High-Resolution Refinement of the Hexagonal A-DNA Octamer d(GTGTACAC) at 1.4 Å

BY NARASAIAH THOTA

Crystallography Laboratory, Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA, and Laboratory of Biological Macromolecular Structure, Department of Chemistry and Ohio State Biotechnology Center, The Ohio State University, 1060 Carmack Road, Columbus, OH 43210, USA

XI HU LI*

Crystallography Laboratory, Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA

AND CRAIG BINGMAN AND MUTTAIYA SUNDARALINGAM†

Crystallography Laboratory, Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA, and Laboratory of Biological Macromolecular Structure, Department of Chemistry and Ohio State Biotechnology Center, The Ohio State University, 1060 Carmack Road, Columbus, OH 43210, USA

Abstract

The hexagonal crystal form of the octamer d(GTGTACAC), grown in the presence of spermine, has unit-cell dimensions $a = b = 32.18$ and $c = 78.51$ Å, space group $P6_122$, with one DNA strand in the asymmetric unit. The structure has been refined starting with the earlier lower resolution model and using high-resolution 1.4 Å data collected on a Siemens–Xentronics area detector at 258 K. There were 4365 unique reflections greater than $2\sigma(F)$ in the resolution range 5–1.4 Å. The model was refitted into $3F_o - 2F_c$ Sim-weighted omit maps and difference maps were used to locate water molecules. The final model with 161 DNA atoms and 37 water molecules gave an R factor of 19.8%. Crystals of the same octamer were also grown in the presence of spermidine instead of spermine, and refinement using nominal 1.45 Å resolution data, 3292 unique reflections, final $R = 19.1\%$, gave virtually identical DNA parameters. No bound spermine or spermidine was detected in either of these structure analyses. The electron density was clear for the DNA and showed holes in the center of the six-membered rings of bases, and also in the center of some of the sugar rings. The high-resolution structure has provided more precise DNA parameters and confirmed the features observed in the earlier 2 Å study including the packing-induced distortion in the A7 (A15) sugar

pucker from C(3′)-*endo* and C(2′)-*endo*. This change causes the end base pairs to bend away from the helix axis while the rest of the duplex is nearly linear. The hydration patterns in the deep and shallow grooves have been characterized. Chains of water molecules were found, but no rings. The familiar intermolecular contact region between the end base pair and the minor groove of a symmetry-related duplex, involving four residues on one strand and two on the other, has been analyzed. One of these interactions is a hydrogen bond.

Introduction

The 2 Å resolution structures of the DNA octamer d(GTGTACAC) in the polymorphic tetragonal ($P4_32_12$) (Jain, Zon & Sundaralingam, 1987, 1989) and hexagonal ($P6_122$) (Jain & Sundaralingam, 1989) crystal forms have been determined earlier in our laboratory. The tetragonal form had a bound spermine (Jain, Zon & Sundaralingam, 1989) while the hexagonal form did not (Jain, Zon & Sundaralingam, 1991); however, the latter showed a flip in the pucker of the penultimate sugar ring to C(2′)-*endo*. Large crystals of the hexagonal crystal form which diffracted to high resolution were grown subsequently. 1.4 Å intensity data were collected for the current refinement study to confirm the structural features observed in the 2 Å model and to look for bound cations such as spermine or spermidine. The present study is at the highest resolution of any A-DNA structures reported so far. High-resolution

---

* Present address: Department of Chemistry, University of Wisconsin–Madison, Madison, WI 53706, USA.

† Author to whom correspondence should be addressed.

Table 1. *Crystallographic data for the spermine octamer* d(*GTGTACAC*)

Parameter	Data
Temperature	258 K
Crystal size	1.2 × 0.8 × 0.8 mm
Crystal system	Hexagonal
$a (= b)$	32.18 Å
c	78.51 Å
Volume	70 409 Å^3
Space group	$P6_122$
Z	12
Asymmetric unit	1 strand
Volume/base pair	1467 Å^3
V_m, Matthew's coefficient	2.50 Å^3:D
Resolution	1.40 Å
Possible reflections	5261
Number of observations recorded	20 889
$R_{sym}(F)$	3.4%
Unique reflections	4854 (92% of possible)
Unique reflections $> 2\sigma(F)$	4365 (83% of possible)
Observed reflections/unique base pair	1091
Number of waters/unique base pair	9.3
Number of reflections/water	118

structures of DNA provide precise DNA parameters and define the location of the solvent molecules with greater confidence.

Materials and methods

The octamer sequence was synthesized as reported earlier (Jain, Zon & Sundaralingam, 1991). The crystals were grown at 283 K from 20%(v/v) 2-methyl-2,4-pentanediol (MPD), 4 mM MgCl$_2$, 2 mM spermine and the DNA duplex, using the vapor diffusion technique against 50% MPD and 20 mM MgCl$_2$ in the reservoir. Crystals formed in a week and continued to grow for about a month with dimensions up to 1.2 × 0.8 × 0.8 mm. The X-ray intensity data were collected at 258 K using a Siemens–Xentronics area detector at Argonne National Laboratory. At this temperature, the unit-cell dimensions were about 1% less than at 277 K (Jain, Zon & Sundaralingam, 1991). A total of 975 frames at three different angular settings were collected from one crystal. Indexing and data processing were performed using the *XENGEN* package version 1.3 (Howard *et al.*, 1987). There were 4854 unique reflections with an R_{sym} of 3.4% on F. Of these, 4365

reflections, 83% of the total possible, were above $2\sigma(F)$. Details of the crystallographic data are given in Table 1. The number of reflections as a function of resolution is shown in Table 2.

The volume per base pair for the crystal is 1467 Å^3, which is slightly smaller than the values for most of the other hexagonal octamer structures (Dickerson, 1992). There are 1091 observed reflections per unique base pair representing the highest number yet observed for an A-DNA structure.

Large crystals of d(GTGTACAC) were also grown in the presence of the oligocation spermidine instead of spermine. The X-ray data were collected out to 1.45 Å resolution on the area detector at Argonne National Laboratory, adopting a procedure similar to that used for the spermine crystal. There were a total of 3292 unique reflections ($F > 2\sigma$) with an R_{sym} of 5.2%. Although 83% of the data to 1.6 Å were observed, only 47% of the data were observed between 1.6 and 1.45 Å.

Structure refinement

The DNA coordinates of the 2 Å resolution model (Jain, Zon & Sundaralingam, 1991) were used to start the refinement. The model was subjected to restrained least-squares refinement using the program *NUCLSQ* (Westhof, Dumas & Moras, 1985). The space group was lowered to $P6_1$ and the full octamer duplex was treated as the asymmetric unit so that the Watson–Crick base-pair hydrogen-bonding constraints can be imposed. The data were expanded to Laue group $6/m$ which nearly doubled the number of reflections. A tight non-crystallographic dyad was maintained to keep the two strands of the octamer nearly identical. Weighting scheme 2 given in the *NUCLSQ* program was used in the refinement. Before every round of map fitting using $3F_o - 2F_c$ Sim-weighted (Sim, 1960) omit maps, the coordinates of the two strands were averaged to make them identical. The average positional deviation between the two strands was no more than 0.005 Å.

Table 2. *Number of observed reflections, with* $F > 2\sigma(F)$, *in various resolution shells and their R factors*

The numbers correspond to the space group $P6_1$.

Resolution range (Å)	Total possible reflections	Spermine Number observed	% Observed	R factor	Spermidine Number observed	% Observed	R factor
∞–2.00	3102	2913	93	0.175	2792	90	0.178
2.00–1.90	516	500	97	0.204	418	81	0.224
1.90–1.80	637	617	97	0.221	522	82	0.212
1.80–1.70	796	757	95	0.220	595	74	0.221
1.70–1.60	1008	899	89	0.241	712	71	0.225
1.60–1.50	1294	1100	85	0.255	766	59	0.260
1.50–1.40	1683	944	56	0.264	212	13	0.280
Total	9036	7730	86	0.198*	6017	67	0.191*

* Final R factor.

Table 3. *Restrained parameters and refinement statistics for spermine-grown crystal of d(GTGTACAC)*

Restraints	Number	R.m.s.	σ	Number $> 2\sigma$
Distance restraints (Å)				
Bonds	304	0.006	0.019	2
Angles	488	0.021	0.050	
PO$_4$ bonds	56	0.044	0.050	
PO$_4$ angles	324	0.030	0.050	
H bonds	20	0.040	0.050	
Plane (Å)	168	0.021	0.030	
Chiral volume (Å^3)	48	0.069	0.100	
Non-bonded contacts (Å)				
Single torsion	3	0.109	0.063	2
Multiple torsion	51	0.213	0.063	30
Thermal factors (Å^2)				
Sugar–base bonds	304	2.386	7.5	
Sugar–base angles	448	3.240	7.5	
PO$_4$ bonds	56	3.832	7.5	
PO$_4$ angles, H bonds	344	3.784	10.0	

Spermine-grown crystal

In the initial round the thermal parameters of the DNA atoms were fixed at 20 Å^2 and the positional parameters were refined using the 2913 reflections in the resolution range 5.0–2.0 Å. In the subsequent rounds the resolution was increased in four stages to 1.4 Å. Spheroidal peaks that appeared in the omit difference maps with a density greater than 2.5σ and within hydrogen-bonding distances to the donor/acceptor DNA atoms or to other water molecules were included as water molecules. After several rounds of refinement the thermal factors for the waters were allowed to vary. A total of 43 water molecules were initially located per strand of the DNA. The B factors of all waters were held fixed and their occupancies were refined and six water molecules attained occupancies of less than 0.5. They were excluded in subsequent cycles and the occupancies of the remaining 37 water molecules were assigned values closest to the ideal values of 1.0, 0.75 and 0.5. In subsequent rounds, the occupancies of the water molecules were held fixed and the B values refined. The final R factor and the correlation coefficient are 19.8 and 96.9% respectively. The statistics from the restrained least-squares refinement are given in Table 3. The Luzzati plot (Luzzati, 1952) shows that the mean positional coordinate error is about 0.16 Å.

Spermidine-grown crystal

The spermidine-grown octamer structure was also refined starting with the 2 Å model. The 6017 reflections were included in two stages, first the 10–2 Å (2792 reflections) data and then the data between 2 and 1.45 Å (3225 reflections) were added. After several iterative rounds of refinement and electron-density-map fitting, and inclusion of 35 water molecules, the final R factor and correlation coefficient

were 19.1, 96.8% respectively. A Luzzati plot (Luzzati, 1952) indicates a mean positional error of about 0.17 Å.

The final coordinates of the DNA octamer d(GTGTACAC) and the water molecules for the crystals grown both in the presence of spermine and spermidine have been deposited in the Protein Data Bank* (Bernstein *et al.*, 1977). The r.m.s. difference

---

* Atomic coordinates and structure factors of the spermine-grown octamer (Reference: 1D78, R1D78SF) and the spermidine-grown octamer (Reference: 1D79, R1D79SF) have been deposited with the Protein Data Bank, Brookhaven National Laboratory, and are available in machine-readable form from the Protein Data Bank at Brookhaven. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37066 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. At the request of the authors, the list of structure factors will remain privileged until 1 September 1993.
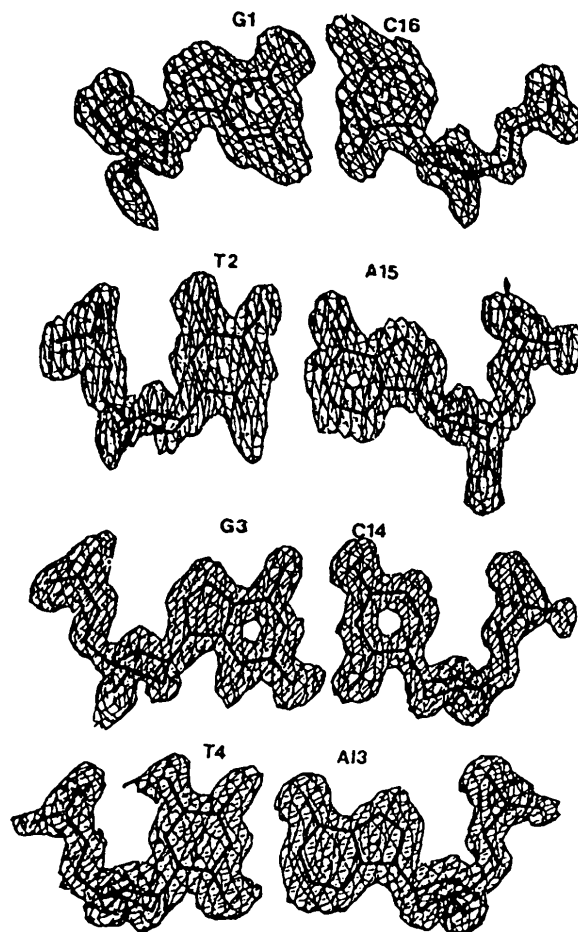

Fig. 1. The electron densities in $3F_o - 2F_c$ omit maps for the four independent nucleotide base pairs G1–C16, T2–A15, G3–C14 and T4–A13. In each map, the two nucleotides shown were omitted from the phasing calculation. Contours are at the 1.1σ level.

Table 4. *Torsion angles of the backbone of* d(GTGTACAC) *for the spermine* (top row), *the spermidine* (middle row) *and the* 2 Å *models* (bottom row)

Sequence	α	β	Backbone and glycosyl angles* ( )					Pseudorotation parameters ( )	
			γ	δ	ε	ζ	χ	P	τ_m
G	—		51	86	210	290	183	9.7	38.9
		..	(56)	(84)	(207)	(290)	(179)	(14.6)	(39.1)
			60	83	210	290	180	17.3	37.7
T	291	185	49	80	205	288	201	18.0	43.5
	(298)	(179)	(45)	(80)	(204)	(289)	(202)	(23.4)	(41.6)
	287	185	52	80	205	293	201	17.8	40.8
G	293	170	65	81	200	284	197	23.7	42.9
	(294)	(162)	(65)	(76)	(196)	(288)	(197)	(26.1)	(44.4)
	298	172	52	81	187	297	203	17.9	41.3
T	297	175	47	78	202	290	206	15.4	44.4
	(291)	(176)	(50)	(78)	(197)	(291)	(207)	(18.5)	(42.3)
	284	180	55	76	199	287	207	20.4	42.1
A	295	166	56	77	210	285	205	19.6	43.3
	(299)	(166)	(53)	(81)	(209)	(282)	(206)	(18.2)	(40.2)
	308	162	49	81	211	286	204	22.2	40.4
C	300	161	59	77	197	282	208	16.6	43.5
	(305)	(158)	(55)	(77)	(201)	(276)	(207)	(18.0)	(45.0)
	289	171	62	73	187	296	209	20.3	45.3
A	306	178	60	138	189	275	240	160.8	36.1
	(316)	(171)	(54)	(139)	(185)	(277)	(240)	(161.1)	(34.6)
	300	177	58	137	187	277	243	163.6	36.0
C	282	176	57	94	—		219	79.6	37.8
	(282)	(178)	(49)	(84)			(220)	(58.9)	(42.4)
	281	173	57	89			216	68.5	34.7

* An earlier nomenclature for the backbone torsion angles is $\omega = \alpha$, $\varphi = \beta$, $\psi = \gamma$, $\psi' = \delta$, $\varphi = \epsilon$, $\omega' = \zeta$ (Sundaralingam, 1969).

for the DNA atoms in the two refinements is 0.08 Å and the atoms deviating more than 0.15 Å are: O(1P)(G3) and O(2P), C(3') and O(3') of the terminal residue C8. 30 water sites are also common. Since the two DNA parameters are very similar, we restrict the discussion to the results obtained with the spermine crystal which employed higher resolution data.

Results and discussion

As generally observed in oligonucleotide duplex structures, the peripheral phosphate groups have larger thermal factors (23 Å²) than the sugars (19 Å²) and bases (15 Å²). The group thermal factors for each nucleotide in the present high-resolution study are lower than the corresponding values in the earlier low-resolution 2 Å study. The electron-density maps were very clear and showed the characteristic central holes in many of the base six-membered rings and the sugar rings (Fig. 1). Also, the electron density for the penultimate residue A7 (A15) clearly showed C(2')-*endo* sugar puckering, confirming the earlier assignment. This figure also shows the electron density for the 3'- and 5'-phosphate groups, where the positions of the phosphate oxygens are very clear. The conformational angles and various helical and base-pair parameters are compared in Tables 4 and 5, respectively, for the crystals grown in sper-

mine, spermidine and our previous 2 Å resolution study. The values were calculated using the NEWHEL91 program (R. E. Dickerson, personal communication).

Molecular conformation

The torsion angles for the backbone, sugar and base–sugar glycosyl bonds are given in Table 4. Except for the change in the sugar pucker of residue A7 (A15) from C(3')-*endo* to C(2')-*endo*, the other conformational parameters are characteristic of an A-DNA nucleotide [C(3')-*endo* sugar, *gauche*⁺ conformation about the exocyclic C(4')—C(5') bond and *anti* conformation about the glycosyl C(1')—N bond].

DNA bending

The DNA duplex is bent at the ends. To characterize the nature of this bend, helical axes were computed for all contiguous three base-pair fragments in the duplex. Whenever the fragments included the terminal base pairs, the fragment axes deviated by about 50° from the overall helical axis. All other internal fragments were within 4° of the overall helical axis. Therefore, we conclude that the central six dyad-related base pairs are straight and the bending in the duplex is between A7 (A15) and C8 (C16). Apparently the puckering change at A7

Table 5. *Comparison of the helical parameters of d(GTGTACAC) for the spermine (top row), the spermidine (middle row) and the 2 Å models (bottom row)*

Sequence	Tip (°)	Inclination (°)	Roll (°)	Propeller twist (°)	Buckle (°)	Twist (°)	Rise (Å)	χ displacement (Å)
G	11.3	12.0		-17.0	14.5			-3.7
	(11.3)	(12.2)		(-18.5)	(15.2)			(-3.6)
	8.0	13.7		-16.8	11.1			-3.6
			-11.46				35.3	2.1
			(-12.56)				(35.2)	(2.2)
			-11.98				35.9	2.1
T	0.2	19.1		-8.3	18.7			-3.8
	(-0.9)	(19.1)		(-8.7)	(17.8)			(-3.7)
	-4.0	17.8		-10.3	19.4			-3.8
			2.08				31.1	3.2
			(3.18)				(31.4)	(3.2)
			3.91				31.3	3.4
G	2.4	18.8		-11.4	8.5			-3.3
	(2.1)	(18.4)		(-12.8)	(8.3)			(-3.2)
	0.5	16.1		-13.9	9.8			-3.4
			-5.97				33.4	2.8
			(-5.58)				(33.3)	(2.8)
			-4.71				32.3	3.0
T	-3.4	18.1		-12.8	1.5			-3.9
	(-3.2)	(18.2)		(-11.9)	(2.0)			(-3.8)
	-2.2	14.7		-11.5	1.3			-3.9
			7.39				32.9	3.0
			(6.73)				(32.3)	(3.0)
			2.73				31.6	3.1
A	3.4	18.1		-12.8	-1.6			-3.9
	(3.2)	(18.2)		(-11.9)	(-2.0)			(-3.8)
	2.2	14.7		-11.5	-1.3			-3.4
			-6.13				33.4	2.8
			(-5.54)				(33.3)	(2.8)
			-4.66				31.3	3.0
C	-2.3	18.8		-11.3	-8.5			-3.3
	(-2.1)	(18.4)		(-12.8)	(-8.4)			(-3.2)
	-0.5	16.1		-13.9	-9.8			-3.4
			2.21				31.1	3.2
			(3.16)				(31.4)	(3.3)
			3.86				32.3	3.4
A	-0.2	19.0		-8.3	-18.8			-3.8
	(0.9)	(19.1)		(-8.8)	(-17.9)			(-3.7)
	4.0	17.8		-10.2	-19.4			-3.8
			-11.42				35.6	2.1
			(-12.53)				(35.2)	(2.1)
			-11.95				35.9	2.1
C	-11.3	12.0		-17.0	-14.6			-3.7
	(-11.3)	(12.2)		(-18.5)	(-15.2)			(-3.6)
	-8.00	13.7		-16.8	-11.1			-3.6
Average		17.0	-12.4			33.3	2.7	-3.7
		(17.0)	(-13.0)			(33.2)	(2.8)	(-3.6)
		15.6	-13.1			32.9	2.9	-3.6

(A15) causes the local helical axis to deviate substantially from the molecular axis around the terminal base pairs. It is interesting that the major bend in the molecule is not at the center of the duplex as has been seen in other related octamers (Jain & Sundaralingam, 1989; Shakked *et al.*, 1981), but rather at the ends. Thus a C(2')-*endo* sugar in one strand of an A-form duplex might cause an abrupt bend at that point.

Groove sizes

The P—P distances across the major and minor grooves and P—P separations within a strand are shown in a cylindrical projection in Fig. 2. The minor groove width is 10.3 Å and the pseudo major groove width between P(2)—P(10) is 5.8 Å.

Helical and base-pair parameters

The distortion in the penultimate sugar pucker is perhaps responsible for many of the deviations seen in the helical parameters of the end base pairs. The average rise and twist of the base pairs are 2.74 Å and 33.3°. The twist and roll angles show an alternation (Table 5 and Fig. 3). The twist exhibits high and low values for the Pu–Py and the Py–Pu steps respectively, while the roll shows the reverse
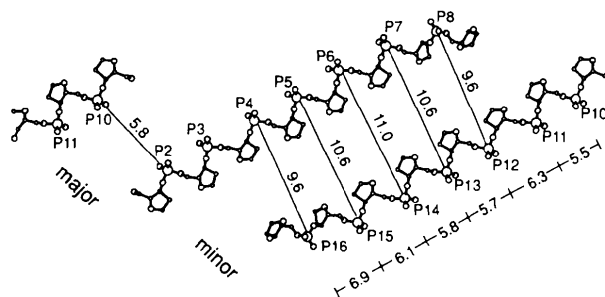

Fig. 2. Cylindrical projection of d(GTGTACAC) showing the P—P separations (Å) in the major and minor grooves and within a strand.

trend. The high twist angles for the first and last steps probably arise from the puckering distortion in the penultimate sugar and end effects. Similar alternation at the Pu–Py steps is observed for the tip. Again, the larger tip angles for the terminal base pairs can be attributed to the flip in the sugar pucker

of the penultimate residues (Jain, Zon & Sundaralingam, 1991). The inclination angles show similar values except for the terminal base pairs. The higher propeller twist angles for the terminal G–C base pairs are probably a result of its maximizing the intramolecular stacking and the intermolecular inter-
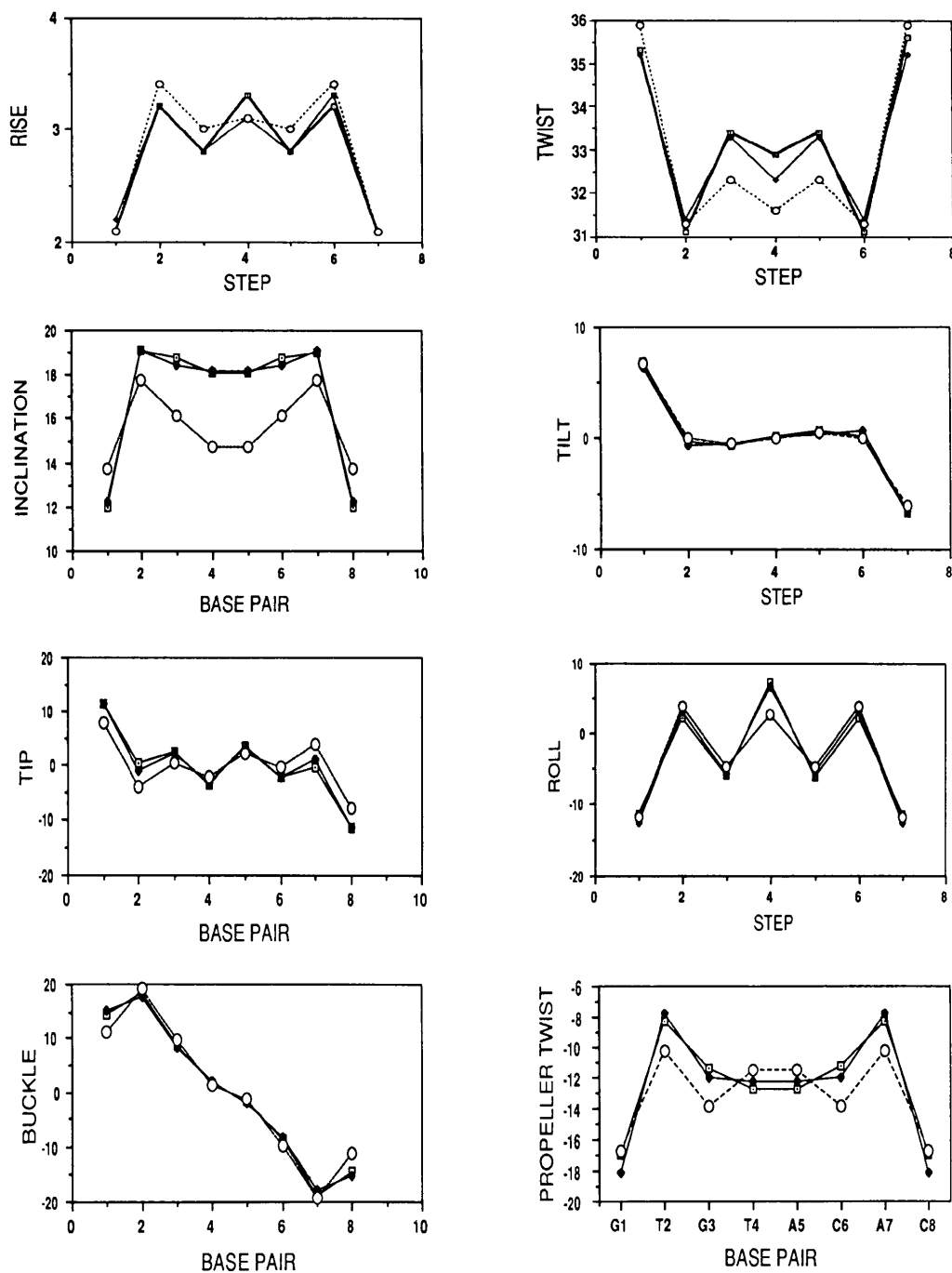
Fig. 3. The helical and base-pair parameters (Å, °) for d(GTGTACAC): solid line, open square, represents spermine-grown crystal; solid line, filled square, represents spermidine-grown crystal; and dashed line, open circle, represents the 2 Å model.
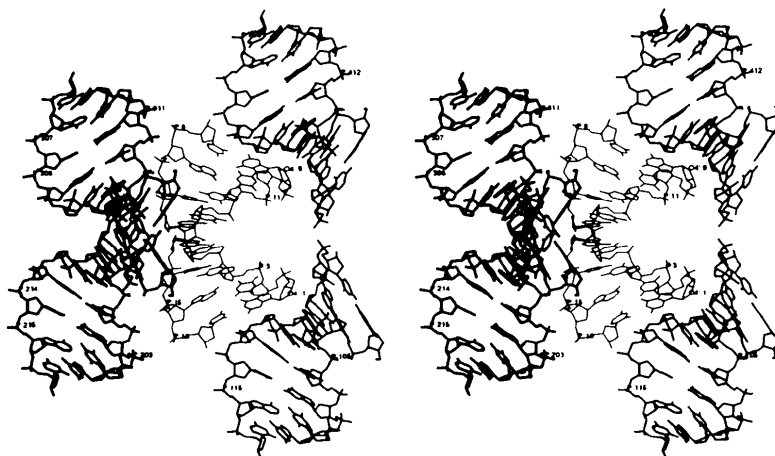
Fig. 4. A stereo figure of the intermolecular interactions between the reference duplex (light bonds) and the four surrounding symmetry-related duplexes (dark bonds).

action with the minor groove of a symmetry-related duplex. For similar reasons, the buckle angle is high for the terminal base pairs.

Intermolecular contacts

In the hexagonal crystal form, each duplex is in direct contact with four neighboring symmetry-related duplexes (Fig. 4). As in other A-DNA oligonucleotide structures, the terminal base pair of the reference duplex interacts with the sugar–phosphate backbone of a 6_1-related duplex and *vice versa*. The extensive van der Waals contacts between the complementary surfaces of the terminal base pair of the reference molecule and the minor groove backbone of the symmetry-related molecule are shown in Fig.
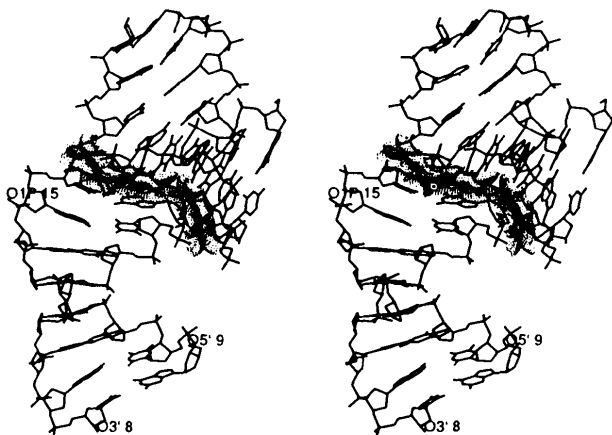

Fig. 5. A stereo figure of the van der Waals contacts between the reference duplex (bottom) and the neighboring duplex (top). The van der Waals radii used are C = 1.7, N = 1.6 and O = 1.5 Å.

Table 6. *Short contacts* (<3.5 Å) *between the terminal G1–C16 base pair and the minor groove of a* 6_1 *symmetry-related* $(x − y, x − 1, 1/6 + z)$ *duplex*

The hydrogen bond is shown in bold.

Reference molecule	Neighboring molecule	Distance (Å)
C(2')(C16)	O(2P)(C14)	3.50
C(2')(C16)	C(5')(C14)	3.50
C(2)(G1)	O(4')(A13)	3.30
N(1)(G1)	O(4')(A13)	3.40
N(2)(G1)	O(4')(A13)	3.40
C(1')(G1)	O(2)(G12)	3.20
O(4')(G1)	**N(2)(G11)**	**2.90**
O(4')(G1)	O(2)(A6)	3.30
C(4')(G1)	O(2)(C6)	3.50
C(5')(G1)	O(4')(A7)	3.15
C(4')(G1)	O(4')(A7)	3.40
O(2P)(P2)	C(5')(A7)	3.30

5. The atomic distances within 3.5 Å between the duplexes in contact are shown in Fig. 6 and Table 6. Notice that the residues of both strands are involved. The terminal nucleotide atoms O(2) and C(2') of C16 in strand 2 interact with the phosphate P(14) and sugar A13 of strand 2 of the symmetry-related duplex. The complementary nucleotide G1 in strand 1 of the reference duplex is in contact with O(4')(A13) and O(2)(T12) of strand 2, while the sugar interacts with O(2)(T12) and N(2)(G11). O(4')G(1)···N(2)(G11) is the only hydrogen bond in this contact zone. At this point the interactions switch across the minor groove and the sugar of G1 interacts with O(2)(C6) and O(4')(A7), while the 3'-phosphate interacts with C(5')(A7), all on strand 1. Thus, the terminal base pair G1–C16 interacts with the tetranucleotide stretch C14–A13–T12–G11 of strand 2 and with the dinucleotide stretch C6–G7 of strand 1 of the symmetry-related duplex in the minor groove.
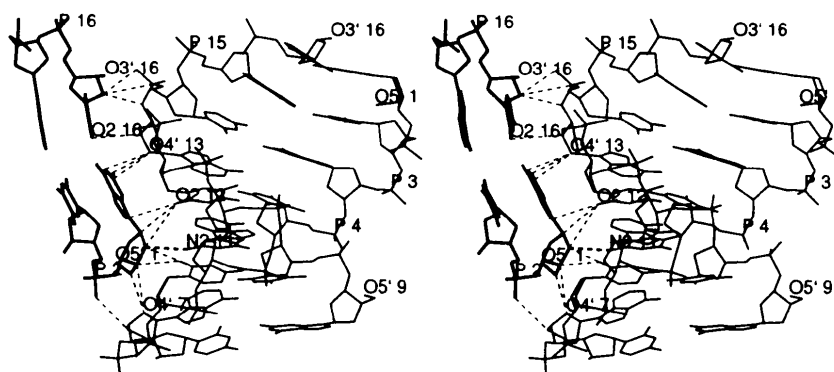
Fig. 6. A stereo figure in which the van der Waals contacts between the terminal nucleotide pair (G1 C16) of the reference duplex and the sugar-phosphate backbone of the neighbouring duplex are shown by dashed lines, see also Table 6.
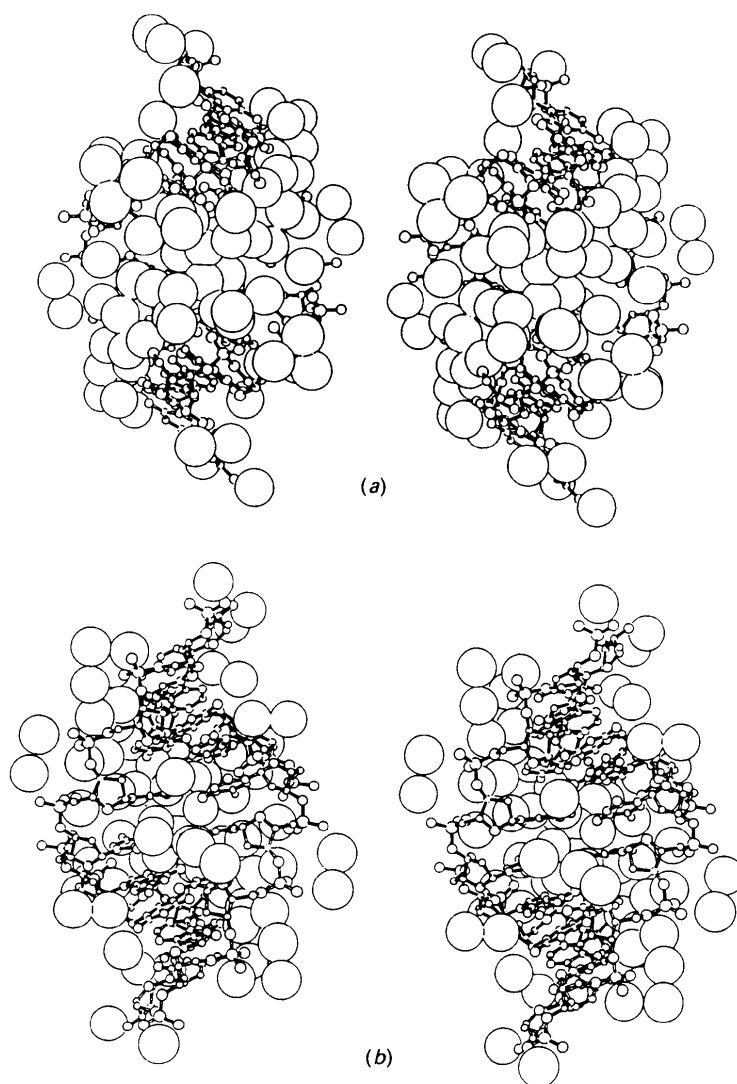

(a)

(b)

Fig. 7. Stereo figures showing the hydration pattern viewed down the dyad axis, (a) into the major groove and (b) into the minor groove. The water molecules are drawn as spheres of radius 1.4 Å, while the DNA is shown by ball-and-stick representation. The bonds in one strand of the duplex are drawn filled.

Hydration

There are 74 water molecules around the octamer duplex (Fig. 7) representing 9.3 water molecules per base pair. About 18 water molecules are in the major groove interacting with the polar atoms of the bases and neighboring water molecules. The shallow minor groove, on the other hand, contains only 10 water molecules. The phosphate anionic oxygen atoms interact with 32 waters, including the phosphate bridges. Some of the water molecules form hydrogen-bonded chains (Table 7). A few of the water molecules are close to the phosphate ester oxygen atoms O(5′) (on the major groove side) and O(3′) (on the minor groove side), with closer contacts to O(3′). The sugar ring oxygen atoms of the four Py nucleotides are found to be close to water molecules which also hydrogen bond to N(3) of the purine base on the 5′-side. The water molecules seem to occupy a cleft on the 5′-side of each pyrimidine nucleotide in this alternating Pu–Py oligonucleotide and appear to be correlated with the relatively low helical twist angles for Py–Pu steps.

Of the six possible inter-phosphate sites only three [P(2)—P(3), P(4)—P(5) and P(5)—P(6)] are bridged by water molecules; all three involve C(3′)-*endo* sugar–phosphate segments. P(4)—P(5) and P(5)—P(6) water bridges are on three successive phosphates at the center of the duplex around the dyad axis and probably stabilize the backbone in the C(3′)-*endo* conformation. These water bridges have been found in all the three refinements: spermine, spermidine and the 2 Å study. The absence of a water bridge between P(3)—P(4), P(6)—P(7) and P(7)—P(8) may be due to the increase in the adjacent P—P distances to 6.3, 6.1 and 6.9 Å respectively. The largest is the P(7)—P(8) distance which is caused by the change in the A7 sugar pucker. It is interesting that there is a water bridge at this site on the opposite strand. The smaller increases in the other two P—P distances are probably contributed by small conformational changes around the other backbone bonds including the sugar, and do not have water bridges on the opposite strand either. P(6) and P(7) are bridged by two water molecules.

P(8) and P(16) of two dyad-related duplexes are bridged by two water molecules (Fig. 8). There are 12 water molecules in the second coordination sphere.

Concluding remarks

The two structure refinements of the same DNA octamer duplex d(GTGTACAC) carried out in our laboratory with crystals grown in the presence of the oligocations spermine and spermidine have shown that the structural and conformational parameters

Table 7. *Water hydrogen bonding* (Å, °) *in the major groove, minor groove and backbone*

B—D···W hydrogen bond is listed if the *D···W* distance is within 3.3 Å and the angle at atom *D* is in the range 90–150°. Waters taking part in the inter-phosphate bridges are identified by asterisks.

| Residue	B	D	W	D···W	B—D···W
T2	C(4)	O(4)	W(85)	2.65	128
T2	C(4)	O(4)	W(45)	2.79	150
G3	C(8)	N(7)	W(60)	2.50	119
G3	C(6)	O(6)	W(31)	2.40	138
G3	C(6)	O(6)	W(48)	2.82	141
T4	C(4)	O(4)	W(32)	2.44	128
A5	C(6)	N(6)	W(68)	2.97	150
A5	C(8)	N(7)	W(72)	2.62	122
C6	C(4)	N(4)	W(30)	3.12	107
A7	C(8)	N(7)	W(17)	2.67	125
C8	C(4)	N(4)	W(43)	3.00	120
G1	C(2)	N(2)	W(84)	2.89	124
G3	C(2)	N(2)	W(24)	3.19	91
G3	C(2)	N(3)	W(24)	2.79	110
T4	C(2)	O(2)	W(71)	2.83	145
A5	C(2)	N(3)	W(75)	3.02	104
A7	C(2)	N(3)	W(29)	2.82	103
T2	P	O(1P)	W(19)	2.50	143
T2	P	O(1P)	W(49)	3.00	92
T2	P	O(2P)	W(49)	2.66	102
T2	P	O(1P)	W(58)*	2.93	105
T2	P	O(2P)	W(77)	2.57	145
G3	P	O(1P)	W(41)	2.88	106
G3	P	O(1P)	W(90)	2.71	106
G3	P	O(2P)	W(58)*	2.78	143
G3	P	O(2P)	W(57)	2.75	129
T4	P	O(1P)	W(59)*	2.77	109
A5	P	O(2P)	W(59)*	2.66	144
A5	P	O(2P)	W(44)*	3.20	103
A5	P	O(1P)	W(62)	2.85	126
C6	P	O(1P)	W(44)*	2.89	145
C6	P	O(1P)	W(70)	3.03	110
C6	P	O(2P)	W(27)	2.66	150
A7	P	O(1P)	W(50)	2.92	121
A7	P	O(2P)	W(55)	2.78	114
A7	P	O(2P)	W(63)	2.50	133
C8	P	O(2P)	W(52)	2.88	105
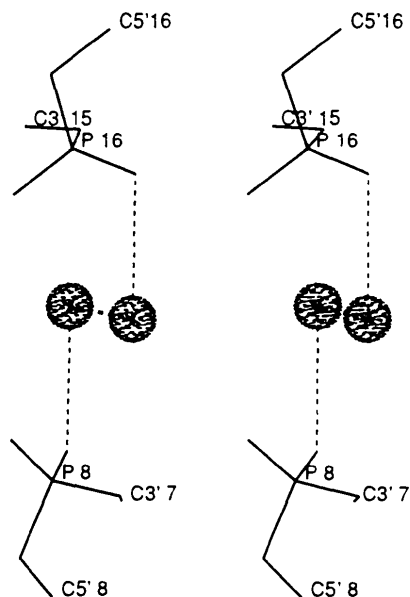
Fig. 8. Stereo figure of the phosphate groups P8 and P16 of two different duplexes related by a dyad axis, which is nearly horizontal in the figure, bridged by two water molecules of full occupancy, shown as dotted circles.

obtained for the DNA in the two separate studies are virtually identical. However, a larger number of solvent sites (50.5) were identified in our low-resolution (2 Å) study compared to 37 solvent sites in the present high-resolution (1.4 Å) study. 29 of these solvent sites are common. The remaining differences in the solvent sites may be due to the genuine differences between crystals or artifacts of the low-resolution study. Thus, the high-resolution study has provided more precise DNA parameters and solvent sites.

Note added in proof: The model for the spermine crystal was again refined in the higher symmetry space group, with one strand in the asymmetric unit. In *NUCLSQ*, the base-pair hydrogen-bond restraints, which are now intermolecular, could not be imposed. Six cycles resulted in a reduction in the *R* value by 0.3%, with r.m.s. shifts in coordinates and *B* values of 0.05 Å and 0.5 Å^2, respectively. The high-resolution data retained the base-pair geometry between the two strands, without any constraints.

The refinement was repeated with the program *X-PLOR* (Brünger, 1990), where the base-pair geometry is restrained. The *R* value fell by 1.8% with r.m.s. shifts in the coordinates and *B* values of 0.16 Å and 1 Å^2, respectively, with the largest shifts for atoms at the 5′-terminus. A superposition of the

two refined models shows an r.m.s. deviation of 0.10 Å, with backbone atoms deviating most (r.m.s. 0.11 Å) and the base atoms the least (r.m.s. 0.07 Å). These deviations are probably the result of the slightly different force fields used in the two programs. The coordinate shifts in both these refinements are less than the estimated positional error in the starting model reported in the paper.

References

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., ROGERS, J. R., KENNARD, O., SCHIMANOUCHI, T. & TASUMI, M. J. (1977). *J. Mol. Biol.* **112**, 535–542.

BRÜNGER, A. T. (1990). *X-PLOR Manual.* Version 2.1. Yale Univ., New Haven, CT, USA.

DICKERSON, R. E. (1992). *Methods Enzymol.* **211**. In the press.

HOWARD, A. J., GILLILAND, G. L., FINZEL, B. C., POULOS, T. L., OHLENDORF, D. H. & SALEMME, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.

JAIN, S. & SUNDARALINGAM, M. (1989). *J. Biol. Chem.* **264**, 12780–12784.

JAIN, S., ZON, G. & SUNDARALINGAM, M. (1987). *J. Mol. Biol.* **197**, 141–145.

JAIN, S., ZON, G. & SUNDARALINGAM, M. (1989). *Biochemistry*, **28**, 2360–2364.

JAIN, S., ZON, G. & SUNDARALINGAM, M. (1991). *Biochemistry*, **30**, 3567–3576.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

SHAKKED, Z., RABINOVICH, D., CRUISE, W. B. T., EGERT, E., KENNARD, O., SALA, G., SALISBURY, S. A. & VISWAMITRA, M. A. (1981). *Proc. R. Soc. London Ser. B*, **213**, 479–487.

SIM, G. A. (1960). *Acta Cryst.* **13**, 511–512.

SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.

WESTHOF, E., DUMAS, P. & MORAS, D. (1985). *J. Mol. Biol.* **184**, 119–145.