Theoret. chim. Acta (Berl.) 9, 242—252 (1968)

Molecular Orbital Calculations on the Preferred Conformation of Nucleosides*

FRANK JORDAN** and BERNARD PULLMAN

Université de Paris, Institut de Biologie Physico Chimique, Laboratoire de Biochimie Quantique, associé au CNRS, 13, rue Pierre Curie, Paris V

Received November 2, 1967

The extended Hückel theory has been applied to the study of the conformation of the nucleosides of the purine and pyrimidine bases of the nucleic acids. Although the evaluation of the total energy as a function of the rotation angle presents in all cases two minima, the calculations predict a preferred *anti* conformation for uridine, cytidine and adenosine and a preferred *syn* conformation for guanosine. These predictions appear to be in agreement with the available experimental data.

Mit der erweiterten Hückel-Theorie wurde die Konformation der Purin- und Pyrimidin-Nucleoside untersucht. Die Gesamtenergie in Abhängigkeit vom Verdrehungswinkel zeigt zwei schwache Minima; für Uridin, Cytidin und Adenosin ist die *anti*-Konformation bevorzugt, für Guanosin die *syn*-Konformation. Die Berechnungen sind in Übereinstimmung mit den verfügbaren experimentellen Daten.

La méthode de Hückel étendue a été appliquée à l'étude de la conformation des nucléosides des bases puriques et pyrimidiques des acides nucléiques. Bien que l'énergie moléculaire totale présente dans tous les cas deux minima en fonction de l'angle de rotation, la théorie prévoit une conformation préférentielle *anti* pour l'uridine, la cytidine et l'adénosine et une conformation préférentielle *syn* pour la guanosine. Ces prédictions paraissent en accord avec les données expérimentales disponibles.

Introduction

An examination of molecular models of pyrimidine and purine nucleosides led DONOHUE and TRUEBLOOD [1] to suggest that the rotation of the base with respect to the ribose creates in these molecules two regions of conformational stability. They defined a rotation angle ϕ_{CN} to classify the relative position of the base to ribose in the course of such rotations around the glycosidic C–N bonds "as the angle formed by the trace of the plane of the base with the projection of the C–O bond of the furanose ring when viewed along the C–N bond. This angle will be taken as zero when the furanose ring oxygen atom is antiplanar to C(2) of the pyrimidine or purine ring, and positive angles will be taken as those measured in a clockwise direction when viewing from C to N″.

The two preferred regions were centered around $\phi_{CN} \approx -30°$ (conformation *anti*) and $\phi_{CN} \approx +150°$ (conformation *syn*).

* This work was supported by grant no. 67-00-532 of the Délégation Générale à la Recherche Scientifique et Technique (Comité de Biologie Moléculaire).

** NATO Postdoctoral Fellow in Science 1967—1968.

A review of the crystallographic determinations of the conformations of mononucleosides reveals, however, that there is only one known *syn* conformation, and that all nucleosides but guanosine prefer the *anti* ϕ_{CN} range [2].

Because of the importance of the conformation of nucleosides in building useful DNA and RNA models, it seemed of interest to investigate more precisely the nature of the rotational barrier around the glycosidic linkage. Recently there have been two attempts in this direction. HASCHEMEYER and RICH studied the problem in a number of structurally defined nucleosides and nucleotides [3] from the point of view of close Van der Waals contacts. TINOCO et al. [4] announce having performed calculations on the intermolecular (Van der Waals-London) interactions between the two rings in uridine, cytidine, adenosine and guanosine, and having found, in agreement with experiment, that while the first three nucleosides prefer an *anti* conformation, guanosine should be most stable in its *syn* range. (This last study employed a "$C_{2'}$ endo" ribose uniformly in all four nucleosides which implies that the best four atom plane of the ribose skeleton has the $C_{2'}$ atom puckered out of the plane [5].)

The present paper undertakes a more thorough and somewhat different quantum chemical approach to estimate the rotational barriers in the four nucleosides through the use of the extended Hückel method, treating each molecule as an unique entity and evaluating its total electronic energy in different conformations. For this purpose four experimentally known geometries were taken to construct the four nucleosides and the bases were then rotated around the ribose portions of the molecules through 360 degrees, at 60 degrees intervals in a counterclock-wise manner, using the C–N glycosidic linkage as the axis of rotation. In what follows the 0° rotations will thus always to the ϕ_{CN} found in the crystal structure of a molecule. The procedure enables then to pick out the preferred conformations and to estimate the rotational barrier between them.

Method of Calculation

The extended Hückel theory [6] was chosen for the calculations because of its simplicity and since it has been repeatedly and successfully applied to a variety of conformational problems [6, 7, 8, 9, 10, 11, 12].

This theory builds molecular orbitals as linear combinations of atomic basis orbitals:

$$\Psi_i = \sum_j C_{ij}\,\phi_j\,.$$

Minimization of the total energy by the variational principle leads to the set of secular equations:

$$\sum_{i=1}^{n} [H_{ij} - ES_{ij}]\,C_{ij} = 0 \qquad j = 1, 2, \ldots n\,.$$

The basis set was built up of 1*s* orbitals of H (Slater orbital exponent 1.00), of one 2*s* and three 2*p* orbitals for each C, N and O atoms present (with 1.625, 1.950, 2.275 as orbital exponents, respectively).

The diagonal matrix elements (coulombic integrals of the type $H_{ii} = \int \phi_i\, H\, \phi_i\, d\tau$) were taken as the valence state ionization potentials of the orbitals: −13.60 eV for H 1*s*; −11.40 eV and −21.40 eV for the C2*p* and C2*s* orbitals, respectively [6]; −13.40 eV for N2*p*, −26.00 eV for N2*s* [13], −17.76 eV for O2*p* and −35.30 eV for O2*s* [14] orbitals.

The off-diagonal matrix elements (resonance integrals of the type $H_{ij} = \int \phi_i H \phi_j d\tau$) were approximated by use use of the Wolfsberg-Helmholtz formula [15]

$$H_{ij} = 0.5 \text{ K } (H_{ii} + H_{jj}) S_{ij}$$

where H_{ii}'s are the orbital valence state ionization potentials and S_{ij} is the overlap integral. A value of $K = 1.75$ was used in this study as suggested by HOFFMANN [6]. Because of the large size of the molecules under investigation no iterative process correcting the diagonal elements for the redistribution of charges was carried out. It is believed nevertheless that the shape of the potential barrier would not change to any large extent with such iterations [7]; in other words, the minima in energy would most likely occur at the same angles of rotation. This assumption remains however to be confirmed for such large heterocyclic molecules by more refined calculations.

Choice of Geometries

It should be emphasized that the results, although quite significant, must be viewed in the light of the geometries that were chosen for the calculations.

First of all, it was assumed that the relationship of the C–N glycosidic bond to the plane of the base does not change during the rotation of the bases around the ribose. Although this may or may not be strictly the case, it must be recognized that the energy for bending this C–N bond in some way during rotation may on the other hand be overcompensated by a reduction in the repulsion energy. This possibility, although a real one, was not investigated since it would add an extra dimension to an already very large problem.

Furthermore, the exact geometry is known experimentally only for one of the four nucleosides, cytidine [16], thus, in constructing those of the others, certain assumptions had to be made.

The geometry of uridine was derived from a detailed work on calcium thymidylate by TRUEBLOOD, HORN and LUZZATI [17]. The methyl group of thymine was replaced by a hydrogen atom and a ribose was constructed from the deoxyribose of calcium thymidylate vectorially.

The geometry of adenosine was taken from the study of adenosine-5-phosphate [18], which meant a simple conversion of the nucleotide to a nucleoside.

The geometry of guanosine was derived from that of the only known guanosine derivative, deoxyguanosine. HASCHEMEYER and SOBELL determined the X-ray structure of the 5-bromo-deoxycytidine-deoxyguanosine complex [2]. To convert deoxyguanosine to guanosine, the deoxyribose of the former had to be converted to the ribose of the latter and all the hydrogen coordinates had to be constructed.

The above list immediately reveals the possibility for errors in at least the structures of uridine and guanosine where we had to assume that the puckering of the ribose as well as the rotation angle φ_{CN} would be similar in the desired nucleoside as in the molecules from which they were derived. The recent review of structural properties of nucleosides and nucleotides [4] helped to justify the problem of uridine, since, whereas calcium thymidylate has the C_3, carbon puckered in the ribose, uridine derivatives were found to have either C_2, or O_3, atoms out of the plane in the various derivatives studied.

The deviation of C_1, from the plane of the base is almost negligible in all uridine structures as well as in calcium thymidylate. The torsion angle ϕ_{CN} in calcium

Table. *Total Orbital Energy at Various Degrees of Rotation in Nucleosides*

Rotation (Degrees)	E_{total} (kcal/mole)			
	Cytidine $\phi_{CN} = -24°$ at 0° rotation	Adenosine $\varphi_{CN} = -18°$ at 0° rotation	Uridine $\phi_{CN} = -43°$ at 0° rotation	Guanosine $\varphi_{CN} = +138°$ at 0° rotation
0	−43532.681	−46122.656	−44345.897	−49498.588
60	−43522.151	−46117.172	−44342.011[a]	−49492.034
120	−43523.590	−46121.105	−44342.532	−49497.284
180	−43508.875	−46105.370	−44335.881	−49497.662
240	−43436.417	−46080.072	−44329.306	−49338.243
300	−43464.723	−46108.884	−44339.892	−49494.787

[a] (50 °)

thymidylate is −43° which is close to the average of all φ_{CN}'s found in uridine derivatives.

Since deoxyguanosine is the only known guanosine derivative, all structural properties of the latter were taken identical to those of the former, with C_2, puckered out of the plane of the other four atoms in the ribose skeleton and a torsion angle $\varphi_{CN} = +138°$.

Results and Discussion

The Table shows the energy of the various conformations which in the extended Hückel theory is a simple sum of the orbital energies. As already stated zero

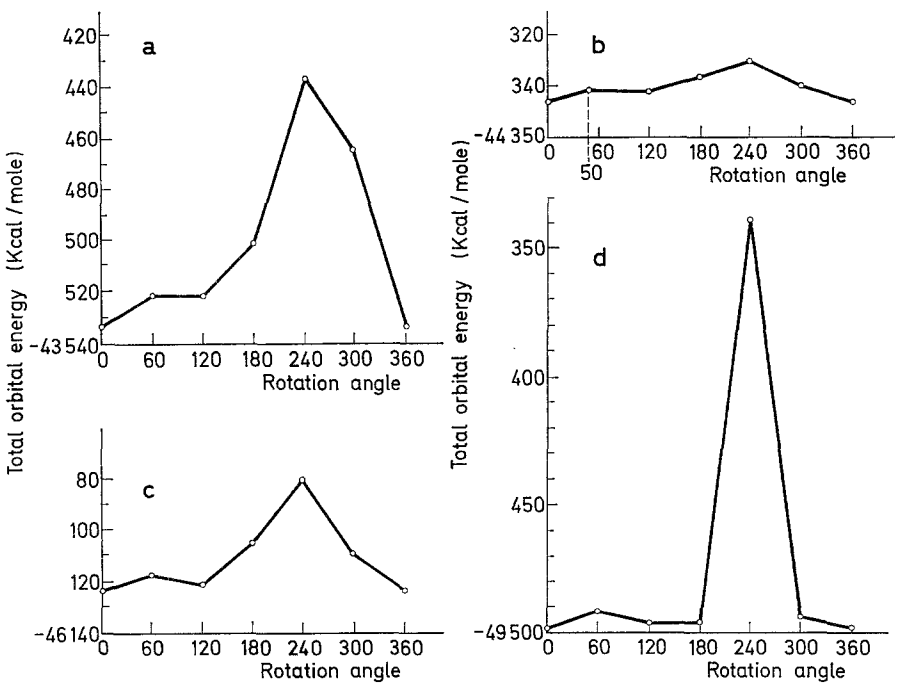


Fig. 1 a—d. Total orbital energy as a function of the rotational angle. a cytidine, b uridine, c adenosine, d guanosine

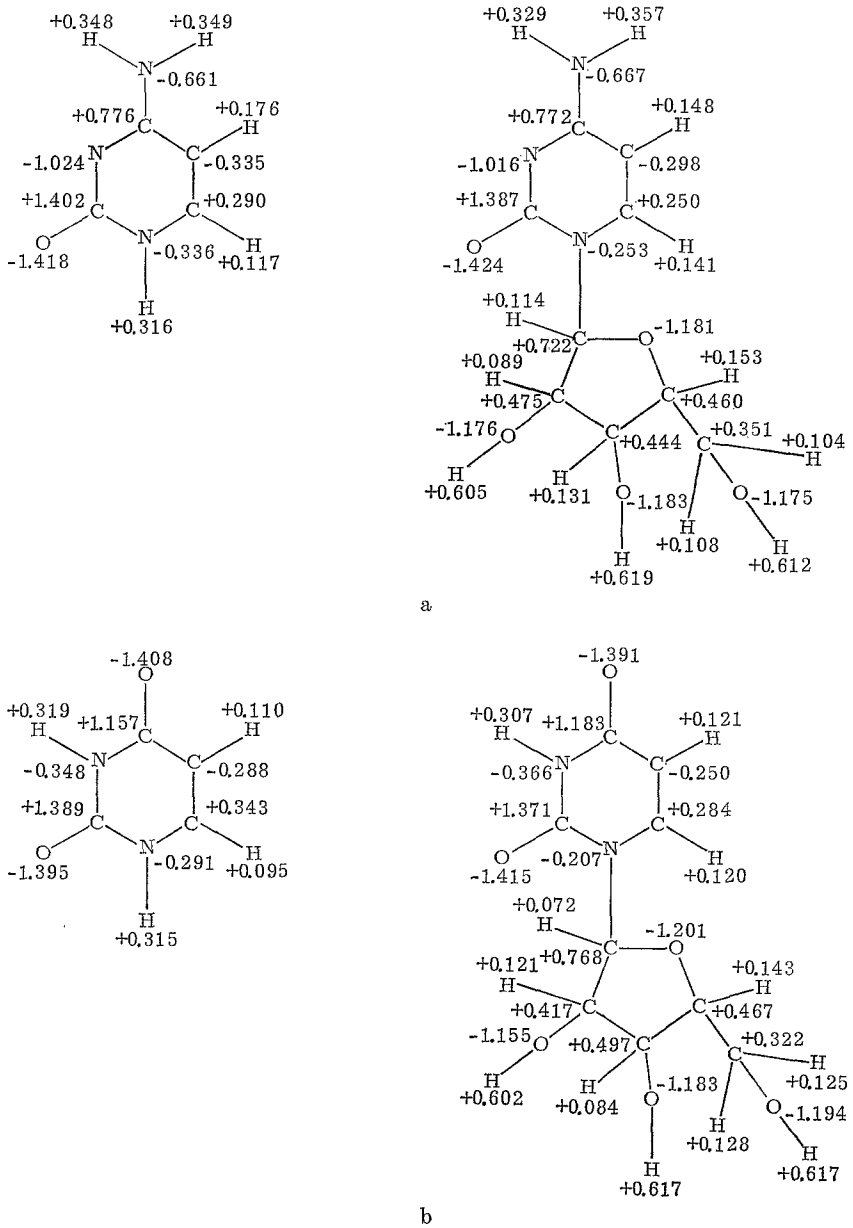


Fig. 2 a—d. Net atomic charges. a cytosine and cytidine, b uracil and uridine, c adenine and adenosine, d guanine and guanosine

rotation angle always refers to the experimental φ_{CN} and rotation was done in a counterclockwise manner. Fig. 1 gives a more pictorial representation of the results.

Perhaps the most striking aspect of the results is that the preferred conformations are correctly predicted in all four cases. Although they do not represent

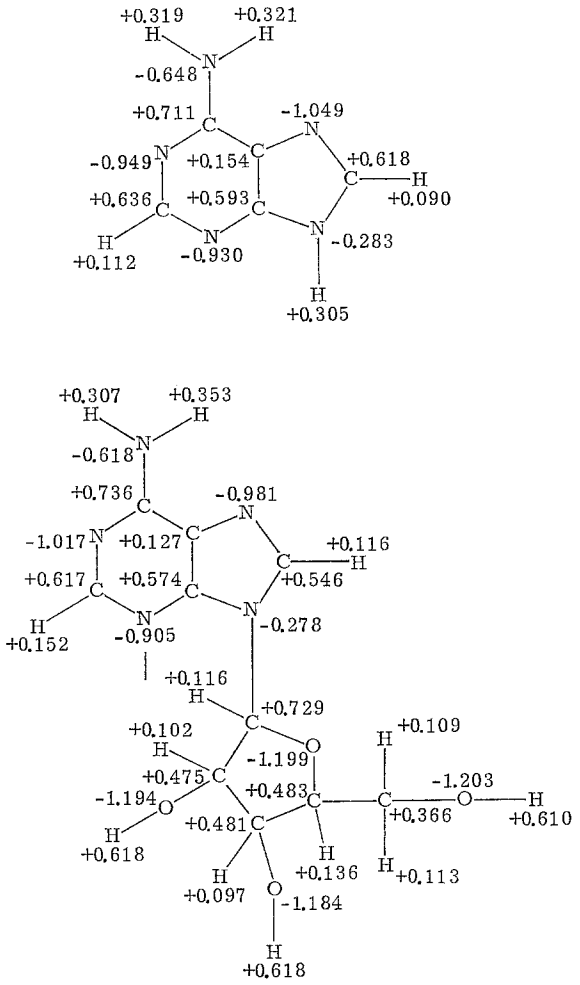


Fig. 2c

absolute energy minima, they do indicate minima in rotation angle for a specific relationship of C_1,–N to the plane of the base.

Thus, the calculations correctly predict cytidine, uridine and adenosine to be in the *anti* range, while guanosine is predicted to be in the *syn* one. What is perhaps still more significant though, is the fact that while the pyrimidine nucleosides show clear preference for the *anti* range, the energy difference in the purine nucleosides between most favorable *anti* and *syn* conformations is only of the order of 1 kcal/mole or less. This indicates a much greater likelihood for finding certain forms of the purine nucleosides in either *syn* or *anti* conformations. As to DONOHUE and TRUEBLOOD'S suggestions [2] concerning the position of these two ranges, they are shown by Fig. 1 to be obeyed very well indeed by the purine nucleosides.

Guanosine which has a $\varphi_{CN} = +138°$ at the 0° rotation at which the principal minimum occurs shows a very clear second minimum (only $1\frac{1}{2}$ kcal/mole less

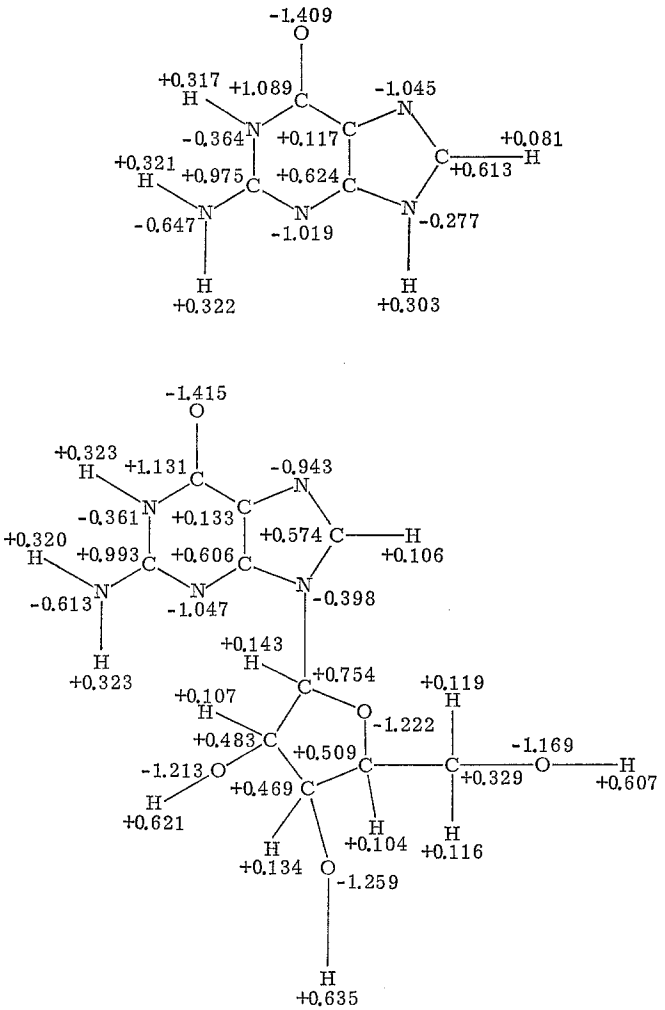
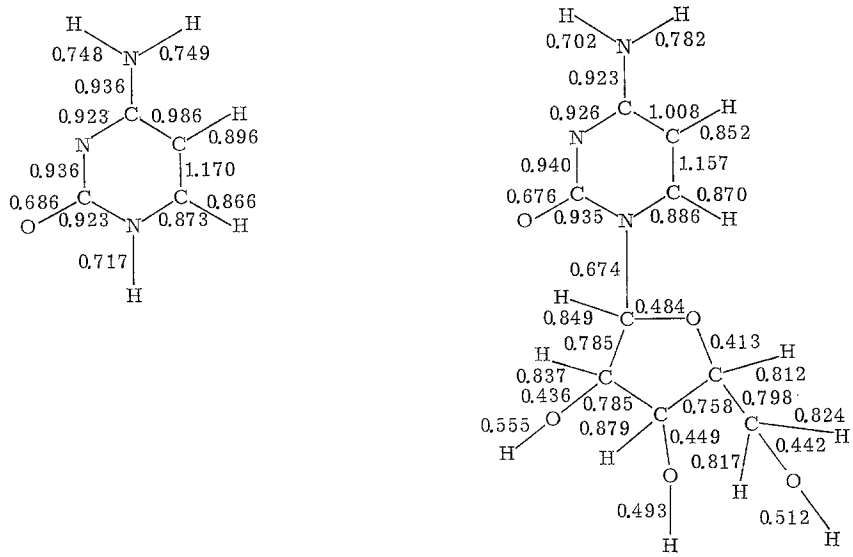


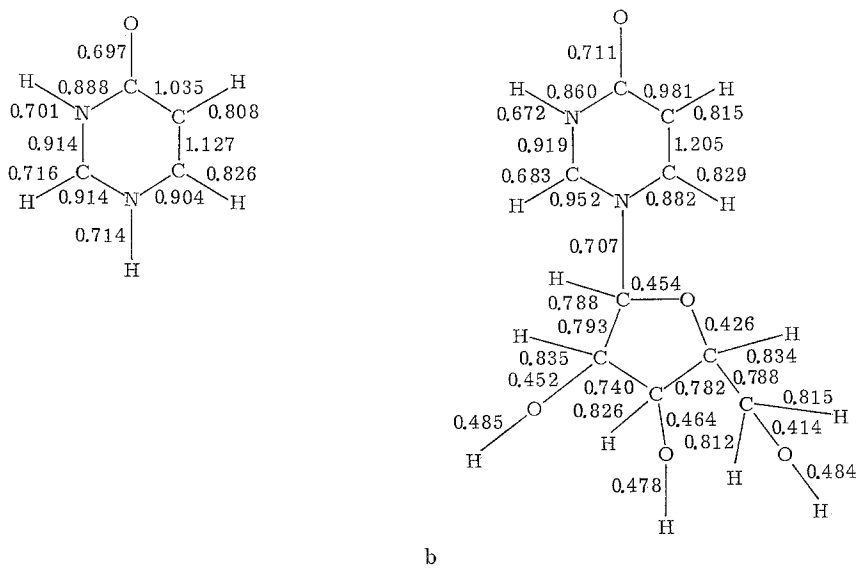
Fig. 2d

stable) at around 180° rotation which is almost equivalent to φ_{CN} of $-42°$ or the *anti* range. Adenosine similarly also has a favorable second minimum at 120° to 150° rotation which is equivalent to a φ_{CN} of $+100°$ to 130° or the *syn* range. That adenosine has a favorable *syn* range is supported by the ease of formation of a cyclic adenosine nucleoside in the *syn* conformation [19]. On the other hand, although there is a slight minimum in the *syn* range of pyrimidine nucleosides as well, the energy of this minimum is much less favorable than for purine nucleosides.

The quantitative aspects of the barriers are more difficult to ascertain. Nevertheless although the differences in the energies of the conformations amount to a very small fraction of the total orbital energies we consider these differences to be significant since bond lengths and bond angles are held constant for each molecule



a



b

Fig. 3 a—d. Total bond overlap populations. a cytosine and cytidine, b uracil and uridine, c adenine and adenosine, d guanine and guanosine

during the rotation and it is only the relative position of the base to the ribose that is changed. The barrier heights are probably somewhat exaggerated in all cases just as they were in the case of simple hydrocarbons [6].

Although, as already mentioned, there was no effort made as of now to optimize the electronic charges for these large molecules with a profusion of heteroatoms, the net charges (net atomic populations) and bond overlap populations are

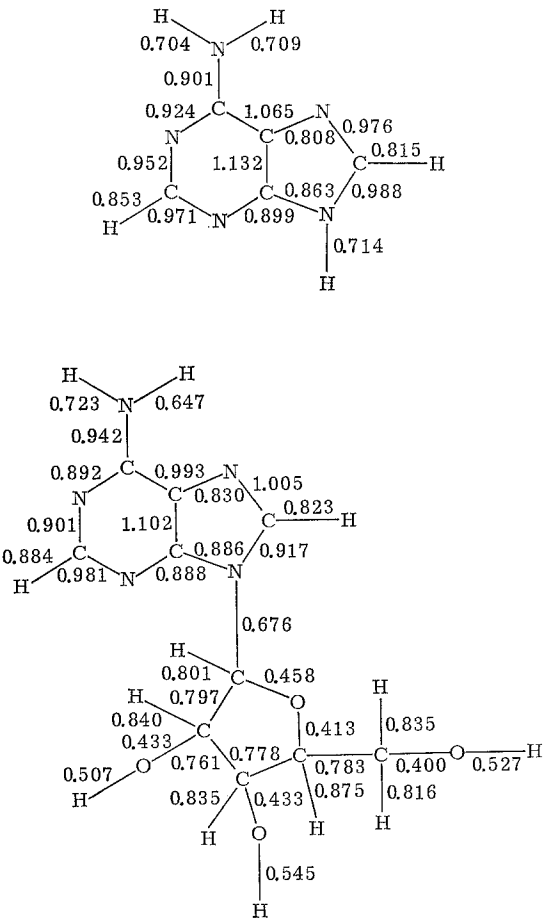


Fig. 3 c

still instructive to look at. For this purpose, calculations on the planar bases (geometry as in Ref. [20]) were also performed with parameters identical to the ones presented before. Fig. 2 shows a comparison of the charges in the planar bases with the charges in the 0° rotated nucleosides. Fig. 3 gives the comparison of the bond overlap populations for the same set of molecules. As can be seen in Fig. 2, the perturbation of the charges in the base caused by the attachment of the ribose spreads out throughout the periphery of the base. The slight changes in charges and in overlap populations between the bases and the nucleosides can for the most part be attributed to the differences in geometry of the substituted and unsubstituted bases.

The problem of the significance of the absolute values of the net charges, some of which appear undoubtedly overestimated is a more complicated one and must be viewed in relation to other types of calculations carried out on the same molecules and to such experimental data as e.g. dipole moments. This investigation wil be carried out separately.

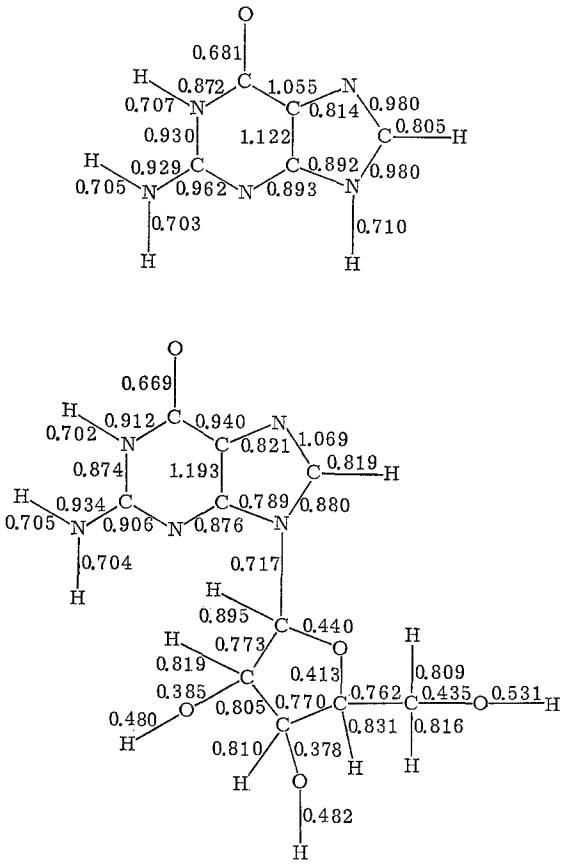


Fig. 3 d

Admittedly, although the results presented here can only be considered as tentative, the satisfactory agreement between the essential experimental and theoretical findings indicates the general applicability of the approach to problems of conformational stability of biomolecules.

References

1. Donohue, J., and K. N. Trueblood: J. molecular Biol. **2**, 363 (1960).
2. Haschemeyer, A. E. V., and H. M. Sobell: Acta Cryst. **19**, 125 (1965).
3. —, and A. Rich: J. molecular Biol. **27**, 369 (1967).
4. Tinoco, I., R. C. Davis, and S. R. Jaskunas: In: Molecular associations in biology, B. Pullman, Ed. New York: Academic Press 1968.
5. Sundaralingam, M.: J. Amer. Chem. Soc. **87**, 599 (1965).
6. Hoffmann, R.: J. Chem. Physics **39**, 1397 (1963).
7. Jordan, F.: Ph. D. Thesis, University of Pennsylvania 1967.
8. Peyerimhoff, S. D., R. J. Buenker, and L. C. Allen: J. Chem. Physics **45**, 734 (1966).
9. Buenker, R. J., S. D. Peyerimhoff, L. C. Allen, and J. L. Whitten: J. Chem. Physics **45**, 2835 (1966).
10. Allen, L. C., and J. D. Russell: J. Chem. Physics **46**, 1029 (1967).
11. Kier, L. B.: Mol. Pharmacol. **3**, 487 (1967).

12. GIORDANO, W., J. R. HAMANN, J. J. HAZKINS, and J. J. KAUFMAN: Mol. Pharmacol. **3**, 307 (1967).
13. HOFFMANN, R.: J. Chem. Physics **40**, 2745 (1964).
14. PRITCHARD, H. O., and H. A. SKINNER: Chem. Reviews **55**, 745 (1955).
15. WOLFSBERG, M., and L. HELMBOLTZ: J. Chem. Physics **20**, 837 (1952).
16. FURBERG, S., C. S. PETERSEN, and C. H. R. ROMMING: Acta Cryst. **18**, 313 (1965).
17. TRUEBLOOD, K. N., P. HORN, and V. LUZZATI: Acta Cryst. **14**, 965 (1961).
18. KRAUT, J., and L. H. JENSEN: Acta Cryst. **16**, 79 (1963).
19. ZUSSMAN, J.: Acta Cryst. **6**, 504 (1953).
20. BERTHOD, H., C. GIESSNER-PRETTRE und A. PULLMAN: Theoret. chim. Acta (Berl.) **5**, 53 (1966).

Professor BERNARD PULLMAN
Université, Laboratoire de Biochimie Quantique
13 rue Pierre Curie
Paris V, France