J. Indian Inst. Sci., Mar.–Apr. 1993, **73**, 181–192.

Recent studies on crystal structures of interest to nucleic acids*

M. A. VISWAMITRA
Department of Physics, Indian Institute of Science and Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 012, India.

Abstract

A detailed knowledge of nucleic acid structures requires a study at the level of the component nucleoside/nucleotide units as well as at the oligomer/polymer level. Crystallographic studies of nucleosides and nucleotides have often provided in the past a wealth of information concerning nucleic acid structures and interactions. We mention here some of our recent studies in this area and also some of the preliminary X-ray studies we have carried out on deoxydecanucleotides, as part of our program on DNA structures.

Key words: X-ray diffraction, crystal structures, nucleosides, nucleotides, oligonucleotides.

1. Introduction

For some years we have been analysing the molecular structures of DNA fragments using the methods of single-crystal X-ray analysis. One of our aims has been to determine whether different DNA sequences possessed any distinctive structural features of their own. A knowledge of the way in which specific base pairs and sequences affect the fine details of the DNA double helix structure is necessary to our understanding of how nucleotide sequences are selectively recognised by proteins. We began our studies[1–4] in the early 1970s at a time when hardly any single-crystal X-ray studies had been done on deoxyribonucleotides although there were several such investigations on ribonucleotides. These studies led to crystallization of the deoxy tetranucleotide d(pATAT). The crystal structure[5] gave the first description of Watson–Crick base pairs in a right-handed DNA duplex fragment at atomic resolution. The X-ray study also showed a base-dependent structure and suggested[5–7] a dinucleotide as a repeat unit for poly (dA–dT) unlike in classical DNA models.

* Text of lecture delivered at the Annual Meeting of the Faculty of the Jawaharlal Nehru Centre for Advanced Scientific Research at Bangalore on November 20, 1992.

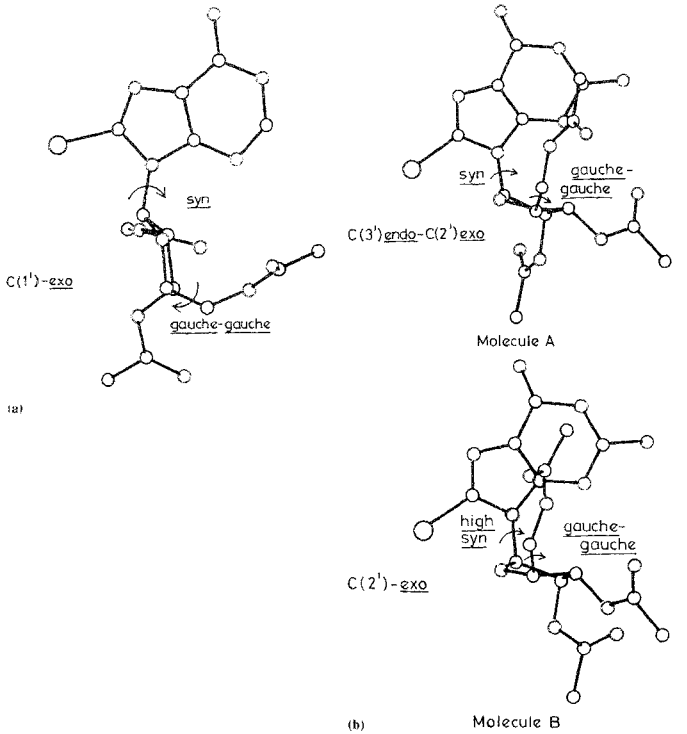

FIG. 1. Molelcular conformations of 8-bromo 2',3',5'-triacetyl adenosine and 8-bromo 2',3',5' triacetyl guanosine. (a) 8-Br Tri A, (b) The two molecules of 8-Br Tri G.

2. Experimental

2.1. *Influence of acetyl and bromine substitutions on stacking: Crystal and molecular structures of 8-bromo 2',3',5'-triacetyl adenosine and 8-bromo 2',3',5'-triacetyl guanosine*

Crystal structure studies on 2',3',5'-triacetyl adenosine and 2',3',5'-triacetyl guanosine (abbreviated Tri A and Tri G) and 8-bromo adenosine and 8-bromo

Table I
Crystal data

		8-Br Tri A	*8-Br Tri G*
Formula	:	C_{16} H_{18} N_5 O_7 Br	C_{16} H_{18} N_5 O_8 Br
Molecular weight	:	472·25	488·25
Space group	:	$P2_12_12_1$	P1
a	:	7·855 (1) Å	8·201 (1) Å
b	:	15·047 (2) Å	8·753 (1) Å
c	:	16·970 (1) Å	15·700 (2) Å
α	:	90·0°	74·20 (1)°
β	:	90·0°	72·54 (1)°
γ	:	90·0°	69·76 (1)°
Volume	:	2013·42 Å^3	990·66 Å^3
z	:	4	2

guanosine (8–Br A and 8–Br G) have recently brought to light[8–11] the individual influence of acetyl groups and bromine atoms on molecular conformation and interactions. The X-ray studies on the present compounds (8–Br Tri A and 8–Br Tri G) were taken up to investigate the effect of these groups when they are present together within the same molecule[12].

Cuboidal crystals of the samples (Sigma Chemicals) were obtained by slow evaporation of ethanol–water mixtures heated to 35°C. The three-dimensional CuKα intensity data for the crystals were collected on a CAD-4 diffractometer and the structures were solved using Patterson and difference Fourier synthesis. The crystal data are given in Table I. Figure 1 shows the molecular conformations. A conformational comparison of 8-Br Tri A and 8-Br Tri G with the unbrominated and unacetylated analogues is given in Table II.

A.A.A. base triplet: With the ribose hydroxyl blocked, the only hydrogen bond

Table II
Conformations of 8-Br Tri A and 8-Br Tri G compared with unbrominated and unacetylated analogues

Compound	*syn/anti*	*sugar pucker*	*C(4′)–C(5′) bond*
8-Br Tri A	*syn*	C(1′)-*exo*	*gauche–gauche*
Tri A	*syn*	C(2′)-*endo* C(3′)-*exo*	*gauche–trans*
8-Br A	A:*syn*	C(2′)-*endo*	*gauche–gauche*
	B: high *syn*	C(2′)-*exo*	*gauche–gauche*
Tri G	A:*anti*	C(2′)-*endo* C(1′)-*exo*	*gauche–gauche*
Orthorhombic	B:*anti*	C(3′)-*exo* C(4′)-*endo*	*gauche–gauche*
Tri G monoclinic	*anti*	C(2′)-*endo*	*gauche–gauche*
8-Br G	*syn*	C(2′)-*endo*	*gauche–gauche*

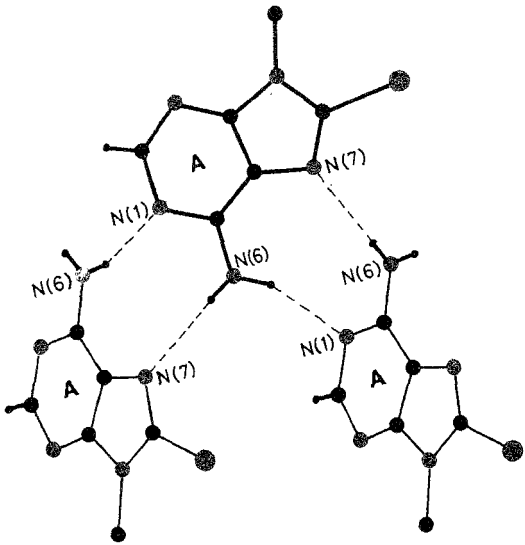

FIG. 2. A.A.A. base triplets observed in 8-bromo triacetyl adenosine.

donors are N(6) in the adenosine and N(1) and N(2) in the guanosine derivative. The acetyl group at O(5′) in the present structures excludes the possibility of O(5′)–H...N(3) hydrogen bond often seen in 8-substituted nucleosides. The N(6) atom of 8-Br Tri A is hydrogen bonded to the N(1) and N(7) atoms of symmetry-related molecules. Thus, both Watson–Crick and Hoogsteen sites of adenine are used giving rise to A.A.A. base triplets (Fig. 2) similar to those found in Tri A, 8-bromo 2′, 3′-O-isopropylidene adenosine[13] and 2′-deoxyadenosine[14]. Recent years have seen considerable interest in nucleic acid structures containing triple helices. The A.A.A. triplet geometry found in the crystals is analogous to U.AU/T.AT/A.AT H-bonding schemes proposed for the triple helices[15–18].

8-Br Tri A does not show any base–base stacking while in 8-Br Tri G, the acetyl group at C(5′) end folds back so that C(35′) methyl folds back over the parent base in a 'scorpion tail' fashion. The observed differences in the nature of acetyl-base stacking in the two structures are an interesting consequence of bromine substitution. We have also recently determined[19] the crystal structure of 5-bromo cytidine

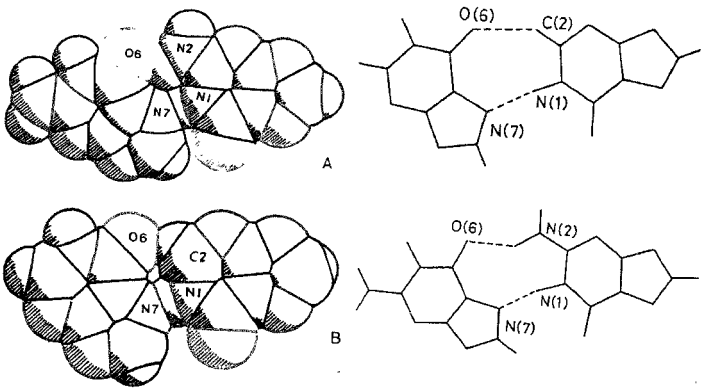

FIG. 3. Hypoxanthine base pair (I.I) observed in the structure of isopropylidene inosine is analogous to the guanine base pair (G.G) observed in isopropylidene guanosine with a C(2)-H . O(6) interaction instead of N(2)-H...O(6) bond contributing to the stability of I.I base pairs

monophosphate. The conformation of the nucleotide is *anti*, C(3')-*endo* and *gauche-gauche* similar to that found in related non-halogenated nucleoside/nucleotides.

2.2. *Base-pairing through C–H..O hydrogen bonds: Crystal structures of 2', 3'-O isopropylidene inosine, 2',3'-O isopropylidene thioinosine and 6,5'-anhydro-6 hydroxy 2',3'-O isopropylidene uridine*

Inosine is a minor constituent of tRNA and is known to occupy the 5'-terminal position of the anticodon triplet in some tRNAs. It differs from guanosine in the absence of an NH_2 group at position 2 of the base. It would be of interest to study the effect of this change on molecular conformation and interactions.

The crystals of the compound were grown by slow evaporation of an aqueous solution. Crystal data are: $C_{13}\ H_{16}\ N_4\ O_5\ 0{\cdot}5\ H_2O$; $P2_12_12_1$; $a = 10{\cdot}895$, $b = 13{\cdot}307$, $c = 40{\cdot}577$ Å. The structure was solved[20] by the use of direct methods using CuKα diffractometer data. The four independent molecules in the unit cell show significant conformational differences. Ribose puckers fall in the O(4')-*exo* region, unfavourable in unsubstituted nucleosides. The hypoxanthines form self base pairs (I.I) analogous to the guanine base pairing (G.G) found in the crystal structures of Iso–G[21], with C(2)–H...O(6) however replacing the N(2)–H...O(6) hydrogen bond as shown in Fig. 3. This unusual base pairing has been seen in only one other inosine structure solved so far[22].

The base pairing pattern in the crystals of 2',3'-O-isopropyledene thioinosine[23,24] is similar to that in Iso-I, with sulphur atom forming C(2)–H...S(6) H–bond as shown

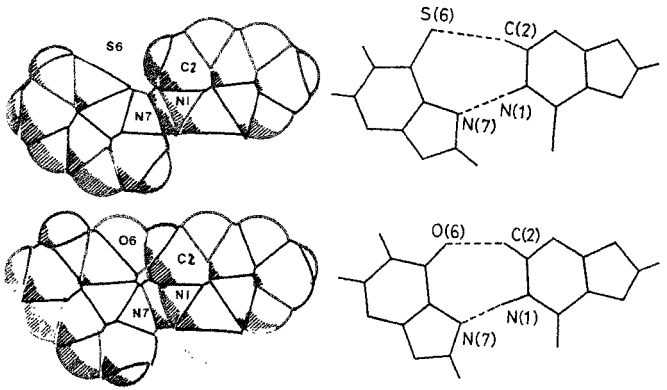

FIG. 4. Crystal structure of 2',3'-O isopropyledene thioinosine. The hypoxanthines form self base pairs through C(2)-H...S(6) and N(1)-H...N(7) hydrogen bonds.

in Fig. 4. We have also observed a base-pairing scheme stabilised by the presence of C–H...O hydrogen bonds, in the crystal structure of 6,5'-anhydro-6-hydroxy-2',3'-O-isopropyledene uridine[25]. This compound crystallizes in the space group $P2_12_12_1$ with a = 10·412, b = 14·936, c = 16·651 Å. The two crystallographically independent molecules in the unit cell exhibit unusual U.U base pairs involving N(3)-H...O(4) and C(5)-H...O(4) hydrogen bonds as shown in Fig. 5, significantly different from the symmetrical pairing scheme (N(3)–H...O(4)/O(2) hydrogen bonds) commonly observed in uracil-containing crystals. The preference of the uracil bases in this crystal to form self base pairs through the weaker C-H...O hydrogen bonds in spite of the availability of N(3) of the base as proton donor is of significance. We have thus in these crystal structures clear evidence of base pairing mechanisms involving C-H...O hydrogen bonds. Presently the framework for generating nucleic acid models involves basically the use of N-H...O and N-H...N hydrogen bonds. Base pairing mechanisms involving C-H...O hydrogen bond, if incorporated into model building, significantly widen the scope for generating novel nucleic acid structures.

3. Crystal structure evidence for O–H...C hydrogen bond

We have also made attempts to understand the nature and significance of hydrogen bonds of the type X–H...C(X=O,N). While C–H...O hydrogen bonds are widespread, the complementary O–H...C interaction is rare because C is not as electronegative as O and also because C atoms are not often situated in sterically unhindered positions (unlike carbonyl and ethereal O atoms which permit easy access by C–H

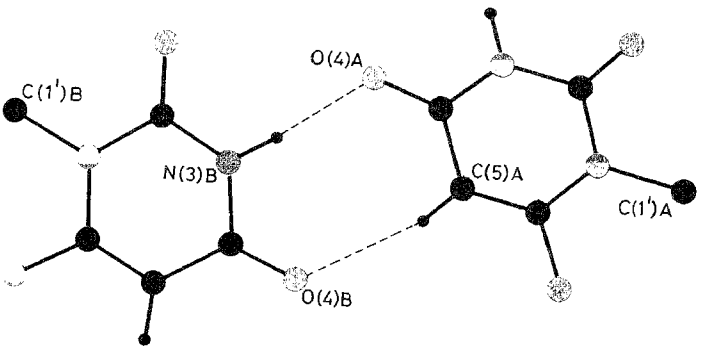

FIG. 5. The two independent molecules in the crystal structure of 2'-3'-O-Isopropylidene uridine exhibit unusual uracil base pair (U.U) involving N(3)-H . . . O(4) and C(5)-H . . . O(4) hydrogen bonds.

groups to form C–H...O bonds). In spite of these limitations, a sufficiently electron-rich C atom (alkyne, alkene, aromatic) has propensity to form a hydrogen bond-like interaction with X–H groups (X=O,N). We have recently seen direct evidence of such an O–H...C hydrogen bond[26] in the crystal structure of the steroidal compound, 17-ethynylylandrost–2, 4-dieno[2,3-D] dihydroxazole–17–ol, donazole.

4. Model compounds for studying protein–nucleic acid interactions: Crystal and molecular structure of 2'-O-monosuccinyl adenosine 3',5'-(cyclic) phosphate monohydrate

The motivation for this work comes from our wider interest in the geometry of interactions between nucleotides and peptides. A knowledge of this is necessary to understand the complex processes of binding of nucleic acids with proteins at a molecular level. Recent crystallographic studies on protein–nucleic acid complexes show the significant progress made in this area of molecular biology. However, the precise geometries of interactions, at atomic level, could be difficult in view of limited resolution of such crystals. There is considerable scope, therefore, for crystal structure studies on complexes of smaller constituents such as nucleotides and amino acids. But, at present the paucity of data on such complexes, reflect the difficulties of growing such crystals suitable for X-ray studies. Structural studies on model compounds, comprising nucleic acid components, covalently linked to an amino acid or amino acid-like moiety, could also make significant contributions to our understanding of the problem.

We have now obtained crystals of two such compounds[27], where the nucleic acid constituent, adenine, is covalently linked to the succinic acid moiety, which resembles the amino acid side chains of glutamic and aspartic acids:

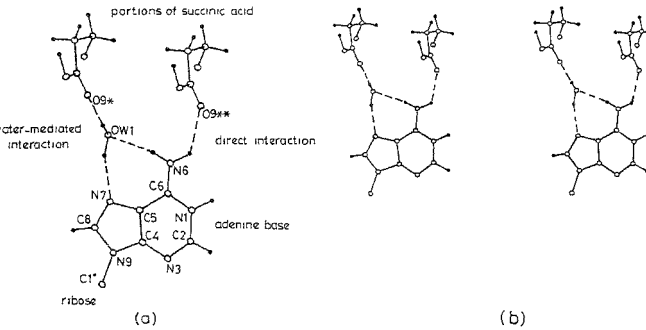

FIG. 6. 2′-O-Monosuccinyl adenosine 3′-5′(cyclic) phosphate monohydrate (cAMP2′): View perpendicular to the base showing the succinic acid moiety interactions with adenine base.

2′-O-Monosuccinyl adenosine 3′-5′-(cyclic) phosphate monohydrate (abbreviated as cAMP2′ suc);

5′-O-Monosuccinyl adenosine dihydrate (abbreviated as A5′suc).

The crystal structure cAMP2′suc[28] exhibits two types of interactions between succinic acid moiety and the cyclic nucleotide: carboxyl group of succinic acid moiety has direct interactions with ester oxygen of the phosphate group and with N(6) of adenine base. It also shows indirect or water-mediated interactions with the nitrogens, N(6) and N(7), of the same base as shown in Fig. 6. Interestingly, similar water-mediated interactions between amino acid side chains and nucleotides have been observed in crystal structures of certain repressor protein–DNA operator complexes. In addition, the structure of A5′suc[29] shows a novel mode of adenine base aggregation through stacking as well as hydrogen bonding. The Hoogsteen sites of adenine base (A1) are found to link two stacked bases (A2 and A3) along the b-axis by a pair of symmetry-related N(6)–H...N(7) hydrogen bonds (Fig. 7). The nature of base–base interactions observed here is totally different from those observed for self-paired adenine bases in nucleoside, nucleotide or polynucleotide structures so far.

5. Results of oligonucleotide crystal structure studies

Crystal structure studies on oligonucleotides, since their first X-ray studies a decade ago[5,30,31], have shown several features of sequence effects in the DNA double helix. However, there is still much to be understood on how such structural changes are brought about. There is also much to be known regarding novel double-helical forms like the parallel helix and the triplex structure which DNA sequences can adopt. Answers to these will take us closer to understanding some of the fundamental

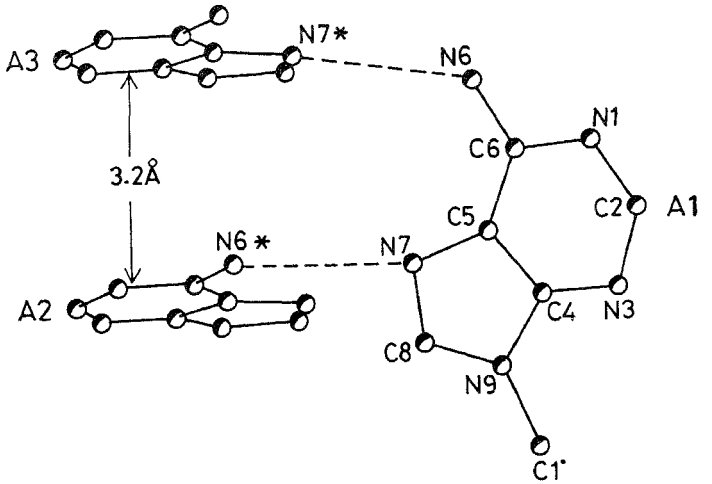

FIG 7. 5'-O-Monosuccinyl adenosine dihydrate (A5'suc) shows novel mode of adenine base aggregation. Adenine base A1 links 2 stacked bases A2 and A3 by a pair of symmetry-related hydrogen bonds.

questions on DNA structure and functions. As part of this study, we have now crystallized in our laboratory several deoxydecanucleotide sequences with systematic changes introduced in their base sequence.

5.1. *d(CCGGGCCCGG) and d(CCCGGCCGGG) crystallize in DNA–A form*

Crystallization was carried out by the vapour diffusion of droplets of solution containing about 0·1 mM DNA, 0·3 mM $MgCl_2$, 0·5 mM sodium cacodylate, 0·1 mM spermine tetrachloride and 15% (v/v) 2-methyl 2,4-pentane diol (MPD) against 80% (v/v) MPD in a reservoir. Precession photographs showed the crystals to be isomorphous with space group $P2_12_12_1$.

The cell dimensions are: $a=24{\cdot}592$, $b=43{\cdot}535$ and $c=46{\cdot}689$ Å for d(CCGGGCCCGG) and $a=24{\cdot}306$, $b=44{\cdot}314$ and $c=45{\cdot}566$ Å for d(CCCGGCCGGG).

The X-ray intensity data were collected on a CAD-4 diffractometer to 2·5 Å resolution. The preliminary structure solution using rotation and translation functions show an A-DNA type conformation for the two sequences. Further work is in progress[32,33].

5.2. *Parallel-stranded double helix: X-ray analysis of cytidylyl-2',5'-adenosine*

X-ray crystallographic studies on 3'–5' oligomers have provided a great deal of information on the stereochemistry and conformational flexibility of nucleic acids and polynucleotides. In contrast, there is very little information available on 2'–5' polynucleotides. We have now obtained the crystal structure of cytidylyl-2',5'-adenosine (C2'p5A') at atomic resolution[34] which suggests interesting conformational deviations between these two classes of polymers.

The dinucleoside phosphate crystallizes in the monoclinic space group C2, with a = 33·912, b = 16·824, c = 12·898 Å and β = 112·35(1)° with two molecules in the asymmetric unit. The two crystallographically independent C2'-p5'A molecules form right-handed miniature parallel-stranded double helices with their respective two-fold symmetry-related molecules. The cytosines and adenines form self pairs with three and two hydrogen bonds, respectively. The conformation of the C and A residues about the glycosyl bond is *anti* same as in the 3'–5' analog but contrasts the *anti* and *syn* geometry of C and A residues in A2'p5'C[35]. A right-handed 2'–5' parallel-stranded double helix having eight base pairs per turn and 45° turn angle between them can be generated using the mini duplex observed in the crystals, as a repeating unit.

Acknowledgement

I thank the Department of Biotechnology, Government of India, for financial support. The work is the result of contributions my colleagues and students have made over the years and it gives me pleasure to acknowledge them. I particularly thank Profs T.P. Seshadri and S. Ramakumar with whom I have collaborated on many of the studies reported here. I thank Prof. C.N.R. Rao for his support and encouragement.

References

1. VISWAMITRA, M. A., REDDY, B. S., LIN, B. H. Y. AND SUNDARALINGAM, N. — *J. Am. Chem. Soc.*, 1971, **93**, 4565–4573.
2. VISWAMITRA, M. A. AND SESHADRI, T P. — *Nature*, 1974, **252**, 176–177.
3. VISWAMITRA, M. A., SESHADRI, T. P. AND POST, M. L. — *Nature*, 1975, **258**, 542–544.
4. REDDY, B. S. AND VISWAMITRA, M. A. — *Acta Crystallogr. B*, 1975, **31**, 19–26.
5. VISWAMITRA, M. A., KENNARD, O., JONES, P. G., SHELDRICK, G. M., SALISBURY, S., FALVELLO, L. AND SHAKKED, Z. — *Nature*, 1978, **273**, 687–688.

6. KLUG, A., JACK, A., VISWAMITRA, M. A., KENNARD, O., SHAKKED, Z. AND STEITZ, T. A. — *J. Mol. Biol.*, 1979, **131**, 669–680.

7. VISWAMITRA, M. A., KENNARD, O., SHAKKED, Z., JONES, P. G., SHELDRICK, G. M., SALISBURY, S. AND FALVELLO, L. — *Curr. Sci.*, 1978, **47**, 289–292.

8. WILSON, C. C., LOW, J. N. AND TOLLIN, P. — *Acta Crystallogr. C*, 1985, **41**, 1123–1125.

9. WILSON, C. C., TOLLIN, P. AND HOWIE, R. A. — *Acta Crystallogr. C*, 1986, **42**, 697–700.

10. LOW, J. N., TOLLIN, P., WILSON, C. C. AND SCRIMGEOUR, S. N — *Acta Crystallogr. C*, 1986, **42**, 700–702.

11. TAVALE, S. S. AND SOBELL, H. M. — *J. Mol. Biol.*, 1970, **48**, 109–123.

12. MANDE, S S., LALITHA, H. N., RAMAKUMAR, S. AND VISWAMITRA, M. A. — *Nucleosides Nucleotides*, 1992, **11**, 1089–1101.

13. FUJII, S., FUJIWARA, T. AND TOMITA, K.–I. — *Nucl. Acids Res.*, 1976, **3**, 1985–1996.

14. RADWAN, M. M. AND WILSON, H. R. — *Acta Crystallogr. B*, 1980, **36**, 2185–2187.

15. STROBEL, S. A. AND DERVAN, P. B. — *Nature*, 1991, **350**, 172–174.

16. PILCH, D. S., LEVENSON, C. AND SHAFER, R. H. — *Proc. Natn. Acad. Sci.*, 1990, **87**, 1942–1946

17. ARNOTT, S., BOND, P. J., SELSING, E. AND SMITH, P. J. C. — *Nucl. Acids Res.*, 1976, **3**, 2459–2470.

18. SANTOS, C., ROSEN, M. AND PATEL, D. — *Biochemistry*, 1989, **28**, 7282–7289.

19. MANDE, S. S., SESHADRI, T. P. AND VISWAMITRA, M A. — *Acta Crystallogr. C*, 1993 (under publication).

20. MANDE, S. S., SHAMALA, N., SESHADRI, T. P. AND VISWAMITRA, M. A. — *Nucleosides Nucleotides*, 1992, **11**, 1103–1114.

21. MANDE, S. S., SESHADRI, T. P. AND VISWAMITRA, M. A. — *Acta Cystallogr. C*, 1989, **45**, 92–94.

22. HAKOSHIMA, T., OMORI, H., TOMITA, K.–I., MIKI, H. AND IKEHARA, M. — *Nucl. Acids Res.*, 1981, **9**, 711–729.

23. MANDE, S. S. — *X-ray Crystallographic studies on modified nucleosides*, Ph.D. Thesis, Indian Institute of Science, Bangalore, India, 1991.

24. MANDE, S. S., SESHADRI, T. P. AND VISWAMITRA, M. A. — Paper under preparation.

25. MANDE, S. S., SESHADRI, T. P. AND VISWAMITRA, M. A. — *Acta Crystallogr. C*, 1988, **44**, 138–141.

26. VISWAMITRA, M. A., RADHAKRISHNAN, R., BANDEKAR, J. AND DESIRAJU, G. R — *J. Am. Chem. Soc.*, 1993 (in press).

27. PADMAJA, N. — *Small molecules of biological and chemical interest: X-ray crystallographic studies and space group frequency analysis*, Ph.D. Thesis, Indian Institute of Science, Bangalore, India, 1991.

28. PADMAJA, N., RAMAKUMAR, S. AND VISWAMITRA, M. A. — *Bull. Chem. Soc. Jap.*, 1991, **64**, 1359–1363.

29. PADMAJA, N., RAMAKUMAR, S. AND VISWAMITRA, M. A. — Paper under preparation.

30. WANG, A. H.-J., QUIGLEY, G. J., KOLPAK, F. J., CRAWFORD, J. L., VAN BOOM, J. H., VAN DER MAREL, G. AND RICH, A. — *Nature*, 1979, **282**, 680–686.

31. WING, R., DREW, H., TAKANO, T., BROKA, C., TANAKA, S., ITAKURA, K. AND DIKERSON, R. E. — *Nature*, 1980, **287**, 755–758.

32. CHAMPION, D., KUMAR, C. S., RAMAKRISHNAN, B. AND VISWAMITRA, M. A. — *Proc. DNA Trends Conf.*, 1993, pp 12, 13, Bangalore.

33. VISWAMITRA, M. A., GAUTHAM, N., CHAMPION, D. AND KUMAR, C. S. — *Proc. Asian Crystallographic Association*, 1992, 16 U-62, Singapore.

34. KRISHNAN, R., SESHADRI, T. P. AND VISWAMITRA, M. A. — *Nucl. Acids Res.*, 1991, **19**, 379–384.

35. PARTHASARATHY, R., MALIK, M. AND FRIDEY, S. M. — *Proc. Natn. Acad. Sci.*, 1982, **79**, 7292–7296.