Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of d(TGCGCG)·d(CGCGCA) in two crystal forms: effect of sequence and crystal packing in Z-DNA

S. Thiyagarajan, S. S. Rajan and N. Gautham*

Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai 600 025, India

Correspondence e-mail: gautham@unom.ac.in

Received 30 March 2005
Accepted 26 May 2005

NDB References: ZD0013; ZD0014.

The sequence d(TGCGCG)·d(CGCGCA) crystallized in two crystal forms, orthorhombic and hexagonal, in the presence of cobalt hexammine chloride, a known inducer of the left-handed Z-form of DNA. The crystal structures have been solved and refined at 1.71 Å resolution in space group $P2_12_12_1$ and 2.0 Å resolution in space group $P6_5$. The orthorhombic structure contains one Z-DNA hexamer duplex, while the hexagonal structure contains two hexamer duplexes in the structure. Of the latter, one is situated on a crystallographic sixfold screw axis, leading to disorder. This paper reports the effects of sequence and crystal packing on the structure of Z-type DNA. The structures lend additional support to the authors' earlier conclusion that a stretch of four C·G base pairs is sufficient to nucleate and define the regular model of the left-handed helix based on the structure of $d(CGCGCG)_2$.

1. Introduction

In our programme to use X-ray crystallography to study the effects of A·T base pairs in Z-DNA, we have undertaken the determination of the structures of a series of hexameric DNA fragments with sequences based upon the 'canonical' Z-DNA sequence $d(CGCGCG)_2$ (Wang *et al.*, 1981). Since this sequence is self-complementary, the duplex has a dyad and there are three ways in which an A·T base pair could replace one of the C·G base pairs. Depending on the position of this replacement, the new sequences are termed HT1, HT3 and HT5, referring to the first, third and fifth positions of thymine. The different substitutions also result in a variation of the extent of the G·C tract in the sequences. In the first sequence, we have four consecutive G·C base pairs, in the second we have three, while in the third there are five consecutive G·C base pairs. We have earlier reported the crystal structures of d(CGCGTG)·d(CACGCG) and d(CGTGCG)·d(CGCACG) (*i.e.* HT5 and HT3, respectively; Sadasivan & Gautham, 1995). HT5 was a high-resolution structure and its conformational parameters matched those of the Z-DNA model (Wang *et al.*, 1981) almost exactly. The structure of HT3 could be obtained only at a lower resolution and its conformational parameters deviated substantially from the model. We have now solved the structure of the third sequence, HT1, in two different crystal forms. These structures show that HT1, with a G·C tract length of 5, is closely similar to the fibre model Z-helix. Taken together with previous reports, both from our laboratory (Karthe & Gautham, 1998; Sadasivan & Gautham, 1995; Thiyagarajan *et al.*, 2002) as well as from others (Harper *et al.*, 1998; Parkinson *et al.*, 1995; Wang *et al.*, 1984; Geierstanger *et al.*, 1991; Coll *et al.*, 1988), the observation of Wang *et al.*

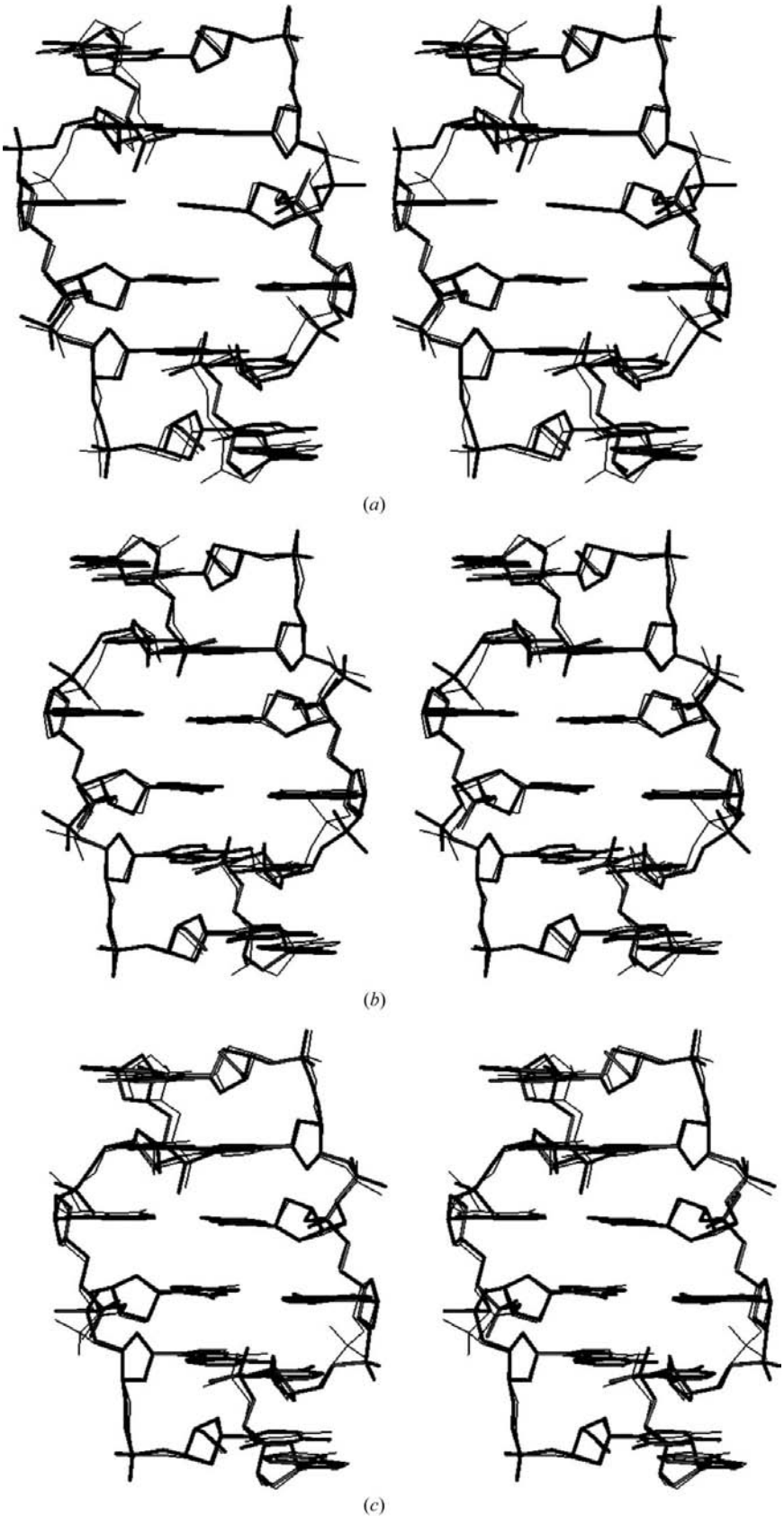

Figure 1
Least-squares superposition diagrams of (*a*) OrthHT1 and (*b*) HexHT1 (thick lines) with fibre-model Z-type hexamer (faint lines) and (*c*) of OrthHT1 (thick lines) with HexHT1 (faint lines).

(1987) that the formation of Z-DNA is dependent on the length of the alternating C·G tract is supported. However, there are a few exceptions to this hypothesis. For example, the sequence $(GCGCGC)_2$ crystallized as A-form DNA (Mooers *et al.*, 1995) even though there were two consecutive CpG steps. In another study, the dinucleotide CpG assumed the Z conformation and formed a pseudo-continuous Z-type helix (Ramakrishnan & Viswamitra, 1988).

Packing interactions also influence the Z-type structure. In the most commonly observed packing mode in Z-DNA crystal structures, the approximately cylindrical hexamers are stacked one over the other along the helical axis to form infinite columns, which are then bundled together to form the crystal. The molecular structure is rigid and all geometrical parameters have values close to those found in the fibre model. However, when the helix adopts a different packing scheme, the structure ceases to be rigid and shows a range of variability (Malinina *et al.*, 1998).

In an earlier paper based on the present HT1 structures (Thiyagarajan *et al.*, 2004), we primarily discussed the role of cobalt hexammine ions in bringing about a change in the tautomeric state of the adenine base, leading to a wobble base-pairing scheme for the A·T base pair. In the present paper, we focus the discussion on the effects of sequence and crystal packing on the structure of HT1.

2. Materials and methods

We have already reported the crystallization and structure solution of both crystal forms of HT1 elsewhere (Thiyagarajan *et al.*, 2004). For completeness, we briefly repeat these descriptions here. Crystals were grown at room temperature (293 K) by the hanging-drop vapour-diffusion method in the presence of cobalt hexammine chloride. Data collection was carried out using a MAR Research imaging-plate system with Cu $K\alpha$ radiation generated by a rotating-anode generator. One data set was indexed in the orthorhombic system. The space group is $P2_12_12_1$, with unit-cell parameters $a = 17.98$, $b = 30.93$, $c = 44.63$ Å. There were 2562 unique reflections to a resolution of 1.71 Å, with an overall $R_{\rm merge}$ of 5.7% and a completeness

of 93.9%. The second crystal was indexed in the hexagonal space group $P6_5$, with unit-cell parameters $a = b = 35.59$, $c = 44.52$ Å. In this data set there were a total of 2686 unique reflections, with an overall $R_{\rm merge}$ of 15% and 99% data completeness to 2.0 Å resolution. The structures were solved by molecular replacement using *AMoRe* (Navaza, 1994) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The starting model was based on the coordinates of the fibre model Z-DNA (Wang *et al.*, 1981), with A·T replacing one of the terminal C·G base pairs.

Using the orthorhombic data set, the structure was refined using *REFMAC*5 (Murshudov *et al.*, 1997) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) to a final R factor of 20.8% ($R_{\rm free}$ = 24%). In the hexagonal crystal, the asymmetric unit contains two hexameric duplexes, one of which is disordered; its helix axis coincides with the crystallographic sixfold screw axis. The refinement converged to an R factor of 27.3%, with a free R value of 35%. In the rest of the paper, the orthorhombic structure is called OrthHT1, and has the following numbering scheme.

```
5′ T1  p G2  p C3  p G4 p C5 p G6 3′
3′ A12 p C11 p G10 p C9 p G8 p C7 5′
```

In the hexagonal structure, the fully ordered hexamer is referred to as HexHT1. An arbitrary hexamer is chosen to represent the other disordered helix and is called DinuHT1. The numbering scheme for HexHT1 is the same as above. The dinucleotide residues have the following numbering scheme.

```
5′ p T13 p G14 3′

3′ p G16 p T15 5′
```

G14 and G16 will be addressed as A14 and A16, respectively, whenever these purines are considered to represent adenine bases. A similar numbering scheme is adopted for other hexameric sequences taken for comparison. Conformational and helical parameter calculations were carried out using the programs *FREEHELIX*98 (Dickerson, 1998) and *CURVES* (Lavery & Sklenar, 1988). Fig. 2 was plotted using *PyMol* (DeLano, 1998). The coordinates and structure factors have been deposited in the NDB (codes ZD0013 and ZD0014).

3. Results and discussion

3.1. Molecular structure

The duplex sequence d(TGCGCG)·d(CGCGCA) has an A·T base pair at the terminus followed by five consecutive G·C base pairs. Although these hexamers crystallize in two different space groups and show variations in crystal packing, in all three independent hexamers the molecular structure is very close to the 'canonical' Z-DNA helix (Wang *et al.*, 1981). This lends support to the hypothesis (Sadasivan & Gautham, 1995) that a stretch of four consecutive G·C base pairs (*i.e.* two consecutive CpG steps) is necessary and sufficient to nucleate the canonical Z-type DNA. This is despite the variations in base pairing in OrthHT1 that arise owing to the binding of cobalt hexammine (Thiyagarajan *et al.*, 2004).

Fig. 1 shows the least-squares superposition of the three hexamers reported in this paper on each other as well as on the fibre model Z-DNA. The root-mean-square deviations in the positions of the common atoms after the superposition are given in supplementary tables.[1] The r.m.s.d. values between the fibre model and the present structures are 1.1 and 0.6 Å for OrthHT1 and HexHT1, respectively. Thus, the structure of HT1 in both crystal forms is similar to the fibre model. The most noticeable change from the fibre model is the Z_{II}-type backbone structure at residues C5 and C9 in OrthHT1. In other words, the backbone conformation is *gauche*$^+$, *trans* about ζ and α at these two residues compared with *gauche*$^-$, *trans* in the fibre model. The Z_{II} conformation was first noticed in the crystal structure of d(CGCGCG)$_2$ (Wang *et al.*, 1979), in which only residue C5 had this conformation. However, from other studies (Harper *et al.*, 1998) it became apparent that the Z-helix could accommodate either Z_I or Z_{II} without any clear sequence specificity. For example, in the structure of d(TGCGCA)$_2$ (Harper *et al.*, 1998; Thiyagarajan *et al.*, 2002), the backbone has the Z_{II} conformation at residues C3, C9 and C11. This sequence was also crystallized in presence of cobalt hexammine chloride, although the presence of the ion is not correlated with the Z_{II} structure, except for its binding to the phosphate group of residue C3. In OrthHT1, the Z_{II} conformations at C9 and C5 may be related to cobalt hexammine interactions. The phosphate group of C9 interacts with the amine group of the ion directly, while the phosphate group of C5 is bound to a hexammine ion through two water molecules (see Fig. 2). The ion also interacts with the N7 atom of a symmetry-related base G8 through another water molecule.

Unlike OrthHT1, HexHT1 has a pure Z_I conformation at all purine–pyrimidine steps. HT5 also showed Z_I conformation at all bases. It is notable that even at the C5 phosphate position, where the crystal structure of d(CG)$_3$ showed Z_{II} conformation (Wang *et al.*, 1979), HexHT1 shows Z_I conformation. DinuHT1 has unconventional backbone conformations such as *gauche*$^-$, *gauche*$^+$ or *gauche*$^-$, *trans*. However, this helix is disordered and has a discontinuous backbone. The conformations seen are therefore averages of different possibilities and we will not discuss these further.

Base-geometry parameters of the helices and the backbone torsion angles of the bases are given in the supplementary tables.[1] As expected for Z-DNA, the helices have dinucleotide repeats with an average twist of -61.3° per dinucleotide in case of OrthHT1, -60.7° for HexHT1 and an ideal -60.0° for DinuHT1 (owing to its placement on a sixfold screw axis). In OrthHT1 the pyrimidine–purine steps have an average twist of -11.3°, while the two purine–pyrimidine steps have an average twist of -50.0°. In HexHT1 the average twist values are -12.4 and -48.3° at the two types of base step, respec-

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5045). Services for accessing these data are described at the back of the journal.

Table 1
Interhelical contacts in the two structures.

Space group	Atom	Distance (Å)	Atom	Distance (Å)	Atom	Distance (Å)	Atom	Distance (Å)	Atom	Remarks
$P2_12_12_1$	O2P (C9)	2.72	$[Co(NH_3)_6]^{3+}$ (N2)							Z_{II} conformation
	O3′ (A12)	2.94	O2P (G2)							Mode 1 packing
	O3′ (G6)	2.74	O1P (C9)							Mode 1 packing
	N7 (G2)	3.09	O24	2.69	O2P (C11)					
	O3′ (G6)	2.84	O25	2.86	O5′ (T1)					
	O2P (G10)	2.92	O27	2.44	O1P (C5)					
	O2P (C3)	2.97	O51	2.83	O2P (G8)					
	O1P (C5)	2.88	O47	3.23	N7 (G8)					
	O5′ (T1)	2.72	O19	3.09	O40	2.83	N2 (G6)			
	O2P (C5)	2.41	O34	3.00	O18	3.08	$[Co(NH_3)_6]^{3+}$ (N2)			Z_{II} conformation
	O1P (G4)	2.91	O44	3.11	O45	3.16	N4 (C3)			
	O2P (C5)	2.41	O34	3.00	O18	2.41	O56	2.68	N4 (C9)	
	N4 (C11)	2.98	O46	2.51	O17	3.06	O43	2.91	O6 (G2)	
$P6_5$	O3′ (G6)	2.98	O2P (A12)							
	O4′ (T14)	2.37	O20	2.90	O2P (G4)					
	O2P (G10)	2.90	O20	3.09	O22	3.05	O2P (T14)			

tively. In DinuHT1 the values are −11.5 and −48.5°. The base-step rise at all the steps in all three helices is similar to the values seen in other crystal structures of Z-type helices (Ho & Mooers, 1997).

In OrthHT1, the terminal T1pG2 base step has a high negative tilt (−5.1°) compared with C1pG2 of $d(CG)_3$ (0.9°). There is a similar change in the value of inclination at the terminal base pair. However, this behaviour is apparently not directly related to the A·T base pair, since in HT5 and HT3 the pattern is quite different. It may instead be attributed to the interaction of cobalt hexammine with the helix (Thiyagarajan *et al.*, 2004). The terminal base pairs in OrthHT1 as well as HexHT1, T1·A12 and G6·C7, show high propeller twists (9.2 and 5.7° in OrthHT1, and 5.5 and 5.7° in HexHT1, respectively) compared with those (−0.2 and 2.4°) in the crystal structure of $d(CG)_3$. This could be attributed to the cross-strand hydrogen bond between the helices stacked one over the other.

Other major differences between HexHT1 and OrthHT1 are the following. (i) A high negative tilt of −4.6° at C11pA12 step compared with the positive value in OrthHT1. (ii) The high roll angle (10.7°) at the G10pC11 step. Such a high value for roll was also seen in HT3 at the A10pC11 base step. (iii) The higher twist of −14.6° at the C11pA12 step. Among the structures discussed here, HT5 has such a higher twist of about −14° at the C11pG12 base step, while other sequences have values less than −12° at this terminal step.

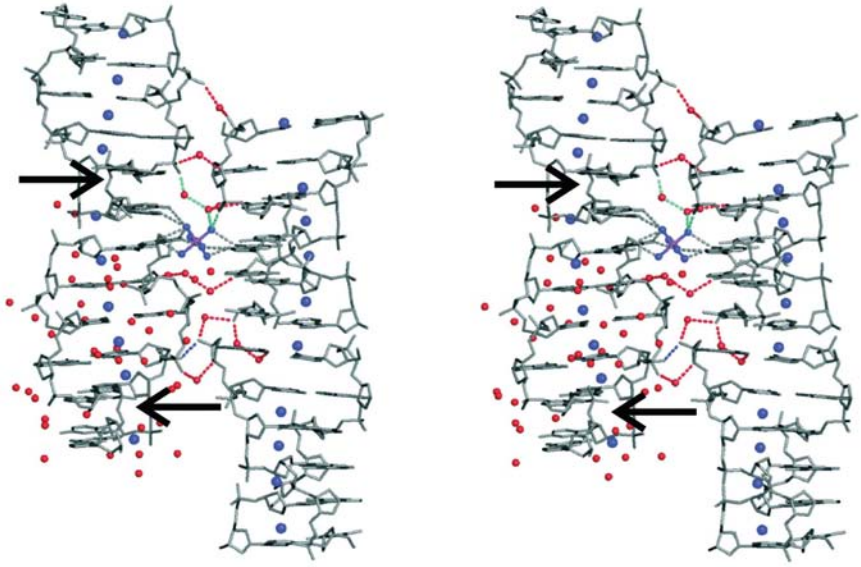

Figure 2
Water network in OrthHA6 structure. The spine of hydration in shown as blue spheres. The arrows point to the discontinuity in the spine. However, this is not related to the presence of the A·T base pair. Interhelical contacts mediated through water are shown as red dotted lines. Cyan dotted lines indicate interactions of the cobalt hexammine ion with the phosphate groups (which have a Z_{II} conformation). Blue dotted lines depict contacts arising from mode 1 packing.

3.2. Crystal packing

In general, there are two kinds of packing modes common to Z-DNA crystal structures in orthorhombic system. The difference between the two modes is in the interhelical interaction. Mode 1, previously termed the 'magnesium form' (Gessner *et al.*, 1989), is seen in crystals grown in the the presence of magnesium. However, later experiments showed that the presence of magnesium is not essential for mode 1 packing. Mode 2, previously called the 'spermine form' (Egli *et al.*, 1991), was adopted when Z-DNA was crystallized in presence of spermine alone. Again, later structures showed that Z-DNA can adopt mode 2 packing even in the presence of magnesium (Moore *et al.*, 1995). The two modes are characterized by specific interhelical interactions observed in the crystal packing (Harper *et al.*, 1998; Kumar & Weber, 1993; Egli *et al.*, 1991). The present crystal

structure did not have magnesium in the crystallization drop and may therefore be expected to adopt mode 2 packing. However, similar to the HT5 structure, OrthHT1 adopts mode 1 packing, adding to existing examples showing that the presence of magnesium is not critical for such packing. A characteristic feature of mode 1 packing is that the O3′ atom of residue 6 interacts with the phosphate group of residue 9 and the O3′ atom of residue 12 interacts with the phosphate group of residue 2, which is seen in the present structure (see Fig. 2). The distances between the atoms involved in these interactions are listed in Table 1.

Probably the strongest packing interactions in both the crystal forms are those mediated by the hexammine ion. These have been described in detail elsewhere (Thiyagarajan *et al.*, 2004). Apart from these, there are several other interhelical interactions mediated by water molecules. For instance, in OrthHT1 the phosphate group of G8 interacts with the phosphate group of C3 of a symmetry-related hexamer through a water molecule (Fig. 2). Similarly, the C5 phosphate group interacts with N7 of G8 of the symmetry-related hexamer (Fig. 2). The other interhelical interactions mediated by water are listed in Table 1.

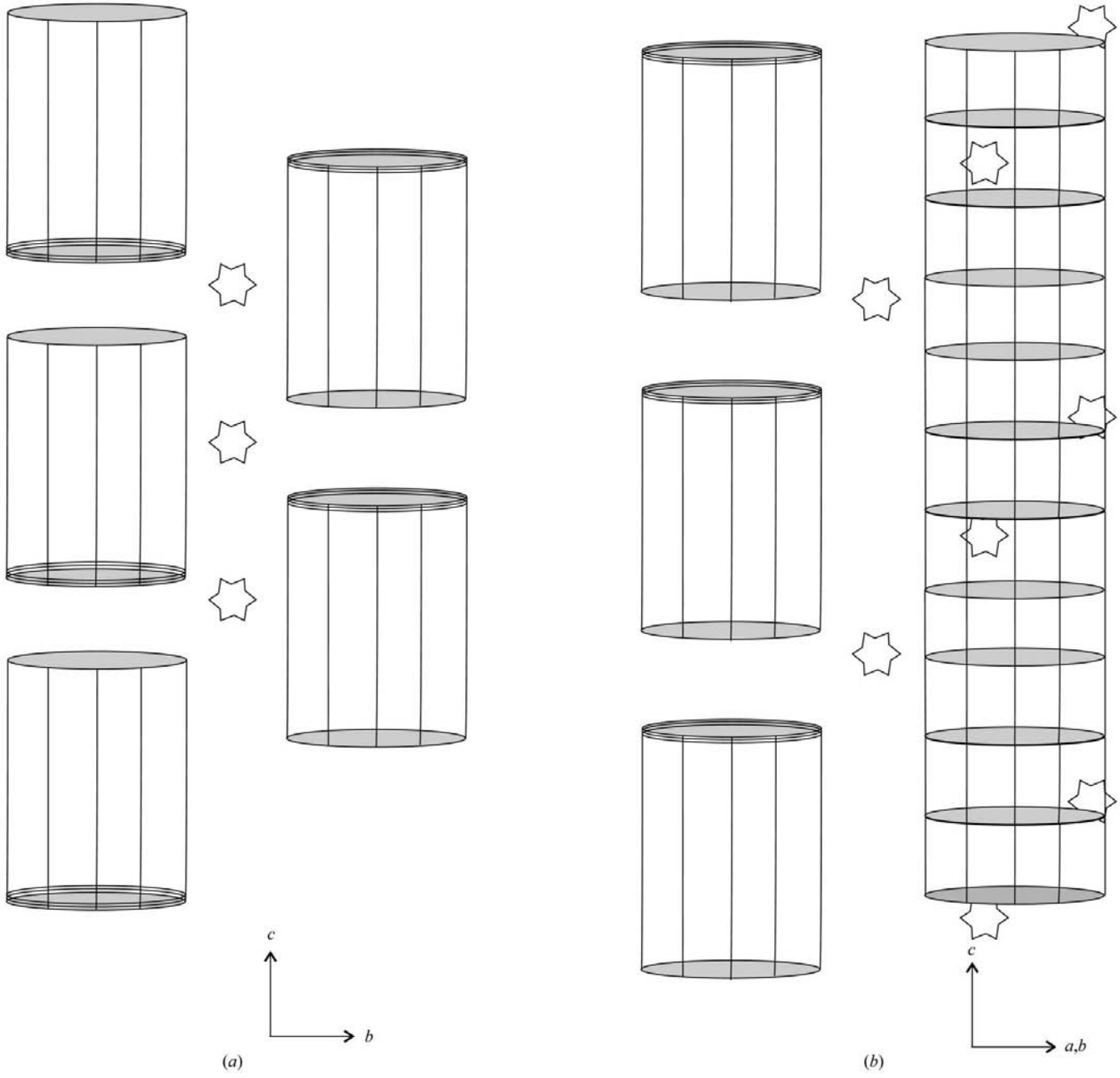

Figure 3
Schematic diagram representing the ion binding in (*a*) the orthorhombic structure and (*b*) the hexagonal structure. The cylinders represent Z-DNA hexamers, while the stars are the cobalt hexammine ions. Note that in the hexagonal structure the ion and one set of hexamers are disordered.

The base stacking in the helices does not show significant difference compared with other Z-DNA structures, except at the virtual step between one hexamer and the next. The twist at the virtual GpT step is −61.1°, higher than in any of the previous Z-DNA crystal structures. This may have arisen owing to the high shear at the terminal T1·A12 base pair. We note that the ApT virtual step between symmetry-related helices in the structure of HAT61 has a twist of only −45.6°.

In the other crystal form, *i.e.* the $P6_5$ structure, there are two kinds of interhelical interactions: the interactions between HexHT1 helices and those between HexHT1 and DinuHT1. HexHT1 helices show interactions between the O3′ atom of G6 and the phosphate group of A12. There are also a few water-mediated interactions between the helices of HexHT1 and DinuHT1. Table 1 again lists some of the interactions between these helices. In HexHT1, the twist at the virtual step GpT step has a surprisingly low value of −19.9°. DinuHT1 has a discontinuous backbone and therefore no single virtual step can be identified. In general, the packing of OrthHT1 is closer to that seen in $d(CGCGCG)_2$ (Wang *et al.*, 1979) than in HexHT1. However, as far as the molecular structures are concerned, it is HexHT1 that is closer to $d(CG)_3$ than OrthHT1.

As mentioned above, the interactions made by the cobalt hexammine ion in both crystal structures are chiefly responsible for the two different crystal forms. In OrthHT1, there is one ion per hexamer, making an extensive network of contacts with neighbouring helices. The ion is at the centre of the interactions between three symmetry-related helices, as shown schematically in Fig. 3(*a*). In the hexagonal crystal form there is only a quarter of an ion per hexamer, reflecting the lower concentration of cobalt hexamine in the crystallization solution. The ion position is disordered, leading to disorder in the position of the neighbouring helix. Thus, alternate columns of helices are disordered along the *a* and *b* axes (Fig. 3*b*).

3.3. Solvent interactions

OrthHT1 shows a very strong water network stabilizing the structure as well as the crystal packing (Fig. 2). Some common features of Z-DNA structures, namely the spine of hydration and the water molecules bound to the backbone phosphates stabilizing the structure, are present in OrthHT1. In general, the O2 of pyrimidines are hydrated in Z-DNA and these water molecules form the spine of hydration. Each water molecule constituting the spine bridges the O2 atoms of pyrimidines of the base pairs above and below it (Gessner *et al.*, 1994). In the present structure, this spine of hydration is present between all the base pairs except between C1·G12 and G2·C11. This may be a consequence of the interaction of the C1·G12 base pair with the cobalt hexammine ion, leading to a wobble-type G·C base pair (Thiyagarajan *et al.*, 2004). N4 of C1 interacts with the cobalt hexammine ion, leading to a high shear of −1.14 Å at this base pair, in contrast to the ±0.2 Å shear at other base pairs. This might have altered the interaction between the base pairs C1·G12 and G2·C11, leading to a break in the spine (Fig. 2).

Other probable hydration regions in Z-DNA are N7 and O6 of the guanine bases. In the present structure, all the pyrimidine bases except C5 have at least one water molecule within 3.1 Å of O2. The N2 amine of all the guanine bases is also hydrated. In OrthHT1, the N7 and O6 atoms of two of the five guanine bases are bound to the cobalt hexammine ion and not to water. In the other three guanines, O6 of G2 and N7 of G8 have one water molecule within 3.2 Å, while G10 does not show any hydration. The adenine (A12) base also has strong interactions with the ion and does not show any hydration at N6 and N7. In the hexagonal crystal not all regions of the Fourier map could be interpreted. No clear density corresponding to the spine of hydration was visible. A few water molecules bound to the backbone phosphate groups were identified. O4′ of T14 is bridged to the phosphate group of G4 through one water molecule, while the phosphate group of T14 interacts with that of G4 through two water molecules.

4. Summary

The structures lend additional support to the hypothesis that two consecutive CpG steps are necessary to nucleate and form a regular Z-DNA helix. The packing of HT1 in orthorhombic form is closer to that seen in $d(CG)_3$ than its hexagonal counterpart, although as far as the molecular structures are concerned it is the hexagonal structure that is closer to $d(CG)_3$ than the orthorhombic structure.

This work was supported by DST-FIST and UGC-SAP funding. ST thanks CSIR, India for financial support as an SRF. The X-ray data-collection facilities at the Structural Biology Unit, National Institute of Immunology, New Delhi, the GNR Center for Structural Biology, Central Leather Research Institute, Chennai and the National Area Detector Facility at Molecular Biophysics Unit, Indian Institute of Science, Bangalore are acknowledged.

References

Coll, M., Fita, I., Lloveras, J., Subirana, J. A., Bardella, F., Huynh-Dinh, T. & Igolen, J. (1988). *Nucleic Acids Res.* **16**, 8695–8705.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (1998). *The PyMol Molecular Graphics System.* DeLano Scientific LLC, San Carlos, CA, USA.
Dickerson, R. E. (1998). *Nucleic Acids Res.* **26**, 1906–1926.
Egli, M., Williams, L. D., Gao, Q. & Rich, A. (1991). *Biochemistry*, **30**, 11388–11402.
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. & Ho, P. S. (1989). *J. Biol. Chem.* **264**, 7921–7935.
Gessner, R. V., Quigley, G. J. & Egli, M. (1994). *J. Mol. Biol.* **239**, 1154–1168.
Harper, N. A., Brannigan, J. A., Buck, M., Lewis, R. J., Moore, M. H. & Schneider, B. (1998). *Acta Cryst.* D**54**, 1273–1284.
Ho, P. S. & Mooers, B. H. M. (1997). *Biopolymers*, **44**, 65–90.
Karthe, P. & Gautham, N. (1998). *Acta Cryst.* D**54**, 501–509.
Kumar, D. V. & Weber, T. I. (1993). *Nucleic Acids Res.* **21**, 2201–2208.
Lavery, R. & Sklenar, H. (1988). *J. Biomol. Struct. Dyn.* **6**, 63–91.

Malinina, L., Tereshko, V., Ivanova, V., Subirana, J. A., Zarytova, V. & Nekrasov, Y. (1998). *Biophys. J.* **74**, 2482–1490.
Mooers, B. H., Schroth, G. P., Baxter, W. W. & Ho, P. S. (1995). *J. Mol. Biol.* **249**, 772–784.
Moore, M. H., van Meervelt, L., Salisbury, S. A., Lin, P. K. T. & Brown, D. M. (1995). *J. Mol. Biol.* **251**, 665–673.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Parkinson, G. N., Arvanitis, G. M., Lessinger, L., Ginell, S. L., Jones, R., Gaffney, B. & Berman, H. M. (1995). *Biochemistry*, **34**, 15487–15495.
Ramakrishnan, B. & Viswamitra, M. A. (1988). *J. Biomol. Struct. Dyn.* **6**, 511–523.
Sadasivan, C. & Gautham, N. (1995). *J. Mol. Biol.* **248**, 918–930.
Thiyagarajan, S., Rajan, S. S. & Gautham, N. (2004). *Nucleic Acids Res.* **32**, 5945–5953.
Thiyagarajan, S., Satheesh Kumar, P., Rajan, S. S. & Gautham, N. (2002). *Acta Cryst.* D**58**, 1381–1384.
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.
Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., van Boom, J. H., van der Marel, G. A. & Rich, A. (1981). *Science*, **211**, 171–176.
Wang, Y., Thomas, G. A. & Peticolas, W. L. (1987). *Biochemistry*, **26**, 5178–5186.