CONFORMATIONAL ANALYSIS OF NUCLEIC ACIDS
Extended Hückel study on the D-ribose (C_3, -exo)-phosphate unit

Dane VASILESCU, Francine CAMOUS and Jean-Noël LESPINASSE
Laboratoire de Biophysique, U.E.R.D.M. Université de Nice, Parc Valrose, 06 Nice, France

Received 13 January 1972

1. Introduction

For some years many works have been dealing with conformational analysis of biomolecules. Most of the results already published have concerned the polypeptides, especially the dipeptides. As for the polynucleotides, the only known quantum calculations deal with the nucleosides (Jordan and Pullman [1], Berthod and Pullman [2]) and a study of the D-ribose (C_3,-endo)-phosphate unit by Govil and Saran [3].

The work presented here is a study by Extended Hückel Theory of the D-ribose (C_3, -exo)-phosphate unit. The allowed conformations of the unit are found to be in agreement with those obtained from empiric calculations and also with observed structures.

2. Experimental

We have studied the unit shown in fig. 1. The D-ribose exists in 4 conformations, investigated by Sundaralingam [4], Sundaralingam and Jensen [5]: C_3, -exo; C_3, -endo; C_2, -exo; and C_2, -endo. These 4 conformeric possibilities result from the puckering of the furanose ring which produces a displacement of the C_2, or C_3, atoms from the plane (C_1, $-O_1$, $-C_4$). In fact, the interconversions: C_3, -exo $\leftrightarrow$ C_2, -endo and C_3, -endo $\leftrightarrow$ C_2, -exo, are energetically favoured [4].

We have chosen first to work on the D-ribose in its C_3, -exo conformation in order to complete Govil and Saran's study [3]. However, it is important to note that the known experimental results for deoxyadenosine and thymidine (see [2] and [6]) show that the C_3, - exo conformation must be associated with a deoxyribose.

Conformational analysis of D-ribose (C_3, -exo) phosphate unit needs energy E to be minimized as a function of 3 conformational angles $\theta_1, \theta_2, \theta_3$. The denomination of these angles is the one given by Sasisekharan et al. [7]. The zero positions and the senses of the rotational angles $\theta_1, \theta_2, \theta_3$ are those used by Lakshminarayanan and Sasisekharan [8].

The nitrogen atom linking the sugar and the base was not taken into account.

We made our calculations on a C.D.C. 3600 computer, using a modified version of the initial program of R. Hoffman for the Extended Hückel Theory (E.H.T.) kindly lent by F. Peradejordi and J.F. Cavelier [9]. In this version, the program can handle up to 110 atomic orbitals. Moreover, we fitted this version of the program to an I.B.M. 7040 computer, with a limitation of 55 atomic orbitals.

The atomic coordinates were calculated by another program, using the coordinate tables published in [8] for the (180°, 180°, 180°) conformation of the ribose phosphate unit.

The adjustable constant K, which appears in the calculation of the off-diagonal terms of the Hamiltonian matrix:

$$H_{ij} = \frac{1}{2} K (H_{ii} + H_{jj}) S_{ij} ,$$

was assigned the usual value 1.75 [10].

The atomic orbitals of the basis set are of the Slater type. The orbital exponents and valence state ionization potentials (diagonal elements of Hamiltonian matrix) are reported in table 1.

At last, as the ribose phosphate unit under study is

Table 1

Atom	Orbital	Orbital	VSIP (eV)	Reference
H	1 s	1.000	– 13.60	[10]
C	2 s	1.625	– 21.40	[10]
	2 p	1.625	– 11.40	
O	2 s	2.275	– 35.30	[11]
	2 p	2.275	– 17.76	
P	3 s	1.600	– 20.20	[12]
	3 p	1.600	– 12.49	

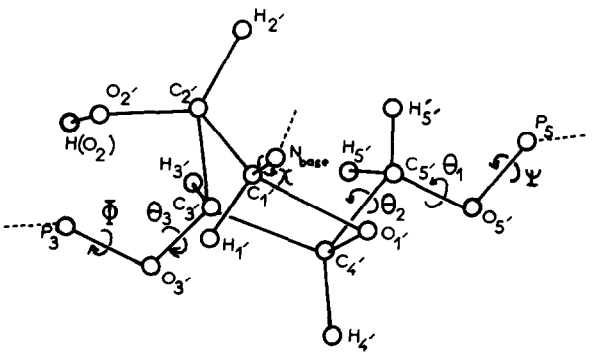

Fig. 1. The D-ribose-phosphate unit.

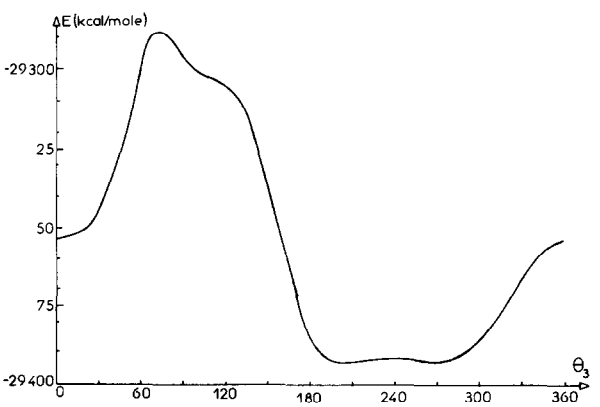

Fig. 2. The energy of the unit as a function of θ_3, the rotational parameters θ_1 and θ_2 being assigned the value 180°.

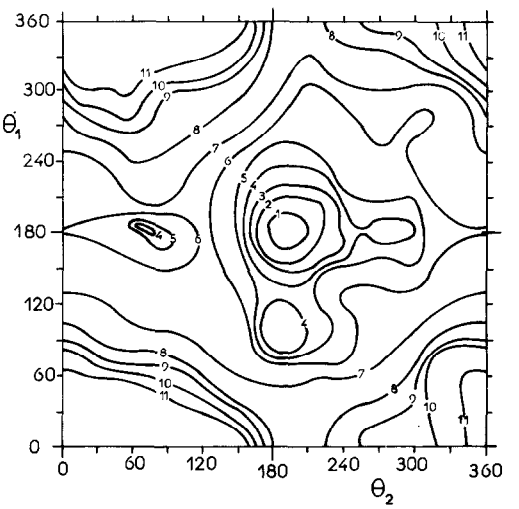

Fig. 3. Variation of the energy with θ_1 and θ_2, the value assigned to θ_3 being 260°. The isoenergetic curves, numbered from 1 to 11, correspond respectively with the following energy values (kcal/mole): – 29,393; – 29,392; – 29,391; – 29,390; – 29,289; – 29,387; –29,380; – 29, 350; 29,300; – 29,250; – 29,100. The curves corresponding to an energy value higher than – 29,100 kcal/mole are not shown.

not a molecule, there is an alternative between the neutral structure of the free radical and a "quasi-molecule" obtained when all non-bonded orbitals of the P_3, P_5 and C_1 atoms are filled with electrons. If we do not take N_{base} into account, the "quasi-molecule" structure is obtained by adding 2 electrons to the 5 valence electrons of P_3 and P_5 and 1 to the 4 valence electrons of C_1. We choose this "quasi-molecule" structure to make our calculations, this choice being justified by the results obtained [3, 13].

3. Results

Assigning different values to θ_1 and θ_2, the energy of the unit has been calculated as a function of θ_3 and the corresponding curves have been drawn. All these curves almost assume the same shape as the one given in fig. 2. It can be seen from fig. 2. that the minima in energy occur when θ_3 is in the range of 200–280°. Clearly, for such values of θ_3, the most probable conformations of the unit under study can be found by drawing the isoenergetic maps $E(\theta_1, \theta_2)$. Several maps were drawn with θ_3 in the range 200–

Table 2

Allowed region of minimum energy	Range of θ_1	Range of θ_2	Range of θ_3	Method used for calculations	Reference
I (our notation)	170–190°	180–210°	$\theta_3 = 260°$	E.H.T.	This work
α (our notation)	150–210°	170–220°	210–260°	Hardsphere model	[8]
2 (notation of [6])	140–230°	130–220°	147–292°	Structures observed	[6]
I_{Pot} (our notation)	centered around 180°	centered around 180°	180–280°	Potential energy calculations	[14]
II (our notation)	centered around 180°	150–300°	$\theta_3 = 260°$	E.H.T.	This work
β (our notation)	150–210°	260–310°	210–260°	Hardsphere model	[8]
3 (notation of [6])	120–300°	250–300°	147–292°	Structures observed	[6]
II_{Pot} (our notation)	centered around 180°	centered around 300°	180–280°	Potential energy calculations	[14]
III (our notation)	80–120°	170–220°	$\theta_3 = 260°$	E.H.T.	This work
5 (notation of [6])	60–110°	130–220°	147–292°	Structures observed	[6]
III_{Pot} (our notation)	60–110°	150–220°	180–280°	Potential energy calculations	[14]
IV (our notation)	centered around 180°	centered around 60°	$\theta_3 = 260°$	E.H.T.	This work
1 (notation of [6])	130–230°	40–100°	147–292°	Structures observed	[6]
IV_{Pot} (our notation)	180±60°	centered around 60°	180–280°	Potential energy calculations	[14]

280°. Their general appearance is given in fig. 3 by the map $E(\theta_1, \theta_2, \theta_3 = 260°)$. Taking this map as a basis of discussion, our results may be summed up in fig. 4.

On this map, 4 types of allowed regions are delimited by the isoenergetic lines: 1) As shown by fig. 4 the central region I delimited by the – 29,393 kcal/mole isoenergetic line lies in the allowed range α obtained by the use of the hard sphere model and in area 2 given by Sundaralingam [6]. (The Watson–Crick DNA model structure lies in area 2). The $E(\theta_1, \theta_2)$ map drawn in [14] from potential functions calculations also shows this central zone. 2) Areas II, III and IV, delimited by the – 29,390 kcal/mole isoenergetic line, spread out the allowed horizontal and vertical strips [$\theta_1 \in (150°, 210°)$ and $\theta_2 \in (170°, 220°)$] given by the hard sphere model. Moreover, they almost lie in areas 1, 2, 3 and 5 of [6] and they appear on the $E(\theta_1, \theta_2)$ map of [14].

We compare in table 2 the results of our extended Hückel calculations on the D-ribose-(C_3, –exo)-phosphate with those obtained for the same unit by hard sphere and potential energy calculations, and also with the observed structures given by Sundaralingam [6]. It is obvious that our results are in agreement with the others. However, E.H.T. predicts that area IV, forbidden by the hard sphere model, is a stable region. This is important to note because most of the observed nucleotides and polynucleotides lie in this region (see [6]). Furthermore, the potential energy calculations also give area IV as a stable one.

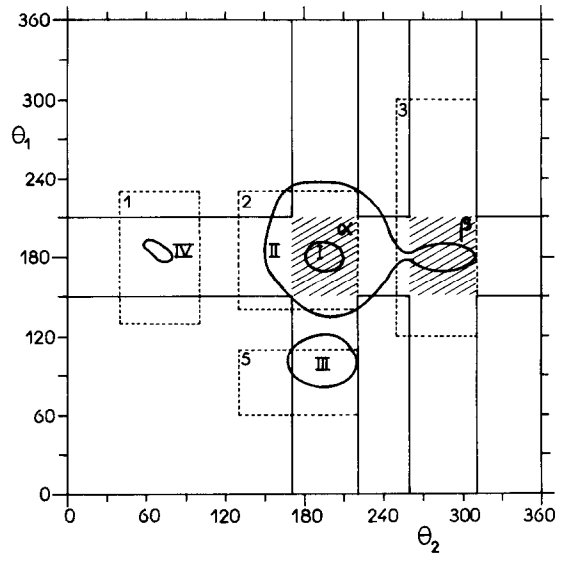

Fig. 4. Comparison of the results given by different methods. The hard-sphere model gives 2 allowed regions, α and β, which are hatched on the figure. Regions I, II, III, IV, that we obtain by Extended Hückel calculations ($\theta_3 = 260°$), are enclosed with a full line. Regions 1, 2, 3 and 5 of [6] are enclosed with a dotted line.

Finally, a more careful study on regions I, II, III and IV, when θ_3 is varied in the range 200–280° at 10° intervals, shows that the most probable conformations of the D-ribose phosphate unit are:
($\theta_1 = 180°, \theta_2 = 190°, \theta_3 = 270°$) $E = -29{,}393.74$ kcal/mole
($\theta_1 = 180°, \theta_2 = 190°, \theta_3 = 210°$) $E = -29{,}393.66$ kcal/mole
It may be pointed out that the energy values of these 2 conformations are just slightly lower than those appearing in fig. 3.

Final remark

In a recent paper, published in the final stage of our work, Saran and Govil [15] made Extended Hückel calculations on a phosphodiester unit. Our results are in agreement with theirs as concerns the $E(\theta_1, \theta_2)$ map.

Acknowledgements

This work was supported in part by grants of D.R.M.E., R.C.P. 195 and University of Nice. The authors express their gratitude to Prof. R. Daudel, Doctor F. Peradejordi and J.F. Cavelier for their help concerning the utilization of E.H.T. programs.

References

[1] F. Jordan and B. Pullman, Theor. Chim. Acta 9 (1968) 242.
[2] H. Berthod and B. Pullman, Biochim. Biophys. Acta 246 (1971) 359.
[3] G. Govil and A. Saran, J. Theor. Biol. 30 (1971) 621.
[4] M. Sundaralingam, J. Am. Chem. Soc. 87 (1965) 599.
[5] M. Sundaralingam and L.H. Jensen, J. Mol. Biol. 13 (1965) 930.
[6] M. Sundaralingam, Biopolymers 7 (1970) 821.
[7] V. Sasisekharan, A.V. Lakshminarayanan and G.N. Ramachandran, in: Conformation of Biopolymers, ed. G.N. Ramachandran (Academic Press, New York, 1967) Vol. 2 p. 641.
[8] A.V. Lakshminarayanan and V. Sasisekharan, Biochem. Biophys. Acta 204 (1970) 49.
[9] F. Peradejordi and J.F. Cavelier, personal communication.
[10] R. Hoffmann, J. Chem. Phys. 39 (1963) 1397.
[11] H.O. Pritchard and H.A. Skinner, Chem. Rev. 55 (1955) 745.
[12] J. Hinze and H.H. Jaffe, J. Am. Chem. Soc. 84 (1962) 540.
[13] G. Govil, J. Chem. Soc. A (1970) 2464.
[14] A.V. Lakshminarayanan and V. Sasisekharan, Biopolymers 8 (1969) 489.
[15] A. Saran and G. Govil, J. Theor. Biol. 33 (1971) 407.