Acta Cryst. (1996). D**52**, 481–485

Structure of the DNA Octanucleotide d(ACGTACGT)$_2$

DEBORAH J. WILCOCK,[a] ADRIENNE ADAMS,[a] CHRISTINE J. CARDIN[b]* AND LAURENCE P. G. WAKELIN[c]

[a]*Department of Chemistry, Trinity College, Dublin 2, Ireland,* [b]*Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, England, and* [c]*St Lukes Institute of Cancer Research, Highfield Road, Rathgar, Dublin 6, Ireland. E-mail: scscarch@reading.ac.uk*

(*Received 15 September 1995; accepted 27 November 1995*)

Abstract

d(ACGTACGT), $C_{78}H_{84}N_{30}O_{32}P_7.20H_2O$, M_r (DNA) = 2170, tetragonal, $P4_32_12$ (No. 96), $a = 42.845\,(1)$, $b = 42.845\,(1)$, $c = 24.804\,(1)$ Å, $V = 45532.5\,(2)$ Å^3, $z = 8, \lambda(\text{Mo}\,K\alpha) = 0.71069$ Å, $\mu(\text{Mo}\,K\alpha) = 0.10\,\text{mm}^{-1}$, $T = 295$ K, $R = 0.18$ for 1994 unique reflections between 5.0 and 1.9 Å resolution. The self-complementary octanucleotide d(ACGTACGT)$_2$ has been crystallized and its structure determined to a resolution of 1.9 Å. The asymmetric unit consists of a single strand of octamer with 20 water molecules. It is only the second example of an octanucleotide having terminal A·T base pairs whose structure has been determined by X-ray crystallography. The sequence adopts the modified A-type conformation found for all octanucleotide duplexes studied to date with the helix bent by approximately 15° and an average tilt angle of 0°. Unusually the data collection was carried out using a 3 kW molybdenum sealed-tube source. The conformational details are discussed in comparison with other closely related sequences.

1. Introduction

The self-complementary DNA octamer d(ACGTACGT)$_2$ was synthesized as the first part of a study of small molecule–DNA interactions. It has been crystallized in its native form to provide a reference point for ligand-binding studies to the same sequence, a study which is now in progress. To date we have obtained crystals of the octamer complexed with a 9-aminoacridine-4-carboxamide (Hudson, Kuroda, Denny & Neidle, 1987) which diffract to 2.4 Å. The lattice is trigonal, and our working hypothesis is that it contains one or two drug molecules intercalated into a B-type DNA. As an additional part of the study we have also prepared the brominated analogue d(ACGTACG$^{5\text{Br}}$U) and have shown that this gives crystals isomorphous to the native form. In the present publication we report details of the structure of the native DNA conformation, the second octameric sequence to be crystallized having terminal A·T base pairs and an example of an oligonucleotide studied using molybdenum radiation. NMR studies of the octamer d(ATGCGCAT)$_2$ have shown it to adopt a B-type conformation in solution, but in the solid state, this and many other DNA octanucleotides crystallize in an A-type conformation (Clark *et al.*, 1990). This behaviour has been attributed to the forces associated with packing self-complementary octanucleotides into their characteristic crystal form (Neidle, 1994). An analysis of conformational parameters shows that many of the modifications to the A-type structure, noted by other authors, are also found in our sequence.

2. Materials and methods

The oligodeoxyribonucleotide of sequence ACGTACGT was synthesized by solid-phase methods on an Applied Biosystems DNA synthesizer using phosphoramidite chemistry and purified using ion-exchange and reverse-phase high-pressure liquid chromatography (HPLC). Crystals were grown at 291 K by vapour diffusion using the hanging-drop technique (McPherson, 1982). A 10 µl drop containing 0.5 m*M* DNA, 1.5 m*M* cacodylate buffer, pH 7.0, 1.25 m*M* magnesium acetate, 2.5 m*M* spermine–4HCl and 6% 2-methyl-2,4-pentanediol (MPD) was equilibrated against at 1 ml reservoir containing 30% MPD. Microcrystals formed within one week but over a period of a further four weeks these redissolved and large tetragonal bipyramids grew which were suitable for X-ray analaysis.

A single crystal of dimensions 0.6 × 0.75 × 0.4 mm was mounted in a glass capillary and sealed together with drop of mother liquor. Data were collected on a MAR research 18 cm image plate at room temperature with graphite-monochromated Mo $K\alpha$ radiation generated by an Enraf–Nonius FR590 generator and a fine-focus sealed 3 kW tube, 50 kV 50 mA settings. 100 images were measured, with a $\Delta\varphi$ of 1° and an exposure time of 900 s per image. The crystal-to-detector distance was 150 mm and data were collected to a resolution of 1.5 Å. The cell dimensions were determined on a CAD-4 diffractometer to be $a = b = 42.845\,(1)$ and $c = 24.804\,(1)$ Å. The space group was found to be either $P4_12_12_1$ (No. 92) or $P4_32_12$ (No. 96). Data processing was carried out using *XDS* (Kabsch, 1988, 1993). A total of 24 514 reflections were measured to 1.5 Å resolution and averaged to 3962 unique reflections.

©1996 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 ©1996

[$R_{int} = 14.96\%$ for $I > 0$, where $R_{int} = \sum |F_o^2 - F_o^2(\text{mean})| \sum(F_o^2)$]. To 1.9 Å resolution a total of 14 470 reflections were measured and averaged to 1994 unique reflections with an R_{int} of 8.83%. Data to 1.9 Å resolution were used for structure solution. Crystal decay during data collection was very small, which was a positive consequence of the use of molybdenum radiation, and absorption was negligible.

The structure was solved using the coordinates of the previously solved octamer d(CTGTACAC)$_2$ (Jain, Zon & Sundaralingam, 1987) as a starting model for refinement. This sequence was chosen because the cell dimensions ($a=b=42.43$ and $c=24.75$ Å) and symmetry ($P4_32_12$) are very close to ours, together with high sequence identity. The symmetry of our octamer crystal form was later confirmed to be $P4_32_12$ by successful refinement. The coordinates were obtained from the Nucleic Acid Database (Berman *et al.*, 1992) and used as the initial model for refinement. The structure was refined using the macromolecular program *X-PLOR* (Brünger, 1992) using data in the resolution shell 5.0–1.9 Å. The necessary changes to the base pairs were made (G1 to A1, T2 to C2, A7 to G7, C8 to T8) and all the temperature factors set to 15.0 Å^2 before beginning the refinement. Refinement was interspersed with examination of $(2F_o - F_c)$ and $(F_o - F_c)$ electron-density maps generated by *X-PLOR* and displayed using *TOM/FRODO* (Jones, 1978) to ensure a good fit between map and model. The two missing atoms from the seventh and eighth base pairs were located from these difference maps. Initially the oligomer was refined as a rigid group of eight bases and then as eight individual rigid groups after which the crystallographic R value was 29.4%. A positional refinement was then carried out and finally a B-factor refinement leading to an R value of 20.0%. A search for water molecules was then initiated. The water molecules were treated in the refinement as solvent O atoms. The criteria for location of a water molecule were (i) a well shaped peak in the difference map between 2.0 and 3.4 Å from a possible hydrogen-bonding partner; (ii) retention of a satisfactory geometry during refinement; and (iii) a final value of the thermal parameter less than 50 Å^2. The refinement converged after the location of 20 water molecules. The final R value was 18.0% for the 1994 reflections between 5.0 and 1.9 Å. The average thermal parameters (Å^2) for phosphate, sugar and bases are approximately 19.5, 15.9 and 11.6, respectively. The values for the solvent molecules have an average value of 35 Å^2. The model fits the electron density well and the geometry of the structure is good with r.m.s. deviations for the bond lengths and angles of 0.01 Å and 3.47°, respectively.

3. Results and discussion

The asymmetric unit in the crystal of d(ACGTACGT)$_2$ is a single strand of DNA. Two strands are related by the crystallographic dyad to form a right-handed double helix. The average rise per base pair is 3.2 Å, the mean helix rotation is 33° and there are 11 base pairs per turn. The bases are oriented *anti* about the glycosidic bond and the furanose ring on average adopts the C3′-*endo* conformation. The bases are displaced away from the helix axis: the minor groove is wide and shallow, while the major groove is narrow and deep (see Fig. 1). The crystal packing is similar to that seen in other A-DNA octamers – molecules are organized around a left-hand fourfold screw axis. The terminal base pair of one octamer is stacked against the almost flat surface of the minor groove of an adjacent symmetry-related molecule, stabilized by van der Waals contacts and hydrogen bonding. A most striking feature of the helix is its pronounced curvature. The helix appears to be bent by approximately 15° towards the minor groove. This has been seen in other octamers solved in the same space group and is most likely a result of crystal packing forces (Bingman, Li, Zon & Sundaralingham, 1992).

While overall the helix appears to be in the A-form, a closer examination reveals structural features which differ from the average and which may be sequence dependent. The nucleotides are labelled A1 to T8 on strand one and A9 to T16 on strand two in the 5′ to 3′ direction.

Geometrical parameters of the base–base interactions were calculated using the program *RNA* (Babcock, Pednault & Olson, 1993) and are presented in Table 1. Propeller twist is present in all base pairs but there is no difference in the average values between CG and AT base pairs, the largest values being found at the end of the helix. This is in contrast to the sequence d(ATGCGCAT)$_2$ which does show a difference in the CG and AT base pairs (Clark *et al.*, 1990). The smallest value of roll is found in steps three and five with a value of −1.9° indicating that adjacent pairs are nearly parallel at these steps. The larget values are found in the adjoining steps two and five. There is some variation in helical twist along the helix with values of 32° at the ends of the duplex. These increase to 36° at step three and then drop to 27° at the central step four. Average values for the tilt angle from fibre diffraction for A-DNA are 20°, but the average in our structure is 0°. This is closer to the values obtained for B-DNA of 4°. Low tilt angles have been found for other A-DNA octamers *e.g.* 1° in d(GTCTAGAC)$_2$ (Cervi, Langlois d'Estaintot & Hunter, 1992) and 0° in d(ATGCGCAT)$_2$ (Clark *et al.*, 1990).

Along the sequence there is a distinct alternation in two of the base-pair step parameters, roll and tilt, with purine–pyrimidine steps having larger roll and smaller tilt angles than pyrimidine–purine steps. Slide values also vary along the sequence with pyrimidine–purine steps having larger values than purine–pyrimidine steps. Finally the rise per base pair, while having an average value of 3.2 Å, varies from 3.0 to 3.3 Å, the smallest value being found at the central TA step.

Table 1. *Base-pair and base-pair step parameters*

Parameters were calculated using the *RNA* program (Babcock, Pednault & Olson, 1993).

Base-pair	Propeller twist (°)	Tilt (°)	Roll (°)	H-twist (°)	Slide (Å)	Inclination (°)	Rise per residue (Å)
A1·T16	−19.4						
		0.4	5.2	32.1	−1.4	9.3	3.3
C2·G15	−13.9						
		−1.5	10.7	33.2	−1.6	18.8	3.2
G3·C14	−14.3						
		−1.9	−1.9	36.3	−1.4	−3.1	3.3
T4·A13	−8.5						
		0.0	3.7	27.4	−1.5	7.7	3.0
A5·T12							
Mean	−14.1	0.0	4.5	32.5	−1.4	8.25	3.2
Standard deviation	3.7	1.3	4.8	2.8	0.1	8.3	0.1

Values for sugar–phosphate torsion angles together with those for the pseudorotation phase angle, P, and sugar pucker parameter, τ_m, are presented in Table 2. Most of the sugar moieties are in the C3′-*endo* conformation as can be seen from the values of δ, (76–89°), the pseudorotation phase angle, (−7 to 26°) and the sugar pucker parameter, (31–50°). One exception is the furanose ring on residue A5 which is in the C2′-*exo*-C3′-*endo* conformation. Values for α and γ are *gauche*$^-$ and *gauche*$^+$, respectively, at all residues with the notable exception of residue A5. Here α has a value intermediate between *gauche*$^+$ and *trans* and γ has a pure *trans* value. This phenomenon has been observed in other octamer structures at the central pyrimidine–purine step: *e.g.* d(GCCCGGGC)$_2$ (Heinemann, Lauble, Frank & Blocker, 1987) and has previously been ascribed to the effect of interstrand stacking interactions.

Fig. 2 shows the extent of stacking shown by successive base pairs. Here we observe distinctly greater interstrand stacking at the pyrimidine-3′,5′-purine steps (steps 2 and 4 in our sequence, CG/CG and AT/AT pairs, respectively) than that reported for the sequence d(ATGCGCAT)$_2$ (Clark *et al.*, 1990). These authors report almost complete unstacking at their pyrimidine-3′,5′-purine steps, whereas we see nearly 50% interstrand overlap of the six-membered rings. This pattern of interstrand overlap at the pyrimidine–purine steps has been observed in two previously solved octanucleotide structures, d(GTGTACAC)$_2$ (Jain, Zon & Sundaralingam, 1991) and d(GTGCGCAC)$_2$, (Bingman *et al.*, 1992), both of which have an alternating purine-pyrimidine sequence. At the purine-3′,5′-pyrimidine steps (steps 1 and 3 in our sequence, AC/GT and GT/AC, respectively) there is almost complete intra-strand overlap of the six-membered rings. Fig. 2 shows the consistency of this alternating intra- and interstrand overlap, which by virtue of the twofold axis, extends the length of the sequence. In this sequence, the value of 135° for the α torsion angle is only found at the central TA/TA step (Fig. 2*d*), which is no more subject to

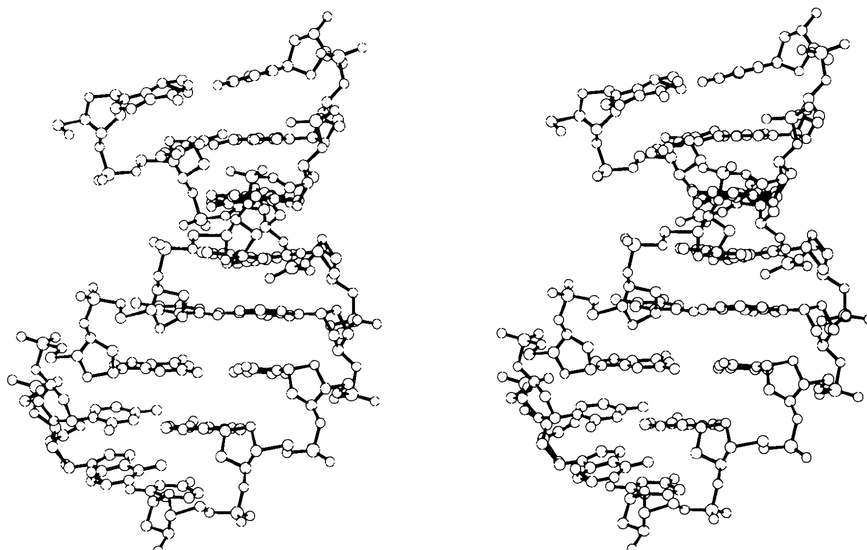

Fig. 1. A stereoview of the d(ACGTACGT)$_2$ sequence, showing both the major and minor grooves.

Table 2. *Backbone torsion angles (°) and sugar-pucker parameters*

The torsion angles are defined as: α = O3′—P—O5′—C5′, β = P—O5′—C5′—C4′, γ = O5′—C5′—C4′—C3′, δ = C5′—C4′—C3′—O3′, ε = C4′—C3′—O3′—P.

	α	β	γ	δ	ε	p	τ_m
A1			67	86	213	13	44
C2	293	160	58	76	197	26	48
G3	294	171	59	88	193	20	38
T4	286	175	59	82	193	11	31
A5	135	−155	−170	89	218	−7	42
C6	284	167	62	80	207	16	50
G7	291	167	62	76	190	5	49
T8	283	175	63	84		15	40

interstrand stacking than the GC/GC step (Fig. 2*b*). This anomalous value seems to be characteristic of the central step in octamer helices and may be a feature of the packing.

The minor groove width varies from 8.9 to 9.7 Å with an average value of 9.2 Å. The major groove width cannot be accurately measured in such a short helix but the P2—P10 distance can give an indication of its value and in our structure this is 8.3 Å. Fibre-diffraction values are about 4 Å for the major groove width of A-DNA and 12 Å for B-DNA. This increase in major groove width is a common feature of other octamer crystal structures *e.g.* d(ATGCGCAT)$_2$ (Clark *et al.*, 1990), d(GTCTAGAC)$_2$ (Cervi *et al.*, 1992), d(GTGCGCAT)$_2$ (Bingman *et al.*, 1992).

In summary, d(ACGTACGT)$_2$ crystallizes in a modified A-type conformation with an average tilt angle of 0° and with the helix bent by 15° towards the minor groove. Close examination reveals other structural features which differ from those expected for A-DNA and which are more typical of B-DNA. The dominant local feature is found at the central TA step where the helical twist drops to 27° and the rise per base pair falls to 3 Å. At this step propeller twist also drops to its lowest value of −8° and base-pair stacking is maximal. All of these features have previously been observed in other octanucleotide duplexes crystallizing in the same group *e.g.* d(GCCCGGGC)$_2$ (Heinemann *et al.*, 1987), d(ATGCGCAT)$_2$ (Clark *et al.*, 1990), d(GTCTAGAC)$_2$ (Cervi *et al.*, 1992), d(GTGCGCAT)$_2$ (Bingman *et al.*, 1992). Comparison with d(ATGCGCAT)$_2$ reveals many similarities between the two structures as might be expected given their similar sequences and identical crystal forms (Clark *et al.*, 1990).*

We would like to thank Mr Paul Tyndall for DNA synthesis and purification, and Professor S. Neidle and Dr Z. Dauter for data collections and for helpful discussions. We thank Professor G. Sheldrick

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 243d). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: AD0013).

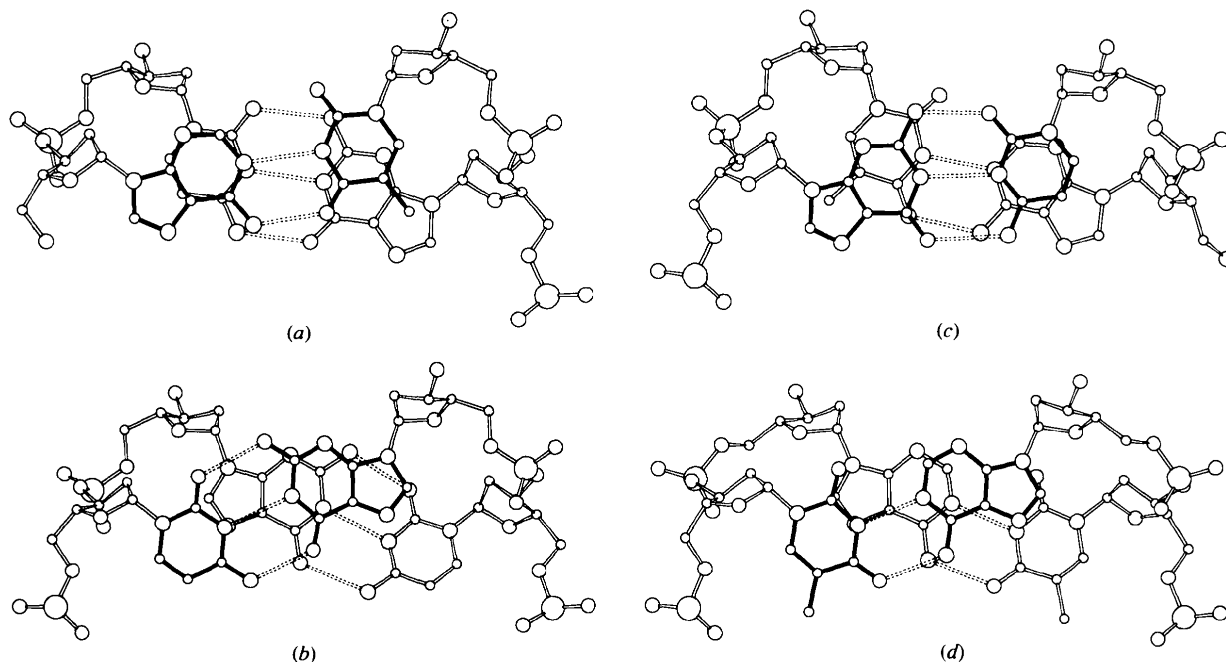

Fig. 2. Views of the four distinct base-pair overlaps in the d(ACGTACGT)$_2$ sequence. The projection is the same for each view, onto the least-squares plane of the upper base pair in the plot, and the first in the list below. The 5′ chain end is at the lower left corner of each figure, and the 3′ end of the opposite strand at the lower right. (*a*) Step 1 (AC/GT); (*b*) step 2 (CG/CG); (*c*) step 3 (GT/AC); and (*d*) step 4 (TA/TA).

(Göttingen) for use of the *SHELXL* program, and for helpful discussions, and Dr E. Keller (Freiburg) for the use of his plotting program *SCHAKAL*. This work was supported by EOLAS, The Irish Science and Technology Agency.

References

Babcock, M. S., Pednault, P. D. & Olson, W. K. (1993). *J. Biomol. Struct. Dynam.* **11**, 596–627.

Berman, H. M., Olson, W. K., Beverirdge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S. H., Srinivasan, A. R. & Schneider, B. (1992). *Biophys J.* **63**, 751–759.

Bingman, C., Li, X., Zon, G., & Sundaralingam, M. (1992). *Biochemistry,* **31**, 12803–12812.

Brünger, A. T. (1992). *X-PLOR Version 3.0: A system for crystallography and NMR,* Yale University, New Haven, Connecticut, USA.

Cervi, A., Langlois D'Estaintot, B. & Hunter, W. N. (1992). *Acta Cryst.* B**48**, 714–719.

Clark, G. R., Brown, D. G., Sanderson, M. R., Chwalinski, T., Neidle, S., Veal, J. M., Jones, R. L. Wilson, W. D., Garman, E. & Stuart, D. I. (1990). *Nucleic Acids Res.* **18**, 5521–5528.

Heinemann, U., Lauble, H., Frank, R. & Blocker, H. (1987). *Nucleic Acids Res.* **15**, 9531–9550.

Hudson, B. D., Kuroda, R., Denny, W. A. & Neidle, S. (1987). *J. Biomol. Struc. Dynam.* **5**, 145–158.

Jain, S., Zon, G. & Sundaralingam, M. (1987). *J. Mol. Biol.* **107**, 141–145.

Jain, S., Zon, G. & Sundaralingam, M. (1991). Biochemistry, **31**, 12803–12812.

Jones, A. (1978). *J. Appl. Cryst.* **11**, 268–272.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley.

Neidle, S. (1994). *DNA structure and recognition.* Oxford IRL Press.