CONCEPT OF CONFORMATIONALLY EQUIVALENT HEMINUCLEOTIDES AND ITS IMPLICATIONS IN THE ANALYSIS OF THE FOLDING OF SECONDARY AND TERTIARY STRUCTURES OF NUCLEIC ACIDS

R. MALATHI AND N. YATHINDRA

Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Madras 600 025, India.

ABSTRACT

The secondary structural features of double-stranded helices, both of poly (mononucleotide) as well as poly (dinucleotide) types are illustrated with the aid of distance diagonal plots using the recently developed heminucleotide scheme which affords a common basis for description and analysis of all conformations of nucleic acids. The types of interactions that characterise the major and minor grooves of nucleic acid helices, the variations in their positions and widths, the minimum number of residues required for their occurrence etc. which follow as a direct consequence of variations in the sugar residue conformation (ψ, ψ') of the nucleotide and the internucleotide phosphodiester (ω', ω) are clearly manifested. These are relevant to an understanding of the possible mode of interactions of molecules like peptides, drugs, metal ions, water etc. in helical grooves of nucleic acids.

INTRODUCTION

THE analysis of conformations of nucleic acids is by several orders more complicated compared to other biopolymers because of their inherent chemistry comprising a six chemical bond repeat with a flexible sugar ring in its backbone. Based on certain preferred characteristics of a few chemical bonds recently we suggested[1] that the repeating nucleotide backbone, at least as a first approximation may be conveniently regarded as made up of very nearly identical conformational entities (figure 1) while accounting for all the major sources of variability which govern the mechanics of folding of both secondary as well as tertiary structures of nucleic acids. The utility of such a scheme in evaluation and interpretation of the random coil properties of polynucleotide chains and short, medium and long range interactions associated with various substructures which make up the yeast $tRNA^{phe}$ structure along with the symmetry aspects is already demonstrated[1-6]. Yet another interesting feature of the scheme is the striking similarity in the magnitude of the heminucleotide and the peptide repeat in proteins. This aspect may become useful in better understanding and simulation of models concerning nucleotide dipeptide and nucleic acid—protein interactions[7].

In the present paper, we show the effectiveness of the heminucleotide probe in better understanding and visualisation of the structural features such as the nature and kinds of the so-called grooves that distin-

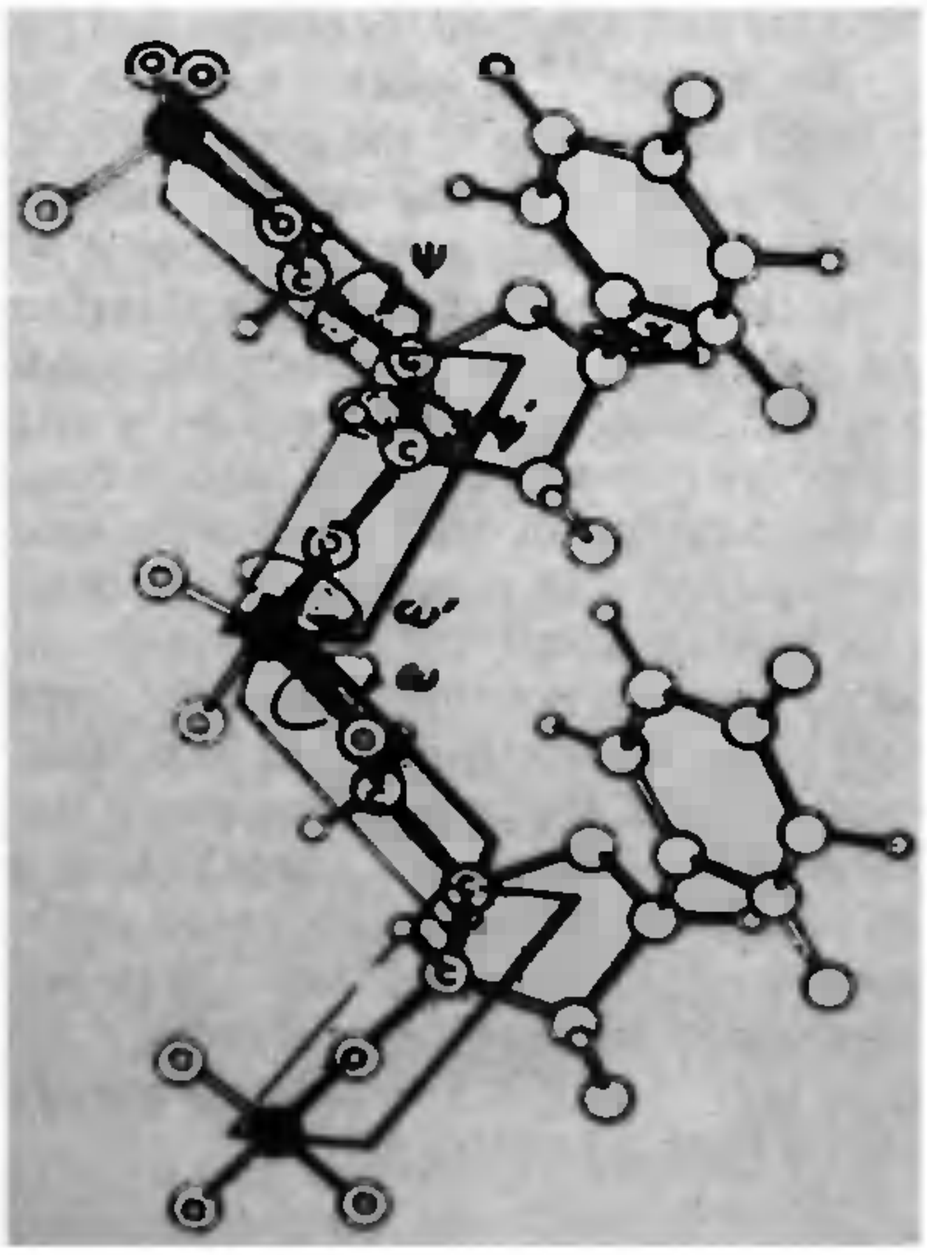

Figure 1: A schematic representation of the concept of conformationally equivalent heminucleotide units characterising the repeating nucleotide of a polynucleotide backbone. The magnitudes of the two heminucleotides are nearly identical (3.9A°) and independent of the backbone (ψ, ψ') and (ω', ω) rotational pairs that determine the mechanics of polynucleotide folding.

guish various secondary helical structures of poly (mononucleotide) and poly (dinucleotide) type. Distance diagonal plots and radial projections of the various classes of known double helical structures have been constructed and compared. The analysis brings out elegantly the properties of the helical grooves and their dependence on the helical parameters and the type of backbone helical repeat the helix is made up of.

METHODS

In constructing the distance plot for a double helix, the two strands are considered as a continuous stretch and accordingly numbered. A double helical segment of twelve base-paired residues (*i.e.* 24 nucleotides and 48 heminucleotides) is considered for the present study. The distances of P and C4′ atoms of every heminucleotide from that of every other heminucleotide of the polynucleotide backbone are plotted and constant value contours of 10, 15 and 20A are drawn. The plot has an inherent symmetry along the diagonal and hence one half would suffice for discussion whereas the other half is utilised for comparison purposes as described earlier[3,5,6] in relation to tRNA plot.

The radial projections of the double helical structures are obtained by cutting open the helix along a line parallel to the z axis, passing through $\phi = 0$, and projecting the atoms on the cylindrical surface. The residual height z is plotted against the product $r\phi$ where r is the radius of the helix and ϕ the azimuthal angle. One full turn of the double helix is considered as this would suffice for description of the secondary structural features such as grooves etc. Since atoms of equal inclination would fall on a straight line the number of residues per turn(n) and the translation height(h) could be directly obtained from these plots. Since both P and C4′ atoms are considered the radial projection plot would appear zig-zag. Nonetheless the loci of successive phosphorous or successive C4′ atoms would result in a straight line. The co-ordinates used to construct distance plots and radial projections for A, B, poly A and Z helical forms are obtained respectively from references 8,9,10.

RESULTS AND DISCUSSION

Distance plots: Figure 2 shows a distance diagonal plot where the lower half and upper half of the diagonal corresponds to A and B form of DNA respectively. In both the forms the central region marked $\mathscr{H}$ encompasses distances of 17-19 A° representing the interaction between base-paired residues. Regions above and below the central strip are due to non

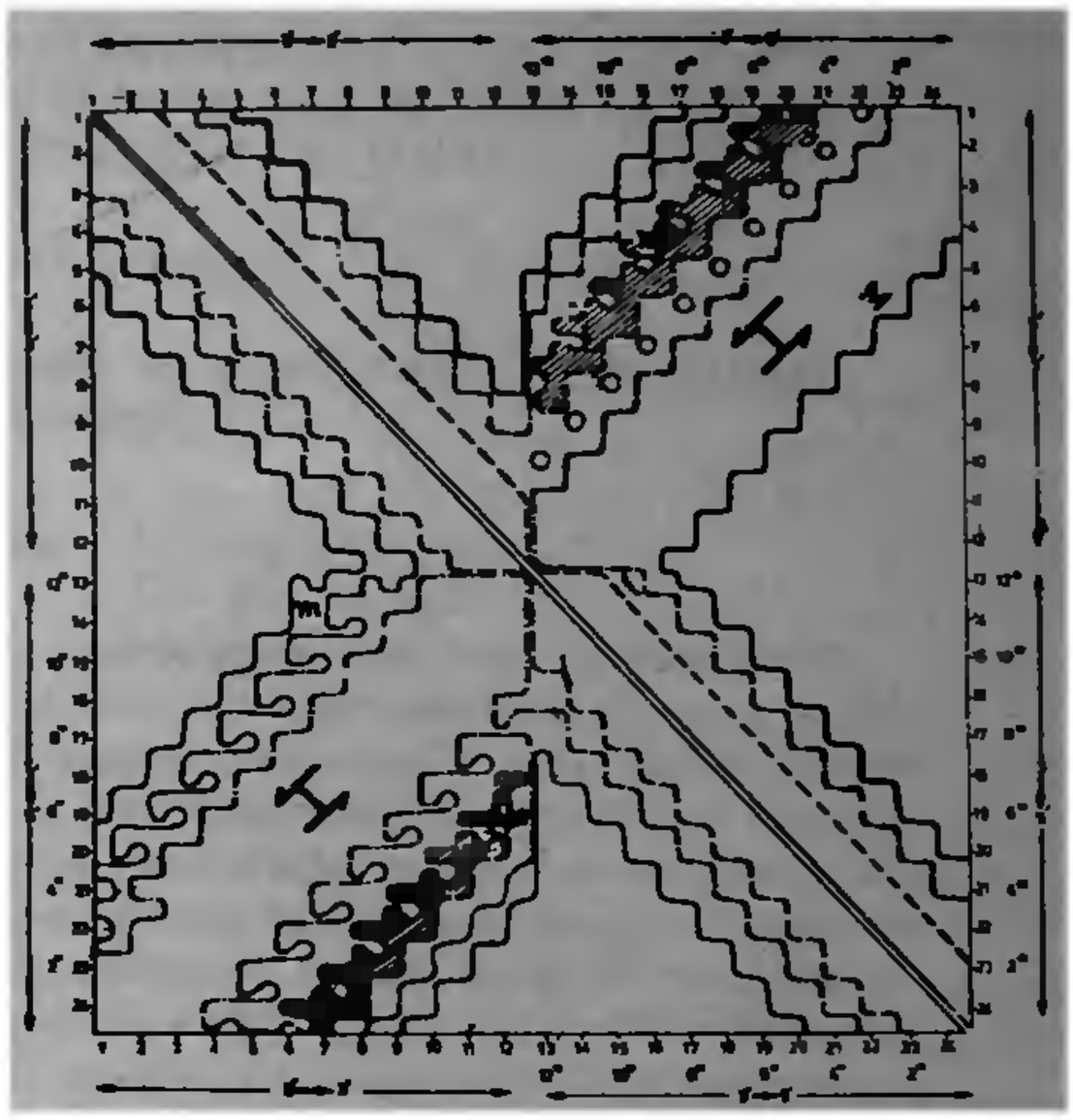

Figure 2: Comparison of the distance diagonal plots obtained for A DNA (lower half and B DNA upper half). In this and the subsequent plots the successive phosphorous atoms are numbered and those with asterisks denote their corresponding complimentary atoms. Interactions between base paired residues are denoted by $\mathscr{H}$ in the central region. The reversal in the width of the major(M) and minor(m) grooves in the A and B forms of DNA is clear.

base-paired interaction between the strands which in fact represent the minor(m) and major(M) grooves in these double helices. The P...P separations in these regions generally define the width of the grooves. A significant observation on the plot is the reversal in the width of minor and major grooves between the two forms reflecting the most important difference between A and B forms of DNA. Distance plot obtained for D DNA would have all the features similar to B DNA (figure 2) but with reduction in the groove widths, a main consequence of the larger twists (45°) of D DNA compared to B form.

To illustrate an example of a parallel double helix, a distance plot (figure 3) has been constructed for polyadenylic acid[9] which exists as a double helix at acidic pH. In contrast to the earlier plot (figure 2), a broad rectangular domain of 20 A° appears parallel to the diagonal instead of being perpendicular. The smaller diameter of the helix is reflected by separations of lower magnitude(12 A°) in the central region $\mathscr{H}$ which denotes the interaction between the base pairs. Yet another feature on the plot is that the separations above and below the base-paired regions are of

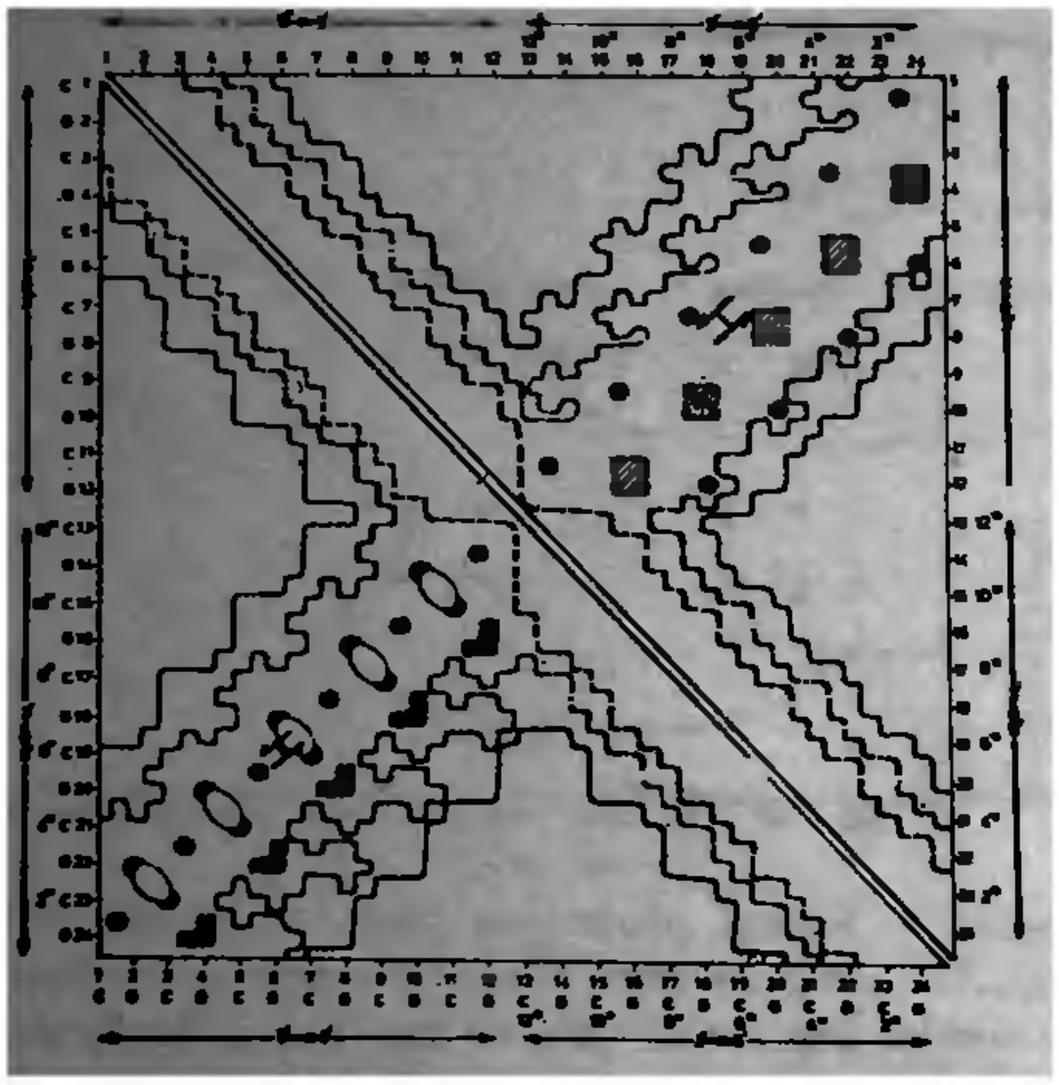

Figure 3: Distance plot for polyadenylic acid. The region $\mathscr{H}$ represents the interaction between the base paired residues. The magnitudes of separation above and below the central region are identical indicating that no distinction such as major or minor groove is possible.

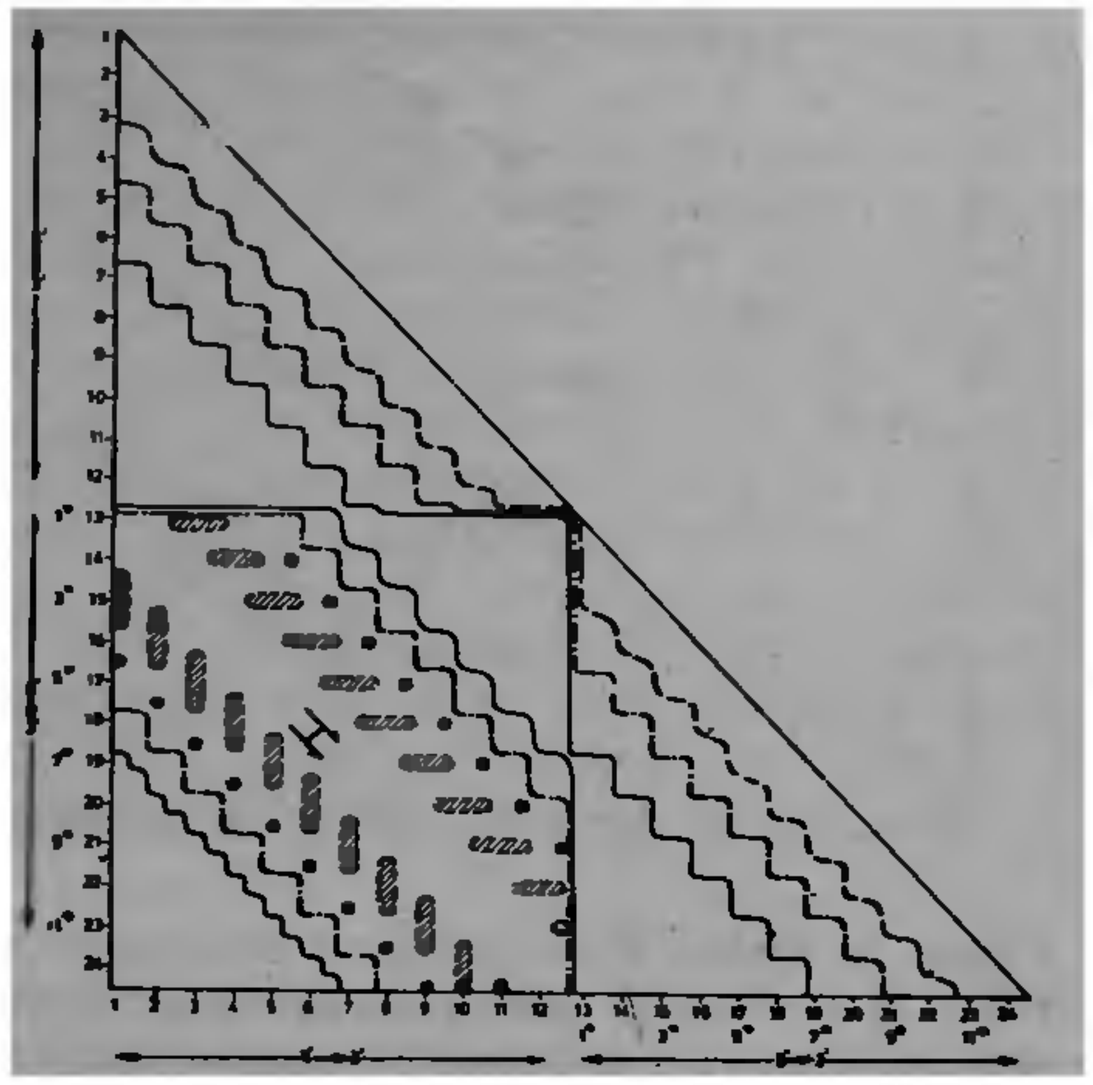

Figure 4: Comparison the diagonal plots of Z_I and Z_{II} forms of DNA. The base sequences are also indicated to facilitate discussion. The central region represents the base paired interactions. Region (m) corresponds to the minor groove. The characteristic dinucleotide' repeat is revealed by repetative patterns occurring after every two residues.

same magnitude and hence no distinction with regard to major or minor groove is possible.

Interestingly, the heminucleotide scheme also finds useful application in the analysis of the recently discovered left handed Z type[10] of poly(dinucleotide) helices since the magnitude of the repeating heminucleotide moiety (P—C4′ and C4′—P) is essentially similar even in all these structures. Distance plot obtained for the Z type of helices is shown in figure 4 where the Z_{II} (lower half) and Z_I (upper half) forms[10] are compared. A striking observation here is the repetition of the patterns after every two residues revealing the characteristic 'dinucleotide' repeat[11]. As in other plots, the central region $\mathscr{H}$ represents the interaction between base-paired residues with the region below this possessing distinct L shaped pattern(10 A°) in Z_{II} form and rectangular blocks (10A°) appearing in the same position in Z_I form. These patterns in fact correspond to the only groove[10] in Z helices namely the minor groove with the P...P separations of 8 and 7 A° in Z_{II} and Z_I respectively. The region above the central strip $\mathscr{H}$ does not manifest distinctly as in A and B, since the distances are similar to the base paired residues. Dumbbell shaped patterns in the central strip $\mathscr{H}$ and also patterns seen between the L shaped patterns in Z_{II} are due to distances slightly greater than 15 A°. Such patterns do not occur in Z_I since the distances are lower than 15 A° and this is indicative of the compactness of Z_I compared to Z_{II}. Note that the minor groove occurs below the region $\mathscr{H}$ in figure 4 unlike in B DNA (figure 2) where the same is seen above the base-paired region.

Radial Projections

For the construction of radial projection plots atoms P and C4′ of the heminucleotide are considered. As in earlier cases, the z-$r\,\phi$ plots are obtained for A, B forms of DNA[8], poly adenylic acid[9] and the Z type[10]. The anionic oxygen atoms are also plotted since their mutual orientations in the major and minor grooves are important for understanding their possible interactions with peptides, drugs, ions, salts, solvents, etc.

All the z-$r\,\phi$ plots are characterised by a central line representing one of the strands of the double helix with sections of the other strand appearing above and below this line. Interactions between the two strands, of non base-paired type, occurring above and below the central line result in the so-called minor and major grooves in these double helices and the P...P separations in these regions would provide a measure of the groove widths. The z-$r\,\phi$ plot constructed for the A form is shown in figure 5. The minor(m) and the major groove(M) are indicated. It could be noted that the

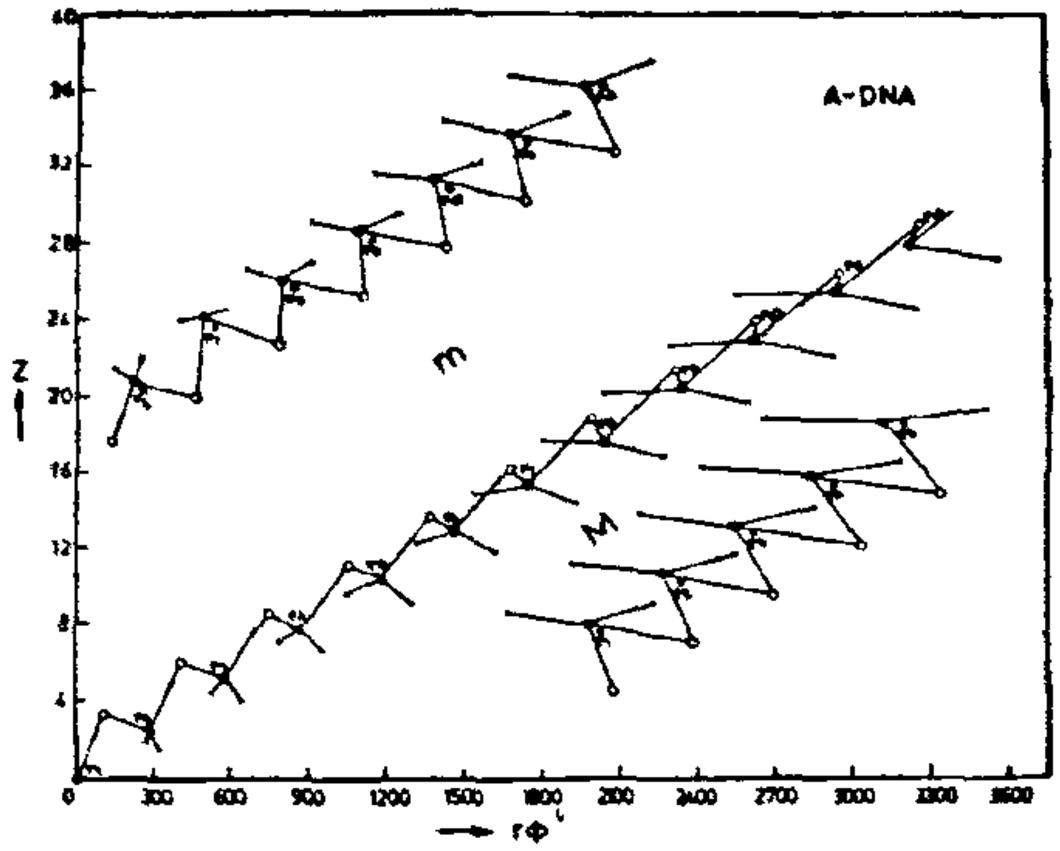

Figure 5: Radial projection plot obtained for A DNA. In this and the other plots to follow the phosphorous atoms are indicated by bigger circles (darkened) and C4′ atom by open circles. The anionic oxygen atoms are also plotted. The minor(m) and major groove(M) are indicated. The larger width of minor groove than the major groove is apparent. Note that the loci of successive phosphorous or carbon C4′ atoms result in a straight line.

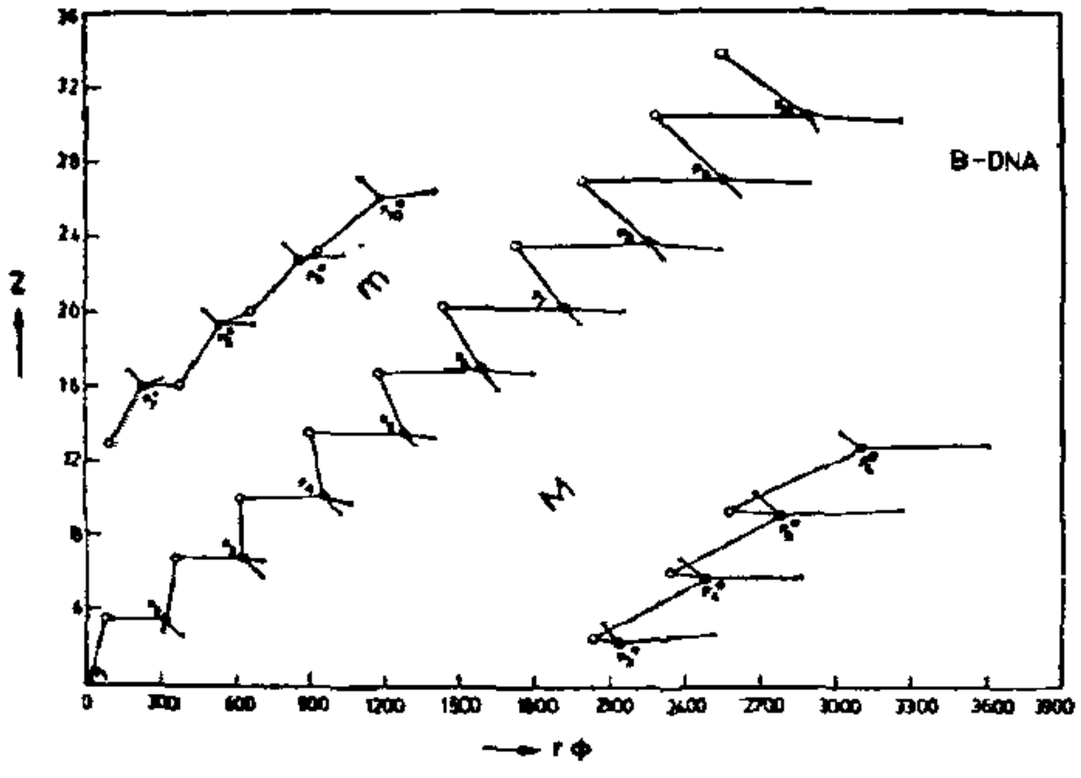

Figure 6: Radial projection plot for B DNA. The reversal in the magnitude of the width of minor(m) and major groove(M) between A and B forms is obvious.

width of minor groove is greater than that of major groove. The orientations of the anionic oxygens facing each other in the major groove is elegantly illustrated in figure 5.

A similar plot constructed for B form of DNA is shown in figure 6. The width of the major groove here is larger than that of minor groove. This is in contrast to A DNA (figure 5) where the width of major groove

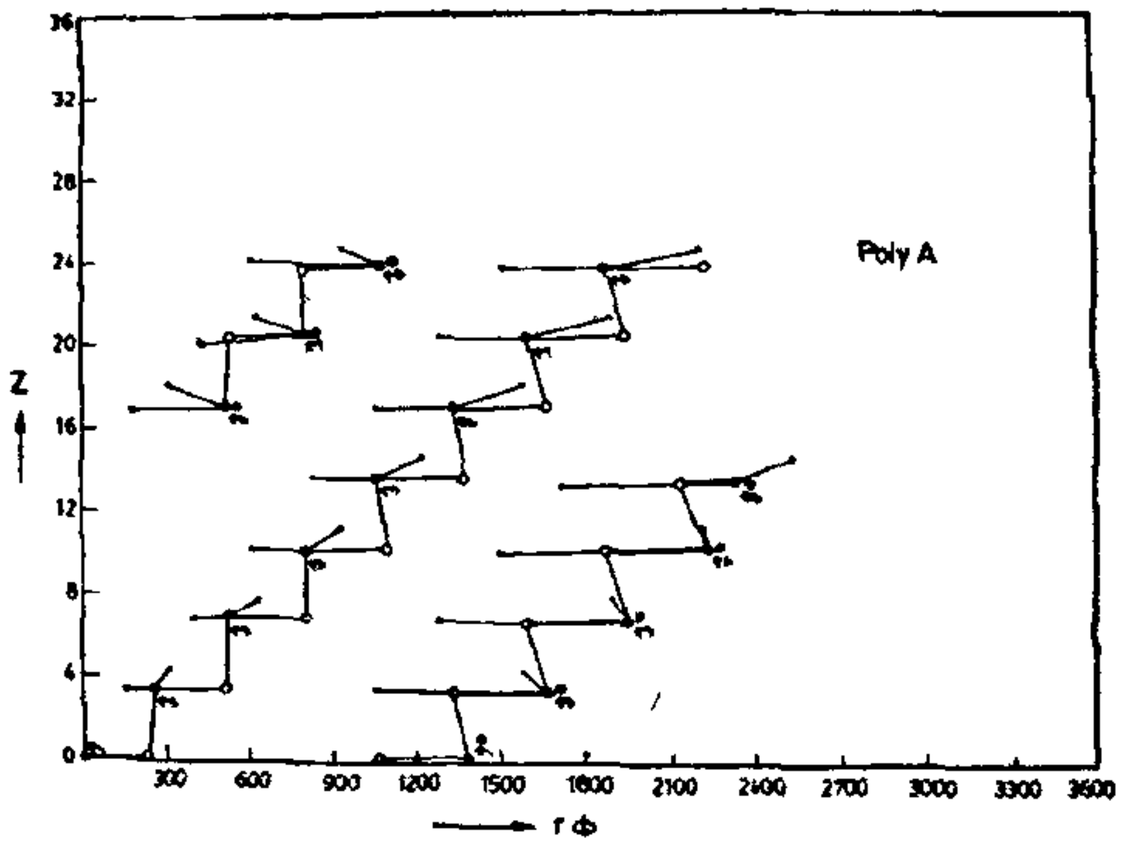

Figure 7: Radial projection plot for polyadenylic acid. The similar widths of the two forms is suggested by similar separations above and below the central line.

is smaller than minor groove. The reversal seen in the width between A and B forms is the main feature that distinguishes A and B type of helices and has been very clearly depicted in the plot. As in the major groove of A DNA (figure 5) the anionic oxygens face each other here also.

It is seen from both the distance plots as well as *z-r φ plots for A, B, D forms that grooves would* become conspicuous, only in the double helical segments having greater than four base-paired repeats.

Figure 7 shows the *z-r φ* plot constructed for polyadenylic acid. The grooves of equal width are revealed by identical separations of the atoms above and below the central line, as was also clear from the distance plot (figure 3). This is the most interesting feature which distinguishes parallel polyadenylic acid from other anti-parallel double helical structures. The *z-r φ* plot (figure 8) is also constructed for the Z_H form of poly(dinucleotide) helices. Since these are sequence specific, the sequences C and G are also indicated on the plot. The phosphorous atoms associated with the cytosine and guanine do not occur along the same line unlike in poly(mononucleotide) helices, where the successive phosphorous or C4′ atoms occur along a straight line. Due to the same reason, the separations across the strands vary for cytosine and guanines; reflecting the characteristic nature of 'dinucleotide' repeat[11] In constrast to all the other plots, the minor groove here (figure 8) is seen below the central line with the shortest P...P separation across the strands occurring between the phosphorous atoms associated with guanines. The P...P separations associated with the cytosines that form

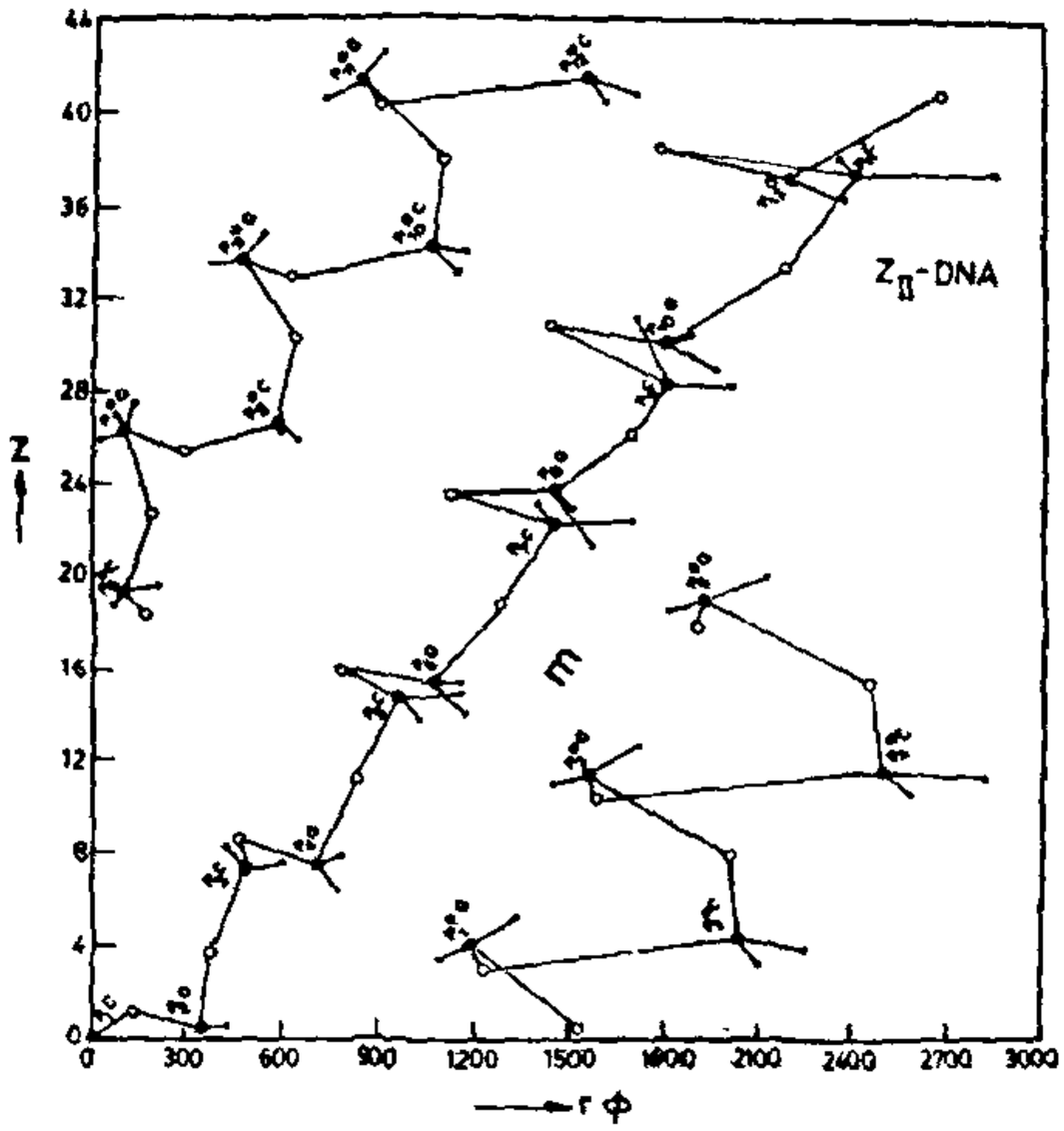

Figure 8: Radial projection obtained for Z_{II} DNA. The atom numbers and the sequences are indicated. The characteristic dinucleotide' repeat is obvious from the occurrence of phosphorous atoms associated with cytosine on one straight line and that of guanines on another. Note that the minor groove occurs below the central line.

the immediate neighbours of the above are larger indicating that the width of the groove is not uniform as is in the case of poly(mononucleotide) helices.

CONCLUSIONS

The differences in the secondary structural features of the various double helices of both poly(mononucleotide) as well as poly(dinucleotide) types arising due to variations in the nucleotide sugar residue (ψ,ψ') and internucleotide phosphodiester (ω',ω) conformations that determine the polynucleotide folding are elegantly brought out and illustrated. Most importantly the revelation of the nature of the grooves, their positions, their widths and number of residues involved in the formation of the grooves and the orientation of the anionic oxygen atoms in the grooves are conspicuous. It is clear from the above analysis that grooves in nucleic acid helices would be conspicuous only in double helical segments having greater than four base-paired residues. Also the plots bring out the distinct differences in the position of occurrence of minor groove between the left handed Z type and right handed A and B types of DNA and the characteristic 'dinucleotide' and 'mononucleotide' type of helical repeat. The distinct advantages of the current diagonal plots compared to the conventional plots that could be obtained using P...P separations are that (1) it is possible to describe all the known secondary helical and tertiary nucleic acid structures on common premises and (2) the interactions associated with the rest of the nucleotide backbone other than the phosphate group could be depicted through C4 atoms which occur almost at the centre of the repeating nucleotides.

ACKNOWLEDGEMENT

R. Malathi thanks C.S.I.R., New Delhi for a fellowship.

1. Malathi, R. and Yathindra, N., *Curr. Sci.*, 1980, **49**, 803.
2. Malathi, R. and Yathindra, N., *Int. J. Quan. Chem.*, 1981, **20**, 241.
3. Malathi, R. and Yathindra, N., *Curr. Sci.*, 1981, **50**, 1051.
4. Malathi, R. and Yathindra, N., *Int. J. Biol. Macromol.*, 1982, **4**, 18.
5. Malathi, R. and Yathindra, N., *Biopolymers.*, 1982, **21**, 2033.
6. Malathi, R. and Yathindra, N., *Biochem. J.*, 1982, **205**, 457.
7. Yathindra, N, Jayaraman, S and Malathi, R. In *Conformation in biology.*, (eds.) R. Srinivasan & R. H. Sarma Adenine Press, New York 1982.
8. Arnott, S and Hukkins, D. W. L. *Biophys. Biochem. Res. Commun.*, 1972, **47**, 1504.
9. Rich, R, Davies, D. R., Crick, F. H. C., Watson, J. D. *J. Mol. Biol.*, 1961, **3**, 71.
10. Wang, A. H. J., Quigley, G. J., Kolpak, P. J., Marel, G. V., van Boom, J. H. and Rich, A *Science.*, 1981, **211**, 171.
11. Yathindra, N. In *Biomolecular structure, conformation, function and evolution.*., (eds) R. Srinivasan, T. Subramanian and Yathindra, N. Pergamon Press, London, 1981.