CONFORMATIONAL CHARACTERISTICS OF THE DIMERIC SUBUNITS OF DNA FROM ENERGY MINIMIZATION STUDIES

MIXED SUGAR-PUCKERED dApdA, dApdT, dTpdA, AND dTpdT

P. K. PONNUSWAMY AND P. THIYAGARAJAN *Department of Physics, Autonomous Postgraduate Center, University of Madras, Tiruchirapalli 620 020, Tamilnadu, India*

ABSTRACT An extensive investigation on the conformational characteristics of four deoxydinucleoside monophosphates, namely, dApdA, dApdT, dTpdA, and dTpdT was carried through calculation of the classical potential energy of the systems. The four major types of sugar-pucker sequences, namely, ^3E-^3E, ^3E-^2E, ^2E-^3E, ^2E-^2E, were included in the study. For each of the units, energies were computed for 96 starting conformations that resulted from the consideration of all possible low energy regions for the relevant seven dihedral angles and the four sugar-pucker sequences, and minimized by permitting all the seven dihedral angles to vary simultaneously. The number and the order of preference of low energy conformations obtained were found to be characteristic of the base sequence of the unit considered. The conformational states close to the A-DNA, B-DNA, C-DNA, and Watson-Crick DNA structures are noted to be preferred for all the units except dTpdT. The ^3E-^2E sugar-pucker sequence is the most favored and the ^2E-^3E sequence is the least favored state in terms of the associated number of local minima. For each unit, there exists a set of specific conformational states with more or less equal stabilities but different sugar-pucker sequences. The mixed sugar-pucker states ^2E-^3E and ^3E-^2E, when incorporated, in the conventional A-DNA and B-DNA conformational states, respectively, have energies that allow them to act as intermediates in the B form $\rightleftharpoons$ A form transitions. Such transitions are most likely to occur at sites with a Thymine-Adenine base sequence. Available experimental results were interpreted in terms of their stabilities.

INTRODUCTION

The intrinsic structural and conformational properties of subunit systems of biopolymers provide a wealth of information for the understanding of their secondary structural features. Theoretical studies on proteins and nucleic acids, provided information on their most and least probable conformations. Studies on various model systems were also utilized for interpreting macromolecular conformations. However, in the case of nucleic acids, a thorough knowledge of all the dimeric subunits has not yet been derived theoretically owing to the complexity introduced by the large number of structural/conformational variables. Earlier studies were made on the monomeric and dimeric subunits by using classical (Lakshminarayanan and Sasisekharan, 1970; Olson and Flory, 1972; Yathindra and Sundaralingam, 1974; 1975) and quantum mechanical calculations (Tewari et al., 1974; Perahia et al., 1977). These investigations, have provided valuable information about the nucleic acid structure. As a result, we

presently have a fairly good understanding of the preferred puckering conformations for the sugar systems and the low energy domains for the other conformational parameters. Based on these results, some interesting new models were recently proposed for DNA (Rodley et al., 1976; Sasisekharan and Pattabiraman, 1976; Sasisekharan et al., 1978) and attempts are in progress to elucidate the chromatin structure. Nevertheless, because of the inherent deficiency in the application of the classical and quantum mechanical techniques, we are still far away from having a complete picture of conformational characteristics of all the nucleic acid subunits even at the dimeric level: the studies cited above were made by varying only one or, at the most, two dihedral angles, fixing the rest at some preferred values while searching for the probable local energy minima for the subunit systems. Consequently, they failed to reveal all the available local minima with their stabilities for the various subunit systems in terms of their base sequence. For deriving such information it is essential to locate the real local minima in the multidimensional energy space by permitting all the variables to adjust simultaneously. A set of such low energy conformations for the dimeric subunits with all possible base and sugar-pucker sequences, would be valuable in building models for polynucleotides as well as to interpret the low resolution electron density maps of the large nucleic acid molecules, such as tRNAs, in their crystalline forms. A search for all the local minima in the multidimensional energy space is possible only through a suitable function minimization study.

Recently studies using multidimensional function minimization procedure were made on the nucleic acid systems (Broyde et al., 1975; 1978; Broyde and Hingerty, 1979; Fujii and Tomita, 1976; Kister and Dashevsky, 1976; Govil et al., 1977; Thiyagarajan and Ponnuswamy, 1978*a*, 1978*b*, 1979; Zhurkin et al., 1978; Levitt, 1978). These studies brought out many interesting aspects of the influence of the bases in the dimer, their actual sequence on the preferred backbone conformations and base stacking properties. In these studies, however, one important aspect was not considered: namely, the mixed sugar-pucker states. Here we report on an extensive investigation on four dimeric subunits of DNA, namely, dApdA, dApdT, dTpdA and dTpdT, representing the four general types of base sequences found in nucleic acid systems, viz, purine-purine, purine-pyrimidine, pyrimidine-purine and pyrimidine-pyrimidine. A similar study on the RNA subunits, ApG, ApU, CpG and CpU is reported in the next article. For studying all these units we used an improved set of atom-atom interaction parameters to calculate the potential energy and an efficient minimization technique to minimize the total potential energy as a function of all the flexible dihedral angles in the units. The ring conformation of the furanose is kept rigid in both of the important puckered conformations ^3E or ^2E.

METHODS

For the molecular systems studied, the total potential energy (E_{tot}) was considered to be the sum of the contributions from nonbonded (E_{nb}), hydrogen bonding (E_{hb}), electrostatic (E_{es}) and torsional (E_{tor}) interactions which were computed by using the expressions

$$E_{nb} = \sum_{i,j} \epsilon_{ij} \left[\left(\frac{\langle r_g \rangle_{ij}}{r_{ij}} \right)^{12} - 2.0 \left(\frac{\langle r_g \rangle_{ij}}{r_{ij}} \right)^{6} \right] \quad (1)$$

$$E_{hb} = \sum \left[\frac{A}{r^{12}_{x...H}} - \frac{B}{r^{10}_{x...H}} \right] \quad (2)$$

where x is N or O,

$$E_{es} = 166.0 \sum_{i,j} \frac{q_i q_j}{r_{ij}} \tag{3}$$

$$E_{tor} = \frac{V_\theta}{2}(1 + \cos 3\theta). \tag{4}$$

The sums in Eqs. 1, 2, and 3 were taken over all pairwise atomic interactions whose distances r_{ij} (or $r_{x\ldots H}$) vary with dihedral angles. In Eq. 1, the constants ϵ_{ij} and $\langle r_g \rangle_{ij}$ for atomic pairs involving other than phosphorus were taken from the work of Scheraga's group (McGuire et al, 1972; Momany et al, 1975). For atomic pairs involving the phosphorus atom we computed ϵ_{ij} and $\langle r_g \rangle_{ij}$ as described in our earlier article (Thiyagarajan and Ponnuswamy, 1978*a*). The sum in Eq. 1 is taken over all atomic pairs except $x \ldots$ H (where x is oxygen or nitrogen and H is hydrogen covalently attached with nitrogen or oxygen) and the sum in Eq. 2 is taken only over the $x \ldots$ H interactions: The A and B values in this Eq. 2 are hydrogen bond interaction parameters, which were taken from McGuire et al, 1972. These interaction parameters were derived from an extensive study on crystal packing of various categories of molecular entities and were tested for their appropriateness in conformational energy calculations. For computing the electrostatic interactions a dielectric constant of 2 was employed. The atomic partial charges q_i are those determined by Renugopalakrishnan et al. (1971). The details of all these parameters as well as the computation of the total potential energy via the above expressions were given in our earlier articles (Thiyagarajan and Ponnuswamy, 1978*a* and *b*). The VMMO1 version of the Fletcher-Powell-Davidon minimization procedure (Davidon, 1959; Fletcher and Powell, 1963) was used to minimize the total potential energy as a function of the dihedral angle variables.

To compare the stabilities of conformations belonging to various sugar-pucker sequence domains, it is necessary to incorporate the internal energies of the sugar ring systems. Hence the internal energy of the sugar ring was computed as the sum of the contributions from nonbonded, electrostatic, hydrogen bonding, and torsional interactions as well as that from the strain experienced through the variations in the bond lengths and bond angles, all within the ring system. The strain of deformation energy was computed by using the expression $E_d = (K_\ell/2)(\Delta\ell)^2 + (K_\theta/2)(\Delta\theta)^2$, where K_ℓ and K_θ are the stretching and bending force constants, and $\Delta\ell$ and $\Delta\theta$ are the deviations in the bond lengths and bond angles from the equilibrium values. The values of K_θ and K_ℓ, as well as the equilibrium bond lengths and bond angles were taken from Ramachandran and Sasisekharan (1968). The internal energies thus calculated for the deoxynucleotide with ^3E and ^2E puckerings are 1.87 and 1.61 kcal/mol, respectively, indicating their more or less equal stabilities. For the ribonucleotide the internal energies assume values of 2.52 and 2.61 kcal/mol for the $^3E-$ and ^2E states, respectively. These theoretical results corroborate very well with the solution studies on mononucleotides of DNA and RNA which predict that the ^2E population is slightly more than that of ^3E in the case of deoxynucleotides whereas the ^2E and ^3E states are equally populated in the case of ribonucleotides. The internal energy terms were added appropriately to E_{tot} of every minimized conformation.

Although the parameters developed by Scheraga's group were intended for polypeptides, they could also be used for other systems as well. In fact, the conformational properties of both the families of biomolecules, proteins and nucleic acids were studied so far by taking the same interaction parameters for the commonly occurring atoms. For the other atoms the interaction parameters were determined appropriately by taking the values of van der Waal's radii, refractivities and the effective number of electrons from Bondi (1964), Ketelaar (1959) and Scott and Scheraga (1965). Such an approach has been adopted with the reasonable assumption that the atoms, regardless of the system in which they occur, will behave in a similar fashion. With the same reasoning we have employed the new parameters for our study on nucleic acid systems.

The major problem in a multidimensional function minimization study is the selection of starting points in the whole space. Enough points should be considered to ensure that all the probable local minima and, consequently, the global minimum are included. We decided the starting points as follows: Fig. 1 depicts the relevant seven dihedral angle variables for the representative molecule, dApdT.

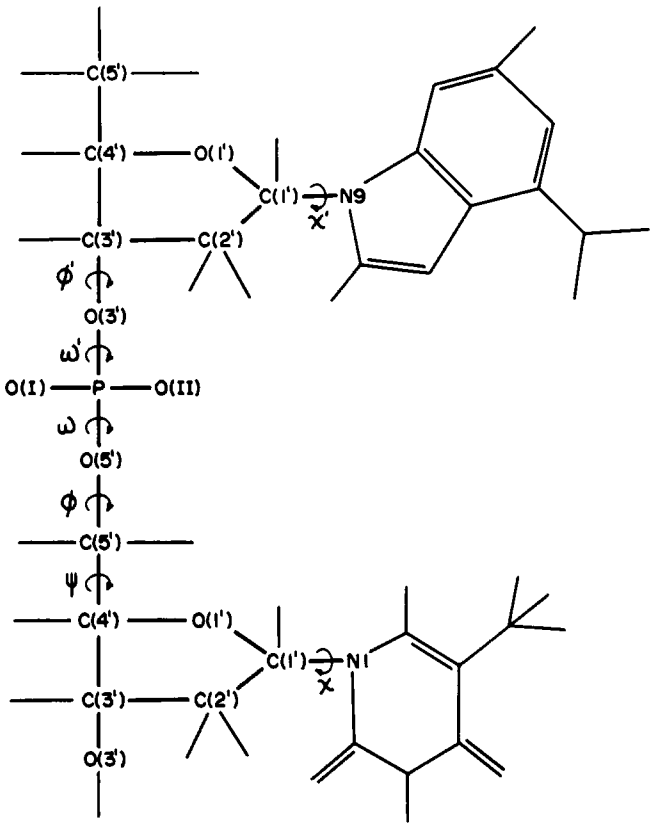

FIGURE 1 The skeleton, numbering convention, and conformational angles for the deoxydinucleoside monophosphate dApdT.

Studies made by Sundaralingam and by Sasisekharan on mono-, and dinucleotides, and also by us on mononucleotides demonstrate that the local minima, and the most probable low energy regions around these minima for the variables marked in Fig. 1 are χ', $\chi \simeq 40° \pm 20°$, $\phi' \simeq -130° \pm 20°$, $\phi \simeq 180° \pm 20°$, $\psi \simeq 60° \pm 20°$, $180° \pm 20°$ and $-60° \pm 20°$.

The dihedral angles ω' and ω (about phosphodiester bonds) have three local minima corresponding to the staggered orientations, i.e., around 60°, 180°, and $-60°$. As a working approximation, we selected the midpoint in each range in each of the dihedral angles to represent the corresponding low energy region. Thus, we have one value in each of the dihedral angles χ', χ, ϕ', and ϕ, and three values in each of ω', ω, and ψ angles. The possible combinations of these selected representative values for the seven dihedral angles resulted in 27 ($=1 \times 1 \times 1 \times 1 \times 3 \times 3 \times 3$) conformations for a dinucleoside monophosphate unit. Three out of these 27 cases, having $(\omega', \omega) \simeq (180°, 180°)$, were omitted bacause of their improbability as revealed by *ab initio* studies (Newton, 1973). The potential functions used by us also do not eliminate these conformations as improbable since we have not included the anomeric effects. (Recently Chandrasekaran and co-workers [1979] found this conformation to be probable in certain polynucleotide helices.) The rest of the 24 conformations were considered as starting points. The two sugar rings in the unit were also allowed to assume the two most-preferred pucker ^3E- and ^2E (consequently fixing the two ψ' angles), which resulted in four sugar-pucker sequences, ^3E-^3E, ^3E-^2E, ^2E-^3E, and ^2E-^2E for each unit (for definition of sugar puckering see Altona and Sundaralingam, 1972). Thus, when the alternative sugar-pucker sequences are also considered, a total of 96 (24×4) starting

points have resulted for each unit. This number is high enough to include all the best local minima and the global minimum for the molecular systems.

Studies treating the sugar ring as a flexible system were made by a few researchers (Cremer and Pople, 1975; Broyde et al, 1978; Broyde and Hingerty, 1979; Levitt and Warshel, 1978). In view of the reported marginal variations in the sugar ring conformation during the intrafamily transitions, and also of the enormous computation time involved, we treated the sugar unit as rigid, either in the ^3E or in the ^2E state, and the results were found to explain satisfactorily the conformational properties of DNA and RNA systems.

The geometrical parameters used in the present study are those of Lakshminarayanan and Sasisekharan (1970) which were given in our earlier article (Thiyagarajan and Ponnuswamy, 1978*a*). The computations were made using an IBM 370/155 system (IBM Corp., White Plains, N.Y.).

DESIGNATION OF PROBABLE CONFORMATIONS

In order to describe the results in a more comprehensive manner, the most probable conformations for the subunits were categorized and designated using the g^+ ($60° \pm 30°$), t ($180° \pm 30°$), and g^- ($-60° \pm 30°$) symbols. For designating a conformation, only the orientations of the three flexible dihedral angles ω', ω and ψ were considered. For example, if a conformation has ω' in the g^+, ω in t, and ψ in g^- regions, it is designated as a g^+tg^- conformer. When conformations similar/close to the helical structures observed for the polynucleotide fibers are encountered, they are referred to as helical conformers. The conditions under which we compare the theoretical results on fragmental units with the experimental observations on polynucleotides were detailed in our earlier article (Thiyagarajan and Ponnuswamy, 1978*a*).

RESULTS

The low energy conformations obtained after the minimization are listed in Tables I–IV for the four units, dApdA, dApdT, dTpdA, and dTpdT. It should be pointed out here that the results of the nonmixed sugar pucker states, namely, ^3E-^3E and ^2E-^2E of these subunits were already reported (Thiyagarajan and Ponnuswamy, 1978*a*). However, for comparison and completeness they are considered along with the results of the mixed sugar-pucker systems of these units.

It is interesting to note from the tables that out of the 96 starting conformations examined for each of the units, only 16, 16, 21, and 15 were found to fall within the considered energy limit (5 kcal/mol above the lowest energy located) for dApdA, dApdT, dTpdA, and dTpdT, respectively. This fact demonstrates the severe restrictions imposed by the bases on the preferred conformations of these units (Thiyagarajan and Ponnuswamy, 1978*a*; Broyde et al., 1978). The three subunits other than dTpdA seem to enjoy more or less equal freedom as far as their adoptable low energy conformations are concerned. Although the four units experience more or less similar restrictions, the order of preference of various probable states depends on the sequence in each unit. A detailed account of the results of the individual subunits is given below.

Preferred Conformations for dApdA Unit

Table I presents the low energy conformations for dApdA: four conformations become low energy cases for dApdA in the ^3E-^3E sugar-pucker sequence domain. The lowest energy

TABLE I

dApdA: ENERGY-MINIMIZED CONFORMATIONS LYING WITHIN 5.0 kcal/mol ABOVE THE LOWEST ENERGY FOUND

No.	Dihedral angles (*degrees*)					Relative energy (*kcal/mol*)	Description
	χ'	ω'	ω	ψ	χ		
^3E-^3E sugar-pucker sequence domain							
1	10	−65	−67	63	16	0.0*	$g^-g^-g^+$ (A-DNA)
2	25	−64	166	170	7	1.2	g^-tt (WC-DNA)
3	105	164	67	62	17	2.7	tg^+g^+
4	105	162	−62	65	21	3.0	tg^-g^+
^3E-^2E sugar-pucker sequence domain							
5	10	−77	178	176	20	0.1	g^-tt (alt-WC)
6	48	−62	−172	69	8‡	1.2	g^-tg^+
7	114	111	173	62	12	2.8	$g_s^+tg^+$§
8	105	162	−61	67	53	3.4	tg^-g^+
9	106	179	61	65	41	3.4	tg^+g^+
10	105	162	75	−60	16	4.2	tg^+g^-
^2E-^2E sugar-pucker sequence domain							
11	41	−93	−70	62	85	3.2	$g^-g^-g^+$ (B-DNA)
12	48	−112	161	167	57	3.3	g_s^-tt (WC-DNA)
13	78	−123	−74	69	92	3.7	$t_sg^-g^+$ (C-DNA)
^2E-^3E sugar-pucker sequence domain							
14	45	−149	−76	63	13	2.9	tg^-g^+
15	43	−92	−74	72	15	4.5	$g_s^-g^-g^+$ (alt-B-DNA)
16	29	−154	67	62	17	4.6	tg^+g^+

*Corresponds to −34.0 kcal/mol.

‡The ϕ' value is ~ −112°; in other conformations it is in the *trans* region.

§s denotes skewed.

conformer (global minimum) resembles the A-DNA form (Arnott and Hukins, 1973), and the Watson-Crick (WC) DNA form (Crick and Watson, 1954) has an energy of ΔE = 1.2 kcal/mol (Arnott et al., [1980*a*] designate the latter conformation as A*-DNA). Both these conformations promote reasonable base stacking. The remaining two conformations in this domain are open-type loop forms with (ω', ω, ψ) ≃ (t, g^+, g^+) and (t, g^-, g^+) and *high anti-anti* base orientation. The $\psi \simeq$ *trans* orientation noted in the case of WC-form is also a favorable and important state in the case of double-stranded polynucleotide structures (Zhurkin et al., 1978; Arnott et al., 1980*a*).

There are six low energy conformations for dApdA in the ^3E-^2E domain. In this category the preferred conformer has backbone dihedral angles similar to the WC-form and hence it is designated as alt-WC-DNA. When compared with the lowest energy A-DNA conformer, it has an energy of ΔE = 0.1 kcal/mol and promotes good base stacking (see Fig. 2). In this sugar-pucker sequence domain, $\psi \simeq g^+$ becomes important. We also note four loop promoting conformations with (ω', ω) around (g^-, t), (g^+, t), (t, g^-), and (t, g^+). It is interesting to note from Table I that conformations 3 and 4 are quite similar to conformations 9 and 8, respectively, except in their sugar-pucker sequence and base orientation of the 5′-nucleotide residue. This residue assumes a higher χ value in the ^3E-^2E domain which can be attributed to

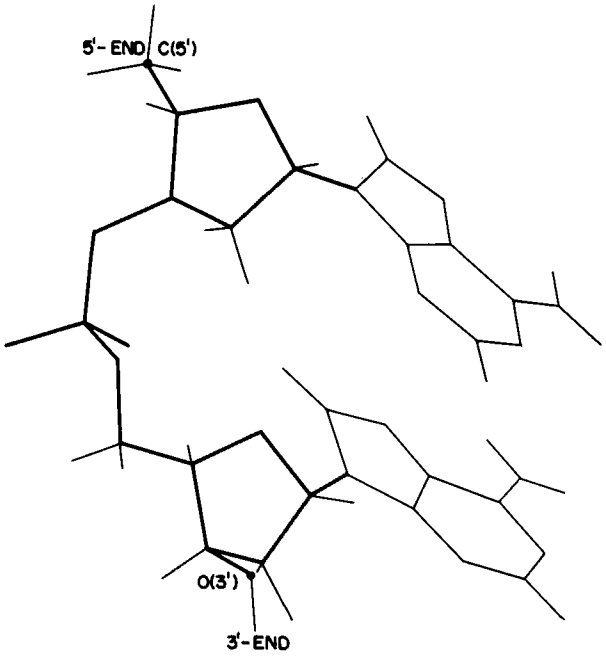

FIGURE 2 The alt-WC-DNA conformation predicted for dApdA (No. 5 in Table I). The base stacking is good.

the presence of the ^2E sugar pucker (Sundaralingam, 1969). One more open-type loop promoting conformation is found to be the low energy case with $(\omega', \omega) \simeq (t, g^+)$ but with $\psi \simeq g^-$ at $\Delta E = 4.2$ kcal/mol.

In the ^2E-^2E domain we find three low energy conformations, all of them, however, lying at energy level of $\Delta E = 3.2$ to 3.7 kcal/mol. In this category the B-DNA conformer is preferred with $\Delta E = 3.2$ kcal/mol. A conformer similar to WC-form has also been noted which, however, does not promote any base stacking feature.

The ^2E-^3E sugar pucker domain also contains only three low energy conformations, of which the tg^-g^+ conformer is the preferred one with $\Delta E = 2.9$ kcal/mol compared to A-form. Although this conformer aligns both bases on the same side, promoting a compact shape for the dimer when compared to a similar backbone conformation having the ^3E-^2E sugar-pucker sequence, it does not exhibit any base stacking. This fact is illustrated in Fig. 3 by superposing the two conformations, one with ^3E-^2E and the other with ^2E-^3E sugar sequences. In this domain a backbone conformation similar to the B-form is found to be preferred and this is designated as alt-B-DNA. This conformational state does not promote any base stacking.

An interesting general observation is the presence of the backbone conformations $(\omega', \omega, \psi) \simeq (t, g^+, g^+)$ and (t, g^-, g^+) in the three sugar-pucker domains ^3E-^3E, ^3E-^2E and ^2E-^3E and the conformation (g^-, g^-, g^+) in the domains ^3E-^3E, ^2E-^2E and ^2E-^3E. It is likely that these conformational states play a role in the phenomenon of dynamic equilibrium between ^2E and ^3E sugar-pucker states noted by Kondo et al. (1972) in their NMR studies on deoxydinucleoside monophosphates.

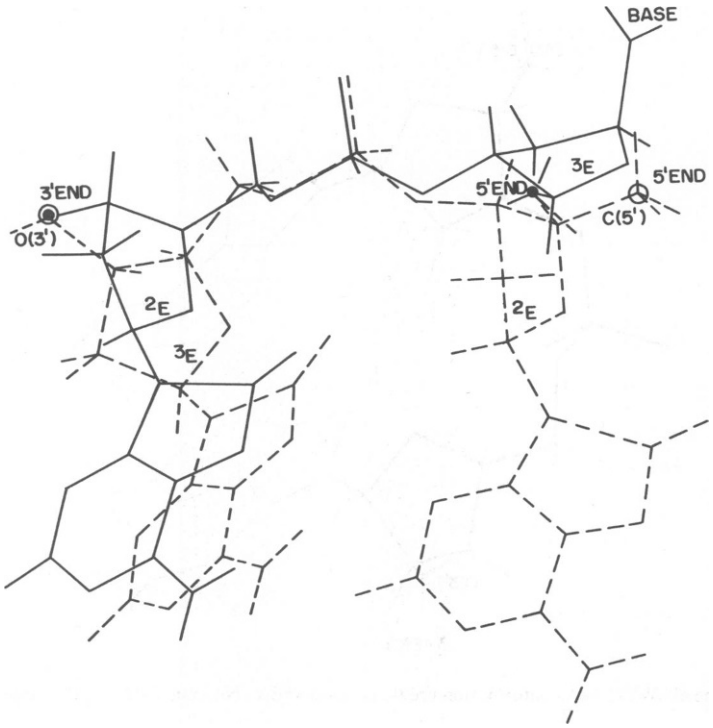

FIGURE 3 In this diagram the difference in the backbone conformation introduced by the two sugar-pucker states, namely, ^2E-^3E and ^3E-^2E is depicted. The broken line diagram depicts the conformation tg^-g^+ – ^2E-^3E (No. 14 in Table I) and the continuous line diagram the conformation tg^-g^+ – ^3E-^2E (No. 8 in Table I); the C(5') . . . O(3') virtual bonds are aligned as closely as possible. See text for the characteristics of these conformations.

Preferred Conformations for dApdT Unit

This unit also enjoys a conformational freedom equal to that of dApdA, as far as the number of local minima is concerned. Table II indicates that the same four conformations that were preferred for dApdA, become low energy cases for this unit in the ^3E-^3E sugar-pucker sequence domain. However, the A-DNA and WC-DNA forms are less stable for this unit compared with dApdA.

The lowest energy conformer (global minimum) for dApdT occurs in ^3E-^2E domain resembling the WC-form. This promotes good base stacking. The alt-π-bend conformation with $(\omega', \omega, \psi) \simeq (g^-, t, g^+)$ depicted in Fig. 4 falls next in energy rank. In this conformation, the backbone is slightly elongated compared to the conventional π-bend (Kim and Sussman, 1976) having ^3E-^3E pucker sequence. Here again we observe two open-type loop promoting conformations with $(\omega', \omega, \psi) \simeq (t, g^-, g^+)$ and (t, g^+, g^+) at ΔE = 1.3 and 1.4 kcal/mol, respectively. To illustrate the effect of change in the sugar-pucker sequence from ^3E-^3E to ^3E-^2E, we superpose the two conformations with $(\omega', \omega, \psi) \simeq (t, g^+, g^+)$, one in the ^3E-^3E and the other in the ^3E-^2E domains in Fig. 5. In another superposition diagram, Fig. 6, we illustrate the structural change produced by the alteration in ω value from g^+ to g^- state in the

TABLE II

dApdT: ENERGY-MINIMIZED CONFORMATIONS LYING WITHIN 5.0 kcal/mol ABOVE THE LOWEST ENERGY FOUND

No.	Dihedral angles (*degrees*)					Relative energy (*kcal/mol*)	Description
	χ'	ω'	ω	ψ	χ		
^3E-^3E sugar-pucker sequence domain							
1	39	−62	169	167	15	0.9	g^-tt (WC-DNA)
2	5	−61	−45	61	11	2.4	$g^-g^-g^+$ (A-DNA)
3	105	163	67	54	8	3.8	tg^+g^+
4	106	166	−67	55	8	4.5	tg^-g^+
^3E-^2E sugar–pucker sequences domain							
5	10	−76	−179	177	28	0.0*	g^-tt (alt-WC)
6	18	−57	−176	63	24‡	1.0	g^-tg^+
7	106	166	−71	58	34	1.3	tg^-g^+
8	106	163	65	57	34	1.4	tg^+g^+
9	101	−92	175	−60	34	3.8	$g_t^-tg^-$
10	104	65	178	−53	32	3.9	g^+tg^-
11	105	160	93	179	36	4.2	tg_t^+t
^2E-^2E sugar-pucker sequence domain							
12	45	−116	−55	67	62	3.9	$g_t^-g^-g^+$ (B-DNA)
13	26	−145	110	155	−6	4.9	$t_tg_t^+t$
14	76	−160	−69	60	32	4.9	tg^-g^+ (C-DNA)
^2E-^3E sugar-pucker sequence domain							
15	61	−120	166	148	18	4.7	$g_t^-tt_t$ (alt-WC)
16	3	−117	−53	51	18	4.9	$g_t^-g^-g^+$ (alt-B-DNA)

*Corresponds to −32.6 kcal/mol.

‡The ϕ' value is −112°.

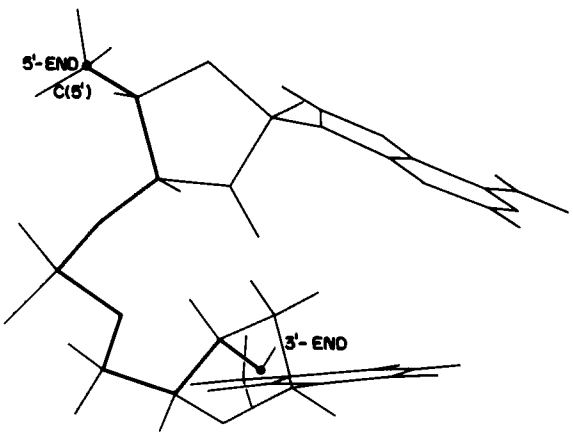

FIGURE 4 The alt-π-bend state predicted for dApdT (No. 6 in Table II); this could introduce a sharp turn in the polynucleotide.

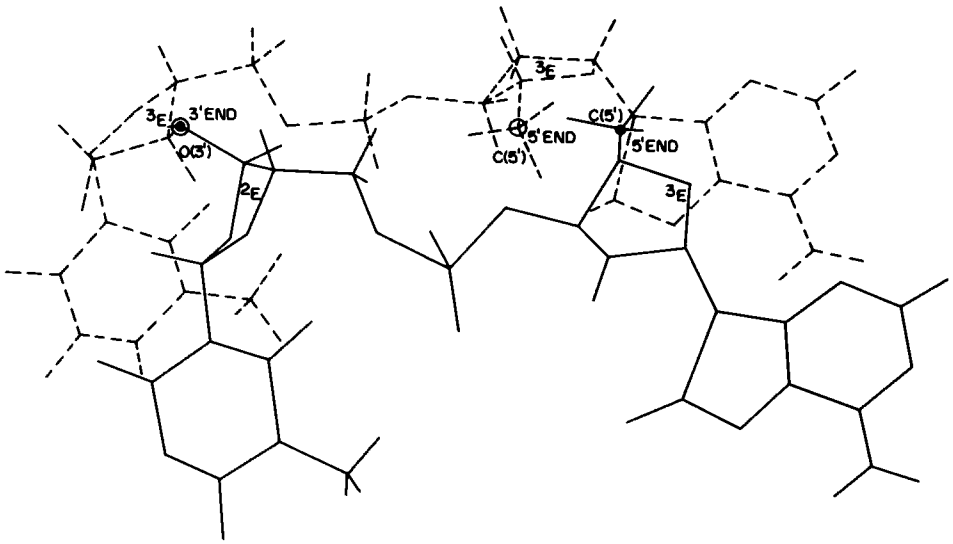

FIGURE 5 In this diagram the difference in the backbone conformation introduced by the two sugar-pucker states, namely, ^3E-^3E and ^3E-^2E is shown. The continuous line corresponds to the tg^+g^+ – ^3E-^2E conformation (No. 8 in Table II) while that in broken line to the tg^+g^+ – ^3E-^3E conformation (No. 3 in Table II). See legend to Fig. 3 for other details.

conformation with $(\omega', \psi) \simeq (t, g^+)$ and ^3E-^2E pucker sequence. We note in Fig. 5 that the backbone is elongated when the sugar pucker changes from ^3E-^3E to ^3E-^2E: this is found to be true in conformations occurring in other (ω', ω, ψ) domains also. Fig. 6 reveals that both the relevant conformations are less compact loop structures; while the bases are disposed on the opposite sides of the backbone in $t\,g^-g^+$ conformer, they are placed on the same side in tg^+g^+

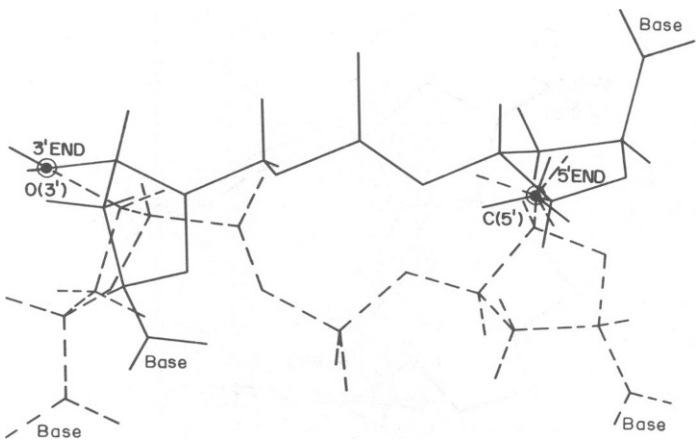

FIGURE 6 The difference introduced in the backbone course due to the movement of ω from g^- to g^+ state is illustrated in this diagram. The continuous line corresponds to the tg^-g^+ – ^3E-^2E conformation (No. 7 in Table II) while that in broken line to the tg^+g^+ – ^3E-^2E conformation (No. 8 in Table II). See legend to Fig. 3 for other details.

conformer. We also find two more conformations with $\psi \simeq g^-$ and one with $\psi \simeq t$ as low energy cases for dApdT in the ^3E-^2E domain.

For this unit too the ^2E-^2E sugar-pucker domain contains only three low energy conformations, of which the B-form (Arnott et al., 1969) is the preferred one with $\Delta E = 3.9$ kcal/mol (recently Arnott et al. [1980*b*] have suggested $[\omega', \omega] \simeq [t, g^-]$ for B-DNA). For this conformer the ^2E-^3E sequence is very much constrained, as we note only two local minima, of which alt-B-DNA is one.

Preferred Conformations for dTpdA

From Table III we note that the same conformations predicted for dApdA and dApdT turn out to be the low energy cases in the ^3E-^3E domain. The A-DNA and the WC-DNA forms are at energies of $\Delta E = 0.8$ and 1.2 kcal/mol compared with the lowest energy conformer. However, the two loop promoting conformers do not have *high anti*-base orientation in the

TABLE III

dTpdA: ENERGY-MINIMIZED CONFORMATIONS LYING WITHIN 5 kcal/mol ABOVE THE LOWEST ENERGY FOUND

No.	Dihedral angles (*degrees*)					Relative energy (*kcal/mol*)	Description
	χ'	ω'	ω	ψ	χ		
^3E-^3E sugar-pucker sequence domain							
1	15	−61	−73	66	26	0.8	$g^-g^-g^+$ (A-DNA)
2	27	−67	160	174	17	1.2	g^-tt (WC-DNA)
3	25	167	68	62	17	3.8	tg^+g^+
4	25	165	−62	65	30	4.2	tg^-g^+
^3E-^2E sugar-pucker sequences domain							
5	16	−72	−178	173	26	0.0*	g^-tt (alt-WC-DNA)
6	32	−69	−65	73	56	0.9	$g^-g^-g^+$ (alt-A-DNA)
7	18	−57	−177	67	16‡	0.9	g^-tg^+
8	21	166	−62	66	55	3.3	tg^-g^+
9	21	167	72	65	41	3.4	tg^+g^+
10	21	59	−168	−58	16	4.5	g^+tg^-
11	19	−80	174	−61	16	4.7	g^-tg^-
12	21	167	73	−59	18	5.0	tg^+g^-
^2E-^2E sugar-pucker sequence domain							
13	55	−89	−73	67	86	0.6	$g^-g^-g^+$ (B-DNA)
14	42	−50	−166	55	4	2.6	g^-tg^+
15	30	−122	−60	60	109	3.0	$t_1g^-g^+$ (C-DNA)
16	37	−158	83	178	19	3.4	tg^+t
17	34	90	145	53	−8	4.3	$g^+t_1g^+$
^2E-^3E sugar-pucker sequence domain							
18	53	−88	−76	57	60	1.4	$g^-g^-g^+$ (alt-B-DNA)
19	45	−119	140	167	22	3.7	$g_1^-t_1t$ (alt-WC-DNA)
20	43	−169	66	62	97	4.1	tg^+g^+
21	41	−149	−74	59	21	4.4	$t_1g^-g^+$

*Corresponds to −33.9 kcal/mol.
‡The ϕ' value is −115°.

3'-nucleotide which is in contrast to the observations made for dApdA and dApdT. This is probably due to the fact that a pyrimidine base remains only in the normal *anti* region.

For dTpdA also, the lowest energy conformation belongs to the ^3E-^2E sugar-pucker sequence domain and resembles the WC-form. This conformation promotes a reasonable base-stacking property. As noted for dApdA and dApdT, the (g^-tg^+), (tg^-g^+) and (tg^+g^+) conformations become low energy cases for this unit also. Thus, these three (ω', ω, ψ) domains become most probable in all the DNA subunit systems. The most interesting feature observed in the case of dTpdA is the preference of alt-A-DNA form with a backbone like the A-DNA, but with ^3E-^2E sugar-pucker sequence at ΔE = 0.9 kcal/mol. Such a conformation, was not found to be a low energy case both in dApdA and dApdT. This observation has a striking relevance to the drug intercalation phenomenon in nucleic acids as we note that this conformation is observed in drug-dinucleotide complex crystals (Tsai et al., 1977; Jain et al., 1977; Westhof and Sundaralingam, 1980). It is also observed that for drug-dinucleotide intercalation to occur, the base sequence should be of the pyrimidine-purine type. Our study explains the specificity for such a phenomenon in terms of the intrinsic stability of such a conformation for the dinucleoside monophosphate when the base sequence is of the pyrimidine-purine type. The intercalation of the drug is further noted to involve alterations in ϕ and χ parameters by Berman et al. (1978).

In the ^2E-^2E sugar-pucker domain there are five low energy conformations in contrast to the cases of dApdA and dApdT, for which we noted only three conformations. The B-form is again the preferred state in this sugar-pucker domain. The C-form (Arnott and Selsing, 1975) has an energy of ΔE = 3.0 kcal/mol. The other conformations that are preferred in this domain have $(\omega', \omega, \psi) = (g^-, t, g^+)$, (t, g^+, t), and $g_s^+, t, g^+)$ (s stands for skewed orientation). In this domain, we find the two conformers B-DNA and C-DNA exhibiting *high anti*-base orientation in the 5'-nucleotide as in the case of dApdA.

The alt-B-DNA conformer is the preferred state for dTpdA in the ^2E-^3E sugar-pucker domain with an energy of ΔE = 1.4 kcal/mol. Three more loop-type conformers with $(\omega', \omega, \psi) \simeq (g_s^-, t_s, t)$, (t, g^+, g^+) and (t_s, g^-, g^+) become low energy cases. The last of the above three conformations keeps both the bases on the same side as was found in dApdA.

On the whole, the dTpdA unit seems to enjoy more conformational freedom compared with dApdA and dApdT, as far as the number of preferred local minima is concerned. Another interesting point is the higher stabilities of A-DNA, B-DNA, and their mixed pucker states. The energetics show that the stability decreases in the order of B-DNA > A-DNA > alt-A-DNA > alt-B-DNA. In the phenomenon of B $\rightleftharpoons$ A transition in DNA, the mixed sugar-pucker states may serve as intermediates so that where T—A sequences occur, they may be recognized as initiating regions in polydeoxynucleotides. The probable conformational states, intermediates and the probable mechanism involved in the B $\rightleftharpoons$ A transitions are indicated in Fig. 7 with a series of four pictures.

Preferred Conformations for dTpdT Unit

Table IV lists the fifteen conformations preferred for dTpdT. Unlike the other three units, this unit prefers only two conformations in the ^3E-^3E domain and they are the WC-form and the (tg^+g^+)-bend at energies of ΔE = 0.6 and 4.3 kcal/mol, respectively. Both A-DNA and (tg^-g^+) conformers become high energy cases.

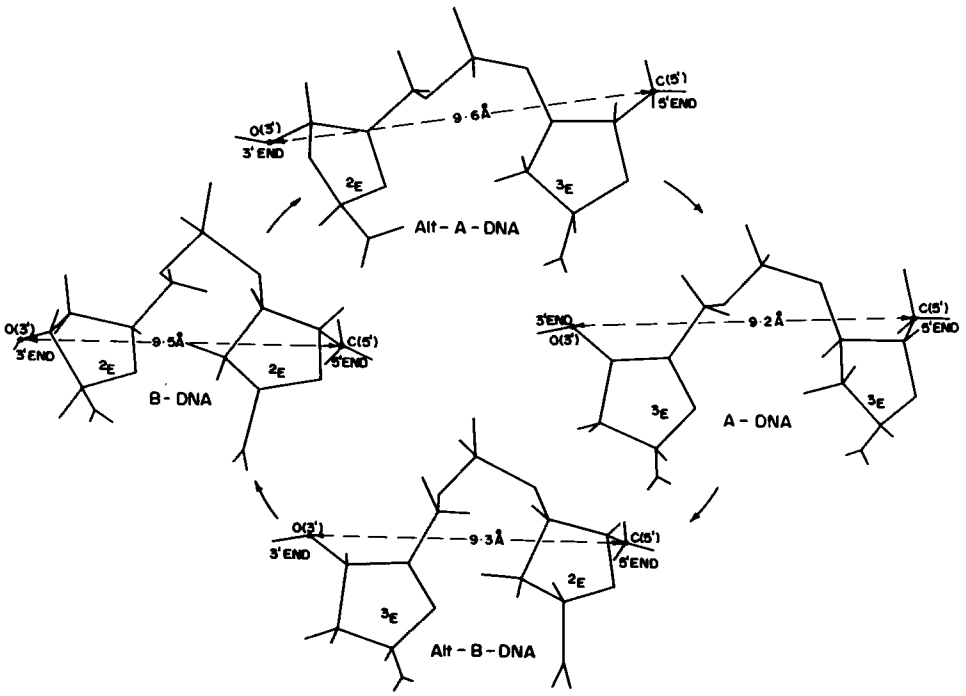

FIGURE 7 The cycle of events predicted to be present in the B form ⇌ A form transition in DNA via a dimeric subunit.

Here again the lowest energy conformation occurs in the ^3E-^2E domain with $(\omega', \omega, \psi) \simeq (g^-, t, g^+)$ which is capable of promoting a sharp turn in the backbone. The rest of the seven conformations are open-type loop structures that include the tg^-g^+ and tg^+g^+ conformations. In contrast to the other three units, the alt-WC-form does not become a low energy case.

In the ^2E-^2E sugar-pucker sequence domain, four conformations become low energy states. The B-form is not preferred. The C-DNA form has an energy of ΔE = 3.0 kcal/mol and it closely resembles the conformation observed in the crystal structure of pdTpdT (Camerman et al., 1976). In the ^2E-^3E sugar-pucker domain only one conformation with $(\omega', \omega, \psi) \simeq (g_s^-, t_s, t)$ turns out to be the low energy case with an energy of ΔE = 3.7 kcal/mol. Two other important features noted are, (*a*) the *high anti*-base orientation is completely absent, and (*b*) the sugar-pucker sequence ^2E-^2E is noted to contain more local minima than ^3E-^3E.

DISCUSSION

Preferred Sugar-Pucker States

An analysis of the number of low energy conformations associated with the four sugar-pucker sequences in the four DNA subunits reveals that the ^3E-^3E, ^3E-^2E, ^2E-^2E, and ^2E-^3E occur respectively, 4, 6, 3, and 3 times in dApdA; 4,7,3, and 2 times in dApdT; 4, 8, 5, and 4 times in dTpdA; and 2, 3, 4, and 1 times in dTpdT. Thus, as far as the number of local minima is concerned the ^3E-^2E is the most preferred and the ^2E-^3E is the least preferred sequences in all

TABLE IV

dTpdT: ENERGY-MINIMIZED CONFORMATIONS LYING WITHIN 5 kcal/mol ABOVE THE LOWEST ENERGY FOUND

No.	Dihedral angles (*degrees*)					Relative energy (*kcal/mol*)	Description
	χ'	ω'	ω	ψ	χ		
^3E-^3E sugar-pucker sequence domain							
1	20	−74	165	163	16	0.6	g^-tt (WC-DNA)
2	25	169	67	54	7	4.3	tg^+g^+
^3E-^2E sugar-pucker sequence domain							
3	13	−56	−177	65	18*	0.0‡	g^-tg^+
4	19	−168	−72	58	18	1.3	tg^-g^+
5	20	167	65	57	35	1.4	tg^+g^+
6	18	56	−158	−54	32	3.2	g^+tg^-
7	20	65	−174	172	33	3.6	g^+tt
8	21	167	74	−59	35	3.7	tg^+g^-
9	42	170	88	180	32	4.0	tg^+t
10	30	−79	167	−66	21	4.9	g^-tg^-
^2E-^2E sugar-pucker sequence domain							
11	37	−144	98	180	2	2.7	$t_as_a^+t$
12	41	−166	−52	62	32	3.0	tg^-g^+ (C-DNA)
13	34	−157	64	58	34	3.4	tg^+g^+
14	42	−71	−177	61	37	4.3	g^-tg^+
^2E-^3E sugar-pucker sequence domain							
15	41	−99	148	163	24	3.7	$g_s^-t_at$

*ϕ' value is −115°.

‡Corresponds to −32.0 kcal/mole

the four units. Considering the two nonmixed sugar-pucker sequences it appears that the ^3E-^3E has a slight edge over the ^2E-^2E in dApdA and dApdT, while the ^2E-^2E is slightly more favored than ^3E-^3E in dTpdA and dTpdT as far as the number of local minima is concerned. However, when the stabilities of the conformations associated with the four sugar-pucker sequence domains in all the units are also considered the ^3E-^3E domain, in general, seems to be more favored than the ^2E-^2E domain in the three units other than dTpdT. These results are pertinent to the respective molecules in an isolated condition. However, in real systems, (crystals or solution) one encounters many other interactions like crystal packing, base pairing, and solvent-solute association, which were not considered in the present study. These additional interactions will certainly influence the preference of sugar-pucker sequences. In one of our earlier studies (Ponnuswamy and Thiyagarajan, 1978) we illustrated how the solvent-solute interaction can play its role in stabilizing similar backbone conformations, but causing differences in their sugar-pucker sequences. Our present results on the nonmixed sugar sequences are consistent with the studies of Broyde et al. (1978). The classical potential energy calculations on dApdT duplex (Kister and Dashevsky, 1976) and on Poly (dA-dT) and Poly (dG-dC) duplexes (Calascibetta et al., 1975) also show similar results on the order of preference of ^3E-^3E and ^2E-^2E domains.

Flexibility of Certain Conformations

From Table I we note that the three triplet conformational states, namely, 2, 5 and 12; 3, 9 and 16; and 4, 8 and 14 have backbone dihedral angles in the domains of $(\omega', \omega, \psi) \simeq (g^-, t, t)$, (t, g^+, g^+), and (t, g^-, g^+), respectively. The conformational states in each triplet, however, are associated with differing sugar-pucker sequences and also differing base orientations. In each of the triplets, the ^3E-^3E and the ^3E-^2E pucker sequences promote mainly a *high anti*-base orientation in the 3'-nucleotide residues, while the ^2E-^2E and ^2E-^3E sequences, usually, exhibit only a normal *anti* base orientation. An analysis of the structural shapes and the end-to-end distances (see Table V) for the three conformations in each triplet shows that, usually, the ^3E-^2E pucker sequence promotes an elongated, less compact backbone compared with the case of ^3E-^3E sequence. On the other hand, the ^2E-^3E sequence produces a slightly more compact structure than that produced by the ^2E-^2E sequence. Among all the sequences the ^3E-^3E produces the most compact structures. In conformations with $\psi \simeq g^-$ the end-to-end distances are smaller for the ^2E-^2E sequence than those for ^3E-^2E. Hence, for a particular set of backbone dihedral angles, the preference of a specific sugar-pucker sequence in a polynucleotide chain may be according to the local requirement.

There also exists ample interconversion flexibility in a few pairs of the low energy conformational states in all the four units. The two conformations 2 and 5 in Table I do not differ much in their stability and so also the cases 4 and 14. The dihedral angles in each of these pairs remain almost the same, but the sugar-pucker sequences are different. Inasmuch as the energy barrier separating the ^3E and ^2E states is very small we suggest that in the above mentioned cases the sugar conformation may oscillate between the two states without imposing much strain on the backbone. Because of the presence of such conformational states, a dynamic equilibrium between the ^2E and ^3E states is probably maintained in solution, as has been noted from the high resolution NMR studies of deoxydinucleoside monophosphates (Kondo et al., 1972; Wood et al., 1975; Cheng and Sarma, 1977). All these features discussed

TABLE V

END-TO-END [C(5') . . . 0(3')] DISTANCES IN CERTAIN TYPICAL CONFORMATIONS FOR DIFFERENT SUGAR-PUCKER SEQUENCES

Type of conformer	[C(5') . . . 0(3')] distances			
	^3E-^3E	^3E-^2E	^2E-^2E	^2E-^3E
		Å		
g^-tt	9.85	10.02	9.74	9.81
g^-tg^+	7.98	8.81	8.21	9.47
$g^-g^-g^+$	9.24	9.64	9.46	9.33
tg^-t	6.63	8.02	9.43	8.11
g^+tg^-	4.36	5.77	6.07	4.57
tg^-g^-	7.53	8.40	9.50	8.43
tg^+t	8.89	9.25	10.20	9.62
tg^-g^+	8.00	8.75	10.00	9.26
tg^+g^+	6.52	7.94	9.19	7.76
g^+g^+t	6.01	6.29	7.24	7.23

$\phi' \simeq -130°$, $\phi \simeq 180°$; $g^- \simeq -70°$; $t \simeq 180°$; $g^+ \simeq 60°$

for dApdA are also noted in the case of other units (see Tables II to IV). These results thus clearly demonstrate the part played by the base and sugar-pucker sequences on the overall conformational properties of the DNA subunits.

Preferred Base Orientations

The present study also throws light on the most probable orientations for the two bases under conditions of mutual adjustments about various backbone dihedral angles in the dimeric subunits. An analysis of χ' and χ values given in Tables I to IV indicates that in most of the predicted nonhelical conformations for dApdA and dApdT adopting ^3E-^3E or ^3E-^2E sugar-pucker sequence (3, 4, 7–10 in Table I and 3, 4, 7–11 in Table II), the base of the 3′-nucleotide residue is found to assume *high anti* orientation. However, such nonhelical conformations adopting the ^2E-^2E and ^2E-^3E sugar sequences do not exhibit this specific feature. In contrast to this distinct observation, the other two units dTpdA and dTpdT (having a pyrimidine in the 3′-nucleotide residue) assume only the *normal anti* base orientation in the 3′-nucleotide residue. This indicates that the base orientation will turn to the *high anti* region only if the base is a purine, which is consistent with the known experimental observations (Sundaralingam, 1969). It is also obvious from Tables II to V that the *high anti* base orientation is almost completely absent in the 5′-nucleotide residue of all the four subunits; the only exception to this observation is the C-DNA conformer predicted for the two units dApdA and dTpdA which adopt such an orientation in the 5′-nucleotide residue. The absence of the *high anti* base orientation in the 5′-nucleotide residue in almost all the four subunits could be attributed to the influence of the 5′-phosphate group which invariably keeps the base (purine) in the *normal anti* region. This 5′-phosphate influence which keeps a corresponding base in the *normal anti* region was designated as the "phosphate effect" (Sundaralingam, 1973; Thiyagarajan and Ponnuswamy, 1979). There is an exception to the above observation as we note *high anti* orientations of the base in the 5′-nucleotide residue in the case of C-DNA conformer predicted for dApdA and dTpdA units which could be interpreted to have the following importance; occasionally a purine unit can adopt *high anti* orientation in a polynucleotide at sites where ^2E sugar-pucker states occur in a sequence, and the C-DNA form may be the most probable conformation in these sites.

Base-Stacking Properties

The base-stacking parameters calculated for the low energy conformations predicted for the four units are given in Table VI. In general, the low energy conformations of the subunits that resemble the regular helical structures observed for polynucleotide fibers usually exhibit reasonable base-stacking property. In dApdA, dApdT and dTpdA, the conformations similar to the A-DNA, B-DNA, and WC-DNA helical forms (excepting the B-DNA conformer for dApdT) exhibit recognizable base-stacking feature. However, for dTpdT, only the conformation corresponding to the WC-DNA exhibits reasonable base-stacking feature and the A-form and B-form conformers become high energy cases. The little stacking believed to be present in dTpdT in solution (Wood et al., 1974) is supported by the present study. In low energy conformations associated with the mixed sugar-pucker sequence domains, the alt-WC conformer with ^3E-^2E sugar state exhibits good base-stacking properties in the three subunits dApdA, dApdT, and dTpdA. The alt-A-DNA conformer with ^3E-^2E sequence predicted for dTpdA does not exhibit any base-stacking character, even though the bases are brought close

TABLE VI

BASE-STACKING PROPERTIES PREDICTED IN DEOXYDINUCLEOSIDE MONOPHOSPHATES

Molecule	Type of conformer	Sugar-pucker sequence	Mean distance between base planes	Angle between base planes	Over-lapping area*	Stacking property‡
			(Å)	*(degree)*	*(%)*	
dApdA	A-DNA	^3E-^3E	3.54	17	31	Good
	WC-DNA	^3E-^3E	3.33	15	21	Reasonable
	B-DNA	^2E-^2E	4.21	15	51	Reasonable
	alt-WC	^3E-^2E	3.76	8	35	Good
dApdT	A-DNA	^3E-^3E	3.53	4	42	Good
	WC-DNA	^3E-^3E	3.81	17	81	Good
	B-DNA	^2E-^2E	4.54	33	83	Poor
	alt-WC	^3E-^2E	3.97	15	40	Good
dTpdA	A-DNA	^3E-^3E	3.69	16	61	Good
	WC-DNA	^3E-^3E	3.59	20	35	Good
	B-DNA	^2E-^2E	3.57	15	73	Good
	alt-WC	^3E-^2E	3.66	15	30	Good
	alt-A-DNA	^3E-^2E	3.80	28	2	Poor
	alt-B-DNA	^2E-^3E	4.13	20	10	Poor
dTpdT	WC-DNA	^3E-^3E	3.77	23	55	Reasonable

*The extent of overlapping is said to be 100% when the areas of two similar bases get completely overlapped and also when the base with smaller area is completely overlapped by the base of larger area.

‡The base stacking is said to be "good" when the mean distance between the base planes is in the range of 3.5–4.0 Å, the angle between the base planes is within 20°, and the extent of base overlap is >30%. The terms "reasonable" and "poor" are used with respect to the definition given to a good base stacking.

to a distance of 4.0 Å. In general, the extent of base stacking in the dimeric subunits of DNA decreases in the order of purine-purine, pyrimidine-purine, purine-pyrimidine, pyrimidine-pyrimidine base sequences. This assessment is consistent with the solution conformation of dinucleoside monophosphates (Ts'o et al., 1969; Lee and Tinoco, 1977) and also with our study on the RNA subunits described in the following article. However, the base-stacking properties revealed by the NMR work of Cheng and Sarma (1977) differ from this observation. When the base stacking properties corresponding to ^3E-^2E and ^2E-^3E sugar sequences are considered it seems that the former promote better stacking than the latter.

Probable Sites for alt-A-DNA and alt-B-DNA Conformations

It is interesting to note that the alt-A-DNA conformer with $(\omega', \omega, \psi) \simeq (g^-, g^-, g^+)$ and ^3E-^2E sugar sequence is a low energy case only for dTpdA. This conformation promotes a slight elongation in the backbone as well as in the base-to-base distance compared with the conventional A-DNA conformer. In a base-paired duplex this conformer with considerable alterations in ϕ and χ can provide enough room for a chromophoric drug molecule to intercalate between the base pairs (Berman et al., 1978). This is the type of conformation observed in the drug-dinucleotide complex crystals (Tsai et al., 1977; Jain et al, 1977; Westhof and Sundaralingam, 1980) wherein the base sequence is pyrimidine-purine type.

In one of our previous articles (Thiyagarajan and Ponnuswamy, 1978*a*) we discussed the

probable sites and the involved energetics in the B ⇌ A transitions in DNA. We predicted that these transitions may be initiated mainly at the sites where py-pu (T-A) base sequence is present. The present study adds further information to this suggestion. According to the present study the probable cycle in which the B ⇌ A transitions might occur is B-DNA → Alt-A-DNA → A-DNA → Alt-B-DNA → B-DNA (see Fig. 7). We propose the following events to occur during the cycle. To start with, the DNA exists in the more stable B-form (ΔE = 0.6 kcal/mol). The B → A transition might occur by way of opening up the chain due to flipping in the sugar-pucker conformation of the pyrimidine residue from ^2E to ^3E at the py-pu base sequence site. This is the alt-A-DNA form having an energy of ΔE = 0.9 kcal/mol. The A-DNA conformer (ΔE = 0.8 kcal/mol) is attained when such disturbance travels through the polynucleotide chain in the 5′-end to 3′-end direction. For the A → B transition to occur once again, the chain may open up and the alt-B-DNA form is induced because of the flipping of the sugar conformation of the pyrimidine residue from ^3E to ^2E state at the same or other py-pu sites. The perpetuation of this disturbance in the 5′-end to 3′-end direction of the polynucleotide chain will restore the B-form conformation for DNA. It is likely that all these processes may be operated by the interactions of some specific proteins with DNA in the nucleus.

Kinked and Nonkinked Conformational States and Chromatin Structure

In order to explain the chromatin structure a variety of conformations for the polynucleotide were recently tried (Crick and Klug, 1975; Sobell et al., 1977; Sussman and Trifonov, 1978; Levitt, 1978). Crick and Klug proposed that a kink promoting conformation with (ω', ω, ψ) ≃ (g^-, g^-, t) and ^2E-^2E sugar-pucker sequence, if it occurs at definite intervals along the polynucleotide backbone, will enable the DNA to fold tightly into a compact shape. This conformation, however, is not found to be a low energy case in the present study. Sobell et al., (1977) suggested a different type of kink, from their x-ray investigations on drug-dinucleoside phosphate complex crystals, with (ω', ω, ψ) ≃ (g^-, g^-, g^+) and ^3E-^2E sugar-pucker sequence. This conformer is found to be a low energy case only for dTpdA (Table III, no. 6) in the present study.

Unlike these two models, the ones proposed by Sussman and Trifonov (1978) and by Levitt (1978) contain nonkinked conformations (without the disruption of base-stacking interaction) to enable the DNA to smoothly bend to attain a compact shape. These two models have slightly altered torsion angles in the sugar phosphate moieties relative to B-DNA, varying gradually along the chains. This is quite consistent with the present study as the predicted B-DNA conformation for the four units have slightly different torsion angles depending on the base sequence. Hence such base-sequence–dependent flexibility noted at the dimeric level, if present in the polynucleotide chain also, would allow the DNA to smoothly bend into a compact shape.

Comparison with Experimental Observations: Crystal Structures of DNA Subunits

In the literature available to date, we find only four crystal structures of oligodeoxynucleotides, one at dimeric, pdTpdT (Camerman et al., 1976), two at tetrameric, d-pApTpApT (Viswamitra et al., 1978) and d-pCpGpCpG (Drew et al., 1980) and one at hexameric, d-pCpGpCpGpCpG (Wang et al., 1979) levels. The conformation observed at A-T portion of

d-pApTpApT and at G-C portion of d-pCpGpCpGpCpG crystals, with $(\omega', \omega, \psi) = (g^-, g^-, g^+)$ and ^3E-^2E sugar-pucker sequence, has an energy of 6.1 kcal/mol in the case of dApdT. The G-C portion of the d-pCpGpCpG crystal (Drew et al., 1980) however has a different conformation with $(\omega', \omega, \psi) = (g^-, g^-, g^+)$ and C(1')-*exo*-C'(2')-*endo* pucker sequence. This is a variant of B-DNA conformation which happens to be a low energy case according to the present study. The conformation observed at the T-A portion of d-pApTpApT crystal with $(\omega', \omega, \psi) = (t, g^-, g^+)$ and ^2E-^3E pucker sequence turns out to be similar to the one predicted for dTpdA (Table III, no. 21) at $\Delta E = 4.4$ kcal/mol. However, the conformation observed at G-C portion of the left handed helical structure conformations found for d-pCpGpCpG and d-pCpGpCpGpCpG crystals with $(\omega', \omega, \psi) = (g^+, g^+, t)$ is not a low energy case according to the present study. This may be due to the completely different base sequence present in these crystals compared to the one studied by us. Furthermore, there are extensive intermolecular hydrogen bonds in the crystalline state that may yield additional energy, so that these states may be of higher energy in their absence.

High Resolution NMR Studies on DNA Subunits in Solution

Recently Cheng and Sarma (1977) predicted that the predominant conformations for dimeric subunits of DNA in solution are those with ^2E sugar pucker, $\psi = g^+$, and (ω', ω) in any one of the orientations such as (g^+, g^+), (g^-, g^-), (g^+, t), (t, g^+), (t, g^-) and (g^-, t). Sizable fraction of conformational states with other sugar puckers also exist. According to the present study all the conformations except that with $(\omega', \omega, \psi) \simeq (g^+, g^+, g^+)$ become low energy cases in the dimeric subunits of DNA. However, their stabilities are base-sequence dependent. As already mentioned, dTpdT shows little base stacking according to the present study. This again is consistent with the studies made by Wood et al., (1974) and Cheng and Sarma (1977). Another important result emerged from the NMR studies of Cheng and Sarma (1977) is that even in solution the ^3E state is important for dimeric subunits of DNA, contributing about 20 to 50% to the conformational blend depending on the base sequence. The present study explains this observation in that, depending on the base sequence the prominent low energy conformations may adopt different sugar-pucker sequence domains and contribute to the conformational blend observed by Cheng and Sarma for deoxydinucleoside monophosphate.

CONCLUSION

The present study throws ample light on the overall conformational behavior of the deoxydinucleoside monophosphate and explains a large amount of experimental data from the point of view of the prominent conformations for the individual units as well as the important local low energy conformations which probably constitute the conformational blend noted in the solution studies. This study also predicts the possible sites of occurrence of the mixed sugar pucker states in the polynucleotides from which the probable mechanism involved in the B $\rightleftharpoons$ A transition is predicted.

Received for publication 23 April 1979 and in revised form 21 April 1981.

REFERENCES

Altona, C., and M. Sundaralingam. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc.* 94:8205–8212.

Arnott, S., S. D. Dover, and A. J. Wonacott. 1969. Least squares refinement of the crystal and molecular structure of DNA and RNA from X-ray data and standard bond lengths and angles. *Acta Crystallogr. Sect. B. Struct. Crystallogr. Cryst. Chem.* B25:2192–2206.

Arnott, S., and D. W. L. Hukins. 1973. Refinement of the structure of A-DNA and implications for the analysis of X-ray diffraction data from fibers of Biopolymers. *J. Mol. Biol.* 81:93–105.

Arnott, S., and E. Selsing. 1975. The conformation of C-DNA. *J. Mol. Biol.* 98:265–269.

Arnott, S., R. Chandrasekaran, P. J. Bond, D. L. Birdsall, A. G. W. Leslie, and L. C. Puigjaner. 1980*a*. The seventh Aharon Katzir-Katchalsky conference on structural aspects of recognition and assembly in biological macromolecules. *Nof Ginossar, Israel.* 2:24–29.

Arnott, S., R. Chandrasekharan, D. L. Birdsall, A. G. W. Leslie, and R. L. Ratliff. 1980*b*. Left handed helices. *Nature (Lond.).* 283:743–745.

Berman, H. M., S. Neidle, and R. Stodola. 1978. Drug-nucleic acid interactions. Conformational flexibility at intercalation site. *Proc. Natl. Acad. Sci. U. S. A.* 75:828–832.

Bondi, A. 1964. van der Waals volumes and radii. *J. Phys. Chem.* 68:441–451.

Broyde, S., R. M. Wartell, S. D. Stellman, B. Hingerty, and R. Langridge. 1975. A and B conformations of DNA and RNA subunits. Potential energy calculations for dGpdC. *Biopolymers.* 14:1597–1673.

Broyde, S., R. M. Wartell, and S. D. Stellman. 1978. Minimum energy conformations of DNA dimeric subunits: potential energy calculation for dGpdC, dApdA, dCpdC, dGpdG and dTpdT. *Biopolymers.* 17:1485–1506.

Broyde, S., and B. Hingerty. 1979. The influence of terminal 3′, 5′ phosphates on conformations of dApdA. *Nucleic Acids Res.* 6:2165–2178.

Calascibetta, F. G., M. Dentini, P. De Santis, and S. Morosetti. 1975. Conformational analysis of polynucleotide chains. Double-strand structures. *Biopolymers.* 14:1667–1684.

Camerman, N., J. K. Fawcett, and A. Camerman. 1976. A molecular structure of a deoxyribose nucleotide, sodium thymidylyl (5′–3′) thymidylate (5′) hydrate (pTpT) and possible structural model for poly thymidylate. *J. Mol. Biol.* 101:601–620.

Chandrasekaran, R., S. Arnott, A. Banerjee, S. Campbell-Smith, A. G. W. Leslie, and L. Puigjaner. 1979. Some new polynucleotide structures and some new thoughts about old structures. *ACS (Am. Chem. Soc.) Symposium Ser.* In press.

Cheng, D., and R. Sarma. 1977. Intimate details of the conformational characteristics of deoxyribodinucleoside monophosphates in solution. *J. Am. Chem. Soc.* 99:7333–7348.

Cremer, D., and J. A. Pople. Molecular orbital theory of the electronic structure of organic compounds. XXIII. Pseudorotation in saturated five membered ring compounds. 1975. *J. Am. Chem. Soc.* 97:1358–1367.

Crick, F. H. C., and A. Klug. 1975. Kinky helix in B-DNA. *Nature (Lond.).* 255:530–533.

Crick, F. H. C., and J. D. Watson. 1954. The complementary structure of deoxyribonucleic acid (DNA) *Proc. R. Soc. (Lond.).* A223:80–96.

Davidon, W. C. 1959. Variable metric method for minimization. *AEC Res. and development report:* ANL-5990.

Drew, H., T. Takano, S. Tanaka, K. Itakura, and R. E. Dickerson. 1980. High-salt d(CpGpCpG), a left-handed Z′DNA double helix. *Nature (Lond.).* 286:567–573.

Fletcher, R., and M. J. D. Powell. 1963. A rapidly convergent descent method for minimization. *Computer J.* 6:163–168.

Fujii, S., and K. I. Tomita. 1976. Conformational analysis of polynucleotides. I. The favourable left handed helical model for the poly (8, 2′-5- cycloadenylic acid) with high anti conformation. *Nucleic Acids Res.* 8:1973–1984.

Govil, G., C. Fisk, F. B. Howard, and H. T. Miles. 1977. Structure of poly 8-bromoadenylic acid. Conformational studies by CpF energy calculations. *Nucleic Acids Res.* 9:2573–2592.

Jain, S. C., C. C. Tsai, and H. M. Sobell. 1977. Visualization of drug-nucleic acid interactions at atomic resolution. II. Structure of an ethidium/dinucleoside monophosphate crystalline complex. Ethidium 5 Iodocytidylyl (3′–5′) guanosine. *J. Mol. Biol.* 114:317–332.

Ketelaar, J. 1958. Chemical Constitution. Elsevier Scientific Publishing Co., Amsterdam. 91.

Kim, S. H., and J. L. Sussman. 1976. π-turn is a conformational pattern in RNA loops and bends. *Nature (Lond.).* 260:645–646.

Kister, A. E., and V. G. Dashevsky. 1976. Calculations of the conformation of mono- and di-, and regular polynucleotides. *Biopolymers.* 15:1009–1013.

Kondo, N. S., K. N. Fang, P. S. Miller, and P. O. P. Ts'o. 1972. Influence of furanose on the conformation of adenine dinucleoside monophosphates in solution. *Biochemistry.* 11:1991–2003.

Lakshminarayanan, A. V., and V. Sasisekharan. 1970. Stereochemistry of nucleic acids and polynucleotides. II. Allowed conformation of the monomer unit for different ribose puckerings. *Biochim. Biophys. Acta.* 204:49–59.

Lee, C. H., and I. Tinoco, Jr. 1977. Studies of the modified dinucleoside phosphate containing 1-N^6 ethanoadenosine and 2'-0 methyl cytidine by 360 MHz Nuclear magnetic resonance spectroscopy. Investigation of the solution conformations of dinucleoside. *Biochemistry*. 16:5403–5414.

Levitt, M. 1978. How many base pairs per turn loss DNA have in solution and in chromatin some theoretical calculations. *Proc. Natl. Acad. Sci. U. S. A.* 78:640–644.

Levitt, M., and A. Warshel, 1978. Extreme conformational flexibility of the furanose ring in DNA and RNA. *J. Am. Chem. Soc.* 100:2607–2613.

McGuire, R. F., F. A. Momany, and H. A. Scheraga. 1972. Energy parameters in polypeptides. V. An empirical hydrogen bond potential function based on molecular orbital calculations. *J. Phys. Chem.* 76:375–393.

Momany, F. A., R. F. McGuire, A. W. Burgess, and H. A. Scheraga. 1975. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* 79:2361–2381.

Newton, M. D. 1973. A model conformational study of nucleic acid phosphate ester bonds. The torsional potential of dimethyl phosphate mono anion. *J. Am. Chem. Soc.* 95:256–258.

Olson, W. K., and P. J. Flory. 1972. Spatial configuration of poly nucleotide chains II. Conformational energies and average dimensions of polyribonucleotides. *Biopolymers*. 11:25–36.

Perahia, D., B. Pullman, B. Vasilescu, R. Cornillon, and H. Broch. 1977. A molecular orbital investigation on the conformation of tRNA. *Biochim. Biophys. Acta*. 478:244–259.

Ponnuswamy, P. K., and P. Thiyagarajan. 1978. Solvent accessibilities in the dimeric subunits of RNA and DNA. *Biopolymers*. 17:2503–2518.

Ramachandran, G. N., and V. Sasisekharan. 1968. Conformation of polypeptides and proteins. *Adv. Protein Chem.* 23:283–438.

Renugopalakrishnan, V., A. V. Lakshminarayanan, and V. Sasisekharan. 1971. Stereochemistry of nucleic acids and polynucleotides III. Electronic charge distribution. *Biopolymers*. 10:1159–1167.

Rodley, G. A., R. A. Scobi, R. H. J. Bates, and R. M. Lewitt. 1976. A possible conformation for double stranded polynucleotides. *Proc. Natl. Acad. Sci. U. S. A.* 73:2959–2963.

Sasisekharan, V. 1973. Conformation of polynucleoside. *In* Conformation of Biological Molecules and Polymers. E. C. Bergmann, and B. Pullman, editors. Fifth Jeresalem Symposia on Quantum Chemistry and Biochemistry. 5:247–266.

Sasisekharan, V., and N. Pattabiraman. 1976. Double stranded polynucleotides: two typical alternative conformations for nucleic acids. *Curr. Sci.* 45:779–782.

Sasisekharan, V., N. Pattabiraman, and G. Guptha. 1978. Some implications of an alternative structure for DNA. *Proc. Natl. Acad. Sci. U. S. A.* 75:4092–4096.

Scott, C., and H. A. Scheraga. 1965. Method for calculating internal rotation barriers. *J. Chem. Phys.* 42:2209–2215.

Sobell, H. M., C. C. Tsai, S. C. Jain, and S. G. Gilbert. 1977. Visualization of drug nucleic acid interaction at atomic resolution III. Unifying structural concepts in understanding drug-DNA interactions and their broader implications in understanding protein DNA interactions. *J. Mol. Biol.* 114:333–365.

Sundaralingam, M. 1969. Stereochemistry of nucleic acids and their constituents. IV. Allowed and preferred conformations of nucleosides, nucleotides, mono-, di-, tri- tetraphosphates nucleic acids and polynucleotides. *Biopolymers*. 7:821–860.

Sundaralingam, M. 1973. The concept of a conformationally 'rigid' nucleotide and its significance in polynucleotide conformational analysis. *In* Conformation of Biological Molecules and Polymers. B. Pullman, and E. Bermann, editors. Academic Press, Inc. New York. 5:417–456.

Sussman, J. L., and E. N. Trifonov. 1978. Possibility of non-kinked packing of DNA in chromatin. *Proc. Natl. Acad. Sci. U. S. A.* 75:103–107.

Tewari, R., R. K. Nanda, and G. Govil. 1974. Spatial configuration of single stranded polynucleotides: calculation of average dimensions and NMR coupling constants. *Biopolymers*. 13:2015–2035.

Thiyagarajan, P., and P. K. Ponnuswamy. 1978*a*. Conformational characteristics of dApdA, dApdT, dTpdA and dTpdT from energy minimization studies. *Biopolymers*. 17:533–553.

Thiyagarajan, P., and P. K. Ponnuswamy. 1978*b*. Conformational characteristics of the RNA subunit ApA from energy minimization studies. *Biopolymers*. 17:2143–2158.

Thiyagarajan, P., and P. K. Ponnuswamy. 1979. Conformational characteristics of the trinucleoside diphosphate d-(ApApA) from energy minimization studies. *Biopolymers*. 18:789–804.

Tsai, C. C., S. C. Jain, and H. M. Sobell. 1977. Visualization of drug-nucleic acid interactions at atomic resolution I. Structure on Ethidium/dinucleoside monophosphate crystalline complex Ethidium 5-iodouridylyl (3'-5') adenosine. *J. Mol. Biol.* 114:301–310.

Ts'o, P. O. P., N. S. Kondo, M. P. Schweizer, and D. P. Hollis. 1969. Studies of the conformation and interaction in dinucleoside mono- and diphosphates by proton magnetic resonance. *Biochemistry.* 8:990–1001.

Viswamitra, E. A., O. Kennard, P. G. Jones, G. M. Sheldrick, S. Salsibury, L. Faluellow and Z. Shalked. 1978. DNA double helical fragment at atomic resolution. *Nature (Lond.).* 273:687–689.

Wang, A. H. J., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich. 1979. Molecular structure of a left handed double helical DNA fragment at atomic resolution. *Nature (Lond.).* 282:680–686.

Westhof, E., and M. Sundaralingam. 1980. X-ray structure of a cytidylyl-3',5'-adenosine-proflavine complex. A self paired parallel-chain double helical dimer with an intercalated acridine dye. *Proc. Natl. Acad. Sci. U. S. A.* 77:1852–1856.

Wood, D., F. Hruska, and K. Ogilvie. 1974. Proton magnetic resonance studies of 2'-deoxythymidylyl (3'-5')-2'-deoxythymidine in aqueous solution. *Can. J. Chem.* 52:3353–3356.

Yathindra, N., and M. Sundaralingam. 1974. Backbone conformations in secondary and tertiary structural units of nucleic acids. Constraint in the phosphodiester conformations. *Proc. Natl. Acad. Sci. U. S. A.* 71:3325–3328.

Yathindra, N., and M. Sundaralingam. 1975. Backbone conformation in deoxyribonucleic acids. Conformational role of the 2'-hydroxyl group on the internucleotide phosphodiester conformations. *Biopolymers.* 11:25–36.

Zhurkin, V. B., Y. U. P. Lysov, and V. I. Ivanov. 1978. Different families of double stranded conformations of DNA as revealed by computer calculations. *Biopolymers.* 17:377–412.