BULLETIN OF THE CHEMICAL SOCIETY OF JAPAN VOL. 40 522—526 (1967)

Semi-empirical Self-consistent Field Molecular Orbital Calculation of the Electronic Structure of the Base Pairs of Deoxyribonucleic Acid

Akira IMAMURA, Hiroko FUJITA and Chikayoshi NAGATA
Biophysics Division, National Cancer Center Research Institute, Tsukiji, Chuo-ku, Tokyo

(Received September 7, 1966)

The electronic structures of adenine-thymine and guanine-cytosine pairs of DNA were calculated using the semi-empirical self-consistent field molecular orbital method. The calculations were carried out for two models, that is, one including hydrogen $2p\pi$ orbitals and the other not. The obtained π electron densities and the π bond orders were then discussed in comparison with those of the corresponding bases. Further, the π electronic stabilization energies due to the hydrogen bond formations were calculated; the π stabilization energy of the G-C pair was found to be larger than that of A-T pair, in agreement with the experimental data. The π electronic structures in the excited states were discussed in connection with the biological behavior of DNA, that is, the misduplication of DNA leading to mutation by irradiation.

Recently, the electronic structure of the base components of nucleic acids has been studied extensively by using molecular orbital methods; the structure has been studied in connection with their physico-chemical properties and biological actions.[1-3] However, the base pairs of deoxyribonucleic acid have not been so extensively investigated. Pullman and Pullman did do calculations using the Hückel MO method,[4] and later Hoffmann and Ladik carried out calculations of the base pairs, taking hydrogen $2p\pi$ orbitals into consideration, by the Hückel MO method.[5] As is often mentioned, however, the Hückel MO method is not reliable for such heterocyclic molecules as the base components of DNA, since the results obtained by Hückel MO method depend upon the choice of parameters used in the calculations. Self-consistent field molecular orbital calculations have been carried out by Rein and Ladik[6] for the guanine-cytosine pair; also discussed the relation between the possibility of mutation. Further, Rein and Harris,[7] in order to obtain the double-well potential for the proton in the hydrogen bond, calculated the electronic structure of the guanine-cytosine pair, taking the σ electrons localized in the middle of the hydrogen bond into consideration explicitly. The adenine-thymine pair, on the other hand, was not treated, probably because of the difficulty of estimating the effect of the hyperconjugation of methyl group in thymine.

In the present paper, the electronic structures of the base pairs of DNA, that is, guanine-cytosine and adenine-thymine, were calculated by the semi-empirical self-consistent field molecular orbital method, in which two models were used, one including the hydrogen $2p\pi$ orbital in the linear combination of the atomic orbitals, and the other not. Hereafter, the latter will be referred to as model 1, and the former, as model 2. The electronic structures obtained in the ground and the excited states and the stabilization energies due to the hydrogen bond formations will be discussed in comparison with those of the base components.

Method

The calculation has been carried out for the π electron systems of the base pairs according to the semi-empirical SCF-LCAO-MO method. The one-center Coulomb repulsion integrals were estimated using the Pariser-Parr approximation,[8] while with regard to the two-center Coulomb repulsion integrals the Mataga-Nishimoto approximation was used.[9] The differential overlaps and the penetration integrals were neglected. The core integrals were evaluated according to the Goeppert-Mayer and Sklar potential,[10] and the resonance core integrals were given as in Eq. (1):

---

[1] A. Veillard and B. Pullman, *J. Theoret. Biol.*, **4**, 37 (1963).

[2] R. K. Nesbet, *Biopolymers Symposia*, **1**, 129 (1964).

[3] C. Nagata, A. Imamura, Y. Tagashira and M. Kodama, This Bulletin, **38**, 1638 (1965).

[4] A. Pullman and B. Pullman, *Biochim. Biophys. Acta*, **36**, 343 (1959).

[5] T. A. Hoffmann and J. Ladik, "The Structure and Properties of Biomolecules and Biological Systems," ed. by J. Duchesne, Interscience, New York (1964), p. 108.

[6] R. Rein and J. Ladik, *J. Chem. Phys.*, **40**, 2466 (1964).

[7] R. Rein and F. E. Harris, *ibid.*, **41**, 3393 (1964).

[8] R. Pariser and R. G. Parr, *ibid.*, **21**, 466 (1953).

[9] N. Mataga and K. Nishimoto, *Z. Physik. Chem., Neue Folge*, **13**, 140 (1957).

[10] M. Goeppert-Mayer and A. L. Sklar, *J. Chem. Phys.*, **6**, 645 (1938).

$$H_{rs} = (K/2)S_{rs}(I_r + I_s) \qquad (1)$$

where H_{rs} designates the core resonance integral between rth and sth atomic orbitals; S_{rs}, the corresponding overlap integral, and I_r, the ionization potential of the rth atomic orbital. K is the parameter that was introduced by Wolfberg and Helmholtz[11]; it was assumed to be unity.

TABLE 1. SEMI-EMPIRICALLY EVALUATED VALUES OF EFFECTIVE NUCLEAR CHARGE, Z^*, IONIZATION POTENTIAL, I_r AND ONE CENTER COULOMB REPULSION INTEGRAL, $(rr \mid rr)$

r	Z^*	I_r (eV)	$(rr \mid rr)$ (eV)
C	3.25	−11.16	11.13
O	4.55	−17.28	15.27
N*(azo)	3.90	−14.12	12.34
N*(amino)	4.25	−28.53	16.57
H_3**	2.83	−10.256	9.326
$H_{2p\pi}$	0.3625	− 1.757	1.241

* Azo nitrogen donates a π electron to the system, and amino nitrogen two π electrons.

** Hyperconjugation in the methyl group is taken into account, and the group orbital of H_3 is assumed to overlap with π-orbitals.

The values of various integrals tabulated in Table 1 are the same as those previously evaluated by the present authors,[3] except for the hydrogen 2pπ orbital. Most of these values were obtained from the paper of Hinze and Jaffé.[12] With regard to the methyl group, however, we used the value of Morita,[13] who treated the benzenium ion by the semi-empirical SCF MO method, taking the effect of hyperconjugation into account. The integrals concerning the hydrogen 2pπ orbital were taken from our previous paper[14] on the assumption that the net charge on hydrogen is 0.25.

The data of the atomic distances were taken from the papers of Spencer[15] and Arnott *et al.*[16] The oscillator strengths of the transition were calculated by using Eq. (2):

$$f = 1.085 \times 10^{-5} \sigma_{0b} \mathbf{M}^2_{0b} \qquad (2)$$

where σ_{0b} is the wave number of the transition and $\mathbf{M}_{0b}$ is the corresponding transition moment. The stabilization energy due to the hydrogen bond formation is the difference between the sum of the total energies of the component bases and the total energy of the base pair. The total energy of a molecule is given as follows:

$$E = \sum_{i=1}^{\text{occ}} (H_i + \varepsilon_i) + E_{rep} \qquad (3)$$

where ε_i is the ith orbital energy and H_i is the ith diagonal element of the core Hamiltonian. E_{rep} denotes the nuclear repulsion energy, which is given as follows:

$$E_{rep} = \sum_{r} \sum_{\substack{s \\ (r \neq s)}} n_r n_s (rr \mid ss) \qquad (4)$$

in which $(rr \mid ss)$ are the two-center Coulomb repulsion integrals between the rth and sth atomic orbitals and where n_r is the number of electrons which the rth atomic orbital donates to the π electron system.

Results and Discussion

The π Electron Densities. The π electron densities of the A-T and G-C pairs are shown in Fig. 1, where they are compared with those of each base. The alteration of the π electron densities due to the base-pair formation is small, on the whole. On the proton-donating atoms, the π electron densities decrease in magnitude upon the base-pair formation; on the other hand, the π electron densities on the proton-accepting atoms increase. The calculation in model 2, in which the hydrogen 2pπ orbitals were taken into account, shows that the π electron densities on the atom participating in the hydrogen-bonding formation decrease compared with those of the calculation in model 1, in which the hydrogen 2pπ orbitals were not included. The alteration in the π electron densities upon the formation of a base pair is larger in the G-C pair than in the A-T pair in both models. Moreover, in model 2 the π electron densities on the hydrogen 2pπ orbital is larger in the G-C pair than in the A-T pair. These results are interesting in view of the fact that the G-C pair is more stable than the A-T pair.

The π Bond Orders. The π bond orders of A-T and G-C pairs are shown in Fig. 2, together with those of each base. The alteration of the π bond orders upon the base-pair formations is also small, on the whole. However, the bonds which are concerned with the hydrogen bond undergo rather a remarkable change in their bond orders. For instance, in C=O, C–NH_2 and C–NH (hydrogen bonding nitrogen of the purine ring) bonds of guanine, the changes in the π bond orders upon the base-pair formation are larger than in other bonds. Further, it is worth noticing that the change in the π bond orders in the G-C pair is larger than that in the A-T pair. This tendency

---

[11] M. Wolfsberg and L. Helmholtz, *ibid.*, **20**, 837 (1952).

[12] J. Hinze and H. H. Jaffé, *J. Am. Chem. Soc.*, **84**, 540 (1962).

[13] T. Morita, This Bulletin, **33**, 1486 (1960).

[14] A. Imamura, M. Kodama, Y. Tagashira and C. Nagata, *J. Theoret. Biol.*, **10**, 356 (1966).

[15] M. Spencer, *Acta Cryst.*, **12**, 66 (1959).

[16] S. Arnott, M. H. F. Wilkins, L. D. Hamilton and R. Langridge, *J. Mol. Biol.*, **11**, 391 (1965).

[17] J. Marmur and P. Doty, *Nature*, **183**, 1427 (1959).

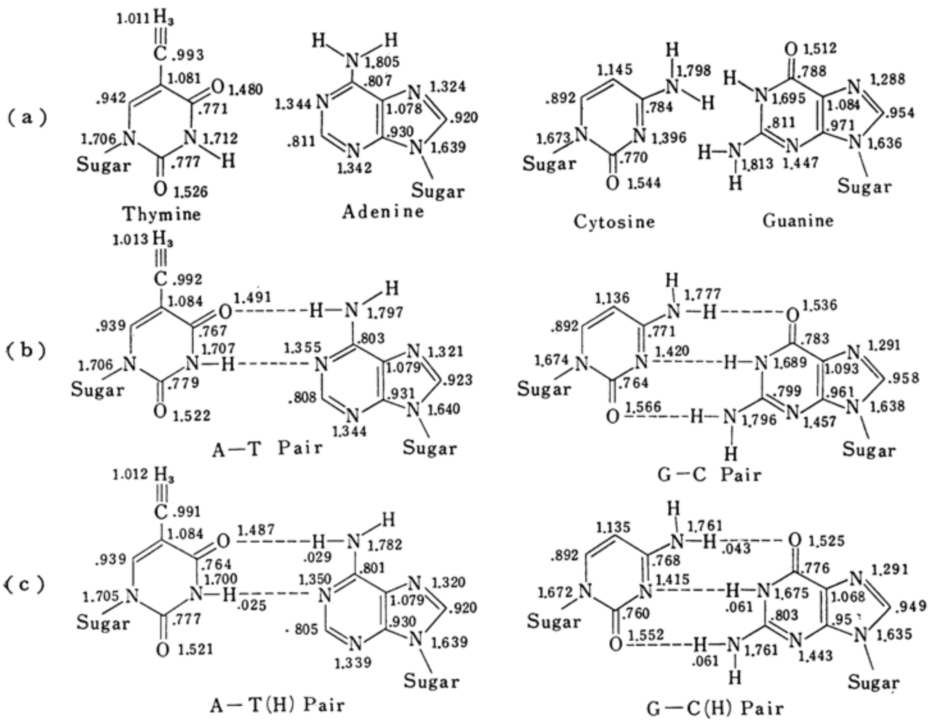

Fig. 1. The π electron densities of the base pairs and the base components of DNA.
(a) base component (b) base pair in model 1 (c) base pair in model 2

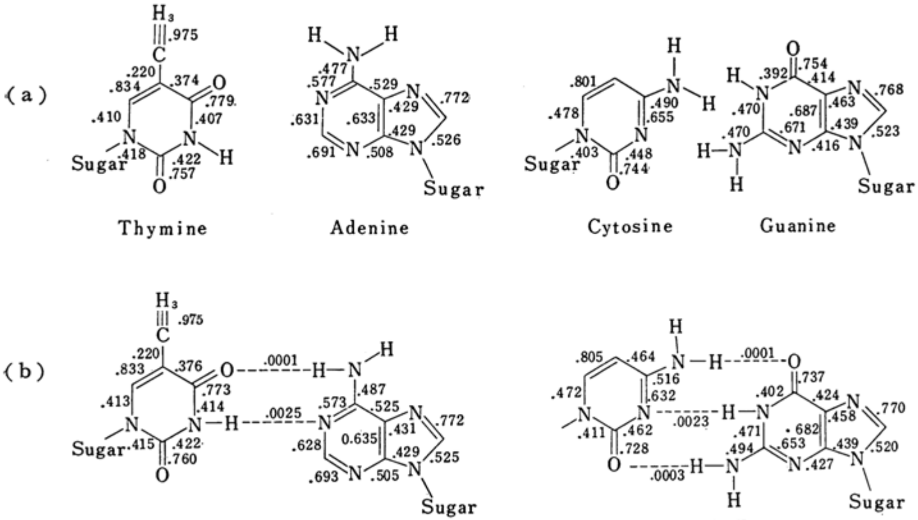

Fig. 2. The π bond orders of the base pairs and the base components of DNA.
(a) the base component (b) the base pair in model 1

was also true for the changes in the π electron densities upon the base-pair formation. The bond orders of the hydrogen bonds are also worth making reference to; the bond orders of the NH-N hydrogen bonds which combine directly the purine ring with the pyrimidine ring are larger than those of other hydrogen bonds in both the A-T and G-C pairs. This result indicates that the NH-N hydrogen bond contributes to stabilizing the base pair more than the other hydrogen bond does as far as the π electrons are concerned.

The π Electronic Stabilization Energy by the Hydrogen Bond Formation. As was described above, the stabilization energies by the

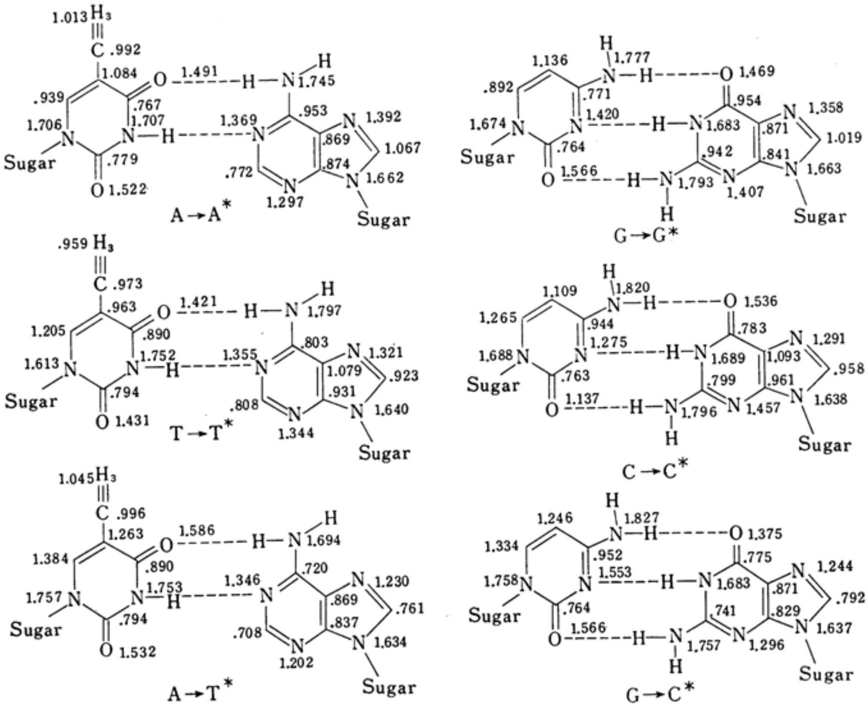

Fig. 3. The π electron densities of the base pairs in the excited states.
A→A* means the excitation localized in adenine base and A→T* the charge transfer excitation from adenine to thymine.

TABLE 2. THE π ELECTRONIC STABILIZATION ENERGIES BY THE BASE PAIR FORMATION

Compound	Total energy eV	Stabilization energy eV
Adenine	−201.18	—
Thymine	−205.78	—
A-T (model 1)	−407.01	0.05 (1.2 kcal/mol)
A-T (model 2)	−407.52	0.56 (13.0 kcal/mol)
Guanine	−246.64	—
Cytosine	−173.55	—
G-C (model 1)	−420.56	0.37 (8.5 kcal/mol)
G-C (model 2)	−421.45	1.26 (28.9 kcal/mol)

hydrogen bond formation can easily be calculated by using Eq. (3). The values of the stabilization energies are tabulated in Table 2, together with the total π electronic energies of the base components for both cases, that is, in models 1 and 2. It is experimentally well known that the G-C pair is more stable than the A-T pair,[17] so that these calculated values agree well with the experimental values if the contributions of σ electrons to the stabilization energy are the same in both the A-T and G-C pairs. In other words, the fact that the G-C pair is more stable than the A-T pair can be explained in terms of the difference between the π electronic stabilization energies of the G-C and A-T pairs. The absolute values of the π electronic stabilization energies in the calculation in model 1 seem to be of a reasonable magnitude; on the other hand, the values in model 2 are rather large considering that the values contain only the contribution of π electrons. This discrepancy, therefore, is mainly due to an inappropriate estimation of the values of the integrals concerning the hydrogen 2pπ orbital.

The π Electronic Structures in the Excited States. Rein and Ladik have calculated the electronic structure of the G-C pair and pointed out the possibility, judging from their calculation, that the lowest excited state in which the charge transfer from guanine to cytosine occurs may be related to the mutagenic effect of the radiation.[6] In our calculation also, the charge-transfer-excited state is the lowest excited state in the G-C pair in which the charge transfers from guanine to cytosine; on the other hand, in the A-T pair the charge-transfer-excited state is not the lowest excited state.[18] However, in the G-C pair the oscillator strength of this transition is less than 10^{-5}, so that in our calculation the direct excitation from the ground state to the charge transfer

---

[18] In the present calculation, the energy levels can be classified into two groups, one belonging to one base and the other belonging to the other base, from the coefficients of the atomic orbitals in a molecular orbital since the delocalization of the π electrons between bases is very small.

excited state can hardly occur. Accordingly, only when the probability of the transition from a higher excited state to the lowest charge-transfer-excited state *via* a non radiative process is not of a negligible order, may the charge-transfer-excited state in the G-C pair be responsible for the mutation by irradiation. That is, in the charge-transfer-excited state the probability of the transfer of the hydrogen bonding proton from one site to the other increases so that the misduplication of the DNA may occur more frequently, thus leading to the mutation.(19,20) Rein and Ladik have obtained the oscillator strength of the charge-transfer transition in the G-C pair as 0.004,(6) which is not a negligible value. The discrepancy between their value and ours may be due to the choice of the integral; that is, in their calculation the core Coulomb integral of the hydrogen bonding atom was evaluated after Pullman and Pullman,(4) who estimated these integrals tentatively considering the effect of the σ electrons in the hydrogen bond. At present, it is difficult to determine whether this estimation of the integral is reasonable or not. However, their value, which is 0.2β lower than the usual integral in the proton-accepting atom and 0.2β higher than the usual integral in the proton-donating atom, seems to be rather overestimated.

Now the other possibility of the occurrence of the mutation by irradiation can be considered. The excitation within the same base also changes the electronic structures of the base pair, and this may lead to the mutation. From this point of view, the π electron densities of the excited states are indicated in Fig. 3 in both cases, that is, the charge-transfer-excited states and the excited states in which the excitation occurs within a base. As is shown in Figs. 1 and 3, in the A-T pair the excitation related to adenine or thymine may increase the possibility of the proton transfer since the π electron density of proton-donating atoms such as the amino group of adenine becomes smaller in the excited state, causing the hydrogen of this group to become more positive, while the π electron densities of the proton-accepting atoms such as carbonyl oxygen in thymine increase in magnitude. On the other hand, the situation is reversed for the G-C pair; that is, the possibility of the proton transfer in the excited state is less than in the ground state. Therefore, it may be said from this calculation, that the proton transfer can occur *via* the charge-transfer-excited state in the G-C pair, but in the A-T pair it occurs *via* the excited state related to the adenine or the thymine base only.

---

19) C. Nagata, A. Imamura, K. Fukui and H. Saito, *Gann*, **54**, 401 (1963).

20) J. Ladik, *J. Theoret. Biol.*, **6**, 201 (1964).