Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Near-atomic resolution crystal structure of an A-DNA decamer d(CCCGATCGGG): cobalt hexammine interaction with A-DNA

Boopathy Ramakrishnan, Chandra Sekharudu, Baocheng Pan and Muttaiya Sundaralingam*

Departments of Chemistry and Biochemistry, The Ohio State University 200 Johnston Laboratory, 176 West 19th Avenue, Columbus, Ohio 43210-1002, USA

Correspondence e-mail: sundaral@chemistry.ohio-state.edu

Received 29 July 2002
Accepted 15 October 2002

PDB and NDB References: d(CCCGATCGGG), 1m77, r1m77sf; d(CCCGATCGGG), ad0026.

The structure of the DNA decamer d(CCCGATCGGG) has been determined at 1.25 Å resolution. The decamer crystallized in the tetragonal space group $P4_32_12$, with unit-cell parameters $a = b = 44.3$, $c = 24.8$ Å and one strand in the asymmetric unit. The structure was solved by the molecular-replacement method and refined to $R_{\rm work}$ and $R_{\rm free}$ values of 16.3 and 18.5%, respectively, for 5969 reflections. The decamer forms the A-form DNA duplex, with the abutting crystal packing typical of A-DNA. The crystal packing interactions seem to distort the local conformation: A5 adopts the *trans*/*trans* conformation for the torsion angles α and γ instead of the usual *gauche*$^-$/*gauche*$^+$ conformations, yielding G*(G·C) base triplets. The highly hydrated $[Co(NH_3)_6]^{3+}$ ion adopts a novel binding mode to the DNA duplex, binding directly to phosphate groups and connecting to N7 and O6 atoms of guanines by water bridges. Analysis of thermal parameters (*B* factors) shows that the nucleotides involved in abutting crystal packing are thermally more stable than other nucleotides in the duplex.

1. Introduction

DNA molecules are polymorphous and can adopt A-form, B-form and Z-form helical structures (Saenger, 1984). The helical form adopted by a DNA molecule is largely related to its sequence and environment. B-form DNA was first observed to transform to the A-form upon dehydration in fibers (Franklin & Gosling, 1953). Crystallographic studies show that the decamers d(CCGGCGCCGG) (Heinemann *et al.*, 1992), d(CCAGGCCTGG) (Heinemann & Alings, 1989), d(CCAGCGCTGG) (Chiu & Dickerson, 2000) and d(CCGCTAGCGG) (Eichman *et al.*, 2000) adopt the B-form, while the decamers d(ACCGGCCGGT) (Frederick *et al.*, 1989) and d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) adopt the A-form. These results indicate that A·T base pairs inside decamers may favour formation of the B-form helix. We therefore substituted the central GC bases in d(CCCGGCCGGG) with AT bases and studied the influence of the A·T base pairs on the DNA conformation.

The stability of nucleic acid structures is crucial in their interaction with proteins in biological processes and depends on the hydrogen-bonding interactions, base stacking, hydration and interaction with metal ions. In order to study the details of hydration, it is important to study nucleic acid structures at very high resolution. The high-intensity X-ray beams of third-generation synchrotrons and new techniques in the synthesis and purification of nucleic acids have allowed much progress to be made in improving the resolution of crystal structures in recent years. For instance, the resolution of the Dickerson–Drew B-DNA dodecamer d(CGCGAA-

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

^{F}TTCGCG) (where ^{F}T represents 2′-deoxy-2′-fluoroarabino-thymine) has been improved to 0.95 Å and the A-DNA decamer d(GCCGT^{M}TACGC) (where ^{M}T represents 2′-methoxy-3′methylene-phosphonatethymine) has been refined to 0.83 Å (Egli *et al.*, 1998). Also, the structure of the B-DNA decamer d(CCAGTACTGG) has been determined at 0.74 Å (Kielkopf *et al.*, 2000). These structures display a large amount of accurate detail of the DNA fine structure and hydration. The highest resolution achieved so far is 0.61 Å for the crystal structure of a parallel-stranded RNA tetraplex r(UGGGGU), which clearly shows the electron density of the H atoms in the nucleic acid (Deng *et al.*, 2001). High-resolution crystal structures can also provide information on the flexibility of nucleotides in terms of their thermal vibrations. In the case of low-resolution structures, the errors arising from structure determination may be expressed in their thermal vibrations; therefore, it is not accurate or reliable to correlate the conformational flexibility of macromolecules with their thermal vibrations as expressed by B factors. Here, we report the crystal structure of an A-DNA decamer d(CCCGATCGGG) at 1.25 Å resolution and discuss the relationship between the thermal parameters (B factors) and the stability of the nucleotides and the cobalt hexammine ion binding to the A-DNA.

2. Materials and methods

2.1. Synthesis, crystallization and data collection

The DNA decamer d(CCCGATCGGG) was synthesized by the phosphoramidite method using an in-house Applied Biosystem DNA synthesizer 381 (Foster City, CA, USA). The deprotected decamer was precipitated by ethanol in the presence of 2.5 M ammonium acetate at 248 K. The lyophilized precipitate was used for crystallization without further purification. Crystals were grown by the hanging-drop vapor-diffusion method in the presence of 1 mM DNA (single-stranded concentration), 40 mM sodium cacodylate buffer pH 6.0 and 0.5 mM cobalt hexammine chloride against 50%(v/v) 2-methyl-2,4-pentanediol (MPD). A crystal of dimensions 0.25 × 0.25 × 0.5 mm was mounted in a glass capillary with some mother liquor at one end for X-ray data collection. Three-dimensional X-ray intensity data to 1.2 Å were collected at room temperature using an in-house Siemens–Nicolet area detector mounted on a four-circle goniometer, with a MaxScience rotating-anode source operating at 50 kV and 100 mA. The crystal-to-detector distance was 12.0 cm. Two φ scans corresponding to low and high resolution and 10 ω scans were performed. The low-resolution data were collected in 0.25° steps, while the high-resolution scans used 0.20° scan steps (90 s per frame). The frames of high- and low-resolution data were processed together using *XENGEN* 2.0 (Howard, 1993). A total of 51 067 reflections were collected, of which 6929 were unique (90% of the possible reflections) with an R_{sym} of 5.5%. The crystal belonged to the tetragonal space group $P4_32_12$, with unit-cell parameters $a = b = 44.3$, $c = 24.8$ Å and $\alpha = \beta = \gamma = 90°$. There is one single strand of decamer in the asymmetric unit, with a volume per base pair of 1250 Å^3.

2.2. Structure solution and refinement

There is one DNA single strand in the asymmetric unit. The dyad of the decamer DNA duplex is expected to coincide with the crystallographic twofold axis present in the crystal lattice. Therefore, a two-parameter search, rotation and translation about the twofold axis, was carried out using a program developed in our laboratory (Jain & Rao, private communication). 100 strong reflections between 8 and 3 Å resolution were used in this search, with right-handed A- and B-DNA fiber decamer search models in both space groups $P4_12_12$ and $P4_32_12$. The search with the A-DNA in space group $P4_32_12$ yielded the solution with the highest correlation coefficient (52%). Further least-squares refinement using 6–3 Å resolution data, followed by omit electron-density maps, confirmed the solution. Once the structure was solved, the model was refined using 6–2 Å resolution data, followed by simulated annealing in order to remove any conformational bias of the starting model by initially heating the system to 673 K and then slowly cooling to 573 K with 0.5 fs sampling intervals. R_{work} and R_{free} were 20.1 and 22.7%, respectively. At this stage, a cobalt hexammine ion was located in the structure from difference Fourier maps and was included in further refinement. Next, the resolution of the data was extended to

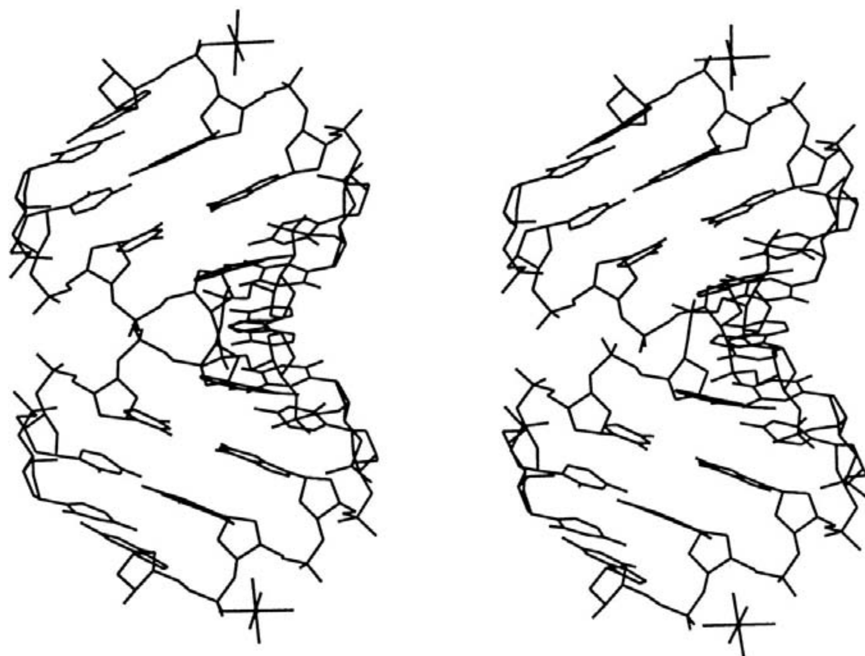


Figure 1
Stereoview of the decamer d(CCCGATCGGG) with each strand bound to a cobalt hexammine ion.

1.8 and then to 1.4 Å for further refinement, with the addition of water molecules located in the difference electron-density maps ($F_o - F_c$). At this stage, the nucleotides were fitted to the $3F_o - 2F_c$ electron-density maps by omitting one nucleotide at a time. Final refinement was carried out using 5969 reflections [$F \geq 2\sigma(F)$] between 8.0 and 1.25 Å resolution to an $R_{\rm work}$ and $R_{\rm free}$ of 16.3 and 18.5%, respectively. 33 water molecules and one cobalt hexammine ion were located as solvent. All refinements were performed using *X-PLOR* (Brünger, 1994). The root-mean-square deviation from ideality is 0.012 Å for bond lengths and 3.4° for bond angles. The *B*-factor root-mean-square deviation is 1.29 Å^2 for the bonded main-chain atoms and 1.46 Å^2 for bonded side-chain atoms. The atomic coordinates and structure factors have been deposited in the Nucleic Acid Database (Berman *et al.*, 1992).

3. Results and discussion

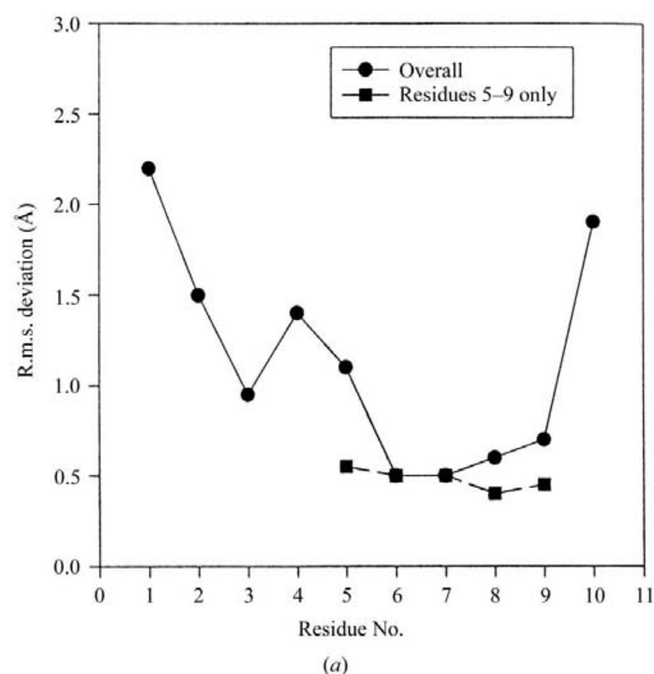
3.1. Overall conformation and helix geometry

The decamer forms an antiparallel A-DNA double helix with its symmetry-related molecule, whose dyad coincides with the crystallographic twofold axis (Fig. 1). All the base pairs are of the Watson–Crick type. The DNA duplex has typical features of an A-DNA: a narrow major groove with an average width of 3.0 Å and a wide and flat minor groove with an average width of 10.5 Å. The average helical twist, rise, base-pair inclination and slide are 33.0°, 2.3 Å, 17° and 1.4 Å, respectively. All the sugar puckers belong to the C3′-*endo* conformation. The sugar–phosphate backbone torsion angles are remarkably similar to the starting fiber A-DNA model, except for the torsion angles α (P—O5′) and γ (C4′—C5′) of A5, which adopt *trans*/*trans* conformations instead of the usual *gauche*$^-$/*gauche*$^+$ conformations.

The overall r.m.s. deviation is 1.34 Å upon superposition with the fiber A-DNA model. The terminal nucleotides have a much higher r.m.s. deviation than the inner nucleotides (Fig. 2*a*). When only the inner structure containing nucleotides 5–9 is used for superposition, the conformations are much more similar (Fig. 2*b*), with an r.m.s. deviation of only 0.5 Å (Fig. 2*a*). This indicates that the inner structure is quite similar to the fiber A-DNA. Comparison with the crystal structure of d(CCCGGCCGGG) (Ramakrishnan & Sudaralingam, 1993) shows that the two terminal nucleotides in the present structure have moved inward and that the major difference lies in the phosphate groups, as indicated by the average r.m.s. deviation of 2.00 Å for the phosphate groups compared with the r.m.s. deviation of 1.15 Å for superposition of the whole molecule.

3.2. Crystal packing

The unit-cell parameters of the present decamer are strikingly similar to those of tetragonal form A-DNA octamers, being 4% greater than the averaged *a* and *b* axes and the same as the *c* axis of the octamers. A decamer has roughly 25% more atoms than an octamer, but the present decamer crystal has only an 8% greater volume than the tetragonal octamer crystals. Thus, the present A-DNA decamer is expected to be more tightly packed than the octamers. The crystal packing of the decamer is quite similar to that observed in the crystal structures of the tetragonal octamers, where the terminal base pairs abut into the minor groove of the symmetry-related A-DNA duplex. In the octamer crystals such packing occurs around the center of the duplexes and results from the low twist angle (∼20°) for the central base pairs. In the present decamer (which is longer in length) the terminal base pairs of the symmetry-related molecules abuts one base pair away from the center of the DNA duplex, where a low twist angle of 25° is observed for the G(4)·C(7) and A(5)·T(6) base pairs and a normal twist angle is observed for the central base pairs.



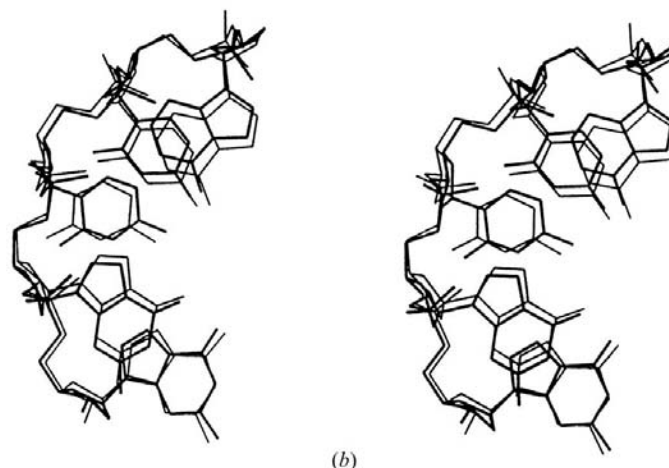


Figure 2
Difference between the present structure and the fiber model. (*a*) R.m.s. deviation between the present structure and the fiber model, showing that the inner nucleotides are much more similar to the fiber model than the terminal nucleotides; (*b*) superposition of nucleotides A5–G9 (thick bonds) with the fiber model (light bonds).

3.3. Intermolecular interactions and their influence on the thermal vibration of the DNA

In the crystal, the terminal base pair C(1)·G(10) interacts with the inner base pairs C(3)·G(8) and G(4)·C(7) of the symmetry-related duplexes, forming two different 'base-paired triplets'. In one base triplet, inter-duplex interaction occurs between two guanine bases N2(G8)···N3(G10) (3.11 Å) and N2(G10)···N3(G8) (3.27 Å) (Fig. 3*a*). In the other base triplet, inter-duplex interaction occurs between two guanine bases N2(G10)···N3(G4) (2.86 Å) and guanine and cytosine bases N2(G4)···O2(C1) (2.97 Å) (Fig. 3*b*), which is similar to that observed in the crystal structure of the A-DNA decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam; 1993). Obviously, G10 is strongly involved in hydrogen-bonding interactions. The present result indicates that purines may be more important than pyrimidines in crystal packing interactions in A-DNA.

The thermal vibrations of atoms in crystal structures are usually measured by thermal parameters (*B* factors). In constrained/restrained refinement methods, however, errors introduced by data of poor quality and related parameters such as bulk solvent are often dumped to the isotropic *B* factors. It is often meaningless to correlate the thermal parameters with the crystal structure at low resolution. The present structure is determined at near-atomic resolution and most of the solvent molecules are located clearly; therefore, the thermal parameters are expected to reflect the mobility/flexibility of the molecule in the crystals. Fig. 4 shows the average isotropic thermal parameters for various groups in the present structure. In contrast to the expectation that terminal nucleotides will have larger *B* factors than the central nucleotides, G10 has the lowest *B* factor. This anomaly can be explained by the fact that the terminal nucleotide G10 is involved in a strong crystal packing interaction (Figs. 3*a* and 3*b*). G4 has lower *B* factors than G8 because G4 has stronger hydrogen bonds than G8 in the abutting interaction (2.86 and 2.97 Å for G4 compared with 3.11 and 3.27 Å for G8). These results show that the hydrogen-bonding interactions hold the nucleotides tightly so that they have less degrees of freedom in thermal vibrations. In the present structure, the thermal vibration and the hydrogen-bonding interaction are very well correlated and the *B* factors are indeed an appropriate measure of thermal vibration.

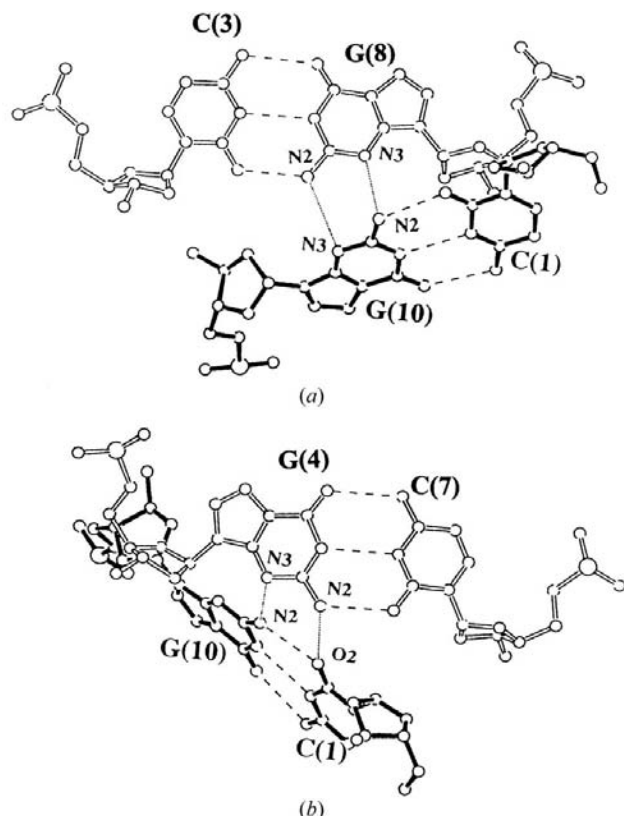


Figure 3
Two different conformations of the base triplets G*(G·C) observed in the minor groove as a result of the abutting crystal packing.

3.4. Cobalt hexammine

One cobalt hexammine ion is located at the edge of the major groove per independent strand in the present structure. It is hydrated with nine water molecules (Fig. 5*a*) and involved in interactions with the sugar–phosphate backbones of four adjacent duplexes (Fig. 5*b*). All the direct interactions are listed in Table 1. The cobalt hexammine ion does not directly interact with the base atoms in the structure. Instead, water molecules bridge the cobalt hexammine ion with the N7 and O6 atoms of two adjacent GpG nucleotides, G9 and G10 (Fig. 5*c*).

It has been shown that cobalt hexammine ion induces formation of the A-form helix for DNA molecules with stretches of GpG sequences and promotes the B–A transition towards the A-form in solution (Braunlin & Xu, 1992; Xu *et al.*, 1993). Cobalt hexammine ions have different binding modes for A-DNA duplexes, either bridging phosphate groups or adhering to the edges of two adjacent guanines (Gao *et al.*, 1995; Nunn & Neidle, 1996). The present binding mode is

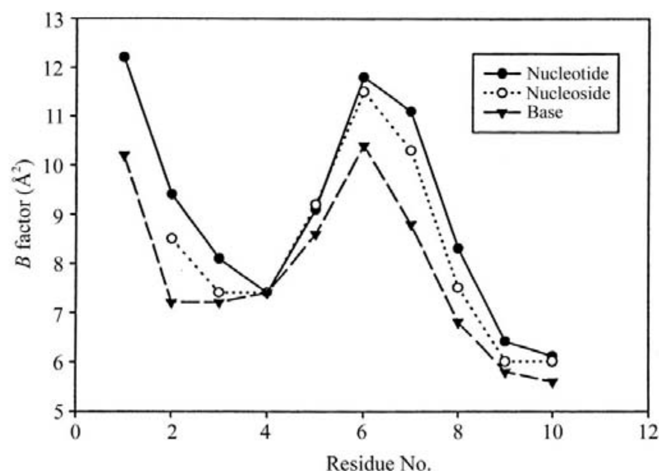


Figure 4
A diagram illustrating the isotropic temperature factor (*B* factor) for various groups, showing that the nucleotides involved in the crystal packing interactions have smaller *B* factors.

different. Cobalt hexammine ion interacts directly with the sugar–phosphate backbone on one side and indirectly with guanine base atoms by mediation of water bridges on the other side. It is located at the edges of the major groove and terminal regions of the duplex, interacting with the phosphate groups of four adjacent duplexes. Thus, the cobalt hexammine ion holds the adjacent duplexes tightly and contributes greatly to the crystal packing.

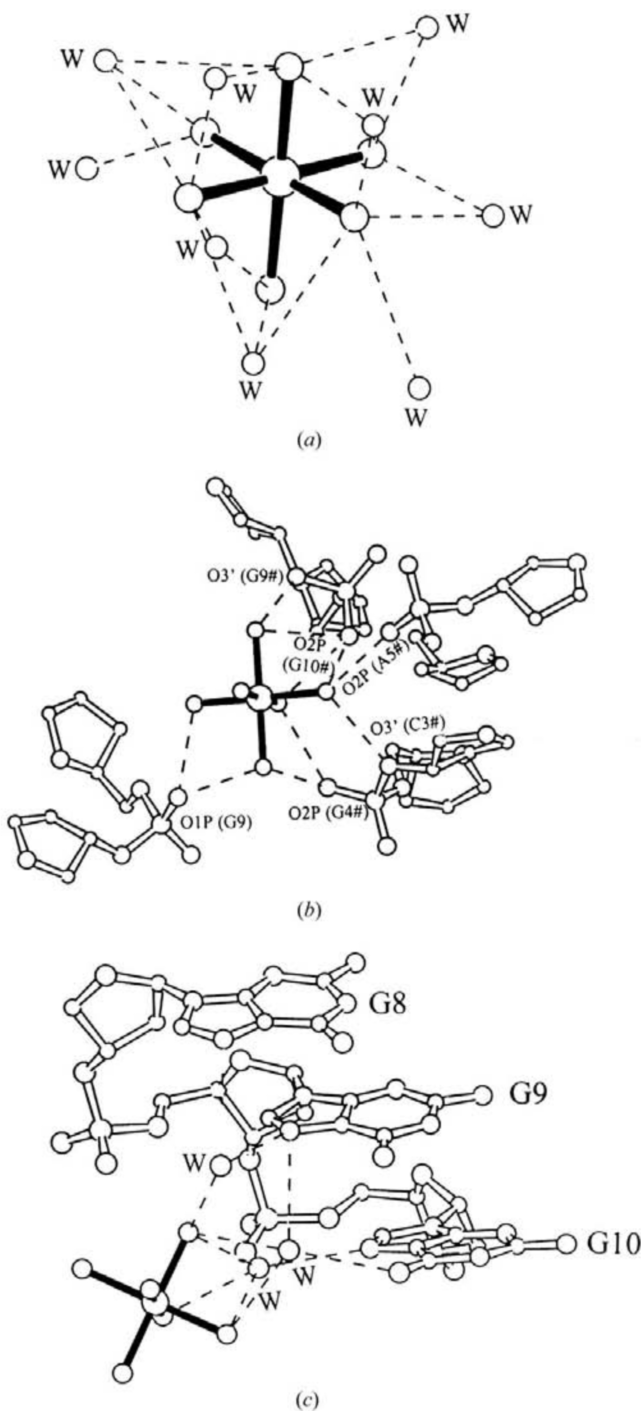


Figure 5
Interactions of the cobalt hexammine ion in the present structure. (*a*) Hydration of the cobalt hexammine ion (W denotes water molecules); (*b*) the cobalt hexammine ion binds to phosphate groups from four different duplexes (# denotes symmetry-related strands); (*c*) the cobalt hexammine ion binds to guanines *via* water bridges.

Table 1
Cobalt hexammine hydrogen-bonding interactions.

CoHex	DNA atom	Distance (Å)	Symmetry†
N1	W103	2.59	
N1	W108	3.16	
N1	W129	3.32	
N1	W123	2.71	5 0 0 1
N2	O3′(G9)	2.95	4 0 0 0
N2	O1P(G10)	3.39	4 0 0 0
N2	W101	3.20	
N2	W102	2.85	
N2	W103	3.22	
N2	W104	2.54	5 0 0 1
N3	O2P(G9)	2.84	
N3	W101	2.85	
N3	W102	2.94	
N3	W108	3.17	
N3	W131	3.12	
N4	O1P(G4)	2.82	7 0 −1 0
N4	O2P(G9)	2.89	
N4	W108	3.21	
N4	W129	3.28	
N5	O1P(G10)	3.32	4 0 0 0
N5	O3′(C3)	3.06	7 0 −1 0
N5	O1P(A5)	3.05	5 0 0 1
N5	W104	2.96	5 0 0 1
N5	W123	2.80	5 0 0 1
N6	O1P(G10)	2.93	4 0 0 0
N6	O1P(G4)	2.94	7 0 −1 0
N6	W101	2.81	
N6	W111	2.57	7 0 −1 0

† The numbers in the symmetry column represent the following symmetry operations: 4, 0, 0, 0 = $\frac{1}{2}+y, \frac{1}{2}-x, \frac{1}{4}+z$; 7, 0, −1, 0 = $\frac{1}{2}-x, -\frac{1}{2}+y, \frac{3}{4}-z$; 5, 0, 0, 1 = $y, x, 1-z$.

Cobalt hexammine ions have been observed to bind in two (Nunn & Neidle, 1996) and in three different sites (Gao *et al.*, 1995) in DNA crystal structures, involving the base atoms of guanines and phosphate groups. In the present structure, only one kind of binding site, at phosphate groups, is observed. This result implies that cobalt hexammine ions may have a preference to bind to phosphate groups over the other binding sites. This is consistent with the electrostatic interaction between the negatively charged phosphate group and the positively charged cobalt hexammine ion.

3.5. Hydration

High-resolution crystal structures can provide a detailed and accurate description of the hydration of DNA molecules. In the present structure, water molecules were well located with an average *B* factor of 30.8 Å^2. Water molecules in the present structure mainly hydrate three different kinds of groups: the backbone phosphates, base atoms and cobalt hexammine ions. Of 34 water molecules, 15 are used to hydrate backbone phosphates, 14 to hydrate base atoms and nine to hydrate the cobalt hexammine ion. It is noted that some water molecules may hydrate more than one kind of atom.

Fig. 6 shows the hydration of the backbone phosphates. Of the 15 water molecules, nine hydrate O1P of the phosphates and six hydrate O2P. It is clear from the figure that the

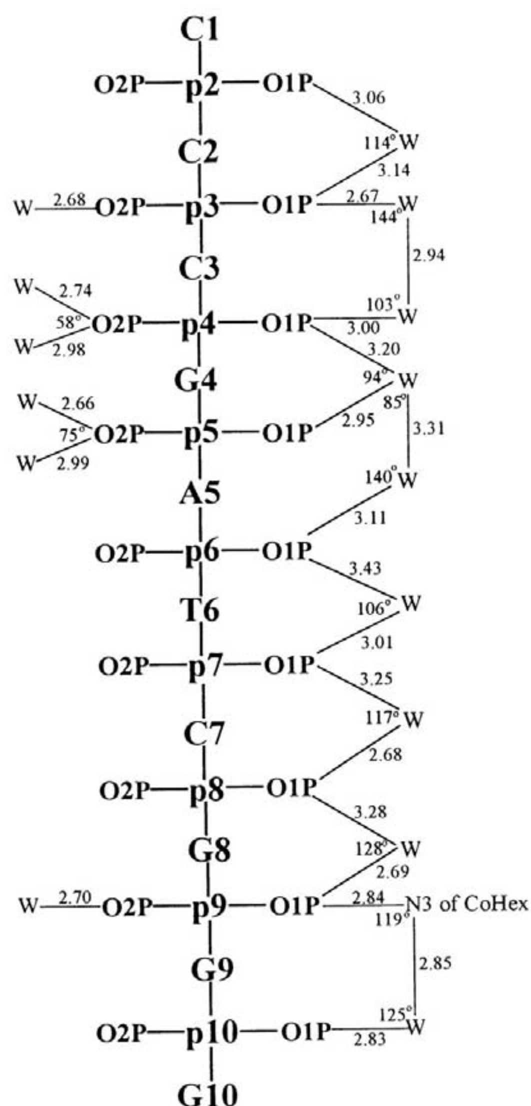


Figure 6
Schematic diagram of hydration of the phosphate groups in the present structure.

hydration of O1P forms a water-bridge pattern along the backbone. The pattern is perturbed at two sites: one is between G4 and A5 and arises from the crystal packing interaction of the abutting terminal bases, while the other is at G9 and G10 and is a consequence of the interaction of the cobalt hexammine ion. The hydration pattern of O2P is less regular because O2P is exposed to the outside of the duplex and is thus influenced more by the environment.

The major groove is more hydrated than the minor groove. O6 and N7 of all guanines, N6 and N7 of all adenines and N4 of all cytosines in the structure are hydrated. O4(T6) is not hydrated. Water bridges are observed in the adjacent purine steps: G4–A5 and G8–G9–G10 (Table 2). No water bridges are found in adjacent pyrimidine steps. In the minor groove, only N3(A5), O2(T6) and O2(C7) are hydrated. All the atoms at both termini are not hydrated, which may be the result of crystal packing interactions.

Table 2
Water bridges in the major groove.

DNA atom	Distance (Å)	H_2O	Distance (Å)	DNA atom
O6(G4)	2.71	W107	3.05	N6(A5)
O6(G8)	2.90	W110	2.63	O6(G9)
N7(G9)	3.21	W102	2.93	O6(G10)

We gratefully thank the National Institutes of Health for supporting this work through grant GM-17378 and the Board of Regents of Ohio for an Ohio Eminent Scholar Endowment to MS.

References

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.
Braunlin, W. H. & Xu, Q. (1992). *Biopolymers*, **32**, 1703–1711.
Brünger, A. T. (1994). *X-PLOR Manual, Version* 3.1. Yale University, New Haven, CT, USA.
Chiu, T. K. & Dickerson, R. E. (2000). *J. Mol. Biol.* **301**, 915–945.
Deng, J. P., Xiong, Y. & Sundaralingam, M. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 13665–13670.
Egli, M., Tereshko, V., Teplova, M., Minasov, G., Joachimiak, A., Sanishvili, R., Weeks, C. M., Miller, R., Maier, M. A., An, H., Cook, P. D. & Manoharan, M. (1998). *Biopolymers*, **48**, 234–252.
Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Ho, P. S. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3971–3976.
Franklin, R. E. & Gosling, R. G. (1953). *Nature (London)*, **171**, 740–741.
Frederick, C. A., Quigley, G. J., Teng, M.-K., Coll, M., van der Marel, G. A., van Boom, J. H., Rich, A. & Wang, A. H.-J. (1989). *Eur. J. Biochem.* **181**, 292–307.
Gao, Y. G., Robinson, H., van Boom, J. H. & Wang, A. H.-J. (1995). *Biophys. J.* **69**, 559–568.
Heinemann, U. & Alings, C. (1989). *J. Mol. Biol.* **210**, 369–381.
Heinemann, U., Alings, C. & Bansal, M. (1992). *EMBO J.* **11**, 1931–1939.
Howard, A. J. (1993). Abstr. Am. Crystallogr. Assoc. Annu. Meet., Abstract C001.
Kielkopf, C. L., Ding, S., Kuhn, P. & Rees, D. C. (2000). *J. Mol. Biol.* **296**, 787–801.
Nunn, C. M. & Neidle, S. (1996). *J. Mol. Biol.* **256**, 340–351.
Ramakrishnan, B. & Sudaralingam, M. (1993). *J. Mol. Biol.* **231**, 431–444.
Saenger, W. (1984). *Principles of Nucleic Acid Structure.* New York: Springer–Verlag.
Xu, Q., Schoemaker, R. K. & Braunlin, W. H. (1993). *Biophys. J.* **65**, 1039–1049.