Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

B-DNA at atomic resolution reveals extended hydration patterns

Dominique Vlieghe,[a] Johan P. Turkenburg[b] and Luc Van Meervelt[a]*

[a]Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium, and [b]Department of Chemistry, University of York, Heslington, York YO1 5DD, England

Correspondence e-mail: luc.vanmeervelt@chem.kuleuven.ac.be

Received 28 January 1999
Accepted 9 June 1999

NDB Reference:
d(GGCCAATTGG), bd0006.

Despite the importance of hydration around DNA in the understanding of its conformation and interactions with other molecules in many biological processes, only limited atomic resolution information is available. Crystal-engineering techniques, which were originally developed to mimic DNA base triplets in a crystal lattice, also eliminate the rotational disorder of oligonucleotides around their helical axis and thereby enhance the resolution of the structure analysis. The low-temperature crystal structure of the synthetic DNA decamer d(GGCCAATTGG) has been determined at atomic resolution (1.15 Å) using 17700 reflections and the highly organized hydration patterns in both grooves have been characterized. The narrow d(AATT) minor groove is occupied by an 'extended hydration spine' alternately bridging base pairs and phosphate O1P atoms of opposite strands, while a distinctive pattern of parallel water ribbons is observed in the major groove. This analysis provides structural insight into the correlation found between narrow minor-groove width and occurrence of the B_I conformation and can be used to design new minor-groove binders. By their location between adjacent helices, two fully hydrated magnesium ions further stabilize the crystal packing. The structure also provides details of the hydration and conformation of G·GC triple helices.

1. Introduction

The solvent environment plays an important role in the conformation of biological macromolecules. In the case of DNA, a high relative humidity favours the B form, while reduced humidity or increased ionic strength promote conformational transitions to, depending on the sequence, A-, C-, D- and Z-DNA. The hydration around DNA double helices is often depicted as being organized in two regions (Saenger, 1984). The first region is impermeable to cations and in close contact with the polyelectrolyte molecule; the second region is bulk solvent. The two regions are separated by a transition zone with only partial ordering, which contains water molecules and ions fast exchanging with the bulk-solvent region.

DNA has many potential hydration sites: the functional groups of the bases facing the grooves and the phosphate and sugar O atoms. In total, about 20 water molecules per nucleotide comprise the primary hydration shell, of which, as indicated by infrared spectroscopy (Falk *et al.*, 1970), about ten waters are in direct contact with DNA. The number of experimental methods suitable for determining the detailed water structure around biological macromolecules is only small. For example, high-definition NMR studies yield infor-

mation on the residence time of interacting water molecules within about 4 Å of protein or nucleic acid atoms, but in most cases an exact localization of these solvent molecules is not possible (Kochoyan & Leroy, 1995). Information about the positions of water molecules can be obtained using X-ray crystallography. Depending on the resolution of the diffraction pattern and the water content of the crystals, most of the water molecules in the first region can be identified. The first B-DNA crystal structure of d(CGCGAATTCGCG), containing the *Eco*RI restriction site embedded in an alternating CG matrix, showed 72 water molecules per dodecamer duplex (Drew & Dickerson, 1981). This number represents only 17% of the 430 solvent molecules estimated from the crystal density. However, the first hydration region plays a key role in understanding the conformation of nucleic acids and the mechanisms by which they interact with other molecules. The remaining solvent molecules are too disordered to be identified in the electron-density maps. In the past, it was common practice to neglect the bulk-solvent contribution by discarding the very low order reflection data, but now most macromolecular refinement programs include appropriate corrections for the diffuse solvent.

Based on the oligonucleotide crystal structures present in the Nucleic Acid Database (Berman *et al.*, 1992), the distribution of water molecules around DNA has been quantified. Using Fourier transform techniques, average building blocks, which represent bases with first hydration shell waters at average positions, have been proposed and used to predict hydration sites for selected oligonucleotides (Schneider *et al.*, 1993). Recently, this study has been extended to the analysis of distributions of metal ions and water around the phosphate group in organic phosphates deposited in the Cambridge Structural Database (Allen & Kennard, 1993*a*,*b*). An asymmetric monodentate position, in which the cation or water only binds to one of the two negatively charged O atoms of the phosphate group, is preferred to a symmetric bidentate position between the two O atoms (Schneider & Kabelac, 1998).

The precise determination of the first hydration region by X-ray methods requires high-resolution diffraction data. In most cases (Berman *et al.*, 1992), the resolution of B-DNA oligonucleotide crystal structures varies between 1.4 and 2.5 Å. An important reason for atomic resolution not being observable for oligonucleotide structures is the rotational disorder of the helices in the crystal lattice. With decreasing distance between neighbouring helices, the rotational disorder reduces and will enhance the resolution that can be obtained. This was observed in the monoclinic structure of d(CCAACITTGG), where the phosphate backbones of one column of helices fit into the grooves of its neighbours, resulting in the highest resolution so far observed (1.3 Å) for longer B-DNA fragments (Lipanov *et al.*, 1993). We have recently shown (Van Meervelt *et al.*, 1995, Vlieghe *et al.*, 1996*a*,*b*) that crystal-engineering techniques used to determine the structure of triple-helical fragments result in a very tight crystal packing for the decamer d(GGCCAATTGG) (Fig. 1). The interaction of unpaired 5′-guanine bases G11 and G12 with adjacent duplexes generates infinitely extending columns through the formation of parallel $(G{\cdot}GC)_2$ triple-helix fragments. The relative position and orientation of these columns is fixed by the formation of antiparallel $(G{\cdot}GC)_2$ triple-helix fragments involving interaction of G1 and G2 in the major grooves of neighbouring columns. This triplex formation leads to a very stable crystal lattice in which an B-DNA AATT sequence is held in suspension (see Fig. 1 of Vlieghe *et al.*, 1996*b*), suggesting that the previously obtained 2.0 Å resolution might not be the limit. Here, we present the analysis of the same sequence and its solvent network in both grooves at an atomic resolution of 1.15 Å obtained using a stronger synchrotron source in combination with flash-cooling techniques. An 'extended hydration spine' in the minor groove with water-mediated inter-strand and intra-strand phosphate–phosphate contacts, and a distinctive pattern of parallel water ribbons in the major groove are described.

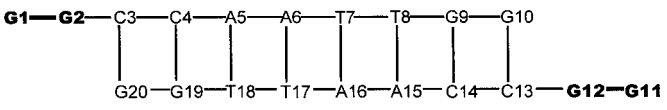

Figure 1
Sequence and residue numbering of the decamer d(GGCCAATTGG). Unpaired 5′-guanine bases are outlined and interact in the major groove of adjacent duplexes to form $(G{\cdot}GC)_2$ triple-helix fragments.

2. Methods

2.1. Crystallization

The purified oligonucleotide d(GGCCAATTGG) was purchased from the Oswel DNA Service (University of Southampton, England). The oligomer was crystallized at 289 K in space group $P2_12_12_1$ (unit-cell parameters a = 26.11, b = 36.46, c = 52.56 Å) with one molecule in the asymmetric unit as previously described (Vlieghe *et al.*, 1996*b*), using the following concentrations: [oligonucleotide] = 0.41 m*M*, [buffer] = 10.6 m*M* (pH = 6), [$MgCl_2$] = 42.6 m*M*, [MPD] = 4.3% and [spermine] = 0.22 m*M* against a 40% MPD reservoir (10 µl). Similar crystals with typical dimensions of 0.5 × 0.2 × 0.02 mm were obtained, but over a period of 7 d instead of 2 d.

2.2. Data collection and reduction

In contrast to the previous data collection (Vlieghe *et al.*, 1996*b*), intensity data were collected at a temperature of 120 K on the 180 mm MAR Research imaging-plate detector at beamline 5.2R of the synchrotron Elettra at Trieste (λ = 1.000 Å) using cryo-cooling techniques. The maximum resolution was limited to 1.15 Å owing to the detector dimensions and wavelength at the beamline. Two data-collection runs at different resolution limits were performed to correctly address the strongest reflections; for the first run at a resolution of 1.15 Å (crystal-to-detector distance 80 mm), the total oscillation range was 90° in steps of 1°; the second run of 36 images with an individual oscillation angle of 2.5° was achieved at 1.6 Å resolution (crystal-to-detector distance = 122 mm). The data set was processed with *DENZO*

Table 1
Data-collection statistics.

Resolution (Å)	1.15
Measured reflections	63114
Unique reflections	17760
Completeness to 1.15 Å (%)	96.7
Completeness in 1.18–1.15 Å shell (%)	95.0
R_{sym}† to 1.15 Å (%)	5.4
R_{sym}† in 1.18–1.15 Å shell (%)	15.8
Multiplicity to 1.15 Å	3.6
Multiplicity in 1.18–1.15 Å shell	3.2
Mean $I/\sigma(I)$	25.0
Reflections with $I > 3\sigma(I)$ (%)	92.0
Reflections with $I > 3\sigma(I)$ in 1.18–1.15 Å shell (%)‡	82.3

† $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections. ‡ Reflects that the real resolution limit has not been attained.

(Otwinowski, 1993) and scaled with the *CCP*4 (Collaborative Computational Project, Number 4, 1994) program *SCALA* (Evans, 1993). Data-collection statistics are given in Table 1.

2.3. Structure refinement

Refinement on all 17700 unique F^2 using *SHELXL*93 (Sheldrick, 1993) reduced the R factor to 0.167 for 16727 observed reflections [$F > 4\sigma(F)$] between 10 and 1.15 Å ($wR_2 = 0.423$) including 116 water molecules and two Mg^{2+} ions. Most of the bonds (1,2 distances) and angles (1,3 distances) were restrained to be approximately equal without applying fixed values (the SADI option in *SHELXL*). The criteria for identification of solvent atoms were that a well shaped peak in the $F_o - F_c$ map greater than 2σ in height was within 2.3–3.3 Å of plausible hydrogen-bonding partners and that well shaped peaks in the $2F_o - F_c$ maps, reasonable geometry and acceptable temperature factors (<45 Å^2) were obtained upon subsequent refinement. Bases were labelled G1 to G10 in the 5′ to 3′ direction on strand 1 and G11 to G20 on strand 2. A double conformation was refined for backbone atoms C2′, C3′ and O3′ for residue A16 and phosphate group T17 (population factor of 0.61 for first set). One water molecule was refined with this same population parameter, as it only has reasonable contacts to the first conformation. All atoms were refined anisotropically; however, restraints on the anisotropic displacement parameters were necessary: for all DNA atoms, the anisotropic displacement components of two atoms along the line joining them and of spatially adjacent atoms were restrained to be equal and 46 solvent and 21 DNA atoms were also restrained to be approximately isotropic. Anisotropic refinement of the double helix and the solvent molecules lowered the R factor by 0.033 and 0.018, respectively. The average equivalent temperature factor for each residue varies from 9.1 to 15.4 Å^2, which is close to the overall temperature factor of 9.0 Å^2 derived from the Wilson plot. H atoms were placed at calculated positions and refined in the riding mode with fixed isotropic temperature factors 1.2 times the U_{eq} of their parent atom. The parameters g (1.015) and U (2.748) were used to model diffuse solvent by Babinet's principle (Moews & Kretsinger, 1975). Using all low-resolution data has only a small influence on both parameters and does not leave an imprint on the final conformation.

The atom-positional estimated standard deviations obtained from a final cycle of full-matrix least-squares refinement (Sheldrick, 1996) range from 0.02–0.18 Å and were used to obtain estimated standard deviations on all geometric parameters. The following averages were observed: 0.04 Å for bond distances (range 0.02–0.14 Å), 2.3° for bond angles (range 1.0–8.0°) and 3.3° for torsion angles (range 1.6–21.0°). For the torsion angles, the lowest e.s.d. values are found for the bases and the highest values for the phosphates. A clear correlation is found between the e.s.d. value and the quality of the electron-density map. Double conformations also show higher e.s.d. values.

2.4. Structure analysis

The program *NEWHELIX*93 was used for calculation of the helical parameters and the program *Curves* (Lavery & Sklenar, 1989) was used for the calculation of groove dimensions and sugar puckering. All calculations have been performed with the backbone conformation with the highest occupancy. Figures were produced using *MOLSCRIPT* (Kraulis, 1991), *Raster3D* (Merritt & Murphy, 1994) and *BOBSCRIPT* (Esnouf, 1997).

3. Results

3.1. Overall features of the duplex

The eight base-pair fragment d(CCAATTGG) adopts a classical B-DNA double helix with a helical twist of 35.2°, a rise of 3.42 Å, 10.2 residues per turn and a diameter of about 19.6 Å. The superposition of the 1.15 and the 2.0 Å structures shows a root-mean-square deviation of 0.49 Å, with closest similarity for the bases (r.m.s. deviation ranging from 0.16 to 0.50 Å), followed by the sugar rings (0.20–0.71 Å) and least similarity for the phosphate groups (0.27 to −2.02 Å), with the phosphate groups of residues C13 and A16 deviating most (2.02 and 1.60 Å, respectively); this is the same order as usually found for the average temperature factors ($B_{bases} < B_{sugars} < B_{phosphates}$). The atomic resolution has made it possible to interpret peaks in the difference electron-density maps close to atoms C2′, C3′ and O3′ of residue A16 and to the phosphate group T17 as a double backbone conformation (with a population parameter of 0.39 for the second conformation).

Table 2 gives the sugar–phosphate backbone angles, glycosyl torsion angles and pseudorotation phase angles. The flexibility of the sugar conformation in B-DNA is illustrated by the pseudorotation phase angle ranging from 96.3 (6)° (O4′-*endo*) to 179.8 (3)° (C2′-*endo*). However, residues C3 and C13 show a C3′-*endo* sugar-puckering mode, a feature which may be related to the triplex formation. The torsion angles α, β, γ and δ are in their usual (−)-*gauche*, *trans*, (+)-*gauche* and (+)-*anticlinal* ranges, respectively. Exceptions are found in the backbone of the triple-helical fragments for the α angle of C13 [(+)-*synclinal*], the γ angle of G1, C3 and

Table 2
Sugar–phosphate backbone, glycosyl torsion angles and pseudorotation phase angles (°) for d(GGCCAATTCC).

Main-chain torsion angles are defined by P–α–O5′–β–C5′–γ–C4′–δ–C3′–ε–O3′–ζ–P. The glycosyl torsion angle χ is defined by O4′—C1′—N1—C2 for pyrimidines and O4′—C1′—N9—C4 for purines.

Residue	α	β	γ	δ	ε	ζ	χ	Phase
G1			186	104	−151	−97	−128	96
G2	−70	169	50	146	−96	−59	−79	153
C3	−59	138	187	90	−146	−73	−137	50
C4	−63	170	54	149	−106	−187	−82	155
A5	−70	152	43	140	−185	−90	−107	157
A6	−64	179	47	127	−170	−76	−112	132
T7	−72	175	65	114	−175	−82	−129	120
T8	−70	175	53	121	−185	−91	−111	140
G9	−68	187	46	124	−179	−92	−114	135
G10	−68	172	53	107			−124	111
G11			45	143	−173	−106	−106	171
G12	−64	184	46	145	−94	−177	−98	153
C13	64	214	186	85	−159	−69	−147	11
C14	−63	181	58	147	−188	−104	−111	164
A15	−57	196	49	147	−187	−91	−111	180
A16*a*†	−69	186	45	119	−193	−105	−114	120
A16*b*†			64	114	−174	−73		153
T17*a*†	−37	160	50	116	−190	−95	−122	121
T17*b*†	−70	181						
T18	−54	183	55	136	−189	−101	−108	160
G19	−61	192	49	139	−179	−92	−104	177
G20	−72	182	48	137			−99	147

† Conformations *a* and *b* for residues A16 and T17 have population parameters of 0.61 and 0.39, respectively.

C13 (*trans*) and the δ angle of C3, C13 [(+)-*synclinal*] and are related to the triplex formation. The twisting away of residues G1 and G2 is reflected by changes in torsion angles β and δ, but mainly in γ of residue G3. The unusual torsion angles for C13 were not previously observed in the lower resolution structure (Vlieghe *et al.*, 1996*b*). However, these torsion angles are compatible with the electron-density maps at 2.0 Å resolution. This illustrates that electron-density maps at 2.0 Å resolution still can contain some ambiguity with respect to backbone conformations. The glycosidic conformations are all (−)-*anticlinal*, with an average torsion angle $\chi = -112°$. Only two phosphate groups (P5 and P13) adopt the B_{II} conformation, which is in agreement with the narrow minor groove in the B_{II} free central AATT region (Dickerson, 1990). The decamer d(GGCCAATTGG) (subsequently referred to as 10-mer) contains the same central AATT tetramer as the dodecamer d(CGCGAATTCGCG) (subsequently referred to as 12-mer) studied in detail by Dickerson and coworkers (Drew & Dickerson, 1981). Minor-groove widths for 12-mer and 10-mer sequences are very similar, indicating that the triplex formation does not change the minor-groove characteristics (Fig. 2).

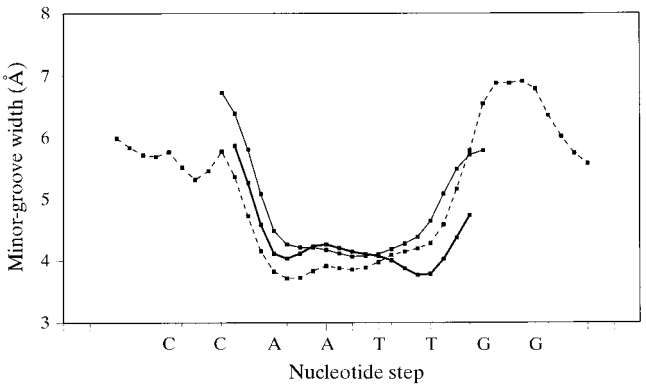

Figure 2
Plot of minor-groove widths based on C1′ atoms for the 1.15 Å (full line, bold) and 2.0 Å (full line, light) structures of d(GGCCAATTGG) (10-mer) and d(CGCGAATTCGCG) (12-mer, dashed line).

3.2. Hydration of the grooves

The most striking difference from the 2.0 Å resolution structure is seen in the hydration of the B-DNA fragment. In total, 116 water molecules and 2 Mg^{2+} ions were located (compared with 44 water molecules for the 2.0 Å resolution structure), of which 85 are in the first coordination shell. The two $Mg(H_2O)_6^{2+}$ clusters bridge three adjacent symmetry-related backbone strands, further stabilizing the crystal lattice and causing temperature factors to be lower for this region. Each phosphate group is individually hydrated by an average of 3.2 molecules. On average, the adenine and thymine bases not involved in triplet formation have 2.25 and 2.00 water molecules per base within 3.20 Å, respectively, which is significantly more than the 1.47 and 0.91, respectively, found in a systematic analysis of 28 crystal structures of oligonucleotides (Schneider *et al.*, 1993).

The major-groove region contains 2–3 water molecules directly hydrogen bonded to the bases per base pair. These solvent molecules lie in the base-pair plane, linking the polar atoms of adjacent bases (Fig. 3). In the backbone strand with the lowest temperature factors, the O2P atoms are heavily hydrated, forming intrastrand phosphate bridges which are further connected to the base-associated hydration.

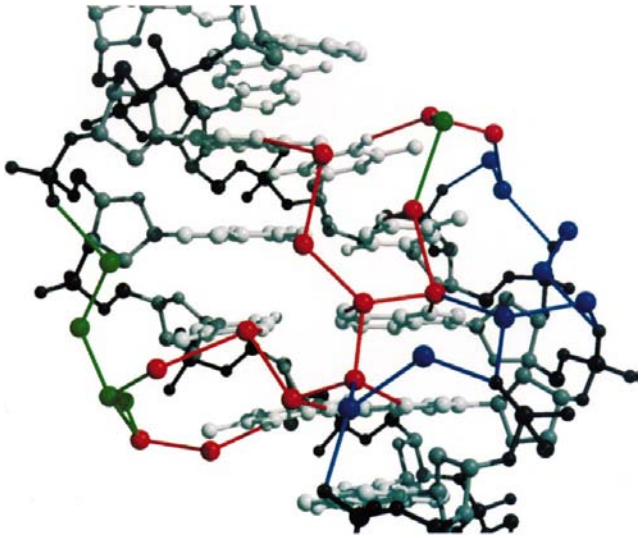

Figure 3
Hydration pattern of the AATT major groove as observed in the 1.15 Å crystal structure of d(GGCCAATTGG). Water molecules forming the intrastrand phosphate bridges are coloured blue, those directly hydrogen bonded to base atoms or in close contact with the thymine methyl group are coloured red and the others are coloured green.

Table 3
Hydrogen-bonding distances (Å) in antiparallel and parallel G·GC triplets.

Subscripts WC refer to Watson–Crick bases, H to third-strand guanines.

Antiparallel triplets

	G1·G10C13	G2·G9C14
$N4(C_{WC})\cdots O6(G_{WC})$	2.87 (2)	2.84 (2)
$N3(C_{WC})\cdots N1(G_{WC})$	2.95 (2)	2.91 (2)
$O2(C_{WC})\cdots N2(G_{WC})$	2.96 (2)	2.85 (2)
$N7(G_{WC})\cdots N1(G_{H})$	3.03 (2)	2.85 (2)
$O6(G_{WC})\cdots N2(G_{H})$	2.82 (2)	2.83 (2)

Parallel triplets

	G12·G20C3	G11·G19C4
$N4(C_{WC})\cdots O6(G_{WC})$	2.85 (2)	3.02 (2)
$N3(C_{WC})\cdots N1(G_{WC})$	2.91 (2)	2.95 (2)
$O2(C_{WC})\cdots N2(G_{WC})$	2.86 (2)	2.80 (2)
$N7(G_{WC})\cdots N2(G_{H})$	2.91 (2)	2.96 (2)
$O6(G_{WC})\cdots N1(G_{H})$	2.79 (2)	2.83 (2)
$N4(C_{WC})\cdots O6(G_{H})$	3.13 (2)	3.11 (2)

As in the 12-mer, the narrow central region of the minor groove is occupied at the bottom by a zigzag hydration spine (Fig. 4; water molecules coloured in red). In the current structure, this spine further extends towards the flanking phosphate groups in the minor-groove walls. Depending on the minor-groove width, two or three water molecules bridge the phosphate O1P atoms of opposite strands. The nine water molecules of the base-associated hydration spine alternatingly bridge base pairs and phosphate groups, and are located closer to strand 1 than to strand 2. The average temperature factor for these water molecules is 15.5 $Å^2$, which is lower than the average for all water molecules (21.2 $Å^2$). The temperature factors for the water molecules bridging the phosphate groups average at 19.0 $Å^2$. At one end, no water-mediated contact between P7 and P20 is observed despite the short interstrand phosphate–phosphate distance. Instead, this region of the minor groove displays two water ribbons which do not bridge base atoms to sugar O4′ atoms, as described for wide B-DNA minor grooves (Privé *et al.*, 1987), but form intrastrand phosphate bridges. As this part of the helix is involved in the triple-helix formation, it is difficult to give an exact explanation for this double ribbon. At the end of the spine, the hydration network breaks up into an irregular pattern.

3.3. Triple-helix fragments

The best-defined regions in the electron-density maps are found for both triple-helix parts (Fig. 5). The lowest temperature factors are also observed in these regions, illustrating the importance of triplet formation in the stabilization of the crystal lattice. Compared with the lower resolution structure determination, the geometry of the triplexes has not changed much (r.m.s. deviations of 0.33 Å and 0.24 Å for the antiparallel and parallel triplexes, respectively). Atomic resolution enables the calculation of the e.s.d. values on the non-restrained triplet hydrogen bonds (Table 3). Despite the great similarity in hydrogen-bonding distances between the third-strand guanines (denoted G_H) and the Watson–Crick bases (denoted as G_{WC} and C_{WC}), the ranges for the structure reported here are always smaller. Two other differences are noteworthy. First, the already weak $CH\cdots O$ hydrogen bond between $C8(G_{WC})$ and $O6(G_H)$ observed in one of the antiparallel triplets at 2.0 Å resolution (3.33 Å; Vlieghe *et al.*, 1996*b*) is even further weakened [3.45 (3) Å]. Secondly, the parallel triplets are further stabilized by inter-triplet hydrogen

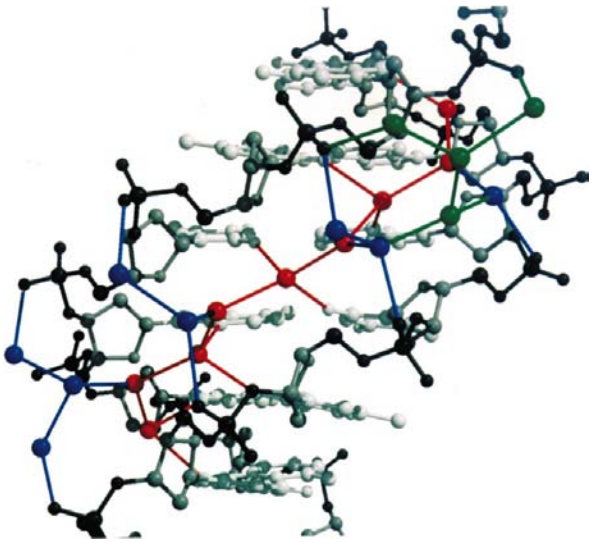

Figure 4
Hydration pattern of the AATT minor groove (colour scheme as in Fig. 3). Red-coloured water molecules form part of the original hydration spine as described by Drew & Dickerson (1981). Angles in the 'extended hydration spine' average at 109.7° (ranging from 74 to 143°) showing a tetrahedral solvent network.

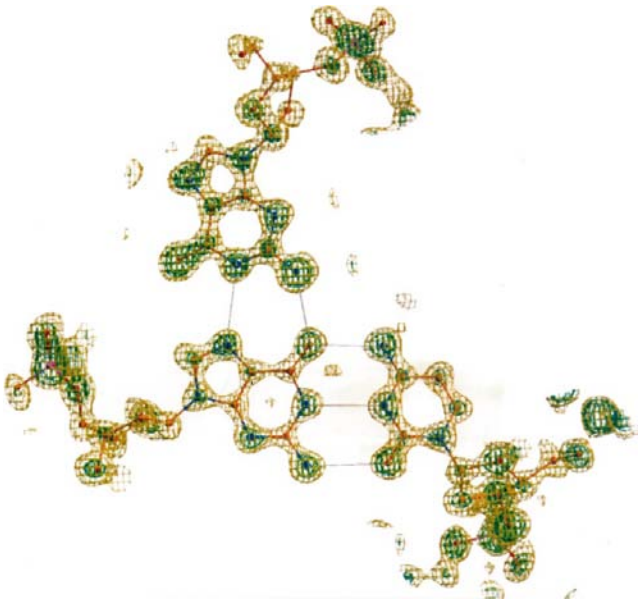

Figure 5
The antiparallel reverse-Hoogsteen triplet G1·G10C13 superimposed on a section of the $(2F_o - F_c)$ electron-density map [contouring levels 2σ (green) and 4σ (yellow)].

bonding (Fig. 6). Although observed, but not emphasized, in the 2.0 Å structure, the inter-triplet hydrogen-bond distance N4(C_{WC})···O6(G_H) of 2.89 (2) Å is shorter than the analogous intra-triplet distances [3.11 (2) and 3.13 (2) Å], in agreement with proposed models for $(G{\cdot}GC)_n$ triplexes generated by molecular modelling (Van Vlijmen *et al.*, 1990). This inter-triplet hydrogen bonding is possible because of the non-planarity of the triplets, with angles between the base planes ranging from 3.9 to 33.5°. For both types of triplex, the largest inclinations are found between G_H and C_{WC} base planes. The third strand divides the major groove into two parts, usually called the Watson–Hoogsteen and Crick–Hoogsteen grooves (Radhakrishnan & Patel, 1993). Comparison of the first and second hydration shells of the different grooves in both triplexes (Fig. 7) shows that most water molecules lie in the Watson–Hoogsteen groove (nine for the parallel and ten for the antiparallel triplex). The Crick–Hoogsteen groove in the antiparallel triplex contains only five water molecules, compared with ten in the parallel triplex. This can be explained by the occurrence of a $Mg(H_2O)_6^{2+}$ cluster close to this groove of the parallel triplex and the smaller dimensions and hydrophobic C8–H8 group of the antiparallel triplex. The Watson–Crick or original minor groove of the double helix is more hydrated in the antiparallel triplex (13 water molecules *versus* six for the parallel triplex) owing to the close proximity of a $Mg(H_2O)_6^{2+}$ cluster. In both triplexes, one water molecule interconnecting the Watson and Hoogsteen strand is found at a comparable position.

4. Discussion

Solvent molecules play an important role in understanding the detailed conformation of DNA and its interactions with other molecules in many biological processes. Structural studies on short DNA fragments using X-ray diffraction can be used to identify most solvent molecules in the first hydration region of the helix. However, the details these studies provide of the hydration pattern depend highly on the resolution of the diffraction, which is so far limited to 1.3 Å for longer B-DNA fragments.

For A-DNA fragments, structured water is seen in the deep major groove as a ribbon of pentagons (Kennard *et al.*, 1986), and/or as phosphate bridges (Wang *et al.*, 1982). In contrast, B-DNA occurring at high water activity contains distinctive hydration patterns in the minor groove, the precise structure being related to the groove width (Privé *et al.*, 1991). Wide minor grooves contain two water ribbons; narrow minor grooves have a hydration spine, to which we should now add water-mediated inter-strand and intra-strand phosphate–phosphate contacts. In the major groove, a distinctive pattern of water molecules also plays a part in stabilizing the structure.

The crystallographic analysis of d(GGCCAATTGG) at the atomic resolution of 1.15 Å reported here shows previously unobserved details of the B-DNA conformation and hydration. For example, the conformational flexibility of the sugar–phosphate backbone, otherwise only represented by thermal parameters in crystal structures, is also visible in this high-resolution structure as a local double conformation in the neighbourhood of phosphate group T17. In the d(AATT) major groove, a base- and phosphate-associated solvent network is identified in great detail. This hydration pattern of the bases follows the pseudo-twofold symmetry of the sequence, but three waters necessary to complete the pattern are missing. Positive density is observed in the final difference electron-density map at their expected positions; however, owing to our strict criteria for solvent selection, these waters were not included in the final model. In trial refinements, either their temperature factor was unrealistically high (>70 $Å^2$) or the peak quality was poor. Using the pseudo-twofold symmetry, we have modelled an ideal pattern for the first hydration shell of the AATT major-groove region in B-DNA (Fig. 8) of four interconnected parallel water ribbons, comprising six- and seven-membered rings.

The narrow minor groove of the central d(AATT) sequence is occupied by an 'extended hydration spine' consisting of a base-associated hydration spine alternatingly bridging base pairs and extending towards the flanking phosphate groups in the minor-groove walls. The 'extended hydration spine' provides insight into the well known correlation found between narrow minor-groove width and the occurrence of the B_I phosphate conformation (Dickerson, 1990). This 'extended hydration spine' requires the B_I conformation for the flanking phosphate groups, with the O1P atoms facing into the minor groove. In contrast to the major groove, the hydration pattern in the minor groove does not reflect the pseudo-twofold symmetry of the sequence. The highly organized hydration pattern in both grooves of the d(AATT) sequence can play an important role in the transmission of information into the bulk-solvent region.

Anti-tumoural minor-groove binders have a crescent shape which allows the molecules to fit the contour surface of the B-DNA minor groove, displacing the well known hydration spine (Wang & Teng, 1990). Several hydrogen bonds between the drug and the O and N atoms at the base of the minor groove, together with close van der Waals contacts, are the stabilizing forces of this DNA–drug complex formation. The 'extended hydration spine' described here should be taken into account in the future design of new generations of minor-groove binders, which not only interact with the DNA bases but also with the phosphates. The approach of using crystal-

Figure 6
View into the Watson–Hoogsteen groove (see Radhakrishnan & Patel, 1993 for definition) of the parallel $(G{\cdot}GC)_2$ triplet showing the inter-triplet N4(C_{WC})···O6(G_H) hydrogen bond (2.89 Å, red stick). The G12·G20C3 triplet is coloured blue and the G11·G19C4 triplet is coloured green; the N4 and O6 atoms involved in intra- and inter-triplet hydrogen bonding are shown as spheres.

engineering techniques in order to enhance the final resolution of a crystal structure can also be applied to DNA–minor-groove complexes, which in general do not diffract further than about 2.2 Å. We have indeed recently obtained a resolution enhancement of 0.5 Å for the d(GGCCAATTGG)–DAPI complex (D. Vlieghe and L. Van Meervelt, unpublished work).

The idea that the first hydration shell around DNA is impermeable to cations was recently questioned by a low-temperature determination of the 12-mer sequence at 1.4 Å resolution (Shui *et al.*, 1998) suggesting that – based on the calculation of the sodium specific bond valence (ν_{Na^+}) – the base-associated spine of hydration is partially occupied by Na atoms. As it is not possible to differentiate between sodium and water during the refinement and as the presence of the Mg^{2+} ions alone cannot compensate for the total negative charge of the oligonucleotide, some waters indeed may be cations. For the base-associated spine water molecules we calculate similar ν_{Na^+} values varying from 0.45 to 0.48 [calculated with the program *WASP* (Nayal & Di Cera, 1996), using a contact distance of 4 Å]. Several water molecules not situated in the minor groove have even higher ν_{Na^+} values (around 0.50). For all water molecules, ν_{Na^+} also has a Gaussian distribution with average 0.26 and standard deviation 0.13. The ν_{Na^+} distribution and maximum value for 332242 water molecules in the PDB (Nayal & Di Cera, 1996) are different from those observed for both oligonucleotides: ν_{Na^+} averages at 0.18 with a standard deviation of 0.13, and 64 water molecules have $\nu_{Na^+} \geq 1.0$ and a distorted octahedral coordination geometry. In this structure, however, the coordination of the base-associated spine water molecules is tetrahedral (see Fig. 4), from which we can conclude that the

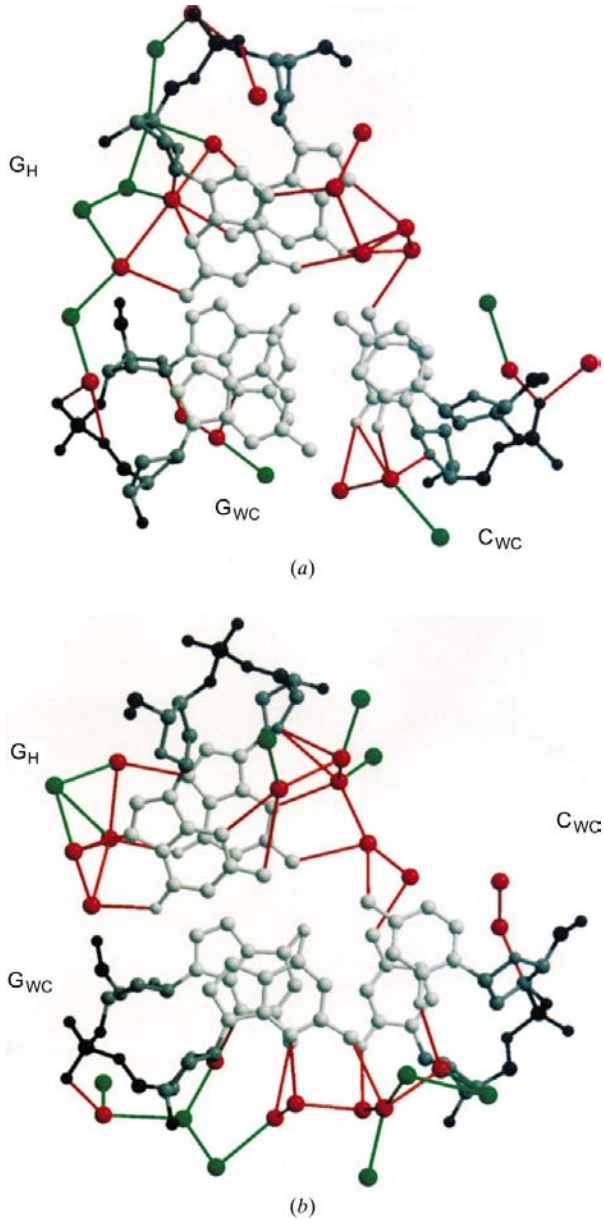

Figure 7
Hydration of the different grooves in (*a*) parallel and (*b*) antiparallel G·GC triplexes. Letters refer to the Watson, Crick and Hoogsteen strands. The same colour scheme is used as in Fig. 3.

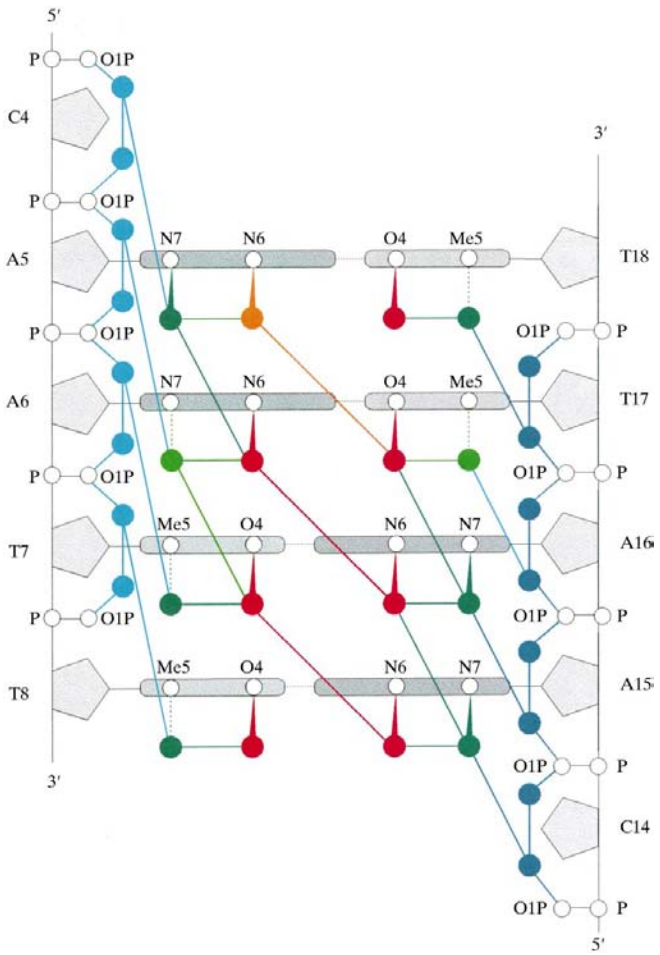

Figure 8
Schematic representation of the ideal model for the first hydration shell in the major groove of a B-DNA AATT region (colour scheme as in Fig. 3, waters in light colours were not observed in the crystal structure). The model was generated by using the observed pseudo-twofold symmetry of the solvent network, using only the red- and blue-coloured waters in Fig. 3. Four water ribbons following the sugar backbone direction are shown in thicker red and blue lines, connected in a network of six- and seven-membered rings.

spine consists of fully occupied water molecules. This is strengthened by the fact that – owing to the tetrahedral arrangement – positioning of donor–acceptor atoms is possible in the minor-groove hydrogen-bonding pattern.

Nucleic acid triplexes are implicated in genetic recombination *in vivo* and have possible applications in genome analysis and antigene therapy. As previously shown (Vlieghe *et al.*, 1996*b*), the unpaired 5′-guanine bases in the crystal structure of d(GGCCAATTGG) interact with available hydrogen-bonding sites in the major groove of adjacent duplexes. As a consequence, valuable high-resolution information on the hydration, hydrogen-bonding scheme, groove dimensions and planarity of parallel and antiparallel $(G{\cdot}GC)_2$ triple helices is obtained[1].

This work was partly supported by the Human Capital and Mobility Programme of the European Community, by the 'Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie (IWT)' and by the Fund for Scientific Research (Flanders). JPT was supported jointly by Glaxo Wellcome and Leiden University (and subsequently by the University of York), and LVM is a Research Director of the Fund for Scientific Research (Flanders). We thank the staff of the Elettra X-ray diffraction beamline for assistance and H. Reynaers for continued support.

References

Allen, F. H. & Kennard, O. (1993*a*). *Chem. Design Autom. News*, **8**, 1.
Allen, F. H. & Kennard, O. (1993*b*). *Chem. Design Autom. News*, **8**, 31–37.
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.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dickerson, R. E. (1990). *Structure and Methods*, edited by R. H. Sarma & M. H. Sarma, Vol. 3, pp. 1–38. New York: Adenine Press.
Drew, H. R. & Dickerson, R. E. (1981). *J. Mol. Biol.* **151**, 535–556.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 133–138.

[1] Supplementary material is available from the IUCr electronic archive (Reference: li0326). Services for accessing these data are described at the back of this issue.

Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Falk, M., Poole, A. G. & Goymour, C. G. (1970). *Can. J. Chem.* **48**, 1536–1542.
Kennard, O., Cruse, W. B. T., Nachman, J., Prange, T., Shakked, Z. & Rabinovich, D. (1986). *J. Biomol. Struct. Dyn.* **3**, 623–647.
Kochoyan, M. & Leroy, J. L. (1995). *Curr. Opin. Struct. Biol.* **5**, 329–333.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Lavery, R. & Sklenar, H. (1989). *J. Biomol. Struct. Dyn.* **6**, 655–667.
Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J. & Dickerson, R. E. (1993). *Biochemistry*, **32**, 1373–1389.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–228.
Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Privé, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L. & Dickerson, R. E. (1987). *Science*, **238**, 498–504.
Privé, G. G., Yanagi, K. & Dickerson, R. E. (1991). *J. Mol. Biol.* **217**, 177–199.
Radhakrishnan, I. & Patel, D. J. (1993). *Structure*, **1**, 135–152.
Saenger, W. (1984). *Principles of Nucleic Acid Structure.* New York: Springer–Verlag.
Schneider, B., Cohen, D. M., Schleifer, L., Srinivasan, A. R., Olson, W. K. & Berman, H. M. (1993). *Biophys. J.* **65**, 2291–2303.
Schneider, B. & Kabelac, M. (1998). *J. Am. Chem. Soc.* **120**, 161–165.
Sheldrick, G. M. (1993). *SHELXL*93. *Program for Crystal Structure Refinement.* University of Göttingen, Germany.
Sheldrick, G. M. (1996). *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey, pp. 47–55. Warrington: Daresbury Laboratory.
Shui, X., McFail-Isom, L., Hu, G. G. & Williams, L. D. (1998). *Biochemistry*, **37**, 8341–8355.
Van Meervelt, L., Vlieghe, D., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1995). *Nature (London)*, **374**, 742–744.
Van Vlijmen, H. W. Th., Ramé, G. L. & Pettitt, B. M. (1990). *Biopolymers*, **30**, 517–532.
Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1996*a*). *Acta Cryst.* D**52**, 766–775.
Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Précigoux, G. & Kennard, O. (1996*b*). *Science*, **273**, 1702–1705.
Wang, A. H.-J., Fujii, S., van Boom, J. H. & Rich, A. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 3968–3972.
Wang, A. H.-J. & Teng, M.-K. (1990). *Crystallographic and Modeling Methods in Molecular Design*, edited by C. E. Bugg & S. E. Ealick, pp. 123–150. New York: Springer–Verlag.