Evidence for Crystal Environment Dominating Base Sequence Effects on DNA Conformation: Crystal Structures of the Orthorhombic and Hexagonal Polymorphs of the A-DNA Decamer d(GCGGGCCCGC) and Comparison with Their Isomorphous Crystal Structures[†,‡]

B. Ramakrishnan and M. Sundaralingam[*]

Laboratory of Biological Macromolecular Structure, Departments of Chemistry and Biochemistry and Ohio State Biotechnology Center, The Ohio State University, 1060 Carmack Road, Columbus, Ohio 43210-1002

Received April 20, 1993; Revised Manuscript Received August 19, 1993[⊗]

ABSTRACT: We have determined the structure of the A-DNA decamer d(GCGGGCCCGC) in two crystal forms, orthorhombic and hexagonal, at 1.7- and 1.8-Å resolution, respectively. In the orthorhombic form, the fifth guanine residue has nearly *trans–trans* conformations for the α–γ backbone torsions, as in the isomorphous orthorhombic structure d(CCCGGCCGGG) [Ramakrishnan, B., & Sundaralingam, M. (1993) *J. Mol. Biol. 231*, 431–444]. However, in the hexagonal form, the eighth cytosine residue adopts the *trans–trans* conformations for the backbone α–γ torsions, as in the isomorphous hexagonal structure d(ACCGGCCGGT) [Frederick, C. A., Quigley, G. J., Teng, M.-K., Coll, M., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Eur. J. Biochem. 181*, 295–307]. Even though the average helix and base-pair parameters are nearly the same in the two polymorphous crystal forms having the same sequence, many of the base-dependent local helix parameters are quite different. However, in the isomorphous crystal forms, in spite of the differing base sequences, the local helix and base-pair parameters of the duplexes are nearly the same. This indicates that, in crystals, the local conformation of a DNA structure is affected severely by the crystal packing environment rather than by the base sequence.

In protein crystals, the crystal environment mainly affects the conformation of the surface residues of the protein, leaving the core structure intact, whereas in oligonucleotide crystals, the entire periphery of the molecule is exposed and accessible to the crystal environment so that it is susceptible to conformational change. Dickerson et al. (1987) have discussed the effect of the crystal environment on the breakdown of the dyad symmetry of the B-DNA dodecamer d(CGCGAATTCGCG). In the isomorphous crystal structure of the dodecamer d(CGBr[5]CAAAAATGCG), DiGabriele et al. (1989) found that two orientations of an asymmetric DNA sequence coexist in the crystal lattice, and when the sequences of these two differently oriented helices are aligned the bends are not identical. Both of these observations provide evidence that the crystal environment plays an important role in determining DNA conformations. Although approximately 80 DNA oligomer crystal structures have been determined to date, only a few sequences are known to crystallize in more than one crystal system. Jain et al. (1991) crystallized the A-DNA octamer d(GTGTACAC) in two crystal forms (tetragonal and hexagonal) and showed that the conformations of the duplexes were different due to crystal packing. Although the tetragonal form has been obtained for many other octamer sequences, the hexagonal form is the only known reported structure with its own unique crystal environment. Similarly, Shakked et al. (1988) showed that the conformations of the A-DNA octamer d(GGGCGCCC) in the two crystal forms (tetragonal and hexagonal) were different. The tetragonal form belongs to the same space group as the Jain et al. (1991) octamer, while the hexagonal form has a different space group. Recently, Lipanov et al. (1993) crystallized the B-DNA decamer d(CCAACITTGG) containing the uncommon residue inosine (I) in two different crystal forms (monoclinic and trigonal), but the monoclinic form is a magnesium complex while the trigonal form is a calcium complex and the metal coordination in the two structures is different. Although many other related monoclinic B-DNA decamer crystal structures are known, the trigonal form is the only known structure. The Z-DNA tetramer d(CGCG) (Crawford et al., 1980; Drew et al., 1980) was also obtained in two crystal forms (hexagonal and orthorhombic).

It will be important to have crystal structures of DNA oligomers close to a full turn in order to obtain detailed information about groove sizes and conformation. The effects of base sequence and crystal packing environment on DNA can be unscrambled if crystal structural information is available for the same sequence in different crystalline environments (polymorphous crystals) and different sequences in the same crystalline environment (isomorphous crystals). We have obtained crystals of the same decamer d(GCGGGCCCGC) in two A-DNA crystal forms, orthorhombic ($P2_12_12_1$) and hexagonal ($P6_122$), which can be compared with the corresponding isomorphous A-DNA decamer crystal structures with differing base sequences in the same orthorhombic, d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam,

[†] We gratefully thank the National Institutes of Health of the U.S. Public Health Service for supporting this research through Grant GM17378, the Regents of The Ohio State University for an Ohio Eminent Scholar Chair to M.S., and the Ohio Supercomputing Center for computer time on the Cray Y-MP/864.

[‡] Presented at the American Crystallographic Association Meeting, Albuquerque, New Mexico, May 23–28, 1993, Abstract BB04.

[*] Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1993.

0006-2960/93/0432-11458$04.00/0

Table I: Crystal Data on the Polymorphous and Isomorphous Crystal Structures

	orthorhombic		hexagonal	
	d(GCGGGCCCGC)[a]	d(CCCGGCCGGG)[b]	d(GCGGGCCCGC)[a]	d(ACCGGCCGGT)[c]
a (Å)	24.90	24.88	39.11	39.23
b (Å)	44.84	44.60	39.11	39.23
c (Å)	47.97	46.97	79.23	78.00
space group	$P2_12_12_1$	$P2_12_12_1$	$P6_122$	$P6_122$
resolution (Å)	1.7	1.65	1.8	2.0
asymmetric unit	one duplex	one duplex	one strand	one strand
vol/base pair (Å^3)	1339	1330	1749	1773
no. of reflections	5445	6035	2300	1434
final R (%)	18.6	18.5	18.2	18.0
no. of solvents	101 waters	103 waters	39 waters	36 waters

[a] This work. [b] Ramakrishnan & Sundaralingam, 1993. [c] Frederick et al., 1989.

1993), and hexagonal, d(ACCGGCCGGT) (Frederick et al., 1989), systems. In our orthorhombic form, the duplex is the asymmetric unit as in the B-DNA dodecamer; however, the hexagonal form has a molecular dyad and only one strand of the duplex represents the asymmetric unit. Thus, we can compare the two strands with each other in the orthorhombic forms and compare both with the hexagonal duplex.

MATERIALS AND METHODS

The DNA decamer d(GCGGGCCCGC) was synthesized using our in-house Applied Biosystem DNA synthesizer (trityl yield at each coupling was ≈99%). The base-deprotected crude oligomer was precipitated by ethanol at −25 °C in the presence of 2.5 M ammonium acetate, and the lyophilized precipitate was used for crystallization without further purification.

Orthorhombic Crystal Form. Crystals were grown by the hanging drop method from a drop containing 2 mM DNA, 40 mM sodium cacodylate buffer (pH 7.0), and 5 mM spermine tetrachloride against 50% (v/v) MPD (2-methyl-2,4-pentanediol) in water. The present crystallization conditions are similar to those used for the isomorphous orthorhombic decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993), with the exception of 75% MPD for the reservoir solution. The crystals belong to the orthorhombic system, space group $P2_12_12_1$, with a = 24.90, b = 44.84, and c = 47.97 Å. The volume per base pair, 1339 Å^3, based on two DNA strands or one DNA duplex in the asymmetric unit, is close to the value (1330 Å^3) observed for the isomorphous orthorhombic decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993). A crystal of size 0.45 × 0.45 × 0.2 mm was mounted along the a-axis in a thin-walled glass capillary with some mother liquid at one end. Three-dimensional X-ray diffraction data from 8- to 1.7-Å resolution were collected from this crystal at room temperature (290 K) on a Siemens area detector using a rotating anode X-ray source and graphite monochromated Cu Kα radiation (λ = 1.5418 Å) operating at 50 kV and 90 mA. The crystal to detector distance was 12.0 cm. One ϕ scan (range 180°) and two ω scans (range 60° each) were performed at 0.25° steps, with each frame exposed for 90 s. The frames were processed using XENGEN 1.3 (Howard, 1985). Of a total of 17 581 reflections collected, 6219 were unique (97% of possible reflections) with an R_{sym} of 2.8%. The crystal unit cell dimensions are quite similar and therefore isomorphous with the A-DNA decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) (Table I), which was recently solved by us.

A rigid-body refinement starting with the coordinates of the latter decamer against 131 reflections between 8- and 5-Å resolution dropped the R value from 48% to 30% (rms displacement in atomic coordinates is 1.07 Å). Further positional and thermal parameter refinement using 3553

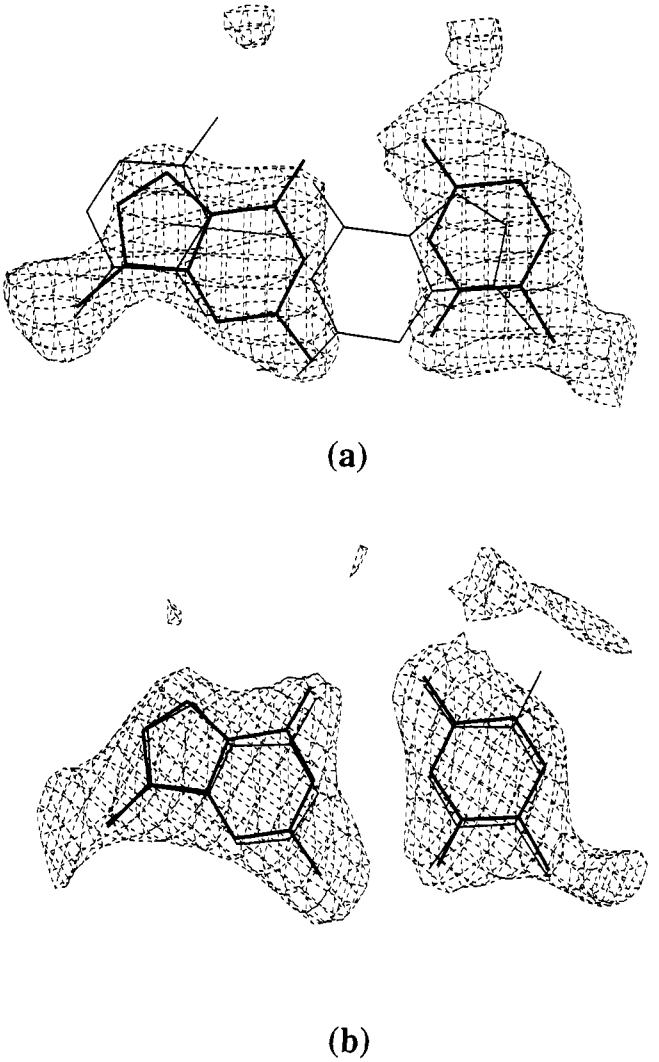


FIGURE 1: $3F_o - 2F_c$ omit maps for the first base pair at the 1.5σ level. The maps were clipped at 1.5 Å radius around all of the atoms displayed. (a) Map omitting only the base-pair atoms of the first base pair C·G (thin lines) in the starting decamer model d(CCCG-GCCGGG) after the first round of refinement. The map clearly shows that the actual base pair is G·C (thick lines) instead of the starting base pair C·G, in accordance with the DNA sequence d(GCGGGCCCGC). (b) Map omitting only the base-pair atoms for the first base pair A·T (thin lines) in the starting decamer model d(ACCGGCCGGT) after the first round of refinement. The map clearly indicates the presence of the N2 amino atom in the guanine base and the absence of the 5-methyl group in the cytosine base, confirming the G·C (thick lines) base pair in accordance with the actual DNA sequence d(GCGGGCCCGC).

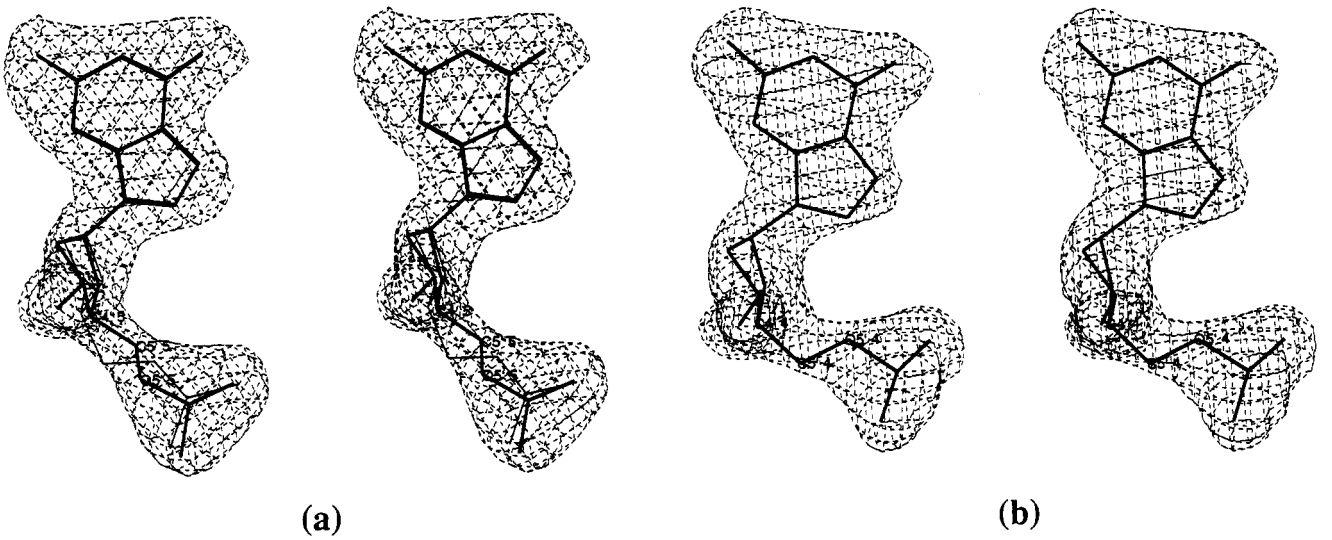


FIGURE 2: $3F_o - 2F_c$ omit map for the fifth (a) and fourth (b) residues in the orthorhombic crystal forms. (a) In the orthorhombic crystal form, the fifth residue has near *trans* conformation for the torsion angles α and γ about the bonds P–O5′ and C5′–C4′, respectively. The residue with the *trans* conformation is shown by the thick line. The residue having the *gauche*$^-$ and *gauche*$^+$ conformations for the α and γ torsions is shown by the thin line. Note that the backbone having a *trans* conformation (thick line) fits into the density better than the one having the *gauche* conformation (thin line). (b) Omit map for the fourth guanine residue in the orthorhombic form, which has the conventional *gauche*$^-$ and *gauche*$^+$ conformations for the torsions α and γ. Note that the distance between the C8 base atom and the phosphorus atom is smaller than that found for the *trans* conformation, as in a.

reflections between 6- and 2-Å resolution dropped the R value to 26.4%. At this stage, $3F_o - 2F_c$ omit maps (Figure 1a) clearly showed the differences in the base sequences in the model at positions 1, 3, 8, and 10 (strand 1) and 11, 13, 18, and 20 (strand 2), conforming to the correct base sequence d(GCGGGCCCGC). This was further confirmed by omitting only the incorrect base pairs (1/20, 3/18, 8/13, and 10/11) one base pair at a time and refining the rest of the structure. At each step, the correct base pair was introduced into the starting model based on $3F_o - 2F_c$ omit maps. Thus, the correct base sequence for the entire decamer d(GCGGGCCCGC) was confirmed. Further, to minimize any structural bias from the starting model, the decamer was subjected to simulated annealing refinement by heating the structure to 4000 K and slowly cooling to room temperature. This dropped the R value to 19.9%. Refinement of the model, including 101 water molecules located from difference maps using 5450 reflections ($F > 3\sigma(F)$), between 8- and 1.7-Å resolution converged to a final R value of 18.5%. $3F_o - 2F_c$ omit maps for some of the residues are shown in Figure 2. In the final model, the rms deviations from the ideal geometry (nucleic acid library file PARAM11.DNA) for bond lengths, angles, torsion angles, and improper deviations are 0.013 Å, 3.6°, 31.9°, and 3.1°, respectively.

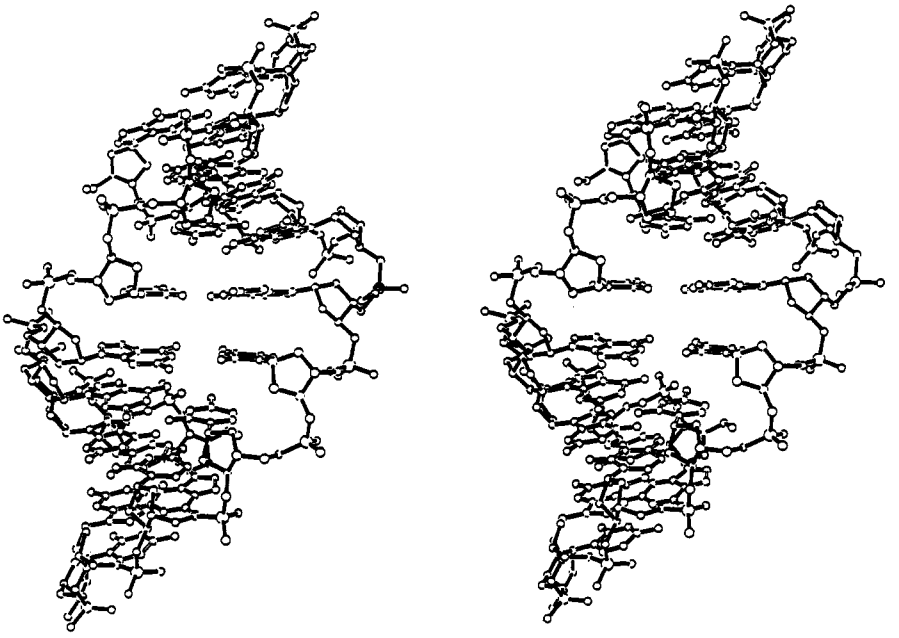
Hexagonal Crystal Form. Hexagonal pyramidal crystals were grown using conditions and methods similar to those used for the orthorhombic form, except that cobalt hexamine (0.5 mM) was used instead of spermine. The present crystallization conditions were quite different from those of its isomorphous crystals, d(ACCGGCCGGT) (Frederick et al., 1989). A crystal of size 0.5 × 0.5 × 0.3 mm was mounted in a thin-walled glass capillary for X-ray data collection with its axis perpendicular to the hexagonal face along with some mother liquid at one end. The hexagonal crystal has a space group of $P6_122$ or $P6_522$, with $a = b = 39.11$ and $c = 79.23$ Å. The molecular dyad coincides with the crystallographic 2-fold symmetry; thus, there is only one DNA strand in the asymmetric unit and the volume per base pair is 1749 Å^3, which is higher than the value for the orthorhombic form (see above) but very close to that (1773 Å^3) observed for the isomorphous decamer d(ACCGGCCGGT) (Frederick et al., 1989). X-ray diffraction data from 15 to 1.8 Å were collected using one ϕ scan (range 180°) at 0.25° steps, using the same experimental setup that was used for the orthorhombic crystal form. Each frame was collected for 90 s. Of a total of 19 222 reflections collected, 3435 were unique (91% of possible reflections) with an R_{sym} of 3.9%. Since the cell dimensions are very similar to those for the crystals of the A-DNA decamer d(ACCGGCCGGT) (Frederick et al., 1989) (Table I), the two crystals were assumed to be isomorphous with the space group $P6_122$.

Table II: Average Helix Parameters[a]

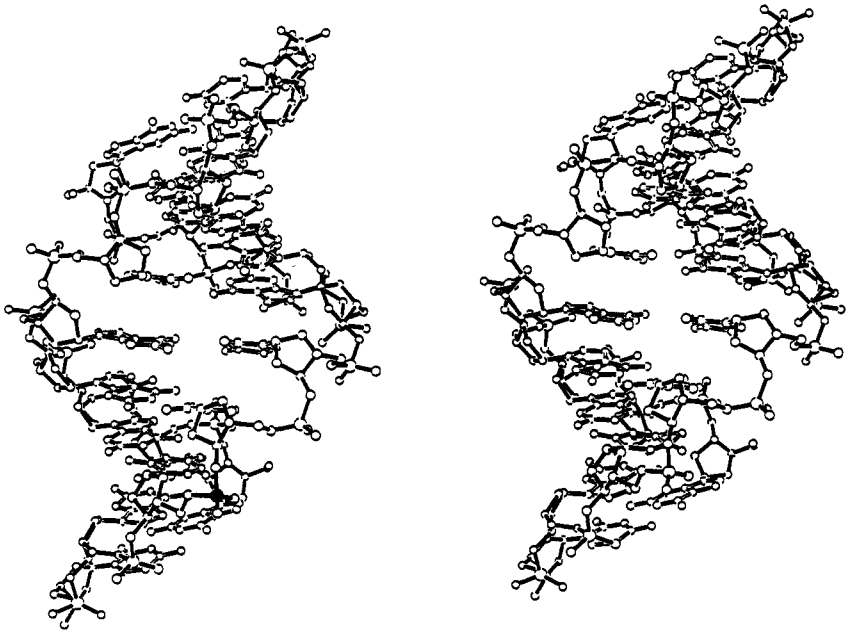
	orthorhombic	hexagonal
helical twist (deg)	32(3)	33(4)
rise/base pair (Å)	2.5(3)	2.7(2)
base-pair inclination (°)	16(3)	17(3)
slide (Å)	−1.7(3)	−1.2(3)
minor groove width (Å)	8.2–10.6	9.7–9.9
range and av value (σ)	9.8(8)	9.8(1)
major groove width (Å)	5.4	3.7

[a] All of the helix parameters are calculated using NEWHELIX 92 (R. E. Dickerson, private communication). Standard deviations are in parentheses on the least significant digit.

A rigid-body refinement was performed using the coordinates of the decamer d(ACCGGCCGGT) (Frederick et al., 1989) against the X-ray data using 192 reflections in the 15–5-Å resolution range, dropping the R value from 35% to 30% with an rms in atomic displacement of 0.61 Å. Positional and thermal parameter refinement dropped the R value to 25% for 2053 reflections between 6- and 2-Å resolution. At this stage $3F_o - 2F_c$ omit maps (Figure 1b) revealed the correct bases for the first, third, eighth, and tenth positions, matching the present decamer sequence d(GCGGGCCCGC) (Figure 1b). As was done in the orthorhombic structure, this was further confirmed by omitting only the base pair atoms, refining the remaining structure one base pair at a time, and introducing the correct base pairs at each step from $3F_o - 2F_c$ omit maps, thus confirming the present decamer sequence d(GCGGGCCCGC). The model having the correct base



(a)



(b)

FIGURE 3: Overall conformation of the decamer in the two crystal forms, (a) orthorhombic and (b) hexagonal (identical view looking into the minor groove; the helix axis is vertical). The orthorhombic structure is broad at the middle compared to the hexagonal structure. The backbone conformation is also different for the fifth residue in the orthorhombic form and the eighth residue in the hexagonal form (the respective phosphorus atoms have been shaded).

sequence was subjected to simulated annealing at 4000 K to eliminate any structural bias from the starting model, which dropped the R value further to 21%. Additional refinement using 2298 reflections ($F > 3\sigma(F)$) between 8- and 1.8- Å resolution gave a final R value of 18.3%, including 39 water molecules. The final model has rms deviations from the ideal geometry for bond lengths, angles, torsion angles, and improper deviations of 0.017, 3.6°, 31.9°, and 3.9°, respectively.

Even though different cations were used during crystallization, which is essential to obtain the different crystal forms, no bound cations were found in the structures. All refinement studies were carried out using XPLOR 2.1 (Brunger et al., 1987). The atomic coordinates for both crystal forms will be deposited in the Brookhaven Protein Data Bank. In the orthorhombic crystal form, the nucleotides are numbered 1–10 from 5′ to 3′ in one chain and 11–20 on the other chain, also in the 5′→3′ direction. In the hexagonal crystal form, there is only one strand in the asymmetric unit and the nucleotides are numbered similarly. The nomenclature and convention used for the torsion angles are in accord with the IUPAC-IUB (1983) recommendations, where the torsion angles of the six backbone bonds of the chain P–O5′–C5′–C4′–C3′–

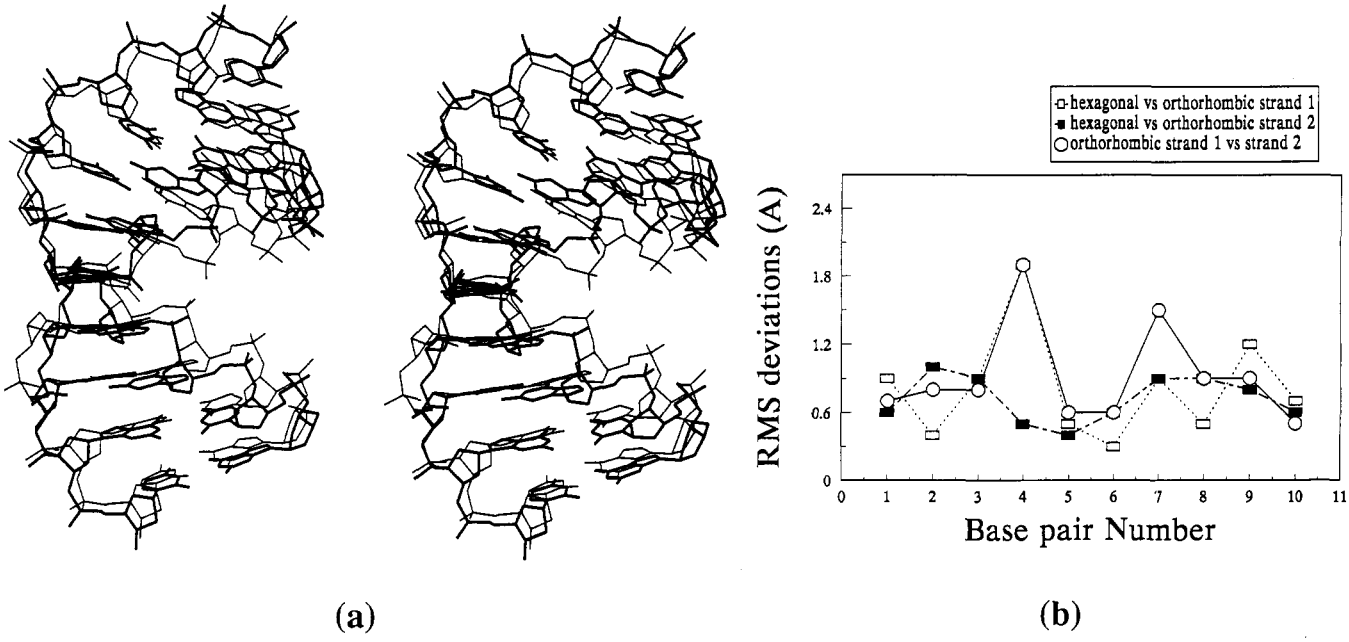


FIGURE 4: (a) Stereoview of the orthorhombic structure (thick line) superimposed on the hexagonal structure (thin line). The superposition was performed using the duplexes of the two crystal forms, and the mean deviation is 1.2 Å with an rms deviation of 1.5 Å. The sugar phosphate backbone atoms have large deviations between the two crystal forms (mean 1.4 Å, rms 1.7 Å) compared to the base pairs where the mean deviation is 0.9 Å and rms deviation is 1.1 Å. (b) Variation of the mean rms deviation between the bases of the strands for the two crystal forms. Note the large deviation for the fourth base of strand 1 of the orthorhombic form compared to the hexagonal form.

O3′–P are denoted as α, β, γ, δ, ϵ, and ζ, respectively.

RESULTS AND DISCUSSION

Comparison of the Orthorhombic and Hexagonal Polymorphous Crystal Forms. The overall helix parameters that characterize the DNA double helix, viz., twist, rise, base-pair inclination, and base-pair slide, are listed in Table II, and they are very similar in both of the polymorphic crystal forms. The DNA duplexes in both the crystal forms (Figure 3) have a narrow and deep major groove and a wide and flat minor groove, with the base pairs inclined to the helix axis characteristic of A-DNA. In the orthorhombic form the minor groove width varies from 8.0 to 10.6 Å (average 9.6(9) Å), being narrow at the terminals and broad in the middle; on the contrary, in the hexagonal form it is more regular with an average value of 9.9(1) Å (Figure 3). A similar characteristic variation in the minor groove size is observed in the respective isomorphous crystal forms of the decamers d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) and d(ACCGGCCGGT) (Frederick et al., 1989).

When the duplexes of the two crystal forms are superposed, the overall mean deviation is 1.2 Å, with an rms deviation of 1.5 Å (Figure 4a). The phosphate backbone in the two crystal forms shows quite large variations and has a mean deviation of 1.4 Å and an rms deviation of 1.7 Å. However, the 10 base pairs of the two crystal forms have a mean deviation of only 0.9 Å and an rms deviation of only 1.1 Å, which are much smaller than the values for the backbone. The detailed rms deviations for the individual bases are shown in Figure 4b. The fourth base pair of the two crystal forms shows large deviations (2.3 Å) when the base pairs of the two duplexes are superposed, but the fourth guanine base in strand 1 shows large deviations (1.9 Å) when the bases of the DNA strand of the hexagonal form are superimposed on the bases of the individual strands of the orthorhombic form. This indicates that in the orthorhombic structure the fourth guanine base in strand 1 is significantly displaced into the minor groove (Figure 4a) from the rest of the duplex and relative to the hexagonal crystal structure. Such a displacement of the fourth guanine base was also observed in the isomorphous crystal structure d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993).

In the orthorhombic form, the fifth residue adopts the nearly extended *trans* conformations for the backbone torsion angles α and γ about the bonds P–O5′ and C5′–C4′, respectively (Figure 2), whereas in the hexagonal form, the eighth residue adopts the nearly *trans–trans* conformation. These differences in the backbone torsions may arise from crystal packing (Jain & Sundaralingam, 1989). However, the fifth residue in the isomorphous orthorhombic A-DNA decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) and the eighth residue in the isomorphous hexagonal A-DNA decamer d(ACCGGCCGGT) (Frederick et al., 1989) adopt the extended *trans–trans* conformation. In both crystal forms, all of the sugars and all of the glycosyl torsion angles adopt the C3′-*endo* pucker and *anti* conformations, respectively.

Distortion in the Duplexes. The comparison of the helix parameters of the orthorhombic and hexagonal forms is shown in Figure 5. The linear correlation coefficients (denoted as R in the figures) between the helix parameters of the two crystal forms are low or even negative. For example, the correlation coefficient is negative for twist angles and is only 34% for inclination, implying that there is no local structural similarity between them. Thus, for the same DNA sequence, while the average helix parameters are nearly the same, the detailed structures in the two crystal forms are very different. This suggests that the observed variations arise mostly from the differences in the crystal packing.

When the components of unit vectors normal to the base-pair planes (x(AX) and y(AY)) of the decamers in the two crystal structures are compared with those of fiber A-DNA, the orthorhombic structure shows a distortion at the fourth and terminal base pairs whereas the hexagonal structure shows a distortion at the middle 4–7 base pairs. In the case of the orthorhombic structure, the distortion at the fourth base pair is due to the displacement of this base pair toward the minor

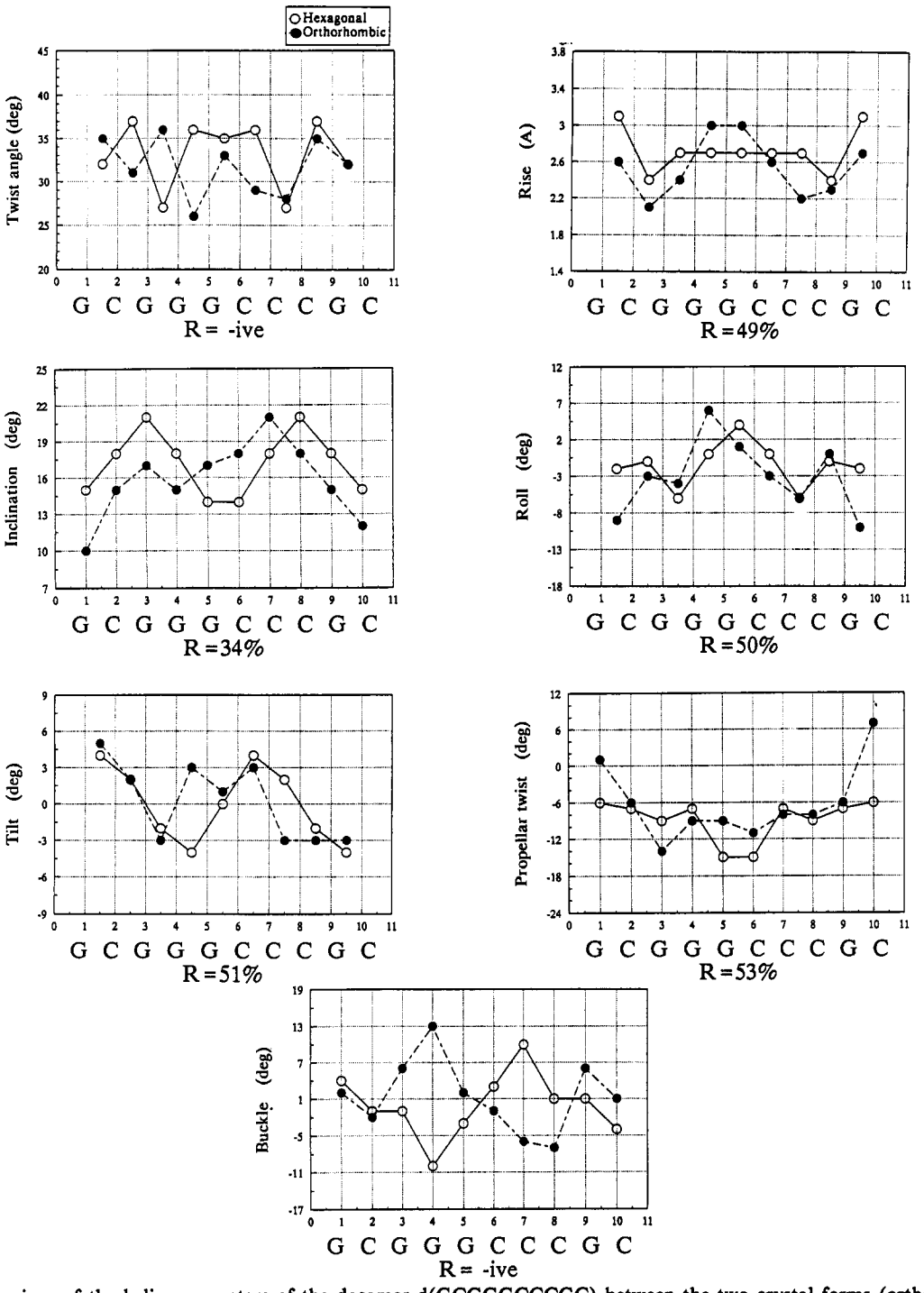


FIGURE 5: Comparison of the helix parameters of the decamer d(GCGGGCCCGC) between the two crystal forms (orthorhombic (●)/ hexagonal (○)). The linear correlation coefficients R for all of the parameters are low or negative. Each box along the y-axis represents one standard deviation (σ).

groove, and the distortion of the terminal base pairs could be due to the end effect arising from the intermolecular hydrogen-bonding and packing interactions to symmetry-related molecules (see below). In the hexagonal structure, the duplex is bent by 17° toward the major groove. A similar characteristic bend was also observed in the isomorphous crystal structure, d(ACCGGCCGGT) (Frederick et al., 1989). A bend in the dyad-related halves of the A-DNA duplex was first reported by Bingman et al. (1992a) in the crystal structure of the dodecamer d(CCGTACGTACGG). The components of the unit vectors normal to individual bases, instead of base pairs, of each strand show that strand 1 in the orthorhombic structure is more distorted than strand 2 and that the central region in the hexagonal structure shows a higher degree of distortion.

Crystal Packing in the Two Crystal Forms. Figure 6 shows the packing around the reference DNA duplex in the two crystal forms. In the orthorhombic form (Figure 6a), the reference molecule interacts with four neighboring molecules: two with the terminal base pairs and the remaining two with the body of the reference molecule on the minor groove side. Similar packing interactions are also seen in the hexagonal form (Figure 6b). It should be mentioned that, in the hexagonal form, there is only one DNA strand in the asymmetric unit: the complementary strand related by the molecular dyad coincides with the crystal dyad symmetry. In the orthorhombic form, the two symmetry-related molecules in the minor groove interact with the body of the reference duplex and are asymmetrically displaced away from the center

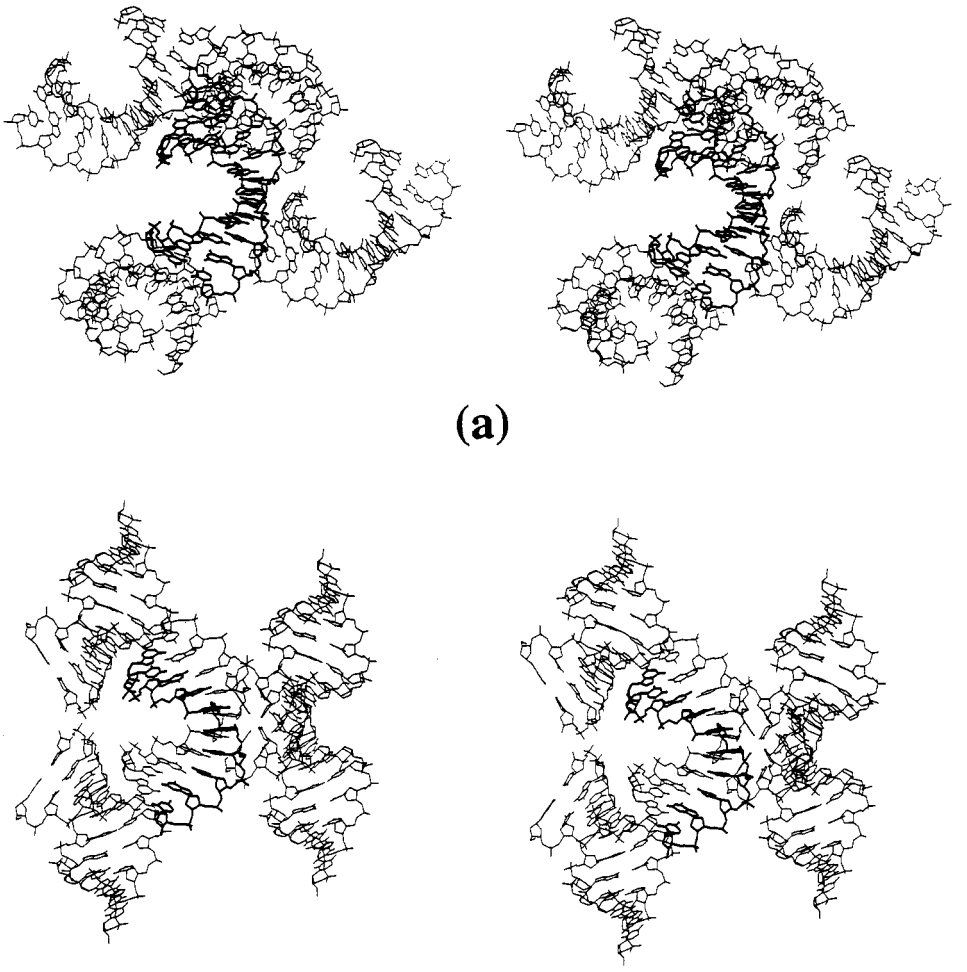


FIGURE 6: Packing environment in the two polymorphic crystal forms: (a) orthorhombic and (b) hexagonal. The reference molecule (duplex) interacts with four neighboring molecules (duplexes) (thin line) in each case. In the orthorhombic form, the two molecules that interact with the body of the reference duplex are asymmetrically tilted away from the center of the duplex and toward the terminal. In the hexagonal form, these two dyad symmetry-related duplexes are crowded around the center of the reference duplex in the minor groove.

of the duplex and toward the termini. In contrast, in the hexagonal form the two dyad symmetry-related duplexes crowd around the center of the reference duplex in the minor groove. As in the present hexagonal form, in the crystal structures of the dodecamers d(CCGTACGTACGG) (Bingman et al., 1992a) and d(GCGTACGTACGC) (Bingman et al., 1992b), two symmetry-related molecules are involved in intermolecular interactions with the central region of the reference molecule; coincidentally, these crystals also belong to the hexagonal system with the same space group ($P6_122$). Such intermolecular crowding around the center of the DNA duplex in the minor groove could be responsible for the bending of the two halves of the duplex toward the major groove both in the present hexagonal decamer and the two dodecamers, d(CCGTACGTACGG) (Bingman et al., 1992a) and d(GCGTACGTACGC) (Bingman et al., 1992b).

Intermolecular Hydrogen-Bonding Interactions: Orthorhombic Structures. As can be seen from Figure 7a, in the top half of the present orthorhombic structure the N2 and N3 atoms of the fourth guanine base form two hydrogen bonds, one with O4′ and the other with N2 of the symmetry-related C(12) and G(11) residues, respectively [the distances and angles are ad follows: G(4)N3···N2G(11), 3.0 Å; G(11)-C2–N2···N3G(4), 120°; G(4)N2···O4′C(12), 3.2 Å; G(4)-C2–N2···O4′C(12), 114°]. Also, the N2 atom of the eleventh symmetry-related guanine base forms a somewhat weaker hydrogen bond with the O2 atom of the C(18) base [distance, G(12)N2···O2C(18) 3.0 Å; angle, G(11)C2–N2···O2C(18) 100°]. The N2 atom of the third residue also forms a hydrogen bond with the O2 atom of the symmetry-related C(10) base [distance, G(3)N2···O2C(10) 2.9 Å; angle, G(3)C2–N2···O2C-(10) 139°]. Note that the hydrogen-bonding scheme observed in the present structure involving the fourth guanine base is quite different from that observed in d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) (Figure 7b). This is attributable to the difference in the first base pair (G·C vs C·G) which does not favor the novel base-paired triplet. Also, in the present orthorhombic decamer both the G(11) and C(12) bases interact with the base G(4), whereas in the related orthorhombic decamer only the terminal base pair G(10)·C(11) interacts with G(4). However, the resulting influence on the duplex is similar in both of the orthorhombic crystal structures.

In the bottom half of the present orthorhombic structure, the N3 atom of the symmetry-related base G(1) makes a somewhat weaker hydrogen bond with the N2 atom of G(15) (Figure 8a) [distance, G(15)N2···N3G(1) 3.3 Å; angle, G(15)-C2–N2···N3G(1) 118°]. Also, the N2 atom of the same G(1) guanine residue forms a weaker hydrogen bond with the O2 atom of C(7) [distance, G(1)N2···O2C(7) 3.3 Å; angle, G(1)-

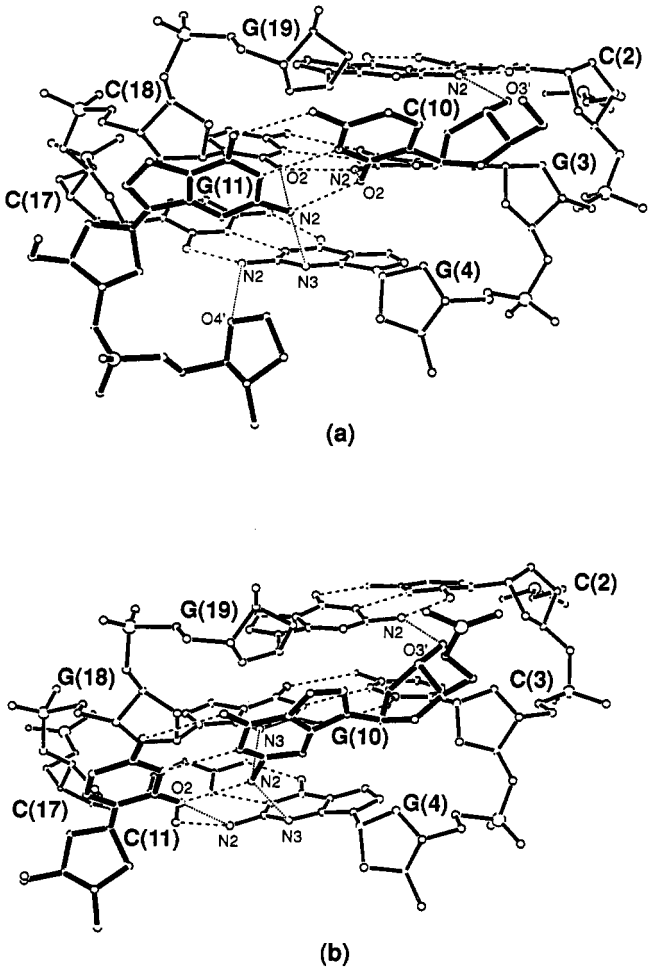


FIGURE 7: Hydrogen-bonding interactions in the top half of the orthorhombic crystal form. (a) In the top half of the present orthorhombic crystal form, the O2 of C(10) and the N2 of G(11) of the symmetry-related terminal base pair G(11)·C(10) are hydrogen-bonded with the N2 of G(3) and the N3 of G(4), respectively. Also, the N2 atom of the G(4) forms a hydrogen bond with the O4′ atom of the same symmetry-related C(12) residue. (b) Intermolecular hydrogen-bonding interactions in the top half of the isomorphous orthorhombic form of the decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993).

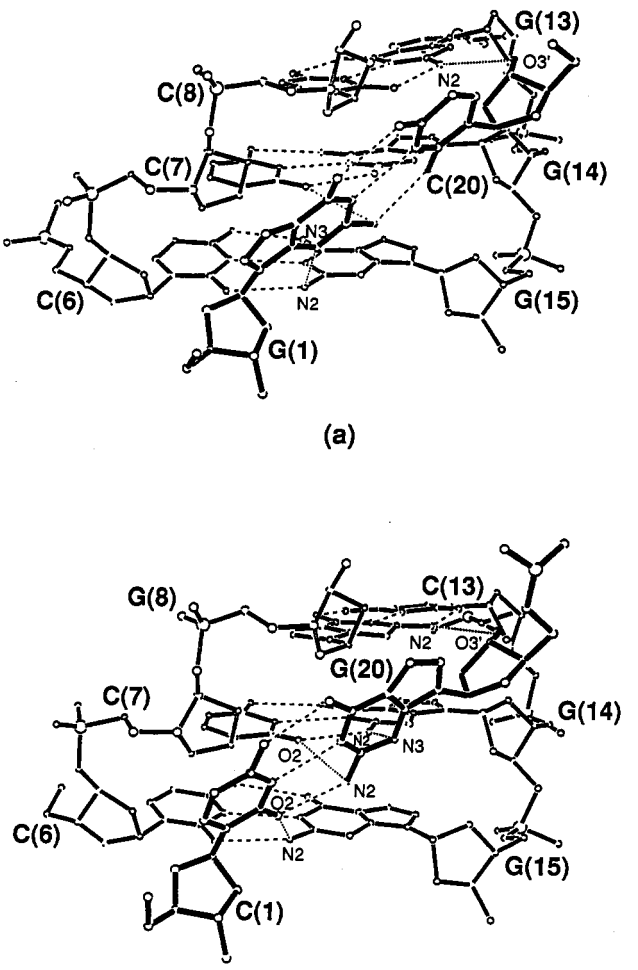


FIGURE 8: Hydrogen-bonding interactions in the bottom half of the orthorhombic crystal form. (a) In the bottom half of the present orthorhombic crystal form, the N2 and N3 atoms of the symmetry-related base G(1) form somewhat weaker hydrogen bonds with O2 of C(7) and N2 of G(14), respectively. (b) Intermolecular hydrogen-bonding interactions in the bottom half of the isomorphous orthorhombic form of the decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993).

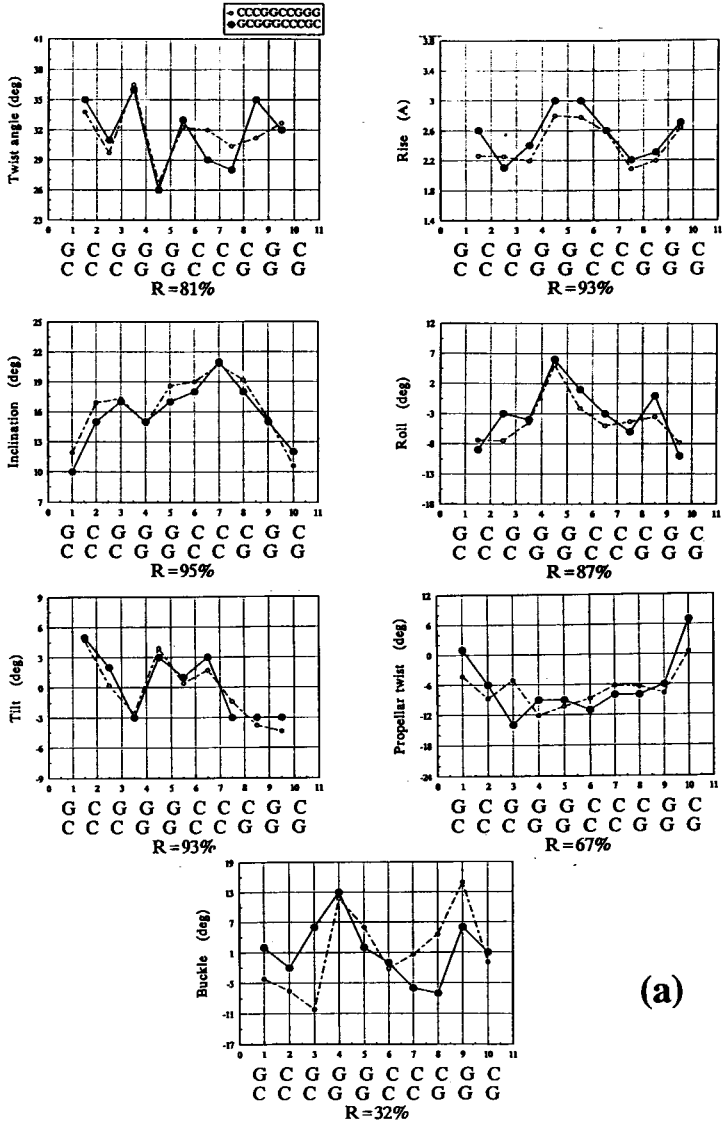
C2–N2···O2C(7) 90°]. Again, this is quite different from the hydrogen-bonding scheme observed in the isomorphous orthorhombic structure d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) (Figure 8b).

Hexagonal Structures. In the hexagonal crystal structure, the N2 atoms of the bases G(4) and G(5) form two hydrogen bonds with the N3 and O4′ atoms, respectively, of the symmetry-related residue G(1) [distance G(4)N2···O4′G(1) 3.0 Å; angle, G(4)C2–N2···O4′G(1) 145°; distance, G(5)-N2···N3G(1) 3.1 Å; angle, G(5)C2–N2···N3G(1) 115°]. In the isomorphous hexagonal decamer d(ACCGGCCGGT) (Frederick et al., 1989), the hydrogen-bonding scheme is the same, except that the symmetry-related base is now A(1).

In the present hexagonal crystal form, the terminal O3′ hydroxyl group of the C(10) residue is not involved in any hydrogen bonds, but in the orthorhombic crystal form, it forms a hydrogen bond with the N2 atom of the symmetry-related guanine base G(19) of the C(2)·G(19) base pair. Similarly, the other O3′ hydroxyl group of C(20) of the orthorhombic crystal forms a hydrogen bond with the N2 atom of the symmetry-related guanine base G(13) of the C(8)·G(13) base pair [distance C(10)O3′···N2G(19) 2.9 Å; angle, C(19)C2–N2···O3′C(10) 132°; distance, C(20)O3′···N2G(13) 3.0 Å; angle, G(13)C2–N2···O3′C(20) 144°].

Comparison of the Polymorphic Crystal Forms with the Isomorphous Crystal Forms. In order to substantiate our observation that crystal packing influences DNA conformation, it is necessary to compare the helix parameters of the two polymorphic forms having the same base sequence with those of the respective isomorphous crystal forms having different base sequences. As mentioned, the present orthorhombic crystal form is isomorphous with the A-DNA decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993), and the present hexagonal form is isomorphous with the A-DNA decamer d(ACCGGCCGGT) (Frederick et al., 1989). The isomorphous A-DNA decamer d(CCCGGC-CGGG) (Ramakrishnan & Sundaralinagm 1993) was determined at resolution comparable (1.65 Å) to that of the present orthorhombic form (1.7 Å), and the isomorphous hexagonal A-DNA decamer d(ACCGGCCGGT) (Frederick et al., 1989) was determined at a slightly lower resolution (2.0 Å) than was the present hexagonal form (1.8 Å). Therefore, the comparison is made between structures that have been determined at similar resolutions or accuracies.

The linear correlation coefficients (R) between the helix parameters of the pairs of isomorphous crystal structures are



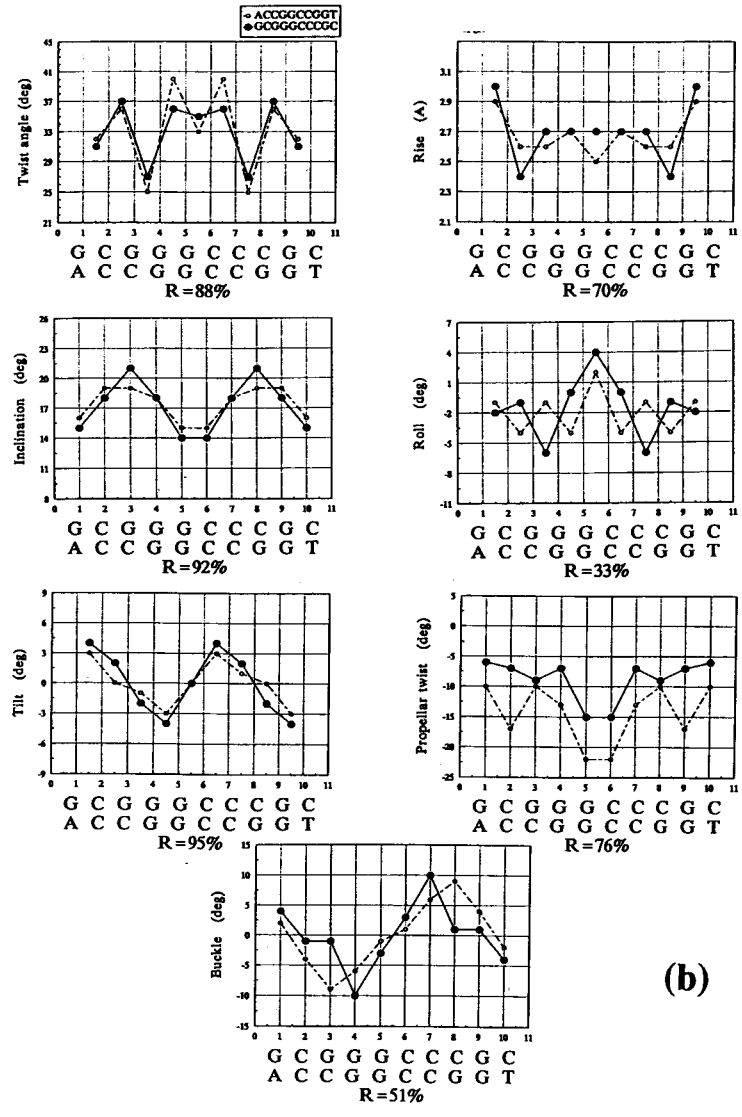


FIGURE 9: Helix parameter comparisons for the two crystal forms. (a) Comparison of the helix parameters for the isomorphous orthorhombic crystal forms: present work d(GCGGGCCCGC) (●) and d(CCCGGCCGGG) (○). The correlation coefficient for the buckle increases to 82% when the values for the third and eighth base pairs are omitted (effect due to the sequence change; see text). Similarly, the correlation coefficient for the propeller twist increases to 87% when the value for the third base is omitted (due to intermolecular interactions). (b) Comparison of the helix parameters for the isomorphous hexagonal crystal forms: present work d(GCGGGCCCGC) (●) and d(ACCGGCCGGT) (○). The poor correlation coefficient for the roll increases to 90% when the value for the 3–4 base-pair step is omitted. The correlation coefficient for the propeller twist increases to 97% when the value for the second base pair is omitted (sequence change). Similarly, the correlation coefficient for buckle increases to 89% when the values for the third and the eighth base pairs are omitted (same reason as above).

quite high (Figure 9a,b), even though the base sequences are different. For example, in the orthorhombic crystal form, the correlation coefficient for the twist angles is 81% and for the inclinations it is 95%; in the hexagonal crystal form they are 88% and 92%, respectively. This clearly indicates that, in crystals, the crystal forces largely determine the local helix parameters of the DNA double helix rather than the base sequence.

However, sequence-dependent variations are seen in some helix parameters, particularly the base-pair buckle for the orthorhombic form and the base-pair buckle and propeller twist for the hexagonal form. In the orthorhombic forms (Figure 9a), the buckle reverses its sense (positive value to negative value) for the third base pair, where the base sequence changes from C(2)**G(3)**G(4) in the present structure, d(GCGGGCCCGC), to C(2)**C(3)**G(4) in d(CCCGGC-CGGG). A similar trend is also seen in the hexagonal form at the same position (third base pair) (Figure 9b) where the sequence changes from C(2)**G(3)**G(4) in d(GCGGGCCCGC) to C(2)**C(3)**G(4) in d(ACCGGCCGGT). Furthermore, in the hexagonal structures, the propeller twist at the second C·G base pair shows a large variation where the base sequence changes from G(1)C(2) to A(1)C(2). Also, as per Calladine (1982), the N2 atom of G(1) generates additional steric interactions with the second C(2)·G(9) base pair.

Intermolecular Base Hydrogen Bonds Modify the Helix Parameters. Having shown that the crystal forces modify the helix parameters of DNA, we observed that the similarities within the isomorphous crystals and the dissimilarities within the polymorphous crystals in the helix parameters might be explained by the similarities or differences in the intermolecular interactions within the respective (isomorphous and polymorphous) crystals. When only the base pairs of the two crystal forms are superposed, the fourth guanine base in strand 1 of the orthorhombic crystal form shows a larger deviation than the rest of the base pairs (Figure 4b). A similar large deviation was observed in the related orthorhombic structure of the decamer d(CCCGGCCGGG) (Ramakrishnan & Sundaralingam, 1993) and was attributed to the intermolecular hydrogen-bonding interactions between the fourth guanine base and a 2-fold screw symmetry-related C(11)·G(10) base pair. Also, in the orthorhombic crystal structures, the spikes observed in the helix parameters (Figure 9a) for inclination, roll, and tilt arise from the sliding displacement of the fourth base toward the minor groove compared to the rest of the duplex.

In the orthorhombic form, the base pairs of the fragments 2–4 and 6–8 are involved in direct intermolecular hydrogen-bonding interactions with the last (C·G) and first (G·C) base pairs of neighboring molecules, respectively. It is possible that the variation introduced in the helical twist angles of the second, third, and fourth base pairs to promote intermolecular hydrogen bonds is compensated by the low twist angle of 26°, involving the fifth base pair G·C (which is not involved in any intermolecular interactions), so that the average twist angle of about 33° is maintained. Due to the weak hydrogen bonds involving the sixth and seventh base pairs (Figure 7a) and the strong hydrogen bond involving the eighth base pair, a large variation in the twist angles is observed involving the sixth and seventh base pairs. Similarly, in the hexagonal crystal structure, the fragment 4–7 is involved in intermolecular hydrogen bonding to compensate for the deviations in the twist angles involving the second, third, eighth, and ninth base pairs, which are devoid of any direct intermolecular hydrogen-bonding interactions. In general, the twist angle involving the base pairs not involved in hydrogen bonding that succeed or precede the fragment involved in the intermolecular hydrogen bonds assumes a low value, so that the average value is close to 33°.

Hydration of the DNA in the Two Crystal Forms. As mentioned earlier, the hexagonal form is more hydrated than the orthorhombic form on the basis of the volume per base pair. It has been reported that, in contrast to fiber DNA, A-DNA crystal unit cells are more hydrated than the B-DNA crystal unit cells (Ramakrishnan & Sundaralingam, 1993; Bingman et al., 1992a). In the hexagonal crystal, the oligomers organize in such a way around the crystallographic 6_1 screw that there is a wide solvent channel nearly 20 Å in diameter running down the *c*-axis about the 6_1 screw, which is responsible for the high volume per base pair and the larger amount of water. However, no such wide channel is found in the orthorhombic form. Similar solvent channels are present in the hexagonal crystal structures of the octamer d(G-GUAUACC) (Doucet et al., 1889) and the dodecamer d(CCGTACGTACGG) (Bingman et al., 1992a). In both of the latter structures, a disordered B-DNA of the same sequence as the A-DNA fills the water hole and coexists with the A-DNA in the crystal. However, in the present hexagonal form there is no evidence for B-DNA coexisting with the A-DNA. Even though in the hexagonal form the volume per base pair is higher than that in the orthorhombic form, the number of water molecules located (7.8 per base pair) is lower than the number in the orthorhombic form (10.1 per base pair), probably because of the degree of structural order in the highly hydrated crystal, which in turn contributes to the lower resolution of the X-ray data observed. Of the 101 water molecules located in the orthorhombic form, 74 are directly bound to the oligomer, whereas of the only 39 water molecules located in the hexagonal form, 27 are bound to the DNA oligomer. Neither structure contains a spine of hydration or other striking motif.

CONCLUSION

We have made a detailed comparison of the polymorphous crystal forms of the same DNA sequence among themselves, as well as with the isomorphous crystal structures with differing base sequences. The similar conformations observed for different base sequences in the same crystal form and different conformations for the same base sequence in different crystal forms show that crystal packing interactions dominate base sequence effects. This simply demonstrates the inherent conformational flexibility of DNA to adapt to the environment and assume a variety of perturbed conformations. If the C3′-*endo*-puckered A-DNA, which is thought to be more rigid compared to the C2′-*endo*-puckered B-DNA, can exhibit such large base-pair displacements and distortions under the influence of intermolecular interactions, B-DNA can be expected to be distorted similarly or even more.

NOTE ADDED IN PROOF

Our recent comparative study of the isoforms of the tetragonal and hexagonal family of A-DNA octamers with different base sequences further confirms the present findings (Ramakrishnan, B., & Sundaralingam, M. (1993) *J. Biomol. Struct. Dyn. 11*, 11–26).

REFERENCES

Bingman C., Zon, G., & Sundaralinagm M. (1992a) *J. Mol. Biol. 227*, 738–756.

Bingman C., Jain, S., Zon, G., & Sundaralingam, M. (1992b) *Nucleic Acids Res. 20*, 6637–6647.

Brunger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.
Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.
Crawford, J. L., Kolpak, F. J., Wang, A. H.-J., Quigley, G. J., van Boom, J. H., van der Marel, G. A., & Rich , A. (1980) *Proc.Natl.Acad.Sci. U.S.A. 77*, 4016–4020.
Dickerson, R. E., Goodsell, D. S., Kopka, M. L., & Pjura, P. E. (1987) *J. Biomol. Struct. Dyn. 5*, 557–579.
DiGabriele, A. D., Sanderson, M. R., & Steitz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.
Doucet, J., Benoit, J. P., Cruse, W. B., Prange, T., & Kennard. O. (1989) *Nature (London) 337*, 190–192.
Drew, H. R., Takano, T., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 286*, 567–573.
Frederick, C. A., Quigley, G. J., Teng, M.-K., Coll, M., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Eur. J. Biochem. 181*, 295–307.
Howard, A. J., Nielsen, C., & Xuong, N. H. (1985) *Methods Enzmol. 114*, 211–237.
IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) *Eur. J. Biochem. 131*, 9–15.
Jain, S., & Sundaralingam, M. (1989) *J. Biol. Chem. 264*, 12780–12784.
Jain, S., Zon, G., & Sundaralingam, M. (1991) *Biochemistry 30*, 3567–3576.
Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J., & Dickerson, R.E. (1993) *Biochemistry 32*, 1373–1389.
Ramakrishnan, B., & Sundaralingam, M. (1993) *J. Mol. Biol. 231*, 431–444.
Shakked, Z., Guerstein-Guzikevich, G., Frolow, F., & Rabinovich, D. (1988) *Nature 342*, 456–460.