Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of d(GCGAAGC) and d(GCGAAAGC) (tetragonal form): a switching of partners of the sheared G·A pairs to form a functional G·A×A·G crossing

Tomoko Sunami,[a] Jiro Kondo,[a] Ichiro Hirao,[b] Kimitsuna Watanabe,[c] Kin-ichiro Miura[d] and Akio Takénaka[a]*

[a]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan, [b]RIKEN GSC, Wako-shi, Saitama 351-0198, Japan, [c]Graduate School of Engineering, University of Tokyo, Tokyo 113-8656, Japan, and [d]Faculty of Science, Gakushuin University, Tokyo 171-8588, Japan

Correspondence e-mail: atakenak@bio.titech.ac.jp

Received 24 July 2003
Accepted 9 December 2003

PDB References:
d(GCGAAGC), 1ub8, r1ub8sf; d(GCGAAAGC), 1ue4, r1ue4sf.

The DNA fragments d(GCGAAGC) and d(GCGAAAGC) are known to exhibit several extraordinary properties. Their crystal structures have been determined at 1.6 and 1.65 Å resolution, respectively. Two heptamers aligned in an antiparallel fashion associate to form a duplex having molecular twofold symmetry. In the crystallographic asymmetric unit, there are three structurally identical duplexes. At both ends of each duplex, two Watson–Crick G·C pairs constitute the stem regions. In the central part, two sheared G·A pairs are crossed and stacked on each other, so that the stacked two guanine bases of the G·A×A·G crossing expose their Watson–Crick and major-groove sites into solvent, suggesting a functional role. The adenine moieties of the A_5 residues are inside the duplex, wedged between the A_4 and G_6 residues, but there are no partners for interactions. To close the open space on the counter strand, the duplex is strongly bent. In the asymmetric unit of the d(GCGAAAGC) crystal (tetragonal form), there is only one octamer chain. However, the two chains related by the crystallographic twofold symmetry associate to form an antiparallel duplex, similar to the base-intercalated duplex found in the hexagonal crystal form of the octamer. It is interesting to note that the significant difference between the present bulge-in structure of d(GCGAAGC) and the base-intercalated duplex of d(GCGAAAGC) can be ascribed to a switching of partners of the sheared G·A pairs.

1. Introduction

Life was established and has evolved on the basis of the self-complementary structure of DNA, proposed 50 years ago by Watson & Crick (1953). It is a highly sophisticated medium for the storage of genetic information. Many structural studies have been reported. An alternative conformation is the A-form (Fuller *et al.*, 1965), which is predominantly observed in RNA structures, although DNA can also have this conformation at low humidity (Leslie *et al.*, 1980). Another discovery was left-handed Z-DNA (Wang *et al.*, 1979; Drew *et al.*, 1980), which is involved in some biological processes (Brown & Rich, 2001). In addition to these duplexes, several kinds of DNA multiplexes [DNA triplexes (Soyfer & Potaman, 1996), a quadruplex with G-quartets (Keniry, 2000), an *i*-motif with $C{\cdot}C^{+}$ pairs (Chen *et al.*, 1994; Snoussi *et al.*, 2001) and a parallel duplex with homo base pairs (Sunami *et al.*, 2002)] have been reported. On the other hand, RNA generally exists as a single strand but is folded to form a complicated three-dimensional structure so that it serves as a functional molecule similar to proteins. Typical examples are found in ribosomal RNAs (Harms *et al.*, 2001; Ban *et al.*, 2000; Yusupov *et al.*, 2001; Wimberly *et al.*, 2000), hammerhead ribozymes (Pley *et al.*, 1994; Scott *et al.*, 1995), hairpin ribozyme (Rupert & Ferre-

Table 1
Crystal data and statistics of data collection of the 7hmt and 8hmt-t crystals.

Crystal sample	7hmt d(GCGAAGC)					8hmt-t d(GCGAAAGC)
X-ray source	Photon Factory					Photon Factory
Space group	$P2_12_12_1$					$I422$
Unit cell (Å)	a = 48.7, b = 48.9, c = 63.8					$a = b$ = 36.9, c = 64.3
Z†	6					1
Wavelength (Å)	1.00 (remote)	1.6046 (peak)	1.6053 (edge)	1.609 (remote)	1.00	1.00
Resolution (Å)	34–1.8	30–2.0	30–2.0	30–2.0	24–1.6	26–1.6
Measured reflections	99782	77754	77593	77619	328413	88780
Unique reflections	13155	10007	10009	9968	20709	3175
Completeness (%)	91.9	95.1	95.0	94.7	100	99.9
Completeness in outer shell (%)	63.9 (1.80–1.90 Å)	73.0 (2.11–2.0 Å)	72.7 (2.11–2.0 Å)	70.6 (2.11–2.0 Å)	100 (1.69–1.6 Å)	100 (1.69–1.60 Å)
$R_{\rm merge}$‡ (%)	3.4	5.5	5.0	4.5	6.7	4.2
$R_{\rm anom}$§ (%)	3.4	7.4	4.8	2.5	—	—
Theoretical f'/f''	0.16/1.78	−5.85/3.92	−6.09/3.92	−7.57/0.47	—	—
Refined f'/f''	−0.607/1.442	−7.948/4.514	−9.298/2.596	−6.731/0.567	—	—

† Number of single strands in the asymmetric unit. ‡ $R_{\rm merge} = 100\times\sum_{hklj} |I_{hklj} - \langle I_{hkl}\rangle|/\sum_{hklj}\langle I_{hkl}\rangle$. § $R_{\rm anom} = 100\times\sum_{hklj} |I_{hklj}(+) - I_{hklj}(-)|/\sum_{hklj}[I_{hklj}(+) + I_{hklj}(-)]$.

D'Amaré, 2001), group I intron ribozymes (Cate *et al.*, 1996) and so on, in which the stem parts formed by Watson–Crick base pairs are folded into a globular form by tertiary interactions involving loops and bulges. The two molecules DNA and RNA are essential to living systems. Recent *in vitro* selection techniques, however, have made it possible to create a functional DNA (Breaker & Joyce, 1994) similar to ribozymes. Therefore, it is expected that natural DNA can also have a specific function, like RNA, when it exists as a single-stranded molecule. Our structural knowledge of DNA, beyond the helical structures containing Watson–Crick complementary base pairs, is limited to the several examples described above. To extend and establish the structural basis for understanding the mechanism of functional DNA and to design new functional DNA, it is necessary to reveal further additional structural motifs of DNA that contain non-complementary bases.

We have found that DNA fragments containing the sequences d(GCGAAGC) and d(GCGAAAGC) exhibit extraordinary properties with (i) abnormal mobility in electrophoresis (Hirao *et al.*, 1988), (ii) high thermostability (Hirao *et al.*, 1989), (iii) unusual CD spectra (Hirao *et al.*, 1989) and (iv) robustness against nuclease digestion (Hirao *et al.*, 1992). From these properties, mini-hairpin structures have been postulated (Hirao *et al.*, 1994; Yoshizawa *et al.*, 1997; Padrta *et al.*, 2002). The sequence d(GCGAAGC) was found in several genes (Arai *et al.*, 1981; Cowing *et al.*, 1985) and several efforts to design drugs to target this sequence have already been reported (Veselkov *et al.*, 2002; Williams *et al.*, 2002; Samani *et al.*, 2001). X-ray analyses of oligonucleotides containing the sequences d(GCGAAGC), d(GCGAAAGC) and d(GCGAAAGCT) were initiated in order to reveal their detailed structures. In the crystal of the latter nonamer (Sunami *et al.*, 2002), the four residues with sequence d(CGAA) form a parallel duplex with those of another nonamer through homo base-pair formation and the remaining four residues d(AGCT) form an antiparallel duplex with those extending from another parallel duplex. In the case of the octamer, two crystal forms were found. In the hexagonal form (hereafter referred to as 8hmt-h), it was found that the two strands of the octamer form a base-intercalated duplex; the central four adenine residues are intercalated and stack on each other between the two strands, and the three duplexes are bundled around two hexamminecobalt cations (Sunami *et al.*, 2004). These structures are quite different from the mini-hairpin structure, suggesting structural versatility of the specific sequence. Another crystal (tetragonal form, hereafter referred to as 8hmt-t) was obtained in the absence of hexamminecobalt cations. It is important to examine ionic effects on the structures by X-ray analysis. In addition, the heptamer d(GCGAAGC) (hereafter referred to as 7hmt) also has a high thermal stability, so that the $T_{\rm m}$ value is the same as that of the octamer (349.5 K; Hirao *et al.*, 1994). In the present study, X-ray analyses of the two DNA fragments 7hmt and 8hmt-t were performed and it was found that a large structural difference occurs at the central adenine residues of 7hmt, which form a G·A×A·G crossing, which may represent a new functional motif.

2. Materials and methods

2.1. Synthesis and crystallization

DNA oligomers with sequences d(GCGAAGC) (7hmt) and d(GCGAAAGC) (8hmt) were synthesized on a DNA synthesizer, purified by HPLC and salts were removed by gel-filtration. Crystallization conditions were surveyed by the hanging-drop vapour-diffusion method at 277 K. A droplet was prepared by mixing 2 µl of 1.5 m*M* 7hmt solution and 2 µl of reservoir solution pH 6.0 containing 20 m*M* NaCl, 140 m*M* KCl, 20 m*M* $Co(NH_3)_6Cl_3$, 18% 2-methyl-2,4-pentanediol and 0.17% *n*-decanoyl-*N*-methylglucamide (purchased from Dojindo Laboratories Co. Ltd) in 40 m*M* sodium cacodylate buffer; crystals of 7hmt that were 300 × 250 µm in size were obtained within 10 d. After macroseeding, crystals grew further to 350 × 270 µm in size over 10 d. Some of them were mounted in nylon cryoloops (Hampton Research) with

reservoir solution containing 30%(*v*/*v*) 2-methyl-2,4-pentanediol as a cryoprotectant and stored in liquid nitrogen.

A new crystal form 8hmt-t (maximum dimensions 300 × 200 µm) was obtained within one week in the absence of hexamminecobalt cations, with droplets prepared by mixing 2 µl of 3 m*M* 8hmt solution and 2 µl crystallization solution containing 40 m*M* sodium cacodylate pH 6.0, 12 m*M* spermine.4HCl, 80 m*M* NaCl, 20 m*M* $MgCl_2$ and 10% MPD equilibrated against 25% MPD. Some of them were mounted in nylon cryoloops and stored in liquid nitrogen.

2.2. Data collection

For phasing using the anomalous effect of the Co atoms, four X-ray data sets were taken from the 7hmt crystal at different wavelengths (λ = 1.00, 1.6046, 1.6053 and 1.609 Å from XAFS measurement) at BL18b, Photon Factory, Japan. A crystal specimen of 7hmt was cooled to 100 K and X-ray diffraction was recorded on a CCD detector (Quantum 4R) positioned at 175 mm from the crystal for λ = 1.00 Å and at 100 mm for λ = 1.6 Å. Diffraction patterns taken with 3° oscillation and 90 and 120 s exposure per frame over a total range of 180° were processed at 1.8 and 2.0 Å resolution using the program *DPS*/*MOSFLM* (Leslie, 1992; Steller *et al.*, 1997; Rossmann & van Beek, 1999; Powell, 1999). The four data sets were scaled separately using the programs *SCALA*, *TRUNCATE* and *SCALEIT* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

For structure refinement at higher resolution, a further data set was collected at 100 K with λ = 1.00 Å at the same beamline, using a different crystal positioned at 125 mm from a CCD detector. Frames were taken at 2° oscillation intervals with 90 s exposure over a total range of 180°. To compensate for overloaded reflections, another data set was taken with 5° oscillation and 60 s exposure. The two data sets (125 frames in total) were separately processed at 1.6 Å resolution with the same programs as described above. 20 709 unique reflections with a 100% completeness were obtained with $R_{\rm merge}$ = 6.7%.

X-ray data were also obtained from an 8hmt-t crystal at the same beamline (λ = 1.00 Å). X-ray diffraction data were collected from a crystal specimen cooled to 100 K using a CCD detector positioned at 120 mm. Diffraction patterns were recorded at 1.5° oscillation intervals for 30 and 10 s exposures over a total range of 180° and were processed at 1.6 Å resolution using the programs described above. 3175 unique reflections with 99.9% completeness were obtained with $R_{\rm merge}$ = 4.2%. Statistics of data collection and crystal data for 7hmt and 8hmt-t are summarized in Table 1.

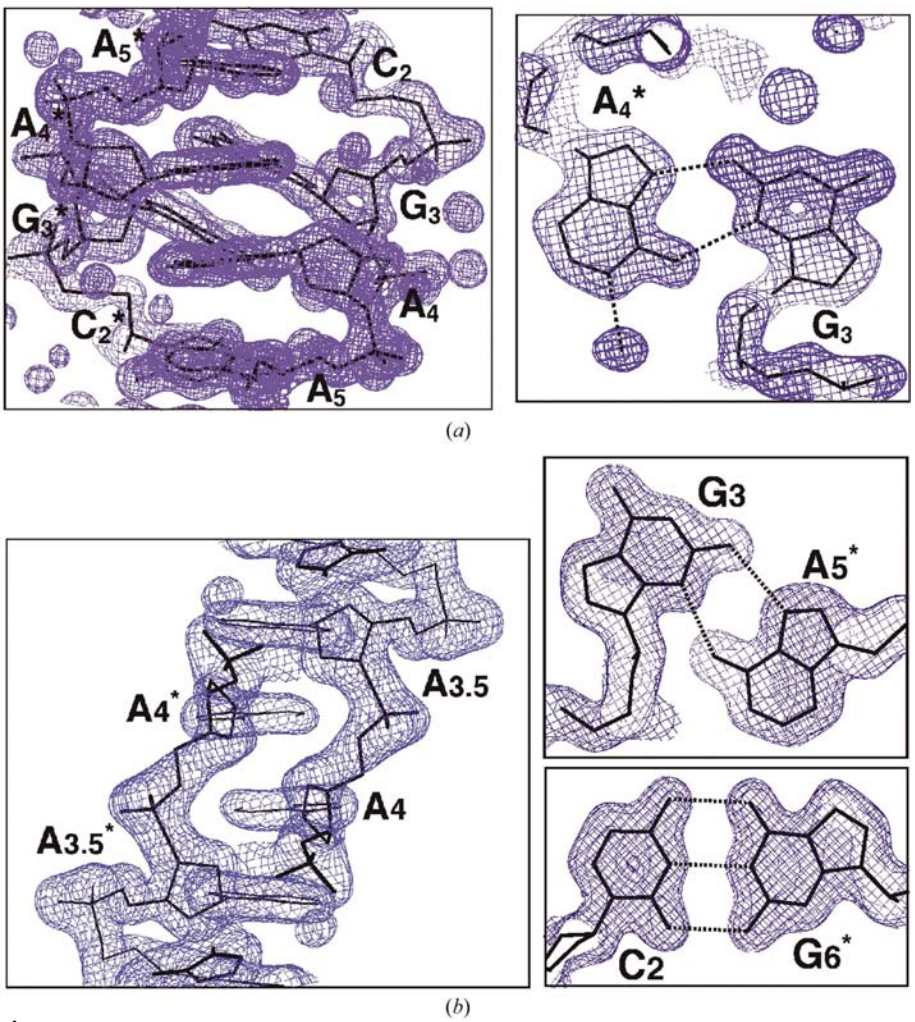

Figure 1
Local $2F_o - F_c$ maps for 7hmt (*a*) and for 8hmt-t (*b*). Densities are contoured at the 1σ level. Broken lines indicate possible hydrogen bonds.

2.3. Structure determination and refinement

The number of 7hmt oligonucleotides in the asymmetric unit was estimated to be six, based on a calibration curve for nucleic acid crystals (Takénaka *et al.*, 1995). The initial crystal structure was determined by the MAD method using the anomalous effect of Co atoms with the program *SOLVE* (Terwilliger, 2002). Five Co atoms were found in the asymmetric unit. Although the figure of merit was 0.65, the resultant electron-density map modified by the solvent-flattening technique (solvent content 55.4%) with the program *CNS* (Brünger *et al.*, 1998) clearly indicated the phosphate–ribose backbones and the individual bases of the six oligomers. The molecular structures were easily constructed on a graphics workstation with the program *QUANTA* (Molecular Simulation Inc.).

Table 2
Statistics of structure refinements of the 7hmt and 8hmt-t crystals.

Crystals	7hmt	8hmt-t
Resolution range (Å)	24–1.6	15–1.65
Reflections used ($F_o > 3\sigma$)	20173	2792
R factor† (%)	19.4	20.9
R_{free}‡ (%)	23.2	23.6
No. DNA atoms	858	164
No. waters	304	66
No. ammonias	54	0
No. cations	9 Co	1 Mg
R.m.s.d. from ideal geometry		
Bond lengths (Å)	0.003	0.004
Bond angles (°)	0.7	0.9
Improper angles (°)	1.1	1.1

† R factor = $100 \times \sum \big||F_o| - |F_c|\big| / \sum |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively. ‡ Calculated using a random set containing 10% of observations that were not included during refinement (Brünger *et al.*, 1998).

The atomic parameters of the six 7hmt oligomers were refined with the program *CNS* (Brünger *et al.*, 1998) through a combination of rigid-body, simulated-annealing and crystallographic conjugate-gradient minimization refinements and *B*-factor refinements, followed by interpretation of the omit map at every nucleotide residue. No restraints were applied between paired nucleotides and sugar puckering. Four additional hexamminecobalt cations and 304 water molecules in total were found on an $F_o - F_c$ map after several steps of refinement and were included in the later refinements. Fig. 1(*a*) shows the final local $2F_o - F_c$ maps for the stacked two ends of the duplexes and the central residues.

The 8hmt-t crystal has unit-cell parameters similar to those of the 8hmt-h crystal (Sunami *et al.*, 2004), but the space group was quite different (8hmt-h: space group $P6_322$, $a = b = 37.4$, $c = 64.6$ Å; 8hmt-t: space group $I422$, $a = b = 36.9$, $c = 64.3$ Å). As a first trial, it was reasonable to assume a duplex structure similar to that of 8hmt-h. Application of molecular replacement with the atomic coordinates of the 8hmt-h crystal gave a unique significant solution using the program *AMoRe* (Navaza, 1994). The atomic parameters were refined following a procedure similar to that for the 7hmt crystal using the program *CNS*. Fig. 1(*b*) shows the final local $2F_o - F_c$ maps for the central part of the duplex and the $G_3{\cdot}A_5^*$ and $C_2{\cdot}G_6^*$ base pairs.

Statistics of structure refinement of the 7hmt and 8hmt-t crystals are summarized in Table 2. All local helical parameters including torsion angles and pseudorotation phase angles of ribose rings were calculated using the program *3DNA* (Lu & Olson, 2003). Fig. 1 was drawn with the program *O* (Jones *et al.*, 1991) and Figs. 2–10 with the program *RASMOL* (Sayle & Milner-White, 1995).

3. Results

3.1. Strand association of d(GCGAAGC)

There are three independent duplexes formed between chains *A* and *B*, between chains *C* and *D* and between chains *E* and *F* in the asymmetric unit. They are very similar to each other, with r.m.s. deviations of 0.4–0.6 Å, as shown in Fig. 2. In each duplex, two heptamers are aligned in an antiparallel fashion, associated with each other through base pairs. The individual duplex also has an approximate molecular twofold symmetry at the centre, within 0.6–0.8 Å r.m.s. deviation. Therefore, only one of the three structures is described in detail below, the others being similar.

As shown in Fig. 3, two Watson–Crick G·C base pairs are formed at each end of the duplex, ($G_1{\cdot}C_7^*$ and $C_2{\cdot}G_6^*$) or ($G_6{\cdot}C_2^*$ and $C_7{\cdot}G_1^*$), where * indicates the counter-chain. In the central part, two non-Watson–Crick sheared G·A base pairs are formed between G_3 and A_4^* and between A_4 and G_3^*,

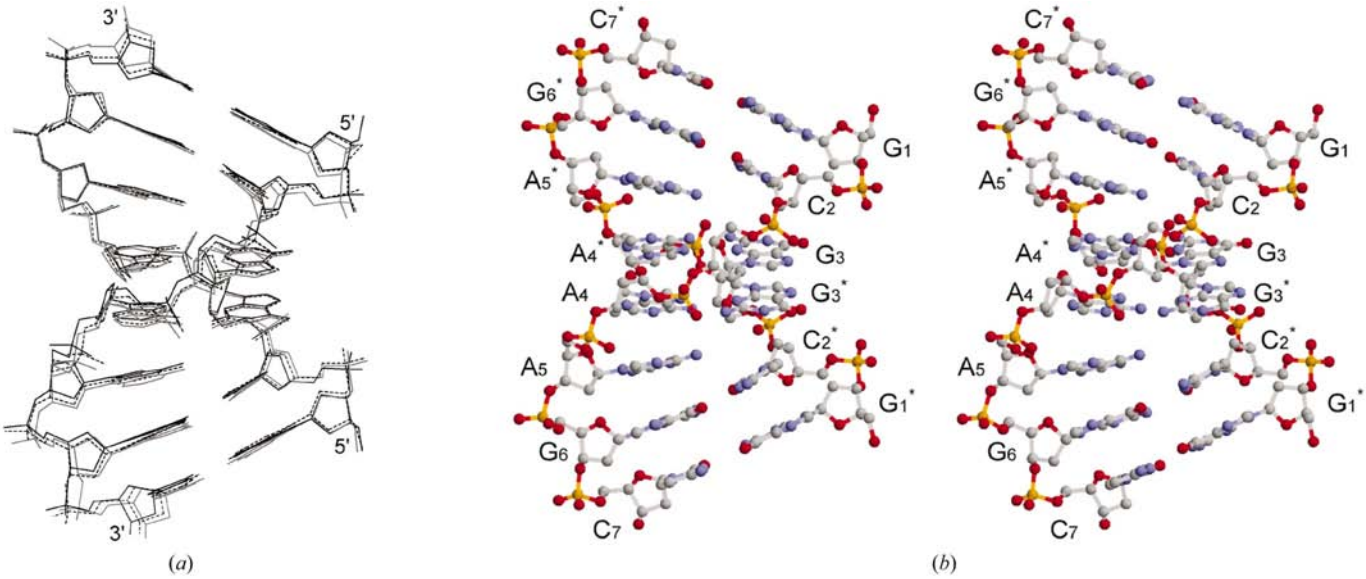

Figure 2
A superimposition of the three independent bulged-in duplex structures of 7hmt, *AB, CD* and *EF* (*a*) and a stereo-pair diagram of the *AB* duplex (*b*). *AB* is drawn in black, *CD* in grey and *EF* in dotted lines.

through two hydrogen bonds, N2H···N7 and N3···HN6. These two G·A pairs are stacked on each other. The remaining two A_5 residues do not participate in any pair formations. They are not bulged-out but stay within the duplex, sandwiched between the A_4 and G_6 residues. Because of these two bulged-in residues, the duplex is strongly curved.

3.2. G·A×A·G crossing

The characteristic feature of the 7hmt duplex is the consecutive sheared G·A and A·G pairs in the central part. The two chains cross at this point so that the two guanine bases are stacked, as well as the two adenine bases, as shown in Fig. 4. By this G·A×A·G crossing, the Watson–Crick sites and the major-groove sites of the G_3 and G_3^* residues are exposed into the surrounding solvent, as well as the minor-groove sites of A_4 and A_4^*. Based on this crossing, there are two base-stacked columns: a short column G_1–C_2–G_3–G_3^*–C_2^*–G_1^* and a long column C_7–G_6–A_5–A_4–A_4^*–A_5^*–G_6^*–C_7^*. The exposed bases are covered with water molecules and hexamminecobalt cations. This G·A×A·G crossing may be useful to exchange the chains, because in the standard double helix of DNA, base stacking occurs in the same chain.

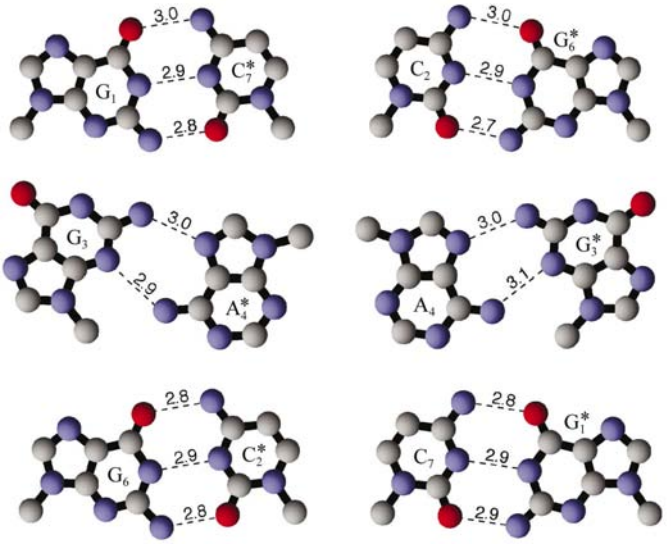

Figure 3
Base-pair formations stabilizing the bulged-in duplex found in the 7hmt crystal. The geometry is the same in the three duplexes.

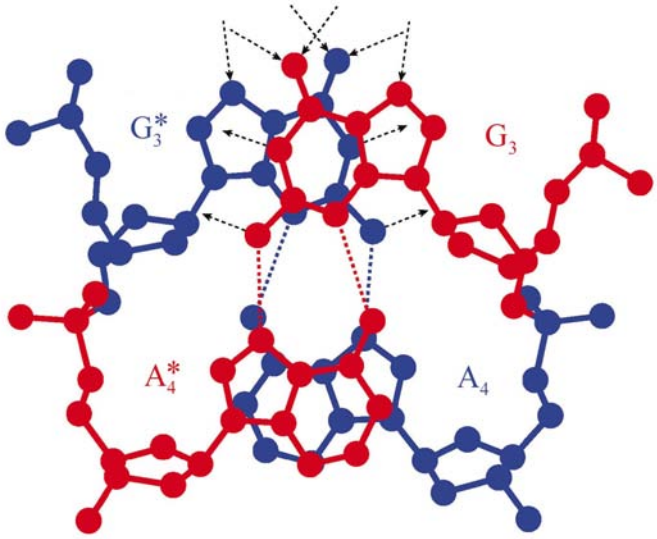

Figure 4
The G·A×A·G crossing viewed down the local helical axis. The two guanine bases are stacked; dotted lines with arrows indicate possible interactions.

3.3. Conformation to stabilize the specific structure

The local helical parameters and the sugar puckers are given in Tables 3 and 4. The two residues at each end of the duplex adopt the canonical B-form conformation to form stems. To form the sheared pairs in the central part, the two C1′···C1′ distances become shorter (8.7–9.1 Å) and the buckle angles larger (33–39°), as the pair formation occurs between the minor-groove site of the guanine base and the major-groove site of the adenine base. The ribose rings of the G_3 and G_3^* residues adopt a C2′-*endo* pucker (C3′-*exo* is close to C2′-*endo*) and the A_4 and A_4^* residues adopt a C3′-*endo* pucker. The other notable features are small twist angles (−3 to 7°) from G_3·A_4^* to A_4·G_3^*, the pairs forming the G·A×A·G crossing. The C3′-*endo* conformation of A_4 makes space for the bulged A_5 residue to stay within the duplex. This puckering occurs for accepting an intercalater in general. The structural features described above are summarized in Fig. 5.

3.4. Roles of metal ions and crystal packing

The three independent duplexes are respectively stacked at the ends of those related by the crystallographic 2_1 symmetry to form three long columns, as shown in Fig. 6. One (*AB*) is along the *b* axis and the other two (*CD* and *EF*) are along the *a*

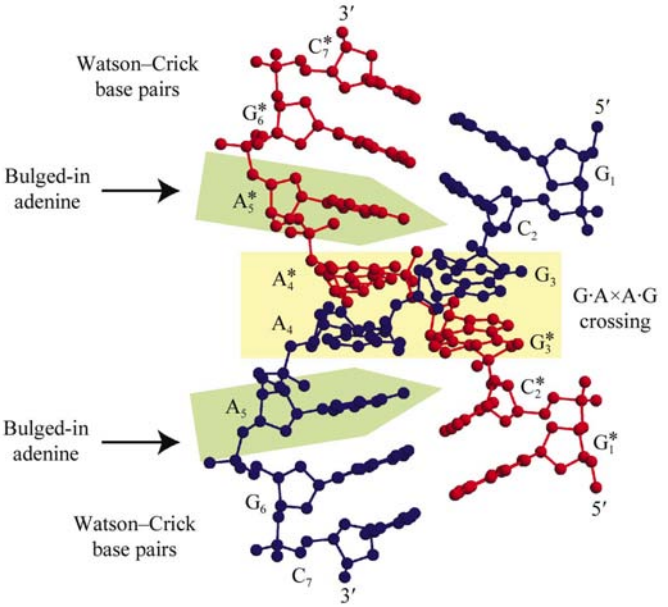

Figure 5
Structural features of the bulged-in duplex of 7hmt.

axis, the latter two columns being separated by 1/2 along the *c* direction.

Fig. 7 shows three hexamminecobalt cations bound in the major groove of each duplex. One is bound to both G_3 and G_3^* of the G·A×A·G crossing and the remaining two cations are bound to G_6 and G_6^*, respectively. As seen in Fig. 7, these cations always form hydrogen bonds between the coordinated ammonia groups and the O6 and N7 atoms of guanine bases and at the same time interact with the phosphate groups of the adjacent duplexes, so that the three columns are linked together within the crystal packing.

3.5. Structure of d(GCGAAAGC) in the tetragonal crystal

The unit-cell parameters of the 8hmt-t crystal are similar to those of the 8hmt-h crystal (Sunami *et al.*, 2004; $a = b = 37.4$, $c = 64.6$ Å for 8hmt-h; $a = b = 36.9$, $c = 64.3$ Å for 8hmt-t), but the space group is quite different (*I*422 for 8hmt-t and $P6_322$ for 8hmt-h). This suggests that the difference is in the crystal packing of similar molecular units. As expected, the octamers form a base-intercalated duplex similar to that found in the 8hmt-h crystal, as shown in Fig. 8 (refer to the nucleotide numbers[1]). The two chains related by crystallographic twofold symmetry are associated to form a duplex. At both ends of the duplex, the two consecutive G·C pairs form the stem parts. These are similar to the stem formations in the 7hmt duplexes. The third A_5 and A_5^* residues, however, form a sheared G·A pair with the G_3^* and G_3 residues of the counter-strand through the two hydrogen bonds, N6H(A_5)···N3(G_3^*) and N7(A_5)···HN2(G_3^*). In contrast, for the 7hmt duplexes described above, this type of sheared pair occurs at the central residues of the duplex. Here, the central $A_{3.5}$ and A_4 residues are not involved in base-pair formation. Their base moieties are respectively stacked on those of the other strand so that A_4 is intercalated between $A_{3.5}^*$ and A_4^*, and A_4^* is intercalated between $A_{3.5}$ and A_4. These four adenine bases, $A_{3.5}$, A_4^*, A_4 and $A_{3.5}^*$, expose their Watson–Crick sites[2] into the solvent

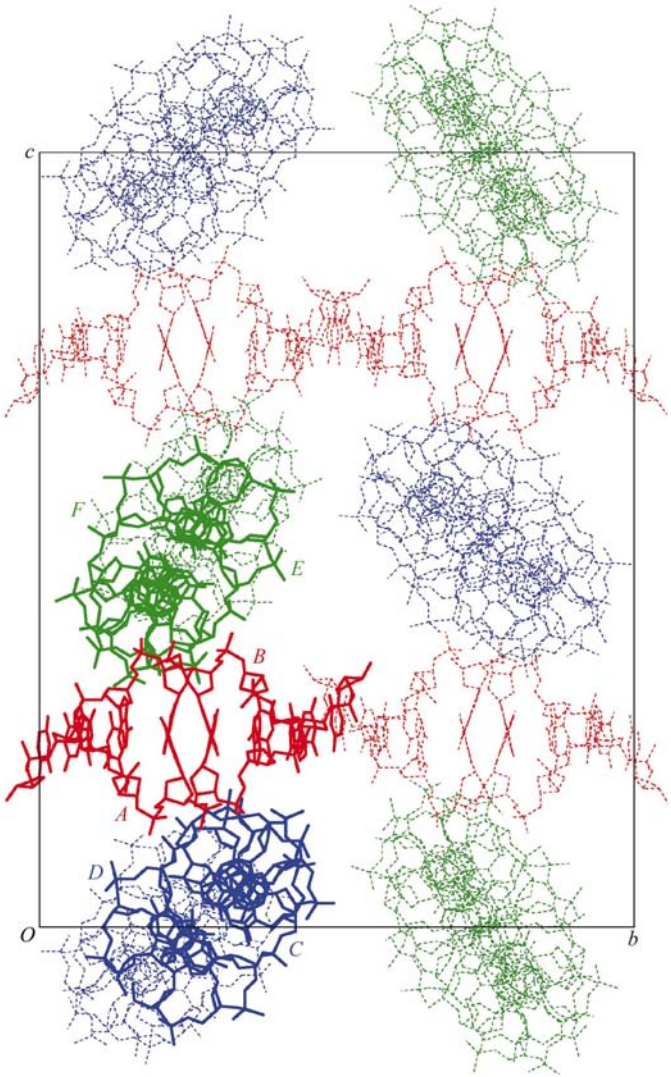

Figure 6
A packing diagram of the 7hmt crystal, viewed down the *a* axis. The three duplexes in the asymmetric unit are shown with full lines (*AB* duplex in red, *CD* duplex in blue and *EF* duplex in green) and crystallographically related duplexes are shown with broken lines.

[1] To compare the two structures of 7hmt and 8hmt-t, the nucleotide-numbering system for 8hmt-t is different from that in 8hmt-h; *e.g.* $A_{3.5}$ is A_4 in 8hmt-h, A_4 is A_5 in 8hmt-h and so on.

[2] The donor and acceptor sites for the hydrogen bonds that form the Watson–Crick base pairs.

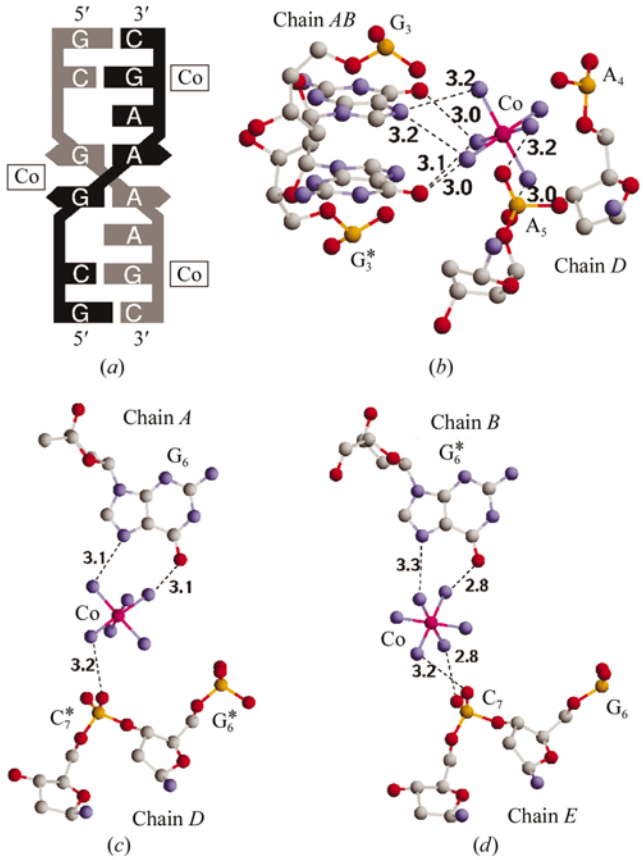

Figure 7
The three hexamminecobalt cations bound at the G·A×A·G crossing (*a* and *b*), between *AB* and *CD* (*c*) and between *AB* and *EF* (*d*). These binding sites are commonly found in the other duplexes.

Table 3
Local helical parameters of the 7hmt duplexes.

Inclin, Prop, Buckl and Open are the inclination, propeller twist, buckle and opening angles (see Lu & Olson, 2003).

Base pair	Chains	Prop (°)	Buckl (°)	Open (°)	C1′···C1′ (Å)	Chains	Inclin (°)	Tip (°)	Twist (°)	Rise (Å)
G_1·C_7^*	*AB*	−2	7	1	10.7					
	CD	−1	−4	−4	10.9					
	EF	−2	−4	0	10.7					
						AB	−2	−4	31	3.3
						CD	6	−1	28	3.0
						EF	3	0	29	3.0
C_2·G_6^*	*AB*	−3	10	−1	10.6					
	CD	−3	15	1	10.5					
	EF	−1	11	1	10.6					
						AB	19	8	85	5.4
						CD	22	11	79	5.1
						EF	18	11	78	5.0
G_3·A_4^*	*AB*	−6	34	−1	8.9					
	CD	−2	38	−3	8.9					
	EF	0	37	−1	8.9					
						AB	3	−5	−7	4.7
						CD	−16	7	−6	5.0
						EF	−14	−5	−3	5.0
A_4·G_3^*	*AB*	2	−33	−2	9.1					
	CD	−1	−39	−1	8.7					
	EF	−1	−37	−3	9.0					
						AB	19	−11	75	5.3
						CD	19	−10	79	5.2
						EF	18	−12	78	5.0
G_6·C_2^*	*AB*	−1	−12	2	10.5					
	CD	−1	−16	5	10.5					
	EF	−3	−6	0	10.6					
						AB	−4	4	32	3.3
						CD	−2	6	35	3.3
						EF	−6	1	33	3.4
C_7·G_1^*	*AB*	−1	−6	1	10.7					
	CD	0	−9	1	10.7					
	EF	1	−6	2	10.6					
A-form		12	0	−2	10.7	A-form	20	0	33	2.3
B-form		−1	0	−2	10.7	B-form	−5	0	36	3.4

region in order to interact with the surrounding water molecules. There are three base-stacked columns, a long column G_1–C_2–G_3–$A_{3.5}$–A_4^*–A_4–$A_{3.5}^*$–G_3^*–C_2^*–G_1^* and two short columns, A_5–G_6–C_7 and A_5^*–G_6^*–C_7^*, in the duplex. These features are very similar to those of the base-intercalated duplex found in the 8hmt-h crystal (Sunami *et al.*, 2004), with an r.m.s. deviation of 0.3 Å.

3.6. Helix conformation

The local helical parameters and the sugar puckers are given in Table 5. To form sheared pairs at the third residues (G_3·A_5^* and G_3^*·A_5), the C1′···C1′ distances become shorter (8.4 Å) and the buckle angles larger (30°). However, the ribose rings of the two residues retain a C2′-*endo* conformation. These G·A pairs form a large platform stabilizing the base-intercalated duplex. In the base-stacked column, the ribose rings of the $A_{3.5}$ residues adopt a C3′-*endo* pucker to make space for accepting an A_4 base of the counter-strand between the $A_{3.5}$ and A_4 residues.

Table 4
Sugar puckers of the 7hmt duplex.

Residue	Chain *A*	Chain *B*	Chain *C*	Chain *D*	Chain *E*	Chain *F*
G_1	C3′-*exo*	C3′-*exo*	C3′-*exo*	C3′-*exo*	C3′-*exo*	C3′-*exo*
C_2	C1′-*exo*	C1′-*exo*	C1′-*exo*	C2′-*endo*	C1′-*exo*	C2′-*endo*
G_3	C3′-*exo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*
A_4	C4′-*exdo*	C3′-*endo*	C3′-*endo*	C3′-*endo*	C3′-*endo*	C3′-*endo*
A_5	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*
G_6	C1′-*exo*	C1′-*exo*	C1′-*exo*	C4′-*exo*	O4′-*endo*	C1′-*exo*
C_7	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*
A-form	C3′-*endo*					
B-form	C2′-*endo*					

3.7. Long columns of duplexes and tetra-assembly of the base-intercalated duplexes

In the 8hmt-h crystal, three columns of base-intercalated duplexes are assembled around the hexamminecobalt cations which are located on a threefold axis, two cations being bound to the G_3 and G_3^* residues through two hydrogen bonds between the O6 atom and the coordinated ammonia and between the N7 atom and the coordinated ammonia. In the present crystal 8hmt-t, however, four columns of base-intercalated duplexes are assembled around the crystallographic fourfold axis. The central four residues, $A_{3.5}$, A_4^*, A_4 and $A_{3.5}^*$, point towards the central axis, but there are no ions around the axis. Instead, several water molecules, some of them disordered, fill the axial space to form water-mediated interactions. A hexahydrated magnesium cation is found between the two adjacent duplexes. To link the four columns, the two coordinated water molecules are hydrogen bonded to the O6 and N7 atoms of the G_6^* residue and another water molecule interacts with the phosphate groups of the G_3 residues of the adjacent duplex, as shown in Fig. 9. The two tetrameric assemblies related by the crystallographic body-centred symmetry are stacked between the two G_1·C_7^* and C_7·G_1^* base pairs. Several water molecules are also found between the tetrameric assemblies, laterally aligned.

4. Discussion

The present study shows that d(GCGAAGC) does not form a mini-hairpin structure in the crystalline state, despite many solution studies suggesting a mini-hairpin. Similar results are also obtained for the octamer d(GCGAAAGC). The structure of 8hmt-t is a base-intercalated duplex even in the absence of hexamminecobalt cation. Potassium, sodium and magnesium ions used in crystallization will have some effect, as discussed in the previous paper (Sunami *et al.*, 2004). In addition, the crystalline state may facilitate duplex formation, because the two types of duplexes in the present crystal structures form long columns by stacking at the ends of other duplexes and the column–column interactions may be a strong driving force to stabilize the crystal lattices.

The bulged-in duplex itself is a new structure. From a comparison with the base-intercalated duplex, it is interesting to note that the addition of an adenine residue in the centre of

Table 5
Local helical parameters and sugar puckers of the 8hmt-t duplex.

Only half of the helical parameters are given owing to the crystallographic twofold symmetry. The corresponding values for 8hmt-h are given in parentheses (Sunami *et al.*, 2004). Inclin, Prop, Buckl and Open are the inclination, propeller twist, buckle and opening angles (see Lu & Olson, 2003).

Base pair	Inclin (°)	Tip (°)	Twist (°)	Rise (Å)	Prop (°)	Buckl (°)	Open (°)	C1′···C1′ (Å)	Residue	Pucker	8hmt-h
									G_1	C3′-*exo*	C3′-*exo*
G_1·C_7^*					3 (9)	−3 (0)	−2 (3)	10.8 (10.5)	C_2	C1′-*exo*	C1′-*exo*
	1 (−2)	−5 (−8)	28 (28)	3.0 (3.0)					G_3	C2′-*endo*	C2′-*endo*
C_2·G_6^*					6 (4)	13 (16)	−1 (1)	10.7 (10.5)	$A_{3.5}$	C3′-*endo*	C3′-*endo*
	9 (10)	−1 (3)	56 (52)	3.3 (3.2)					A_4	C2′-*endo*	C1′-*exo*
G_3·A_5^*					−10 (−4)	30 (31)	11 (8)	8.4 (8.4)	A_5	C2′-*endo*	C2′-*endo*
	†	†	†	†					G_6	C1′-*exo*	C1′-*exo*
									C_7	C4′-*exo*	C4′-*exo*

† $A_{3.5}$ and A_4 are not paired.

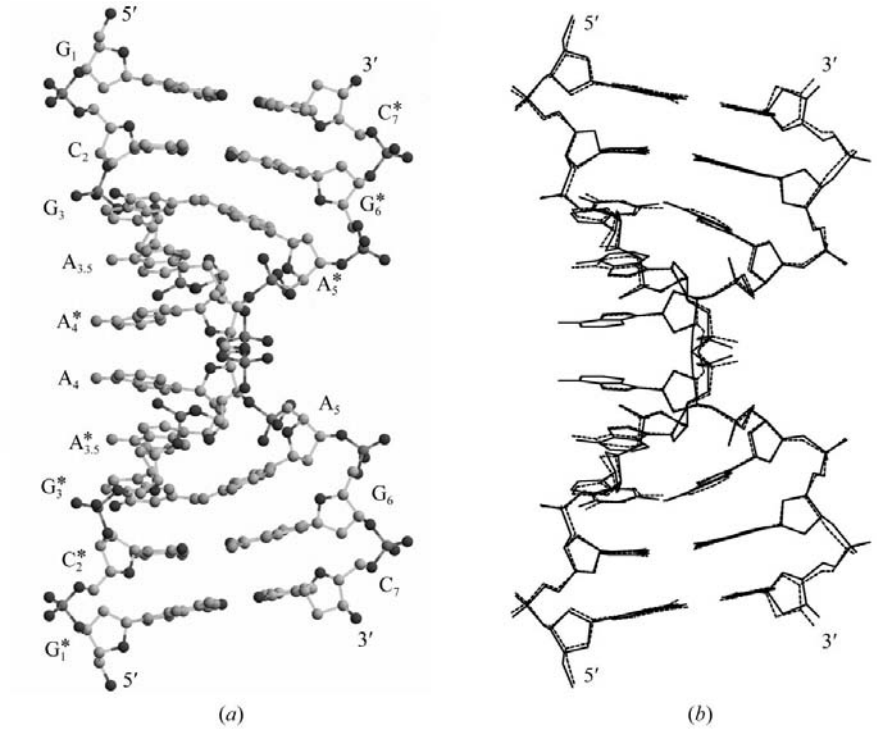

Figure 8
The base-intercalated duplex of 8hmt-t (*a*) and its superimposition on that of 8hmt-h, with an r.m.s.d. of 0.3 Å (*b*).

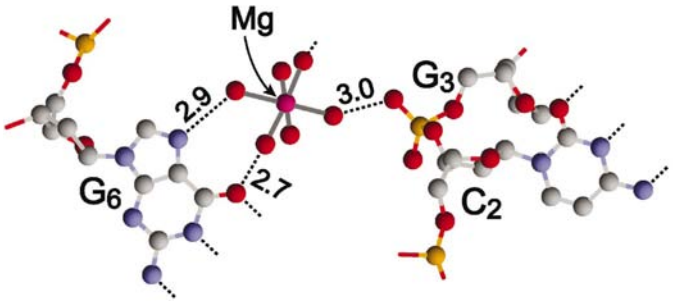

Figure 9
The hexahydrated magnesium cation bound to the major groove of the G_7 residue in 8hmt-t. The coordinated water molecule is hydrogen bonded to the phosphate group of the neighbouring duplex.

the 7hmt duplex induces drastic changes of the whole structure. Fig. 10 shows a superimposition of the two stem parts (G_1·C_7^* and C_2·G_6^* in 7hmt and G_1·C_7^* and C_2·G_6^* in 8hmt-t). In the base-intercalated duplex, the third residue G_3 forms a sheared pair with the A_5^* residue, which is the third residue from the 3′-terminus. The pairing occurs between the minor-groove site of G_3 and the major-groove site of A_5^*. Therefore, the pair itself is more wound up so that the twist angle of the sheared pair becomes larger (56°). In the bulged-in duplex, however, a similar sheared pair occurs with the A_4^* residue, which is the fourth residue from the 3′-terminus. As the A_5^* residue is not bulged out, but stays in the duplex as a bulged-in residue, the ribose–phosphate backbone allows the A_4^* residue to wind more, with a large twist angle of 75–85°. As shown in Fig. 10(*b*), insertion of $A_{3.5}$ into the heptamer would break the G_3·A_4^* pair, force the A_4^* residue to flip up and force the G_3 residue to move down to form a sheared pair with A_5^*. Thus, switching the partner of G_3 from A_4^* in 7hmt to A_5^* in 8hmt makes the central sequence change from the G·A×A·G crossing to the base-intercalated $A_{3.5}$–A_4^*–A_4–$A_{3.5}^*$ column and *vice versa*. This is an important property to keep in mind for designing functional nucleic acids, as pointed out by Chou *et al.* (2003).

The base-intercalated $A_{3.5}$–A_4^*–A_4–$A_{3.5}^*$ column looks like a functional region which can bind to other molecules, as described in the previous paper (Sunami *et al.*, 2004). The G·A×A·G crossing also looks like a functional region. Indeed, similar crossings have been found in centromere sequences (Gao *et al.*, 1999), U4-snRNA (Vidovic *et al.*, 2000), ribosomal RNA (Ban *et al.*, 2000; Wimberly *et al.*, 2000) and the hammerhead ribozyme (Pley *et al.*, 1994; Scott *et al.*, 1995). Their base–base stacking geometry is very similar to that of 7hmt. In the case of U4-snRNA, the spliceosomal 15.5 kDa protein of the human U4/U6·U5 tri-snRNA is bound to the exposed major groove and the Watson–Crick site of the two G residues.

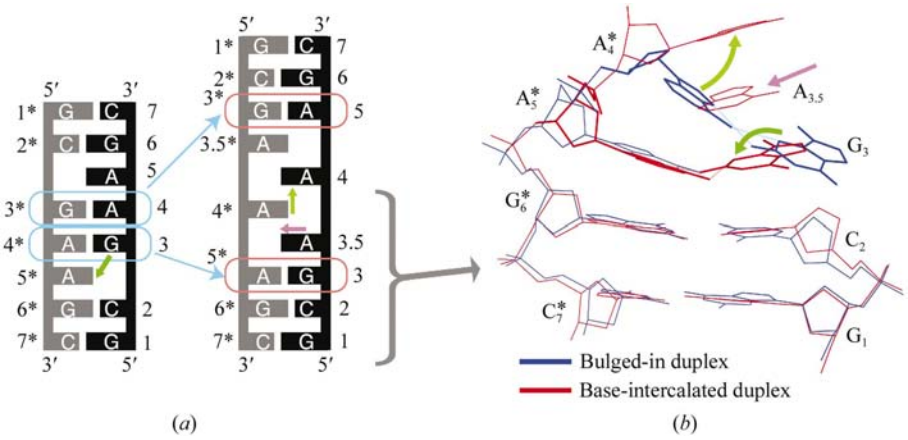

Figure 10
A schematic drawing of the two duplexes (*a*) and the switching of partners of the sheared G·A pair (*b*). An addition of $A_{3.5}$ to the heptamer induces structural changes in the octamer. The sheared $G_3{\cdot}A_4^*$ pair is broken so that A_4^* is flipped up and G_3 forms a sheared pair with A_5^*.

In general, the bulged A residue is flipped out from the duplex (Joshua-Tor *et al.*, 1992). The A_5^* residue in 7hmt is sandwiched between the A_4^* and G_6^* residues and stays within the duplex without any base–base interactions. The tight sheared-pair formation between G_3 and A_4^* may not allow A_5^* to flip out from the duplex. The bulged-in conformation will serve to bend the nucleic acid frame. When a local DNA sequence is modified by insertion of an adenine residue, as in 7hmt, or deletion of an adenine residue, as in 8hmt, a switching of partners will occur to change the functional structure.

We thank N. Sakabe, M. Suzuki, N. Igarashi and A. Nakagawa for facilities and help during data collection and T. Simonson for proofreading the original manuscript. This work was supported in part by Grants-in-Aid for Scientific Research (Nos. 12480177 and 14035217) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by the Structural Biology Sakabe Project.

References

Arai, K., Low, R., Kobori, J., Shlomai, J. & Kornberg, A. (1981). *J. Biol. Chem.* **256**, 5273–5280.

Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). *Science*, **289**, 905–920.

Breaker, R. R. & Joyce, G. F. (1994). *Chem. Biol.* **1**, 223–229.

Brown, B. A. & Rich, A. (2001). *Acta Biochim. Pol.* **48**, 295–312.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996). *Science*, **273**, 1678–1685.

Chen, L., Cai, L., Zhang, X. & Rich, A. (1994). *Biochemistry*, **33**, 13540–13546.

Chou, S. H., Chin, K. H. & Wang, A. H. (2003). *Nucleic Acids Res.* **31**, 2461–2474.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowing, D. W., Bardwell, J. C. A., Craig, E. A., Woolford, C., Hendrix, R. W. & Gross, C. A. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 2679–2683.

Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. (1980). *Nature (London)*, **286**, 567–573.

Fuller, W., Wilkins, M. H. F., Wilson, H. R. & Hamilton, L. D. (1965). *J. Mol. Biol.* **12**, 60–80.

Gao, Y. G., Robinson, H., Sanishvili, R., Joachimiak, A. & Wang, A. H. (1999). *Biochemistry*, **38**, 16452–16460.

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F. & Yonath, A. (2001). *Cell*, **107**, 679–688.

Hirao, I., Kawai, G., Yoshizawa, S., Nishimura, Y., Ishido, Y., Watanabe, K. & Miura, K. (1994). *Nucleic Acids Res.* **22**, 576–582.

Hirao, I., Naraoka, T., Kanamori, S., Nakamura, M. & Miura, K. (1988). *Biochem. Int.* **16**, 157–162.

Hirao, I., Nishimura, Y., Naraoka, T., Watanabe, K., Arata, Y. & Miura, K. (1989). *Nucleic Acids Res.* **17**, 2223–2231.

Hirao, I., Nishimura, Y., Tagawa, Y., Watanabe, K. & Miura, K. (1992). *Nucleic Acids Res.* **20**, 3891–3896.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Joshua-Tor, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D. & Sussman, J. L. (1992). *J. Mol. Biol.* **225**, 397–431.

Keniry, M. A. (2000). *Biopolymers*, **56**, 123–146.

Leslie, A. G. W. (1992). *Crystallographic Computing 5. From Chemistry to Biology*, edited by D. Moras, A. D. Podjarny & J.-C. Thierry. Oxford University Press.

Leslie, A. G. W., Arnott, S., Chandrasekaran, R. & Ratliff, R. L. (1980). *J. Mol. Biol.* **143**, 49–72.

Lu, X.-J. & Olson, W. K. (2003). *Nucleic Acids Res.* **31**, 5108–5121.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Padrta, P., Stefl, R., Kralik, L., Zidek, L. & Sklenar, V. (2002). *J. Biomol. NMR*, **24**, 1–14.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). *Nature (London)*, **372**, 68–74.

Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.

Rossmann, M. G. & van Beek, C. G. (1999). *Acta Cryst.* D**55**, 1631–1640.

Rupert, P. B. & Ferre-D'Amaré, A. R. (2001). *Nature (London)*, **410**, 780–786.

Samani, T. D., Jollès, B. & Laigle, A. (2001). *Antisense Nucleic Acid Drug Dev.* **11**, 129–136.

Sayle, R. A. & Milner-White, E. J. (1995). *Trends Biochem. Sci.* **20**, 374–376.

Scott, W. G., Finch, J. T. & Klug, A. (1995). *Cell*, **81**, 991–1002.

Snoussi, K., Nonin-Lecomte, S. & Leroy, J. L. (2001). *J. Mol. Biol.* **309**, 139–153.

Soyfer, V. N. & Potaman, V. N. (1996). *Triple-helical Nucleic Acids.* New York: Springer–Verlag.

Steller, I., Bolotovsky, R. & Rossmann, M. G. (1997). *J. Appl. Cryst.* **30**, 1036–1040.

Sunami, T., Kondo, J., Hirao, I., Watanabe, K., Miura, K. & Takénaka, A. (2004). *Acta Cryst.* D**60**, 90–96.

Sunami, T., Kondo, J., Kobuna, T., Hirao, I., Watanabe, K., Miura, K. & Takénaka, A. (2002). *Nucleic Acids Res.* **30**, 5253–5260.

Takénaka, A., Matsumoto, O., Chen, Y., Hasegawa, S., Chatake, T., Tsunoda, M., Ohta, T., Komatsu, Y., Koizumi, M. & Ohtsuka, E. (1995). *J. Biochem.* **117**, 850–855.

Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1937–1940.

Veselkov, A. N., Eaton, R. J., Semanin, A. V., Pakhomov, V. I., Dymant, L. N., Karavaev, L. & Davies, D. V. (2002). *Mol. Biol. (Moscow)*, **36**, 880–890.
Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. & Ficner, R. (2000). *Mol. Cell*, **6**, 1331–1342.
Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G. & Rich, A. (1979). *Nature (London)*, **282**, 680–686.
Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K. & Hirao, I. (1997). *Biochemistry*, **36**, 4761–4767.
Watson, J. D. & Crick, F. H. (1953). *Nature (London)*, **171**, 737–738.
Williams, H. E., Colgrave, M. L. & Searle, M. S. (2002). *Eur. J. Biochem.* **269**, 1726–1733.
Wimberly, B. T., Brodersen, D. E., Clemons, W. M. Jr, Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. (2000). *Nature (London)*, **407**, 327–339.
Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. & Noller, H. F. (2001). *Science*, **292**, 883–896.

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

addenda and errata

Structures of d(GCGAAGC) and d(GCGAAAGC) (tetragonal form): a switching of partners of the sheared G·A pairs to form a functional G·A×A·G crossing. Erratum

Tomoko Sunami,[a] Jiro Kondo,[a] Ichiro Hirao,[b] Kimitsuna Watanabe,[c] Kin-ichiro Miura[d] and Akio Takénaka[a]*

[a]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan, [b]RIKEN GSC, Wako-shi, Saitama 351-0198, Japan, [c]Graduate School of Engineering, University of Tokyo, Tokyo 113-8656, Japan, and [d]Faculty of Science, Gakushuin University, Tokyo 171-8588, Japan. Correspondence e-mail: atakenak@bio.titech.ac.jp

In the paper by Sunami *et al.* (2004), an asterisk was inadvertently missed out from the last sentence of §3.6 on p. 428. The corrected sentence should read as follows: In the base-stacked column, the ribose rings of the $A_{3.5}$ residues adopt a C3′-*endo* pucker to make space for accepting an A_4* base of the counter-strand between the $A_{3.5}$ and A_4 residues.

Table 4 of this paper also contained an error and a corrected verison of the table is given below.

Table 4
Sugar puckers of the 7hmt duplex.

Residue	Chain *A*	Chain *B*	Chain *C*	Chain *D*	Chain *E*	Chain *F*
G_1	C3′-*exo*	C3′-*exo*	C3'-*exo*	C3′-*exo*	C3′-*exo*	C3′-*exo*
C_2	C1′-*exo*	C1′-*exo*	C1′-*exo*	C2′-*endo*	C1′-*exo*	C2′-*endo*
G_3	C3′-*exo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*
A_4	C4′-*exo*	C3′-*endo*	C3′-*endo*	C3′-*endo*	C3′-*endo*	C3′-*endo*
A_5	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*	C2′-*endo*
G_6	C1′-*exo*	C1′-*exo*	C1′-*exo*	C4′-*exo*	O4′-*endo*	C1′-*exo*
C_7	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*	C4′-*exo*
A-form	C3′-*endo*					
B-form	C2′-*endo*					

References

Sunami, T., Kondo, J., Hirao, I., Watanabe, K., Miura, K. & Takénaka, A. (2004). *Acta Cryst.* D**60**, 422–431.

  DOI: 10.1107/S0907444904005104