Theoret. chim. Acta (Berl.) 14, 242—249 (1969)

Electronic Aspects of Photodimerization of the Pyrimidine Bases and of their Derivatives

V. I. Danilov, Yuri A. Kruglyak*, V. A. Kuprievich*, and V. V. Ogloblin

Division of Quantum Biology, Institute of Microbiology and Virusology
Academy of Sciences of the Ukrainian SSR, Kiev, USSR

Received December 2, 1968

We present the results of a computation for the lowest excited triplet and singlet states of the pyrimidine bases and of some of their derivatives by the SCF open shell theory and CI method on SCF virtual orbitals. It is shown that excitation to the lowest singlet S_1 and triplet T_1 states mostly affects the bond C_5–C_6 and that triplet excitation is almost completely localized on this bond. The latter effect leads to exceptionally strong weakening of the bond C_5–C_6 in the triplet state T_1. It is concluded that under otherwise equal circumstances the triplet state T_1 presents more favourable conditions for dimerization reaction than the lowest excited singlet state S_1 does. Correlation between the degree of the triplet excitation localization on the bond C_5–C_6 and the easiness of the photodimerization is established. An explanation of the experimental facts concerned with the formation of the thymine free radicals in DNA is also suggested.

Wir präsentieren Resultate von der Berechnung der niedrigsten angeregten Triplett- und Singulett-Zustände der Pyrimidin Basen und einiger ihrer Derivate mittels der SCF-Theorie der offenen Schalen sowie der CI-Methode angewandt auf virtuelle SCF-Orbitale. Es wird gezeigt, daß die Anregung der niedrigsten Singulett S_1- und Triplett T_1-Zustände meistens die C_5 und C_6-Bindung berührt und daß die Triplett-Anregung fast vollständig auf diese Bindung lokalisiert ist. Der letztere Effekt führt zu einer außerordentlich starken Schwächung der C_5–C_6-Bindung im Triplett-Zustand T_1. Es wird daraus geschlossen, daß unter sonst gleichen Umständen der Triplett-Zustand T_1 günstigere Bedingungen für eine Dimerisationsreaktion liefert als der niedrigste angeregte Singulett-Zustand S_1. Ein Zusammenhang zwischen dem Grad der Lokalisierung der Triplett-Anregung auf die C_5–C_6-Bindung und der Photodimerisierbarkeit wird hergestellt. Eine Erklärung der experimentellen Tatsachen, welche die Anordnung der freien Thymin Radikale in der DNA betreffen, wird ebenfalls angegeben.

Nous présentons les résultats d'un calcul pour les états excités singulet et triplet les plus bas des bases pyrimidiques et de certains de leurs dérivés, effectués par les méthodes SCF à couche ouverte et interaction de configurations. L'excitation dans ces plus bas états excités affecte essentiellement la liaison C_5–C_6, l'excitation triplet étant pratiquement totalement localisée sur cette liaison. Ceci conduit à un affaiblissement important de la liaison C_5–C_6 dans l'état triplet. On en conclut qu'à circonstances égales l'état triplet présente des conditions plus favorables pour une réaction de dimérisation que le plus bas état excité singulet. Une corrélation est établie entre le degré de localisation 5–6 de l'excitation triplet et la facilité de photodimérisation. On suggère aussi une explication des faits expérimentaux liés à la formation de radicaux libres thymine dans le DNA.

Introduction

The lethal action of ultraviolet light upon bacteria, phages, and animal cells is mainly caused by the inter- and intrastrand formation of dimers from different pyrimidine bases (PB). Thus the synthesis of nucleic acids is blocked entailing the loss of the reproductive ability. Available experimental [1–3] and theoretical

---

* Permanent address: Quantum Chemistry Group, Division of Theory of Chemical Structure and Reactivity, L. V. Pisarzhevsky Institute of Physical Chemistry, Academy of Sciences of the Ukrainian SSR, Kiev, USSR.

[4–6] data suggest that the PB dimerization in diluted solutions proceed via the lowest triplet state T_1.

In this connection the importance of theoretical studies of the electronic structure of the PB and their derivatives in the ground and lowest excited states is evident. The first attempt was made by Mantione and Pullman [7, 8] who evaluated the properties of the first excited state of the PB and some of their derivatives and established a correlation between the easiness of the photodimerization and the unpaired electron densities on atoms C_5 and C_6[1]. However, the simple LCAO MO method used by these authors does not distinguish between triplet and excited singlet states which is important for the discussion of the dimerization.

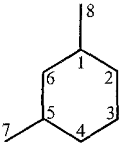

Fig. 1. Numbering of atoms in molecules of the pyrimidine bases and of their derivatives

Calculations by the CI SCF method have shown that the bond order between C_5 and C_6 decreases substantially in the lowest excited singlet S_1 as well as in the triplet T_1 state compared with the corresponding ground state S_0 [9–11]. Nevertheless this procedure led to similar electron densities and bond orders for both excited states so that both types of investigation did not permit to answer the question whether the singlet or the triplet state favours the dimerization more.

In previous papers [4, 6] we have shown by SCF open shell (OS) and CI SCF methods that the lowest triplet state T_1 of thymine presents better conditions for the dimerization than the lowest singlet state S_1 does. Here we present computation results for some excited singlet and triplet states of the PB and some of their derivatives which were obtained by SCF OS or CI SCF methods the latter being restricted to all singly excited configurations.

Details of Computation

The SCF OS and CI SCF methods and the parameters used are given in details in Refs. [12, 13]. In the present work we have used the conjugative model for the methyl group. The computations were performed with two approximations for the two-center Coulomb integrals γ: by Mataga-Nishimoto (MN) and by Ohno.

Electron densities, bond orders $P_{\mu\nu}$, and spin densities ϱ_μ in the CI SCF method were calculated by expressions deduced in [13–15]:

$$ {}^{1,3}P_{\mu\nu} = 2 \sum_{i=1}^{n} C_{\mu i} C_{\nu i} + \sum_{\substack{i\to k\\ j\to l}}^{1,3} X_{i\to k} X_{j\to l} (C_{\mu l} C_{\nu k} \delta_{ij} - C_{\mu i} C_{\nu j} \delta_{kl}) \,, \tag{1} $$

$$ \varrho_\mu = \sum_{\substack{i\to k\\ j\to l}} X_{i\to k} X_{j\to l} (C_{\mu l} C_{\mu k} \delta_{ij} + C_{\mu i} C_{\mu j} \delta_{kl}) \,, \tag{2} $$

---

[1] The numbering of atoms is given in Fig. 1.

where indices i and j denote the occupied, k and l the virtual MO's, $2n$ the number of π-electrons, and $C_{\mu i}$ and $X_{i \to k}$ the expansion coefficients of the MO's on the AO's and of the wave function on the singly excited configurations, resp. The left upper indices indicate the multiplicity of the state.

Results and Discussion

In Table 1 we give the computed transition energies to three excited singlet and triplet states. Those of singlet transitions agree satisfactorily with experimental values determined from the absorption bands of the bases [16, 17]. In the same way, experimental energies of excitation to the state T_1 [17, 18] are well reproduced by theoretical results computed with γ integrals by Ohno. Although all triplet transition energies with the γ's by MN generally are 0.6–0.8 eV lower than those with the γ's by Ohno the relative order of them in a given compound is not altered. It may be noted that, with the same integral system, SCF OS and CI SCF methods lead to similar results.

Table 2 contains the quantities which, in this context, are the most important ones: bond orders P_{56}, spin densities ϱ_5 and ϱ_6, and electron densities P_5 and P_6 calculated with γ integrals by MN[2]. The bond C_5–C_6 in the ground state of all compounds studied appears to be practically double, a conclusion which has been first made by Mantione and Pullman [7] using the simple Hückel method. In the states S_1 or T_1 (see CI SCF results) the bond order P_{56} is substantially decreased, however in considerably less extent in the state S_1 than in T_1. The bond orders according to the CI SCF method show that the triplet excitation is almost completely localized on the bond C_5–C_6 whereas the singlet excitation is "delocalized" on several bonds. As an example for this the bond orders of thymine in the states S_0, S_1, and T_1, computed by both CI SCF and SCF OS methods [4] are given in Table 3 [3]. Both methods give such a substantial weakening of the bond C_5–C_6 in the PB as well as in some of their derivatives that this bond may be considered as broken in the state T_1. Since the dimerization reaction means breakage of the π-bond C_5–C_6 followed by the formation of a cyclobutane ring the state T_1 under otherwise equal circumstances presents more favourable conditions for this reaction than the state S_1 does.

Comparison of P_5 and ϱ_5 as well as P_6 and ϱ_6 in the state T_1 shows that, irrespective of computational methods, the electron density on atoms C_5 and C_6 is mainly influenced by the unpaired electron density. This increases the reactivity significantly because such a T_1-state is a state of valence unsaturation [19].

The strong localization of the triplet excitation leads to the important fact that the free valence indices [4, 6] on atoms C_5 and C_6 of thymine in the state T_1 are extremely high: 1.177 for C_5 and 1.316 for C_6. This means that these atoms are highly reactive with respect to the attack by free radicals and explains the experimental fact that a H atom is added to atom C_6 of the thymine rest when DNA is irradiated by UV light [20]. Moreover, free valence indices on these atoms are

---

[2] It should be noted that computation of these properties with the γ's of Ohno leads to practically the same results.

[3] One can also see from Tables 2 and 3 that the SCF OS method leads to a somewhat stronger localization of the triplet excitation on the bond C_5–C_6.

Table 1. *Computed excitation energies of the pyrimidine bases and of their derivatives to three lowest excited singlet and triplet states (in eV)*[a]

State		Uracil	6-Methyl-uracil	Thymine	Orotic acid	5-Amino-uracil	Cytosine	5-Methyl-cytosine	2-Thio-thymine	Iso-cytosine	5-Nitro-uracil[b]	6-Aza-thymine
S_1	*a*	5.305 (5.260)	5.045 (5.057)	5.048 (4.982)	4.919 (4.931)	4.626 (4.623)	4.496 (4.608)	4.423 (4.522)	4.790 (4.766)	4.787 (4.747)	4.243 (4.607)	4.723 (4.491)
S_2	*a*	6.017 (6.686)	5.888 (6.585)	5.754 (6.440)	6.054 (5.973)	5.341 (5.809)	5.834 (6.101)	5.642 (6.002)	5.289 (5.520)	5.304 (5.562)	4.797 (5.191)	5.873 (6.235)
S_3	*a*	6.603 (6.754)	6.510 (6.751)	6.584 (6.587)	6.126 (6.259)	6.688 (6.115)	6.590 (6.731)	6.346 (6.400)	5.919 (6.227)	5.724 (6.127)	5.269 (5.566)	6.079 (6.325)
T_1	*a*	1.882 (2.633)	1.889 (2.614)	1.823 (2.517)	1.729 (2.711)	1.567 (1.970)	2.026 (2.721)	1.969 (2.612)	1.939 (2.677)	1.868 (2.651)	0.075 (2.144)	1.289 (1.872)
	b	2.010 (2.713)	2.023 (2.696)	1.937 (2.572)	2.056 (2.786)	1.641 (1.969)	2.206 (2.817)	2.014 (2.686)	2.163 (2.767)	1.984 (2.789)		1.318 (1.875)
T_2	*a*	3.301 (4.119)	3.305 (4.130)	3.287 (4.095)	3.200 (4.118)	3.266 (4.001)	3.122 (3.907)	3.075 (3.864)	2.781 (3.565)	2.652 (3.319)	1.653 (2.485)	3.136 (3.896)
T_3	*a*	3.477 (4.364)	3.479 (4.375)	3.450 (4.333)	3.321 (4.602)	3.438 (4.267)	3.962 (4.539)	3.929 (4.501)	3.932 (4.749)	3.702 (4.414)	2.949 (3.919)	3.374 (4.150)

[a] Values in parenthesis are computed with the γ's by Ohno, otherwise – by MN; a = CI SCF method, b = SCF OS method.

[b] Results for 5-nitrouracil by SCF OS method are omitted because for the reason of output unreprodisibil we are not sure that in the case of this molecule process of self-consistency converges to the state T_1.

Table 2. *Bond order P_{56}, electron densities P_5, P_6, and spin densities ϱ_5, ϱ_6 in the states S_0, S_1, and T_1 of the pyrimidine bases and of their derivatives*[a]

Property state	Uracil	6-Methyl-uracil	Thymine	Orotic acid	5-Amino-uracil	Cytosine	5-Methyl-cytosine	2-Thio-thymine	Iso-cytosine	5-Nitro-uracil	6-Aza-thymine
P_{56} S_0	0.900	0.870	0.875	0.886	0.861	0.846	0.825	0.876	0.890	0.812	0.760
S_1 *a*	0.444	0.407	0.422	0.422	0.479	0.553	0.522	0.597	0.663	0.509	0.405
T_1 *a*	0.191	0.201	0.189	0.206	0.184	0.307	0.283	0.260	0.365	0.827	0.148
b	0.085	0.100	0.106	0.119	0.134	0.144	0.165	0.314	0.388		0.107
ϱ_5 T_1 *a*	0.684	0.676	0.628	0.654	0.575	0.487	0.469	0.553	0.477	0.005	0.677
b	0.822	0.820	0.723	0.787	0.616	0.664	0.561	0.254	0.169		0.735
ϱ_6 T_1 *a*	0.775	0.704	0.783	0.709	0.804	0.681	0.702	0.713	0.624	0.113	0.754
b	0.875	0.802	0.871	0.811	0.864	0.842	0.837	0.768	0.722		0.778
P_5 S_0	1.064	1.080	1.047	1.014	1.044	1.090	1.072	1.006	1.010	1.080	0.884
S_1 *a*	0.796	0.784	0.838	0.787	0.872	0.964	0.940	0.759	0.874	0.815	0.891
T_1 *a*	0.917	0.916	0.922	0.919	1.008	0.936	0.926	0.917	0.919	1.082	0.931
b	0.926	0.923	0.928	0.888	1.021	0.924	0.920	0.908	0.958		0.933
P_6 S_0	0.942	0.927	0.958	0.973	1.037	0.901	0.912	0.994	0.979	0.861	1.192
S_1 *a*	1.038	1.015	1.051	0.932	0.990	1.126	1.121	0.749	0.829	0.931	1.294
T_1 *a*	1.092	1.078	1.092	1.049	1.093	1.110	1.113	1.112	1.113	0.971	1.185
b	1.086	1.074	1.085	1.046	1.035	1.102	1.101	1.106	1.118		1.197

[a] See footnotes to Table 1.

much higher in the state T_1 than in the state S_1 what explains why the first triplet state T_1 is considered as a precursor of free radical formation [20].

In accordance with the SCF OS results the compounds studied here can be arranged (Table 2) in the order of increasing P_{56} and simultaneously of decreasing ϱ_5 and ϱ_6, that is, in the order of increasing excitation localization on the bond C_5–C_6. There appear to be three groups[4]. The first one containing uracil, 6-methyluracil, thymine, and orotic acid is known to be easily dimerizable [21]. The second group contains 5-aminouracil, cytosine, and 5-methylcytosine which dimerize not so easy [21]. In the third group we have 2-thiothymine (thiol-form), isocytosine, 5-nitrouracil, and 6-azathymine which do not dimerize at all or only with considerable difficulty. There are no quantitative data on the isocytosine photodimerization in the literature except the note that isocytosine dimerizes to a smaller extent than 6-methyluracil and uracil [21].

Table 3. *Bond orders $P_{\mu\nu}$ in the states S_0, S_1, and T_1 of thymine computed by* CISCF *and* SCFOS *methods*[a]

$\mu\nu$	S_0	S_1	T_1
12	0.317 (0.325)	0.335 (0.288)	0.311 (0.305)
16	0.317 (0.293)	0.446 (0.430)	0.324 (0.310)
23	0.321 (0.319)	0.280 (0.336)	0.317 (0.320)
27	0.846 (0.843)	0.816 (0.845)	0.849 (0.851)
34	0.306 (0.308)	0.212 (0.192)	0.273 (0.269)
45	0.322 (0.333)	0.458 (0.410)	0.447 (0.449)
48	0.841 (0.836)	0.646 (0.525)	0.724 (0.778)
56	0.875 (0.903)	0.422 (0.424)	0.189 (0.106)

[a] The latter are given in parenthesis. Small difference between the results for the ground state S_0 is connected with different description of methyl group (see [4]).

Our results for isocytosine are essentially different of those from Ref. [7]. We obtained a smaller localization of the unpaired electron densities on atoms C_5 and C_6 for isocytosine than for the compounds of the first and second group because of increasing densities on atom O ($\varrho_O = 0.781$ with the γ's by MN and $\varrho_O = 0.649$ – by Ohno). Mantione and Pullman [8] have not found such a redistribution and believed therefore that the degree of dimerization of isocytosine is intermediate. Our results are also different from [8] with respect to 2-thiothymine. Despite the fact that both computations show no indication towards dimerization

---

[4] Distribution of the bases among groups according to CI SCF results leads to a poorer agreement with available experimental data [21].

there is nevertheless an important difference between them. Considerable decrease of ϱ_5 is connected in Ref. [8] with the concentration of the unpaired electron density on atom S, in our computation, however, with the concentration on atom O ($\varrho_O = 0.649$).

Thus these considerations show a correlation between the easiness of dimerization of the PB and their derivatives on the one hand and the unpaired electronic densities on atoms C_5 and C_6 on the other, as well as between the easiness of photodimerization and the bond order of the bond C_5–C_6.

In the third group 5-nitrouracil is the most stable towards the action of UV light. Its triplet excitation is almost completely localized at the nitrogroup. Thus, $\varrho_N = 0.315(0.416)$, $\varrho_O = 0.845(0.750)$, and $\varrho_O = 0.818(0.652)$ where the first values arise from the SCF OS theory with the γ's by MN and the values in brackets from the CI SCF method with the γ's by Ohno. Because of this the bond order P_{56} is not changed when a transition from S_0 to T_1 occurs. This seems to be the reason why nitrouracil does not dimerize. Mantione and Pullman [7] have received similar results for unpaired electron densities by the Hückel method considering the first excited state as formed by a transition from the highest occupied MO φ_7 to the lowest empty φ_8. However, our more sophisticated CI computations show that the configuration $\phi_{7\to 8}$ gives the main contribution to the state T_2 but not to T_1. Thus, the resemblance between our results and those from Ref. [7] is accidental.

We separated 6-azathymine putting it as the last compound in Table 2 because its UV stability does not behave as regularly as in the other cases. Neither our results nor those of Mantione and Pullman [7] can explain its lack of dimerization. It may happen that the N atom in thymine with its pair of n-electrons gives rise for new singlet and triplet n-$\pi^\star$-transitions which are located lower than the corresponding π-$\pi^\star$-transitions. As a result an excitation may lead to a n-$\pi^\star$-triplet state without localization of the excitation at the bond C_5–C_6.

Our results for the bond order P_{56} in the PB are qualitatively different from those published by Imamura *et al.* [9] and Pullman *et al.* [10, 11]. In the first case the difference is easily explained: formula (1) is different from that used in [9] and converts to the latter only when the summation in (1) is restricted by the condition $i = j$ and $k = l$, i.e. when many terms are rejected. It should be noted that the singlet excited states have been computed in [9] including also a small number of doubly excited configurations. In this case too the formula for $P_{\mu\nu}$ used in [9] is

Table 4. *Bond order P_{56} in the states S_1 and T_1 of the pyrimidine bases computed by* CI SCF *method by different authors (with the γ's by* MN*)*

Computed by	Thymine		Uracil		Cytosine	
	S_1	T_1	S_1	T_1	S_1	T_1
us, formula (1)	0.422	0.189	0.444	0.191	0.553	0.307
us, formula (1) with $i = j$ and $k = l$	0.355	0.364	0.365	0.363	0.526	0.532
Imamura *et al.* [9]	0.459	0.457	0.401	0.480	0.588	0.611
Pullman *et al.* [10, 11]	0.411	0.451	0.413	0.449	0.665	0.826

not correct which becomes clear if one compares it with the full expression given in [14, 15]. (Calculation formula for $P_{\mu\nu}$ and ϱ_μ in [10, 11] are not given.)

In Table 4 we compare the results of Ref. [9–11] with the values of P_{56} calculated by us as well with Eq. (1) as with the incomplete one. It can be seen that the use of the latter for P in fact leads to practically idential results for ${}^1P_{56}$ and ${}^3P_{56}$, which is also characteristic for results of [9][5]. We assume therefore that the authors of [10, 11] have also used the expressions for ${}^{1,3}P_{\mu,\nu}$ and ϱ_μ given in [9].

Acknowledgements. Authors are grateful to Dr. G. G. Dyadyusha and Mr. O. V. Shramko for usefull discussion of present results.

References

1. Beukers, R., and W. Berends: Biochim. biophysica Acta **49**, 181 (1961).
2. Lamola, A. A., and J. P. Mittal: Science **154**, 1560 (1966).
3. —, and T. Yamane: Proc. nat. Acad. Sci. USA **58**, 443 (1967).
4. Danilov, V. I.: Photochem. Photobiol. **6**, 233 (1967).
5. Dyadyusha, G. G., V. I. Danilov, and O. V. Shramko: Molekulyarnaya Biologiya (USSR) **1**, 539 (1967).
6. Danilov, V. I., O. V. Shramko, and G. G. Dyadyusha: Biofizika (USSR) **12**, 544 (1967).
7. Mantione, M. J., and B. Pullman: Biochim. biophysica Acta **91**, 387 (1964).
8. Pullman, B., and M. J. Mantione: Biochim. biophysica Acta **95**, 668 (1965).
9. Imamura, A., H. Fujita, and C. Nagata: Bull. chem. Soc. Japan **40**, 21 (1967).
10. Berthod, H., C. Giessner, and A. Pullman: Int. J. quant. Chem. **1**, 123 (1967).
11. Denis, A., and A. Pullman: Theoret. chim. Acta (Berl.) **7**, 110 (1967).
12. Kruglyak, Yu. A., G. G. Dyadyusha, V. A. Kuprievich, L. M. Podolskaya, and G. I. Kagan: Methods of computation of electronic structure and spectra of Molecules, ch. 3. Kiev: Publishing House "Naukova Dumka" 1969.
13. — On calculation of electronic structure of conjugated molecules and radicals by Self-Consistent Field and Configuration Interaction methods. In: Kayushin, L. P., M. K. Pulatova, and K. M. Lvov: Investigation of paramagnetic centers in γ-irradiated proteins. Moscow: Publishing House "Nauka" 1969.
14. Danilov, V. I., Yu. A. Kruglyak, and V. I. Pechenaya: Theoret. i Eksperim. Khim. (USSR) (in press).
15. — — — Theoret. chim. Acta (Berl.) **13**, 288 (1969).
16. Clark, L., and I. Tinoco, Jr.: J. Amer. chem. Soc. **87**, 11 (1965).
17. Kleinwachter, V., J. Drobnik, and L. Augenstein: Photochem. Photobiol. **5**, 579 (1966).
18. Longworth, J. W., R. O. Rahn, and R. G. Shulman: J. chem. Physics **45**, 2930 (1966).
19. Terenin, A. N.: Photochemistry of dyes and parent organic compounds. Moscow: Publishing House of the Academy of Science of the USSR 1947.
20. Rahn, R. O., R. G. Shulman, and J. W. Longworth: J. chem. Physics **45**, 2955 (1966).
21. Wacker, A.: In: Progress in nucleic acid research, Vol. 1, p. 369. New York: Academic Press 1963.

Assoc. Prof. Yuri A. Kruglyak
Quantum Chemistry Group
Institute of Physical Chemistry
Academy of Sciences of the Ukrainian SSR
Bol'shaya Kitaevskava Street 97
Kiev-28, USSR

---

[5] The difference between results obtained by different authors is obviously connected with different choise of integrals.