Nucleic Acids Research, 1999, Vol. 27, No. 2 543–550

The intrinsic structure and stability of out-of-alternation base pairs in Z-DNA

Brandt F. Eichman, Gary P. Schroth[+], Beth E. Basham[§] and P. Shing Ho*

Department of Biochemistry and Biophysics, ALS 2011, Oregon State University, Corvallis, OR 97331, USA

Received September 16, 1998; Accepted November 10, 1998

NDB/PDB accession no. ZD0001

ABSTRACT

Alternating pyrimidine-purine sequences typically form Z-DNA, with the pyrimidines in the *anti* and purines in the *syn* conformations. The observation that dC and dT nucleotides can also adopt the *syn* conformation (i.e. the nucleotides are out-of-alternation) extends the range of sequences that can convert to this left-handed form of DNA. Here, we study the effects of placing two adjacent d(G•C) base pairs as opposed to a single d(G•C) base pair or two d(A•T) base pairs out-of-alternation by comparing the structure of d(m^{5}CGGCm^{5}CG)$_2$ with the previously published structures of d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) and d(m^{5}CGATm^{5}CG)$_2$. A high buckle and loss of stacking interactions are observed as intrinsic properties of the out-of-alternation base pairs regardless of sequence and the context of the dinucleotide. From solution titrations, we find that the destabilizing effect of out-of-alternation d(G•C) base pairs are identical whether these base pairs are adjacent or isolated. We can therefore conclude that it is these intrinsic distortions in the structure of the base pairs and not neighboring effects that account for the inability of out-of-alternation base pairs to adopt the left-handed Z conformation.

INTRODUCTION

Since its discovery in 1979 (1), nearly every physical property of left-handed Z-DNA has been characterized. Its structure has been studied by numerous physical methods and the thermodynamic rules for its formation in solution and within plasmid DNAs have been studied in detail for a large number of different sequences (reviewed in 2). Still, many of the detailed structural parameters for Z-DNA are yet to be defined, particularly for sequences that do not conform to the standard alternating pyrimidine-purine (APP) sequence motif of Z-DNA and which are therefore considered to be out-of-alternation. In this study, we attempt to determine the structural perturbations to the Z-DNA structure that are inherent in out-of-alternation base pairs and how these affect the ability of these sequences to form Z-DNA.

The detailed structure of Z-DNA was first defined by the single crystal structure of d(CGCGCG)$_2$ (1). This proved to be the conformation responsible for the inversion of the circular dichroism (CD) spectra of poly(dG•dC) at high salt (3). In the nearly two decades since these early studies, the Z-DNA conformation has been studied in crystals by X-ray diffraction and in solution by a variety of spectroscopic techniques, including CD, NMR and infrared spectroscopy (reviewed in 4). The culmination of these studies show that the nucleotide bases of Z-DNA alternate between the *anti* and *syn* conformations, which accounts for the preference of the structure for APP sequences. In addition, replacing d(G•C) base pairs with d(A•T) base pairs and placing purine bases in the *syn* conformation were found to destabilize Z-DNA. Studies of Z-DNA induced by negative superhelical strain confirm these general rules for the sequence-dependent stability of this left-handed conformation (5,6) and further suggest that the structure can form under physiological conditions (7) and indeed *in vivo* (8).

The generality of Z-DNA as an alternative to B-DNA in genomes requires that the structure not be limited to only APP sequences, but to all possible combinations of nucleotides. The first indication that Z-DNA could be formed in non-APP sequences was found in the single crystal structure of d(Br^{5}CGATBr^{5}CG)$_2$ (9). The two adjacent d(A•T) base pairs were out-of-alternation, with the adenine nucleotides in the *anti* conformation and the thymidines adopting the *syn* conformation. These conformations around the glycosidic bonds were also observed by NMR spectroscopy in the solution structure of d(Br^{5}CGBr^{5}CGATBr^{5}CGBr^{5}CG)$_2$ (10). The out-of-alternation base pairs in the crystal structure were highly buckled as compared with other base pairs in Z-DNA. Placing a pyrimidine in the *syn* conformation is thought to be disfavored by the steric collision between the six-membered ring of the base and the five-membered ring of the ribose (11). Thus, this buckle in the base pair may be largely responsible for relieving this steric strain. The more recent structure of d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) demonstrated that out-of-alternation d(G•C) base pairs can also be accommodated by the left-handed conformation (12). The structure of the out-of-alternation d(G•C) base pair also appeared to be highly buckled. However, the structural features of d(Br^{5}CGATBr^{5}CG)$_2$ and d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) could not be directly compared since both the type and the number of base pairs that

*To whom correspondence should be addressed. Tel: +1 541 737 2769; Fax: +1 541 737 0481; Email: hops@ucs.orst.edu

Present addresses: [+]Genelabs Technologies, 505 Penobscot, Redwood City, CA 94063, USA and [§]Department of Molecular Biology, DNAX Research Institute of Molecular and Cellular Biology Inc., 901 California Avenue, Palo Alto, CA 94304-1104, USA

are out-of-alternation differed between the two sequences. Thus, we needed a structure that helps to bridge these two previous non-APP Z-DNA sequences in order to determine the intrinsic properties of out-of-alternation base pairs.

In the current study, we address the question of whether the distortions to out-of-alternation base pairs are intrinsic features or induced by sequence or neighboring effects. We have determined the structure of the self-complementary hexamer sequence d(m^{5}CGGCm^{5}CG)$_2$ as Z-DNA, where the two central base pairs are out-of-alternation. Comparing this structure with that of d(Br^{5}CGATBr^{5}CG)$_2$ allows us to determine the contribution of sequence to the structure of out-of-alternation base pairs in Z-DNA. To study the effects of neighboring base pairs on non-APP sequences, we compare the current structure with d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) and study the effect of placing two out-of-alternation d(G•C) base pairs either adjacent to each other or separated by a distance on the ability of Z-DNA to form in solution. In this way, we attempt to define the intrinsic structure of out-of-alternation base pairs and their effects on the ability of sequences to adopt this left-handed form of the DNA duplex.

MATERIALS AND METHODS

Crystallization and X-ray diffraction studies

The self-complementary hexanucleotide sequence d(m^{5}CGGCm^{5}CG) was synthesized using phosphoramidite chemistry on an Applied Biosystems DNA synthesizer in the Center for Gene Research and Biotechnology at Oregon State University. Size exclusion chromatography on a Sephadex G-25 column was used to remove salts, blocking groups and prematurely terminated oligonucleotides. The oligonucleotide was lyophilized, redissolved in 30 mM sodium cacodylate buffer (pH 7.0) and used for crystallization without further purification. Crystals of the sequence were grown at room temperature by vapor diffusion in sitting drop set-ups with initial solutions containing 0.3 mM DNA (single-stranded), 12 mM sodium cacodylate (pH 7.0), 25 mM $MgCl_2$, 1 mM spermine tetrachloride and 5% (v/v) 2-methyl-2,4-pentanediol (MPD), equilibrated against a reservoir of 30% MPD.

Crystals were mounted in sealed glass capillaries and X-ray diffraction data collected at room temperature on a Siemens P4 diffractometer with CuK_α radiation from a sealed tube source. The data was collected and reduced using the software XSCANS. The crystal diffracted to a resolution better than 1.65 Å (3197 unique reflections) and belongs to the orthorhombic space group $P2_12_12_1$ with unit cell dimensions of a = 17.79 Å, b = 30.90 Å and c = 44.36 Å. The space group and unit cell dimensions are consistent with a Z-DNA hexamer aligned along the crystallographic c-axis.

Structure solution and refinement

The structure of d(m^{5}CGGCm^{5}CG)$_2$ was solved by molecular replacement using as a starting model the 1.3 Å resolution crystal structure of d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) (12) with the d(G•C) base pair at position 4 (underlined) replaced with a d(C•G) base pair. The structure was subsequently refined using a parameter file specific for DNA (13) in the program X-PLOR (14). The initial refinement included data observed above 2σ(F) in the resolution range 8.0–2.0 Å. A model with an R value of 26.9% was obtained after simulated annealing (starting temperature of 2000 K). Data in the resolution range from 8.0 to 1.65 Å [1875 unique reflections observed above 3σ(F)] were included in subsequent rounds of simulated annealing and conventional positional refinement. The refinement converged to a final R value of 19.6% (R_{free} = 24.8%). Although data to 1.65 Å was included in the refinement, the nominal resolution limit for this structure is 1.9 Å for a thin shell completeness of at least 50%.

In the final structure of d(m^{5}CGGCm^{5}CG)$_2$, the bond length and angle r.m.s. deviations from ideality were calculated in X-PLOR to be 0.008 Å and 1.473°, respectively, with average B factors of 10.3 for the 244 non-hydrogen DNA atoms and 32.3 for 35 water molecules. A coordinate error of <0.2 Å was estimated from a Luzzati plot (15).

Helical parameters were analyzed using the program NASTE (Nucleic Acid STructure Evaluation), which was developed in this laboratory specifically for analysis of Z-DNA duplex structures (16). The final coordinates and structure factors for the structure of d(m^{5}CGGCm^{5}CG)$_2$ have been deposited in the Nucleic Acid Database (17) under the reference code ZD0001.

Salt/ethanol titrations

The oligonucleotides d(GC)$_{12}$, d[(GC)$_3$CCGC(GC)$_2$GCGG(GC)$_3$] and d[(GC)$_5$GGCC(GC)$_5$] were synthesized and purified in the same manner as described above for d(m^{5}CGGCm^{5}CG). Stock solutions of $MgCl_2$ were prepared by serial dilutions to give concentrations of 0, 0.1, 0.25 and 0.5–3.5 M in increments of 0.5 M. To induce Z-DNA formation, 10 µl of a 40–60 OD oligonucleotide sample was diluted into 990 µl of each $MgCl_2$ solution, heated to 67°C in a water bath for 10 min and allowed to cool overnight to achieve equilibrium and to ensure proper annealing. CD spectra were collected on a Jasco J-720 spectropolarimeter at room temperature in a 1 cm path length cell.

The two non-APP sequences d[(GC)$_3$CCGC(GC)$_2$GCGG(GC)$_3$]$_2$ and d[(GC)$_5$GGCC(GC)$_5$]$_2$ (the underlined bases are those that are out-of-alternation) could not be induced to form Z-DNA with $MgCl_2$ alone, even at 3.5 M $MgCl_2$. Solutions of these sequences were therefore made with 1.8 M $MgCl_2$, a concentration at which d[(GC)$_{12}$]$_2$ is 100% Z-DNA, and titrated with ethanol concentrations ranging from 0 to 45% (v/v).

RESULTS

Z-DNA duplex structure of d(m^{5}CGGCm^{5}CG)$_2$

We have determined the crystal structure of the self-complementary sequence d(m^{5}CGGCm^{5}CG)$_2$ to a nominal resolution of 1.9 Å. In this sequence, the two center base pairs form an out-of-alternation d(G3pC4)•d(G9pC10) dinucleotide which breaks the APP sequence motif favored by Z-DNA. The resulting structure (Fig. 1) is in the left-handed Z-DNA conformation and, despite the distortions resulting from the d(G3pC4)•d(G9pC10) dinucleotide step, displays the defining characteristics of previous crystal structures of Z-DNA hexanucleotides (1,18; reviewed in 2). The helical twist between base pairs alternate between being highly left-handed (–46.5 ± 0.4°) and only slightly left-handed (–13.1 ± 1.8°), giving the overall structure the characteristic zig-zagged pattern to the backbone (Table 2). The glycosidic bond angles alternate between *anti* (χ = –150.7 ± 5.1°) and *syn* (χ = 68.0 ± 9.2°) and the sugar puckers alternate between C2′-*endo* and C3′-*endo* (Table 1). Nucleotides with the base in the *anti* conformation adopt the C2′-*endo* sugar conformation, while those in the *syn* conformation

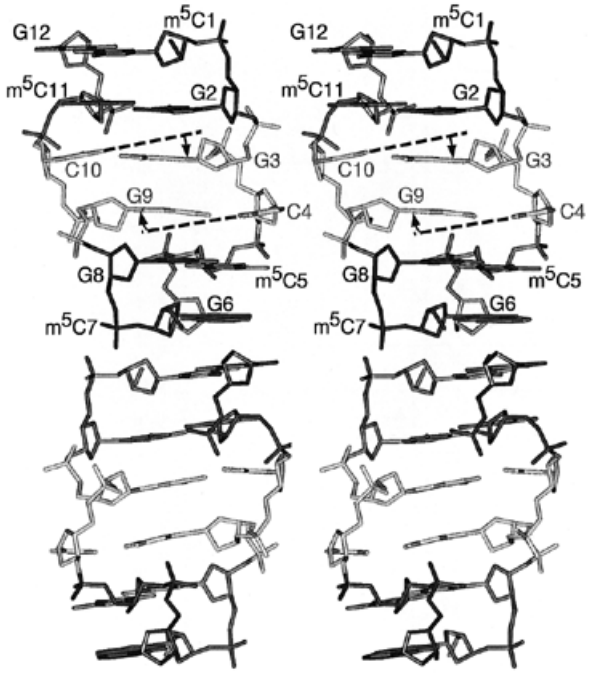

Figure 1. Stereo diagram of the d(m^{5}CGGCm^{5}CG)$_2$ structure at 1.9 Å resolution. Two hexamers are shown as they appear in the crystal, stacked end-on-end along the helix axis to give one pseudocontinuous turn of Z-DNA. Residues are numbered for one asymmetric unit only. The out-of-alternation d(G•C) base pairs are lightly shaded. The high degree of base pair buckle at the cytosines of the out-of-alternation base pairs is indicated by the dashed lines and arrows.

adopt the C3′-*endo* sugar conformation. As with previous Z-DNA hexanucleotide crystal structures, however, the 3′-terminal guanines of each strand are exceptions in that these nucleotides in the *syn* conformation adopt a C2′-*endo* sugar pucker (2).

Structure of the out-of-alternation d(G3pC4)•d(G9pC10) dinucleotide step

The most striking features of the current structure are found in the out-of-alternation base pairs, d(G3•C10) and d(C4•G9), of the central dinucleotide. We observe that the structures of these are similar to those in the non-APP structures d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) and d(m^{5}CGATm^{5}CG)$_2$ in terms of (i) the conservation of alternating *anti-p-syn* glycosidic conformations of the dinucleotides, (ii) the high buckle of the out-of-alternation base pairs and (iii) the base stacking at this central dinucleotide step (Fig. 2).

Although base pairs d(G3•C10) and d(C4•G9) disrupt the APP sequence motif for Z-DNA, the glycosidic bonds along each chain still alternate between the *anti* and *syn* conformations (Table 1). The guanine nucleotides at positions 3 and 9, which are typically occupied by pyrimidines in Z-DNA hexamer sequences, have χ angles of –155.4° and –151.4°, placing them in the *anti* conformation. Likewise, cytosines at positions 4 and 10, which are typically occupied by purines, have χ angles of 57.0° and 64.9° associated with nucleotides in the *syn* conformation.

Table 1. Glycosidic torsion angles (χ) and sugar puckers in the structure of d(m^{5}CGGCm^{5}CG)$_2$

Nucleotide	χ (°)	Sugar pucker
m^{5}C1	–142.1 (*anti*)	C2′-*endo*
G2	59.4 (*syn*)	C3′-*endo*
G3	–155.4 (*anti*)	C2′-*endo*
C4	57.0 (*syn*)	C3′-*endo*
m^{5}C5	–152.6 (*anti*)	C1′-*exo*
G6	78.7 (*syn*)	C2′-*endo*
m^{5}C7	–147.6 (*anti*)	C2′-*endo*
G8	69.8 (*syn*)	C4′-*exo*
G9	–151.4 (*anti*)	C1′-*exo*
C10	64.9 (*syn*)	C4′-*exo*
m^{5}C11	–155.2 (*anti*)	C1′-*exo*
G12	78.2 (*syn*)	C2′-*endo*

All glycosidic and pseudorotation angles (used to assign sugar pucker conformations) were calculated using NASTE (Nucleic Acid STructure Evaluation), a program developed in this laboratory for structural analysis of Z-DNA. C1′-*exo* and C4′-*exo* sugar puckers belong to the C2′-*endo* and C3′-*endo* families, respectively.

Placing cytosines in the unfavorable *syn* conformation perturbs both the planarity and the stacking of the base pairs in Z-DNA. The d(G3•C10) and d(C4•G9) base pairs (Fig. 1) are highly buckled, evident from the relatively large angle κ between the cytosine and guanine bases within the base pairs. In the d(G3•C10) base pair, κ = 12.7°, while κ = –12.3° for the d(C4•G9) base pair. The typical value for a standard base pair in Z-DNA is κ = 3.1 ± 2.4° (Table 2). The out-of-alternation base pairs in d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) (12) and d(Br^{5}CGATBr^{5}CG)$_2$ (9) are also highly buckled (Table 2). Thus, this structural feature appears to be characteristic of out-of-alternation base pairs in Z-DNA, regardless of the type of base pair that is out-of-alternation and regardless of whether the adjacent base pairs are in- or out-of-alternation. This large buckle in the out-of-alternation base pairs results primarily from perturbations to the position of the cytosine base in the duplex. As we can see in Figure 1, the guanines in the *anti* conformation lie virtually parallel with the other base pairs in the duplex. This is consistent with the premise that it is the cytosine that must accommodate the steric strain of being in the *syn* conformation.

In addition, the cytosines of the out-of-alternation base pairs protrude into the major groove of the duplex, leaving the bases largely unstacked and exposed to solvent (Fig. 2). The purines of these base pairs, however, remain well stacked within the duplex, much in the same manner as cytosine bases in normal Z-DNA d(CpG) base steps. The protrusion of the out-of-alternation pyrimidines into the major groove are also observed in the two other non-APP crystal structures. The instability of out-of-alternation base pairs in Z-DNA therefore could arise from either the perturbation to the planarity of the base pair plane or from the additional exposure of the pyrimidine bases resulting from a loss of stacking interactions. In an attempt to distinguish between these two effects, we compared the specific solvent interactions around the out-of-alternation base pairs of the current and previous non-APP structures and compared the effects of pairing or isolating the out-of-alternation base pairs on the ability of sequences to adopt the Z-conformation.

Table 2. Comparison of base step and base pair helical parameters of Z-DNA crystal structures containing out-of-alternation base steps[a]

	d(CGCGCG)	d(m^{5}CGm^{5}CGm^{5}CG)	d(m^{5}CGGCm^{5}CG)	d(m^{5}CGGGm^{5}CG)• d(m^{5}CGCCm^{5}CG)	d(Br^{5}CGATBr^{5}CG)
Twist					
(C1pG2)•(C11pG12)	–8.5	–14.4	–12.8	–13.6	–13
(G2pN3)•(N10pC11)	–48.8	–43.6	–46.8	–46.8	
(N3pN4)•(N9pN10)	–9.1	–14.5	–11.5	–12.4	–9
(N4pC5)•(G8pN9)	–51.4	–44.5	–46.2	–46.8	
(C5pG6)•(C7pG8)	–10.6	–16.1	–15.1	–14.7	–12
Average *anti-p-syn* steps	–9.4 ± 1.1	–15.0 ± 1.0	–13.1 ± 1.8	–13.6 ± 1.2	–11 ± 2
Average *syn-p-anti* steps	–50.1 ± 1.8	–44.1 ± 0.6	–46.5 ± 0.4	–46.8	–49
Rise					
(C1pG2)•(C11pG12)	3.8	3.9	3.9	3.9	
(G2pN3)•(N10pC11)	3.7	3.8	3.6	3.6	
(N3pN4)•(N9pN10)	3.8	3.7	3.7	3.6	
(N4pC5)•(G8pN9)	3.6	3.8	3.7	3.8	
(C5pG6)•(C7pG8)	4.3	3.9	3.6	3.8	
Average *anti-p-syn* steps	4.0 ± 0.3	3.8 ± 0.1	3.7 ± 0.2	3.8 ± 0.2	
Average *syn-p-anti* steps	3.7 ± 0.1	3.8	3.7 ± 0.1	3.7 ± 0.1	
Propeller twist					
C1•G12	0.8	2.0	–2.0	0.9	
G2•C11	2.1	3.4	3.6	1.8	
N3•N10	5.6	4.8	1.6	3.0	
N4•N9	3.4	1.2	–4.7	2.2	
C5•G8	0.6	0.3	1.3	2.3	
G6•C7	3.2	2.1	0.5	2.8	
Average[b]	2.6 ± 1.9	2.3 ± 1.6	2.3 ± 1.6	2.2 ± 0.7	
Buckle					
C1•G12	0.3	6.2	4.1	2.6	
G2•C11	–4.8	–4.8	–0.8	–0.8	
N3•N10	2.8	2.1	12.7	14.8	
N4•N9	–5.9	–5.7	–12.3	–5.4	
C5•G8	0.1	5.3	1.9	2.3	
G6•C7	4.4	–3.8	–2.0	–3.2	
Average[b]	3.1 ± 2.4	4.7 ± 1.5	5.6 ± 5.4	4.8 ± 5.1	
Reference	1	28	This work	12	9

[a]Base step and base pair parameters for crystallized Z-DNA structures containing out-of-alternation base pairs (underlined) are compared with the alternating pyrimidine-purine sequences d(CGCGCG) and d(m^{5}CGm^{5}CGm^{5}CG). Parameters for d(Br^{5}CGATBr^{5}CG) were taken from Wang *et al.* (9). All other values were calculated using the program NASTE. All values are in degrees, except rise, which is in Å.
[b]Averages for base pair propeller twist and buckle were calculated from the magnitudes of the values listed [e.g. $\langle\kappa\rangle = (\Sigma|\kappa_i|)/6$, where κ_i is the buckle at base pair *i*].

Solvent structure

The structure of Z-DNA in crystals is stabilized by highly conserved networks of water molecules that span the helix in both the major and minor grooves (19; reviewed in 2). A detailed analysis of the perturbations to these networks provides insight into the stability of the DNA duplex (20). In the structure presented here, we observe that the out-of-alternation bases interrupt the hydration pattern typically seen in Z-DNA. This is true in both the major and minor grooves of the Z-DNA duplex.

The standard d(CpG) dinucleotide in the *anti-p-syn* conformation typically shows two solvent motifs that bridge the stacked bases in the duplex. Two hydrogen bonded waters form a bridge by hydrogen bonding to the N4 amino nitrogens of the stacked cytosine bases from opposite strands of the d(CpG) dinucleotide. Similarly, a single water is typically observed to bridge the O6 oxygens of two guanine bases from opposite strands of the d(GpC) dinucleotide step that bridges two d(CpG) dinucleotides. In the out-of-alternation structure presented here, the interactions in the major groove are significantly disrupted at the out-of-alternation base pairs, but not significantly at the surrounding base pairs.

At the central out-of-alternation d(GpC) dinucleotide, water molecules are observed to hydrogen bond to the N7 nitrogen of guanine 3 and the O6 keto oxygen of guanine 9. These two waters

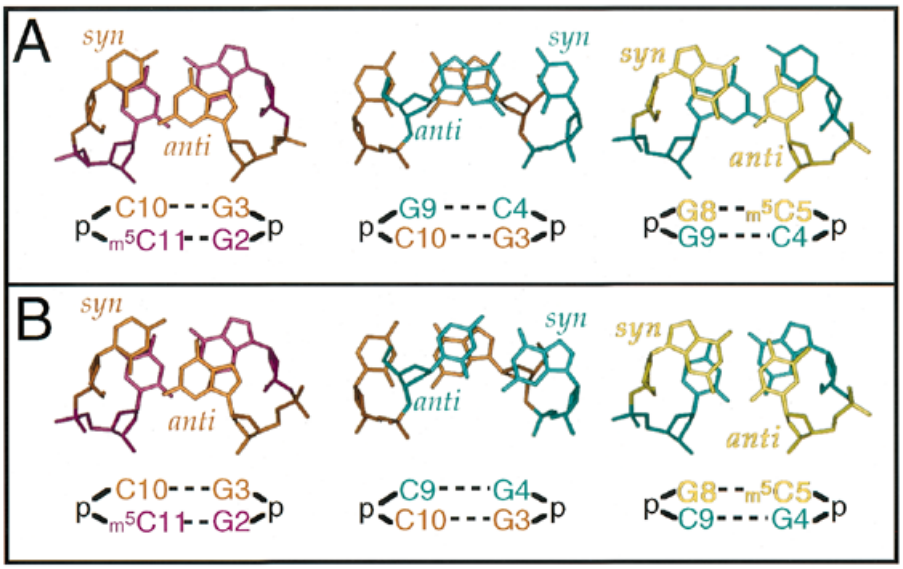

Figure 2. Comparing the stacking of the central base pairs in the crystal structures of d(m^{5}CGGCm^{5}CG)$_2$ (A) and d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) (B). The base pairs are viewed down the helix axes. In each structure, the base pairs are color coded according to their position in the duplex. The dinucleotide steps in the *syn-p-anti*, *anti-p-syn* and *syn-p-anti* arrangements are shown from left to right. (**A**) The out-of-alternation base pairs in d(m^{5}CGGCm^{5}CG)$_2$ show that cytosines C4 and C10 are unstacked and protrude into the major groove. The complementary guanines G3 and G9 of the out-of-alternation base pairs, however, remain stacked in the same manner as cytosine residues of APP *anti-p-syn* dinucleotides. (**B**) The out-of-alternation base pair d(G3•C10) in d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG) also has the cytosine exposed to solvent in the major groove.

are separated by a distance >3.2 Å and, therefore, cannot be considered to be analogous to the hydrogen bonded bridge normally observed for stacked cytosine bases of d(CpG) dinucleotides. At the d(C4pC5)•d(G8pG9) dinucleotide, a single water was observed to bridge the O6 oxygen atoms of guanines at positions 8 and 9. This is similar to the O6-bridging water linking opposite strands at d(GpC) dinucleotides, but differs in that it connects two guanines on the same strand. This interaction is not observed at the opposite end of the duplex at d(G2pG3)•d(C10pC11), which results from differences in lattice interactions between the two halves of the duplex.

The minor perturbations from the normal solvent motifs in the major groove at the flanking d(m^{5}CpG) dinucleotides in the structure here are due primarily to crystal lattice effects and are not inherent in the out-of-alternation sequence. The two water molecules hydrogen bonding to the N4 amino atoms of 5-methylcytosines m^{5}C5 and m^{5}C7 are not themselves hydrogen bonded. One of these waters is slightly displaced relative to the standard hydration pattern of Z-DNA and forms an intermolecular hydrogen bond with the O1P phosphate oxygen of guanine G9 of an adjacent duplex in the lattice. This is similar to the hydration of m^{5}C1 and m^{5}C11, in which both waters contacting the N4 amino nitrogens of the cytosines are also hydrogen bonding to the phosphate oxygen from a symmetry-related duplex. Therefore, the disruption to the waters bridging the cytosines of the d(m^{5}CpG) dinucleotides appears to be induced by crystal lattice interactions and is not relevant to the out-of-alternation base pair effects.

There are two types of water interactions typically observed in the minor groove of Z-DNA. The most striking of these is the continuous hydrogen bonded network that forms a spine of waters lining the minor groove. This spine is defined by hydrogen bonding of the waters to the O2 oxygen of the cytosine bases. In the current structure, the spine of hydration in the Z-DNA minor groove is entirely disrupted by the out-of-alternation base pairs. No waters are observed bound to the O2 keto oxygens of cytosines C4 and C10 in the minor groove of the d(GpC) dinucleotide. This results from the protrusion of the cytosines toward the major groove, which leaves the O2 atoms largely inaccessible to solvent (Fig. 2). Surprisingly, the spine is also disrupted at the flanking d(m^{5}CpG) dinucleotides. Water molecules are observed to hydrogen bond to the O2 oxygens only at 5-methylcytosines m^{5}C7 and m^{5}C11, but not to m^{5}C1 and m^{5}C5.

In addition to the continuous spine of hydration, water molecules are typically observed to help stabilize guanine residues in the *syn* conformation in Z-DNA by forming a bridge from the N2 amino nitrogen of the guanine base to the phosphate oxygens of the backbone. In the structure presented here, the N2 amino nitrogen of guanine G3 is bridged to the O2P phosphate oxygen of cytosine C10 by a single water molecule. This interaction is similar to what is typically observed in APP sequences, but differs in that the bridge occurs across strands. No water interactions, however, were observed at guanine G9 of this same dinucleotide. Finally, all of the stabilizing water bridges between the N2 amino nitrogens of the guanine bases and the O2P oxygens on the backbone are still observed at these flanking dinucleotides.

Differences in the solvent organization between a non-APP sequence and an APP sequence appear to be concentrated primarily at the out-of-alternation base pairs and are not propagated out to the flanking in-alternation base pairs. This was observed in the present structure, as well as in the structure of d(m^{5}CGGGm^{5}CG)•d(m^{5}CGCCm^{5}CG), in which only the solvent structure of the out-of-alternation d(C•G) base pair is disrupted, while that of the flanking in-alternation base pairs remains intact (12). Although slight perturbations to the positions of waters at these flanking base pairs were observed in this structure, they

were not regarded as significant because they could be attributed to either a crystal lattice effect or by limitations in the crystallographic data. How the specific patterns of hydration affect the stability of Z-DNA, however, is not entirely obvious (20).

In order to estimate the effects that these solvent interactions have on the relative stability of Z-DNA versus the standard B-DNA form, we compared solvation free energies for the out-of-alternation base pairs (underlined) in the crystal structures of d(m^{5}CG<u>G</u>Cm^{5}CG)$_2$ and d(m^{5}CG<u>G</u>Gm^{5}CG)•d(m^{5}CGC<u>C</u>m^{5}CG). We had previously shown that the differences in solvation free energy between a duplex in the Z- and B-forms (SFE_{Z-B}) were well correlated with the relative stabilities of APP dinucleotides as Z-DNA (21,22). From the SFE_{Z-B} calculated for d(<u>G</u>pC) and d(<u>G</u>pG)•d(Cp<u>C</u>) (Table 3), we would predict that pairing two adjacent out-of-alternation base pairs should be more stable as Z-DNA than two out-of-alternation base pairs that are isolated along a sequence. The SFE_{Z-B} for a dinucleotide with a single out-of-alternation d(G•C) base pair is more positive than having two tandem base pairs out-of-alternation [as in the dinucleotide step d(GpC)].

Table 3. Solvent free energies (SFEs) calculated for APP and non-APP dinucleotide sequences in an *anti-p-syn* dinucleotide step[a]

Dinucleotide (dn)	SFE (kcal/mol•dn)[b]
d(m^{5}CpG)	–0.87
d(CpG)	0.29
d(CpA)•d(TpG)	0.33
d(<u>G</u>pC)	**0.49**
d(<u>G</u>pG)•d(Cp<u>C</u>)	**0.64**
d(<u>A</u>p<u>T</u>)	**1.28**
d(TpA)	1.35

[a]APP, alternating pyrimidine-purine. Out-of-alternation base pairs are underlined.

[b]SFEs were taken from Kagawa *et al.* (22), except those in bold, which were calculated from hexanucleotides with the sequence d(m^{5}CGNNm^{5}CG)$_2$, where NN refers to the dinucleotide listed. Terminal base pairs were not included in the calculation to eliminate crystallographic end-effects. Values for d(<u>G</u>pC) (this work) and d(GpG)•d(Cp<u>C</u>) (12) were calculated from the crystal structures of these sequences and d(ApT) values were calculated from a model built by replacing the central d(GpC) dinucleotide in d(m^{5}CGGCm^{5}CG)$_2$ with a d(ApT) dinucleotide.

Stabilities of isolated and paired out-of-alternation d(C•G) base pairs in oligomers

There are apparently large differences in the solvent structure and the exposure of the bases in the out-of-alternation base pairs as opposed to the standard base pairs in Z-DNA. To determine whether these differences are significant in affecting the ability of non-APP sequences to form Z-DNA, we compared the ability of isolated and paired out-of-alternation d(G•C) base pairs to form Z-DNA. The results in Table 3 suggest that coupling these non-standard base pairs in Z-DNA greatly reduces the exposure of the unstacked cytosines and, therefore, pairing the out-of-alternation base pairs to form a single non-APP dinucleotide would be less detrimental to the ability of Z-DNA to form than having two isolated out-of-alternation base pairs. Alternatively, if the stability of Z-DNA is affected primarily by the inherent properties of the individual base pairs, for example the high buckle observed in both the isolated and the paired out-of-alternation

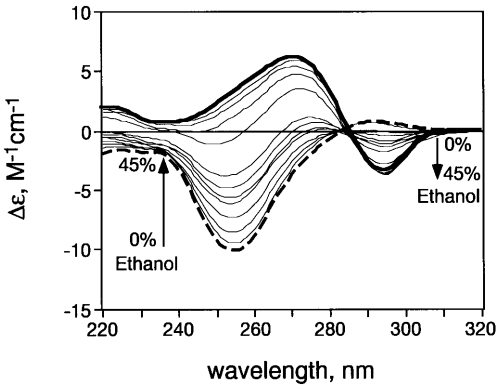

Figure 3. The transition of d[(GC)$_5$GGCC(GC)$_5$]$_2$ from B- to Z-DNA monitored by circular dichroism. Spectra were obtained from samples containing 0.4–0.6 OD oligonucleotide, 1.8 M $MgCl_2$ and ethanol at increasing concentration [0 (dashed curve), 5, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40 and 45% (bold solid curve)]. The CD spectrum at 0% ethanol is indicative of the B-DNA conformation. Inversion of the spectra at high ethanol concentrations with a single isodichroic point indicates a simple two-state transition from B- to Z-DNA.

base pairs, then we would not expect to observe any differences in the behavior of sequences in which two of these base pairs are paired or isolated. In these studies, we compare the ability of salt and ethanol to induce a B- to Z-DNA transition in the 24 bp deoxyoligonucleotides d[(GC)$_{12}$]$_2$, d[(GC)$_3$<u>C</u>CGC(GC)$_2$GC-G<u>G</u>(GC)$_3$]$_2$ and d[(GC)$_5$G<u>G</u>C<u>C</u>(GC)$_5$]$_2$ (where the underlined nucleotides represent positions in the sequence where out-of-alternation base pairs are introduced). The sequence d[(GC)$_{12}$]$_2$ is an APP sequence that should undergo a standard transition to Z-DNA, while d[(GC)$_3$<u>C</u>CGC(GC)$_2$GCG<u>G</u>(GC)$_3$]$_2$ and d[(GC)$_5$G<u>G</u>C<u>C</u>(GC)$_5$]$_2$ have two out-of-alternation base pairs isolated or paired in the sequences, respectively. The transitions were monitored using CD spectroscopy (Fig. 3).

As expected, solutions of d[(GC)$_{12}$]$_2$ under low salt conditions (<1 M $MgCl_2$) produced CD spectra indicative of the B-DNA conformation. As the salt concentrations in the solutions increased, the spectra inverted, consistent with the inversion of the structure to Z-DNA (3,23,24). However, the spectra along the titration did not show a single isodichroic point at 282 nm as the salt concentration was increased. This suggested that the DNA did not undergo a simple two-state transition from B- to Z-DNA. Closer examination of the spectra showed that significant amounts of single-stranded DNA were present in the initial solutions at low salt. The spectra were thus deconvoluted into their double-stranded components by the relationship $CD_{obs} = f_{ss}*(CD_{ss}) + f_{ds}*(CD_{ds})$, where CD_{obs}, CD_{ss} and CD_{ds} refer to the observed, single-strand and double-strand CD spectra, respectively. The transition from B- to Z-DNA could thus be shown to occur with a midpoint at 0.74 M salt for the sequence d[(GC)$_{12}$]$_2$ (Fig. 4a).

The two non-APP sequences d[(GC)$_3$<u>C</u>CGC(GC)$_2$GC-G<u>G</u>(GC)$_3$]$_2$ and d[(GC)$_5$G<u>G</u>C<u>C</u>(GC)$_5$]$_2$, however, were not observed to undergo a B- to Z-DNA transition even at $MgCl_2$ concentrations as high as 3.5 M. This result is consistent with previous biochemical data and the SFE calculations showing that out-of-alternation base pairs greatly destabilize Z-DNA. These sequences were therefore induced to invert to the Z-form with ethanol at a $MgCl_2$ concentration which gave 100% Z-DNA in d[(GC)$_{12}$]$_2$ (Fig. 3). As shown in Figure 4b, the two sequences undergo nearly identical transitions to Z-DNA, with midpoints at

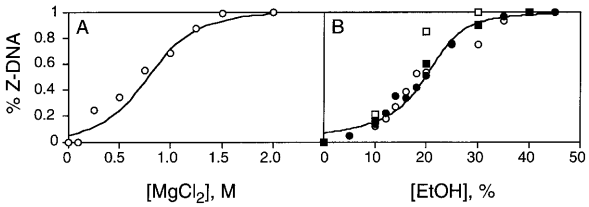

Figure 4. Titration of APP (A) and non-APP (B) 24mer deoxyoligonucleotides with $MgCl_2$ and ethanol to induce Z-DNA formation. The formation of Z-DNA was monitored by following the difference in circular dichroism at 257 and 294 nm ($CD_{257} - CD_{294}$). Data was simulated using a standard statistical mechanics treatment of structural transitions in linear DNA duplexes (31). Circles and squares represent data from two independent experiments for the non-APP sequences. (**A**) Titration of the APP sequence $d[(GC)_{12}]_2$ with $MgCl_2$ shows the midpoint of the transition at 0.74 M salt. (**B**) Non-APP oligomers $d[(GC)_5GGCC(GC)_5]_2$ and $d[(GC)_3CCGC(GC)_2GCGG(GC)_3]_2$ could not be induced to Z-DNA by salts alone. The oligonucleotides $d[(GC)_5GGCC(GC)_5]_2$ (filled symbols), which contains two adjacent out-of-alternation base pairs and $d[(GC)_3CCGC(GC)_2GCGG(GC)_3]_2$ (open symbols), which contains two isolated out-of-alternation base pairs, were titrated with ethanol (EtOH). Midpoints for both transitions were found to occur at 20.3% ethanol. Each sample contained 1.8 M $MgCl_2$, which induces 100% Z-DNA in $d[(GC)_{12}]_2$.

20.3% ethanol. Thus, there is no discernible difference in the ability of sequentially isolated and adjacent out-of-alternation d(G•C) base pairs to form Z-DNA. The SFE calculations predict that stacking adjacent out-of-alternation d(G•C) base pairs into a non-APP dinucleotide should be less deleterious to the stability of the Z-conformation and, therefore, predict that the sequence $d[(GC)_5GGCC(GC)_5]_2$ should form Z-DNA more readily. Since this was not observed, we can conclude that it is the inherent structural distortions to the base pair, more than differences in solvent interactions, that define the inability of out-of-alternation base pairs to adopt the Z-conformation.

DISCUSSION

In the current study, we have solved the structure of $d(m^5CGGCm^5CG)_2$, a sequence that contains two adjacent d(G•C) base pairs that break the pattern of alternating pyrimidine and purine nucleotides typically observed for Z-DNA. It is the cytosine of the out-of-alternation d(G•C) base pairs in this structure that accounts for the perturbations to the Z-DNA structure. This pyrimidine base is buckled out of the plane of the base pair and protrudes into the major groove surface, increasing the exposure of the base to solvent. These structural features were also observed in the out-of-alternation pyrimidine base pairs of the earlier non-APP sequences $d(Br^5CGATBr^5CG)_2$ [which includes two adjacent out-of-alternation d(A•T) base pairs; 9] and $d(m^5CGGGm^5CG)•d(m^5CGCCm^5CG)$ [which contains a single out-of-alternation d(G•C) base pair; 12]. Thus, we can conclude that the distortions are inherent to out-of-alternation base pairs regardless of the base composition or whether the base pairs are isolated or paired within the sequence.

We observed that out-of-alternation d(G•C) base pairs have identical effects on the ability of sequences to undergo a B- to Z-DNA transition regardless of whether these base pairs are isolated or paired. Differences in the stacking of the base pairs and the associated effects on solvent interactions do not appear to play a significant role in defining the ability of out-of-alternation base pairs to form Z-DNA, as they do with standard base pairs in APP sequences (21,22). Thus, the inability of such base pairs to form Z-DNA is a consequence of the inherent structural perturbations to the base pairs, such as the high buckle caused by the steric constraints of placing a pyrimidine in the *syn* conformation.

When trying to assess the ability of a sequence to form Z-DNA, we can now reduce the thermodynamic rules to a set of simple nucleotide propensities (Table 4), even though the basic repeating unit of the Z-DNA structure is a dinucleotide (2 bp). This assumes that the basic structural motif of alternating *anti-p-syn* conformations remains intact along the sequence (1,2). This greatly simplifies the procedures designed to analyze the ability of sequences to form Z-DNA in various genomes (25,26), including the human genome (27).

Table 4. Z-DNA propensities in terms of the free energy to induce a transition from B- to Z-DNA (ΔG°_T) for d(C•G) and d(A•T) base pairs in their *anti•syn* and *syn•anti* conformations[a]

Base pair (bp)	Conformation	ΔG°_T (kcal/mol•bp)
d(C•G)[b]	*anti•syn*	0.3
d(T•A)[c]	*anti•syn*	1.1
d(C•G)[d]	*syn•anti*	2.1
d(T•A)[e]	*syn•anti*	2.2

[a]B–Z transition free energies (ΔG°_T) were calculated from those values determined from negatively supercoiled ccDNA.
[b]Calculated as $0.5\Delta G^\circ_T$ of the d(CpG) dinucleotide (7).
[c]Calculated as $0.5\Delta G^\circ_T$ for d(TpA) (29,30) and ΔG°_T of d(CpA)•d(TpG) (26) minus ΔG°_T of a d(C•G) base pair.
[d]Calculated as ΔG°_T of the d(CpC)•d(GpG) dinucleotide minus ΔG°_T of a d(C•G) base pair (5).
[e]Calculated as ΔG°_T of the d(CpT)•d(ApG) dinucleotide minus ΔG°_T of a d(C•G) base pair (5).

ACKNOWLEDGEMENTS

This work has been supported by grants from the National Institute of Health (RO5GM54538A), the National Science Foundation (MCB-9728240) and the National Institute of Environmental Health Sciences (ES00210). We wish to thank Dr Blaine Mooers for his suggestions during the structure refinement and Jeannine Lawrence for help with the circular dichroism spectroscopy.

REFERENCES

1 Wang,A.H.-J., Quigley,G.J., Kolpak,F.J., Crawford,J.L., van Boom,J.H., van der Marel,G.A. and Rich,A. (1979) *Nature*, **282**, 680–686.
2 Ho,P.S. and Mooers,B.H.M. (1997) *Nucleic Acid Sci.*, **1**, 65–90.
3 Pohl,F.M. and Jovin,T.M. (1972) *J. Mol. Biol.*, **67**, 375–396.
4 Jovin,T.M., Soumpasis,D.M. and McIntosh,L.P. (1987) *Annu. Rev. Phys. Chem.*, **38**, 521–560.
5 Ellison,M.J., Kelleher,R.J., Wang,A.H.-J., Habener,J.F. and Rich,A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 8320–8324.
6 McLean,M.J., Lee,J.W. and Wells,R.D. (1988) *J. Biol. Chem.*, **263**, 7378–7385.
7 Peck,L.J. and Wang,J.C. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 6206–6210.
8 Lukomski,S. and Wells,R.D. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 9980–9984.
9 Wang,A.H.-J., Gessner,R.V., van der Marel,G.A., van Boom,J.H. and Rich,A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 3611–3615.
10 Feigon,J., Wang,A.H.-J., van der Marel,G.A., van Boom,J.H. and Rich,A. (1985) *Science*, **230**, 82–84.
11 Haschemeyer,A.E.V. and Rich,A. (1967) *J. Mol. Biol.*, **27**, 369–384.
12 Schroth,G.P., Kagawa,T.F. and Ho,P.S. (1993) *Biochemistry*, **32**, 13381–13392.

13 Parkinson,G.N., Vojtechovsky,J., Clowney,L. and Berman,H.M. (1996) *Acta Crystallogr.*, **D52**, 57–64.
14 Brünger,A. (1992) *X-PLOR Manual, Version 3.1*. Yale University Press, New Haven, CT.
15 Luzzati,P.V. (1952) *Acta Crystallogr.*, **5**, 802–810.
16 Basham,B., Eichman,B.F. and Ho,P.S. (1998) In Neidle,S. (ed.), *The Oxford Handbook of Nucleic Acid Structure*. Oxford University Press, Oxford, UK, Vol. I, pp. 199–252.
17 Berman,H.M., Olson,W.K., Beveridge,D.L., Westbrook,J. Gelbin,A., Demeny,T., Hsieh,S.-H., Srinivasan,A.R. and Schneider,B. (1992) *Biophys. J.*, **63**, 751–759.
18 Gessner,R.V., Frederick,C.A., Quigley,G.J., Rich,A. and Wang,A.H.-J. (1989) *J. Biol. Chem.*, **264**, 7921–7935.
19 Gessner,R.V., Quigley,G.J. and Egli,M. (1994) *J. Mol. Biol.*, **236**, 1154–1168.
20 Berman,H.M. (1994) *Curr. Opin. Struct. Biol.*, **4**, 345–350.
21 Kagawa,T.F., Stoddard,D., Zhou,G. and Ho,P.S. (1989) *Biochemistry*, **28**, 6642–6651.
22 Kagawa,T.F., Howell,M.L., Tseng,K. and Ho,P.S. (1993) *Nucleic Acids Res.*, **21**, 5978–5986.
23 Arnott,S., Chandrasekaran,R., Birdsall,D.L., Leslie,A.G.W. and Ratliff,R.L. (1980) *Nature* (*Lond.*), **283**, 743–745.
24 Patel,D.J., Canuel,L.L. and Pohl,F.M. (1979) *Proc. Natl Acad. Sci. USA*, **76**, 2508–2511.
25 Ho,P.S., Frederick,C.A., Quigley,G.J., van der Marel,G.A., van Boom,J.H., Wang,A.H.-J. and Rich,A. (1985) *EMBO J.*, **4**, 3617–3623.
26 Vologodskii,A.V. and Frank-Damenetskii,M.D. (1984) *J. Biomol. Struct. Dyn.*, **1**, 1325–1333.
27 Schroth,G.P., Chou,P.-J. and Ho,P.S. (1992) *J. Biol. Chem.*, **267**, 11846–11855.
28 Fujii,S., Wang,A.H.-J., van der Marel,G., van Boom,J.H. and Rich,A. (1982) *Nucleic Acids Res.*, **10**, 7879–7892.
29 Ellison,M.J., Feigon,J., Kelleher,R.J., Wang,A.H.-J., Habener,J.F. and Rich,A. (1986) *Biochemistry*, **25**, 3648–3655.
30 McLean,M.J., Blaho,J.A., Kilpatrick,M.J. and Wells,R.D. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5884–5888.
31 Ivanov,V.I. and Krylov,D.Y. (1992) *Methods Enzymol.*, **211**, 111–127.