STRUCTURE OF POLYNUCLEOTIDES AND NUCLEIC ACIDS

by R. K. NANDA, R. TEWARI *and* G. GOVIL, *Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay-400005*

(*Received 9 July 1973*)

Theoretical investigations using molecular orbital approaches (EHT, CNDO and PCILO) have shown that the backbone structure of polynucleotides is fairly rigid. The theoretically predicted model for the helical structures in nucleic acids is different from the Watson and Crick model but it can explain all the experimental facts on which Watson and Crick structure is based. The theoretical model is strongly supported by recent ^{1}H and ^{13}C NMR experiments on polyuridylic acid in aqueous solutions. The "randomness" of the poly-U strands arises because at room temperature, a small proportion of nucleotide units acquire conformational angles, different from those corresponding to minimum energy. NMR properties of such a "random coil" which are time and space averages have been calculated using a statistical approach. The mean square end to end distance for poly U has also been calculated and compared with measurements based on light scattering data.

INTRODUCTION

The conformational structure of nucleic acids plays a dominant role in many biological processes. However, details of their structure have remained obscured in the past. It is well known from X-ray diffraction that complementary strands of DNA form a double helix. However, such low resolution fiber data does not permit a direct determination of the atomic positions. Previous information on nucleic acid structure (*vide*, Sundaralingam 1969, 1972) is essentially based on single crystal X-ray diffraction analysis of nucleosides and nucleotides coupled with model building. Such an approach has obvious limitations. First, there is no *apriori* reason to believe that polynucleotides in solutions may acquire the configuration of low molecular weight compounds in solid state. Secondly, it does not tell the nature of forces which stabilise the polynucleotide structures. One would like to know, for example, the relative contributions of interbase hydrogen bonding and non-bonded interactions between the atoms in the nucleotide backbone to the stability of double helix. What is the conformational structure of single strand polynucleotides? Recent studies based on molecular orbital theory, statistical mechanics, nuclear magnetic resonance and light scattering data have provided some vital information on these aspects. In this paper, we shall review this information and discuss some of the questions posed above.

GENERAL CONSIDERATIONS

A molecule of nucleie acid has several degrees of freedom arising from the possibility of internal rotation about the various single bonds in the molecule

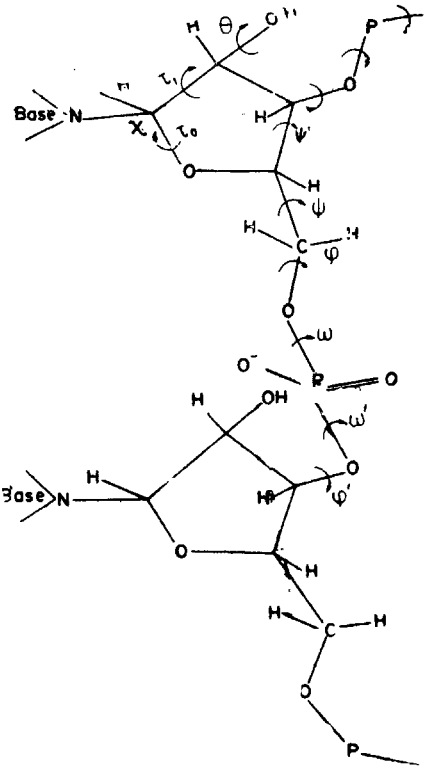

FIG. 1. The molecular structure of a polynucleotide with the various torsional angles which fix its secondary structure.

(Saran & Govil 1971; Govil 1972). In the backbone itself, there are five angles Φ', ω', ω, Φ and Ψ (Fig. 1) which determine the stereo-chemistry of the polynucleotide chains. Here, the angle Φ' defines rotation about the *C3′—O3′* bond, ω' about *O3′—P* bond, ω about *P—O5′* bond Φ about *O5′—C5′* bond and Ψ about *C5′—C4′* bond. The ribose ring can exist in several puckered conformations, due to the changes in τ_0 and τ_1. Two additional angles χ and θ determine the side chain configuration. The nucleic acid bases can exist in several tautomeric forms. Thus the conformational studies on nucleic acids are manifold and more complex than, e.g., polypeptides where one has to consider, in general, only two angles in the backbone and one in side chain. In this paper, we shall restrict only to the structure of the nucleotide backbone. Other aspects of nucleic acid structure have been discussed elsewhere (Govil & Saran 1970, 1971; and Berthod & Pullman 1971*a*, *b*).

MOLECULAR ORBITAL CALCULATIONS

Determination of conformational structure of the nucleotide backbone requires minimisation of energy with respect to the five torsional angles discussed above. In principle, one can attempt calculations on a long polynucleotide chain and

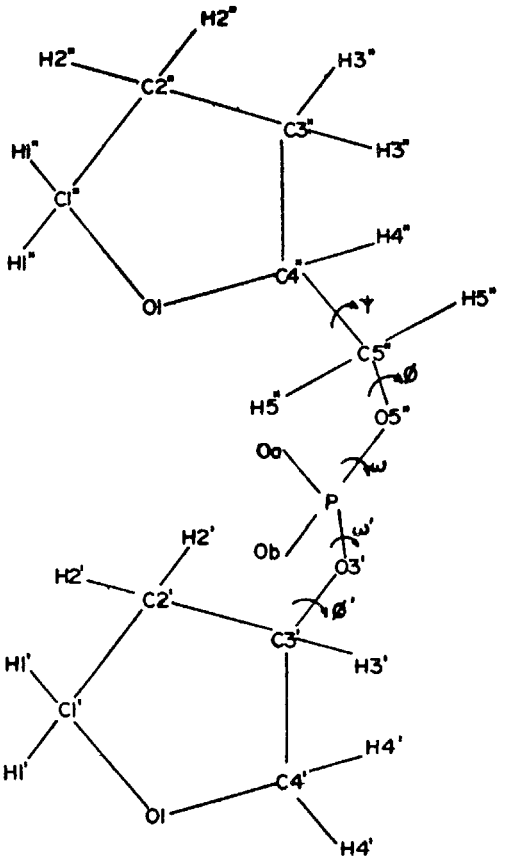

FIG. 2. The phosphodiester unit on which the present calculations have been carried out. The angles Φ', ω', ω, Φ and Ψ are considered positive for a right handed rotation.

minimise energy with respect to all the variables. This would however require access to a large computer and availability of long computer time. Fortunately, since only the nearest neighbour interactions are important in a conformational problem, one can get the basic information by working on a molecular fragment as shown in Fig. 2. This unit contains all the bonds which fix the backbone structure of polynucleotides. The torsion angles Φ', ω', ω, Φ and Ψ are measured in the sense indicated in Fig. 2 relative to the fully folded chain (Sundaralingam 1969, 1972). Standard values of various bond lengths and bond angles have been used (Saran & Govil 1971) in calculations described below and the sugar ring has been assigned a *C3'—endo* geometry.

Three different semi-empirical M.O. techniques are available for calculations involving all valence electrons for large molecules. These are, Extended Hückel Treatment (EHT), Complete Neglect of differential Overlap (CNDO) and Perturbative Configuration Interaction using Localised Orbitals (PCILO). All the three

have been applied to the problem under consideration. Saran and Govil (1971) using EHT method showed that the conformational freedom of nucleotide backbone is restricted. This was confirmed by Pullman *et al.* (1972) who used PCILO method and by Tewari *et al.* (1974) by CNDO method.

The results of these calculations are shown in Figs. 3-7. In making these calculations, the following values are assigned to the angles which are held constant: $\Phi' = 240°$, $\omega' = 290°$, $\omega = 290°$, $\Phi = 180°$ and $\Psi = 60°$.

A glance on the results shows that despite the fact that each chemical bond in the nucleotide backbone is formally a single bond, the rotation about them is highly restricted. In the Φ' space for example, the energy goes down sharply around 200° and rises again around 290°. Similarly, low energy regions in the Φ space are limited to 160°—260° region. Three minima occur when energy calculations are made with respect to Ψ, with the deepest one corresponding to $\Psi \sim 50°$. Minima at 180° and 300° have slightly higher energy. The actual value of the angle corresponding to minimum energy and the energy variations differ somewhat with the M.O. method used, but there is a general agreement among the three techniques. Another significant point to be noted here is the fact that the available single-crystal data on nucleosides, nucleotides and phosphodiesters (Sundaralingam 1969, 1972) show that in the majority of cases the preferred conformations are close to $\Phi' \sim 240°$, $\Phi \sim 180°$ and $\Psi \sim 60°$.

The (ω', ω) map shows some very interesting features and also brings up the importance of the more refined approaches like PCILO and CNDO as compared to EHT. The EHT map shows seven regions of stability which occur around (60°, 80°), (60°, 180°), (180°, 60°) (180°, 180°), (180°, 300°), (300°, 180°) and (290°, 290°). All these regions are interconnected by contours of low energy indicating that the barrier to interconversion from one structure to another is low. Incidentally, the majority of experimental points obtained from single crystal studies of nucleotides and phosphodiesters lie in the (290°, 290°) region, along with some points in (60°, 60°), (180°, 300°) and (300°, 180°) regions. This observation of a preference for a *guache-guache* arrangement about the *O–P* bond is indeed brought out very clearly when calculations are performed by CNDO and PCILO. The CNDO map for example (Fig. 7) shows global minima in the region (290°, 290°) and (60°, 80°). However, all the three methods show a comparatively higher flexibility with respect to rotation about the *O–P* bond, with almost 50 per cent of the area in the (ω', ω) map having energy within 3 Kcal mole^{-1} relative to the global minima.

Nucleotide Rigidity and Models of Nucleic Acids

The calculations discussed in the previous section show that the conformational freedom of the nucleotide backbone is highly restricted. The energy calculations and single-crystal X-ray diffraction studies on nucleosides and nucleotides indicate that the preferred values of the conformational angles lie around $\Phi' \sim 240°$, $\omega' \sim 290°$, $\omega \sim 290°$, $\Phi \sim 180°$, $\Psi \sim 60°$. *A* polynucleotide structure with these angles lead to right handed helices. It may be pointed out that the proposed models of DNA double helix are based on two important observations from fiber-diagrams, namely, repeat distance (pitch) of 34Å and approximately 10 nucleotide

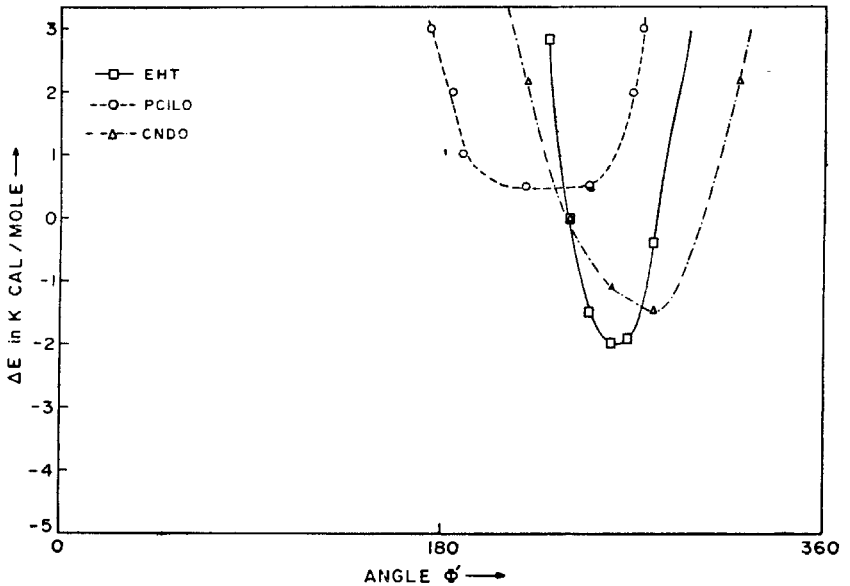

FIG. 3. The potential energy of nucleotide backbone as a function of angle Φ'. Values of other angles have been fixed as follows : $\omega' = 290°$, $\omega = 290°$, $\Phi = 180°$, $\Psi = 60°$ (The global minima of the curves are not comparable).

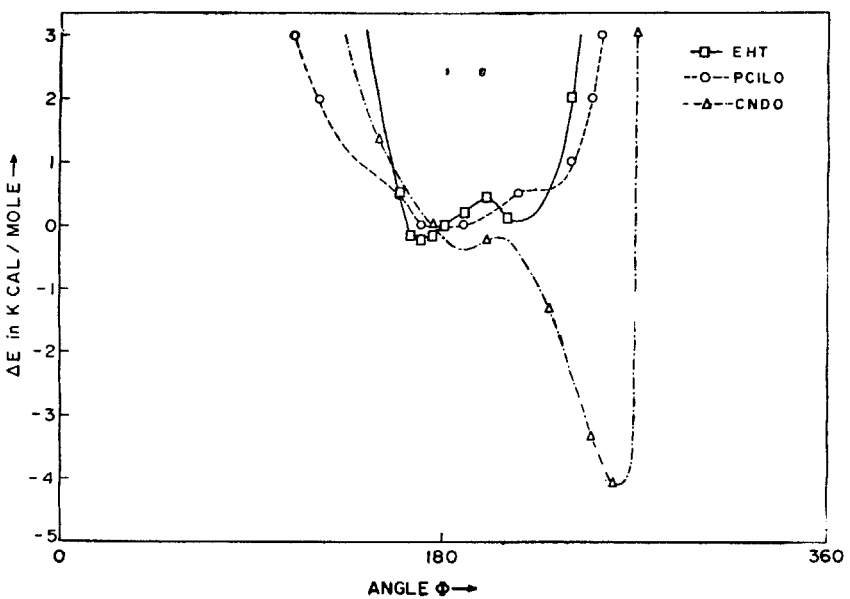

FIG. 4. The potential energy of the nucleotide backbone as a function of angle Φ. Values of other angles have been fixed as follows : $\Phi' = 240°$, $\omega' = 290°$, $\omega = 290°$ $\Psi = 60°$ (The global minima of the curves are not comparable).

units per turn of the helix. Both these criterion can be satisfied with only minor variations in the torsional angles or the primary geometry of the nucleotide back-

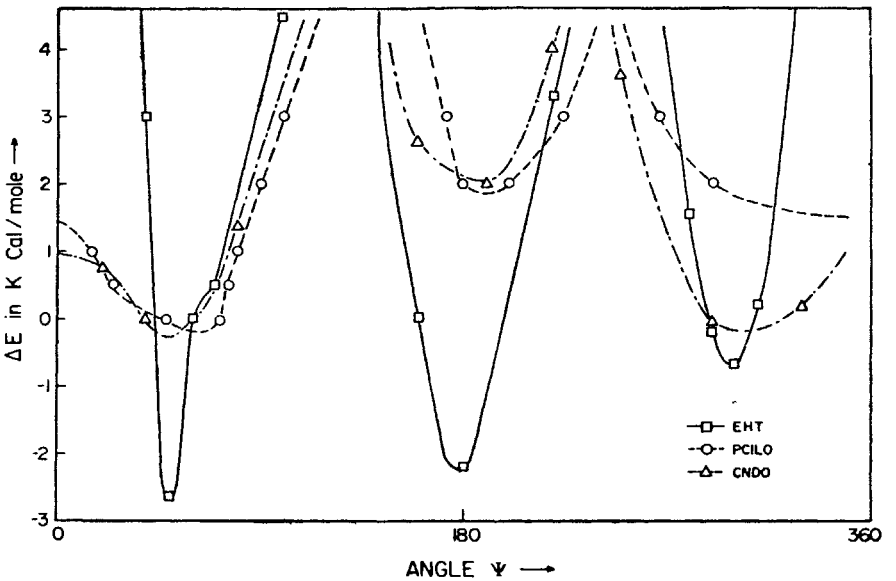

FIG. 5. The potential energy of the nucleotide backbone as a function of angle Ψ. Values of other angles have been fixed as follows : $\Phi' = 240°$, $\omega' = 290°$, $\omega = 290°$, $\Phi = 180°$ (The global minima of the curves are not comparable).

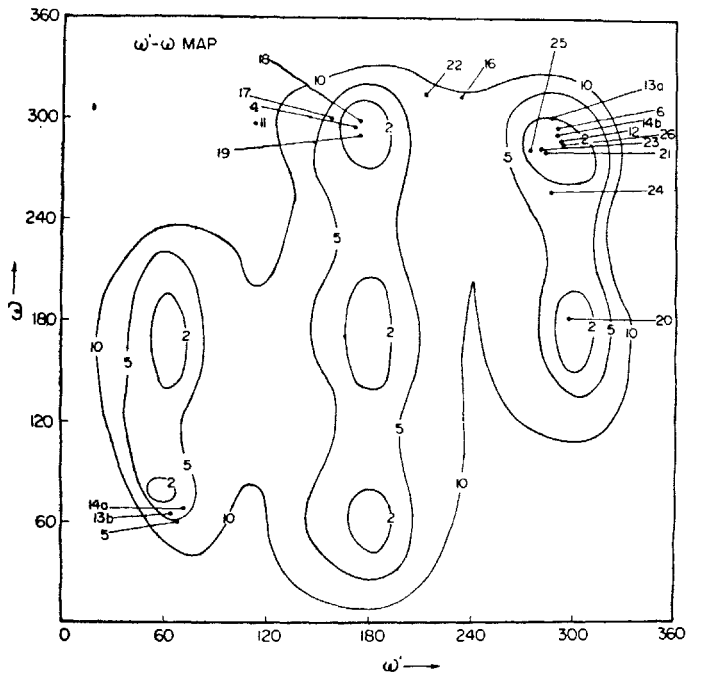

FIG. 6. Isoenergy (conformational) map in (ω', ω) space as calculated by EHT. The points represent experimental observations based on single-crystal X-ray analysis of nucleosides and nucleotides. Values of other angles are $\Phi' = 240°$, $\Phi = 180°$ and $\Psi = 60°$. The map shows seven minima of almost equal energy.

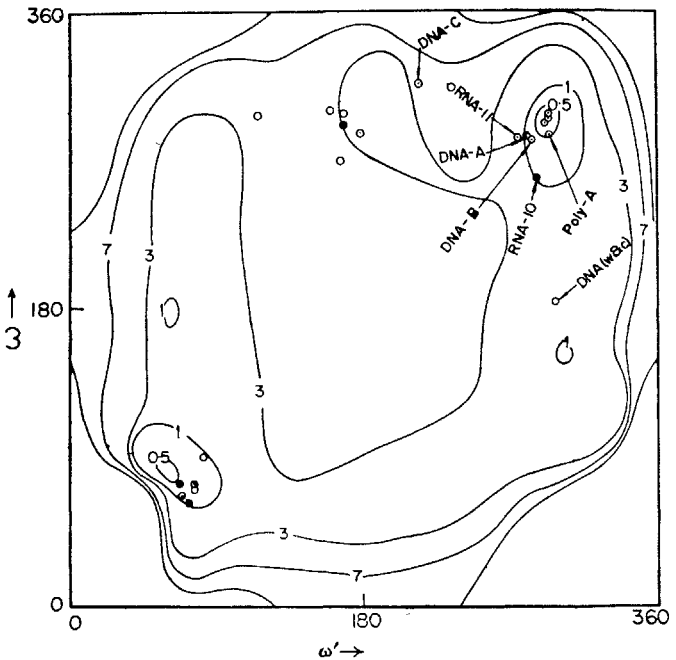

FIG. 7. (ω', ω) map as calculated by CNDO. Note that the *gauche-gauche* arrangement about the O–P bonds [(290°, 290°) and (60°, 80°) points] turn out to be more stable than the other low energy regions predicted by EHT. The PCILO map is very similar to the CNDO map.

bone. Fig. 8 shows such a single strand helix generated through an interactive computer graphic facility at our centre.

Table I gives torsion angles for some of the models proposed for double helix along with the calculated energy of the phosphodiester unit with the same torsional angles. The conformational angles of at least four of the seven proposed models (DNA-A, RNA-10, RNA-11 and Poly A) lie close to the values predicted theoretically. Poly A with angles closest to above values is predicted to be most stable structure. In the Watson and Crick model values of Φ', ω and Ψ are significantly different from the theoretically predicted value. The values of Φ' in DNA-B and of Φ in DNA-C fall in high energy regions.

It is clear from these considerations that the nonbonded interactions play a dominant role in determining the conformational structure of the individual strands of the nucleic acids. The structure of single strand polynucleotides and molecules such as *t*-RNA is likely to be dominated by helical segments with angles listed above. The specific hydrogen bond pairing between the bases play the role of locking two such strands together and removing any "randomness" that may arise in these strands from the finite population of other low energy regions in the (ω', ω) space, The unwinding of double helix (for example, during the process of replication and transcription) is also likely to take place through a rotation about the *O–P* bonds (Govil 1972).

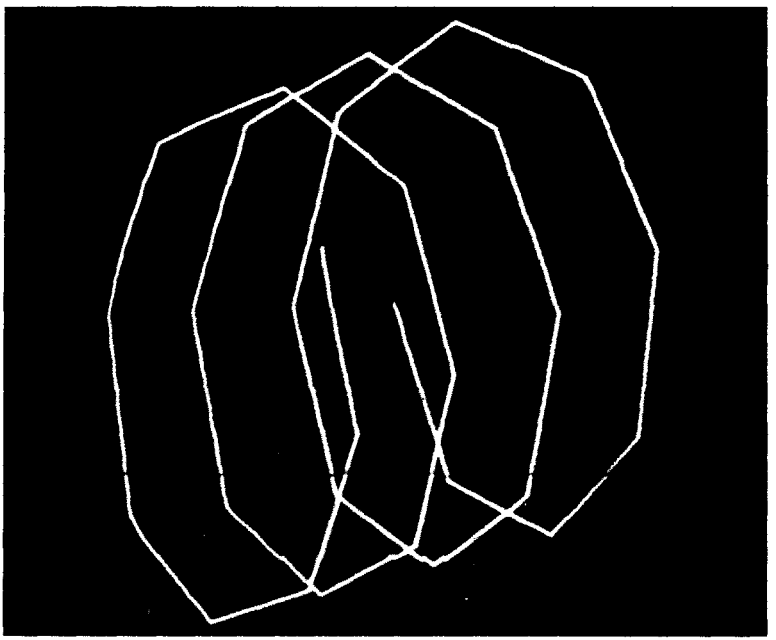

FIG. 8. A polynucleotide helix generated on the computer TV console using the torsion angles $\Phi' = 240^\circ$, $\omega' = 290^\circ$, $\omega = 290^\circ$, $\Phi = 180^\circ$ and $\Psi' = 60^\circ$. Only the phosphorus positions are shown which are connected through straight lines. There are 10 nucleotide units in one turn and the pitch of the helix is 34 Å.

TABLE I

Comparison of the stability of the backbone structure of proposed models of nucleic acids and polynucleotides as computed by EHT, CNDO and PCILO methods. The energies have been expressed relative to the energy for Poly—A structure

Models	Φ'	ω'	ω	Φ	Ψ'	Energy in PCILO	K. Cal/mole	
							EHT*	CNDO
DNA (W and C)	186	296	182	179	155	3.6	24.2	4.9
DNA-B	147	282	281	212	58	3.8	156.3	24.7
DNA-C	211	212	315	143	48	1.7	—2.3	3.5
DNA-A	221	279	283	167	87	0.3	—3.3	0.7
RNA-10	203	285	257	188	88	2.8	1.8	4.6
RNA-11	216	273	282	165	74	1.2	1.5	2.3
Poly-A	216	292	285	168	69	0.0	0.0	0.0

*In the original calculations of Saran and Govil (1971), one of the hydrogen atoms has been wrongly placed. The values reported here are after the necessary correction.

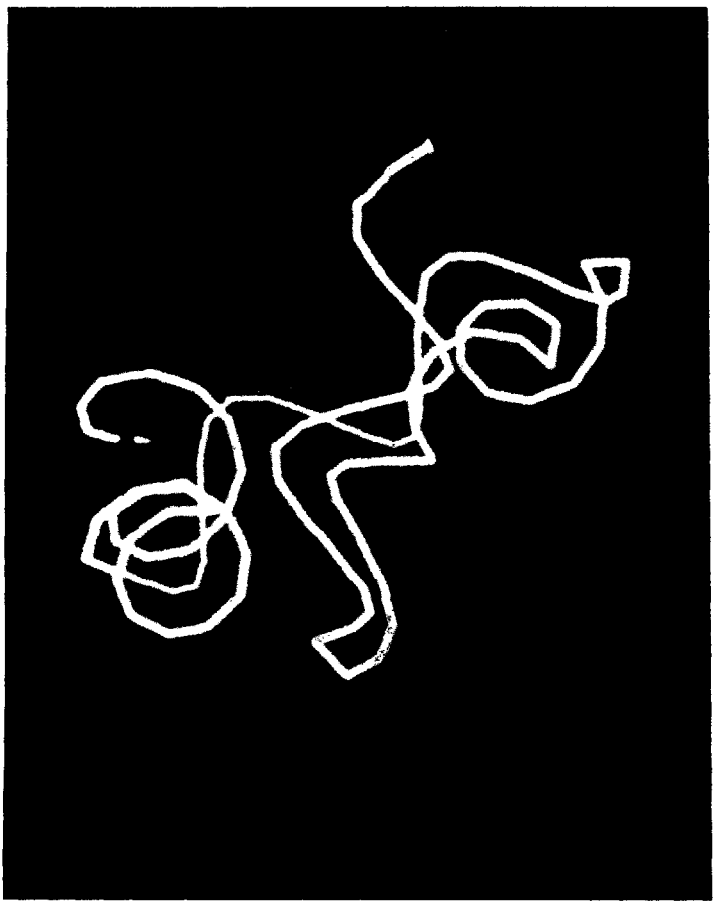

FIG. 9. A computer simulation of the 'random' coil of poly U. Note that the coil is characterised by ordered helical segments followed by breaks.

NMR COUPLING CONSTANTS IN POLY U AND STATISTICAL TREATMENT OF NUCLEOTIDE BACKBONE

A direct verification of the stereochemistry of the polynucleotide backbone in aqueous solutions, is possible through the *vicinal* (three bond) NMR coupling constants (J) which depend on the dihedral angle (θ) through Karplus type relations

$$J = A\ Cos^2\ \theta + B\ Cos\ \theta$$

Unfortunately, the ordered structures such as double helix, give very broad NMR lines due to dipole-dipole interactions. However for single strands polynucleotides which occur as "random" coils, it is possible to measure such coupling constants and enquire if the backbone shows any conformational preference. Detailed ^1H and ^{13}C magnetic resonance study of polyuridylicacid (poly U) have been made recently (Kreishman & Chan 1971; and Govil & Smith 1973) and the results provide clue to the conformational structure of poly U with respect to Φ', Φ, and Ψ.

In the analysis which follows, the following Karplus-type relations have been used

$$^{J_1}\mathrm{HCC^1H} = 9.8\ Cos^2\theta - 0.5\ Cos\ \theta$$

$$^{J_{13}}\mathrm{CCO^{31}P} = 9.5\ Cos^2\theta - 0.6\ Cos\ \theta$$

$$^{J_1}\mathrm{HCO^{31}P} = 16.3\ Cos^2\theta - 4.6\ Cos\ \theta$$

The observed coupling-constants $<J>$ in "random" coil are time and space averages. These are related to the 'rigid' angle coupling constants through the relation

$$<J> = \int_{\theta=0}^{360^\circ} J(\theta)Z(\theta)d\theta$$

where $Z(\theta)$ is the "statistical" weight of the states having the conformation with dihedral angle θ

$$Z_\theta = \frac{\exp\ (-E_\theta/RT)}{\int \exp\ (-E_\theta/RT)d\theta}$$

If the potential energy (E) with respect to θ is known and for convenience of calculation the θ space is divided into equally spaced intervals then it is possible to write

$$<J> = \frac{\Sigma\, J_\theta \exp\ (-E_\theta/RT)}{\Sigma \exp\ (-E_\theta/RT)}$$

The potential energy curves as predicted by EHT, CNDO and PCILO methods have been used in conjunction with Karplus type relations to predict values of average three bond ^{1}H–^{1}H, ^{1}H–^{13}C and ^{13}C–^{31}P coupling constants and have been compared with the experimental values in poly U (Table II). We have also listed θ values in a 'rigid' model which best explain the observed coupling constants. These results clearly show that the nucleotide backbone in poly U is fairly 'rigid' with respect to angles Φ', Φ and Ψ and their mean values are close to those which correspond to minimum energy in the EHT, CNDO and PCILO calculations.

'Random' Coil Structure of Poly U

According to the theoretical calculations, the nucleotide backbone is comparatively flexible with respect to rotation about the O–P bonds. None of the observed NMR couplings are sensitive to the angles ω' and ω. However, the mean square end to end distance ($<h^2>$) in poly U is known from light scattering measurements (Inners and Felsenfeld 1970) and this can be used to test if the nucleotide backbone has any conformational preference in the (ω', ω) space. The observed value ($in\mathring{A}^2$) of $<h^2>$ can be represented by the formula

$$<h^2> = 240\ N$$

where N is the number of nucleotide units in the chain. Having shown that the nucleotide backbone is rigid in the Φ', Φ and Ψ spaces, it is possible to use light

TABLE II

Predicted and Experimental Values of NMR Coupling Constants (<J>) in poly U

Coupling constant	Relevant Torsion Angle	EHT	CNDO	PCILO	'Rigid' Model $\Phi' = 240°$ $\Phi = 180°$ $\Psi = 60°$	Observed values	Remarks
J ($^{13}C_2'$ — ^{31}P)	Φ'	7.1	8.4	2.4	3.5	5.0	
J ($^{13}C_4'$ — ^{31}P)	Φ'	0.5	0.6	4.9	2.7	3.0	
J ($^{1}H_3'$ — ^{31}P)	Φ'	9.5	7.1	9.5	11.7	8.2	
J ($^{13}C_4'$ — ^{31}P)	Φ	8.5	0.8	8.3	1.01	7.0	The predicted
J ($^{1}H_5'$ — ^{31}P)	Φ	2.1	12.0	2.8	1.8	4.0	value by CNDO
J ($^{1}H_5''$ — ^{31}P)	Φ	4.6	10.3	4.3	1.8	4.0	is 'off' byabout 60°
$\lvert J(H_4'—H_5')+J(H_4'—H_5'')\rvert$	Ψ	8.3	9.7	6.1	4.4	7.5	
$\lvert J(H_4'—H_5')—J(H_4'—H_5'')\rvert$	Ψ	4.5	1.6	1.7	0	±1.0	

scattering data coupled with energy calculation to investigate the average situation in the (ω', ω) space. One can discuss first two limiting situations. If the nucleotide backbone is completely rigid with respect to all the five torsional angles, then the end to end distance will be proportional to N. On the other hand, if the coil is completely random (isoenergy map in the ω', ω space is flat), then $<h^2> \sim 27N$. The observed values of $<h^2>$ therefore indicate a high degree of ordering in the (ω', ω) space also.

The molecular orbital isoenergy maps in the (ω', ω) space have been employed to predict the mean square end to end distance. (Tewari *et al.* 1974). For this purpose, the isoenergy map is divided into 9 equally spaced points and the statistical weight of each point Z_i is calculated. The transformation matrices corresponding to each point in (ω', ω) space are also calculated. From this data, it is possible to calculate $<h^2>$. Random coils were simulated on the computer using the probability distribution Z_i as given by energy calculations based on M.O. theory, The values of $<h^2>$ have been estimated and the results are shown in Table III. It is seen that all the M.O. maps predict mean square end to end distance which are shorter than what is actually observed. This means that the nucleotide backbone in poly U is more rigid than what is predicted by theory. This behaviour can arise because of two factors. Either the predicted energy for the (290°, 290°) point by M.O. methods is a little higher than the actual value. Alternatively long range interactions between negatively charged phosphate groups lead to a further stabilisation of the (290°, 290°) region. Therefore, the calculations have been repeated by decreasing the energy of the point (290°, 290°) by different amounts. A value of 2 Kcal mole^{-1} reproduces the correct end to end distance. Such a "random" coil (Fig. 9) is marked by fairly long helical sections followed by "random" breaks. In other words, a short range order is present in poly U, even at higher temperatures.

TABLE III

The mean square end to end distance for chains of 500 nucleotides each

Methods	Mean square end to end distance in $(\text{Å})^2$	Statistical weight for the structure having ($\omega' = 290°$; $\omega = 290°$)
Completely Random Coil (Flat ω', ω)	13900	0.111
EHT	16700	0.143
CNDO	11800	0.310
PCILO	19500	0.276
PCILO Map but with (290°, 290°) point assigned an energy of —2 Kcal mole^{-1}	84200	0.806
Observed value for Poly U (Interpolated for chains of 500 nucleotides)	122000	—

CONCLUSION

It is clear that even in absence of interbase hydrogen bonding the nucleotide backbone has a strong preference for the following torsional angles : $\phi' \sim 240°$, $\omega' \sim 290°$, $\omega \sim 290°$, $\phi \sim 180°$ and $\Psi \sim 60°$. The "randomness" of single stranded nucleic acids arises because a small percentage of nucleotide units acquire a conformation different from the one discussed above. In a double helix, such "random" turns are prevented by interbase hydrogen bonds. The unwinding of the double helix, most likely takes place through rotations about the *O–P* bonds. These concepts are likely to prove very important in solving the structures of t-RNA molecules which are hitherto unknown.

ACKNOWLEDGEMENT

This paper forms a part of the research program in Molecular Biophysics which was started at the initiative of late Professor C. R. Kanekar. We consider ourselves lucky to have been associated with him and to have received his guidance and encouragement in our career.

REFERENCES

Berthod, H., and Pullman, B. (1971 *a*). Molecular Orbital Calculations on the Conformation of Nucleic Acids and their Constitutuents. Part II. *Biochim. Biophys. Acta.*, **246**, 359.

Berthod, H., and Pullman, B. (1971 *b*). Molecular Orbital Calculations on the Conformation of Nucleic Acids and their Constituents. Part I. *Biochim. Biophys. Acta.*, **232**, 595.

Flory, P. J. (1969). Statistical Mechanics of Chain Molecules. Interscience Publications, John Wiley and Sons, New York.

Govil, G. (1972). Nucleic Acid Conformations and Biological activity in Conformation of Biological Molecules and Polymers. Proceedings of the 5th Jerusalem Symposium *ed.* by Bergmann, E. D., Pullman, B., Academic Press, New York.

Govil, G., and Saran, A. (1970). Quantum Chemical Studies on the Conformational Structure of Nucleic Acids Part I. *J. Theor. Biol.*, **30**, 621.

Govil, G., and Saran, A. (1971). Quantum Chemical Studies on the Conformational Structure of Nucleic Acids Part II. *J. Theor. Biol.*, **33**, 399.

Govil, G., and Smith, I. C. P. (1973). A Carbon-13 Magnetic Resonance Study of the helix-coil transition in Poly-uridylic Acid. *Biopolymers*, **12**, 2589.

Inners, D. L., and Felsenfeld (1970). Conformation of Polyribouridylic Acid on Solution. *J. Mol. Biol.*, **50**, 373.

Kreishman, G. P., and Chan, S. I. (1971). A Proton Magnetic Resonance Study of the Interaction of Adenylyl (3′—5′) Adenosine with Polyuridylic Acid. Biopolymers, **10**, 159.

Tewari, R., Nanda R. K., and Govil, G. (1974). Quantum Chemical Studies on the Conformational Structure of Nucleic Acids Part IV. *J. Theor. Biol.*, **46**, 229.

Puliman, B., Perahia, D., and Saran, A. (1972). Molecular Orbital Calculations on the Conformation of Nucleic Acids and their Constituents Part III. *Biochim. Biophys. Acta*, **269**, 1.

Saran, A., and Govil, G. (1971). Quantum Chemical Studies on the Conformational Structure of Nucleic Acids Part III. *J. Theor. Biol.*, **33**, 407.

Sundaralingam, M. (1969) Stereochemistry of Nucleic Acids and their Constituents Part IV. *Biopolymers*, 7, 821.

Sundaralingam, M. (1972). The Concept of a Conformationally rigid nucleotide and its significance in Polynucleotide Conformational Analysis in "Conformation of Biological Molecules and Polymers. Proceedings of the 5th Jerusalem Symposium *ed.* by Bergmann, E. D., Pullman, B., p. 417, Academic Press, New York.

Tewari, R., Nanda, R. K., and Govil, G. 1974. *Biopolymers*, **13**, 2015.