Article No. jmbi.1999.2934 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 291, 83–99

JMB

Atomic-resolution Crystal Structures of *B*-DNA Reveal Specific Influences of Divalent Metal Ions on Conformation and Packing

George Minasov, Valentina Tereshko and Martin Egli*

Department of Molecular Pharmacology and Biological Chemistry and The Drug Discovery Program, Northwestern University Medical School, Chicago, IL, 60611-3008, USA

Crystal structures of *B*-form DNA have provided insights into the global and local conformational properties of the double helix, the solvent environment, drug binding and DNA packing. For example, structures of the duplex with sequence CGCGAATTCGCG, the Dickerson-Drew dodecamer (DDD), established a unique geometry of the central A-tract and a hydration spine in the minor groove. However, our knowledge of the various interaction modes between metal ions and DNA is very limited and almost no information exists concerning the origins of the different effects on DNA conformation and packing exerted by individual metal ions.

Crystallization of the DDD duplex in the presence of Mg^{2+} and Ca^{2+} yields different crystal forms. The structures of the new Ca^{2+}-form and isomorphous structures of oligonucleotides with sequences GGCGAATTCGCG and GCGAATTCGCG were determined at a maximum resolution of 1.3 Å. These and the 1.1 Å structure of the DDD Mg^{2+}-form have revealed the most detailed picture yet of the ionic environment of *B*-DNA. In the Mg^{2+} and Ca^{2+}-forms, duplexes in the crystal lattice are surrounded by 13 magnesium and 11 calcium ions, respectively.

Mg^{2+} and Ca^{2+} generate different DNA crystal lattices and stabilize different end-to-end overlaps and lateral contacts between duplexes, thus using different strategies for reducing the effective repeat length of the helix to ten base-pairs. Mg^{2+} crystals allow the two outermost base-pairs at either end to interact laterally *via* minor groove H-bonds, turning the 12-mer into an effective 10-mer. Ca^{2+} crystals, in contrast, unpair the outermost base-pair at each end, converting the helix into a 10-mer that can stack along its axis. This reduction of a 12-mer into a functional 10-mer is followed no matter what the detailed nature of the 5′-end of the chain: C-G-C-G-A-..., G-G-C-G-A-..., or a truncated G-C-G-A-... Rather than merely mediating close contacts between phosphate groups, ions are at the origin of many well-known features of the DDD duplex structure. A Mg^{2+} coordinates in the major groove, contributing to kinking of the duplex at one end. While Ca^{2+} resides in the minor groove, coordinating to bases *via* its hydration shell, two magnesium ions are located at the periphery of the minor groove, bridging phosphate groups from opposite strands and contracting the groove at one border of the A-tract.

Keywords: crystal packing; DNA bending; metal ions; hydration; X-ray crystallography

*Corresponding author

Introduction

X-ray crystallography and solution NMR have allowed visualization of the remarkable conformational versatility and deformability of DNA. Beyond the double-helix families (Kennard & Hunter, 1991; Dickerson, 1992; Egli, 1994), these techniques have provided insights on triplexes (Sklenar & Feigon, 1990; Van Meervelt *et al.*, 1995; Vlieghe *et al.*, 1996), tetraplexes (Kang *et al.*, 1992;

E-mail address of the corresponding author: m-egli@nwu.edu

0022-2836/99/310083–17 $30.00/0

Laughlan *et al*., 1994) and parallel-stranded arrangements (Gehring *et al*., 1993; Chen *et al*., 1994; Berger *et al*., 1996). Moreover, proteins often cause dramatic distortions of the canonical *B*-form duplex geometry to recognize (Steitz, 1993; reviewed by Allemann & Egli, 1997), process (Kim *et al*., 1990; Winkler *et al*., 1993; Klimasauskas *et al*., 1994) or efficiently pack DNA (Luger *et al*., 1997). Despite the fact that metal ions are a basic component of all such structural studies, X-ray crystallography has only rarely revealed substantial numbers of ions. High-resolution data were shown to be of crucial importance in this respect, but have remained the exception with crystals of DNA fragments so far (Gessner *et al*., 1989; Bancroft *et al*., 1994; Laughlan *et al*., 1994). Although metal ion coordination can drastically alter the geometry of the DNA double helix (Takahara *et al*., 1996), the conformational consequences of metal ion coordination can be expected to be rather subtle in many cases and only very precise structures in combination with the corresponding reference data will then allow their detection.

Most crystallographic studies of *B*-form DNA were conducted with dodecamer and decamer duplexes (Grzeskowiak, 1996). The resolutions of dodecamer structures were generally lower than 2.2 Å, while the more tightly packed crystals of decamer duplexes yielded data with resolutions of up to 1.3 Å. More recent structures of the Dickerson-Drew dodecamer (DDD) duplex with sequence CGCGAATTCGCG at resolutions of around 1.5 Å revealed one bound Mg^{2+} per duplex (Berger *et al*., 1998; Shui *et al*., 1998). Similarly, single metal ions were discovered in the lattice of DNA decamer crystals (Privé *et al*., 1991). The crystal structure of an octamer duplex with overhanging G residues (sequence GCGAATTCG, comprising the central octamer of the DDD) solved to 2 Å resolution featured two magnesium ions (Van Meervelt *et al*., 1995). One of the ions is located in the vicinity of the base triples that are formed between terminal C-G base-pairs and the dangling G bases. Using improved crystallization protocols and third-generation synchrotron radiation, we recently managed to improve the resolution of the orthorhombic Mg^{2+}-form of the DDD duplex to 1.1 Å (Tereshko *et al*., 1999a). The current resolution of the structure is 0.95 Å with 95 % completeness of the data (G.M., V.T. & M.E., unpublished results). These data furnished five magnesium ions per crystallographic asymmetric unit and each duplex is thus surrounded by 13 divalent metal ions. This structure will serve as our reference in the analysis of the dependence of DNA conformation and packing on metal ion coordination and for simplicity we will refer to it as the CGMg dodecamer.

Crystallization of the same DNA oligomer in different crystal lattices can provide information regarding the intrinsic geometry of a DNA duplex and its tendency to be deformed by the packing forces of a particular lattice. Comparisons between the helix structures of methylated and unmethylated decamers that crystallized in two different space groups showed a correlation between the extent of the geometric variations and the degree of local changes between the crystallographic environments (Heinemann & Alings, 1989, 1991; Heinemann & Hahn, 1992). Because the packing arrangements of stacked duplexes in the two different space groups showed close resemblance, the local helix parameters were virtually identical. Conversely, more extensive deviations in the packing modes between two lattices triggered changes in the local helical geometries (Grzeskowiak *et al*., 1991; Baikalov *et al*., 1993). Studies that allowed a comparison of the helix geometry of one and the same *B*-form DNA duplex in two different space groups were conduced for the decamer CGAACITTCG (I = inosine) (Lipanov *et al*., 1993) and for the decamer CGCAATTGCG (Spink *et al*., 1995; Wood *et al*., 1997). In the first study, the different packings of duplexes in the monoclinic and trigonal crystal lattices led to different local geometries and demonstrated that some DNA sequences may be easily deformable. Interestingly, the two crystal forms were obtained by crystallizing the DNA decamer in the presence of either magnesium or calcium acetate, while the concentrations of the other crystallization ingredients were virtually identical in the two cases. The resolutions of the monoclinic Ca^{2+}-form and the trigonal Mg^{2+}-form structures were 1.3 Å and 2.2 Å, respectively, and only one divalent metal ion per structure could be found. Therefore, while metal cations are obviously an important determinant of DNA packing, the details of how they affect packing and possibly conformation remain to be worked out.

For a number of years, we had observed growth of a new crystal form for the CGCGAATTCGCG oligomer in the presence of Ca^{2+}. For example, crystallization screens for DDD duplexes with incorporated carbocyclic A and T residues had yielded a rhombohedral Ca^{2+} crystal form of the dodecamer (Egli, 1996; Portmann *et al*., 1997). However, the standard orthorhombic Mg^{2+} crystal form was not obtained under the familiar magnesium acetate/spermine conditions. At the time, these Ca^{2+}-form crystals were not subjected to a detailed crystallographic investigation, since they diffracted to only relatively low resolution (ca 3 Å). Similar observations were made by others and the structure of the rhombohedral crystal form of the dodecamer at 3 Å resolution was reported (Liu *et al*., 1998). More recently, the native CGCGAATTCGCG oligomer and dodecamers containing 2′-deoxy-2′-fluoro-arabinofuranosyl thymine residues (Berger *et al*., 1998) were observed to crystallize in a rhombohedral crystal form (called CGCa here) when magnesium acetate in the crystallization medium was replaced by calcium chloride. Initial analysis of the structure indicated the lack of electron density for 5′-terminal cytidine bases in these crystals. Consequently, the 5′-terminal C in the DDD sequence was replaced by G to examine whether a different residue would

lead to more ordered duplex ends. Although the resolution of the Ca^{2+}-form crystals of this GGCGAATTCGCG dodecamer (called GGCa here) was improved relative to the CGCa structure, the 5′-terminal G residues were not defined in the electron density maps. Finally, it was tested whether the 11mer GCGAATTCGCG (called GCCa here) would yield Ca^{2+}-form crystals. This was indeed the case and in the structure refined to 1.3 Å resolution, six Ca^{2+} per crystallographic asymmetric unit were observed. Thus, each duplex is surrounded by 11 divalent metal ions in the GCCa lattice. The GCCa structure is not strictly isomorphous with those of CGCa and GGCa, although the arrangement of duplexes in the former displays close similarity to that in the other two crystals.

The availability of very high resolution data for two different crystal forms of the DDD duplex, 1.1 Å and 1.3 Å, respectively, allows an assessment of the dependence of its particular features on the crystal environment. The narrow minor groove in the A-tract, asymmetric kinking of the duplex, tight contacts between phosphate groups and the hydration spine are of particular interest in this respect. The growth of the two different crystal forms is directly related to the use of either Mg^{2+} or Ca^{2+} in the crystallizations. With more ions observed in both lattices than ever before with a *B*-form DNA, it becomes possible for the first time to examine at atomic resolution the basis for the stabilization of different lattices by divalent metal ions. Because the packing arrangements in the two lattices bear some resemblance, in both of them duplexes form semi-continuous rods that are aligned in parallel, differences between the helical geometries may be coupled with specific ion-DNA interactions. Here, we report the structure of the new rhombohedral Ca^{2+}-form of the DDD duplex based on analyses of the CGCa, GGCa and GCCa oligomers. The ion-DNA interactions in the high-resolution CGMg and GCCa structures are compared and the observed differences between the geometries of the duplex in the two lattices are analyzed, taking into account the coordinated divalent metal ions.

Results and Discussion

Geometry of the Ca^{2+}-form helix

Three crystal structures of the $[d(CGCGAATTCGCG)]_2$ duplex (CGCa) and duplexes formed by two related oligomers, one with a C to G mutation at the 5′ terminus (GGCa) , the other lacking the 5′-terminal cytidine base (GCCa), all grown in the presence of Ca^{2+} instead of Mg^{2+}, were determined at resolutions of 2.2, 1.7 and 1.3 Å, respectively (see Table 4). Because of its higher resolution, the ensuing analysis and discussion of the Ca^{2+}-form DDD duplex will be based on the GCCa structure. An example of the quality of the $(2F_o - F_c)$ sum electron density surrounding the final GCCa model is depicted in Figure 1. The uniqueness of this study lies in the availability of two crystal forms of the same DNA oligonucleo-

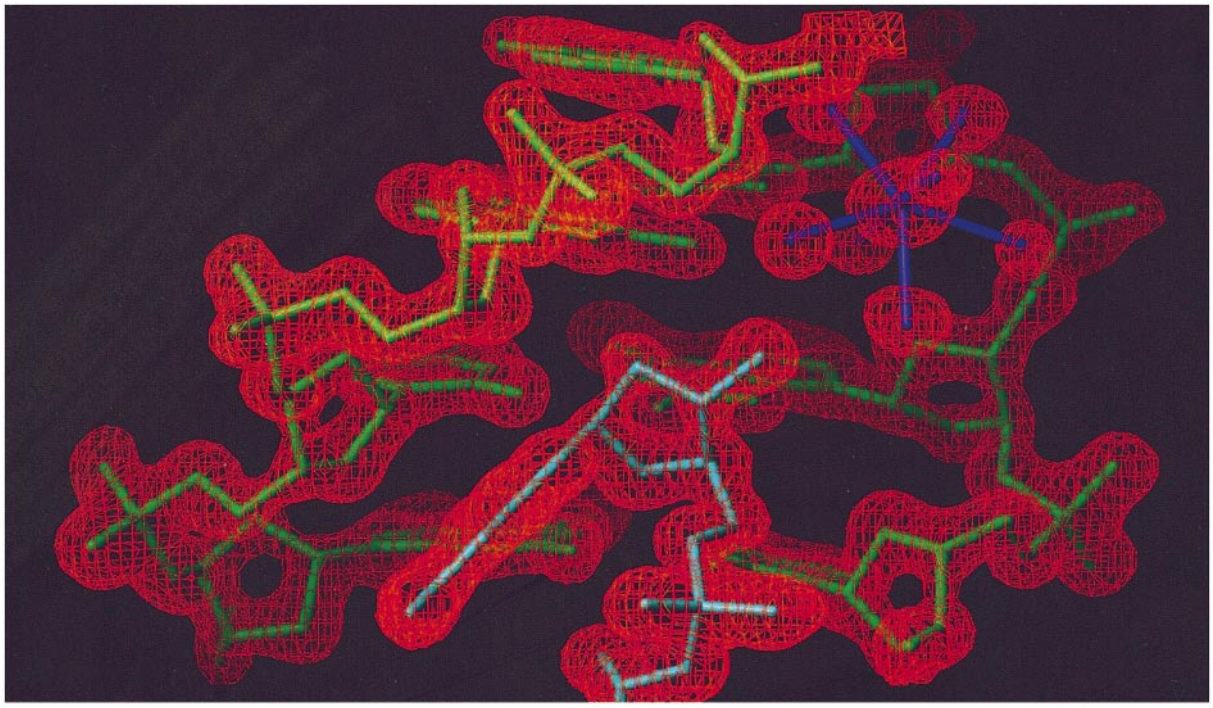

Figure 1. Quality of the GCCa structure. The $(2F_o - F_c)$ sum electron density is contoured at 1.5 σ surrounding a tetramer portion of the duplex (green). The view is into the minor groove and illustrates the insertion of a 3′-terminal G from a stacked neighboring duplex (cyan) into the groove. A Ca^{2+} (Ca3) with seven coordinated water molecules bound in the minor groove is drawn with dark blue bonds.

tide whose structures were determined at high resolution and have revealed unprecedented details of the ionic environment of a *B*-form DNA. Therefore, the properties of the Ca^{2+}-form structure will be discussed in the context of the 1.1 Å CGMg reference structure throughout the following sections.

In the rhombohedral GCCa crystal, a dimer made up of 22 residues constitutes the crystallographic asymmetric unit. The 11mer GCGAATTCGCG forms a decamer duplex with G overhangs at both 3′ termini. To simplify comparison with the reference dodecamer, we have numbered the residues in the 11mer from G2 to G12 for strand 1 and from G14 to G24 in strand 2. The G12 and G24 residues are unstacked from the neighboring C11-G14 and G2-C23 base-pairs, respectively, and are inserted into the minor grooves of adjacent duplexes from the same helix column (Figure 1). The geometry of these base-pairs appears only moderately affected by the looped out G residues. The G2-C23 pair displays the strongest buckling (−15°) and inclination (3.1°) among all base-pairs and a propeller twist of 8°, while the C11-G14 pair at the other end remains more or less planar. Selected helical parameters for the GCCa and CGMg structures are shown in Figure 2.

There is no drastic difference in the geometries of the $[d(GCGAATTCGC)]_2$ decamer portions in the two duplexes. For example, in both of them the G4-C21 and G16-C9 base-pairs located at the border of the A-tract are characterized by considerable buckling (ca 12°). The helical twists between them and the neighboring base-pairs A5-T20 and A17-T8, respectively, are above average, while their arrangements relative to the other neighbors C3-G22 and C15-G10, respectively, are characterized by a low twist and a significant slide. A further feature shared by the GCCa and CGMg duplexes is the large propeller twist in the six central base-pairs, which is accompanied by the narrowing of the minor groove in that region (Figure 3). Overall, the GCCa duplex displays a higher internal sym-

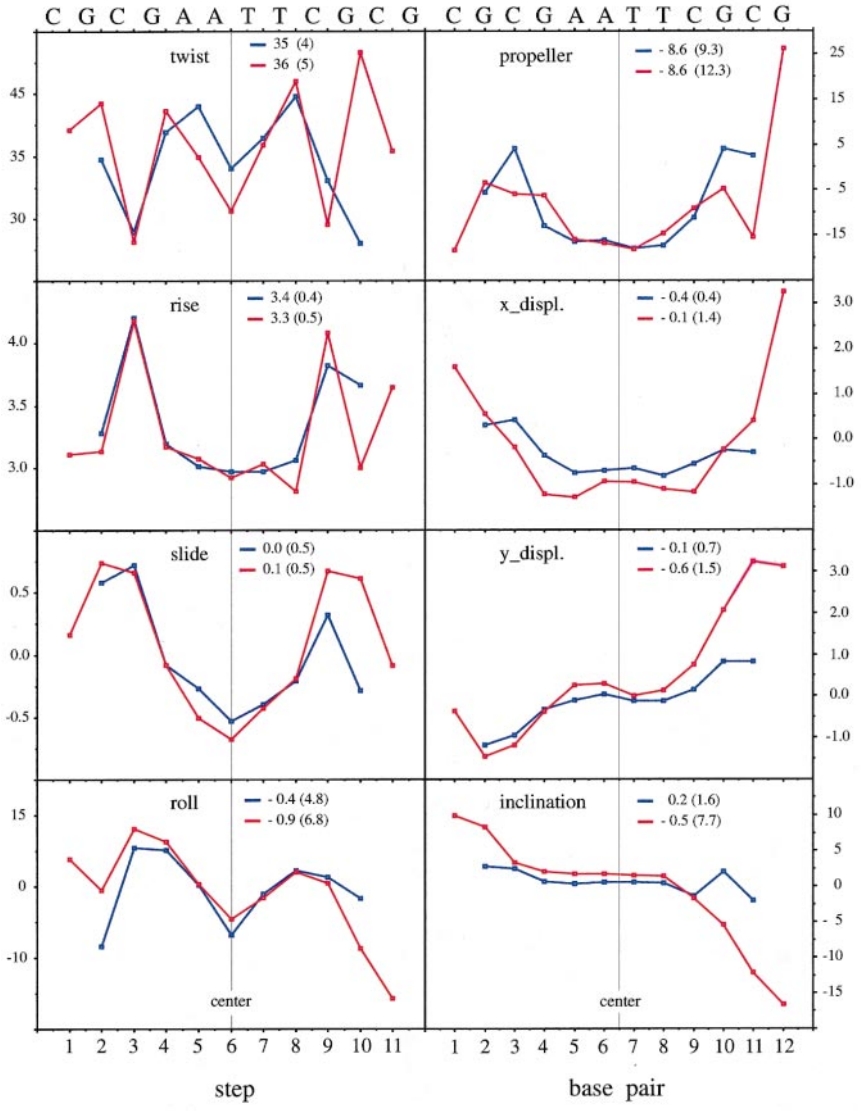

Figure 2. Helical parameters for the GCCa (blue) and CGMg (red) duplexes. Rise, slide, *x*-displacement and *y*-displacement are in Å, and twist, propeller twist, roll and inclination are in degrees. Average values and standard deviations for individual parameters in both duplexes are listed at the upper right. The overall helical axis was calculated based on the central hexamer duplexes, using C1′ and N1 and N9 atoms of pyrimidines and purines, respectively. All parameters were calculated with the program NEWHELIX93, distributed by Dr Richard E. Dickerson.

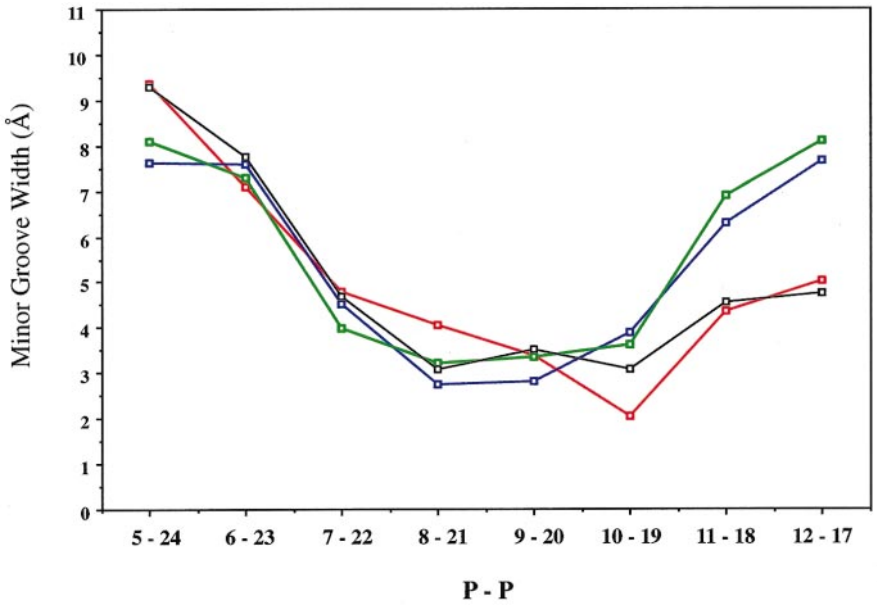

Figure 3. Minor groove width. Comparison between the minor groove widths of the GCCa (blue), the CGMg (red), the GGCa (green) duplexes and a DDD duplex, crystallized at high spermine concentration (Shui *et al.*, 1998) (black, Nucleic Acid Database (NDB) code BDL084 (Berman *et al.*, 1992)).

metry compared with the CGMg one. The r.m.s. (root-mean-square) deviation between the two strands in the GCCa duplex (0.43 Å) is three times smaller than that between strands in the CGMg one (1.29 Å). The r.m.s. deviation between the central decamer duplexes of the GCCa and CGMg structures is 0.84 Å for one relative orientation and 0.99 Å for the other. Considering only base atoms, the deviations for the two possible relative orientations are 0.66 Å and 0.70 Å. Similarly, the r.m.s. deviations between the central hexamer duplexes in the two orientations amount to 0.51 Å and 0.54 Å, respectively, and, considering only base atoms, these values are reduced to 0.31 Å and 0.34 Å. The relatively small deviations demonstrate the similar conformations adopted by the DDD duplex (not taking into account terminal residues) in the orthorhombic and rhombohedral lattices. A superposition of the GCCa and CGMg duplexes is depicted in Figure 4.

Despite different lateral contacts between duplexes in the two lattices, certain features in the sugar-phosphate backbone geometries are retained. The sugar pucker ranges observed in the two structures are listed in Table 1. Thus, the deoxyribose moieties of residues C3 in both duplexes adopt an Eastern pucker, with pseudorotation phase angle amplitudes of 78° and 53° in GCCa and CGMg, respectively. In addition, there are residues with non-standard puckers at the 5′-end of strand 2 in both duplexes. The phosphate group of residue C23 displays a B_{II} geometry (Dickerson *et al.*, 1982) in both lattices (data not shown), while P11 adopts a B_{II} conformation only in the CGMg duplex. In the GCCa duplex all other phosphate groups show the more common B_I conformation, and in the CGMg structure, the backbone around P14 also displays the B_{II} geometry. An interesting difference between the geometries of the GCCa and CGMg duplexes is the asymmetric kink in the latter (Figure 4). Both duplexes show positive rolls at

Table 1. Sugar puckers in the GCCa and CGMg duplexes

Residues	GCCa		CGMg	
C1/G24	-	C2′-*exo*	C2′-*endo*	C3′-*endo*
G2/C23	C2′-*endo*	C2′-*endo*	C2′-*endo*	C1′-*exo*
C3/G22	O4′-*endo*	C2′-*endo*	C4′-*exo*	C2′-*endo*
G4/C21	C2′-*endo*	C1′-*exo*	C2′-*endo*	C1′-*exo*
A5/T20	C2′-*endo*	C1′-*exo*	C2′-*endo*	C1′-*exo*
A6/T19	C2′-*endo*	C1′-*exo*	C1′-*exo*	O4′-*endo*[a]
T7/A18	C1′-*exo*	C2′-*endo*	O4′-*endo*[a]	C1′-*exo*
T8/A17	C1′-*exo*	C2′-*endo*	C1′-*exo*	C2′-*endo*
C9/G16	C1′-*exo*	C2′-*endo*	C2′-*endo*	C2′-*endo*
G10/C15	C2′-*endo*	O4′-*endo*	C2′-*endo*	C4′-*exo*
C11/G14	C2′-*endo*	C3′-*exo*	C2′-*endo*	O4′-*endo*
G12/C13	C2′-*exo*	-	O4′-*endo*	C2′-*endo*

Sugar puckers were calculated with the program CURVES (Lavery & Sklenar, 1989).
[a] Specific pucker is preferred by 2′-deoxy-2′-fluoro-arabinofuranosyl thymine (Berger *et al.*, 1998).

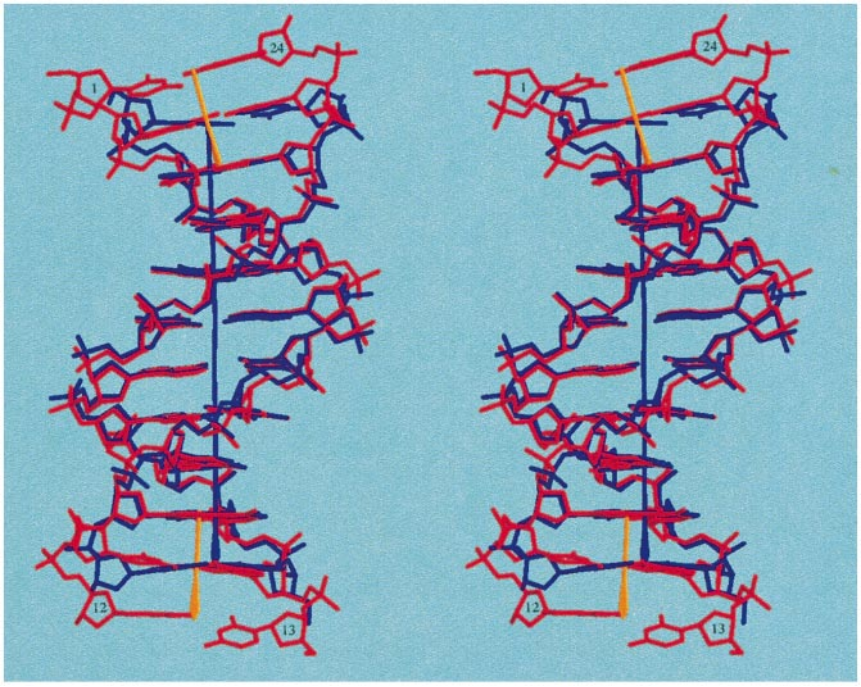

Figure 4. Superposition of the GCCa (blue, only ten base-pairs shown) and CGMg duplexes (red), corresponding to the relative orientation with the smallest r.m.s deviation (see the text). The helix axes based on the central hexamer duplexes practically coincide (straight thick blue line). In the CGMg duplex, the helix axis (orange) defined by base-pairs G10-C15, C11-G14 and G12-C13 at one end is straight and parallel with that of the central hexamer (bottom). Conversely, the helix axis defined by base-pairs C1-G24, G2-C23 and C3-G22 at the other end is inclined relative to the helix axis of the hexamer and illustrates the Mg^{2+}-induced asymmetric kink of the DDD duplex in the orthorhombic lattice (top). All calculations were conducted with the program CURVES (Lavery & Sklenar, 1989).

their C3pG4 and G4pA5 steps (Figure 2). While this conformational feature is offset by a negative roll at the G2pC3 step in the GCCa duplex, it manifests itself in the form of a kink in the case of the CGMg duplex. At the other end of the duplex, at the T8pC9 and C9pG10 step, the base rolls are only minor.

End-to-end overlaps between duplexes in the Ca^{2+} and Mg^{2+}-form crystals

In the Ca^{2+}-form lattice, duplexes form infinite columns along the crystallographic *c*-axis, with more-or-less continuous major and minor grooves (Figures 5(a) and 6(a)). At the duplex interface, the terminal base-pairs G2-C23 and C11-G14 are stacked with a relatively large rise (3.8 Å, for average values see Figure 2) and a large slide (2.7 Å, Figure 6(b)). In the observed packing mode in a rhombohedral lattice, each duplex must undergo a net rotation of $\pm 360^\circ/3 = \pm 120^\circ$. Hence the inter-helix step has to be $\pm 120^\circ$ - 310.5° (overall twist of nine base-pairs with an average twist of 34.5°, Figure 2) $= -70.5^\circ$ (modulo 360°). The minor groove is widened at the duplex interface and the 3′-terminal G residues from opposite strands of two adjacent duplexes are inserted into the groove. As shown in Figure 5(a), the non-crystallographic dyad located between the stacked terminal base-pairs is maintained by the particular orientations adopted by G12 and G24. The backbones are curled back, rotating the guanine bases such that their N2/N3 edges face the equivalent side of residues G2 and G14, respectively, at the stacking interface. The resulting C-G*G triples are stabilized by two N2···N3 hydrogen bonds each between guanine residues (average length 3.0 Å) . The G*G pairs are characterized by considerable propeller twisting, allowing the deoxyribose moieties of residues C3 and C15 to stack on the base planes of G12 and G24, respectively (Figure 5(a)). In addition, the 4′-oxygen atoms of inserted G12 and G24 residues form weak hydrogen bonds to the N2 atoms of residues G2 and G14, respectively (average length 3.3 Å). The intricate network of hydrogen bonds stabilizing the two guanosine residues in the minor groove is completed by interactions between the N1 atoms of residues G12 and G24 and O2P oxygen atoms of the preceding residues C11 and C23, respectively (average length 2.9 Å). Very similar arrangements of terminal guanosines as seen in GCCa were observed in the crystal structures of the *B*-DNA decamer CGCAATTGCG (Spink *et al.*, 1995; Nunn *et al.*, 1997; Wood *et al.*, 1997).

The most striking feature of the end-to-end interaction between duplexes in the GCCa lattice is the extensive stabilization provided by metal ions. As shown in Figure 5(a), no fewer than six calcium ions form a bracket around the duplex ends, linking the four strands and spanning the entire minor groove. Only four of the six metal ions (numbered from 1 to 6) present per asymmetric unit actually participate in the stabilization, namely Ca1, Ca2, Ca3 and Ca6. The two remaining ions coordinating at the duplex interface are symmetry mates of Ca1 (Ca1*a*, symmetry mates are designated with lowercase letters *a*, *b*, *c*, etc.) and Ca2 (Ca2*a*) (Figure 5(a)). Ca2 and Ca3 are heptacoordinated and the other two are hexacoordinated. Table 2 provides details for the coordination spheres of calcium ions and lists distance data for the interactions between ions and DNA in the GCCa lattice. Ca1 is located on the non-crystallographic dyad that relates stacked

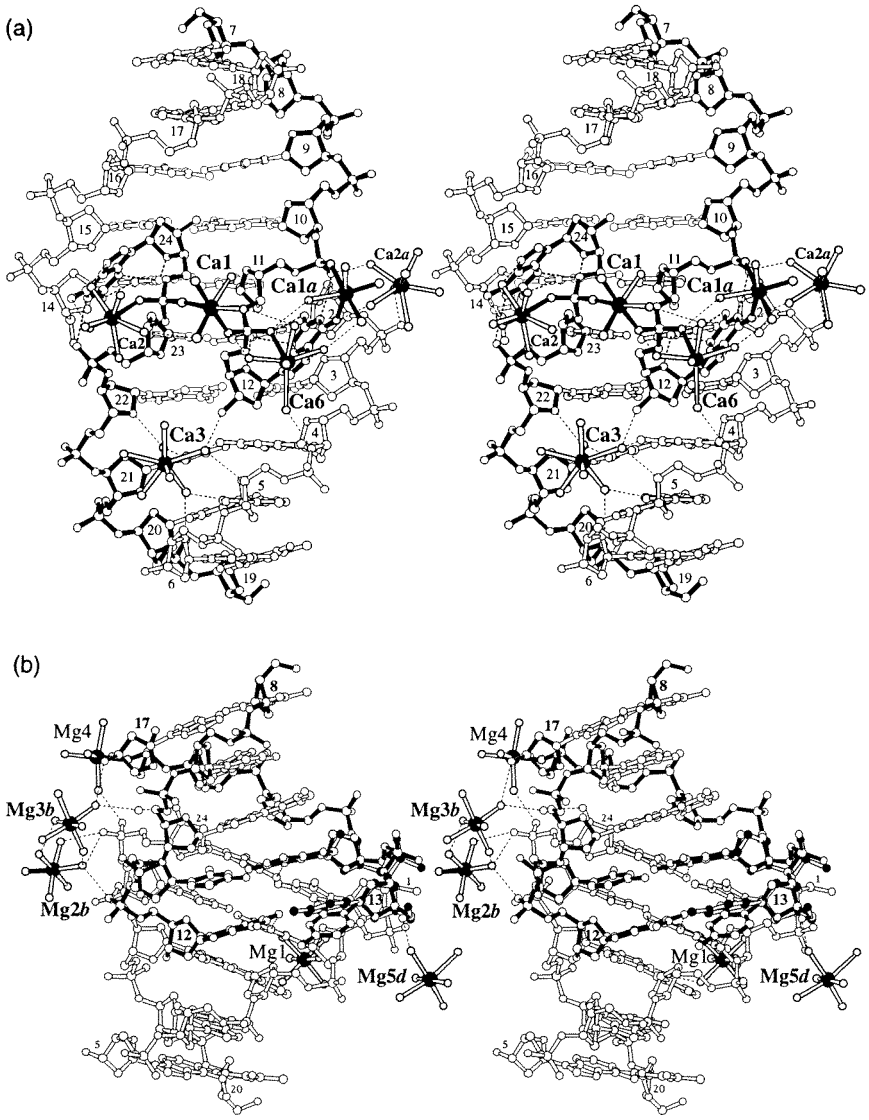

Figure 5. Stereo views of the end-to-end overlaps between duplexes in (a) the GCCa and (b) the CGMg lattice. Backbone atoms of strands inserting their 3′-terminal G residues into the minor grooves of adjacent GCCa duplexes are drawn with filled bonds. In the CGMg lattice, one duplex is drawn with filled and the other one with open bonds. Divalent metal ions are indicated by filled circles and the coordinating water molecules are drawn with smaller circles. Metal ions are labelled and their symmetry mates are designated with lower-case letters. Filled bonds indicate direct ion-DNA contacts, broken lines indicate hydrogen bonds and selected residues are labelled. The orientation of the C13 residue in the low Mg^{2+}-form duplex (Berger *et al.*, 1998) is shown with open bonds (see the text).

duplexes in the GCCa lattice (Figure 5(a)). Among others, Ca1 mediates a close contact between phosphate groups P12 and P24 from two different duplexes. Ca2 stabilizes the backfolding of G24 on one side of the minor groove, while Ca1*a* and Ca6 fulfil this role with G12 on the other side. Calcium ions are inserted between the four strands in such a way that phosphate groups are involved in either direct or water-mediated contacts to ions *via* both their O1P and O2P oxygen atoms. While a coordination mode that relieves close contacts between phosphate groups prevails, several of the ion-DNA interactions involve base atoms. It is noteworthy that Ca1 is engaged in four inner-sphere contacts to DNA, three to phosphate oxygen atoms and a fourth one to O6 of guanine (Figures 5(a) and 6(a)). Compared with Ca1, Ca1*a*, Ca2, Ca2*a* and Ca6 that traverse the minor groove, Ca3 is the only ion that is located in a groove. It is coordinated at the boundary between the A-tract and the GCG section at one end of the duplex (Figures 1 and 5(a)). In addition to contacting three 4′-oxygen atoms and a terminal hydroxyl group from a third strand, Ca3 also forms hydrogen bonds to the N2 and O2 atoms of the C21-G4 base-pair and to N3 of A5 *via* coordinated water molecules (Table 2).

In the orthorhombic Mg^{2+}-form crystal lattice, duplexes also form infinite columns along the crys-

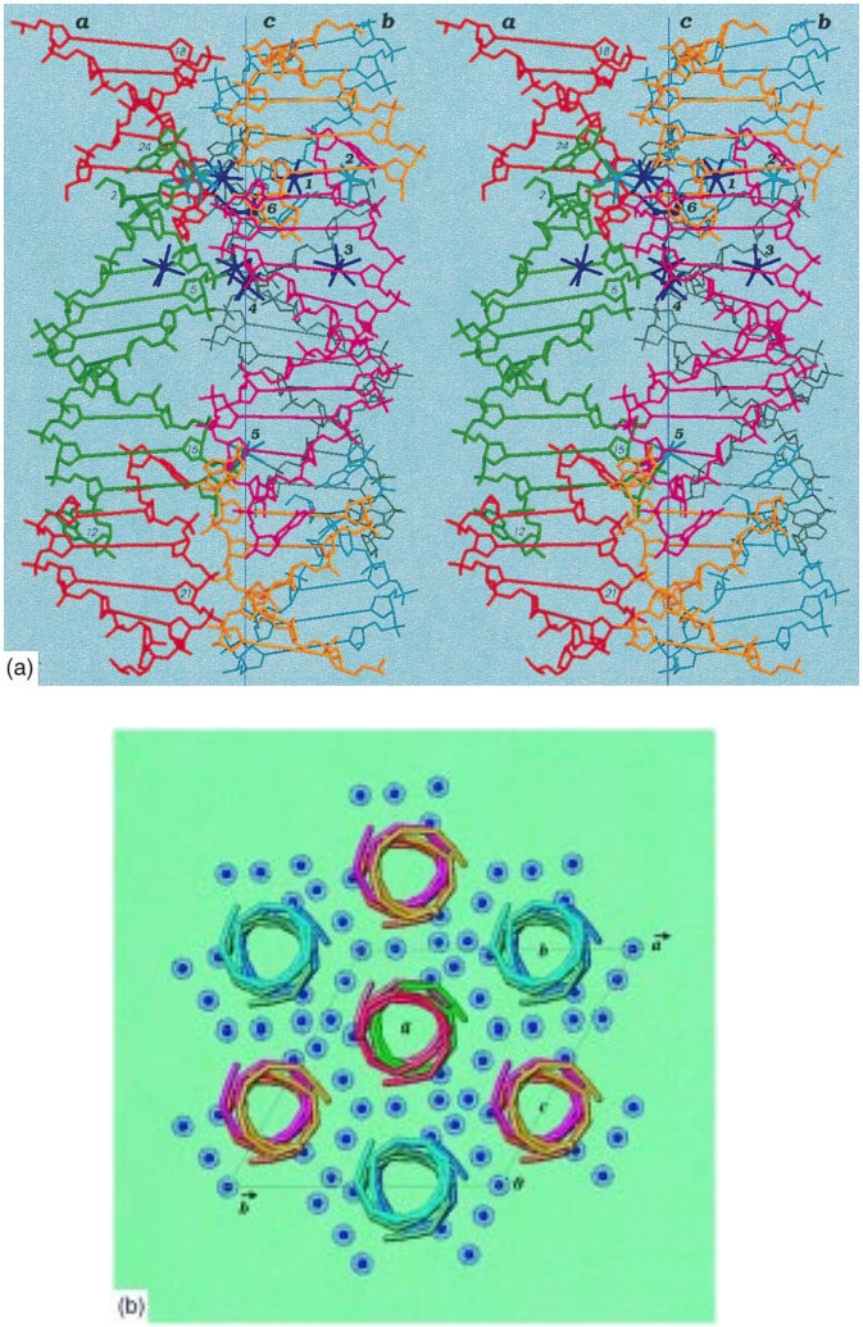

Figure 6. Packing interactions. (a) Stereo view of three columns of stacked duplexes (labelled *a*, *b*, *c*; the unit cell contains nine duplexes) in the GCCa lattice, viewed perpendicular to the 3-fold rotation axis (this continuous line). Only five base-pairs are shown for the duplexes forming the top and bottom layers. Calcium ions are drawn as small blue circles with their coordination spheres indicated by blue lines. (b) The GCCa unit cell (hexagonal setting) projected along the crystallographic *c*-axis. Each unit cell contains nine duplexes and a single duplex constitutes the asymmetric unit. The drawing illustrates the slide between duplexes in the stacks and the layer of Ca^{2+} separating them. Duplexes are drawn as a series of vectors between adjacent C1′ atoms and the color scheme matches that in (a). Ca^{2+} are drawn as dotted surfaces.

tallographic *c*-axis, but the topology of the end-to-end interactions is distinct from that in the GCCa structure. Duplexes do not stack but interact *via* inserting their ends into each other's minor grooves, generating a two base-pair overlap and sugar-base stacking interactions between backbones and terminal base-pairs (Wing *et al.*, 1980) (Figure 5b). Due to the particular conformation of the G overhangs in the GCCa structure, the C-G*G base triples formed in the two lattices are rather similar, despite the different topologies of duplex overlaps. In both cases, Gs are paired *via* two hydrogen bonds between the N2 and N3 atoms and are displaying large propeller twisting (Figure 5). Another shared feature is the extensive participation of ions in the stabilization of inter-duplex lattice contacts. In the CGMg structure, three Mg^{2+} clamp together the ends of duplexes by contacting either phosphate groups (Mg2*b* and Mg4) or terminal hydroxyl groups (Mg3*b*) (Figure 5(b)). The interacting duplexes are tilted relative to one another, generating close contacts mainly between phosphate groups from strands on one side of the joint (P···P distance ca 5.6 Å). On the other side, the distances between phosphate groups are larger (P···P distance ca 6.2 Å) and two further ions (Mg1 and Mg5*d*) bind in close vicinity of the backbones, but do not coordinate

Table 2. Geometry of Ca^{2+} coordination in the GCCa structure

Ion-P dist. (Å)	P-P dist. (Å)	Base	**Ligand DNA**	··· dist. (Å)	**Ion**	··· dist. (Å)	**Ligand water**	··· dist. (Å)	**DNA-atom**	Base	P-P dist. (Å)	Ion-P dist. (Å)
		G12 (b1)	O6	2.4		2.4	W103	3.0	O2P	G12 (b1)[a]		
3.5		C11 (b1)	O1P	2.4								
	5.2—[											
3.7		G24	O1P	2.2	**Ca1**							
	5.3—[											
3.6		G12 (c1)	O1P	2.4		2.3	W102					
						2.4	W105	3.2	O6	G12 (b1)		
						2.2	W111	2.8	O2P	C11 (b1)		5.6
											]—-5.2	
					Ca2	2.4	W107	2.8	O6	G24		
3.7		G24	O2P	2.3		2.3	W108	2.9	N7	G24		
						2.2	W110	2.8	O1P	C3 (b1)		5.5
						2.5	W106					
						2.5	W114	3.1/3.0/ 2.8	N2/O4′/ O3′	G4/A5/ G12(c1)		
						2.3	W113	2.7/2.8	O2/O4′	C21/G22		
					Ca3	2.4	W118	2.7/2.8	N3/O4′	A5/A6		
						2.4	W115					
						2.4	W116					
						2.3	W117					
						2.6	W119					
3.5	5.3	A5	O1P	2.3	**Ca4**[b]	2.3	W121	2.9	O2P	A5		
3.7	5.6	C15	O2P	2.2	**Ca5**[b]	2.6	W216					
						2.3	W124	2.6/2.6	O2P	G12 (a/b)		5.3/5.7
		G12 (c1)	O2P	2.5	**Ca6**	2.6	W197	2.8/2.6	O1P/N7	G12 (a1/c1)		
						3.3	W150	3.3	O3′	G4		
						2.6	W133					

[a] a and b designate symmetry mates and mark residues as belonging to duplexes within the central layer of the a and b stacks (Figure 6). Similarly, a1, b1 and c1 designate residues that belong to duplexes of the top layer within stacks a, b and c, respectively. The original molecule is located at the central level of the c stack.
[b] Ca4 and Ca5 are located on a 3-fold rotation axis.

directly to phosphate groups. An interesting difference between the ion-DNA interactions stabilizing inter-duplex contacts in the GCCa and CGMg lattices is the absence of Mg^{2+}-base contacts in the latter structure.

Comparison between the conformations of terminal base-pairs C1-G24 and C13-G12 in the CGMg duplex reveals elongated Watson-Crick hydrogen bonds in the case of the latter, accompanied by unstacking of the cytosine base (Figure 5(b)). In the structure of the same duplex crystallized at a lower Mg^{2+} concentration (low Mg^{2+}-form (Berger *et al.*, 1998)), the C13-G12 base-pair is intact. In the low Mg^{2+}-form crystals, only Mg1 was observed. In the CGMg lattice, Mg4 mediates contacts between duplexes not only along the crystallographic *c*-axis, but also along the *a*-axis (in Figure 5(b), the *a*-axis runs roughly perpendicular to the paper plane). The *a* cell constant of the CGMg crystal is reduced by 0.6 Å relative to the low-Mg^{2+} form. As a consequence, the N4 atom of C13 can engage in a lateral hydrogen bond to the O2P oxygen atom of residue G22 from a neighboring duplex along the *a*-direction. We believe that this causes the observed distortions of the terminal base-pair in the CGMg duplex.

Ion-mediated lateral contacts between duplexes

In the GCCa structure, calcium ions form a coulombic shield between duplexes (Figure 6(b)). A projection of the unit cell along the *c*-axis gives the impression that duplexes are essentially coated with metal ions. Ion-DNA contacts stabilizing end-to-end duplex interactions or occurring laterally at the level of the stacking interface are more extensive than lateral contacts between the central parts of duplexes (Figure 6(a)). The closest lateral contacts between duplexes are observed around the 3-fold rotation axis, which is practically parallel with the helical axes of duplexes. Three such close inter-duplex approaches per duplex are stabilized by a Ca^{2+}. In addition to relieving a close contact between phosphate groups of residues G12 and G24 in the minor groove, Ca1 also coordinates to the O6 and O1P atoms of residues G12 and C11, respectively, of a third duplex (Figure 6(a), Table 2). Ca4 and Ca5 are both located on the 3-fold axis and relieve tight contacts between phosphate groups P5 and P15, respectively, from three different duplexes. These duplexes face each other with their major grooves (Figure 6(a)). The two ions display octahedral geometry and have three water ligands in addition to directly coordinating to

phosphate oxygen atoms. A further Ca^{2+} (Ca6) with occupancy 1/3 is also located between three duplexes and coordinates to a water molecule which itself sits on the 3-fold rotation axis (Figure 6(a)). This ion displays a direct contact to O2P of residue G12 (Table 2).

Ca1 and Ca4 were observed in the CGCa and GGCa lattices as well and appear to be crucial for lattice formation. Conversely, Ca2, Ca5 and Ca6 are present only in the GCCa lattice. Compared with the CGCa and GGCa crystals, the *a* and *b* cell constants of the GCCa crystals are reduced by around 3 Å. This is obviously a result of the lack of the terminal cytidine residues. Thus, the lateral contacts between duplexes in the GCCa lattice are tighter and allow bridging of phosphate groups from neighbors by calcium ions. The tighter packing is consistent with the higher resolution of the diffraction data collected for GCCa crystals.

DNA conformation and metal ion coordination

In the GCCa lattice, each duplex interacts with 11 calcium ions (Figure 7(a)). Thus, every ion forms contacts to two or more duplexes (Table 2). As described above, Ca1, Ca2 and Ca6 are primarily engaged in mediating end-to-end overlaps and Ca4 and Ca5 stabilize lateral contacts between duplexes. Ca3 is the only ion that interacts with base functions in a groove (Figure 7(a)). Along with seven coordinated water molecules, the Ca^{2+} fits snugly into the minor groove. Figure 3 illustrates that the widths of the minor groove in the GCCa and GGCa duplexes are similar at one end (P6-P23), while the groove of the GCCa duplex (blue line) is 1 Å narrower at the other end (P11-P18). In the GGCa duplex (green line), calcium ions reside in the minor groove at either end of the duplex, while the GCCa duplex only features one (Ca3 at the P6-P23 end). Thus, the observed differences between the minor groove widths of the two duplexes are likely the result of Ca^{2+} coordination. Another feature apparent from the comparison in Figure 3 is the more symmetrical appearance of the minor groove width in the case of the Ca^{2+}-form duplexes (blue and green lines, Figure 3) compared with the CGMg duplex (red line). The minor groove in the A-tract portion (P8-P21 and P9-P20) of GCCa (blue line, Figure 3) is 0.6 Å narrower, on average, relative to CGMg (red line). Several phosphate groups (residues 8, 9, 20, 21) lining the A-tracts of the GCCa duplexes are engaged in tight inter-duplex phosphate-phosphate contacts around the 3-fold rotation axis. The distance of such phosphate interactions is 6.3 Å, on average, and does not involve Ca^{2+} coordination. In the case of the GGCa lattice with less tightly packed duplexes (see above), the corresponding distances are about 1 Å longer on average. The resulting A-tract minor groove width is intermediate compared with the GCCa and CGMg duplexes (Figure 3). Thus, certain differences between the minor groove widths in the GCCa, GGCa and CGMg duplexes can be rationalized by individual packing features.

In the CGMg lattice, each duplex is surrounded by 13 Mg^{2+} (Figure 7(b)). As in the Ca^{2+}-form lattice, each ion is thus engaged in interactions to two or more duplexes (Table 3). While Mg1, Mg2 and Mg3 have full occupancy, Mg4 and Mg5 are only partially occupied (Tereshko *et al.*, 1999a). Mg2 and Mg3 and the symmetry mates Mg2*a* and Mg3*a* form two bridges accross the minor groove. However, unlike Ca3 in the minor groove of the GCCa duplex, the magnesium ions remain at the periphery of the groove and do not penetrate it to interact with base functions. Mg2 and Mg3 bridge phosphate groups P19 and P10, respectively, and Mg2*a* and Mg3*a* bridge phosphate groups P12 and P17, respectively, from opposite strands (Figure 7(b)). This ion coordination affects the width of the minor groove (Figure 3). At the site where the Mg^{2+} tandem crosses the minor groove in the CGMg duplex (P10-P19), its width is reduced by almost 2 Å compared with the GCCa duplex. The CGMg minor groove is also contracted by 1 Å at this location relative to another DDD duplex (Shui *et al.*, 1998) (Figure 3). This duplex was crystallized at low Mg^{2+} concentration but in the presence of higher amounts of spermine and lacks the Mg2 and Mg3 binding sites. Thus, Mg^{2+} coordination can extend the narrow section of the minor groove in the DDD duplex beyond the central A-tract.

In the major groove, Mg1 bridges O6 atoms of G2 and G22 from opposite strands *via* coordinated water molecules. In addition, the ion interacts with two phosphate oxygen atoms from the backbone of a neighbor duplex (Mg1*b*, Figure 7(b), Table 3). The coordination of Mg1 in the major groove affects the conformation of the DDD duplex in a distinct manner. Both the CGMg and the GCCa duplex display a high positive roll of around 10° between base pairs C3-G22 and G4-C21 (Figure 2). At the other end, base-pairs C9-G16 and G10-C15 also show a positive but less pronounced roll. At both sites the rolls are accompanied by a strong slide and a reduced twist. The clamping by the Mg^{2+} of the exocyclic keto oxygen atoms of two adjacent guanine residues induces a kink of the duplex toward the major groove (Figure 4). This is a long-noted feature of the DDD duplex in the orthorhombic lattice that was earlier referred to as facultative bending (Dickerson *et al.*, 1994, 1996). However, at the other end of the CGMg duplex, in the absence of Mg^{2+} binding, a negative roll between base-pairs G10-C15 and C11-G14 neutralizes the above positive roll and straightens the duplex in order to allow continuous stacking. In the GCCa duplex, in the absence of metal ion coordination in the major groove and stacking between adjacent duplexes, no kinking is observed (Figure 6(a)). Thus, the prominent kink at one end

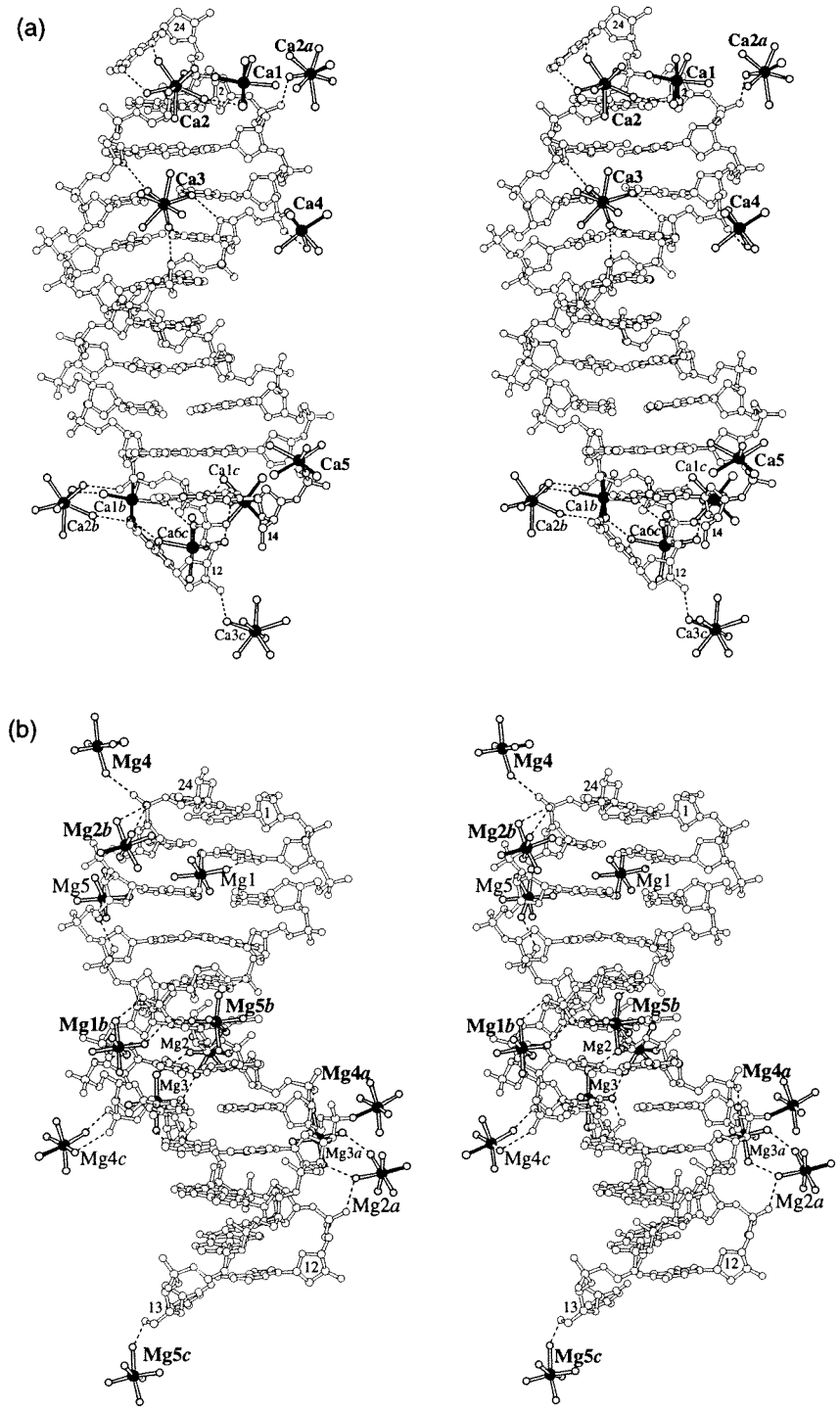

Figure 7. Ionic environment of individual duplexes. (a) Stereo view of the GCCa duplex surrounded by 11 Ca^{2+} (filled circles). (b) Stereoview of the CGMg duplex surrounded by 13 Mg^{2+} (filled circles). Symmetry mates of ions are designated with lower-case letters. Filled bonds indicate direct ion-DNA contacts, broken lines indicate hydrogen bonds and terminal residues are labelled. The orientations of the two duplexes provide optimal views of Ca3 and Mg1 coordinated in the minor and the major groove, respectively, and differ by a rotation of ca 60° around the vertical.

of the CGMg duplex can be directly attributed to the presence of a magnesium ion in the major groove.

Hydration

The high-resolution structures of the GCCa and CGMg duplexes provide a more complete picture of the water structure in the minor groove. The well-known spine of hydration is maintained but the ordered water structure involves more solvent molecules from higher shells as well (Tereshko *et al*., 1999a,b). Rather than just a zig-zag shaped spine winding down the floor of the minor groove (Drew & Dickerson, 1981; Kopka *et al*., 1983), the water structure takes on the appearance of a ribbon composed of four nearly planar fused water hexagons (Figure 8(b)). In the CGMg structure

Table 3. Geometry of Mg^{2+} coordination in the CGMg structure

Ion-P dist. (Å)	P-P dist. (Å)	Base	Ligand DNA	... dist. (Å)	Ion	... dist. (Å)	Ligand water	... dist. (Å)	DNA-atom	Base	P-P dist. (Å)	Ion-P dist. (Å)
						2.1	W29	2.8	O1P	A6 (c)[a]		
						2.1	W27	2.8	O2P	A6 (c)		4.6
					Mg1	2.1	W31	2.8	O2P	T7 (c)		5.3
											]– – –6.7	
						2.1	W26	2.8	N7	G2		
						2.1	W28	2.7	O6	G2		
						2.1	W30	2.7	O6	G22		
						2.1	W34	2.7/2.7	O1P/O1P	G24 (c)/ A18(d)	5.6	4.9/5.4
3.5		T19	O1P	2.1	**Mg2**	2.0	W33					
						2.0	W35					
						2.1	W36					
						2.0	W37					
						1.9	W44	2.7	O1P	G10		4.6
											]– – –6.2	
					Mg3	2.2	W42	2.7	O1P	A18 (d)		5.2
						2.1	W39					
						2.1	W40					
						2.0	W41					
						2.1	W43					
						2.0	W49	3.2	O2P	C9 (f)		
3.6		A17 (e)	O1P	2.2	**Mg4**	2.1	W46	3.0	O1P	C9 (f)		5.0
	5.7– – –[											
		G24				2.1	W50	2.8	O2P	G24		5.8
						2.2	W47					
						2.2	W48					
						1.9	W122	2.7	O1P	A5 (c)		5.1
						2.1	W179	2.7	O2P	A6 (c)		5.5
					Mg5	2.4	W57	2.7	O2P	C21		5.8
						2.7	W53					
						2.1	W54					
						2.2	W55					

[a] a, b, c, etc. designate symmetry mates.

refined to 1.1 Å resolution, the hydration pattern is more complete compared with the GCCa structure. In the latter, some water molecules from the second and third hydration shells are missing or show only low occupancy (Figure 8(a)). It is likely that some of these differences are due to the slightly lower resolution of the GCCa structure. However, another reason for the changes is the coordination of calcium ions in the minor groove of the GCCa and GGCa duplexes. The Ca3 ion binding adjacent to the hexagons distorts the water structure only minimally. First shell water molecules that coordinate to base atoms of C-G and T-A pairs become part of its hydration sphere. At the other end of the minor groove in the GGCa duplex, a second Ca^{2+} coordinates with partial occupancy (Ca7, Figure 8(a)). The resulting rearrangement of solvent molecules disrupts two water hexagons.

In the CGMg duplex, the hydrogen bonding network that links water molecules that are part of the hexagons to phosphate oxygen atoms reflects the topology of the minor groove (Figures 3 and 8(b)). Near the base-pairs A5-T20 and A6-T19, the minor groove is about 2 Å wider than near base-pairs T7-A18 and T8-A17. As a consequence, P7 and P22 are each linked to the second shell water W1′ *via* a single water molecule. Similarly, P8 and P21 are each linked to the second shell water W3′ *via* a single water. In the second half of the A-tract, where the minor groove is narrower as a result of Mg^{2+} coordination (Figures 3 and 7(b)), water molecules from hexagons can directly hydrogen bond to phosphate oxygen atoms (Figure 8(b)). For example, W5′ forms a hydrogen bond to O1P of T20 and W7′ is hydrogen bonded to O1P of G10. The hydrogen bonds to phosphate groups from the other strand are still mediated by single water molecules. At the narrowest site of the minor groove (P10-P19, Figure 3), the bridging water molecule is coordinated to Mg2 (Figure 8(b)). The similarities between the overall water structures in the minor grooves of the GCCa, GGCa and CGMg duplexes provide further evidence that DNA water structure is not the result of a particular crystal lattice, but appears to be intimately related to DNA sequence (Schneider *et al*., 1992, 1993) and duplex topology (Gessner *et al*., 1994).

Water hexagons are also found in the major groove of the CGMg duplex (Figure 9). The hydration pattern comprises two central hexagons flanked by pentagons and covers the central GAATTC hexamer portion. Both the hexagonal

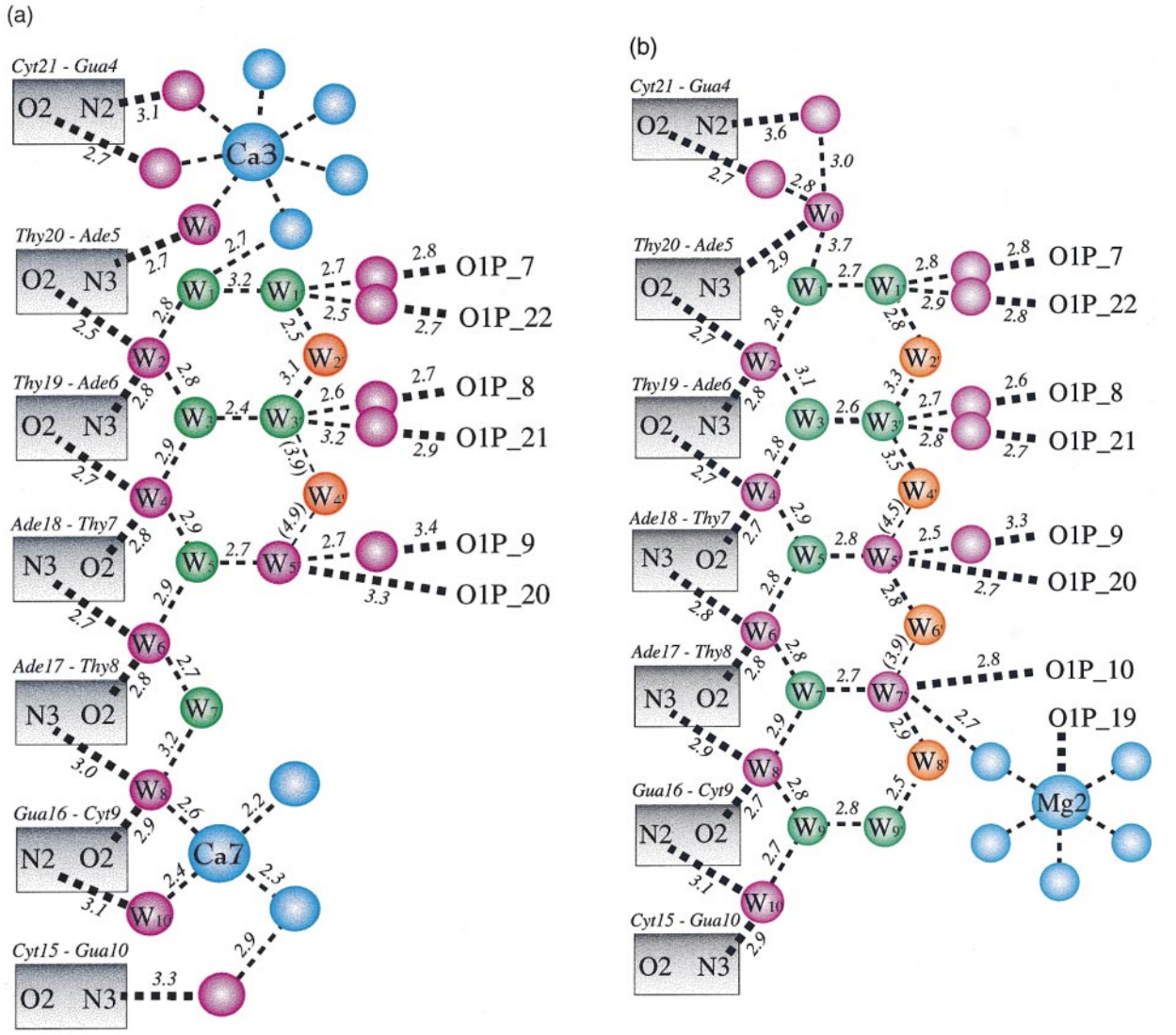

Figure 8. Minor groove hydration. Representations of the water structures in the minor grooves of (a) the GCCa and GGCa duplexes and (b) the CGMg duplex. The partially occupied Ca^{2+} coordinated near base-pair C9-G16 (Ca7) is visible only in the GGCa structure. Base-pairs are drawn as boxes and bases are labelled. Water molecules of the first, second and third hydration shells are drawn as spheres colored in magenta, green and orange, respectively. Ca^{2+} and Mg^{2+} are drawn as larger cyan spheres and ion-coordinated water molecules that are not part of the first hydration shell are colored in cyan as well. Thick broken lines indicate hydrogen bonds between DNA bases or phosphate groups and first shell water molecules. All other hydrogen bonds are indicated by thin broken lines with distances in Å.

and pentagonal arrangements stretch over three base-pairs and each of the participating water molecules hydrogen bonds to a base function of a purine or a pyrimidine base. Due to the different groove topologies, the orientations of the water hexagons in the major and minor grooves are not the same. In the minor groove, the ribbon of hexagons dissects the groove in a fashion that renders hexagons roughly perpendicular to the direction of Watson-Crick hydrogen bonds (Tereshko *et al.*, 1999a,b). In the major groove, the ribbon is rotated and pentagons and hexagons are now oriented more or less parallel to the direction of Watson-Crick hydrogen bonds (Figure 9). In the GCCa duplex, the major groove water structure is only partially defined, most likely due to the slightly more limited resolution of that structure.

Conclusion

The crystallographic results presented here attest to the power of accurate structures and the potential benefits of diffraction data collected to atomic resolution. Cocrystallizations of the Dickerson-Drew dodecamer with Mg^{2+} or Ca^{2+} yield different crystal forms. The located Mg^{2+} ions account for neutralization of ten phosphate groups or almost half the total negative charge of −22 present in the dodecamer duplex. In the case of the Ca^{2+}-form, the located ions neutralize 13 of the 20 phosphate

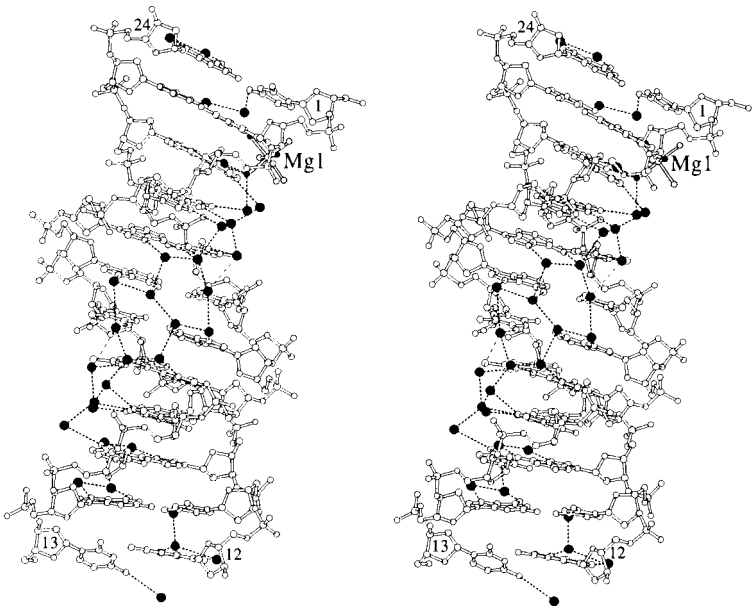

Figure 9. Major groove hydration. Stereo view of the ribbon defined by two central water hexagons and flanking pentagons. Water molecules are drawn as filled circles, hydrogen bonds are indicated by broken lines and terminal residues of the duplex are labelled.

groups. The remaining cations required for charge neutrality may form mostly outer-sphere complexes and the higher flexibility then renders more difficult their detection in electron density maps. It is possible that some of the positive charges are contributed by sodium ions which are more difficult to locate than Mg^{2+} or Ca^{2+}. However, crystallizations of the CGCGAATTCGCG dodecamer in the presence of rubidium or cesium cacodylate only revealed the presence of a single alkali metal ion in the center of the minor groove (Tereshko *et al*., 1999b). It is reasonable to conclude that the cations not revealed by our structures do not crucially affect the conformation of the duplex, since they probably do not directly coordinate to DNA atoms.

Clusters of Mg^{2+} and Ca^{2+} stabilize specific end-to-end and lateral contacts between duplexes in the respective lattices. However, the role of ions is not limited to merely interacting with backbone phosphate groups. Our structures allow a dissection of the contributions of packing and ion coordination to DNA conformation. Several conformational features of the *B*-form duplex are a direct result of ion coordination. For example, the asymmetric kink of the Dickerson-Drew duplex present in one crystal form is due to Mg^{2+} coordination in the major groove. Whereas calcium ions penetrate the minor groove and interact with bases, two magnesium ions bridge phosphate groups from opposite strands, but remain at the periphery of the groove. Ca^{2+} coordination appears to widen the groove outside the A-tract. Conversely, Mg^{2+} coordination contracts the groove and extends the narrow zone of the minor groove beyond the A-tract. Thus, in addition to crystal packing and base sequence, ion coordination appears to have to be taken into account when analyzing the conformational properties of DNA. Many similarities between the overall features of the DNA duplexes in the two crystal forms attest to the validity of the crystallographic results and argue against packing forces as a major determinant of DNA conformation. DNA sequence and ion coordination are clearly more important in this respect. The similarity of the minor groove hydration patterns observed in the two different lattices confirms that sequence is the major determinant of DNA conformation and the organization of the ordered solvent molecules.

The Mg^{2+} and Ca^{2+} coordination modes exhibit some important differences. Besides residing in the major and the minor groove, respectively, calcium ions are engaged in more inner-sphere complexes. This is consistent with the lower hydration enthalpy of Ca^{2+} relative to Mg^{2+} (Cowan, 1995). In both lattices, metal ions mediate contacts between multiple strands, three in the case of Mg^{2+}-form crystals and four in the case of the Ca^{2+}-form crystals. The versatility of the ion-DNA interactions observed in the structures provides an indication for the important role that ions could play in the stabilization of other DNA structures, for example triple-stranded arrangements. In addition, it is possible that metal ions mediate specific contacts between DNA and small molecules (minor groove binding drugs) or proteins. Future high-resolution structures of such complexes will shed light on these aspects.

Materials and Methods

Synthesis, crystallization and data collection

The DDD used for the CGMg and CGCa crystallizations was synthesized and purified as described (Berger *et al*., 1998). The oligonucleotide contains a 2′-deoxy-2′-fluoro-arabinofuranosyl thymine residue at position 7 and the conformational properties of the chemically modified residue and the resulting

consequences for the local helix geometry were discussed in an earlier publication (Berger *et al*., 1998). The GGCa and GCCa oligomers were obtained form Oligos Etc., Wilsonville, OR. Both were purified by RP-HPLC and desalted prior to adjusting the concentrations of the stock solutions to about 5 mM. The conditions for growing the CGCa, GGCa and GCCa crystals are summarized in Table 4. All data collections except those with CGCa crystals were conducted on the insertion device beamline (5-ID-B) of the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) of the Advanced Photon Source (APS) at Argonne National Laboratory, Argonne, IL, using a MARCCD detector. A summary of the crystal data, data resolutions and reflection statistics is provided in Table 4. The data were processed and scaled in the DENZO/SCALEPACK suite (Otwinowski & Minor, 1997).

Structure solution and refinement

The crystal structure of the Ca^{2+}-form of the DDD duplex was initially determined based on data collected with CGCa crystals using the Molecular Replacement method (Navaza, 1994). The resulting model was then refined in combination with the nearly isomorphous GGCa and GCCa data, both of higher resolution than the CGCa data. The initial refinements were conducted with the program CNS (Brünger, 1998), using the most recent parameter files (Parkinson *et al*., 1996). For the high-resolution GCCa structure (1.3 Å), all subsequent refinements were carried out with the program SHELX-97 (Sheldrick & Schneider, 1997), using standard restraints and setting 10 % of the data aside to calculate the R_{free} value (Brünger, 1992). Solvent molecules were picked automatically and the correctness of the assignments was then examined by visualizing electron density maps on the graphics display. All DNA atoms, metal ions and ordered solvent molecules were treated anisotropically. The anisotropies were restrained to adopt similar shapes and directions for bonded atoms (DELU and SIMU commands, respectively). Final *R*-factors for the three structures as well as r.m.s. deviations from standard bond lengths and angles are listed in Table 4.

Database accession numbers

The Nucleic Acid Database accession code for the CGMg structure is BD0007. The coordinates and structure factors for the CGCa, GGCa and GCCa structures have been deposited in the NDB (entry codes BD0020, BD0019 and BD0018, respectively).

Acknowledgements

This work was supported by the National Institutes of Health (grant R01 GM-55237). We thank the referees for helpful comments and suggestions. The DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center located at Sector 5 of the Advanced Photon Source at Argonne National Laboratory, Argonne, IL, is supported by the E.I. DuPont de Nemours & Co., The Dow Chemical Company as well as the U.S. National Science Foundation and the State of Illinois.

Table 4. Crystal data and refinement parameters

Structure	GCCa	GGCa	CGCa
A. *Crystallization*			
DNA (mM) (single strand)		0.8-1.2	
pH		6.9	
$CaCl_2$ (mM)		10-40	
MPD (%)		40	
B. *Crystal data*			
Space group		R3	
$a = b$ (Å)	38.76	41.99	42.31
c (Å)	99.08	99.55	99.68
C. *Data collection*			
X-ray source/detector	Synchrotron-APS/MARCCD	Synchrotron-APS/MARCCD	Rigaku RU-200/R-axis IIc
Temperature		120 K	
Total no. reflections	54,240	39,012	37,558
No. unique reflections	13,911	7,036	3,457
Resolution (Å)	1.30	1.70	2.20
Completeness (%)	99.5	99.4	99.9
R_{sym}[a] (%)	8.9	5.0	8.8
D. *Refinement statistics*			
No. of DNA atoms	448	456	458
No. of water molecules	125	151	96
No. of Ca^{2+}	6 (Ca1 - Ca6)	4 (Ca1, Ca3, Ca4, Ca7)	3 (Ca1, Ca3, Ca4)
R.m.s distance (Å)	0.006	0.009	0.010
R.m.s. angles (°)	1.57	1.35	1.64
No. reflections [$F > 0$]	13,120	6228	3,264
R-factor[b] (work set)	0.182	0.194	0.197
R-factor[c] (test set)	0.220	0.235	0.262

[a] $R_{sym} = \Sigma_{hkl}\Sigma_i|I(hkl)_i - \langle I(hkl)_i\rangle|/\Sigma_{hkl}\Sigma_i\langle I(hkl)_i\rangle$.
[b] R-factor $= \Sigma_{hkl}|F(hkl)_o - F(hkl)_c|/\Sigma_{hkl}F(hkl)_o$.
[c] For 10 % of the data.

References

Allemann, R. & Egli, M. (1997). DNA bending and recognition. *Chem. Biol.* **4**, 643-650.

Baikalov, I., Grzeskowiak, K., Yanagi, K., Qunitana, J. & Dickerson, R. E. (1993). The crystal structure of the trigonal decamer C-G-A-T-C-G-(6Me)A-T-C-G - a *B*-DNA helix with 10.6 base-pairs per turn. *J. Mol. Biol.* **231**, 768-784.

Bancroft, D., Williams, L. D., Rich, A. & Egli, M. (1994). The low-temperature crystal structure of the pure-spermine form of Z-DNA reveals binding of a spermine molecule in the minor groove. *Biochemistry,* **33**, 1073-1086.

Berger, I., Egli, M. & Rich, A. (1996). Inter-strand C-H···O hydrogen bonds stabilizing four-stranded intercalated molecules: stereoelectronic effects of O4′ in cytosine-rich DNA. *Proc. Natl Acad. Sci. USA,* **93**, 12116-12121.

Berger, I., Tereshko, V., Ikeda, H., Marquez, V. E. & Egli, M. (1998). Crystal structures of B-DNA with incorporated 2′-deoxy-2′-fluoro-arabino-furanosyl thymines: implications of conformational preorganization for duplex stability. *Nucl. Acids Res.* **26**, 2473-2480.

Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R. & Schneider, B. (1992). The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* **63**, 751-759.

Brünger, A. T. (1992). The free *R* value: a novel statistical quantity for accessing the accuracy of crystal structures. *Nature,* **355**, 472-475.

Brünger, A. T. (1998). *Crystallography & NMR System (CNS)*, Version 0.3, Yale University, New Haven, CT.

Chen, L., Cai, L., Zhang, X. & Rich, A. (1994). Crystal structure of a four-stranded intercalated DNA: d(C_4). *Biochemistry,* **33**, 13540-13546.

Cowan, J. A. (1995). Introduction to the biological role of magnesium ion. In *The Biological Chemistry of Magnesium* (Cowan, J. A., ed.), pp. 1-23, VCH Publishers Inc., New York, NY.

Dickerson, R. E. (1992). DNA structure from *A* to *Z*. *Methods Enzymol.* **211**, 67-111.

Dickerson, R. E., Kopka, M. L. & Drew, H. R. (1982). Structural correlations in *B*-DNA. In *Conformation in Biology* (Srinivasan, R. & Sarma, R. H., eds), pp. 227-257, Adenine Press, New York, NY.

Dickerson, R. E., Goodsell, D. S. & Neidle, S. (1994). ''... the tyranny of the lattice ...''. *Proc. Natl Acad. Sci. USA,* **91**, 3579-3583.

Dickerson, R. E., Goodsell, D. & Kopka, M. L. (1996). MPD and DNA bending in crystals and in solution. *J. Mol. Biol.* **256**, 108-125.

Drew, H. R. & Dickerson, R. E. (1981). Structure of a *B*-DNA dodecamer. III. Geometry of hydration. *J. Mol. Biol.* **151**, 535-556.

Egli, M. (1994). Structural patterns in nucleic acids. In *Structure Correlation* (Bürgi, H.-B. & Dunitz, J. D., eds), vol. 2, pp. 705-749, VCH Publishers Inc., Weinheim, Germany.

Egli, M. (1996). Structural aspects of nucleic acid analogues and antisense oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **35**, 1894-1909.

Gehring, K., Leroy, J.-L. & Guéron, M. (1993). A tetrameric DNA structure with protonated cytosine·cytosine base pairs. *Nature,* **363**, 561-565.

Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A. & Wang, A. H.-J. (1989). The molecular structure of the left-handed *Z*-DNA double helix at 1.0-Å atomic resolution - geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem.* **264**, 7921-7935.

Gessner, R. V., Quigley, G. J. & Egli, M. (1994). Comparative studies of high resolution Z-DNA crystal structures. I. Comparative hydration patterns of alternating dC-dG. *J. Mol. Biol.* **236**, 1154-1168.

Grzeskowiak, K. (1996). Sequence-dependent structural variation in *B*-DNA. *Chem. Biol.* **3**, 785-790.

Grzeskowiak, K., Yanagi, K., Privé, G. G. & Dickerson, R. E. (1991). The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G - the effect of base pair reversals. *J. Biol. Chem.* **266**, 8861-8883.

Heinemann, U. & Alings, C. (1989). Crystallographic study of one turn of G-C-rich *B*-DNA. *J. Mol. Biol.* **210**, 369-381.

Heinemann, U. & Alings, C. (1991). The conformation of a *B*-DNA decamer is mainly determined by its sequence and not by crystal environment. *EMBO J.* **10**, 35-43.

Heinemann, U. & Hahn, M. (1992). C-C-A-G-G-C-M5C-T-G-G - helical fine-structure, hydration and comparison with C-C-A-G-G-C-C-T-G-G. *J. Biol. Chem.* **267**, 7312-7341.

Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature,* **356**, 126-131.

Kennard, O. & Hunter, W. N. (1991). Single crystal X-ray diffraction studies of oligonucleotides and oligonucleotide-drug complexes. *Angew. Chem. Int. Ed. Engl.* **30**, 1254-1277.

Kim, Y., Grable, J. C., Love, R., Greene, P. J. & Rosenberg, J. M. (1990). Refinement of *Eco*RI endonuclease crystal structure: a revised chain tracing. *Science,* **249**, 1307-1309.

Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994). *Hha*I methyltransferase flips its target base out of the DNA helix. *Cell,* **76**, 357-369.

Kopka, M. L., Fratini, A. V., Drew, H. R. & Dickerson, R. E. (1983). Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* **163**, 129-146.

Laughlan, G., Murchie, A. I. H., Norman, D., Moore, M. H., Moody, P. C. E., Lilley, D. M. J. & Luisi, B. (1994). The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science,* **265**, 520-524.

Lavery, R. & Sklenar, H. J. (1989). Defining the structure of irregular nucleic acids - conventions and principles. *J. Biomol. Struct. Dyn.* **6**, 655-667.

Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J. & Dickerson, R. E. (1993). Structure of the B-DNA decamer C-C-A-A-C-I-T-T-G-G in two different space groups: conformational flexibility of *B*-DNA. *Biochemistry,* **32**, 1373-1389.

Liu, J., Malinina, L., Huynh-Dinh, T. & Subirana, J. A. (1998). The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG). *FEBS Letters,* **438**, 211-214.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* **389**, 251-260.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.

Nunn, C. M., Garman, E. & Neidle, S. (1997). Crystal structure of the DNA decamer d(CGCAATTGCG) complexed with the minor groove binding drug netropsin. *Biochemistry,* **36**, 4792-4799.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. T. & Berman, H. M. (1996). New parameters for the refinement of nucleic acid containing structures. *Acta Crystallog. sect. D,* **52**, 57-64.

Portmann, S., Altmann, K.-H., Reynes, N. & Egli, M. (1997). Crystal structures of oligodeoxyribonucleotides containing 6′-α-methyl and 6′-α-hydroxy carbocyclic thymidines. *J. Am. Chem. Soc.* **119**, 2396-2403.

Privé, G. G., Yanagi, K. & Dickerson, R. E. (1991). Structure of the *B*-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G. *J. Mol. Biol.* **217**, 177-199.

Schneider, B., Cohen, D. & Berman, H. M. (1992). Hydration of DNA bases: analysis of crystallographic data. *Biopolymers,* **32**, 725-750.

Schneider, B., Cohen, D. M., Schleifer, L., Srinivasan, A. R., Olson, W. K. & Berman, H. M. (1993). A systematic method for studying the spatial distribution of water molecules around nucleic acid bases. *Biophys. J.,* **65**, 2291-2303.

Sheldrick, G. M. & Schneider, T. R. (1997). SHELX-97: high-resolution refinement. *Methods Enzymol.* **276**, 319-343.

Shui, X., McFail-Isom, L., Hu, G. H. & Williams, L. D. (1998). The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry,* **37**, 8341-8355.

Sklenar, V. & Feigon, J. (1990). Formation of a stable triplex from a single strand. *Nature,* **345**, 836-838.

Spink, N., Nunn, C. M., Vojtechovsky, J., Berman, H. M. & Neidle, S. (1995). Crystal structure of a DNA decamer showing a novel pseudo four-way helix-helix junction. *Proc. Natl Acad. Sci. USA,* **92**, 10767-10771.

Steitz, T. A. (1993). *Structural Studies of Protein-Nucleic Acid Interaction: The Sources of Sequence-specific Binding*, Cambridge University Press, Cambridge, UK.

Takahara, P. M., Frederick, C. A. & Lippard, S. J. (1996). Crystal structure of the anticancer drug cisplatin bound to duplex DNA. *J. Am. Chem. Soc.* **118**, 12309-12321.

Tereshko, V., Minasov, G. & Egli, M. (1999a). The Dickerson-Drew B-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* **121**, 470-471.

Tereshko, V., Minasov, G. & Egli, M. (1999b). A ''hydrat-ion spine'' in a *B*-DNA minor groove. *J. Am. Chem. Soc.* **121**, 3590-3595.

Van Meervelt, L., Vlieghe, D., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1995). High-resolution structure of a DNA helix forming (C·G)*G base triplets. *Nature,* **374**, 742-744.

Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1996). Parallel and antiparallel $(G \cdot GC)_2$ triple helix fragments in a crystal structure. *Science,* **273**, 1702-1705.

Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. & Dickerson, R. E. (1980). Crystal structure analysis of a complete turn of *B*-DNA. *Nature,* **287**, 755-758.

Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K. & Wilson, K. S. (1993). The crystal structure of *Eco*RV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12**, 1781-1795.

Wood, A. A., Nunn, C. M., Trent, J. O. & Neidle, S. (1997). Sequence-dependent crossed helix packing in the crystal structure of a *B*-DNA decamer yields a detailed model for the Holliday junction. *J. Mol. Biol.* **269**, 827-841.

Edited by T. Richmond

(Received 5 February 1998; received in revised form 28 May 1999; accepted 7 June 1999)