, G.A., van Boom, J.H. and Rich, A. (1984) *Nucl. Acids Res.* **12**, 1243-1263.
24. Orbons, L.P.M., van der Marel, G.A., van Boom, J.H. and Altona, C. (1986) *Eur. J. Biochem.* **160**, 131-139.
25. Feigon, J., Wang, A.H.-J., van der Marel, G.A., van Boom, J.H. and Rich, A. (1984) *Nucl. Acids Res.* **12**, 1243-1263.
26. Patel, D.J., Kozlowski, S.A., Nordheim, A. and Rich, A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1413-1417.
27. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley-Interscience, New York, 292 pp.
28. Orbons, L.P.M. and Altona, C. (1986) *Eur. J. Biochem.* **160**, 141-148.
29. Kearns, D. R. (1977) *Annu. Rev. Biophys. Bioeng.* **6**, 477-523.
30. Kilpatrick, M.W., Wei, C.-F., Gray, H.B., Jr. and Wells, R.D. (1983) *Nucl. Acids Res.* **11**, 3811-3822.
31. Johnston, B.H. and Rich, A. (1985) *Cell* **42**, 713-724.

32. Nejedly, K., Matyasek, R. and Palacek, E. (1988) *J. Biomol. Str. & Dyn.* **6**, 261-273.
33. Kilpatrick, M.W., Klysik, J., Singleton, C.K., Zarling, D.A., Jovin, T.M., Hanau, L.H., Erlanger, B.F. and Wells, R.D. (1984) *J. Biol. Chem.* **259**, 7268-7274.
34. Kohwi-Shigematsu, T., Manes, T. and Kohwi, Y. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2223-2227.
35. McLean, M.J., Larson, J.E., Wohlrab, F. and Wells, R.D. (1987) *Nucl. Acids Res.* **15**, 6917-6935.
36. Azorin, F., Hahn, R. and Rich, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5714-5718.
37. Wells, R.D., Brennan, R., Chapman, K.A., Goodman, T.C., Hart, P.A., Hillen, W., Kellogg, D.R., Kilpatrick, M.W., Klein, R.D., Klysik, J., Lambert, P.F., Larson, J.E., Miglietta, J.J., Neuendorf, S.K., O'Connor, T.R., Singleton, C.K., Stirdivant, S.M., Veneziale, C.M., Wartell, R.M. and Zacharias, W. (1983) *Cold Spring Harbor Symposia on Quantitative Biology*, Vol. XLVII, 77-84.
38. Sheardy, R.D. and Winkle, S.A. (1989) *Biochem.* **28**, 720-725.
39. Mirau, P.A. and Kearns, D.R. (1983) *Nucl. Acids Res.* **11**, 1931-1941.