

## Structure of d(TGCGCA)<sub>2</sub> and a Comparison with Other DNA Hexamers

ANNE HARPER,<sup>a</sup> JAMES A. BRANNIGAN,<sup>a</sup> MARTIN BUCK,<sup>b</sup> LORRAINE HEWITT,<sup>a</sup> RICHARD J. LEWIS,<sup>a</sup> MADELEINE H. MOORE<sup>a</sup> AND BOHDAN SCHNEIDER<sup>c\*</sup>

<sup>a</sup>Department of Chemistry, University of York, Heslington, York YO1 5DD, England, <sup>b</sup>Department of Biology, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BB, England, and <sup>c</sup>J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, CZ-182 23 Prague, Czech Republic. E-mail: bohdan.schneider@jh-inst.cas.cz

(Received 17 November 1997; accepted 13 February 1998)

### Abstract

The X-ray crystal structure of d(TGCGCA)<sub>2</sub> has been determined at 120 K to a resolution of 1.3 Å. Hexamer duplexes, in the Z-DNA conformation, pack in an arrangement similar to the 'pure spermine form' [Egli *et al.* (1991). *Biochemistry*, **30**, 11388–11402] but with significantly different cell dimensions. The phosphate backbone exists in two equally populated discrete conformations at one nucleotide step, around phosphate 11. The structure contains two ordered cobalt hexamine molecules which have roles in stabilization of both the Z-DNA conformation of the duplex and in crystal packing. A comparison of d(TGCGCA)<sub>2</sub> with other Z-DNA hexamer structures available in the Nucleic Acid Database illustrates the elusive nature of crystal packing. A review of the interactions with the metal cations Na<sup>+</sup>, Mg<sup>2+</sup> and Co<sup>3+</sup> reveals a relatively small proportion of phosphate binding and that close contacts between metal ions are common. A prediction of the water structure is compared with the observed pattern in the reported structure.

### 1. Introduction

Of all the single-crystal X-ray structures of DNA duplexes, hexamers in the left-handed Z-DNA conformation have provided some of the most well determined structures (Wang *et al.*, 1979; Gessner *et al.*, 1989). The majority of these hexamers are in an orthorhombic crystal form which diffracts X-rays to high resolution. The first structure obtained was d(CGCGCG)<sub>2</sub> (Wang *et al.*, 1979) and since then variations including both alternating pyrimidine–purine (Wang, Hakoshima *et al.*, 1984; Zhou & Ho, 1990; Sadasivan & Gautham, 1995) and non-alternating sequences (Wang *et al.*, 1985; Schroth *et al.*, 1993) have been determined. The d(CGCGCG)<sub>2</sub> Z-DNA hexamer has also been used as a duplex framework for high-resolution studies of non-Watson–Crick mismatched base pairs (Brown *et al.*, 1986) and modified base-pairing geometries (Coll *et al.*, 1989; Ginell *et al.*, 1990; van Meervelt *et al.*, 1990; Ohishi

*et al.*, 1991; Cervi *et al.*, 1993; Moore *et al.*, 1995). It has been proposed that the Z-DNA conformation is stabilized by methylation of cytosine (Fujii *et al.*, 1982) and demethylation of thymine (Zhou & Ho, 1990). The hexamer helix within the same orthorhombic lattice can accommodate significant base and backbone modifications (Wang *et al.*, 1985; van Meervelt *et al.*, 1990; Kagawa *et al.*, 1991; Schneider *et al.*, 1992; Moore *et al.*, 1995) whilst always forming pseudo-infinite helices by stacking of end base pairs. Interestingly, certain slight base modifications of an alternating pyrimidine–purine sequence adopt a completely different packing arrangement (Parkinson *et al.*, 1995).

Hexamer DNA sequences, in the absence of intercalator drug molecules, tend to crystallize more readily as left-handed helices. Out of a total of 96 hexamer structures reported in the Nucleic Acid Database (NDB) (Berman *et al.*, 1992), 43 are Z-DNA and 46 are hexamer–anthracycline complexes. Only three different native hexamer sequences crystallize in right-handed DNA forms, one as A-DNA (Mooers *et al.*, 1995) and two as B-DNA (Tari & Secco, 1995; Wahl *et al.*, 1996). There are four different sequences which alone crystallize as Z-DNA but with a drug intercalated form a distorted right-handed B-type helix: CGCGCG (Wang *et al.*, 1991), CGTACG (Wang *et al.*, 1984), CGT(NH<sub>2</sub>)ACG–daunorubicin (Wang *et al.*, 1991) and CGATCG (Moore *et al.*, 1989).

The sequence d(TGCGCA)<sub>2</sub> is one of a number of hexamer DNA sequences we have used to investigate sequence selectivity of the anthracycline antibiotic nogalamycin (Smith *et al.*, 1995, 1996). Footprinting experiments using calf-thymus DNA indicate that 5'-TGC-3' is a favorable triplet binding site for nogalamycin (Fox & Waring, 1984, 1986) but initial crystallization experiments carried out with this binding site in the hexamer d(TGCGCA)<sub>2</sub> and nogalamycin produced crystals of the DNA with no drug bound; the crystallization conditions of native d(TGCGCA)<sub>2</sub> were optimized in the absence of nogalamycin.

Z-DNA structures have not only provided important detailed information on both helical conformation and

base-pairing geometry but also revealed information on the binding of cations. Various metal cations affect the equilibrium between right-handed B-DNA and left-handed Z-DNA conformations in solution in different ways (Pohl & Jovin, 1972; Thamann *et al.*, 1981; Rich *et al.*, 1984). Polyvalent metals are usually necessary to stabilize Z-DNA and cobalt hexammine is known to be particularly effective (Behe & Felsenfeld, 1981). The structural role of metal stabilization of Z-DNA has been explored in several structures containing Mg<sup>2+</sup> (Bancroft *et al.*, 1994; Egli *et al.*, 1991; Ohishi *et al.*, 1991, 1996) and one containing cobalt hexammine (Gessner *et al.*, 1985).

The structure of d(TGCGCA)<sub>2</sub> was determined in order to examine the role cobalt hexammine plays in stabilization of the crystal lattice and the DNA conformation. This structure was then compared with other DNA hexamers for which there are coordinates available in the NDB (Berman *et al.*, 1992) primarily from the point of their packing and interactions with metal cations and water.

## 2. Methods

### 2.1. Synthesis, purification and crystallization

The oligomer d(TGCGCA)<sub>2</sub> was synthesized on an Applied Biosystems 392 DNA/RNA Synthesizer and HPLC purified using a Pharmacia reversed-phase Resource column.

The initial crystallization conditions were found using the NACRYST sparse-matrix screen (Baeyens *et al.*, 1994) and further optimized. Crystals were grown at 291 K in about two weeks by the vapour-diffusion method in Cryschem sitting-drop trays. The stock solution contained 2 mM single-stranded hexamer in water. The mother liquor contained 50 mM sodium cacodylate buffer (pH = 6.0), 1 mM cobalt hexammine, 5 mM argininamide chloride, 30% (v/v) MPD and 5% (v/v) PEG 400. Drops of 6 µl total volume obtained by combining 3 µl of stock solution with 3 µl of mother liquor were equilibrated against 500 µl of mother liquor diluted with 200 µl water.

### 2.2. Data collection

d(TGCGCA)<sub>2</sub> crystallized in space group *P*2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with unit-cell dimensions *a* = 21.16, *b* = 28.67, *c* = 44.34 Å. The crystal, of dimensions 0.20 × 0.25 × 0.20 mm, was transferred directly from the mother liquor to a nitrogen stream at a temperature of 120 K, using a rayon loop. A data set complete to 1.3 Å resolution was collected using a MAR Research image plate at the station PX 9.6 at the SRS, Daresbury, at a wavelength of 0.898 Å. Three data sets obtained with one crystal at different crystal-to-detector distances, exposure times and oscillation ranges were merged to obtain the final data set. Data were processed and reduced using *DENZO* and *SCALEPACK*, respectively (Otwi-

nowski, 1991). A summary of the data-collection and refinement statistics is given in Table 1.

### 2.3. Structure solution and refinement

The structure was solved by molecular replacement in *AMoRe* (Navaza, 1994) using a Z-DNA hexamer model generated in *QUANTA* (Molecular Simulations, Burlington, MA) as the search model. This model was based on d(CGCGCG)<sub>2</sub> (Gessner *et al.*, 1989). All molecular-replacement calculations were performed using all data between 24 and 3.0 Å (630 reflections). Rigid-body refinement converged at *R* = 44.7% with a correlation coefficient of 0.58. This solution generated a readily interpretable electron-density map. Graphical analyses of the model and electron-density maps were carried out using *QUANTA*.

All subsequent refinement was carried out in *SHELXL* (Sheldrick, 1993) using all intensities in the 10–1.3 Å range (Table 1) and the previously reported restraints (Schuerman *et al.*, 1996; Smith *et al.*, 1996). For cross-validation purposes approximately 10% of the data were used for the *R*<sub>free</sub> calculation. After isotropic refinement of all DNA non-H atoms and inclusion of 15 water molecules, *R* was reduced to 0.32 (*R*<sub>free</sub> = 0.38). Two of the water molecules refined to very low temperature factors and each was approximately 2 Å from six octahedrally arranged peaks in the electron-density maps. A crystal structure of cobalt hexammine (Takusagawa *et al.*, 1988) taken from the Cambridge Crystallographic Data Centre (Allen *et al.*, 1991) was used as the search molecule in *X-LIGAND* in *QUANTA* (Oldfield, 1994) and the two independent positions were unambiguously identified in the density map. Subsequent refinement after inclusion of bond length and angle restraints for the cobalt hexammine reduced *R* to 0.28 (*R*<sub>free</sub> = 0.35). Solvent molecules were identified using *X-SOLVATE* (Oldfield, 1994) applying the same criteria for their identification and confirmation as reported previously (Schuerman *et al.*, 1996; Smith *et al.*, 1996).

Throughout refinement density maps indicated the DNA backbone around the phosphate P11 was in both ZI and ZII conformations. Following inclusion of 58 water molecules and the two cobalt hexammine complexes in refinement, P11 was modelled in both conformations and their occupancies refined to 0.56 and 0.44 (Table 1). After a total of 78 water molecules were located and all H atoms were inserted in calculated positions and refined using a rigid model, restrained anisotropic temperature-factor refinement was carried out on all non-H atoms. At this stage, the decrease in both *R* (3.4%) and *R*<sub>free</sub> (1.4%) values was considered to be significant. Finally full-matrix restrained least-squares refinement was carried out to convergence using all the data between 10 and 1.30 Å, 7005 unique reflections. Standard deviations on atomic positions and

$U^{ij}$  parameters were then obtained by carrying out three cycles of unrestrained full-matrix least-squares refinement. For this the structure was divided into three blocks, two containing  $xyz$  and  $U^{ij}$  parameters for the DNA only and the other containing those for the cobalt hexammine and water molecules. Atomic coordinates and structure factors have been deposited with the NDB (Berman *et al.*, 1992), code ZDF060.†

#### 2.4. Structure analysis and comparisons

Nucleotide residues were labelled T1 to A6 in the 5' to 3' direction on strand *A* and T7 to A12 on strand *B* as illustrated in Fig. 1(*a*). Helical parameters were calculated by the program *Newhel93* (Dickerson, 1993). The NDB was used to search for known DNA hexamer structures and for obtaining information on these structures. A previously described method (Schneider *et al.*, 1993) was used to determine distributions and the most probable positions of water molecules around unmodified bases in the Z-type conformation.

### 3. Results and discussion

The reflection sphere, virtually complete up to 1.3 Å, allowed for restrained anisotropic temperature-factor refinement, which improved map quality, and a final three cycles of unrestrained full-matrix refinement. Therefore, the geometric parameters of this structure are comparable to small-molecule structures. The average values of estimated standard deviations of bond lengths and angles are 0.025 Å (range 0.012–0.05 Å) and 2° (range 0.6–4°). A comparison with other Z-DNA hexamers (Table 3) shows that the reported structure is based on more reflections than any other structure with a resolution of 1.3 Å and that only four structures solved at a nominal resolution of 1.0 Å were refined using more reflections. The quality of the refined geometry and the completeness of the reflection sphere confirms the value of synchrotron radiation for obtaining diffraction data from single crystals at cryo-temperatures and the advantage of utilizing intensities for refinement.

#### 3.1. Helix structure

An overall view of the structure and packing is given in Fig. 1. Helical and base geometry parameters are listed in Table 2. The structure of d(TGCGCA)<sub>2</sub> is in the Z-DNA conformation and all its base pairs have Watson–Crick geometry. The average helical twist is –62° per dinucleotide and all the helical parameters fall within the range of those previously reported for Z-DNA structures. Also, the values of the backbone torsions are within the ranges expected for a Z-type structure (Schneider *et al.*, 1997). The purine nucleotides

Table 1. Summary of X-ray data-collection and crystallographic-refinement statistics

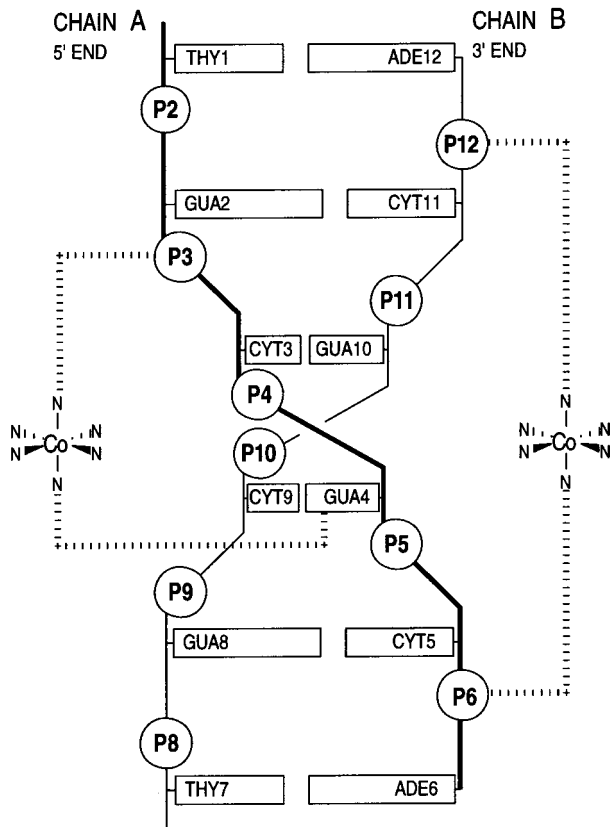
Data collection.			
Values in parentheses correspond to the highest resolution shell, 1.32–1.30 Å (305 reflections).			
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>		
Unit-cell dimensions (Å)	$a = 21.162, b = 28.670, c = 44.335$		
Solvent content (%)	38		
Resolution (Å)	45–1.3		
Unique observations	7049		
Average redundancy	4		
Completeness (%)	99.2 (87.4)		
Average $I/\sigma I$	>20 (23.6)		
$R = \sum  I - \langle I \rangle  / \sum  I $	0.043 (0.078)		
Refinement			
Stage	iso†	ani‡	ani‡
	CGLS§	CGLS§	FMLS¶
	ZI/ZII††	ZI/ZII††	ZI/ZII††
	H‡‡	H‡‡	H‡‡
No. of waters/ No. of Co(NH <sub>3</sub> ) <sub>6</sub>	78/2	78/2	78/2
Data/parameter	4.7	2.1	2.3
$R_{\text{cryst}}§§$	0.154	0.120	0.117
No. of data	6310	6310	7005
$R_{\text{free}}¶¶$	0.187	0.173	—

† Isotropic temperature-factor refinement. ‡ Restrained anisotropic temperature-factor refinement. § Conjugate gradient least-squares refinement. ¶ Full-matrix least-squares refinement. †† Both ZI and ZII conformations refined for P11. ‡‡ 'Riding' H atoms placed in calculated positions using the following parameters:  $C_{sp2}-H = 0.93$  Å,  $C_{sp3}-H = 0.96$  Å,  $U_{\text{iso}}(H[C_{sp2}]) = 1.2 U_{\text{eq}}(C_{sp2})$  and  $U_{\text{iso}}(H[C_{sp3}]) = 1.5 U_{\text{eq}}(C_{sp3})$ . §§  $R_{\text{cryst}} = \sum ||F_o| - |F_c|| / \sum |F_o|$  for all data in 10–1.3 Å range. ¶¶ As for  $R_{\text{cryst}}$  but calculated for a test set comprising 695 reflections not used in the refinement.

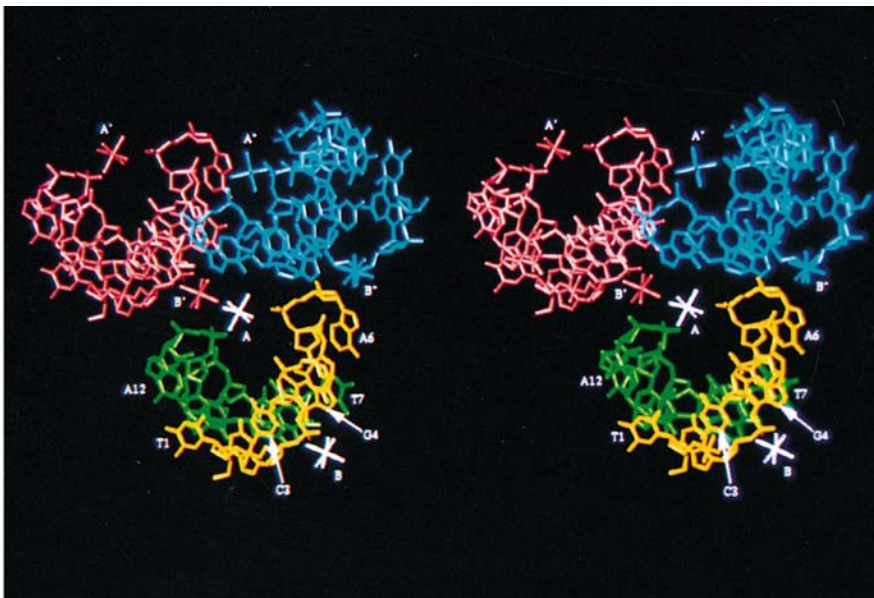
have the glycosidic torsion,  $\chi$ , in the *syn* range from 59 to 75°. Their deoxyribose rings are in the C3'-*endo* conformation, except for A12 which is in the C2'-*endo* conformation, possibly due to a 3' end effect. Indeed, all but four deoxyriboses of terminal purines have C2'-*endo* pucker in Z-DNA hexamers with deposited coordinates while the sugar pucker in purines inside strands is exclusively C3'-*endo*. The pyrimidine nucleosides in d(TCCGCA)<sub>2</sub> have *anti* glycosidic  $\chi$  angles and C2'-*endo* sugar puckers in agreement with other Z-DNA structures.

This structure maintains the previously described ZI/ZII conformational mix (Wang *et al.*, 1991) which is characterized by a difference in the conformations of the two strands in the duplex. In ZI there is a marked difference between values of torsions  $\alpha$ ,  $\beta$ ,  $\epsilon$  and especially  $\zeta$  in purine and pyrimidine residues. In ZII the regular alternation of these torsions is broken and they acquire similar values in both residues. The ZII conformation has previously been found mostly at phosphate 5 or, by an approximate twofold symmetry, at P11 (Table 3). In d(TGCGCA)<sub>2</sub>, P11 and neighbouring atoms on strand *B* are disordered so that the ZI and ZII

† Atomic coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 362D).



(a)



(b)

Fig. 1. (a) The d(TGCGCA)<sub>2</sub> crystal structure. Bases are represented as rectangles and labelled according to residue type and number, PO<sub>2</sub> groups are drawn as circles and labelled by residue numbers. The two cobalt hexamine complexes are marked and their connection to the DNA residues 3 and 4 and to residues 6 and 12 are indicated. (b) A stereoview of the d(TGCGCA)<sub>2</sub> duplex with strand A (residues 1-6) in yellow, strand B (residues 7-12) in green and the two cobalt hexamines, A and B, in white. Two symmetry-related asymmetric units, in blue ( $\frac{1}{2}-x, -\frac{1}{2}-y, -z$ ) and brown ( $1+x, y, z$ ), are drawn to illustrate the immediate packing environment of the cobalt hexamine complexes. The key residues are labelled, the symmetry-related complexes are indicated by primes. Water O atoms are omitted for clarity.

conformations coexist in approximately the same occupancies rendering the two strands conformationally distinct. Although there is at least one report of a phosphate disordered into ZI and ZII conformations (Egli *et al.*, 1991), there are no coordinates publicly available which have both conformations at any one step.

Besides the disordered ZI/ZII stretch at phosphate 11, the ZII conformation is also located at quite untypical positions at P3 and P9 (Table 3). Such an unusual position might be explained by the stabilizing effect of one of the two cobalt hexammine complexes which links the charged O atoms of P3 and of symmetry-related P9. It perhaps supports the observation that the ZII conformation is stabilized by nearby polyvalent metal cations (Wang *et al.*, 1979). However, their presence is not strictly required because ZII has also been observed in the 'pure spermine' form of Z-DNA (Egli *et al.*, 1991). The ZII conformation has been observed at P9 in only two other Z-DNA hexamers (Kagawa *et al.*, 1991; Bancroft *et al.*, 1994) and at P3 in another structure (Schroth *et al.*, 1993). No special feature, such as a metal complex attached to these residues, analogous to the cobalt hexammine in the present structure, has been observed in any of these structures.

Apart from the difference between conformations at phosphates 5 and 11 of this and most other Z-DNA hexamers, the backbone torsion angles approximately respect the twofold symmetry relating two palindromic strands. The values of parameters describing base-pair morphology as buckle and propeller twist comply only very approximately with the twofold axis and appear to be much more dependent on environment. Among the Z-DNA sequences prevail CG and GC steps. If base-pair parameters were primarily determined by sequences they should display more regular behaviour.

### 3.2. Packing

Like most other Z-DNA hexamer structures reported to date, d(TGCGCA)<sub>2</sub> crystallizes in the orthorhombic space group  $P2_12_12_1$ . All these structures crystallize in such a way that the stacking between terminal base pairs of each double-stranded hexanucleotide forms a pseudo-continuous helix parallel to the *c*-axis and d(TGCGCA)<sub>2</sub> is no exception. The packing of d(TGCGCA)<sub>2</sub> proved to be one of the two already known packing modes, 1 and 2, employed by Z-DNA hexamers in  $P2_12_12_1$  (Table 3). The reported structure has shorter *b* and *c* and longer *a* axes when compared with cell dimensions of structures crystallized in mode 1. This is in agreement with the trend observed for other structures in packing mode 2.

There is a total of 30 hexamer Z-DNA structures with deposited coordinates in the NDB (Table 3), 25 of which are in the space group  $P2_12_12_1$ . A majority of these, 20 structures, are in one packing mode, mode 1, previously

called the 'magnesium form' (Gessner *et al.*, 1989). The remaining five pack in a distinct mode 2, previously referred to as the 'spermine form' (Egli *et al.*, 1991). A transition from one mode to the other requires rotation of a DNA hexamer molecule around the crystallographic axis *c* by about 70° and its shift along this axis by 3 Å (Egli *et al.*, 1991). The rotation and translation of a Z-DNA molecule inside the cell leads to different contacts between symmetry-related Z-DNA double helices in both modes. The differences can be illustrated by contacts of phosphate groups to neighbouring symmetry-related DNA molecules. In mode 1, phosphate 5, P5, always contacts some atoms of symmetry-related residue 8, P8 contacts residue 4, and P9 and P10 contact residue 6. Mode 2 is characterized by always containing contacts between P3 and P4 and symmetry-related atoms of residue 6, between P5 and P6 and residue 4, and between P9 and residue 8.

The existence of two closely related packing modes for Z-DNA hexamers was rationalized by the presence or lack of inorganic polyvalent cations, mainly Mg<sup>2+</sup> and Co<sup>3+</sup> (Egli *et al.*, 1991) in the crystallization solution. The presence of Mg<sup>2+</sup> undoubtedly contributes to the stability of the first packing mode and Mg<sup>2+</sup> is observed in crystallographically ordered positions in several of these structures (Table 3). However, the presence of polyvalent metal cations is not a sufficient condition for obtaining a structure in mode 1 as there are structures crystallized with Mg<sup>2+</sup> and spermine (Van Meervelt *et al.*, 1990) and with Mg<sup>2+</sup> alone (Moore *et al.*, 1995) which adopt the second packing mode. Moreover, the reported structure was crystallized with a complex of another polyvalent metal, Co(NH<sub>3</sub>)<sub>6</sub>, and this complex is well localized and involved in stabilizing the structure in its 'pure spermine' packing mode 2.

Another hypothesis, put forward in an attempt to explain the packing of Z-DNA hexamers, related packing to the ZI/ZII backbone irregularity and to the water structure (Schneider *et al.*, 1992). The phosphate groups 5 and 11, which significantly disrupt the potential twofold palindromic symmetry of double helices in many Z-DNA structures, are possibly important for packing into a crystallographic lattice. Since ZII is at the same position in 14 of the 21 structures in packing mode 1 (Table 3) there is probably a positive feedback between this particular position of the ZII bulge in the backbone and packing mode 1. The relation is, however, more of a statistical than a causal nature because ZII does not have to be present in a structure at all or it can exist at P3 (Schroth *et al.*, 1993) or P9 (Kagawa *et al.*, 1991). Five particular water bridges linking these very phosphates, P5 and P11, to symmetry-related DNA molecules are conserved in about 70% of the structures adopting packing mode 1. Again, these water bridges probably support packing into mode 1 but they are not observed in all of these structures. Packing mode 2 shows no regularity in distribution of the ZII confor-

Table 2. Geometrical characteristics of the double helix.

For definition of base-pair parameters and torsion angles see Dickerson *et al.* (1989)

## Torsion angles

Code	$\alpha$	$\beta$	$\gamma$	$\delta$	P	$\epsilon$	$\zeta$	$\chi$
T1	nd	nd	49.2	146.0	155.9	266.2	82.0	215.7
G2	61.4	188.5	177.2	95.5	13.7	179.5	60.6	63.3
C3	168.9	158.6	45.7	139.9	147.4	270.4	79.2	210.1
G4	61.3	195.4	177.9	90.6	36.9	235.4	302.0	60.4
C5	208.0	227.0	58.2	133.0	145.5	265.6	74.9	200.1
A6	66.8	188.8	178.6	87.8	48.2	nd	nd	70.4
T7	nd	nd	62.4	144.3	150.9	269.1	76.6	198.5
G8	61.9	191.5	177.1	94.6	22.2	177.5	58.4	58.8
C9	164.8	176.2	47.3	140.8	154.0	268.7	69.9	214.8
G10	65.8	183.7	175.4	91.8	33.7	254.8	278.7	66.1
C11	204.8	243.9	58.7	135.8	151.5	266.0	79.1	202.1
C11†	197.2	58.7	135.8	266.0	151.5	79.1	74.3	202.1
A12	74.3	192.7	178.5	144.0	147.2	nd	nd	75.2
A12†	192.7	178.5	144.0	352.3	147.2	nd	nd	75.2

## Base-step parameters

Step	Roll	Tilt	Cup	Slide	Twist	Rise
T1–G2	2.92	–4.01	6.06	3.76	13.9	3.55
G2–C3	2.21	–0.40	0.74	–0.43	53.4	3.34
C3–G4	1.72	2.41	–0.26	5.28	10.0	3.76
G4–C5	5.27	–0.38	–9.66	–1.21	46.5	3.94
C5–A6	0.65	–9.03	16.89	3.38	9.3	3.56

† Second conformation of the disordered backbone.

mation along the helix and ZII may or may not be present (Table 3).

The most enigmatic is a unique trigonal packing of the sequence d(CGTA'CG)<sub>2</sub> (Parkinson *et al.*, 1995). It is triggered by a single methyl group at the thymine residue 3; uracil at this position brings about the most common packing mode 1 in the structure ZDFB31 (Schneider *et al.*, 1992). In the trigonal structure the double-stranded hexamers do not stack to form continuous helices, but instead the end base pair stacks on the surface of the major groove of a neighbouring helix. Another unique feature of the trigonal structures is that they have just one strand in the asymmetric unit. Despite its fundamentally different and rare packing the crystals of this sequence have diffraction quality comparable with other Z-DNA structures.

Despite its apparent simplicity, packing of Z-DNA hexamers is still elusive. There are factors which are known to contribute to the stability of a certain packing mode but their presence is not strictly required and, on the other hand, another packing can be adopted even when they are present.

### 3.3. Interactions with metal cations and water

**3.3.1. Cobalt hexammine.** Two Co(NH<sub>3</sub>)<sub>6</sub> complexes were localized in the asymmetric unit of the reported structure. Both are involved in packing interactions (Fig. 1). One of them, called *B*, might stabilize the ZII conformation by interacting with a charged O atom of

phosphate 3 and guanine atoms O6 and N7 of the neighboring residue G4. The other complex, *A*, is located at the periphery of the minor groove and is involved in cross-strand bridging of phosphate groups 6 and 12 of the same duplex. The complex *A* and the symmetry-related complex *B*, *B'*, are located close to each other with a Co–Co distance of 6.7 Å, and their neighbouring ammonium groups are bridged by a water molecule (Fig. 1).

The only other cobalt hexammine localized in a Z-DNA structure is coordinated to an analogous position between the third phosphate and the fourth purine base but of the other chain, *B*, and the structure, NDB code ZDF019 (Gessner *et al.*, 1985), crystallizes in mode 1. The interacting P9 is also in the ZII conformation as is P3 in the reported structure. No further comparison can be made since the coordinates of the structure ZDF019 are not publicly available. One decamer structure in untypical conformation, NDB code UDJ032 (Nunn & Neidle, 1996), also has two cobalt hexammine complexes in the asymmetric unit. Two symmetry-related complexes are in relatively close contact (Co–Co distance 8.0 Å) but they are not directly linked by a water bridge.

**3.3.2. Magnesium and sodium.** There are 20 ordered Mg<sup>2+</sup> and Na<sup>+</sup> cations reported in ten Z-DNA hexamers. None of these cation positions coincides with either of cobalt hexammine sites.

All cations but two Na<sup>+</sup> have been localized in packing mode 1 (Table 3) and these 18 cations are

Table 3. *Z-DNA hexamers with coordinates deposited in the NDB*

ZII, location of the ZII conformation in a structure. Packing, packing mode of the structure crystallized in the  $P2_12_12_1$  space group.  $R$ , crystallographic  $R$  factor. Refl, number of reflections used for refinement. Resol, reported resolution in Å. Ref, reference.

NDB code	ZII†	Packing	Ordered cations‡	Other cations§	$R$	Refl	Resol	Hydr¶	Ref	Notes††
ZDF001	5	1	Mg; spermine/2	no	14.0	15000	0.90	Y	<i>a</i>	
ZDF002	5	1	Mg/4	Na	17.5	10893	1.00	Y	<i>b</i>	
ZDF052	5	1	Mg/3; Na; spermidine	Na	19.1	6301	1.00	Y	<i>c</i>	
DDF027	5	1	Mg/3; Na; AmBu‡‡	Na	16.1	8921	1.00	Y	<i>d</i>	
ZDFB11	5	1	Mg	Na	21.7	3739	1.30	Y	<i>e</i>	Chains A/B swapped
ZDFB14	5	1	Mg	Na, spermine	15.6	1105	2.25	N	<i>f</i>	
ZDFB24	5	1	Mg/2	Na	20.8	2870	1/30	Y	<i>g</i>	
ZDF028	5, 9	1	Cu/6; Na	Mg, spermine	19.8	4107	1.20	Y	<i>h</i>	
ZDFB03	5	1	no	Na	15.6	4208	1.30	N	<i>i</i>	
ZDFB06	5	1	no	Mg, Na, spermine	16.0	5412	1.20	N	<i>j</i>	
ZDFB10	5	1	no	Mg, Na, Cu	20.9	2587	1.30	Y	<i>k</i>	
ZDFB21	5	1	no	Mg, Na, spermine	19.0	1217	1.90	Y	<i>l</i>	
ZDFB31	5	1	no	Mg, Na, spermine	13.8	4588	1.30	Y	<i>m</i>	
ZDF039	no	1	no	no	19.9	2625	1.60	Y	<i>n</i>	
ZDFB04	no	1	no	Na	13.2	3195	1.60	Y	<i>o</i>	
ZDFB05	no	1	no	Na	12.4	3765	1.40	Y	<i>o</i>	
ZDFB12	no	1	Mg	Spermine, tris	17.2	2770	1.50	Y	<i>p</i>	
ZDFB36	no	1	no	Mg, Na, spermine	18.9	1495	1.80	Y	<i>q</i>	
ZDFB37	3,5	1	no	Mg, spermidine	19.3	5029	1.30	N	<i>r</i>	Disordered, duplex 1
ZDFB37	no	1	no	—	—	—	—	—	<i>r</i>	Disordered, duplex 2
ZDFB51	no	1	no	no	17.0	3151	1.65	Y	<i>s</i>	
ZDF029	no	2	Spermine	no	17.0	8183	1.00	Y	<i>t</i>	
ZDF035	9	2	Na/2; spermine/2	no	18.0	9575	1.00	Y	<i>u</i>	
ZDFB25	no	2	no	Mg, Na, spermine	18.1	2559	1.70	Y	<i>v</i>	
ZDFB43	no	2	no	Mg, Na	17.9	2798	1.70	Y	<i>w</i>	
ZDF060	3,9,11	2	Co(NH <sub>3</sub> ) <sub>6</sub> /2	Arginina-mide	11.7	7005	1.30	Y		Reported structure
UDF025	no	—	no	Na, Arg, spermine	18.5	2297	1.90	Y	<i>x</i>	Open end bases
ZDF038	§§	—	no	Na, Ba	16.1	530	2.50	N	<i>n</i>	$P2_1$
ZDF040	no	—	no	Mg, Na, spermine	19.9	816	2.20	N	<i>y</i>	$P1$ -, racemic
ZDFB41	no	—	no	Mg, Na, spermine	17.4	2957	1.36	Y	<i>z</i>	$P3_221$
ZDFB42	no	—	Spermine, Pt	Mg, Na, spermine	17.0	1583	1.60	Y	<i>z</i>	$P3_221$

† 'no' means that no nucleotide has been observed in the ZII conformation. ‡ Cations observed in ordered positions. The numbers indicate how many positions were refined. § Cations present in crystallization batch but not observed in the refined structure. ¶ Structures used to study hydration of the bases are labelled 'Y'. †† Space groups of structures that do not crystallize in the space group  $P2_12_12_1$  are indicated. ‡‡ N-(2-aminoethyl)-1,4-diaminobutane. §§ Structure with less regular backbone conformation than other Z-DNA structures. References: (a) Wang *et al.* (1979), (b) Gessner *et al.* (1989), (c) Ohishi *et al.* (1996), (d) Ohishi *et al.* (1991), (e) Coll *et al.* (1986), (f) Brown *et al.* (1986), (g) Zhou & Ho (1990), (h) Kagawa *et al.* (1991), (i) Fujii *et al.* (1982), (j) Wang *et al.* (1984), (k) Geierstanger *et al.* (1991), (l) Ginell *et al.* (1990), (m) Schneider *et al.* (1992), (n) Sadisivan & Gautham (1995), (o) Chevrier *et al.* (1986), (p) Coll *et al.* (1989), (q) Cervi *et al.* (1993), (r) Schroth *et al.* (1993), (s) Peterson *et al.* (1996), (t) Egli *et al.* (1991), (u) Bancroft *et al.* (1994), (v) Meervelt *et al.* (1990), (w) Moore *et al.* (1995), (x) Malinina *et al.* (1995), (y) Doi *et al.* (1993), (z) Parkinson *et al.* (1995).

observed in five different positions labelled *A*, *B*, *C*, *D* and *E* in Table 4. Positions *A* and *B* are located in the major groove, and cations are coordinated mainly by

guanine hydrophilic atoms. Of those contacts only one in site *A* is a direct  $Mg^{2+}$ -base contact forming a so-called 'inner-sphere complex'. All other metal-base contacts

Table 4. Coordination of metal cations in the reported structure (ZDF060) and other Z-DNA hexamers

The six different locations of Mg<sup>2+</sup> and Na<sup>+</sup> are labelled alphabetically under 'Site' (A–G) and the cobalt hexammine sites in the reported structure are labelled COI and COII. Contacts to symmetry-related helices are indicated by asterisk(s). Contacts in parentheses were not observed in all structures.

Site	Contacts†	Metal	Structures‡
A	N7 G6 (direct), O6 G12*, N4 C1*, (O1P C5)	Mg <sup>2+</sup>	ZDF002, ZDF052, ZDF001, DDF027
B	O6/N7 G10, O6/N7 G12*, (O6 G6**)	Mg <sup>2+</sup>	ZDF002, ZDF052, ZDFB11, DDF027
C	O2P pu6, O1P pu10, N7 pu8*, (O6 pu4*)	Mg <sup>2+</sup>	ZDF002, ZDF052, ZDFB24, DDF027
D	O2P py5, O1P pu6, O1P py9* (may be direct), O2P pu10*, (O1P pu8*)	Mg <sup>2+</sup>	ZDF002, ZDFB14, ZDFB24
D	O2P py5, O1P pu6, O1P py9* (may be direct), O2P pu10*, (O1P pu8*)	Na <sup>+</sup>	ZDF052, DDF027
E	N4 py1, N4 py11, N7 pu10, O2P py11*	Na <sup>+</sup>	ZDF028
F	O1P py5, O2P pu12*, N7 pu4**	Na <sup>+</sup>	ZDF035
G	O2P py5 (direct), O3' pu4 (direct), O2P pu6*, C8 pu4**	Na <sup>+</sup>	ZDF035
COI	O2P pu6, O1P pu12	Co <sup>3+</sup>	ZDF060, ZDF019
COII	O1P py3, O6 pu4	Co <sup>3+</sup>	ZDF060

† Listed are either direct metal–DNA contacts or contacts mediated by a water molecule (metal–water–DNA). The upper limit for Mg<sup>2+</sup> and Co<sup>3+</sup> contacts was 2.40 Å, 2.70 Å for Na<sup>+</sup> and 3.40 Å for O–O and O–N hydrogen-bonding interactions. ‡ NDB codes of structures. For references see Table 3.

are mediated by water molecules so that 'outer-sphere contacts' prevail. Mg<sup>2+</sup> cations bound directly to the guanine N7 lie in the base plane with a rather long Mg<sup>2+</sup>–N7 distance of about 2.30 Å. Water positions around guanine N7 are, on the other side, below and above the base plane (see §3.3.3 on base hydration). Different behaviour of Mg<sup>2+</sup> and water can be explained by different optimal distances at which Mg<sup>2+</sup> and water form their non-bonding contacts: water can bridge two stacked bases while Mg<sup>2+</sup> cannot.

In position C, Mg<sup>2+</sup> is coordinated *via* water to two phosphate groups and symmetry-related purine and pyrimidine bases. It is only the fourth position, D, in which metals bind exclusively to phosphate groups. Of all contacts between five metal cations bound in position D and phosphates, only one Mg<sup>2+</sup> and one Na<sup>+</sup> bind to phosphate O atoms directly forming 'inner-sphere complexes'. These two sites lie close to the OP–P–OP plane with the M<sup>+</sup>–OP–P angle about 150°. This is in general agreement with the averaged positions of these cations determined for structures of organic phosphates (Schneider & Kabeláč, 1998). Two Na<sup>+</sup> cations observed in packing mode 2 in structure ZDF035 (positions F and G) bind close to each other (4.3 Å) between the same base, the same symmetry-related phosphate and another phosphate which is different for each metal.

A small proportion of phosphate binding and very few direct metal–phosphate contacts (inner-sphere binding) is a surprising feature of Z-DNA crystal structures. Thermodynamic studies (Sigel & Sigel, 1996) as well as a database study of crystal structures of small molecule phosphate salts (Schneider & Kabeláč, 1998) have indicated a preference for direct binding between Mg<sup>2+</sup> or Na<sup>+</sup> and charged O atoms of a phosphate group.

Direct binding of Mg<sup>2+</sup> to a guanine N7 may be a corollary of crystal packing when the neighbouring

double helices sterically accommodate an [Mg(H<sub>2</sub>O)<sub>6</sub>]<sup>2+</sup> complex. Another mechanism supporting interactions between Mg<sup>2+</sup> and guanine N7 is the strengthening of the guanine–cytosine pairing by cation binding to N7 or O6. The exocyclic O6 can participate in the coordination sphere of the cation. On the other hand, Mg<sup>2+</sup> never directly binds to adenine not only because it is rare in Z-DNA structures but because it has a much weaker dipole moment than guanine. Adenine–metal interaction is also destabilized by the presence of the adenine amino group N6 (Šponer *et al.*, 1997) when the pyramidalization of N6 H atoms induced by the interaction with a cation destabilizes the adenine–thymine base pairing (Šponer *et al.*, 1996).

Close contacts between metal cations are quite common in Z-DNA structures. Very close contacts have been observed between Mg<sup>2+</sup> cations in structures ZDF002, DDF027 and ZDF052 where Mg<sup>2+</sup> in position A is 2.8 Å from Mg<sup>2+</sup> in position B and the interaction is mediated by two shared water molecules. Two cobalt hexammine complexes which have formal charges +3 are located at a distance of 6.7 (present structure) and 8.0 Å (Nunn & Neidle, 1996). Such a clustering of positively charged metal complexes seemingly defies electrostatics. Electrostatic repulsion could perhaps be compensated by forces of the periodic crystal field (Suhai, 1995) so that compact crystal arrangements become stable.

3.3.3. *Water.* Z-DNA crystal structures contain many crystallographically ordered water molecules and, therefore, offer a good opportunity to study the structure of the first hydration shell. Here we survey hydration around bases. A first-shell water molecule was defined as being no further than 3.40 Å from any base atom. The 24 structures used for the study are indicated in Table 3.

The analysed structures contain 110 unmodified guanines interacting with 349 water molecules,



Table 5. *Geometry of the hydration sites around guanine and cytosine in Z-DNA*

Atoms			Hydration site (S/W) <sup>†</sup>	Distance C—S (Å) <sup>‡</sup>	Angle B—C—S (°)	Torsion A—B—C—S (°)	Relative density <sup>§</sup>
A	B	C					
Guanine							
N1	C2	N2	S1	2.98	123	160	24
C5	C6	O6	W1a	2.78	113	53	12
C5	C6	O6	W1b	2.88	134	-39	7
N9	C8	N7	W2a	2.67	120	156	8
N9	C8	N7	W2b	2.98	110	-149	5
			W1a	W1b	3.48		
			W2a	W2b	2.41		
			W1a	W2b	3.02		
			W1b	W2a	2.48		
Cytosine							
N1	C2	O2	S1a	2.85	144	69	16
N1	C2	O2	S1b	2.87	136	-70	12
N3	C4	N4	W1	3.00	114	175	24
		S1a	S1b	3.46			

<sup>†</sup> Hydration sites are labelled 'S' in the minor and 'W' in the major groove. The same nomenclature is used as in Schneider *et al.* (1993). <sup>‡</sup> Distances are listed when shorter than 3.40 Å. <sup>§</sup> Relative density is the peak height measured as a multiple of the estimated standard deviation of a map.

accounting for 3.2 stoichiometric waters per base. The Fourier-averaged distribution of water around these guanines shows a single high-density peak near the minor-groove atom N2 (Fig. 2a) and a structured cloud of water density hydrating major-groove atoms O6 and N7. The minor-groove site has a large density (Table 5) which indicates high probability of water occurrence at or near this position. Major-groove hydration is less focused with a clear preference for non-planar hydration sites. O6 is hydrated by two hydration sites, the stronger of which is situated above the base plane in the 5' direction of the polynucleotide chain. N7 is also hydrated by two hydration sites but the second one is of marginal significance. The distances between the four major-groove hydration sites (Table 5) indicate that four water molecules can simultaneously occupy these positions but, as suggested by the low density of the fourth site, hydration by three water molecules is more common.

Density of water around cytosine was calculated from 228 water molecules which were found in proximity to the 93 cytosines, corresponding to 2.5 stoichiometric waters per base. The minor-groove atom O2 has two hydration sites, one below and one above the base plane (Fig. 2b, Table 5) but, as the lower density envelope wrapping up these two sites suggests, water can be distributed in a quite extended region perpendicular to the base plane. The major groove has just one well defined planar-site hydrogen bonding to N4.

The reported distributions of water around guanine and cytosine are similar to their values obtained in an older analysis based on a much smaller set of structures (Schneider *et al.*, 1993). A qualitative similarity of water densities as well as quantitative agreement between the new and old positions of hydration sites further confirms robustness of the protocol used for compiling structural data on the first hydration shell.

A low number of adenines and thymines in Z-DNA oligonucleotides allows only qualitative discussion of their water distributions. Adenine major-groove

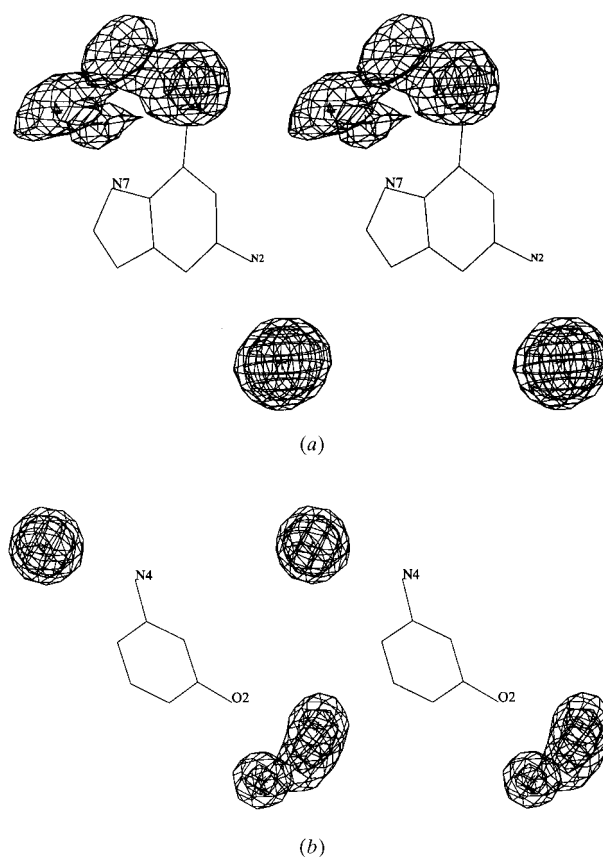


Fig. 2. A stereoview of water distributions around (a) guanine and (b) cytosine in Z-DNA structures. The minor-groove atoms N2/O2 are on the bottom right, the major-groove N7/N4 on the top.

hydration has similar features to guanine, namely an extensive cloud of water density and peak densities far from the base plane. A significant difference is that the site with the highest density which hydrated O6 in guanine moves closer to N7 in adenine so that it hydrates both N6 and N7. The shift of the position of the main hydration site between adenine and guanine is in agreement with their different electrostatic properties. A potential to form hydrogen bonds can be satisfied by only one water in adenine while guanine requires separate hydration of O6 and N7 by up to four water molecules. A similar difference has been observed in the B conformation where two major-groove hydration sites can simultaneously be occupied by two water molecules, but they become much closer in adenine because many waters actually prefer the position between N6 and N7 (Schneider & Berman, 1995).

The water distributions determined for isolated bases can be incorporated into any sequence (for details see Schneider *et al.*, 1993) so we modelled base hydration of the reported structure ZDF058 and compared predicted hydration sites to crystallographic water molecules. The most dominant feature of minor-groove hydration is the spine of water density in the centre of the minor groove (Fig. 3*a*) which is generated by water molecules hydrogen bonded to pyrimidine O2 atoms. The hydration sites of high density immersed in a tube of lower density are hydrogen bonded to two pyrimidines from consecutive steps. The spine is flanked by spherical hydration sites hydrogen bonded to the guanine N2 atoms. A slight shift of guanine water density from the base plane in the 5' direction renders 5'-CG-3' and 5'-GC-3' steps with different hydration patterns. The guanine sites are in all cases too far from the cytosine sites to form direct hydrogen bonds. Prediction of

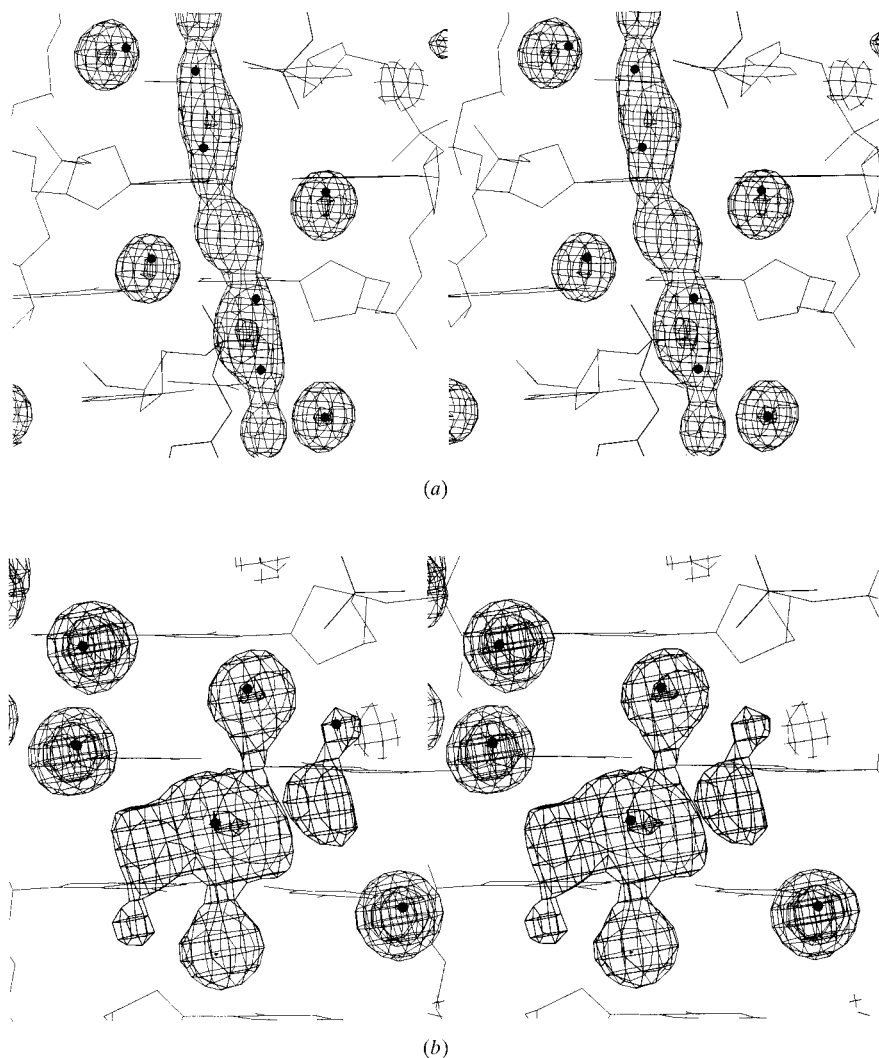


Fig. 3. Predicted water density and actual water positions (dots) in the reported structure. A stereoview (a) shows the minor groove of the central part 5'-GCGC-3' with the 5' ends on the top left and bottom right. A view into the major groove (b) shows the 5'-GCG-3' triplet, the 5' end guanine is on the lower left. The guanine from the top base pair is partially outside the figure frame. All bases are approximately perpendicular to the paper plane. Due to the depth clipping not all DNA atoms are visible.

guanine waters is very good (Fig. 3*a*) and all predicted sites are occupied by crystallographic water molecules. Prediction of cytosine water molecules inside the spine is more ambiguous but still satisfactory: water molecules are located at the periphery of high-density regions of the spine.

In the major groove, the 5'-CG-3' and 5'-GC-3' steps have completely different hydration patterns (Fig. 3*b*). The CG steps (upper left in Fig. 3*b*) have spherical cytosine hydration sites aligned along the vertical helical axis. Water molecules in the positions of these two sites form hydrogen bonds to each other. Hydration of the GC steps is dominated by water density between the two guanines from the opposite strands. The central part of Fig. 3*b*) shows that water distributions around individual guanines largely overlap so that their densities fuse into one cloud of density common to both guanines. Prediction is successful for the cytosine N4 sites and reasonable for sites in a complex hydration pattern of guanines.

This research was funded by the BBSRC to the Protein Structure Group at the University of York and in part by CEC funding to MB. BS was supported by the EU programme Copernicus (CIPA-CT94-0189) and the Grant Agency of the Czech Republic (204/95/0036).

#### References

- Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.
- Baeyens, K. J., Jancarik, J. & Holbrook, S. R. (1994). *Acta Cryst.* **D50**, 764–767.
- Bancroft, D., Williams, L. D., Rich, A. & Egli, M. (1994). *Biochemistry*, **33**, 1073–1086.
- Behe, M. & Felsenfeld, G. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 1619–1623.
- Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys. J.* **63**, 751–759.
- Brown, T., Kneale, G., Hunter, W. N. & Kennard, O. (1986). *Nucleic Acids Res.* **14**, 1801–1809.
- Cervi, A. R., Guy, A., Leonard, G. A., Teoule, R. & Hunter, W. N. (1993). *Nucleic Acids Res.* **21**, 5623–5629.
- Chevrier, B., Dock, A. C., Hartmann, B., Leng, M., Moras, D., Thuong, M. T. & Westhof, E. (1986). *J. Mol. Biol.* **188**, 707–719.
- Coll, M., Saal, D., Frederick, C. A., Aymami, J., Rich, A. & Wang, A. H.-J. (1989). *Nucleic Acids Res.* **17**, 911–923.
- Coll, M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1986). *J. Biomol. Struct. Dyn.* **4**, 157–172.
- Dickerson, R. E. (1993). *Newhel93. A Program for Calculation of Helix Parameters of Nucleic Acids*. University of California at Los Angeles, Los Angeles, CA, USA.
- Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H. C. M., Olson, W., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., Wang, A. H.-J. & Zhurkin, V. B. (1989). *EMBO J.* **8**, 1–4.
- Doi, M., Inoue, M., Tomoo, K., Ishida, T., Ueda, Y., Akagi, M. & Urata, H. (1993). *J. Am. Chem. Soc.* **115**, 10432–10433.
- Egli, M., Williams, L. D., Gao, Q. & Rich, A. (1991). *Biochemistry*, **30**, 11388–11402.
- Fox, K. R. & Waring, M. J. (1984). *Biochim. Biophys. Acta*, **802**, 162–168.
- Fox, K. R. & Waring, M. J. (1986). *Biochemistry*, **25**, 4349–4356.
- Fujii, S., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1982). *Nucleic Acids Res.* **10**, 7879–7892.
- Geierstanger, B. H., Kagawa, T. F., Chen, S.-L., Quigley, G. J. & Ho, P. S. (1991). *J. Biol. Chem.* **266**, 20185–20191.
- Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A. & Wang, A. H.-J. (1989). *J. Biol. Chem.* **264**, 7921–7935.
- Gessner, R. V., Quigley, G. J., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1985). *Biochemistry*, **24**, 237–240.
- Ginell, S. L., Kuzmich, S., Jones, R. A. & Berman, H. M. (1990). *Biochemistry*, **29**, 10461–10465.
- Kagawa, T. F., Geierstanger, B. H., Wang, A. H.-J. & Ho, P. S. (1991). *J. Biol. Chem.* **266**, 20175–20184.
- Malinina, L., Urpi, L., Salas, X., Huynh-Dinh, T. & Subirana, J. A. (1994). *J. Mol. Biol.* **243**, 484–493.
- Meervelt, L., van, Moore, M. H., Lin, P. K. T., Brown, D. M. & Kennard, O. (1990). *J. Mol. Biol.* **216**, 773–781.
- Moers, B. H., Schroth, G. P., Baxter, W. W. & Ho, P. S. (1995). *J. Mol. Biol.* **249**, 772–784.
- Moore, M. H., Hunter, W. N., D'Estaintot, B. L. & Kennard, O. (1989). *J. Mol. Biol.* **206**, 693–705.
- Moore, M. H., van Meervelt, L., Salisbury, S. A., Lin, P. K. T. & Brown, D. M. (1995). *J. Mol. Biol.* **251**, 665–673.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
- Nunn, C. M. & Neidle, S. (1996). *J. Mol. Biol.* **256**, 340–351.
- Ohishi, H., Kunisawa, S., van der Marel, G. A., van Boom, J. H., Rich, A., Wang, A. H.-J., Tomita, K. & Hakoshima, T. (1991). *FEBS Lett.* **284**, 238–244.
- Ohishi, H., Nakanishi, I., Inubushi, K., van der Marel, G. A., van Boom, J. H., Rich, A., Wang, A. H.-J., Hakoshima, T. & Tomita, K. (1996). *FEBS Lett.* **391**, 153–156.
- Oldfield, T. J. (1994). *Proceedings of the CCP4 Study Weekend*, edited by S. Bailey, R. Hubbard & D. Waller, pp. 15–16. Warrington: Daresbury Laboratory.
- Otwinowski, Z. (1991). *DENZO. A Film Processing Program for Macromolecular Crystallography*. Dallas, TX.
- Parkinson, G. N., Arvanitis, G. M., Lessinger, L., Ginell, S. L., Jones, R., Gaffney, B. & Berman, H. M. (1995). *Biochemistry*, **34**, 15487–15495.
- Peterson, M. R., Harrop, S. J., McSweeney, S. M., Leonard, G. A., Thompson, A. W., Hunter, W. N. & Helliwell, J. R. (1996). *J. Synchrotron Rad.* **3**, 24–34.
- Pohl, F. M. & Jovin, T. M. (1972). *J. Mol. Biol.* **67**, 375–396.
- Rich, A., Nordheim, A. & Wang, A. H.-J. (1984). *Annu. Rev. Biochem.* **53**, 791–846.
- Sadasivan, C. & Gautham, N. (1995). *J. Mol. Biol.* **248**, 918–930.
- Schneider, B. & Berman, H. M. (1995). *Biophys. J.* **69**, 2661–2669.

- Schneider, B., Cohen, D. M., Schleifer, L., Srinivasan, A. R., Olson, W. K. & Berman, H. M. (1993). *Biophys. J.* **65**, 2291–2303.
- Schneider, B., Ginell, S. L., Jones, R., Gaffney, B. & Berman, H. M. (1992). *Biochemistry*, **31**, 9622–9628.
- Schneider, B. & Kabeláč, M. (1998). *J. Am. Chem. Soc.* **120**, 161–165.
- Schneider, B., Neidle, S. & Berman, H. M. (1997). *Biopolymers*, **42**, 113–124.
- Schroth, G. P., Kagawa, T. F. & Ho, P. S. (1993). *Biochemistry*, **32**, 13381–13392.
- Schuerman, G. S., Smith, C. K., Turkenburg, J. P., Dettmar, A. N., van Meervelt, L. & Moore, M. H. (1996). *Acta Cryst. D* **52**, 299–314.
- Sheldrick, G. M. (1993). *SHELXL93. Program for Crystal Structure Refinement*. University of Gottingen, Germany.
- Sigel, A. & Sigel, H. (1996). Editors. *Interactions of Metal Ions with Nucleotides, Nucleic Acids and Their Constituents*, Vol. 32. New York: Marcel Dekker.
- Smith, C. K., Brannigan, J. A. & Moore, M. H. (1996). *J. Mol. Biol.* **263**, 237–258.
- Smith, C. K., Davies, G. J., Dodson, E. J. & Moore, M. H. (1995). *Biochemistry*, **34**, 415–425.
- Šponer, J., Burda, J. V., Mejzlik, P., Leszczynski, J. & Hobza, P. (1997). *J. Biomol. Struct. Dyn.* **14**, 613–628.
- Šponer, J., Leszczynski, J. & Hobza, P. (1996). *J. Biomol. Struct. Dyn.* **14**, 117–135.
- Suhai, S. (1995). *J. Chem. Phys.* **110**, 7030–7039.
- Takusagawa, F., Shaw, J. & Everett, G. W. (1988). *Inorg. Chem.* **27**, 3107–1988.
- Tari, L. W. & Secco, A. S. (1995). *Nucleic Acids Res.* **23**, 2065–2073.
- Thamann, T. J., Lord, R. C., Wang, A. H.-J. & Rich, A. (1981). *Nucleic Acids Res.* **9**, 5443–5457.
- Wahl, M. C., Rao, S. T. & Sundaralingam, M. (1996). *Biophys. J.* **70**, 2857–2866.
- Wang, A. H.-J., Gao, Y.-G., Liaw, Y.-C. & Li, Y.-K. (1991). *Biochemistry*, **30**, 3812–3815.
- Wang, A. H.-J., Gessner, R. V., van der Marel, G. A., van Boom, J. H. & Rich, A. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 3611–3615.
- Wang, A. H.-J., Hakoshima, T., van der Marel, G. A., van Boom, J. H. & Rich, A. (1984). *Cell*, **37**, 321–331.
- Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G. A. & Rich, A. (1979). *Nature (London)*, **282**, 680–686.
- Zhou, G. & Ho, P. S. (1990). *Biochemistry*, **29**, 7229–7236.