Acta Cryst. (1994). D**50**, 192–196

Space-Group Degeneracy in the Packing of a Non-Selfcomplementary Z-DNA Hexamer

BY C. SADASIVAN, P. KARTHE AND N. GAUTHAM*

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

(*Received* 25 *May* 1993; *accepted* 8 *October* 1993)

Abstract

The X-ray diffraction pattern of the crystals of the non-selfcomplementary hexadeoxyribonucleotide d(CGCACG)·d(CGTGCG) can be indexed in four different space groups: (i) $P6_5$ and $P2_1$, with cell parameters $a = 17.75\,(1)$, $b = 17.76\,(1)$, $c = 42.77\,(3)$ Å, $\alpha = 90$, $\beta = 90$, $\gamma = 120°$, and (ii) $P2_12_12_1$ and $C2$, with cell parameters $a = 17.75\,(1)$, $b = 30.74\,(2)$, $c = 42.77\,(3)$ Å, $\alpha = 90$, $\beta = 90$, $\gamma = 90°$. While the R_{merge} for the equivalent reflections in the different space groups indicates that $P2_1$ is the correct choice in the present case, it is demonstrated that the near degeneracy of the space groups arises out of the fact that the DNA molecule is nearly cylindrical. A perfect cylinder would show perfect degeneracy.

Introduction

Over the past one and a half decades, crystallography of short mainly selfcomplementary DNA fragments has shown that DNA helices 4–12 base pairs long pack in a few recurrent motifs in the crystals of oligodeoxyribonucleotides (Dock-Bregeon & Moras, 1992). The selection of motifs appears to be chiefly based on the helix type (*i.e.* A-, B- or Z-type DNA). A-type oligomers pack such that the terminal base pair of one helix stacks on the flat minor groove of a neighbouring helix (Wang, Fujii, van Boom & Rich, 1982; Frederick *et al.*, 1989; Jain & Sundaralingam, 1989; Eisenstein, Frolow, Shakked & Rabinovich, 1990). B-type helices pack in three different ways – with groove–groove contacts as in the structure of d(CGCGAATTCGCG) (Wing *et al.*, 1980), with groove–backbone contacts as in the non-selfcomplementary duplex d(ACCGGCGCCACA)·d(TGTGGCGCCGGT) (Timsit, Westhof, Fuchs & Moras, 1989) and with backbone–backbone contacts as in the decamer d(CCAAGATTGG) (Privé *et al.*, 1987). In all three patterns, the oligomers pack end-to-end to form infinite fibre-like DNA chains. The packing of Z-DNA oligomers (Wang *et al.*, 1979; Brennan & Sundaralingam, 1985; Brennan, Westhof & Sundaralingam, 1986) is very similar to that of the B-DNA decamers in that the molecules stack to form continuous helices with backbone–backbone contacts. These patterns, in general, do not appear in crystals of complexes of the oligomers with intercalating drugs or other such ligands.

Preferences in the mode of interhelical interactions appear to lead to space-group preferences. A-type helices crystallize either in hexagonal space groups $P6_1$ or $P6_122$ or in tetragonal space groups $P4_32_12$ or $P4_3$. There is also one instance reported of A-type packing in the space group $P2_12_12_1$ (Wang *et al.*, 1982). B-type helices crystallize in space groups $P2_12_12_1$, $R3$, $P3_221$ or $C2$. Z-DNA helices have been reported so far in space group $P2_12_12_1$ except in the case of the tetramer d(CGCG) in $C222_1$ (Drew, Takano, Tanaka, Itakura & Dickerson, 1980) and of the disordered structures of the octamers d(CGCGCGCG) and d(CGCATGCG), and the decamer d(CGTACGTACG), all the three of which crystallize in $P6_5$ (Fujii *et al.*, 1985; Brennan & Sundaralingam, 1985).

Crystals of Z-DNA hexamers are the best ordered, as indicated by the resolution of the X-ray data, which is seldom poorer than 1.5 Å for the d(CGCGCG) family sequences (Fujii *et al.*, 1985). The presence of A·T base pairs unbalanced by other factors such as methylation of the cytosine bases (Fujii, Wang, van der Marel, van Boom & Rich, 1982), and of rare bases such as ^{Br}U (Brown, Kneale, Hunter & Kennard, 1986), perturbs the packing and makes the resolution worse.

As part of our studies on the factors influencing packing of molecules, and in order to elucidate the effect on the packing of A·T base pairs in Z-DNA duplexes and of non-selfcomplementarity of the two strands of the helix, we have grown and analysed crystals of the oligonucleotide d(CGCACG) duplexed with its complement d(CGTGCG). We report here an interesting space-group degeneracy in these crystals, arising out of the possibility of

* Author for correspondence.

describing, to a first approximation, the packing of DNA helices in terms of the packing of cylinders.

Experimental

The hexanucleotides were synthesized by the phosphoramidite method on an Applied Biosystems 381 DNA synthesizer. Well formed crystals were grown in about eight weeks at room temperature (296 K) by vapour diffusion from a hanging droplet containing 1 m*M* DNA, 50 m*M* sodium cacodylate at pH 6.8, and 10 m*M* $BaCl_2$ equilibrated against 25% 2-methyl-2,4-pentanediol in the reservoir. A crystal of dimensions $0.5 \times 0.2 \times 0.2$ mm was mounted in a glass capillary and used for the data collection. The crystal diffracted up to a nominal resolution of 2.5 Å. Lattice parameters were obtained from a least-squares fit of 20 reflections within the range $6 \le \theta \le 13°$.

The X-ray diffraction pattern for these crystals could be indexed in four different ways, all of them compatible with nearly the same packing mode of the DNA helices. The data were collected over a complete hemisphere, $-7 \le h \le 7$, $-12 \le k \le 12$, $0 \le l \le 17$, in the larger of the two possible sets of cell parameters (see Table 1) on an Enraf–Nonius CAD-4 diffractometer with graphite-monochromated Cu $K\alpha$ radiation ($\lambda = 1.5418$ Å) at 40 kV, 32 mA rating. Three standard reflections were monitored every hour for crystal decay and showed no significant change in intensity over the period of data collection. Data were corrected for Lorentz and polarization factors. An empirical absorption correction (North, Phillips & Mathews, 1968) was also applied with minimum and maximum correction factors 0.75 and 1.0, respectively.

Determination of the space group

As indicated in Table 1, the X-ray data could be indexed in two sets of cell dimensions and in four different space groups. Set I is compatible with space group $P6_5$ (or $P6_1$). However, the volume of the cell is insufficient to accommodate six hexamer duplexes. The space group $P2_1$ with c as the unique axis is also compatible with this set. The volume of the cell would allow the presence of two hexamer duplexes. Table 1 also gives a second set of possible cell parameters. The transformation of the reciprocal cell from set I to set II is given by the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Previous reports (Fujii *et al.*, 1985; Kennard & Hunter, 1989) have shown that the family of Z-DNA hexamers whose 'native' sequence is d(CGCGCG) all crystallize in the orthorhombic space group $P2_12_12_1$ with cell dimensions very close to those given by set II. The X-ray pattern in the present example could also be indexed in the space group $P2_12_12_1$. The diffraction data for the crystal collected for the hemisphere of reflections in the larger of the two possible cells show distinct C centring. The volume of the orthorhombic cell is insufficient to pack one hexamer duplex in the asymmetric unit if a space group with orthorhombic C-centred symmetry were to be chosen. A C-centred monoclinic cell, however, is a possibility and the X-ray data are consistent with the choice of the space group $C2$.

Table 1. *The two possible cells and four possible space groups*

	Cell parameters (Å, °)	Absences	Space group	R_{merge}
Set 1	$a = 17.75$			
	$b = 17.76$		$P2_1$	0.059
	$c = 42.77$	$0\,0\,l,\ l \ne 6n$		
	$\alpha = 90$			
	$\beta = 90$		$P6_1$	0.092
	$\gamma = 120$			
Set 2	$a = 17.75$			
	$b = 30.74$	$h\,0\,0,\ h = 2n+1$	$P2_12_12_1$	0.087
	$c = 42.77$	$0\,k\,0,\ k = 2n+1$		
	$\alpha = 90$	$0\,0\,l,\ l = 2n+1$		
	$\beta = 90$	$h\,k\,l,$	$C2$	0.077
	$\gamma = 90$	$h + k = 2n+1$		

Residuals obtained by merging equivalents in the monoclinic, orthorhombic and hexagonal systems (after making appropriate geometrical transformations) are also given in Table 1. Fig. 1 is a plot of R_{merge} in shells of $\sin\theta/\lambda$ as a function of resolution. Only data with $I > 1\sigma(I)$ were used in these calculations. Overall and at low resolution the data clearly indicate $P2_1$ to be the correct choice. Throughout the resolution range, hexagonal $P6_5$ has the worst of the residuals. The two other possible space groups $P2_12_12_1$ and $C2$ also have higher R_{merge} values, especially in the low-resolution shells. The correct

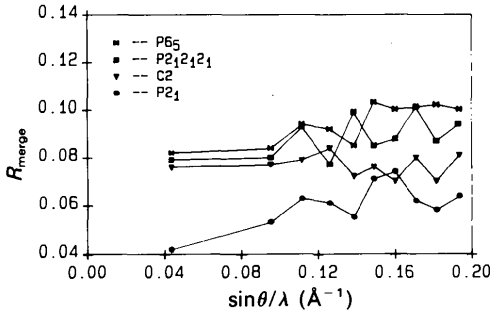

Fig. 1. The R_{merge} for equivalent reflections as a function of resolution. $R_{merge} = (\sum||F| - \langle F\rangle|)/(\sum F|)$, where $\langle F\rangle$ is average over multiple observations of the symmetry-related reflections.

space group is thus $P2_1$ (c axis unique) with a = 17.75 (1), b = 17.76 (1), c = 42.77 (3) Å, γ = 120.05 (3)°. The solution of the structure, described below, and its partial refinement in $P2_1$, confirmed the choice of the cell and space group.

Structure solution and molecular packing

All four space groups indicated by the diffraction pattern could be explained on the basis of a single packing mode. The hexagonal cell indicated by the first set of cell parameters would accommodate a model for packing in which Z-type hexamers are located exactly one above the other to form a continuous helix with its axis coincident with the crystallographic c axis and possessing the 6_5 symmetry that is present in Z-DNA. The unit cell would then contain two hexameric units. The space group $P6_5$ can be ascribed to the crystal only approximately due to the absence of the phosphate group between the two hexamers and the presence of an A·T base pair where C·G would be required for the perfect symmetry. An implication diagram (Buerger, 1946) calculated in the space group $P6_5$ nevertheless showed the approximation to be remarkably good, since the helices could be easily indentified in this projection.

The two hexamers in the unit cell could stack along the c axis in one of the two ways shown in Fig. 2. Both these models would lead to an approximate 6_5 symmetry. Pattern (i) would also be in exact consonance with space group $P2_1$, with the duplex in the asymmetric unit related to the adjacent one along the c axis by an exact 2_1 screw.

Orthorhombic $P2_12_12_1$ is another space group in which the diffraction pattern can be indexed and which is compatible with the same packing mode (Fig. 3). The difference between the packing of the helices in space group $P2_1$ and in space group $P2_12_12_1$ lies in the relative directions of the axes of neighbouring helical columns. In $P2_1$ all columns would be aligned in parallel, *i.e.* with the A·T base pair occurring at the fourth and tenth position of each 12-mer. In $P2_12_12_1$ the columns of helices marked A (Fig. 3) would have the A·T base pairs at the fourth and tenth position of the 12-mer, whilst in the columns marked B, the helices run in the opposite direction with the A·T base pairs at the third and the ninth positions. Thus, the columns A would be anti-parallel to the columns B. An additional difference arises out of alignment of the helices along the c axis. If the hexamers are aligned as shown in Fig. 4(*a*), with the terminal base pairs in all the columns in the same ab plane, both $P2_1$ and $P2_12_12_1$ are possible. If however they are arranged as shown in Fig. 4(*b*), only the space group $P2_12_12_1$ is possible. This is the situation, for example, in the structure of d(CGCGCG) (Wang *et al.*, 1979). The present packing mode is also consistent with the C-centred monoclinic cell, provided the stacking of the two duplexes along the c axis is as shown in pattern (ii) (Fig. 2) with hexamers packed in alignment as illustrated in Fig. 4(*a*).

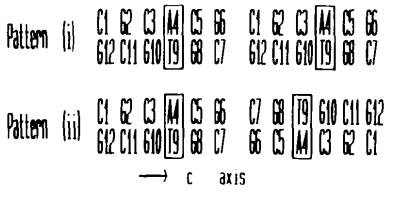

Fig. 2. The two possible modes of stacking of the hexamer duplexes.

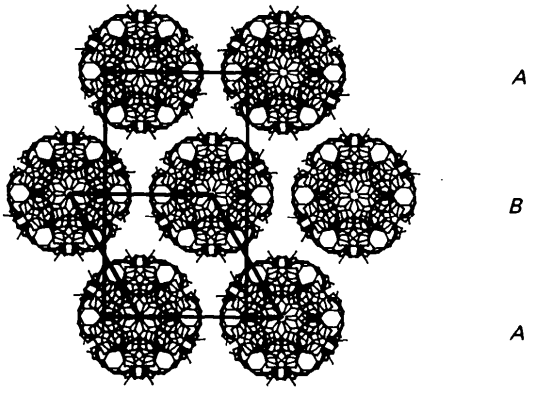

Fig. 3. The two possible unit cells (view down the c axis).

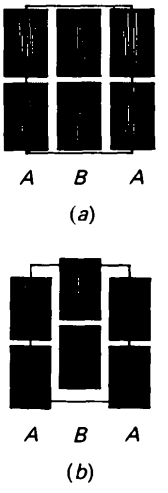

Fig. 4. (*a*), (*b*) Schematic illustrations of the alignment of the helices compatible with the different possible space groups (view perpendicular to the c axis).

Thus, the space-group degeneracy is explained by the packing pattern. Since the symmetry of the reciprocal lattice clearly indicated $P2_1$ as the correct choice, a model of the structure in this space group was constructed from the coordinates of d(pCpG) (Wang *et al.*, 1981) and positioned in the cell after taking into consideration the packing principles described above. The orientation of the molecule around the z axis remained to be determined. The model was rotated in 5° steps about the z axis and the R factor for the data in the 12–8 Å resolution shell was calculated each time. The position with the lowest R factor was then subjected to rigid-body refinement using *CORELS* (Sussman, Holbrook, Church & Kim, 1977). Subsequent refinement was carried out using *NUCLSQ* (Westhof, Dumas & Moras, 1985). All computations were made on a MicroVax II computer system. The R factor is 0.25 for 491 [with $F_o \geq 2\sigma(F_o)$] reflections up to 2.5 Å resolution. The correlation coefficient* is 0.90.

At the present level of refinement, which is adequate to unambiguously establish the packing model though not the finer details of the molecular structure, the overall conformation of the molecule appears to be closer to the Z-I type structure than to the Z-II type structure (Wang *et al.*, 1981). The phosphate groups at the purine–pyrimidine steps, however, are further away from the helix axis than in Z-I. The helix is slightly compressed compared with that in both Z-I and Z-II structures. The chief feature of the packing of the helices in the crystal is that the terminal base pairs are all in the same *ab* plane, unlike in the structure of d(CGCGCG) (Wang *et al.*, 1979) where, in alternate columns, the terminal base pairs are displaced by half the unit cell along the c axis. In the present structure there are many contacts between the backbone atoms of the molecules related by unit-cell translations. These interactions may be responsible for the stability of the packing. They occur between the backbone atoms belonging to residues C1 to C3 of one molecule and those belonging to residues C11 to G12 of the molecule related by a translation along the x axis (see Fig. 2 for numbering). Along the y axis the contacts are between atoms of residue A4 and those of residues T9 to G10 of the translated molecule. The presence of the A·T base pair may thus serve to 'lock' the packing of the helices into the observed pattern.

Discussion

Ignoring sequence effects, perfect DNA helices have four surfaces to offer for interaction with neighbouring helices in the condensed state: the major and minor grooves, the backbone and flat ends. Table 2 lists the ten possible interacting pairs of surfaces. It also indicates which pairs have been observed in crystals.

Table 2. *Possible interaction modes in DNA helices with an example given for each mode observed in the crystals*

	Major groove	Minor groove	Backbone	End
Major groove				
Minor groove		(1)		
Backbone	(2)		(3)	
End		(4)		(5)

References: (1) Wing *et al.* (1980); (2) Timsit *et al.* (1989); (3) Privé *et al.* (1987); (4) Wang *et al.* (1982); (5) Wang *et al.* (1979).

The helices in the present crystals are involved in backbone–backbone and end–end interactions with their neighbours. End–backbone or end–groove interactions found in A-DNA crystals are precluded in packing modes which produce infinite parallel columns. Apart from the Z-DNA hexamer family, decanucleotides of B-DNA also exhibit such a packing pattern (Prive *et al.*, 1987). A small rearrangement of the packing in these would lead to the same type of degeneracy as observed in the present structure.

While the packing of DNA hexamers in the present case exhibits near degeneracy, that of perfect cylinders (Fig. 5) has exact space-group degeneracy. If the cylinders pack as indicated in the figure, it is apparent that the X-ray pattern of such a 'crystal' can be indexed in space groups $P6_1$ or $P6_5$ (with some disorder owing to the gap between the cylinders), $P2_1$, $P2_12_12_1$, $C2$ and $C222_1$. The 'asymmetric unit' in each case would be different. It would

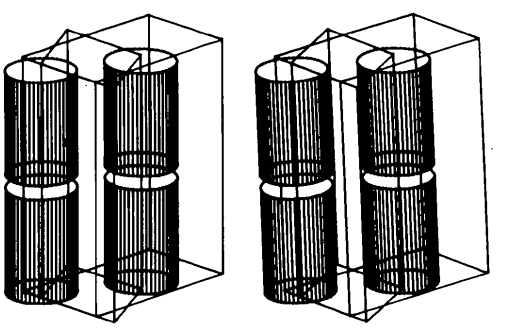

Fig. 5. An idealization involving perfect cylinders of the packing hexamers in the structure.

* Correlation coefficient = $[\sum(F_o - \langle F_o\rangle)(F_c - \langle F_c\rangle)]/[\sum(F_o - \langle F_o\rangle)^2 \times \sum(F_c - \langle F_c\rangle)^2]^{1/2}$.

be a third of the cylinder for $P6_1$, one half for $C222_1$ and one cylinder for $P2_1$, $P2_12_12_1$ and $C2$. The degeneracy would increase if the achiral space groups were considered which, however, is unnecessary in the context of DNA helices. Some of the degeneracy is trivial. For instance, any space group can be considered degenerate with triclinic $P1$. Similarly a crystal with $P6_1$ symmetry can also be dealt with in the less symmetrical space group $P2_1$. However, the degeneracy exhibited by the packing of perfect cylinders is of a different type. For the arrangement shown in the Fig. 5, the crystallographic asymmetric unit is of the same size irrespective of whether space group $P2_1$, $P2_12_12_1$ or $C2$ is chosen. Hence, from a practical point of view, one may consider these space groups to be exactly degenerate.

We thank Professor G. Govil, TIFR, Bombay, for DNA synthesis facilities, Professor M. A. Viswamitra, IISc., Bangalore, and Professor Kunthala Jayaraman, Anna University, Madras, for encouragement. We thank Professor Vasantha Pattabhi, University of Madras, for a critical reading of the manuscript. We also thank the referees for valuable suggestions. The financial aid received under the project DST/SO/SP/89 of The Department of Science and Technology, Government of India, is gratefully acknowledged. CS thanks UGC, India, for the award of a Fellowship.

References

BRENNAN, R. G. & SUNDARALINGAM, M. (1985). *J. Mol. Biol.* **181**, 561–563.

BRENNAN, R. G., WESTHOF, E. & SUNDARALINGAM, M. (1986). *J. Biomol. Struct. Dynam.* **3**, 649–665.

BROWN, T., KNEALE, G., HUNTER, W. N. & KENNARD, O. (1986). *Nucl. Acids Res.* **14**, 1801–1809.

BUERGER, M. J. (1946). *J. Appl. Phys.* **17**, 539–595.

DOCK-BREGEON, A.-C. & MORAS, D. (1992). *Crystallization of Nucleic Acids and Proteins – A Practical Approach*, edited by A. DUCRUIX & R. GIEGÉ, pp. 145–174. Oxford Univ. Press.

DREW, H. R., TAKANO, T., TANAKA, S., ITAKURA, K. & DICKERSON, R. E. (1980). *Nature (London)*, **286**, 567–573.

EISENSTEIN, M., FROLOW, F., SHAKKED, Z. & RABINOVICH, D. (1990). *Nucl. Acids Res.* **18**, 3185–3194.

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.

FUJII, S., WANG, A. H.-J., QUIGLEY, G. J., WESTERINK, H., VAN DER MAREL, G., VAN BOOM, J. H. & RICH, A. (1985). *Biopolymers*, **24**, 243–250.

FUJII, S., WANG, A. H.-J., VAN DER MAREL, G., VAN BOOM, J. H. & RICH, A. (1982). *Nucl. Acids Res.* **10**, 7879–7892.

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

KENNARD, O. & HUNTER, W. N. (1989). *Q. Rev. Biophys.* **22**, 327–379.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

PRIVÉ, G. G., HEINEMANN, U., CHANDRASEGAREN, S., KAN, L.-S., KOPKA, M. L. & DICKERSON, R. E. (1987). *Science*, **238**, 498–504.

SUSSMAN, J. L., HOLBROOK, S. R., CHURCH, G. M. & KIM, S.-H. (1977). *Acta Cryst.* A**33**, 800–804.

TIMSIT, Y., WESTHOF, E., FUCHS, R. P. P. & MORAS, D. (1989). *Nature (London)*, **341**, 459–462.

WANG, A. H.-J., FUJII, S., VAN BOOM, J. H., VAN DER MAREL, G. A., VAN BOECKEL, S. A. A. & RICH, A. (1982). *Nature (London)*, **299**, 601–604.

WANG, A. H.-J., FUJII, S., VAN BOOM, J. H. & RICH, A. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 3968–3972.

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.

WANG, A. H.-J., QUIGLEY, G. J., KOLPAK, F. J., VAN DER MAREL, G. A., VAN BOOM, J. H. & RICH, A. (1981). *Science*, **211**, 171–176.

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

WING, R., DREW, H., TAKANO, T., BROKA, C., TANAKA, C., ITAKURA, K. & DICKERSON, R. E. (1980). *Nature (London)*, **287**, 755–758.