MOLECULAR GEOMETRIES — ACCURACY TEST OF THE PCILO METHOD

Zdeněk HAVLAS and Petr MALOŇ

Institute of Organic Chemistry and Biochemistry,
Czechoslovak Academy of Sciences, 166 10 Prague 6

Received May 24th, 1979

Applicability of the PCILO method for investigation of stationary points on the energy hypersurface has been tested with a set of 11 small organic molecules. From comparison of the molecular geometries, completely optimized with the use of the Payne algorithm, with experimental data and with the results obtained by the MINDO/3 and MNDO methods it follows that the PCILO method is suitable for semiquantitative estimate of molecular geometry (mean error in the bond length determination $5 . 10^{-3}$ nm, valence angles 3·4°). The Payne algorithm is little efficient, if the starting geometry is far from the stationary point and if the assumption of quadratic form of the hypersurface is not fulfilled.

Study of equilibrium molecular geometry of organic compounds represents one of the most important applications of quantum-chemical methods. Use of semiempirical procedures which, by various simplifications and introduction of empirical parameters, reach higher speed of calculations with maintained acceptable level of reliability of results make it possible to investigate relatively large systems interesting from the bioorganic chemistry viewpoint. In this respect the PCILO method seems very promising (Pullman and coworkers[1,2]) which was suggested, first of all, for conformation studies of organic molecules and was used successfully for calculations of fragments of peptide chain[3], components of nucleic acids[4], acetylcholine[5], valinomycin[6], *etc.* Perturbation treatment of energy calculation along with the concept of localized orbitals results in the PCILO method being faster than the analogous variation (CNDO) procedure by orders of magnitude.

So far most of the PCILO calculations were carried out by tabulation of molecular energy in dependence on a variable geometric parameter, most frequently on dihedral angle. The obtained dependences, represented often by so called conformation maps, were then used for qualitative considerations and comparisons with experimental data. Inaccuracy of the given procedure (consisting in that most geometric parameters remain constant throughout the calculation) can be removed by complete minimization of energy with respect to all internal coordinates. Such calculations by the PCILO method were carried out in a series of methylcycloheptanones[7].

The present paper compares the molecular geometries optimized by the PCILO method with experimental data and with the published results obtained by other semi-empirical methods. This basic information should serve for proposing a procedure suitable for optimization of geometries of greater molecules.

CALCULATIONS

For minimization of molecular energy within the PCILO approach we used the modification of the Newton–Raphson procedure described by Payne[8]. The first and the second partial derivatives of energy with respect to internal coordinates were calculated numerically by the method of finite differences Δx ($\Delta x = 1 \,.\, 10^{-4}$ nm for bond lengths, $\Delta x = 0{\cdot}05°$ for bond angles, $\Delta x = 0{\cdot}5°$ for dihedral angles). Bond polarities were optimized in each energy calculation, $0{\cdot}5N \,.\, (N + 3) + 1$ calculations ($N =$ number of the variable geometric parameters) being necessary for one optimization step of the algorithm used. Payne[8] gives a criterion δE_Q enabling evaluation of accuracy of the found position of the stationary point. It is based on comparison of real calculated energy of the presumed stationary point with the energy calculated on the basis of approximation of shape of the energy hyprsurface by quadratic form. In our calculations the search for the stationary point was iteratively repeated, if the value δE_Q was greater than a chosen limit (usually 5 J mol^{-1}). The standard PCILO program[1] was modified for the calculations.

The calculations were carried out for eleven small organic molecules chosen with the purpose of describing the most frequent bonding situations in organic chemistry. The selection was limited by the formalism of localized orbitals used in the PCILO method (aromatic compounds and compounds with pentacoordinated atoms are not involved, ref.[2]). Symmetry of molecules was not considered explicitely, in some cases non-symmetrical initial geometrical arrangements were chosen intentionally, so that expression of the symmetry point group might form one of the testing criteria of the method used.

RESULTS

Hydrocarbons. The optimized molecular geometries of ethane, ethylene, acetylene, 1,3-butadiene and cyclobutane are given in Table I. Except for cyclobutane, symmetry of the calculated structures corresponds to experiment. The C—C bonds are shorter in average by $3 \,.\, 10^{-3}$ nm, whereas the C—H bond lengths are longer by $3 \,.\, 10^{-3}$ nm. Mean absolute error of bond angles is 2·6°. If these values are compared with those obtained by the MINDO/3 (refs[9,10]) and MNDO (refs[11,12]) methods which are parametrized for calculations of heats of formation and molecular geometries (MINDO/3, C—C bonds shorter by $2 \,.\, 10^{-3}$ nm, C—H bonds longer by $5 \,.\, 10^{-4}$ nm, mean absolute error in determination of bond angles 4·4°; MNDO, C—C bonds shorter by $3 \,.\, 10^{-4}$ nm, C—H bonds longer by $5 \,.\, 10^{-4}$ nm, mean absolute error in bond angle determination 1·6°; the data are related to the same set of the tested molecules), then it can be stated that the bond lengths are reproduced more precisely by the MINDO/3 and MNDO methods, whereas the PCILO method is somewhat more precise than the MINDO/3 method in the bond angles values. *trans*-1,3-Butadiene results to be more stable than the *cis* isomer by 3·35 kJ mol^{-1}. For cyclobutane the PCILO method gives planar ring conformation (D_{4h}) as the energetically most favourable. This conclusion contradicts both the experimental data[13] and results of the *ab initio* calculation carried out in the 6–31G* basis (ref.[14]) which indicate the presence of puckered conformation (D_{2d}). A similar discrepancy (planar – D_{4h} conformation) is also shown by the MNDO calculation[12].

Compounds type CH_3—X (X = NH_2, OH, OCH_3, SH, Cl). With this group of molecules the PCILO method was tested for its ability to express molecular geometry of compounds containing other atoms besides carbon and hydrogen. Formalism of the method necessitates, for the individual heteroatoms, explicit definition of directional properties of lone electron pairs of σ-type by means of fictive atoms (FA). We used the values recommended in ref.[2], the distances X—FA and the angles C—X—FA, H—C—X—FA being kept constant in the course of the optimization. From comparison of the results (Table II) with experimental data it follows, that the bond lengths type C—X are reproduced with approximately the same accuracy by the PCILO and the MINDO/3 methods[17–19] (in the both cases they are shorter by $6 . 10^{-3}$ nm for the same set of molecules), being less accurate than those by MNDO method[12] (which are shorter by $3 . 10^{-3}$ nm). For the X—H bond lengths the PCILO method exhibits a greater error than the MINDO/3 (X—H longer by $9 . 10^{-3}$ and $1 . 10^{-3}$ nm, respectively). In determination of bond angle on heteroatom all the three method reach comparable accuracy (the mean absolute error 4·3° (PCILO), 4·1° (MINDO/3), 3·5° (MNDO)). Geometric parameters of methyl groups are determined with similar accuracy as in the case of hydrocarbons. The staggered conformation of methanol results to be more stable than the eclipsed conformation by 5·2 kJ mol^{-1}. Change of lone electron pair of σ-type at nitrogen of methylamine to the π-type leads to planar sp^2 hybridization of the nitrogen atom.

Urea. This molecule containing two different types of heteroatoms was used for study of optimization of positions of fictive atoms determining the directional properties of lone electron pairs of σ-type at the oxygen atom and choice of type of lone electron pairs on the nitrogen atoms. The results are given in Table III. Optimization of the O—FA distance results in monotonous approaching of fictive atoms towards oxygen atom of the carbonyl group with simultaneous great energy decrease (about 10^2 kJ mol^{-1}). The C—O—FA angles are increased during the optimization so that the both lone electron pairs begin to overlap. Spatial arrangement of the real atoms remains, however, practically unchanged in spite of these non-realistic changes. Type change of the lone electron pair on the nitrogen atoms affects markedly the resulting geometry of the studied molecule (Table III, Fig. 1). For the choice of π-type lone electron pairs the calculated conformation (Fig. 1*a*) is significantly nonplanar with pyramidal arrangement of the bonds on the nitrogen atoms (similar to calculation of formamide[25]). In case of σ-type the molecule shows but slight nonplanar deformation (Fig. 1b). This conformation is less stable by 117 kJ mol^{-1} and does not represent the minimum but only a stationary point (the matrix of the second derivatives of energy has two negative eigenvalues)[26,27]. Optimization of directional properties of lone electron pairs of σ-type did not result in any significant change of geometry of the molecule.

TABLE I

Optimized Molecular Geometries of Hydrocarbons

Parameter	Molecular geometry[a]		
	initial	optimized	experimental
	Ethane, ref.[15]		
Symmetry	C_3	D_{3d}	D_{3d}
C—C	0·154	0·146	0·154
C—H	0·109	0·113	0·109
C—C—H	109·5	112·1	110·9
H—C—C—H	10·0	60·0	60·0
	Ethylene, ref.[15]		
Symmetry	D_{2h}	D_{2h}	D_{2h}
C—C	0·134	0·132	0·134
C—H	0·108	0·112	0·109
C—C—H	120·0	123·9	121·2
	Acetylene, ref.[15]		
Symmetry	C_1	$D_{\infty h}$	$D_{\infty h}$
C—C	0·120	0·120	0·120
C—H	0·106	0·110	0·106
C—C—H_1	175·0	180·0	180·0
C—C—H_2	170·0	180·0	180·0
	trans-1,3-Butadiene, ref.[16]		
Symmetry	C_{2h}	C_{2h}	C_{2h}
C_1—C_2, C_3—3_4	0·134	0·134	0·134
C—H	0·108	0·112—0·113	—
C_2—C_3	0·154	0·146	0·146
C_1—C_2—C_3	120·0	127·1	123·3
H—C_1—C_2, H—C_4—C_3	120·0	123·8—124·6	—
H—C_2—C_3	120·0	115·2	—

TABLE I
(*Continued*)

Parameter	Molecular geometry			
	initial		optimized	experimental
	cis-1,3-Butadiene			
Symmetry	C_{2v}[b]		C_{2v}	
$C_1—C_2$	0·134		0·134	
$C_2—C_3$	0·154		0·145	
C—H	0·108		0·112—0·113	
$C_1—C_2—C_3$	120·0		129·3	
$H—C_1—C_2$, $H—C_4—C_3$	120·0		123·6—124·6	
$H—C_2—C_3$	120·0		113·9	
	Cyclobutane, ref.[13]			
Symmetry	D_{4h}	D_{2d}	D_{4h}[c]	D_{2d}
C—C	0·154	0·154	0·150	0·155
C—H	0·109	0·108	0·113	0·109

---

[a] The bond lengths in nm, the bond angles in degrees; [b] the symmetry was slightly distorted (the dihedral angles deviations ±3°); [c] the both initial forms give the same resulting geometry.

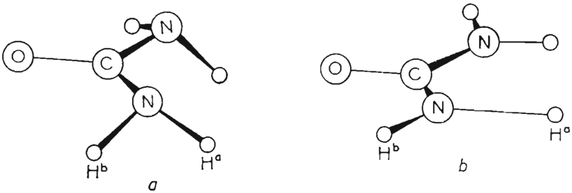

FIG. 1

Optimized Molecular Geometry of Urea

a π-Type lone electron pairs on nitrogen atoms (the energy minimum), *b* σ-type (this conformation does not represent an energy minimum but only a stationary point).

TABLE II

Optimized Molecular Geometries of Compounds Type CH_3X (X = NH_2, OH, OCH_3, SH, Cl)

Parameter	Molecular geometry	
	optimized	experimental
	Methylamine, ref.[20]	
Symmetry	C_s	C_s
C—N	0·141	0·147
N—H	0·113	0·101
C—H^a	0·113	0·109
C—H^{b,c}	0·109	0·109
C—N—H	112·8	112·1
H—N—H	97·4	105·9
Θ^d	2·0	3·5
	Methanol^e, ref.[21]	
Symmetry	C_s	C_s
C—O	0·137	0·143
O—H	0·105	0·095
C—H	0·113	0·109
C—O—H	105·1	108·5
O—C—H^a	108·6	107·0
H^b—C—H^c	111·8	108·6
Θ^d	1·9	3·3
	Dimethyl ether, ref.[22]	
C—O	0·138	0·141
C—O—C	104·8	111·7
	Methanethiol, ref.[23]	
Symmetry	C_s	C_s
C—S	0·175	0·182
S—H	0·139	0·133
C—H	0·113	0·110
C—S—H	102·1	100·3
S—C—H	112·1	108·5
Θ^d	0·7	0·5

TABLE II
(*Continued*)

Parameter	Molecular geometry	
	optimized	experimental
	Chloromethane, ref.[24]	
Symmetry	C_{3v}	C_{3v}
C—Cl	0·168	0·178
C—H	0·113	0·110
Cl—C—H	112·2	108·0

[a,b,c] Differentiation of hydrogen atoms of methyl groups; [d] angle between axis of methyl and. C—X bond; [e] staggered conformation.

TABLE III
Optimized Molecular Energy of Urea (*A* π-type of lone electron pairs at nitrogen atoms; B σ-type)

Parameter	Molecular geometry		
	optimized		experimental (ref.[28])
	A	B	
C—O	0·128	0·129	0·124
C—N	0·140	0·137	0·135
N—H[a]	0·108	0·107	0·099
N—H[b]	0·109	0·107	0·100
C—N—H[a]	108·3	122·8	119·8
C—N—H[b]	106·5	121·8	118·1
O—C—N	121·5	121·5	121·5
O—C—N—H[a]	27·0	2·0	0·0
O—C—N—H[b]	138·0	176·0	180·0
N—O—C—N′	186·5	187·0	180·0

[a,b] Differentiation of hydrogen atoms, see Fig. 1.

DISCUSSION

From the given results it follows that the PCILO method reproduces the molecular geometries of the tested compounds with acceptable accuracy. The bond lengths (except for the bonds hydrogen–heteroatom) are only slightly less accurate than those found by the MINDO/3 method, the bond and dihedral angles are calculated with about the same accuracy. The present version of the PCILO method uses the parameters taken from the CNDO/2 method which decidedly do not represent the optimum solution with respect to the fact that the PCILO method calculates explicitly a part of the correlation energy. Improved agreement between the calculated geometry and experimental data can be achieved by calibration of the parameters, as it is the case with the MINDO/3 and MNDO methods[9,11]. In the present version the method is applicable for semi-quantitative study of conformation of organic compounds. In accordance with the proposal of St-Jacques and coworkers[7] it is possible to use a higher accuracy of calculation of bond and dihedral angles by the optimization procedure using constant experimental values of bond lengths and optimizing only the bond and dihedral angles. Optimization of positions of fictive atoms which determine the directional properties of lone electron pairs of σ-type results in their nonrealistic distortion, the overall molecular geometry being practically unchanged. Besides that, involvement of positions of the fictive atoms among the variable geometric parameters makes the calculation longer. Hence it seems useful to give the positions of fictive atoms according to the rules proposed by the authors of the method[2] and keep them constant throughout the optimization course.

Advantage of the used optimization algorithm lies in explicit construction of the matrix of the second derivatives of energy with respect to variable geometric parameters, which can be used for analysis of the character of the found stationary point[26,27]. However, the algorithm is relatively little efficient. To obtain sufficiently accurate localization of the stationary point (fulfilment of the condition $\delta E_Q < < 5 \text{ J mol}^{-1}$) it was necessary, in all the studied cases, to repeat iteratively the optimization cycles 3 to 6 times. Thus the speed of the PCILO method is not exploited. If the initial geometry does not represent a sufficiently close approximation of the stationary point (especially so in case of inaccurate assessment of dihedral angles), then the described procedure does not converge.

REFERENCES

1. Claverie P., Daudey J. P., Diner S., Giessner-Prettre C., Gilbert M., Langlet I., Malrieu J. P., Pincelli U., Pullman B.: *Quantum Chemistry Program Exchange* No 220. Bloomington, Indiana, U.S.A.
2. Malrieu J. P.: Modern Theor. Chem. *7*, 69 (1977).
3. Pullman B., Pullman A.: Advan. Protein Chem. *28*, 347 (1974).
4. Kwiatkowski J. S., Pullman B.: Advan. Heterocycl. Chem. *18*, 199 (1975).
5. Pullman B., Courrière P.: Mol. Pharmacol. *8*, 371 (1972).

6. Maigret B., Pullman B.: Theor. Chim. Acta *37*, 17 (1975).
7. St-Jacques M., Vaziri C., Frenette D. A., Goursot A., Fliszár S.: J. Amer. Chem. Soc. *98*, 5759 (1976).
8. Payne P. W.: J. Chem. Phys. *65*, 1920 (1976).
9. Bingham R. C., Dewar M. J. S., Lo D. H.: J. Amer. Chem. Soc. *97*, 1285 (1975).
10. Bingham R. C., Dewar M. J. S., Lo D. H.: J. Amer. Chem. Soc. *97*, 1294 (1975).
11. Dewar M. J. S., Thiel W.: J. Amer. Chem. Soc. *99*, 4899 (1977).
12. Dewar M. J. S., Thiel W.: J. Amer. Chem. Soc. *99*, 4907 (1977).
13. Almenningen A., Bastiansen O., Skancke P. N.: Acta Chem. Scand. *15*, 711 (1961).
14. Cremer D.: J. Amer. Chem. Soc. *99*, 1307 (1977).
15. Herzberg G.: *Molecular Spectra and Molecular Structure* III. *Electronic Spectra and Electronic Structure of Polyatomic Molecules.* Van Nostrand, New York 1966.
16. Kuchitsu K., Fukuyama T., Morino Y.: J. Mol. Struct. *1*, 463 (1968).
17. Bingham R. C., Dewar M. J. S., Lo D. H.: J. Amer. Chem. Soc. *97*, 1302 (1975).
18. Bingham R. C., Dewar M. J. S., Lo D. H.: J. Amer. Chem. Soc. *97*, 1307 (1975).
19. Dewar M. J. S., Lo D. H., Ramsden C. A.: J. Amer. Chem. Soc. *97*, 1311 (1975).
20. Lide D. R., jr: J. Chem. Phys. *27*, 343 (1957).
21. Lees R. M., Barker J. G.: J. Chem. Phys. *48*, 5299 (1968).
22. Blukis U., Kasai P. H., Myers R. J.: J. Chem. Phys. *38*, 2753 (1963).
23. Kilb R. W.: J. Chem. Phys. *23*, 1736 (1955).
24. Costain C. C.: J. Chem. Phys. *29*, 864 (1958).
25. Maloň P., Bystrický S., Bláha K.: This Journal *43*, 781 (1978).
26. Pancíř J.: This Journal *40*, 2726 (1975).
27. Dewar M. J. S.: Far. Discuss. Chem. Soc. No *62*, 197 (1977).
28. Sutton L. E.: *Tables of Interatomic Distances and Configuration in Molecules and Ions.* Supplement. The Chemical Society, Burlington House, London 1965.

Translated by J. Panchartek.