EQUILIBRIUM GEOMETRIES OF URACIL AND ITS C- AND N-METHYLATED DERIVATIVES

LÁSZLÓ HARSÁNYI*

Hungarian Academy of Sciences, Research Laboratory for Inorganic Chemistry, Department of Structural Studies, P.O.B. 117, H-1431 Budapest (Hungary)

ATTILA CSÁSZÁR and PÁL CSÁSZÁR**

Laboratory of Theoretical Chemistry, Eötvös Loránd University, Muzeum krt. 6-8, H-1088 Budapest (Hungary)

(Received 25 June 1985)

ABSTRACT

Equilibrium geometries, heats of formation and dipole moments of free uracil (U), 1-methyl-uracil (1-MeU), 3-methyl-uracil (3-MeU), 5-methyl-uracil (5-MeU, thymine), 6-methyl-uracil (6-MeU), 1,3-dimethyl-uracil (1,3-diMeU) and 5,6-dimethyl-uracil (5,6-diMeU) have been determined by semiempirical MINDO/3 calculations using Pulay's force method. Monosubstitutions by the methyl group manifested themselves in mainly local changes in bond distances and angles, while effects of disubstitutions, in good approximation, proved to be additive. Additivity also holds for heats of formation and dipole moments.

INTRODUCTION

Molecules containing NH and CO bonds, especially those having a peptide linkage, play a fundamental role in basic biological processes. Such molecules are pyrimidine bases, constituents of the genetic material, and at the same time, advantageously, have an appropriately moderate size from the quantum chemical point of view.

At present, available experimental data on the geometry of uracil are limited. Investigations in the solid phase revealed its crystal structure [1, 2] and showed that the molecular structure is influenced by intermolecular H-bondings. Hence, information for free molecules can be taken from gas-phase studies or from theoretical calculations.

The only available experimental study for free uracil by gas-phase electron diffraction (ED) is currently in press [3]. No experiments for structure determination of methylated derivatives in the gas phase have been reported.

---

*Author to whom correspondence should be addressed.
**On leave at: Department of Chemistry, The University of Texas at Austin, Austin, TX 78712, U.S.A.

As for the calculations, the first results by MINDO/2 [4] are poorly reliable due to the imperfections of the method. Similarly, CNDO/2 optimizations [5], yield some bond lengths and angles far from reality. Minimal basis set (STO-3G) ab initio calculations [6, 7] resulted in geometry parameters of semiempirical quality. Better predictions are obtained with a procedure proposed by Pulay et al. [8], in which geometry parameters calculated by ab initio 4-21 optimization are corrected empirically (see refs. 9 and 10 for uracil and thymine, respectively).

An ED study on the molecular structure of maleimide [11] demonstrated that geometry optimization by the MINDO/3 method yields equilibrium bond lengths with an accuracy similar to ab initio calculations at the 4-21 level (see [8]). This inspired us to repeat our former CNDO/2 calculations [5] for the series of methylated uracils but at the MINDO/3 level. Note that the higher quality of the semiempirical MINDO/3 geometries is intrinsically related to the parametrization of the method [12].

In the present paper, a MINDO/3 study of methylation effects on the molecular structure of uracil is given. Calculated heats of formation and dipole moments are also discussed.

COMPUTATIONAL DETAILS

Geometry optimizations were done in the MINDO/3 approximation [12] using a modified version of Rinaldi's GEOMO program system [13] with corrections as suggested in ref. 14. Energy gradients were calculated according to Pulay's force method [15]. For the automatic geometry optimization the GDIIS method [16] was applied successfully 7 to 9 full SCF gradient calculations were sufficient to reach the optimum (r_e) for each molecule. Optimizations were terminated when predicted changes in the internal coordinates became smaller than 0.05 pm and 0.05° for bond stretchings and bendings respectively.

The planar form of the diketo tautomers of methylated uracils with one methyl C—H bond in the plane of the uracil ring were considered for the following reasons: (a) The preference for the diketo form of uracil is shown by all quantum chemical calculations (except that of Zielinski and Rein [4] by MINDO/2) and various spectroscopic results [17—19]. (b) The planarity of the uracil ring is widely accepted based on extensive force field studies on uracil (CNDO/2 [20], MINDO/3 [21], ab initio 4-21 [9]) and thymine (CNDO/2 [22], MINDO/3 [23], ab initio 4-21 [10]). (c) The C atom of the methyl group (C_{methyl}) should be in the plane of the ring since separate optimizations with a starting position of C_{methyl} out of the plane of the ring result in a return to the planar structure. (d) The methyl group can be fixed at a position symmetric with respect to the plane of uracil for the barrier of methyl group rotation must be very small (ca. 20—40 J mol^{-1} cf. experiments [24] and ab initio calculations on toluene [25] and a MINDO/3 study on 4-methyl-pyridine [26]) and thus there is no reason to distinguish the various rotamers at the MINDO/3 level.

RESULTS AND DISCUSSION

The available theoretical and experimental geometries for free uracil are collected in Table 1. MINDO/3 predicts reliable geometrical parameters for uracil, as was similarly observed for maleimide [11].

The bond angles and bond lengths calculated by MINDO/3 for the six methylated uracils are summarized in Tables 2 and 3, respectively. For numbering of atoms see Fig. 1.

Bond angles

Substitution of H by electron donating groups, such as - CH_3 induce a characteristic lengthening in benzene type rings and the decrease of the *ipso*

TABLE 1

Bond distances (in pm) and angles (in °) in free uracil (r_e)

Parameter	CNDO/2[a]	Calculations (ab initio)		MINDO/3[d]	Electron diffraction[e] (r_g)
		STO-3G[b]	4-21[c]		
Bond lengths					
N_1-C_2	138.4	142.7	136.9[f]	139.0	139.5
C_2-N_3	137.9	141.9	136.5[f]	138.0	139.1
N_3-C_4	138.9	143.2	138.9[f]	140.1	141.5
C_4-C_5	143.5	149.2	145.8	147.0	146.2
$C_5=C_6$	134.3	132.3	134.5	135.9	134.3
C_6-N_1	137.9	140.5	137.0[f]	136.7	139.6
N_1-H_7	106.4	101.9	100.0	102.7	100.2[g]
$C_2=O_8$	127.8	121.9	121.0	121.6	121.2
N_3-H_9	106.7	101.9	100.3	103.4	100.2[g]
$C_4=O_{10}$	127.8	122.1	120.9	121.1	121.1
C_5-H_{11}	111.2	107.2	107.0	110.2	107.2[g]
C_6-H_{12}	111.7	108.6	107.4	111.3	107.2[g]
Bond angles					(r_α)
$C_2-N_1-C_6$	121.7	122.7	123.3	126.6	123.2
$N_3-C_2-N_1$	114.9	112.7	113.6	109.9	114.6
$C_4-N_3-C_2$	126.2	127.9	128.5	131.6	126.0
$C_5-C_4-C_3$	115.1	112.6	113.4	111.1	115.5
$C_6=C_5-C_4$	119.3	121.2	119.3	121.2	119.7
$N_1-C_6=C_5$	122.4	122.6	121.8	119.6	122.1
$C_2-N_1-H_7$	118.8	116.5	115.7	117.6	115.7[g]
$O_8=C_2-N_1$	121.9	123.8	122.7	123.0	123.8
$H_9-N_3-C_2$	115.7	115.4	115.4	112.8	115.4[g]
$O_{10}=C_4-N_3$	118.0	120.3	120.7	120.4	120.2
$H_{11}-C_5-C_4$	119.9	117.4	118.1	119.1	118.1[g]
$H_{12}-C_6=C_5$	123.8	121.9	122.8	123.8	122.8[g]

[a]Ref. 5. [b]Ref. 6. [c]Ref. 9. [d]Present results. [e]Ref. 3. [f]Uncorrected value. [g]Fixed parameter.
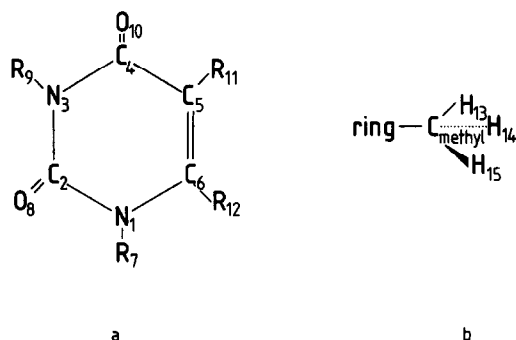
a   b

Fig. 1. Numbering of atoms in the molecules. (a) Index of R represents index of H resp. C involved in the bond to the ring. R=H for all R in the case of uracil; respective methylated derivatives are named following the index of ring atom bonded to the methyl group (i.e., R_7=CH$_3$ and R_9=R_{11}=R_{12}=H for 1-MeU etc.). (b) Smallest indexes of H in the methyl group (13 resp. 16) correspond to the H in the plane of the ring.

TABLE 2

Comparison of bond angles (°) of methylated uracils (MINDO/3 calculations)

Angles	Compounds						
	U	1-MeU	3-MeU	5-MeU	6-MeU	1,3-diMeU	5,6-diMeU
Ring							
C_2—N_1—C_6	126.6	121.3	127.3	127.0	129.1	121.7	130.3
N_3—C_2—N_1	109.9	112.4	112.9	108.9	110.0	115.7	108.7
C_4—N_3—C_2	131.6	132.4	125.7	132.0	130.8	126.2	131.0
C_5—C_4—N_3	111.1	109.9	114.1	113.3	111.1	113.0	113.9
C_6=C_5—C_4	121.2	121.5	121.7	116.6	123.8	122.1	118.9
N_1—C_6=C_5	119.6	122.5	118.2	122.2	115.1	121.3	117.7
C_2—N_1—H_7	117.6	—	117.0	117.6	115.4	—	114.9
O_8=C_2—N_1	123.0	123.2	120.0	123.5	123.0	120.2	123.7
H_9—N_3—C_2	112.8	112.2	—	112.8	113.2	—	113.4
O_{10}=C_4—N_3	120.4	121.1	121.1	119.0	120.6	121.8	118.7
H_{11}—C_5—C_4	119.1	119.3	118.5	—	117.2	118.6	—
H_{12}—C_6=C_5	123.8	120.9	124.5	122.7	—	121.5	—
C_{Me_1}—ring	—	122.1	114.4	122.1	127.0	122.1	118.7
C_{Me_2}—ring	—	—	—	—	—	114.3	128.3
Substituent							
H_{13}—C_{Me_1}—ring	—	115.6	114.7	115.3	112.4	115.6	115.5
H_{14}—C_{Me_1}—ring	—	112.6	113.4	112.6	113.6	112.5	112.8
H_{15}—C_{Me_1}—ring	—	112.6	113.4	112.6	113.6	112.5	112.8
H_{13}—C_{Me_1}—H_{14}	—	105.0	104.0	105.3	105.9	105.0	104.9
H_{13}—C_{Me_1}—H_{15}	—	105.0	104.0	105.3	105.9	105.0	104.9
H_{16}—C_{Me_2}—ring	—	—	—	—	—	114.5	113.6
H_{17}—C_{Me_2}—ring	—	—	—	—	—	113.5	113.4
H_{18}—C_{Me_2}—ring	—	—	—	—	—	113.5	113.4
H_{16}—C_{Me_2}—H_{17}	—	—	—	—	—	106.3	105.9
H_{16}—C_{Me_2}—H_{18}	—	—	—	—	—	106.3	105.9

TABLE 3

Comparison of bond lengths (pm) of methylated uracils (MINDO/3 calculations)

Bond	Compounds						
	U	1-MeU	3-MeU	5-MeU	6-MeU	1,3-diMeU	5,6-diMeU
Ring							
N_1-C_2	139.0	140.8	138.9	138.7	138.8	140.8	138.5
C_2-N_3	138.0	138.1	140.1	137.8	137.8	140.2	137.5
N_3-C_4	140.1	139.7	142.2	140.0	139.8	141.8	139.5
C_4-C_5	147.0	146.7	147.0	149.1	146.8	146.7	149.3
$C_5=C_6$	135.9	135.7	135.6	137.6	137.5	135.4	139.5
N_1-C_6	136.7	138.8	136.4	136.4	138.3	138.4	138.3
N_1-H_7	102.7	—	102.8	102.8	103.0	—	103.1
$C_2=O_8$	121.6	121.6	121.7	121.5	121.6	121.7	121.6
N_3-H_9	103.4	103.5	—	103.4	103.4	—	103.4
$C_4=O_{10}$	121.1	121.1	121.1	121.2	121.2	121.1	121.3
C_4-H_{11}	110.2	110.3	110.3	—	110.4	110.3	—
C_6-H_{12}	111.3	111.5	111.2	111.5	—	111.5	—
$C_{ring}-C_{Me}$	—	—	—	149.3	149.1	—	149.7 149.6
$N_{ring}-C_{Me}$	—	144.1	145.3	—	—	144.2 145.4	
Substituent							
$C_{Me}-H_{Me}$							
$C_{Me_1}-H_{13}$	—	111.3	111.5	110.8	111.0	111.3	110.8
$C_{Me_1}-H_{14}$	—	112.0	111.9	111.3	111.1	112.0	111.4
$C_{Me_1}-H_{15}$	—	112.0	111.9	111.3	111.1	112.0	111.4
$C_{Me_2}-H_{16}$	—	—	—	—	—	111.5	111.0
$C_{Me_2}-H_{17}$	—	—	—	—	—	111.9	111.1
$C_{Me_2}-H_{18}$	—	—	—	—	—	111.9	111.1

bond angles [27, 28]. Of course, this effect must be accompanied by other changes in the ring geometries especially if ring planarity is conserved.

Table 4 demonstrates angle distortions due to methylation in the benzene—pyridine—uracil series. Note that: (a) The character of the distortions is the same in the series. (b) Changes at the point of substitution (*ipso* angle) become larger with decrease of conjugation. (c) MINDO/3 predicts more significant distortions than other methods. (d) C and N-methylation effects calculated by MINDO/3 are regular (unlike those by CNDO/2 [5]).

Superposition of monosubstitution effects is expected [5] for N,N- and C,C-dimethylations (1,3-diMeU and 5,6-diMeU, respectively). For example, for 5,6-diMeU, additivity was found for the $C_6C_5C_4$ and $C_5C_6N_1$ angles: $\angle 5,6\text{-diMeU} = \angle U + \angle[(5\text{-MeU}) - U] + \angle[(6\text{-MeU}) - U]$ $\angle C_6C_5C_4 = 121.2° - 4.6° + 2.6° = 119.2°$ $\angle C_5C_6N_1 = 119.6° + 2.6° - 4.5° = 117.7°$ since the corresponding angles in 5,6-diMeU are 118.9° and 117.7°, respectively (cf. Tables 2 and 4).

TABLE 4

Distortions (Δ) of angles due to methylation (in °)

Molecules	Substitution	Method[a]	Type of angle relative to substitution			
			ipso	ortho	meta	para
[Toluene—Benzene]	C-methylation	ED	−1.4	+0.9	0.0	−0.5
		a.i.	−1.4	+0.8	+0.2	−0.5
		CNDO/2[b]	−3.9	+2.3	+0.1	−0.9
		MINDO/3	−4.5	+2.5	+0.1	−0.8
[4-MePyridine[c]—Pyridine]	C-methylation	CNDO/2[b]	−4.3	+2.6	+0.2	−1.1
		MINDO/3	−4.6	+2.6	+0.1	−0.8
[5-MeU-U]	C-methylation	CNDO/2	−3.0	+1.7	0.0	−0.6
		MINDO/3	−4.6	+2.4	+0.4	−1.0
[6-Me-U-U]		CNDO/2	−4.4	+2.6	0.0	−1.0
		MINDO/3	−4.5	+2.6	0.1	−0.8
[1-MeU-U]	N-methylation	CNDO/2	−2.0	+1.2	0.0	−0.4
		MINDO/3	−5.3	+2.7	+0.5	−1.2
[3-MeU-U]		CNDO/2	−2.0	+1.2	0.0	−0.4
		MINDO/3	−5.9	+3.0	+0.6	−1.4

[a]Notation and references: ED = gas electron diffraction [3], a.i. = 4-21 basis ab initio, corrected [25], CNDO/2 [5], MINDO/3 (present) [b]Present results. [c]Geometry parameters can be found in ref. 26.

Bond lengths

Changes in bond lengths are relatively less significant than those in bond angles due to the higher force constants involved. However, relative changes differ by nearly an order of magnitude when passing from the bonds adjacent to substitution (X_{ipso}—X_{ortho}, X = C or N) to the other bonds in the ring. Methylation effects are mainly local; N-methylations yield a distortion of 2.1 pm, while C-methylations result in an increase by 1.7 pm of C=C, 2.1 pm of C—C and 1.6 pm of C—N for the X_{ipso}—X_{ortho} bonds (cf. Table 3).

Here again additivity is proved for the dimethylation, e.g. the length of the C=C bond in 5,6-diMeU can be correctly predicted based on the above data: 135.9 + 1.7 + 1.7 = 139.3 pm vs. 139.5 pm in Table 3.

Another question of interest is the relative length of the C—H bonds in the methyl group and on the uracil ring (cf. a similar study for toluene [25]). In Table 3 it is seen that, apart from 5-MeU, the methyl C—H bonds are somewhat longer than those attached to the ring. Based on chemical evidence, the latter order is expected; available experimental [29] and ab initio data [25] give further support to this finding.

A comparison of the geometrical parameters in Table 3 with those of other molecules calculated by the same (MINDO/3) method permits one to draw some qualitative conclusions on the ring structure of the pyrimidine bases considered (Table 5). Only a small delocalization can be observed in uracil and in its methylated derivatives. Though this delocalization seems to

TABLE 5

Comparison of some characteristic calculated bond lengths (in pm) with those of uracil (all calculations by the MINDO/3 method, experimental values are in parentheses)

Type of bond	Reference compound		Ref.	Uracil[a]
C—C	Cyclopropane	150.4(151.0)	12	
	Maleimide	150.8(150.6)	11	147.0(146.2)
C···C	Benzene	140.7(139.7)	21	
	Pyridine	140.6(139.2)	12	
C=C	Propene	133.3(133.6)	12	135.9(134.1)
	Maleimide	134.8(134.3)	11	
C=O	Formaldehyde	118.1(120.3)	12	
	t-Acrolein	119.5(120.8)	21	121.1(121.1)
	Maleimide	120.3(120.5)	11	

[a]Cf. Table 1.

be more pronounced than that for maleimide (1H-pyrrole-2,5-dione) [11] it is still of reduced degree and far from that of systems such as benzene, pyridine or pyrimidine. Consequently, the "pyrimidine" bases considered here have a structure dissimilar to that of pyrimidine itself.

Heats of formation and dipole moments

In most cases heats of formation and dipole moments calculated by MINDO/3 agree well with experiments, and are comparable to ab initio calculations with large basis sets [12, 30].

Using the optimized geometries of Tables 2 and 3, these data were calculated (see Table 6). Additivity holds excellently for dipole moments and, less perfectly, also for heats of formation (i.e. the dipole moment of 5,6-diMeU (4.78 D) is the same as that obtained for uracil (4.52 D) plus the effect due to the 5- and 6-methylations, cf. the differences 5-MeU-U and 6-MeU-U 4.98 − 4.52 = 0.46 and 4.32 − 4.52 = −0.20 D, respectively).

TABLE 6

Heats of formation (in kJ mol^{-1}) and dipole moments (in D) of methylated uracils by MINDO/3

Molecules	Heats of formation	Dipole moments (10^{30}C m)
U	−476.7	4.52(15.1)
1-MeU	−398.9	4.45(14.8)
3-MeU	−387.1	4.23(14.1)
5-MeU	−501.2	4.32(14.4)
6-MeU	−523.5	4.98(16.6)
1,3-diMeU	−306.7	4.15(13.8)
5,6-diMeU	−533.4	4.78(15.9)

CONCLUSIONS

Theoretical structures by MINDO/3 optimizations proved to be useful in predicting methylation effects in the series of methylated uracils. Changes in bond angles (mainly *ipso* and *ortho*) and bond lengths are suggested to be mainly local and to be additive in the case of disubstitution.

Qualitatively, this conclusion is the same as that found in our earlier CNDO/2 study, confirming that semiempirical methods of even limited accuracy may lead to a correct qualitative picture of structural changes if the method applied is used systematically within a series of molecules of similar structure. More generally, since the errors of calculation due to approximations of the semiempirical method applied are fairly well conserved for all the investigated molecules in a series, the comparison of data within the series may lead to reasonable chemical conclusions.

Note also that compensation of errors due to the MINDO/3 approximation by introducing empirical correction terms (like the offset values in ab initio calculations [8]), would be possible within one family of molecules only (cf. ref. 31). Extension of a given set of correction terms to other types of molecules is hindered since discrepancies between some experimental and MINDO/3 geometries are found to be of different sign for given molecular fragments in various types of molecules, i.e. $r(C-C)$ is slightly underestimated by MINDO/3 for saturated hydrocarbons [12], correct for maleimide [11] and overestimated for conjugated rings, such as benzenederivatives [21].

Besides the proof of the local character and additivity of methylation effects on the structure, the present study revealed that additivity also holds for heats of formation and dipole moments.

REFERENCES

1 R. F. Stewart and L. H. Jensen, Chem. Commun., (1968) 1019.
2 K. Ozeki, N. Sakabe and J. Tanaka, Acta Crystallogr., Sect. B, 24 (1969) 1038.
3 Gy. Ferenczy, B. Rozsondai. L. Harsányi and I. Hargittai, J. Mol. Struct., in press.
4 T. J. Zielinski and R. Rein, Int. J. Quant. Chem., 14 (1978) 851.
5 L. Harsányi and P. Császár, Acta Chim. Hung., 111 (1982) 351.
6 T. J. Zielinski, M. Shibata and R. Rein, Int. J. Quant. Chem., 19 (1981) 171.
7 T. J. Zielinski, Int. J. Quant. Chem., 22 (1982) 639.
8 P. Pulay, G. Fogarasi, F. Pang and J. E. Boggs, J. Am. Chem. Soc., 101 (1979) 2550.
9 P. Császár, L. Harsányi, A. Császár and J. E. Boggs, J. Mol. Struct., to be published.
10 P. Császár, L. Harsányi and J. E. Boggs, to be published.

11 L. Harsányi, E. Vajda and I. Hargittai, J. Mol. Struct., 129 (1985) 315.

12 R. C. Bingham, M. J. S. Dewar and D. H. Lo, J. Am. Chem. Soc., 97 (1975) 1285, 1294, 1302, 1307.

13 D. Rinaldi, Comp. Chem., 1 (1976) 109.

14 I. Mayer and M. Révész, Comp. Chem., 6 (1982) 153.

15 P. Pulay, Mol. Phys., 17 (1969) 197.

16 P. Császár and P. Pulay, J. Mol. Struct., 114 (1984) 31.

17 A. R. Katritsky and A. J. Waring, J. Chem. Soc., (1962) 1540. J. S. Kwiatkowski and B. Pullman, Adv. Heterocycl. Chem., 18 (1975) 199.

18 M. Szczesniak, M. J. Nowak, H. Rostkowska, K. Szczepaniak, W. B. Person and D. Shugar, J. Am. Chem. Soc., 105 (1983) 5969; M. Maltese, S. Passerini, S. Nunziante-Cesaro, S. Dobos and L. Harsányi, J. Mol. Struct., 116 (1984) 49; A. J. Barnes, M. A. Stuckey and L. Le Gall, Spectrochim. Acta, Part A, 40 (1984) 419.

19 M. J. Nowak, K. Szczepaniak, A. Barski and D. Shugar, Z. Naturforsch., Teil C, 33 (1978) 876.

20 L. Harsányi and P. Császár, Acta Chim. Hung., 113 (1983) 257.

21 A. Császár, Ph.D. Thesis, Budapest, 1984.

22 L. Harsányi and G. Pongor, unpublished results.

23 A. Császár, P. Császár and L. Harsányi, J. Mol. Struct., to be published.

24 J. L. Breuil, D. Cavagnat, J. C. Corhut, M. T. Forel, M. Fouassier and J. Lascombe, J. Mol. Struct., 57 (1979) 35.

25 F. Pang, J. E. Boggs, P. Pulay and G. Fogarasi, J. Mol. Struct., 66 (1980) 271.

26 F. Y. Hansen and H. L. McMurry, J. Mol. Struct., 57 (1979) 209.

27 A. Domenicano and P. Murray-Rust, Tetrahedron Lett., (1979) 2283, and refs. therein.

28 E. von Nagy-Felsöbüki, R. D. Topsom, S. Pollack and R. W. Taft, J. Mol. Struct., 88 (1982) 255, and refs. therein.

29 R. Seip, Gy. Schultz, I. Hargittai and Z. G. Szabó, Z. Naturforsch., Teil A, 32 (1977) 1178.

30 P. Császár and A. Császár, J. Mol. Struct. (Theochem.), 110 (1984) 405.

31 D. Van Hemelrijk, W. Versichel, C. Van Alsenoy and H. J. Geise, J. Comput. Chem., 2 (1981) 63.