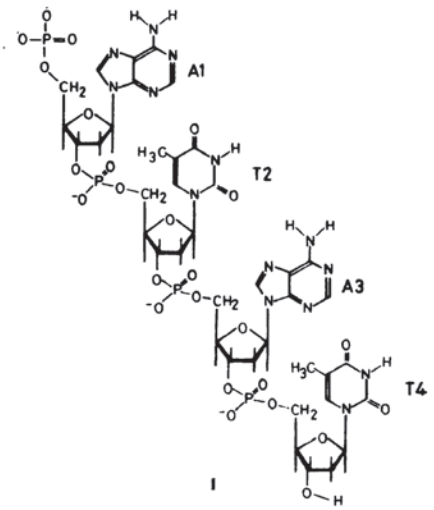
Nature Vol. 273 22 June 1978

15. Shaila, S., Gobleu, B., Brown, G. E., Sen, G. C. & Lengyel, P. *J gen. Virol.* **37**, 535–546 (1977).
16. Hickey, E. D., Weber, L. A. & Baglioni, C. *Biochem. biophys Res. Commun* **80**, 377–383 (1978).
17. Carlesby, A. *Proc. R. Soc.* A **224**, 120 (1954).
18. Ratner, L. *et al. Eur. J. Biochem.* **79**, 565–577 (1977).
19. Calster, R. L. & Lengyel, P. *Nucleic Acids Res.* **3**, 581–598 (1976).
20. Lenz, J. R., Chatterjee, G. E., Maroney, P. A. & Baglioni, C. *Biochemistry* **17**, 80–87 (1978).
21. Hickey, E. D., Weber, L. A. & Baglioni, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 19–23 (1976).

DNA double helical fragment at atomic resolution

DETAILS of the molecular architecture of double helical ribonucleic acids at atomic resolution have recently become available from single crystal X-ray diffraction studies of dinucleotides composed of complementary bases[1–3]. No similar studies have been published for deoxynucleotides. We report here the structure of the deoxytetranucleotide d-pApTpApT (5′-P-adenylyl-(3′-5′)-thymidylyl-(3′-5′)-adenylyl-(3′-5′)-thymidine) (I) at a resolution of 1.0 Å. This is the first tetranucleotide whose structure has been elucidated by X-ray diffraction. The work is part of an investigation of protein–nucleic acid interactions using single-crystal studies of small model compounds. d-pApTpApT was chosen as one of the first compounds to be examined because there is evidence that poly(dA–dT) has unusual binding properties and that A.T-rich regions in certain DNAs have specific biological roles[4–7]. We hope that these investigations will provide information about the influence of specific base pairs and sequences on the fine details of the DNA structures, and thus aid the understanding of the selective recognition of nucleotide sequences of the DNA double helix by proteins.



I

d-pApTpApT was synthesised in this laboratory and crystallised as the ammonium salt with two molecules of the oligomer and 62 water molecules per NH_4^+ ion in a monoclinic cell; space group $P2_1$. The analysis was carried to a resolution of 1.0 Å. Figure 1 shows a view of the molecule.

d-pApTpApT has a self-complementary base sequence, and could, theoretically, crystallise as a self-paired duplex of four base pairs. In the present structure the two molecules in the unit cell are related by a twofold screw axis which eliminates this possibility. A segment of a right-handed antiparallel

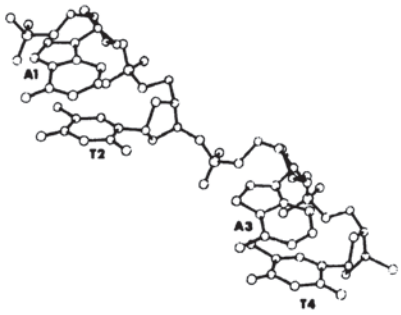


Fig. 1 A molecule of d-pApTpApT viewed perpendicular to *a* and showing the extended backbone conformation and residue labelling.

double helix with two base pairs is, however, formed with complementary hydrogen bonding between adenine and thymine bases related by the 2_1 axis (A1·T4*, T2·A3* where the starred bases are at $1-x, -\frac{1}{2}+y, -z$). The converse pairs (T4·A1* and A3·T2*) have the starred bases at $1-x, \frac{1}{2}+y, -z$; thus, each molecule contributes to two mini-helices, one with each of two other molecules related by *y* translation.

Figure 2 shows a view of the base pairs, held by hydrogen bonds with classical Watson–Crick geometry[8] (observed for adenine-thymine for the first time in a single-crystal study).

Within each pair the bases are inclined at a dihedral angle of approximately 13°. The mean planes through A1·T4* and T2·A3* are nearly parallel (3°), with a vertical separation of 3.34 Å. The angular orientation of the base pairs is 31° and the C1′–C1′ interstrand separation is 10.2 Å. Such mini-helices have previously been obtained from single-crystal studies of RNA double helical fragments[1–3] but this is the first description of Watson–Crick base pairs of a DNA fragment at atomic resolution.

A striking feature of the d-pApTpApT molecule and hence of the mini-helix is that the conformation of the deoxyribose is C3′-*endo* when attached to a purine and C2′-*endo* when attached to a pyrimidine base. This is in contrast to double helical fragments of ApU (ref. 1) and GpC (refs 2, 3), where only the C3′-*endo* conformation was found, and the classical interpretation of DNA fibre patterns which were based on monotonic sugar geometry, C3′-*endo* in DNA-A and C2′-*endo* in DNA-B (ref. 9). In the present structure the relative orientation of the base and sugar (χ_{CN}), which is correlated with the sugar conformation[10], also varies between low and high values (χ_{CN} = 5°(A1), −9°(A3), 69°(T2), 75°(T4)).

Our results suggest that both the sugar pucker and sugar base orientation are 'soft' parameters with low energy barriers between different conformations. It is apparently necessary to consider irregular or periodic variations in these parameters, at least over local regions of double helical structures, depending on the nature of the nucleotide sequence.

An important feature of the structure is the sugar–phosphate backbone (Fig. 1). The conformation within the two halves of the molecule is remarkably similar, with torsion angles ω and ω' about the P-O5′ and P-O3′ ester bonds in the d(A–T) fragments being close to those allowed for right-handed double helical polynucleotides. The phosphodiester linkage between the two halves, however, is very different, with *trans-gauche* conformation giving rise to an extended backbone similar to that observed in the crystal structure of UpA (refs 11, 12) and pTpT (ref. 13).

The base paired duplex (A1–T2)·(A3*–T4*) found in the present structure provides a reasonable starting point for generating a model for the dA–dT copolymer.

Model building studies showed that a helical structure could not be generated from the duplex by helical rotation and translation only. It was necessary to alter the conformation of the

phosphodiester bridge in the TA stacked region, changing ω' from 168° to about −100°. The orientation of the base pairs between successive duplexes was also changed from 31° to about 40°. Figure 3 gives a general view of the model. The central portion is a segment of the poly(dA–dT)·poly(dA–dT) model and regions *a* and *b* are the duplex fragments in the positions found in the crystal structure. The model is a qualitative one and a variety of double-helical structures can be built from the ApT fragment by varying the torsion angle ω' (O3′-P) linking adjacent dinucleotides. The relative plausibility of a number of such models has been investigated by energy minimisation calculations (A. Klug, A. Jack, M. A. V., O. K., Z. S. and T. A. Steitz, in preparation).

The most interesting feature of the model is the sugar–phosphate backbone, where the deoxyribose conformation alternates between C2′-*endo* and C3′-*endo* and the conformation of the phosphodiester bridge changes significantly between the AT and TA regions. The glycosidic angles for purine and pyrimidine bases are also significantly different (0° and 70°). The model is characteristic of a deoxyribose structure, as the substitution of a hydroxyl group on C2′ of the thymidine moiety of the (dT–dA) fragment is severely hindered sterically. An analogous DNA–RNA hybrid structure is not, therefore, feasible. This would suggest that stretches of poly(dA–dT)·poly(dA–dT) regions in satellite DNAs[14] have a structural rather than a transcriptional role[9,15].

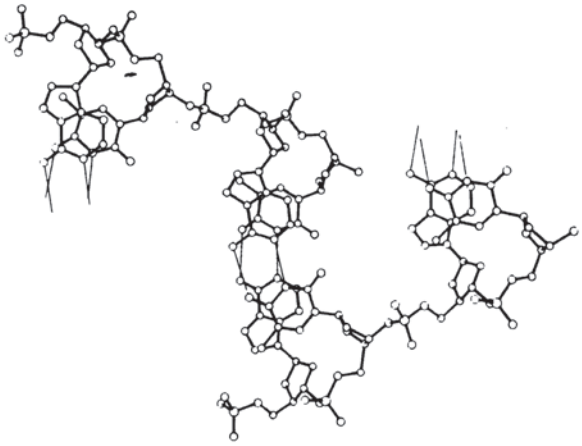


Fig. 2 The mini-helix viewed perpendicular to the base pairs. Hydrogen bonds are indicated by thin lines.

There is evidence that AT-rich regions have other specific biological roles, for example, the preferential binding of *lac* repressor to such regions or to poly(dA–dT)·poly(dA–dT) itself[4].

It has been postulated[16,17] that these regions may have specific geometric features which can be selectively recognised by proteins involved in the coding process or its gene regulation. The crystal structure suggests a possible way in which such geometric features may arise. A future paper (A. Klug, A. Jack, M. A. V., O. K., Z. S. and T. A. Steitz, in preparation) brings together the evidence of the tetranucleotide crystal structure, the fibre diffraction data and various physico-chemical properties, all of which point to a rather special structure for double-stranded poly(dA–dT). A rationale for this structure is given, which provides an explanation for the binding properties of poly(dA–dT)·poly(dA–dT) to the *lac* repressor of *Escherichia coli*.

Our analysis was based on 2,717 independent, significant structure factors (4,528 diffractometer measurements) and a unit cell of dimensions $a = 21.121(10)$, $b = 21.294(14)$, $c = 770(4)$ Å, $\beta = 97.84(4)°$. The structure was solved by a combination of indirect and direct methods and refined to $R = 15\%$ (isotropic). The water structure is disordered. Full details will be published elsewhere.

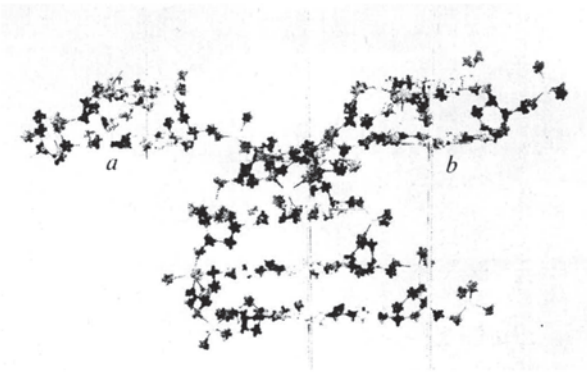


Fig. 3 A photograph of a model of poly(dA–dT) based on the mini-helix: regions *a* and *b* indicate duplex fragments found in the crystal structure.

We thank the MRC for a long-term grant, including a visiting fellowship to M. A. V., the SRC for the diffractometer, the University of Cambridge for computer time, the DST and DAE, India (M. A. V.), and the Weizmann Institute (Z. S.) for financial support, our colleagues in Cambridge, particularly Drs D. Brown, A. Jack, A. Klug and J. R. Rodgers for advice and discussions and Dr F. H. C. Crick for suggesting the problem. O. K. is on the external staff of the MRC. Programs written by Dr W. D. S. Motherwell were used for the illustrations.

M. A. VISWAMITRA
Department of Physics and
ICMR Unit on Genetics and Cell Biology,
Indian Institute of Science,
Bangalore, India

OLGA KENNARD
PETER G. JONES
GEORGE M. SHELDRICK
STEPHEN SALISBURY
LARRY FALVELLO
University Chemical Laboratory,
Lensfield Road,
Cambridge, UK

Z. SHAKKED
Department of Structural Chemistry,
Weizmann Institute of Science,
Rehovot, Israel

Received 24 November 1977; accepted 28 April 1978.

1. Rosenberg, J. M. *et al. Nature* **243,** 150–154 (1973).
2. Day, R. O., Seeman, N. C., Rosenberg, J. M. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **70,** 849–853 (1973).
3. Hingerty, B. *et al. Acta Crystallogr.* B**32,** 2998–3013 (1976).
4. Lin, S. Y. & Riggs, A. D. *Nature* **288,** 1185–1186 (1970).
5. Riggs, A. D., Lin, S. & Wells, R. D. *Proc. natn. Acad. Sci. U.S.A.* **69,** 761–764 (1972).
6. Lin, S. & Riggs, A. D. *Biochem. biophys. Res. Commun.* **45,** 1542–1547 (1971).
7. Wingert, L. & von Hippel, P. H. *Biochim. biophys. Acta* **157,** 114–126 (1968).
8. Watson, J. D. & Crick, F. H. C. *Nature* **171,** 737–738 (1953).
9. Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C. & Watts, L. *J. molec. Biol.* **88,** 523–533 (1974).
10. Sundaralingam, M. *Jerusalem Symp. Quantum Chem. Biochem.* **5,** 417–455 (1973).
11. Sussman, J. L., Seeman, N. C., Kim, S. H. & Berman, H. M. *J. molec. Biol.* **66,** 403–421 (1971).
12. Rubin, J., Brennen, T. & Sundaralingam, M. *Biochemistry* **11,** 3112–3228 (1972).
13. Camerman, N., Fawcett, J. K. & Camerman, A. *J. molec. Biol.* **107,** 601–621 (1976).
14. Laskowski, M. *Prog. Nucleic Acid Res. molec. Biol.* **12,** 161–188 (1972).
15. Arnott, S., Fuller, W., Hodgson, A. & Prutton, I. *Nature* **220,** 561–564 (1968).
16. Bram, S. *Nature new Biol.* **232,** 174–176 (1971).
17. Richmond, T. J. & Steitz, T. A. *J. molec. Biol.* **103,** 25–38 (1976).