Theoret. chim. Acta (Berl.) 10, 231—239 (1968)

All Valence-Electrons Calculations on Purines and Pyrimidines

II. An Iterative Extended Hückel Computation*

A. Pullman, E. Kochanski, M. Gilbert, and A. Denis
Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, associé au C.N.R.S.
13, rue Pierre Curie, Paris 5e, France

Received January 11, 1968

An iterative extended Hückel calculation (IEHC) has been performed for the purine and pyrimidine bases of the nucleic acids. The results presented concern the electron distribution and the dipole moments, the atomic orbital populations and the "hybridization" ratios, the energies of the molecular orbitals and the ionization potentials. In particular, the relative orders of the dipole moments and of the π-ionization potentials are correctly reproduced but their absolute values are too great.

Die Purin- und Pyrimidinbasen der Nukleinsäuren wurden mit einer IEHC-Näherung berechnet. Es werden die Elektronenverteilungen, die Dipolmomente, die AO-Population, die Hybridisierungskonstanten, die MO-Energien und die Ionisationspotentiale angegeben. Die Reihenfolge der Dipolmomente und π-Ionisationspotentiale wird richtig erhalten.

La méthode de Hückel étendue itérative a été utilisée pour l'étude de la structure électronique des bases puriques et pyrimidiques des acides nucléiques. Les résultats présentés concernent la distribution des charges électriques et les valeurs des moments dipolaires, les populations des orbitales atomiques et les constantes d'hybridation, les énergies des orbitales moléculaires et les valeurs des potentiels d'ionisation. En particulier, les valeurs des moments dipolaires et des potentiels d'ionisation π tout en étant dans le bon ordre relatif sont trop élevées en valeur absolue.

A) Introduction

The last few years have seen the appearance of a number of approximations of the method of molecular orbitals aiming at the simultaneous treatment of all valence electrons in large molecules. The pioneering contribution in this domain is the "extended Hückel theory" of Hoffmann [1] which has met with a considerable success in particular in problems connected with shape and conformation [2, 3]. On the other hand, however, the prediction of properties connected with the coefficients of the atomic orbitals or the energy location of the individual molecular orbitals seems to be less reliable [2, 4], and improvements of the Hoffmann procedure have been proposed which essentially amount to make the diagonal matrix elements of the Hamiltonian depend on the net charges and to solve the problem by an iterative process [5, 6, 7]. Moreover the use of Clementi's atomic orbitals [8] is advocated by many authors [7, 9] as well as the replacement of the Wolfsberg-Helmholtz expression of the non-diagonal matrix elements by Cusachs' relation [9]. All these improvements have recently culminated in an "iterative extended Hückel" scheme which is advocated to yield results in satisfactory agreement with experiment [10, 11] for quantities such as orbital energies and dipole mo-

---

* This work was supported by grant No. GM-12229 of the U.S. Public Health Service (National Institute of General Medical Sciences).

ments. This and some results recently obtained in our laboratory [12] has encouraged us to use the procedure for the calculation of the fondamental bases of the nucleic acids.

The electronic structure of these compounds has been thoroughly investigated by different refinements of the π-electron approximation [13] since the first Hückel calculation in 1956 [14]; and the σ-framework has been explicitely introduced [15, 16] in a semi-empirical fashion, so as to permit dipole moment predictions. The simultaneous treatment of all σ and π-electrons, with no hypothesis on the hybridization of the atomic orbitals is however extremely desirable. The extended Hückel theory and the self consistent CNDO method have been recently applied for this purpose to these compounds in our laboratory [17, 18]. This work reports the results obtained by an iterative extended Hückel method.

Table 1. *Diagonal parameters used in the iterative extended Hückel scheme (eV).* $(-H_{\mu\mu} = A + Bq)$

atoms	$A(s)$	$A(p)$	B
H	13.8		14.3
C	21.2	11.4	11.9
N	25.5	14.4	13.7
NH	27.4	14.4[a]	13.7
O	32	16.45	15.2

[a] 12.5 for p_z.

The procedure and its parametrization have been described elsewhere [12, 19]. We shall refer to it as the IEHC procedure, the C indicating that the Cusachs formula is used for non-diagonal matrix elements and that diagonal matrix elements are choosen according to Cusachs and Reynolds [6]. Table 1 summarizes the numerical values utilized.

We shall focus our attention on the results obtained for uracil, cytosine, guanine and adenine (U, C, G, A). The results for purine (P) are given for comparison.

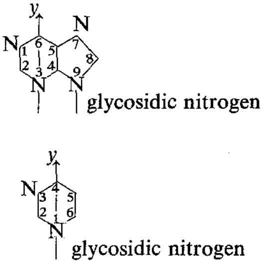

Fig. 1. Numbering of the atoms on the puric and pyrimidic skeletons and direction of the y axis

B) Electron Distribution and Dipole Moments

Fig. 2, a, b, c, gives the distributions of the gross atomic populations corresponding to Mullikens' partition [20] for total, σ and π-charges respectively.

Some interesting features of these electron distributions can be observed:

a) *Total Net Populations:* as would be expected, the more electronegative atoms bear the larger negative charges. Thus all the carbonyl oxygens carry

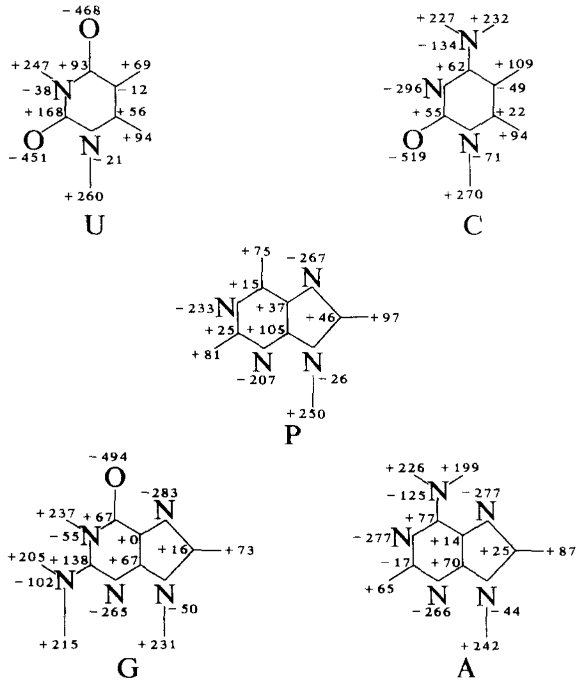

Fig. 2a

Fig. 2 a–c. Net atomic charges. a the total net charges; b the σ-net charges; c the π-net charges (10^{-3} electron units)

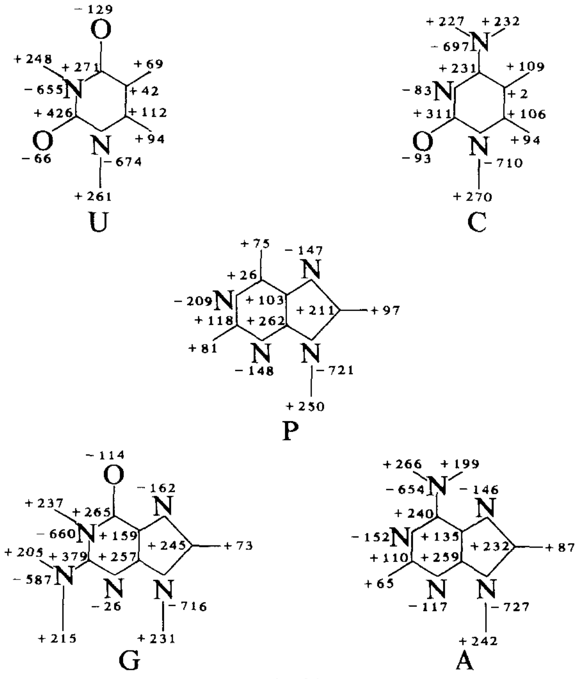

Fig. 2b

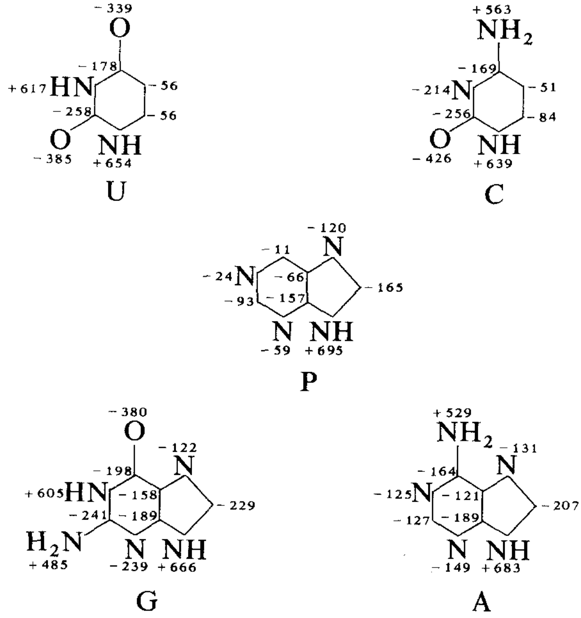

Fig. 2c

charges of about 0.5 e; the pyridine-type nitrogens attract from 0.2 to 0.3 electron units whereas pyrrole-like nitrogens are clearly less negative. The carbon atoms are generally globally positive with the exception of C_2 in adenine and C_5 in both pyrimidines. All the hydrogens are positively charged, the loss of electron occuring more easily along NH than along CH bonds. The NH_2 groups show an interesting constancy in global appearance. *Qualitatively*, all these findings are in fair agreement with the image of the total net populations obtained by the semi-empirical approach [15, 16][1]. A separate examination of the σ and π-distributions reveals however some interesting features.

b) *Net σ- and π-Populations:* Insofar as the IEHC method treats all valence electrons simultaneously including the interplay of charges one upon another by an iteration procedure, it is all the more interesting to decompose the resulting populations into their σ- and π-contributions. Then, it appears that the large global negativity of the carbonyl oxygens and of the pyridinic nitrogens is the result of a very small negative σ-charge with a larger excess in π-population. On the other hand the final small negative pyrrolic nitrogens appear as such as a result of a very high σ-population resulting in a strong repelling of the π-electrons by this center; a similar, although less marked, pattern is shown by the NH_2 nitrogens. A very drastic consequence of this NH behavior is a σ-discharge of all the carbon atoms which in turn attract the π-electrons as a result of both their own σ-positivity and of the repulsion of the NH-nitrogens for their π-electrons. Thus the image of the π-populations appears rather non-conformist since practically all the ring atoms bear an excess of π-electrons obviously coming from the

---

[1] Diagrams of the σ- and π-distributions which were not detailed in the original publications can be found in the review paper [13].

high delocalization of the pyrrollic π-lone-pairs, with some carbon atoms even richer in π-electrons than pyridinic nitrogens. It would seem that this situation is not quite reasonnable, and that it is partly due to a poor choice of the basic matrix elements of the Hamiltonian for the nitrogen atom in the state $sxyz^2$ [21].

c) *Dipole Moments:* Notwithstanding the aforementioned probable deficiencies of the calculated electron distributions it is interesting to consider the dipole moments which they generate. We have evaluated them in the point-charge approximation although we do not forget the objections which may be raised in this connection nor the defects of the equipartition of the overlap populations between the atoms[2]. With the above-mentioned reservations in mind, it is however interesting to observe that the calculated values (Table 2) follow a trend qualitatively similar to that displayed by the values calculated previously by Berthod, Giessner and Pullman [16] and already present in the early simpler evaluations [15, 23]: cytosine and guanine are predicted to have dipole-moments of the same order of magnitude, much larger than uracil, itself with a larger moment than adenine. The values themselves, however, appear too large with respect to the known experimental values and to the calculated values quoted, a feature which is probably a consequence of the exaggeration of the electron displacements previously mentioned. A very striking features of the results is the similarity in the directions of the total moments obtained by different methods (see Ref. [19] for a detailed comparison).

Table 2. *Dipole moments (debye units). In every column the first number is the dipole length, the second number is the angle in degrees of the moment with the y axis (counted counterclockwise)*

molecule	μ_σ		μ_π		μ_{sp}		$\mu_{tot.}$		μ_{exp} [a]
A	0.48	−1.2	3.94	132.3	1.76	50.4	4.39	104.3	3.0
P	0.53	172.4	5.11	47.6	2.48	34.8	7.22	56.6	4.3
U	2.16	75.3	4.76	19.4	1.39	34.0	7.63	35.7	3.9
C	2.39	106.0	8.12	127.4	3.48	90.0	12.43	124.6	
G	0.66	13.0	11.2	−34.8	2.85	−32.4	14.50	−32.4	

[a] Values from Ref. [23] for 9-methyl-adenine, 9-methyl-purine, and 1,3-dimethyl-uracil respectively.

The examination of the different components of the moments shows that i) the important trends both in relative values and in directions are largely due to the π-components, ii) the hybrid moments are not very different from those obtained by the CNDO method [18], iii) the σ-components add up in general with the other contributions. This is to be opposed to the CNDO results as has been discussed elsewhere [21].

C) Atomic Orbital Populations

Table 3 gives the populations in each atomic orbital $2s$, $2p_x$, $2p_y$ at the end of the iterative procedure. (The x and y axis are oriented as shown in Fig. 1.) Quite interesting constancies appear:

i) In all carbonyl compounds a typical "hybridization" on oxygen is $s^{1.7}p^{3.4}$. Moreover in each case there appears a nearly complete pure p lone-pair on the axis perpendicular to the $C = 0$ bond (this is directly visible on the CO bond of U

---

[2] Since the point-charge approximation does not include the hybrid moments, we have evaluated them separately as in the CNDO method [22].

Table 3. *Atomic orbital populations*

	atom	*s*	*x*	*y*	*z*
U	N_1	1.39	1.13	1.15	1.35
	N_3	1.39	1.14	1.13	1.38
	$O(C_2)$	1.65	1.60	1.82	1.38
	$O(C_4)$	1.70	1.96	1.47	1.34
	C_2	1.06	0.75	0.76	1.26
	C_4	1.10	0.88	0.76	1.18
	C_5	1.10	0.94	0.92	1.06
	C_6	1.13	0.83	0.93	1.06
C	N_1	1.41	1.14	1.16	1.36
	N_3	1.34	1.55	1.19	1.21
	$N(C_4)$	1.47	1.16	1.07	1.44
	O	1.67	1.62	1.80	1.43
	C_2	1.13	0.74	0.82	1.26
	C_4	1.11	0.90	0.76	1.17
	C_5	1.09	0.94	0.96	1.05
	C_6	1.13	0.83	0.94	1.08
G	N_1	1.39	1.13	1.14	1.40
	C_2	1.07	0.81	0.75	1.24
	N_3	1.32	1.02	1.68	1.24
	C_4	1.09	0.69	0.96	1.19
	C_5	1.08	0.81	0.96	1.16
	C_6	1.10	0.89	0.75	1.20
	N_7	1.42	1.14	1.61	1.12
	C_8	1.14	0.85	0.76	1.23
	N_9	1.42	1.19	1.11	1.33
	$N(C_2)$	1.42	1.06	1.11	1.52
	$O(C_6)$	1.70	1.96	1.46	1.38
A	N_1	1.37	1.59	1.20	1.12
	C_2	1.19	0.87	0.83	1.13
	N_3	1.36	1.04	1.72	1.15
	C_4	1.09	0.69	0.97	1.19
	C_5	1.08	0.81	0.97	1.12
	C_6	1.08	0.91	0.76	1.16
	N_7	1.41	1.12	1.61	1.13
	C_8	1.14	0.86	0.76	1.21
	N_9	1.42	1.20	1.10	1.32
	$N(C_6)$	1.45	1.14	1.06	1.47
P	N_1	1.39	1.61	1.20	1.02
	C_2	1.18	0.87	0.83	1.09
	N_3	1.36	1.06	1.72	1.06
	C_4	1.08	0.69	0.97	1.16
	C_5	1.09	0.83	0.97	1.07
	C_6	1.16	0.88	0.94	1.01
	N_7	1.41	1.12	1.62	1.12
	C_8	1.15	0.87	0.77	1.17
	N_9	1.42	1.19	1.11	1.31

and on the CO bond of G) where the decomposition $x^{1.96} y^{1.46}$ appears. Thus, this IEHC-calculation favors a structure of the oxygen atoms akin to the classical $di^2\, di\, x^2 z - (s^{1.5} y^{1.5} x^2 z)$ hybridization.

ii) The nitrogens fall into two different categories with NH_2 and NH nitrogens having typical ratios $s^{1.4}(xy)^{2.3}$, and pyridine-like nitrogens staying around the value $s^{1.4}(xy)^{2.7}$. The first type is quite far from the classical tr tr trπ^2 valence-state $(s(xy)^2)$, whereas the second type corresponds roughly to tr$^4\pi(s^{1.3}(xy)^{2.7})$ with, however, no individualization of the trigonal hybrids.

iii) Typical ratios for carbon are $s^{1.2}(xy)^{1.7}$ or $s^{1.1}(xy)^{1.8}$ to be compared to the classical trigonal hybrids $s(xy)^2$.

All "hybridization ratios" show a larger 2s-character than the usual trigonal hybrids.

D) Energies of the Molecular Orbitals

Table 4 gives the calculated energies of all the occupied molecular orbitals obtained.

The following observations can be made on the data:

i) There is a very large intermingling of the σ and π-levels with no appearance of a superficial "π-shell" as was sometimes implicitely admitted in π-electron calculations. Although the details may differ, the interspersing of the σ-levels

Table 4. *Occupied orbitals, sign reversed (eV). Unlabeled orbitals are σ-levels*

U	C	G	A	P
33.09	31.35	32.36	31.56	31.72
29.53	28.56	30.60	29.33	28.44
27.96	26.76	28.73	26.97	26.12
25.98	23.87	25.77	24.94	24.21
23.63	22.89	25.13	23.67	22.78
20.03	18.94	24.84	22.07	19.49
19.14	18.28	21.54	18.86	18.63
17.70	16.92	19.12	18.49	18.06
16.56	16.23	18.67	17.98	16.98
16.33	15.63	18.24	16.81	π 16.04
π 16.29	π 15.25	17.02	15.99	15.60
π 14.53	14.28	16.50	π 15.86	14.58
14.29	13.72	π 16.10	15.34	π 14.27
13.85	π 13.68	15.54	π 14.28	14.23
π 13.70	13.11	15.40	14.20	13.83
12.19	π 12.99	π 14.74	14.13	π 13.44
11.46	11.17	14.68	13.59	13.19
π 11.31	10.86	π 14.14	π 13.44	π 12.42
π 11.08	π 10.54	13.71	13.01	12.04
9.89	π 10.39	12.87	π 12.15	π 11.59
9.43	9.07	π 12.46	11.84	11.33
		11.97	π 11.67	10.53
		π 11.73	11.07	
		11.52	π 10.43	
		π 10.93	10.26	
		10.43		
		π 10.31		
		9.21		

among the π-orbitals seems to be a constant feature in all calculations including all valence-electrons [21]. The same mixing appears also in recent non-empirical SCF calculations made in a Gaussian basis set on pyrrole [24], pyridine [25] and pyrazine [26]. It may be added that this mixing of σ- and π-levels does not occur in the IEHC virtual orbitals in which the π^*-levels are grouped below the σ^*-levels.

ii) The homo (highest occupied molecular orbital) is a σ-orbital, the first π-level coming below. (In uracil there are even two such σ-orbitals above the first π-level.) It seems probable, on the basis of comparison with other all valence electron calculations that this particular result is an artefact of the parametrization of IEHC [21]. Whatever the situation is, it is interesting to look more carefully at the highest occupied π-level. The order of its increasing energy is (in absolute values):

$$\mathrm{G} < \mathrm{C} < \mathrm{A} < \mathrm{U}\,.$$

Thus insofar as the orbital energy gives a measure of the electron-donor ability of a molecule, guanine appears as the best π-donor and uracil as the worst one among the four bases. This result agrees with our previous Pariser-Parr-Pople π-calculations [16] and with the order of the recently measured values of the ionization potentials for the bases [27]:

$$A = \mathrm{C} < \mathrm{T} < \mathrm{U}\,.$$

It is very interesting to remark that in all the carbonyl compounds, (U, C, G), the *σ-homo*('s) are essentially made of the oxygen p_x and p_y orbitals. A similar feature has been noted [18] in the highest σ-molecular orbital obtained by the CNDO-method, which however locates it immediately below the highest π-level. It seems that both in energy and in localization, one may indeed characterize "lone-pairs" of the carbonyl oxygens [21]. The same cannot be said for the nitrogen "lone-pairs" since no molecular orbital appears as made of pure or nearly pure nitrogen atomic orbitals.

Generally speaking it appears that the IEHC method points out to a number of interesting intrinsic features of the electronic structure of the molecules studied. It seems probable that its value and precision may be increased by refining its parametrization. Work along these lines is in progress in our laboratory.

Bibliography

1. Hoffmann, R.: J.chem. Physics **39**, 1397 (1964).
2. Allen, L. C., and J. D. Russel: J. chem. Physics **46**, 1029 (1967).
3. Hoffmann, R.: J. Amer. chem. Soc. **86**, 1259 (1964).
4. Kato, H., T. Yonezawa, K. Morokuma et K. Fukui: Bull. chem. Soc. Japan **87**, 1710 (1964).
5. Viste, A., and H. B. Gray: Inorg. Chem. **3**, 1113 (1964).
6. Cusachs, L. C., and J. W. Reynolds: J. chem. Physics **43**, S 160 (1965).
7. Berthier, G., H. Lemaire, A. Rassat et A. Veillard: Theoret. chim. Acta (Berl.) **3**, 213 (1965).
8. Clementi, E., and D. L. Raimondi: J. chem. Physics **38**, 2686 (1963).
9. Cusachs, L. C.: J. chem. Physics **43**, S 157 (1965).
10. Caroll, D. G., A. T. Armstrong, and S. P. McGlynn: J. chem. Physics **44**, 1865 (1966).
11. Rein, R., N. Fukuda, H. Win, G. A. Clarke, and F. Harris: J. chem. Physics **45**, 4743 (1966).
12. Denis, A., et M. Gilbert: Theoret. chim. Acta (Berl.), in press.
13. A review is given in A. Pullman, and B. Pullman: Adv. in Quantum Chemistry **5** (1968) (in press).
14. Pullman, A., B. Pullman, and G. Berthier: Compt. Rend. Acad. Sci. Paris **243**, 380 (1956).

15. Berthod, H., and A. Pullman: J. Chim. physique **55**, 942 (1965).
16. — C. Giessner-Prettre, and A. Pullman: Theoret. chim. Acta (Berl.) **5**, 53 (1966).
17. Jordan, F., and B. Pullman: Theoret. chim. Acta (Berl.) **9**, 242 (1968).
18. Giessner-Prettre, C., and A. Pullman: Theoret. chim. Acta (Berl.) **9**, 279 (1968).
19. Pullman, A.: International symposium on "Electronic Aspects of Biochemistry". New York Acad. of Sciences, New York, December 1967 (in press).
20. Mulliken, R. S.: J. chem. Physics **23**, 1833 (1955).
21. Pullman, A.: "Yale Seminar on σ-electron calculations", Dec. 1967 (in press).
22. Pople, D. A., and M. Gordon: J. Amer. chem. Soc. **89**, 4253 (1967).
23. De Voe, H., and I. Tinoco jr.: J. Molecular Biol. **4**, 500 (1962).
24. Clementi, E., H. Clementi, and D. R. Davis: J. chem. Physics **46**, 4725 (1967).
25. — J. chem. Physics **46**, 4731 (1967).
26. — J. chem. Physics **46**, 4737 (1967).
27. Lifschitz, C., E. D. Bergmann, and B. Pullman: Tetrahedron Letters **1967**, 4583.

Dr. A. Pullman
Institut de Biologie Physico-Chimique
13 rue Pierre et Marie Curie
Paris V/France