Theoret. chim. Acta (Berl.) 7, 110—116 (1967)

Theoretical Study of the Electronic Properties of Biological Purines and Pyrimidines

III. The Representation of the Methyl Group and the Structure of Thymine *

ARMELLE DENIS and ALBERTE PULLMAN

Université de Paris, Institut de Biologie Physico-Chimique, 13, Rue Pierre Curie, Paris Vè

Received November 30, 1966

In view of a study of the atypical bases of the nucleic acids, the integral values appropriate for the methyl substituent have been determined in the SCF approximation previously described [*1*]. Both the inductive and the hyperconjugation effects have been introduced. An application to thymine is described.

En vue d'une étude détaillée des bases atypiques des acides nucléiques, les intégrales appropriées du substituant méthyle ont été déterminées dans le cadre de l'approximation SCF précédemment décrite [*1*]. L'effet inductif et l'effet d'hyperconjugaison sont introduits simultanément. Le procédé est appliqué à la thymine.

Im Rahmen einer Untersuchung über atypische Basen von Nukleinsäuren sind die Werte der Integrale, die im Zusammenhang mit Methylsubstitutenten auftreten, wie früher angegeben bestimmt worden. Die Methode berücksichtigt sowohl den induktiven als auch den Hyperkonjugationseffekt und wurde auf Thymin angewendet.

In the framework of a refined study of the structure of the nucleic acid bases and their analogs [*1, 2, 3*] we present here the results of an investigation concerning thymine (I) (5-methyluracil). The procedure employed is the Pariser-Parr-Pople approximation of the SCF method, appropriately scaled as described in Ref. [*1*], and supplemented by the inclusion of configuration mixing with all singly-excited configurations. The structure of the σ-bonds has been approximated by the procedure of DEL RE [*4*], adapted for conjugated molecules [*5*]:

Most of the published calculations which refer to thymine, deal in fact with uracil (II), thus escaping or neglecting the methyl problem [*6, 7*]**.

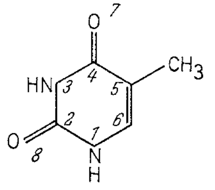
Apart from the interest of permitting to deal with the structure of thymine itself, the explicit introduction of the methyl group in the calculations has also

---

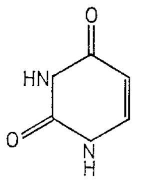
* This work was supported by grant No GM 12289-01 of the United States Public Health Service (National Institute of General Medical Sciences).

** One recent paper treats explicitely both uracil and thymine [*8*]. Unfortunately the large discrepancy with experimental results and an error in the chemical structure adopted for uracil [*1*] decreases greatly its significance.

Note added in proof. A recent paper by TANAKO and NAGAKURA [This journal, **6**, 320 (1966)], although entitled "Electronic Structure od Adenine and Thymine", deals in fact with Uracil.



I

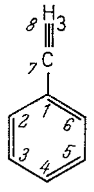


II

the advantage of allowing the study of other methylated bases which occur in a rather widespread fashion in the structure of nucleic acids [9] particularly in transfer RNA [10, 11, 12].

The Representation of the Methyl Group

According to the conventionnal representation of hyperconjugation [13] in the molecular orbital method, the methyl group is treated as contributing two electrons to the π-electron system, and two atomic orbitals to the basis of development of the molecular orbitals, namely one normal π orbital from the carbon atom and one pseudo-π orbital from the H_3 group. In the Pariser-Parr-Pople approximation, this representation necessitates essentially the determination of three integral values: the one-center coulomb integral $\gamma_{H_3H_3}$ the corresponding core parameter U_{H_3} and the two-center core integral β_{CH_3}. Previous determinations of this kind exist for $\gamma_{H_3H_3}$ and β_{CH_3} [14, 15, 16]. Since, however, our scheme differs from the usual approach particularly in the systematic lowering of the one-center core integrals [17], we redetermined the whole set of values. On the other hand we have also introduced an inductive effect of the methyl group through a modification of the effective orbital exponent of the π atomic orbital on the substituted carbon atom. The necessity of the introduction of an inductive effect *together with* hyperconjugation has been demonstrated a long time ago in the Hückel approximation in the study of "natural hypsochromic shifts" [18, 19] and recently reemphasized [20, 21] in connection with the interpretation of the EPR spectrum of the toluene and other anions. The deficiency of a purely hyperconjugative model has been shown to persist in a Pariser-Parr-Pople procedure [21].



III

Toluene (III) has been used as a reference compound. The pseudoorbital χ_8 has been written in terms of the hydrogen 1*s* orbitals h_1 and h_2 as the combination

$$\chi_8 = \frac{h_1 - h_2}{\sqrt{2(1-S)}}$$

so that the one-center coulomb integral γ_{88} is expressed in terms of coulomb and exchange integrals over 1*s* orbitals $(h_1 h_1/h_1 h_1)$, $(h_1 h_1/h_2 h_2)$, $(h_1 h_2/h_1 h_2)$, and $(h_1 h_1/h_1 h_2)$. The one-center 1*s* coulomb integral has been reduced from its theoretical value 17.003 eV to 12,86 eV by the Pariser relation [22]. This value corresponds to an effective orbital exponent $\xi^0_{1s} = 0.756$ which was then used to calculate the reduced values of the other 1*s* integrals, yielding finally $\gamma_{88} = 10.65$ eV.

The value $U_{H_3} = -8{,}9$ eV was then evaluated by using the approximate linear relationship which exists between γ and U for the π orbitals of one-electron contributing carbon, nitrogen and oxygen [*1*].

As concerns the inductive effect on the substituted carbon it has been introduced *via* a reduction of 0.04 of the Slater orbital exponent of the substituted carbon. The corresponding $\gamma^0_{C_1}$ and U_{C_1} values are listed in Tab. 1. In addition, this inductive effect has been reflected on the β values of the bonds involving C_1 through the relation:

$$\frac{\beta}{\beta_0} = \frac{S}{S_0}$$

where β_0 and S_0 are the resonance and overlap integrals for a normal carbon-carbon bond ($\beta_0 = -2.39$ eV for $r = 1.39$ Å). β_{CH_3} has been determined by trial and error so as to reproduce as many properties as possible for the reference molecule, toluene. The final list of integral values is given in Tab. 1.

Table 1. *Integral values for the methyl group (eV)*

$\gamma^0_{H_3}$	10.65	U_C	-9.5
γ_{CH_3}	6.67	β_{12}	-2.46
$\gamma^0_{C_1}$	10.87	β_{17}	-1.44
U_{H_3}	-8.9	β_{78}	-3.40
U_{C_1}	-9.0		

Table 2. *Comparison of toluene to benzene*

	Dipole moment (D)		Ionization potential (eV)		Ultra-violet transitions[a]			
					Theor.[f]			Exp.[g]
	Theor.	Exp.	Theor.	Exp.	ΔE (eV)	f	θ	ΔE (eV)
Toluene	0.45[b]	0.43[c]	9.04	8.82[d]	4.7	0.002	90	4.65
				9.18[e]	5.9	0.15	0	5.86
					6.4	1.34	0	6.57[i] – 6.79
					6.5	1.03	90	
					7.6	0.003	0	
					7.7	0.006	0	
Benzene[h]			9.46	9.25[d]	4.75	0		4.72
				9.56[e]	6.0	0		6.05
					6.6	1.1		6.74[i] – 6.97
					6.6	1.1		

---

[a] First is given the transition energy ΔE from the ground state, then the intensity, f, and finally the direction of polarization θ with respect to the symmetry axis of toluene.

[b] Directed from the methyl group to the ring.

[c] McClellan: Tables of experimental dipole moments (1963). Freeman and Co., San Francisco.

[d] Watanabe, K., T. Nakayama, and J. Mottl: J. Quant. Spect. Rad. Transfer **2**, 369 (1962).

[e] Crable, C. H., and L. Kearns: J. physic. Chem. **66**, 436 (1962).

[f] After inclusion of all singly-excited configurations (SCFCM).

[g] Petruska, J.: J. chem. Physics **34**, 1120 (1961) (vapor unless otherwise stated).

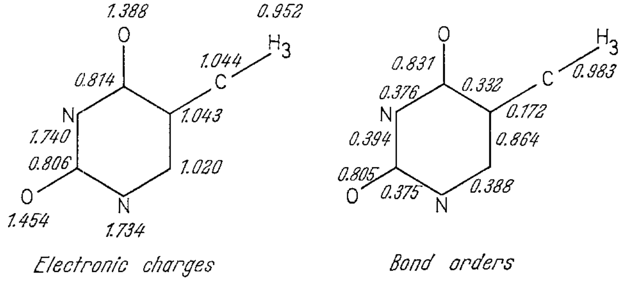
[h] Giessner-Prettre, C., H. Berthod, and A. Pullman (unpublished results).

[i] In heptane.

The main results relating to toluene are summarized in Tab. 2. The agreement with the experimental data for the ionization potential, dipole moment and the absorption spectrum, as well as the comparison with the corresponding data for benzene, are both sufficiently good for encouraging the further application to other methylated molecules.

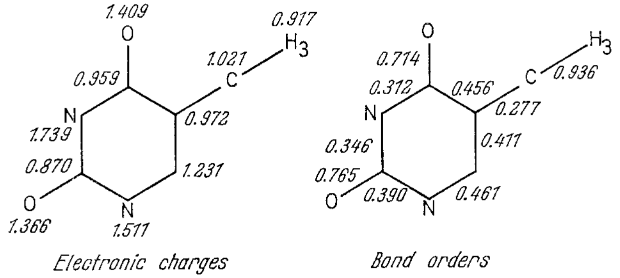
The Structure of Thymine

The SCF calculation of thymine yields the molecular orbitals, among which the highest filled in the ground state corresponds to the molecular ionization potential. The value obtained is 8.8 eV, distinctly lower than the corresponding value 9.15 eV calculated for uracil [*1*]. This result correlates very satisfactorily with the recent

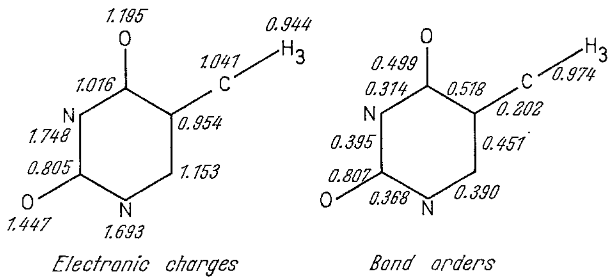


Ground state (SCF)

Fig. 1



First singlet (SCFCM)



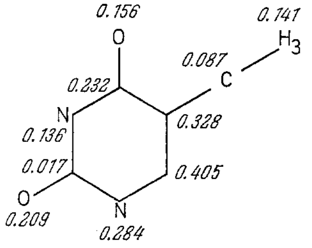
First triplet (SCFCM)

Fig. 2

determination of the ionization potentials of these compounds from a mass spectrometric study (8.94 eV for thymine and 9.47 eV for uracil) [*23*]. It may be advantageously compared with values suggested in previous SCF studies of these compounds [*7*, *8*] which ranged between 10 and 12 eV.

The lowest empty molecular orbital is 0.41 eV, very slightly changed with respect to the value 0.38 for uracil.

The calculated value of the dipole moment is 4.0 D. The only related available experimental value concerns 1,3-dimethyl uracil where it equals 3,9 D. Our calculated value for thymine appears therefore very reasonable. The distribution of the π-charge densities and bond orders in the ground state is given in Fig. 1. The corresponding quantities in the first excited singlet and triplet after configuration mixing are shown in Fig. 2 whereas the spin densities in the lowest triplet appear in Fig. 3.



First triplet (SCFCM)
Spin densities

Fig. 3

The results concerning the electronic transitions in thymine are given in Tab. 3. The experimental values of the observable two maxima of absorption in thymine are 4.68 and 6.0 eV according to Voet et al. [*24*], to be compared with the corresponding values in uracil, 4.8 and 6.1 eV. Oscillator strength for these two bands have been evaluated from the absorption curve of thymidine to be 0.2 and 0.3 respectively [*25*]. No intermediate absorption is apparent in the spectrum of

Table 3. *Electronic transitions in Thymine (eV)*

SCF					SCFCM					
singlets				triplets	singlets					triplets
orbital jump	ΔE	f	θ^a	ΔE	orbital jump	weight (%)	ΔE	f	θ	ΔE
6 – 7	4.9	0,8	162	2.8	6 – 7 5 – 8	0.84 0.04	4.65	0.4	156	2.0
5 – 7	5.5	0.04	–166	4.7	5 – 7 6 – 7	0.91 0.03	5.45	0.1	162	2.9
6 – 8	5.8	0.2	– 99	4.8	6 – 8 6 – 7	0.90 0.03	5.6	0.2	–127	3.2
5 – 8	6.4	0.02	+ 36	5.0	5 – 8 4 – 7	0.68 0.13	6.2	0.5	148	4.5_5
4 – 7	6.4	0.4		6.0_5	4 – 7 5 – 8	0.82 0.09	6.3	0.06		5.2
6 – 9	7.0	0.1			6 – 9 6 – 8	0.75 0.08	6.8	0.3		

---

[a] Angle counted from the N_1C_4 axis (counterclockwise).

thymine or uracil although the possibility of existence of a hidden band around 5.4 eV has been suggested in uracil [*26*].

Examination of the theoretical results and comparison with those previously obtained for uracil [*1, 2, 3*] show the following features:

a) For the first maximum, the bathochromic shift observed from uracil to thymine is correctly reproduced, so is also the numerical position of the maximum in particular after configuration mixing. In T as in U the 4th configuration has more weight in the mixing than the second.

b) A calculated transition with a small intensity appears towards 5.5 eV. This transition is not visible in the spectrum.

c) The next region of absorption includes a series of calculated maxima from 5.6 to 6.8 eV (5.8 to 7 eV before configuration mixing) which manifestly correspond to the wide second region of the observed spectrum.

d) The polarization direction of the first band is at 24° from the N_1C_4 axis towards the methyl group, a result comparing very favorably with the 19° — value determined for 1-methylthymine [*27*].

Among other structural characteristics of thymine worth underlining are:

1. The high electronic density of its C_5, the high value of the bond order of its $C_5 - C_6$ bond and the high value of the free valence of its C_6. These features are responsible for the high reactivity of the $C_5 - C_6$ bond towards addition reactions such as bromination or permanganate oxidation [*28*] and the preferential reactivity of C_6 towards the fixation of free radicals [*29, 30, 21*] (while in uracil this last fixation occurs at C_5 [*32*] in agreement with the value of the free valence greater at C_5 than at C_6 of this molecule [*33*]).

2. The great increase of the reactional properties (electronic charge and free valence) of C_6 in the first excited singlet and triplet and the relatively very large concentration of spin densities at the carbons of the $C_5 - C_6$ bond in the first excited triplet. These features may be related to the location of a number of photoreactions of thymine on the $C_5 - C_6$ bond and in particular to its involvement in the photodimerization of this and related molecules [*34*].

References

[*1*] Berthod, H., C. Giessner-Prettre und A. Pullman: Theoret. Chim. Acta **5**, 53 (1966).
[*2*] — — — Compt. Rend. Acad. Sci. Paris **262**, 2657 (1966).
[*3*] — — — Intern. J. Quant. Chem. (1967). In press.
[*4*] Del Re, G.: J. chem. Soc. (Lond.) **1958**, 4031.
[*5*] Berthod, H., and A. Pullman: J. chem. Physics **62**, 942 (1965).
[*6*] De Voe, H., and I. Tinoco, jr.: J. mol. Biol. **4**, 500 (1962).
[*7*] Nesbet, R. K.: Biopolymers Symposia **1**, 129 (1964).
[*8*] Nagata, C., A. Imamura, Y. Tagashira, and M. Kodama: Bull. chem. Soc. Jap. **38**, 1638 (1965).
[*9*] Srinivasan, P. R., and E. Borek: Science **145**, 548 (1964).
[*10*] Zamir, A., R. W. Holley, and A. Marquisee: J. biol. Chem. **246**, 1267 (1965).
[*11*] Holley, R. W., J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. K. Merril, J. R. Penwick, and A. Zamir: Science **147**, 1462 (1965).
[*12*] Zachau, H. G., G. Dutting und H. Feldmann: Angewandte Chemie **7**, 392 (1966).
[*13*] Pullman, B., et A. Pullman: Les théories électroniques de la chimie organique. Paris: Masson 1952.
[*14*] Morita, T.: Bull. chem. Soc. Jap. **33**, 1486 (1960).
[*15*] Inuzaka, K.: Bull. chem. Soc. Jap. **36**, 1045 (1963).

[16] Nordio, P., M. V. Pavan und G. Giacometti: Theoret. Chim. Acta **1**, 302 (1963).
[17] Pullman, A., and M. Rossi: Bioch. biophysic. Acta 88, 211 (1964).
[18] Bergmann, E. D., G. Berthier, Y. Hirshberg, E. Loewenthal, A. Pullman, and B. Pullman: Bull. Soc. Chim. **18**, 669 (1951).
[19] Pullman, B., M. Mayot, and G. Berthier: J. chem. Physics **18**, 257 (1950).
[20] Lazdins, D., and M. Karplus: J. Amer. chem. Soc. **87**, 920 (1965).
[21] Malrieu, J. P., A. Pullman und M. Rossi: Theoret. Chim. Acta **3**, 261 (1965).
[22] Pariser, R.: J. chem. Physics **21**, 568 (1953).
[23] Akopyan, M. E., and F. I. Vilesov: Doklady Akademii Nauk SSSR **161**, 1110 (1965).
[24] Voet, D., W. B. Gratzer, R. A. Cox, and P. Doty: Biopolymers **1**, 193 (1963).
[25] De Voe, H., and I. Tinoco, jr.: J. mol. Biol. **4**, 518 (1962).
[26] Clark, L. B., and I. Tinoco, jr.: J. Amer. chem. Soc. **87**, 11 (1965).
[27] Stewart, R. F., and L. H. Jensen: J. chem. Physics **40**, 2071 (1964).
[28] Pullman, B.: In Molecular Biophysics (Pullman, B., and M. Weissbluth eds.), p. 117. New York: Academic Press 1965.
[29] Gordy, W., B. Pruden, and W. Snipes: Proc. Natl. Acad. Sci. U.S. **53**, 751 (1965).
[30] Eisinger, I., and R. G. Shulman: Proc. Natl. Acad. Sci. U.S. **50**, 694 (1963).
[31] Pullman, B., et M. J. Mantione: Compt. Rend. Acad. Sci. Paris **260**, 5643 (1965).
[32] Herak, J. N., and W. Gordy: Science **153**, 1649 (1966).
[33] Pullman, B., et M. J. Mantione: Compt. Rend. Acad. Sci. Paris **261**, 5679 (1965).
[34] Mantione, M. J., and B. Pullman: Biochem. biophys. Acta **91**, 387 (1964); **95**, 663 (1965).

Prof. A. Pullman
Institut de Biologie Physico-Chimique
13, Rue Pierre Curie
Paris 5ème France