A HYPOTHESIS FOR A DOUBLE HELIX IN WHICH THE TWO POLYNUCLEOTIDE STRANDS INTERCALATE INSTEAD OF FORMING BASE-PAIRS

M. A. VISWAMITRA AND J. PANDIT
Department of Physics and ICMR Centre on Genetics and Cell Biology, Indian Institute of Science, Bangalore 560 012, India.

ABSTRACT

A proposal is made for a double helical structure in which the two polynucleotide strands intercalate. The repeat unit consists of two nucleotides which show interstrand overlap instead of *forming base-pairs.*

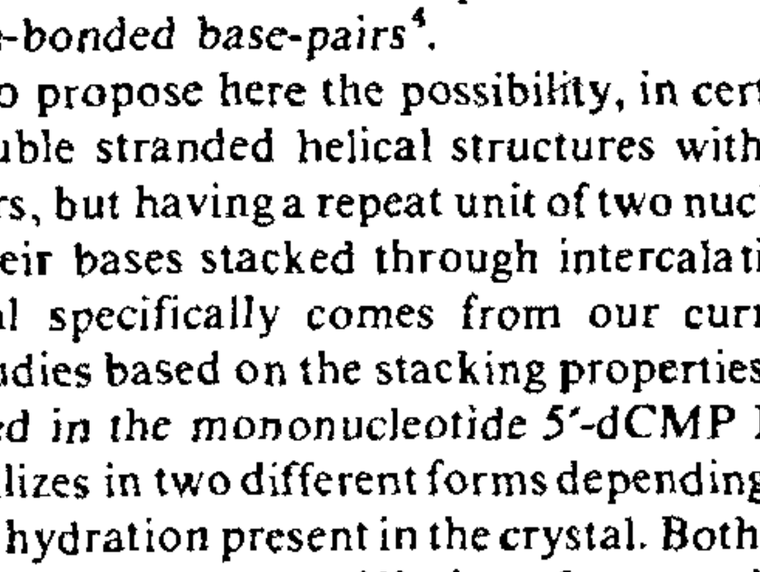
INTRODUCTION

DOUBLE helices since the celebrated discovery of the DNA double helix by Watson and Crick[1] represent the single most important secondary structural form of nucleic acids. In generating these structures one normally uses a mononucleotide as a repeat unit although dinucleotide repeats have now been indicated in certain cases as was first suggested by the crystal structures of the tetranucleotide d(pATAT)[2] and the hexamer d(CGCGCG)[3]. The secondary structures of a variety of polynucleotide helices have now been well characterised based on repeat units of suitable *hydrogen-bonded base-pairs*[4].

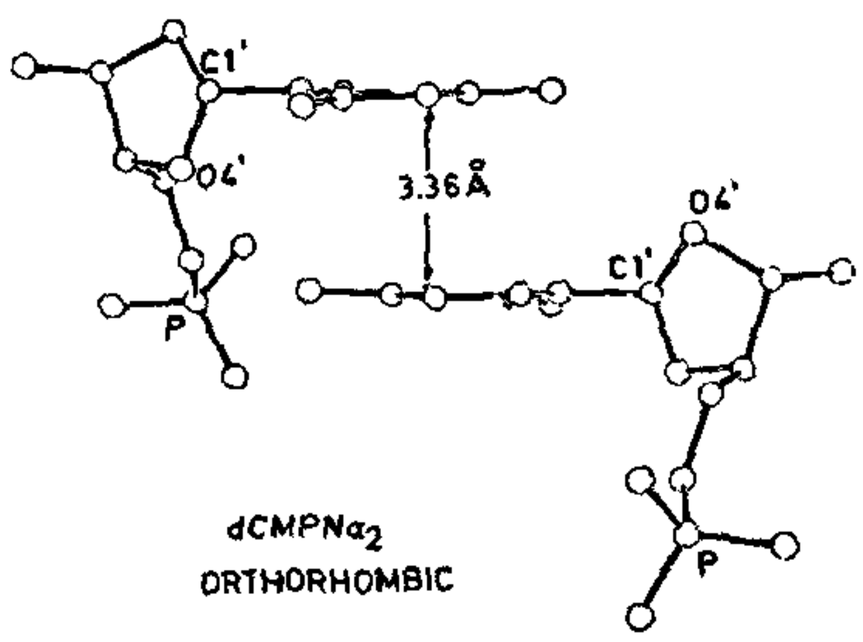
We wish to propose here the possibility, in certain cases, of double stranded helical structures without any base-pairs, but having a repeat unit of two nucleotides with their bases stacked through intercalation. The proposal specifically comes from our current modelling studies based on the stacking properties we have observed in the mononucleotide 5′-dCMP Na_2 which crystallizes in two different forms depending on the degree of hydration present in the crystal. Both the crystal forms were grown by diffusion of acetone into aqueous solutions of the nucleotide.

5′-dCMP Na_2 7H_2O

The crystals are orthorhombic with space group $P2_12_12_1$, a = 6.719, b = 29.614 and c = 9.559 Å. The structure is solved[5] to R = 4.5% using CuKα diffrac-



1a



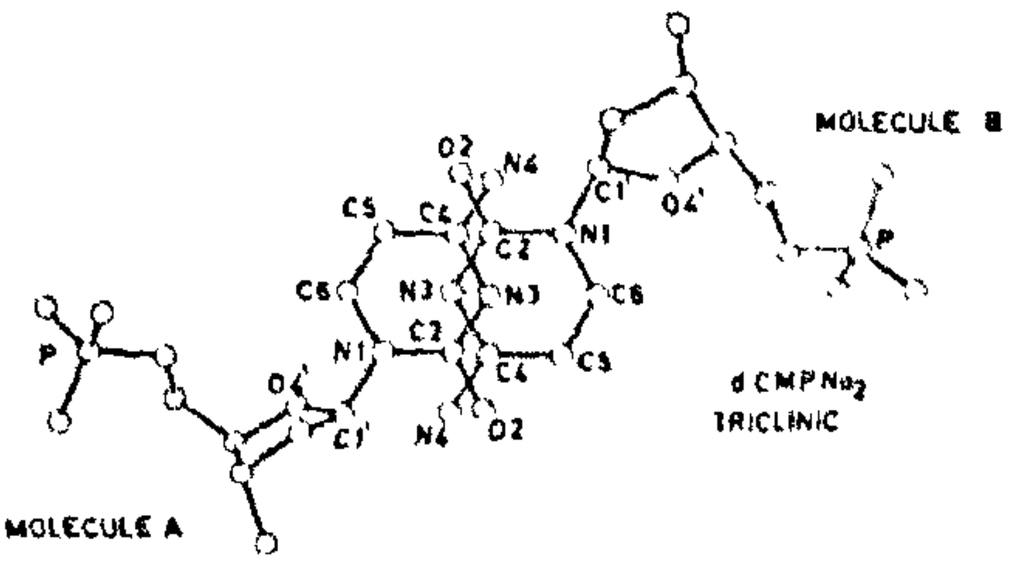
1b

Figure 1 *Two views of the base-stacking seen in the crystal structure of* $dCMPNa_2 . 7H_2O$ (orthorhombic).

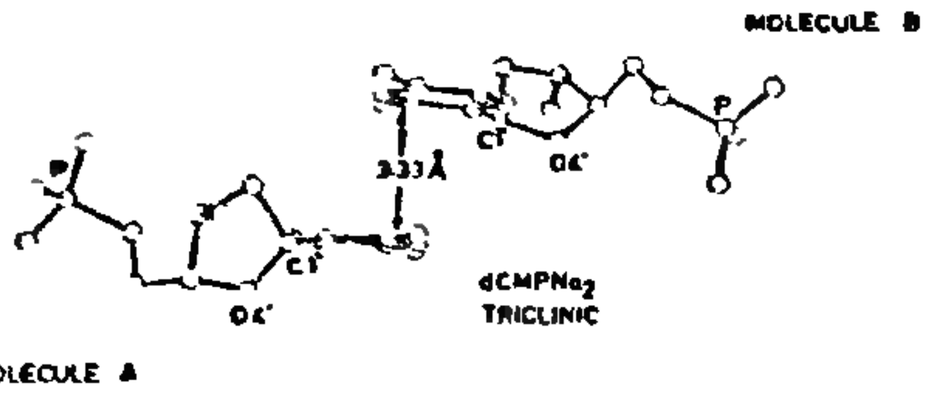
tometer data. The cytosine bases related by the 2-fold screw along the a-axis are stacked at 3.36 Å with considerable overlap as shown in figure 1. Besides this base-base interaction, the structure shows a base-ribose contact with N(4)-O(4′) = 2.915 Å.

5′-dCMP Na_2 11 H_2O

The crystals have the triclinic space group P1 with a = 7.306, b = 10.055, c = 16.670 Å, α = 101.93, β = 93.07 and γ = 90.89°. The present R factor for 4851 CuKα reflections[6] is 0.074. There are two independent dCMP molecules in the cell. The cytosine rings overlap with an average separation of 3.33 Å as shown in figure 2. There are no base-sugar contacts.



2a



2b

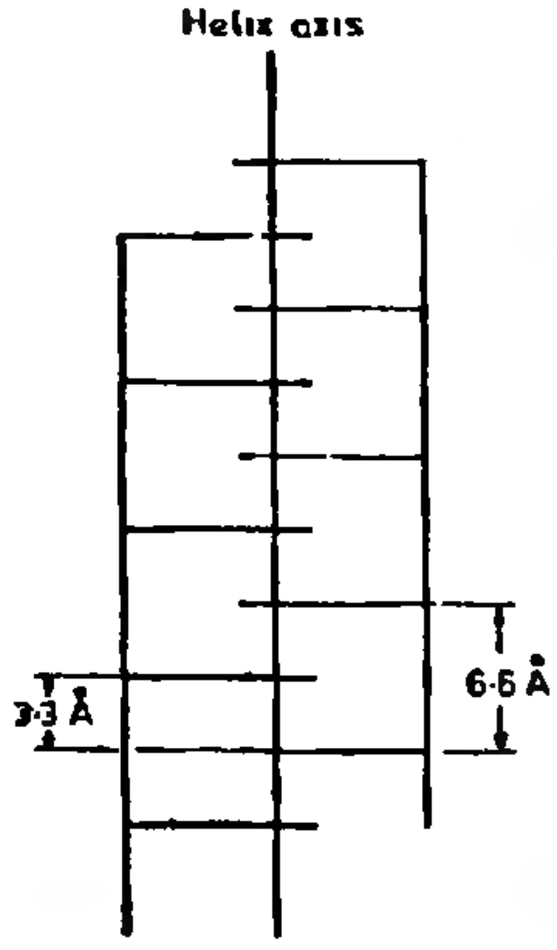
Figure 2 Two views of the base-stacking seen in the crystal structure of $dCMPNa_2.11H_2O$ (triclinic).

Bases related by a cell translation form the hydrogen bond, N(4)-O(2)=3.0 Å. The overlap geometry observed in the crystal is an interesting variant of the interstrand overlap of the cytosine bases observed in the left handed Z DNA double helix[3]. It is very similar to the stacking pattern found energetically favourable for the cytosine free base[7].

In both the crystal structures, the solvent water molecules extensively form hydrogen bonds with themselves as well as the nucleotides. The Na^+ ions are completely surrounded by water oxygens and do not have any nucleotide atoms in their coordination shells.

DOUBLE HELICAL MODELS WITH INTERSTAND BASE-BASE OVERLAP THROUGH INTERCALATION

We have now tried to exploit the nature of interactions observed in the crystals of 5'-dCMP to build a possible double stranded model for poly (dC), using a stacked pair of nucleotides as the repeat unit. Our interest in this exercise partly stemmed from the structural studies available on the helical forms of the RNA polymer poly (C) Several well ordered secondary structures have now been proposed for this polymer. Arnott *et al.*[8] have recently shown that the crystal form of poly (C) fibres contain single stranded helices instead of the double stranded molecules proposed by Langridge and Rich[9]. They find that the double helix models of poly (C) obtained by incorporating different base-pairing schemes are all stereochemically unacceptable because of many overshort interatomic contacts between neighbouring molecules. Such helices have diameters invariably greater than 16 Å much great than the diameter of 13.4 Å found in the crystals of poly (C). Double stranded models with bases interleaved as proposed here have much smaller diameters than base-paired helices and we hoped therefore might have some relevance. For example, with the



Figures 3 & 4 Right and left handed intercalating models for poly (dC) obtained by using the base overlap as seen in the tricilinic crystal.

Inset Schematic drawing of the intercalating double helix.

cytosine rings overlapping as observed in the crystal structures, the P-P distance between the opposing nucleotides is only about 13.5 Å. In the present models, the helix is stabilized entirely by base-stacking without involving any base-pairing scheme. However, unlike in single stranded polynucleotide models where adjacent bases on the same strand overlap, in the present duplex, the overlap of bases is entirely interstrand.

Figures 3 and 4 illustrate the models generated using the base stacking pattern seen in the triclinic crystal. Both right and left handed models with parallel chains were found to build themselves in a smooth fashion with a turn angle of 30° per repeat unit. The cytosine bases are perpendicular to the helix axis which passes through a point midway between the N3 atoms of the overlapping bases. The separation between successive bases is 3.3 Å the distance found in the crystal structures. The models shown no unreasonable stereochemical contacts. The glycosidic torsion for the RH model is about 60° while that of the LH structure is high anti (about 120°). Left handed duplexes with low anti χ values (about $-10°$) also seem to be possible. These values are only approximate estimates. The sugar pucker is C2′-*endo* for all residues and the base separation along each strand is 6.6 Å because of intercalation. The conformation about C4′-C5′ bond is *gauche-trans*, a geometry found in the triclinic crystals of 5-dCMP. The changes in the internucleotide phospho-diester geometry are somewhat analogous to those found at the intercalation site in some drug-dinucleoside crystal structures, where the backbone is stretched in order to separate the base-pair from 3.5 to 7.0 Å.

We have outlined here only the broad structural features of the initial models we have generated using base-stacked rather than base-paired repeat units. These structures are considerably more compact compared to base-paired helices. The present models have a diameter of about 13.5 Å. We plan to investigate their stereochemical details further by computer modelling with variations of base tilt, base overlap, turn angles etc. and to examine any possible relevance of these structures to the great deal of physicochemical and structural data now available on polynucleotides such as poly(C) where optimal base stacking can be a major force in stabilizing the secondary structures.

The discovery of the left handed Z-DNA[3] has dramatically extended the range of helical conformations that polynucleotides can adopt. Crystallographic studies of drug-dinucleoside complexes further point to the intrinsic torsional flexibility of nucleic acids to extend in order to accommodate a variety of molecules through intercalation. The present models make use of this inherent conformational flexibility of the sugar-phosphate backbone. Base-base stacking interactions through intercalation can play a significant role in bringing polynucleotide strands together. A hint of this contribution is seen in the structure of tRNA where extensive intercalation of bases is found to stabilize certain segments of the nucleic acid[9].

We thank the Department of the Science and Technology, India, for financial support and Prof. S. Ramaseshan for his kind interest and encouragement.

14 February 1983

1. Watson, J. D., and Crick, F.H.C., *Nature* (*London*)., 1953, **171**, 737.
2. Viswamitra, M. A., Kennard, O., Jones, P. G., Sheldrick, G.M.S., Salisbury, S., Falvello, L. and Shakked, Z., *Nature* (*London*)., 1978, **273**, 687.
3. Wang, A.H-J., Quigley, G. J., Kolpak, F. S., Crawford, J. H., van Boom, H.H., van der Marel, G. and Rich, A., *Nature* (*London*)., 1979, **282**, 680.
4. Arnott, S., Chandrasekaran, R. and Selsing, E. In *Structure and conformation of nucleic acids and protein-nucleic acid interactions*. Edited by M. Sundaralingam and S. T. Rao, University Park Press, Baltimore, 1975, p.577.
5. Pandit, J., Seshadri, T. P. and Viswamitra, M.A., *Acta Crystallographica*, 1983, (In press).
6. Pandit, J., Seshadri, T. P. and Viswamitra, M.A. (to be published).
7. Gupta, G. and Sasisekaran, V., *Nucleic Acids Res.*, 1978, **5**, 1639.
8. Arnott, S., Chandrasekaran, R. and Leslie, A.G.W., *J. Mol. Biol.*, 1976, **106**, 735.
9. Langridge, R. and Rich, A., *Nature* (*London*)., 1963, **198**, 725.
10. Rich, A., Quigley, G. J. and Wang, A. H. J. In *Nucleic acid geometry and dynamics*. Edited by R.H. Sarma, Pergamon Press, New York, 1980, p. 273.