FAST COMMUNICATIONS

Contributions intended for this section should be submitted to any of the Co-editors of Acta Crystallographica *or* Journal of Applied Crystallography. *In the letter accompanying the submission authors should state why rapid publication is essential. The paper should not exceed two printed pages (about 2000 words or eight pages of double-spaced typescript including tables and figures) and figures should be clearly lettered. If the paper is available on 5.25" IBM PC compatible or 3.5" Apple/Macintosh diskettes it would be helpful if these could be sent with the manuscript together with details of the word-processing package used. Papers not judged suitable for this section will be considered for publication in the appropriate section of* Acta Crystallographica *or in* Journal of Applied Crystallography.

Acta Cryst. (1990). A**46**, *FC9–FC12*

Crystal structure analysis of an A(DNA) octamer d(GTACGTAC)

BY C. COURSEILLE[1], A. DAUTANT[1], M. HOSPITAL[1], B. LANGLOIS D'ESTAINTOT[1], G. PRECIGOUX[1], D. MOLKO[2] AND R. TEOULE[2]

[1] *Laboratoire de Cristallographie, URA 144, CNRS, Université Bordeaux I, 33405 Talence, France*

[2] *Centre d'Etudes Nucléaires de Grenoble, 85X, F38041 Grenoble, France*

(Received 22 February 1990; accepted 5 March 1990)

Abstract. The synthetic deoxyoctanucleotide d(GTACGTAC) crystallizes as an A-type DNA double helix. The structure has been refined to an R factor of 17% at 2.4 Å resolution with 56 solvent molecules located. The tetragonal unit cell, space group $P4_32_12$ has dimensions $a=42.50$ (7) and $c=24.79$ (5) Å. The asymmetric unit consists of a single strand of four base pairs.

Introduction. Uracil does not occur as a usual component of DNA, but may appear either when dUTP residues are incorporated in place of dTTP by DNA polymerases during replication (Laval & Laval, 1970; Tye & Lehman, 1977; Lindahl, 1982), or by the *in situ* deamination of cytosine residues (Shapiro & Klein, 1966; Lindahl, 1979). The physiological role of correction, or repair, of deaminated cytosine residues in DNA is played by the uracil DNA glycosylase which releases a free uracil by hydrolytic cleavage of the uracil–deoxyribose bond and yields an apyrimidic site without changing the sugar–phosphate backbone (Hanawalt, Cooper, Ganesan & Smith, 1979; Lindahl, 1979, 1982).

A few years ago, using synthetic oligonucleotides containing native bases (Delort, Duplaa, Molko, Teoule, Leblanc & Laval, 1985), or modified bases such as m^5dC (Chen, Cohen & Behe, 1983; Sanderson, Mellema, Van der Marel, Wille, Van Boom & Altona, 1983; Feigon, Wang, Van der Marel, Van Boom & Rich, 1984; Prettre-Giessner, Pullman, Tran-Dinh, Neumann, Huynh-Dinh & Igolen, 1984; Taboury, Adam, Taillandier, Neumann, Tran-Dinh, Huynh-Dinh, Langlois d'Estaintot, Conti & Igolen, 1984), 2-amino dA (Taboury *et al.*, 1984), m^6dA (Fazakerley, Teoule, Guy, & Gushlbauer, 1984) it was shown, by enzymatic processes (Delort, Duplaa *et al.*, 1985) and ^{1}H NMR or other spectroscopic analyses (Chen *et al.*, 1983; Sanderson *et al.*, 1983; Fazakerley *et al.*, 1984; Feigon *et al.*, 1984; Prettre-Giessner *et al.*, 1984) that the excision of the uracil is not a random phenomenon. The sequence of the bases surrounding the uracil has an influence on the rate of excision (Delort, Duplaa *et al.*, 1985).

Some ^{1}H NMR studies carried out on d(GTACGTAC) and d(GTACGUAC) showed that the introduction of a uracil in place of a thymine does not affect the global structure of DNA and, therefore, that the recognition of the uracil defect by uracil–DNA glycosylase was not based upon a bulky distorsion (Delort, Neumann, Molko, Hervé, Teoule & Tran-Dinh, 1985).

Precise details on the conformation of these oligonucleotides can come only from single-crystal X-ray studies, thus we undertook the crystallization experiments and structure resolution of the native d(GTACGTAC).

Materials and methods. The oligomer d(GTACGTAC) was synthesized by phosphoramidite methodology on an Applied Biosystems synthesizer (381A). Complete cleavage and deprotection was carried out following standard procedures and the crude material was purified

0108-7673/90/04FC009-04$3.00

by ion-exchange and reverse-phase high-pressure liquid chromatography.

Crystals were grown at 293 K by vapour phase diffusion from droplets sitting in Corning glass depression plates (McPherson, 1982). Well formed tetragonal crystals reached a suitable size over a period of 3–4 weeks from aqueous solutions containing 2 m*M* oligonucleotide, 25 m*M* sodium cacodylate buffer (pH 6.0), 15 m*M* $MgCl_2$, 8 m*M* spermine hydrochloride and 5% (*v*/*v*) 2-methyl-2,4-pentanediol (MPD) diffused against 30% (*v*/*v*) MPD.

A single crystal of dimensions $0.28\times0.20\times0.19$ mm, used for the data collection, was mounted and sealed in a glass capillary, containing a drop of mother liquor and was stable for weeks at room temperature. The space group was determined as $P4_32_12$ or $P4_12_12$ with unit-cell dimensions $a=b=42.50$ and $c=24.79$ Å ($V=44777$ Å^3). Our results are in agreement with some preliminary crystallographic studies of the same oligonucleotidic sequence mentioned by Takuzagawa (1988). Measurements were made at 279 K on a Nicolet detector (Cu *Kα* radiation) to an upper resolution limit of 2.4 Å. Intensities were corrected for Lorentz and polarization factors, and time-dependent decay. Symmetry-equivalent reflections were averaged. In total 2682 measurements were reduced to 976 with $R_{sym}=5.5\%$. Of these data, 894 have $I > o(I)$.

Cell dimensions and symmetry suggested that the crystal structure was isomorphous with other self-complementary DNA oligomers. One of them, selected as a starting model for our structure analysis, was that of d(CTCTAGAG) in space group $P4_32_12$ (Hunter, Langlois d'Estaintot & Kennard, 1989). The asymmetric unit in the crystal structure is a single strand of four base pairs. Assuming the stereochemistry of the helical structure of d(CTCTAGAG) and swapping the bases for those of d(GTACGTAC) using the molecular modelling system *MMS* (Dempsey, 1986), a useful model of the single strand was built. All the data with $I > o(I)$ were included in the calculations and convergence was obtained with an *R* factor of 33%. Difference density ($F_o - F_c$) and electron density ($2F_o - F_c$) maps were calculated and displayed on a Silicon Graphics system using the program *TOM FRODO* (Jones, 1985; Cambillau & Horjales, 1988). Solvent positions were identified from difference and electron density maps with the criteria of well shaped peaks in the difference density maps, situated within 2.2 to 3.4 Å of plausible hydrogen-bonding partners (O or N atoms).

The refinement was terminated with negligible restraints on the atoms of the sugar–phosphate backbone, using the 1.5 version of *XPLOR* (Brünger, Karplus & Petsko, 1989). For the whole structure, including the water molecules, an excellent fit is observed between atomic positions and electron density. The final structure contains 56 solvent molecules in the asymmetric unit. Four of them have a special position on a twofold axis. For all of them, the geometry of their associated hydrogen bonds agrees with parameters expected for water molecules. Most of these molecules display *B* values in the range 12–89 Å^2. Three of them, with lower *B* values, could be counter-ions but, owing to the limited resolution (2.4 Å), cannot be unequivocally assigned. No partial occupancies were allowed for the solvent structure, as far as general positions are concerned. The temperature factors take up the effect of incomplete occupancy of water sites and do not strictly reflect thermal motion. The final *R* factor based on all 894 reflections with $I > o(I)$ is 17%. The structure factors and atomic coordinates have been sent to the Protein Data Bank, Brookhaven National Laboratory.*

Results and discussion. The asymmetric unit in the crystal structure of d(GTACGTAC) is a single strand of four base pairs. Two strands coil about each other around a diagonal crystallographic dyad in order to form a right-handed double helix of the A-DNA family. The helix, illustrated in Fig. 1, has a shallow minor groove and a deep major groove.

In the duplex, nucleotides are numbered 1–8 on strand 1, 9–16 on strand 2 in the 5' to 3' direction. The Watson–Crick base pairs are therefore G1–C16, T2–A15 *etc.*

The overall helix structure can be considered as an A-type DNA, with the C3'-*endo* sugar conformation in general, and is similar to those observed in the other octanucleotides [d(GCCCGGGC) (Heinemann, Lauble, Frank & Blocker, 1987) and d(CTCTAGAG) (Hunter *et al.*, 1989)] crystallizing in the same tetragonal ($P4_32_12$) space group. The average rise per base is 3.2 Å, with an average rotation per base value of 31.5°, which results in 11.4 base pairs per helix turn. The average displacement of the bases from the helix axis is 3.1 Å. The roll, the angle by which adjacent base pairs open up to the minor groove, shows that the relevant base pairs of steps 1, 3

* Lists of structure factors and atomic coordinates have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 52787 (12 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

and 4 are nearly parallel to each other. Step 2 displays a roll value of 15°, which is the highest roll value in this structure.

Nevertheless, some mean parameters slightly differ from the usual values. These divergences can be explained by a perturbed local structure in the centre of the helix, at step 4 (C4–G13/G5–C12).

The small mean helix twist angle (30.5°) is basically caused by a partial unwinding of the helix at the central step (twist = 23°). The strong cross-strand stacking forces the conformation of the sugar–phosphate backbone to make a shift in order to minimize the perturbation. So, the intrastrand P4–P5 atom distance is increased up to 7.0 Å. This constraint also has some long-range effects such as a small displacement between A3–T14/C4–G13 (2.8 Å), a small base-pair inclination (10°) and a wide minor groove (9.7 Å). The main geometrical properties of base-pair steps and base pairs are reported in Table 1. The predominant intermolecular interaction in the crystal structure involves the packing of the first base pair of one duplex against the minor groove of another duplex. The molecular packing produces infinite channels in the *c* direction. The channels are approximately 12×27 Å wide, and are filled with water molecules.

Table 1. *Geometrical properties of base-pair steps and base pairs*

Helical parameters are calculated with the *NEWHELIX* routine using the 1989 Cambridge Nomenclature Convention (Diekmann, 1989).

Base	Step	Roll (°)	Tilt (°)	Tip (°)	Incl. (°)	Twist (°)	Rise (Å)	Propeller twist (°)
G1–C16				−8	12			16.8
	1	7	1.5			30.3	3.2	
T2–A15				−7	8			9.5
	2	15	0.6			35.2	3.3	
A3–T14				4	9			15.1
	3	2	0.8			33.7	3.0	
C4–G13				−2	11			5.2
	4	6	0.0			22.9	3.2	
G5–C12								

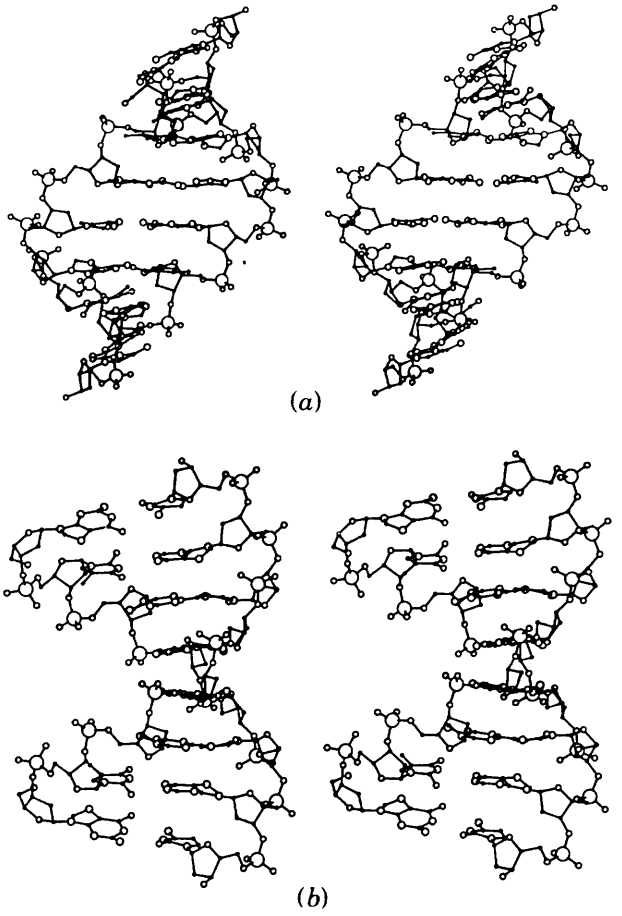

Fig. 1. Stereo representation of the structure of d(GTACGTAC). Carbon, nitrogen, oxygen and phosphorus are drawn with different radii. (*a*) View into the minor groove along the dyad symmetry axis. (*b*) View perpendicular to the minor groove.

Every solvent position has one or more contacts with hydrogen-bonding partners in the range 2.4–3.4 Å. The phosphate groups are the main hydration sites. In addition to these first-shell water molecules, the shorter distances between hydrogen-bonded atoms are located at the O3' and O5' ends.

The crystal structure of d(GTACGTAC) can be compared with the structure observed in solution by NMR. This oligomer gave a NMR spectrum attributed to a conformation characteristic of a B double helix (Delort, Neumann *et al.*, 1985) whereas the observed conformation in the crystal is a representative model of A-DNA. It is known that the conformations of DNA fragments are very dependent on the hydration (Drew & Dickerson, 1981). In the crystal structure, the hydration of the double helix is seriously perturbed by intramolecular packing effects. Such a disruption of an ordered network of hydration could be very important in the mechanism allowing the B to A transformation. Despite the disparity observed between the X-ray and NMR conformations, the results are converging on the same solution of well organized structures without any bulky distorsion.

Consequently, it is probably not such a distorsion which is responsible for the recognition of the uracil defect by the repair enzyme: uracil–DNA glycosylase.

The authors wish to express their thanks to Drs Dino Moras and Jean-Claude Thierry from the Institut de Biologie Moléculaire et Cellulaire, Strasbourg, for the data collection on their Nicolet diffractometer and helpful discussions. We are grateful to Drs Eric Westhof and Christian Cambillau for sending computer programs, and to Dr William N. Hunter for providing the coordinates of the model d(CTCTAGAG).

References

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* A**45**, 50–61.

CAMBILLAU, C. & HORJALES, E. (1988). *J. Mol. Graphics*, **5**, 174–177.

CHEN, C., COHEN, J. S. & BEHE, M. (1983). *Biochemistry*, **22**, 2136–2142.

DELORT, A. M., DUPLAA, A. M., MOLKO, D., TEOULE, R., LEBLANC, J. P. & LAVAL, J. (1985). *Nucleic Acids Res.* **13**, 319–335.

DELORT, A. M., NEUMANN, J. M., MOLKO, D., HERVÉ, M., TEOULE, R. & TRAN DINH, S. (1985). *Nucleic Acids Res.* **13**, 3343–3355.

DEMPSEY, S. (1986). *MMS*. Univ. of California, San Diego, USA.

DIEKMANN, S. (1989). *J. Mol. Biol.* **205**, 787–791.

DREW, H. R. & DICKERSON, R. E. (1981). *J. Mol. Biol.* **151**, 535–555.

FAZAKERLEY, G. V., TEOULE, R., GUY, A., GUSHLBAUER, W. (1984). *FEBS Lett.* **176**, 449.

FEIGON, J., WANG, A. H. J., VAN DER MAREL, G. A., VAN BOOM, J. H. & RICH, A. (1984). *Nucleic Acids Res.* **12**, 1243–1263.

HANAWALT, P. C., COOPER, P. K., GANESAN, A. K. & SMITH, CH. A. (1979) *Annu. Rev. Biochem.* **48**, 783–736.

HEINEMANN, U., LAUBLE, H., FRANK, R. & BLOCKER, H. (1987). *Nucleic Acids Res.* **22**, 9531–9550.

HUNTER, W. N., LANGLOIS D'ESTAINTOT, B. & KENNARD, O. (1989). *Biochemistry*, **28**, 2444–2451.

JONES, T. A. (1985). *Methods Enzymol.* **115**, 157–171.

LAVAL, J. & LAVAL, F. (1970). *IARC Sci. Publ.* **27**, 53–72.

LINDAHL, T. (1979). *Proc. Natl Acad. Sci. USA*, **82**, 135.

LINDAHL, T. (1982). *Annu. Rev. Biochem.* **51**, 61–87.

MCPHERSON, A. (1982). *Preparation and Analysis of Protein Crystals*. San Francisco: Wiley.

PRETTRE-GIESSNER, C., PULLMAN, B., TRAN-DINH, S., NEUMANN, J. M., HUYNH-DINH, T. & IGOLEN, J. (1984). *Nucleic Acids Res.* **12**, 3271–3281.

SANDERSON, M. R., MELLEMA, J. R., VAN DER MAREL, G. A., WILLE, G., VAN BOOM, J. H. & ALTONA, C. (1983). *Nucleic Acids Res.* **11**, 3333.

SHAPIRO, R. & KLEIN, R. S. (1966). *Biochemistry*, **5**, 2358.

TABOURY, J. A., ADAM, S., TAILLANDIER, A., NEUMANN, J. M., TRAN-DINH, S., HUYNH-DINH, T., LANGLOIS D'ESTAINTOT, B., CONTI, M. & IGOLEN, J. (1984). *Nucleic Acids Res.* **12**, 6291.

TAKUZAGAWA, F. (1988). Am. Crystallogr. Assoc. Meet., Philadelphia. Abstracts, p. 107.

TYE, B. K. & LEHMAN, I. R. (1977). *J. Mol. Biol.* **117**, 293–306.