Theoret. chim. Acta (Berl.) 25, 372–381 (1972)

Theoretical Study of the Photodimerization of Pyrimidine Bases

J. BERTRAN

Instituto de Biologia Fundamental, Universidad Autónoma de Barcelona,
Avd. San Antonio Mª Claret 171, Barcelona-13, Spain

Received January 7, 1972

The experimental results are correctly interpreted as from the delocalized model of the transition state and the theory of perturbations. The surfaces of the energy of interaction between both molecules in function of different coordinates of reaction make it possible to discuss the concerted nature of photodimerization and permit the two conjugated systems to be added between carbon atoms 5 and 6 of both molecules.

Die experimentellen Resultate werden mit Hilfe eines delokalisierten Modells des Übergangszustandes und der Störungstheorie richtig interpretiert. Die Energiehyperfläche der Wechselwirkung zwischen den beiden Molekülen als Funktion verschiedener Reaktionskoordinaten erlaubt die Diskussion des konzertierten Ablaufs der Photodimerisation und der Addition der beiden konjugierten Systeme, die zwischen den Kohlenstoffatomen 5 und 6 der beiden Moleküle stattfindet.

Les résultats expérimentaux sont correctement interprétés à partir du modèle delocalisé de l'état de transition et de la théorie des perturbations. Les surfaces de l'énergie d'interaction entre les deux molécules en fonction aux différentes coordonnées de réaction permettent de discuter le caractère concerté de la photodimerisation et que l'addition des deux systèmes conjugués se réalise entre les carbones 5 et 6 des deux molécules.

1. Introduction

The research activities in photochemistry of the pyrimidine bases of the DNA during the last few years can be seen in frequent review articles [1–5]. From the results thus obtained, one can cite the following.

The photodimerization of the thymine, discovered some ten years ago [6] has also been observed in the other pyrimidine bases, uracil [7, 8] and cytosine [9, 10].

The formation of mixed dimers has also been observed [9, 11].

Experimentation with DNA of different composition has shown the following relative efficiency in the formation of dimers: $\widehat{\mathrm{TT}} > \widehat{\mathrm{CT}} > \widehat{\mathrm{CC}}$ [12].

Theoretically, four conformations of final products are possible (Fig. 1). In this study, the nomenclature *cis-syn*, *trans-syn*, *cis-anti*, and *trans-anti* has been adopted to designate them [13]. In certain conditions, the four conformations are obtained [14, 15], whilst in other cases, the *cis-syn* conformation is predominant [16, 17].

The analysis by *x*-ray diffraction techniques has permitted the determination of the geometry of the dimers [18–23]. In the *cis-syn* conformation, the planes of both bases are not parallel, the bases have turned around the bonds 5–6, opening themselves, furthermore, the cyclobutane formed in the dimerization

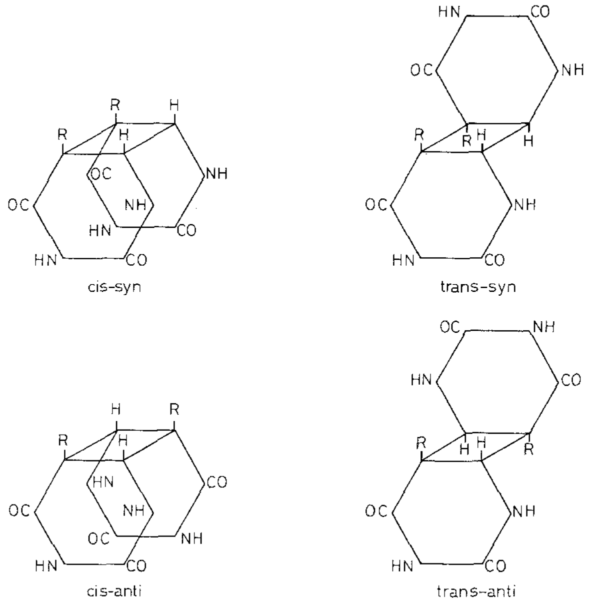

Fig. 1. Possible conformations of thymine and uracil dimers ($R = CH_3$ thymine dimers; $R = H$ uracil dimers)

is twisted by a turn of some 28° around the axis which unites the two midpoints of the 5–6 bonds.

The reaction seems to be concerted with both new bonds, closing itself simultaneously [24]. The bimolecular process results when one base, in a fundamental electronic state, interacts with another, in its first excited state. A metastable dimer in an excited state appears to be the intermediate product [25].

$$P + P^* \xrightarrow[a]{} \widehat{PP}^* \xrightarrow[b]{} \widehat{PP}\,.$$

One supposes that the determining step of the velocity of reaction is the *a*. The intervention of the first singlet excited state or the first triplet state depends upon the conditions of reaction [15]. In diluted solutions, one deals, without doubt, with the first excited triplet state, due to its longer life span [26–28]. In a solid phase and in concentrated solutions, the intervention of the first excited singlet state seems assured [15, 29, 30].

Following the pioneering works of Pullman and Mantione [31, 32] the reaction of the photodimerization of the pyrimidine bases has been interpreted from the diminishing with the excitation of the bond index between the carbons 5 and 6 [33–36]. This carries a notable increase of the free valence above these centres in the first excited state. This static index is characteristic for reactions of addition. Also, one has underlined the high spin density about the carbons 5 and 6 in the first excited state. This corresponds to grand coefficients in these centres in the semioccupied orbitals. The electronic excitation is found to be highly localized in the double bond between the carbons 5 and 6, which are the points of photodimerization.

Other theoretic works try to interpret the photodimerization of the pyrimidine bases by evaluating the energy of interaction between the monomers through the perturbation theory, limiting itself to the frontier orbitals [37], or including all the orbitals π, and the overlaps, explicitly [38].

In this present work, one is trying to interpret the experimental facts previously enumerated from the dynamic indices. One adopts the delocalized model of the transition state, and one studies the interaction of one base in a fundamental state with another in an excited state, at the beginning of the reaction, through the theory of perturbations. By always using a monoelectronic operator, the distinction between the first single excited state, and the triplet is not considered. Finally, one has to underline the kinetic nature of the present study, and this differentiates it from the energetic studies made on the stability of the different conformations of the final products [13, 19].

2. Method of Calculation

The calculations based on the delocalized model of the transition state were carried out according to the method previously described [40]. The conditions under which this method can be applied to cycloadditions of two conjugated systems have been exposed, by studying the photodimerization of the anthracene [41]. Faced with the Wheland model, this model of the transition state corresponds to an extension of the delocalization. This is suitable for rapid reactions in which the transition state does not separate itself far away from the reactants. At intermolecular distances which are great, the Hamiltonian operator in each conjugated molecule differs only slightly from the original, in a manner in which the σ/π separation would still be practically valid. The wave function of the transition state can be constructed from molecular orbitals covering all the system for the electrons π of both molecules, whilst the electrons σ remain localized in each molecule, acting as a hard core which forbids too close an approach. The extension of the Hückel method to the whole delocalized system of electrons is based on an orbital overlap argument, and not on a symmetry one. This can be done because the idea of delocalized bond does not include any element of symmetry.

The delocalization energy is the difference in energy, in units of β, between the transition state containing the whole delocalized system in its first excited state and the inital state of two molecules of pyrimidine bases, one in its ground state and the other in its first excited state. The larger the value of delocalization energy, the lower in energy is the transition state for formation of the photodimer. In some calculations the resonance integrals between carbons 5 and 6 of both molecules have been only taken into account. The resonance integrals are assumed to be equal to 0.5. This value corresponds to "σ" type overlap of 0.125 between two carbon $2p$ SCF orbitals with atoms 3 Å apart. This distance is common for all charge transfer electronic delocalization. In other calculations all the resonance integrals between atomic $2p_z$ orbitals of both molecules have been taken into account. These resonance integrals are assumed to be proportional to overlap integrals [42]. The intermonomer overlaps were obtained in the following way: for every pair of atoms, spreading over both monomers, the $2p_z$ orbitals were decomposed into components of σ and π mutual symmetry. The overlap of the components with a $\sigma-\sigma$ symmetry was calculated by equations suggested by Roothaan [43]. The contribution of the components with the $\pi-\pi$ symmetry to the overlap was disregarded. Finally eigenvalues and wavefunctions of bases in a Hückel approximation have been obtained by using the parameters suggested by Hoffman and Ladik [44].

The calculations based on the theory of perturbations has been done using the approximation proposed by Salem [45]. The energy of interaction between the conjugated systems is expressed in function of the overlaps between the atomic orbitals $2p_z$ of both molecules. Two limitations are imposed. Each atomic centre only interacts with one centre of the other molecule. The resonance integrals are small when compared with the unit (i.E. $\leqq 0.2$).

One supposes that both molecules have orbital bands σ and π well separated. One only evaluates the energy of interaction coming from the molecular orbitals π. The contribution to the stabilisation of the orbitals σ will be depreciable, since it comes from terms of perturbation of the 2° order between molecular orbitals occupied by one molecule and vacant orbitals of the other. In the denominator the difference of energy between the orbitals considered will be elevated, and a small term will figure in the numerator, which corresponds to a weak overlapping.

The extended Hückel method [46] is also used for calculating comparatives.

3. Results and Discussion

The Table 1 presents the energy of delocalization for the photodimerization of the pyrimidine bases. Only the resonance integrals between the carbons 5 and 6 of both molecules have been taken into account.

Table 1. Delocalization energies

	Syn	*Anti*
Uracil	0.703	0.608
Thymine	0.699	0.618
Cytosine	0.588	0.464
Cytosine-thymine	0.600 C*	0.503 C*
	0.700 T*	0.603 T*

The delocalization energy of the anthracene in the photodimerization is 0.566 [41]. Attending only to this factor, the photodimerization is presented as more favourable for the pyrimidine bases. The relative efficiency in the formation of dimers between the cytosine and the thymine [12] agrees well with the delocalization energies evaluated. Take note that the delocalization energy of the mixed photodimer will distinct, following that the excited reactant molecule will be that of the cytosine (C*), or the thymine (T*). One must signal the difference between theoretic results, and experimental values. The latter correspond to equilibrium constants, whilst the former are in regard to constants of velocity of the direct reaction of the dimerization.

By virtue of the delocalization energies obtained, the formation of *syn* conformations (head-to-head) will be kinetically more favourable than the formation of *anti* conformations (head-to-tail).This result, already previously obtained for the thymine [37], has a general character. To study the conformations *cis* (endo) and *trans* (exo) it is necessary to consider the interactions between all the centres of both molecules. The Table 2 presents the results obtained for the uracil and cytosine. The uracil was selected to avoid the problems stemming from the methyl group of the thymine. The calculations have been done with the two bases in parallel planes at a distance of 2.8 Å. Later on, this selection will be justified. Using Slater orbitals at this distance, one obtains lower values for the overlap integrals with respect to those obtained with atomic orbitals SCF. The Table 2 also contains the interaction energy evaluated by the theory of perturbations, and the increase of energy between the transition state and reagents evaluated by the extended Hückel method.

Table 2

	Delocalization energies		Interaction energies		$\Delta E^{\neq}$ computed by EHMO	
Uracil	Exc.	Fund.	Exc.	Fund.	Exc.	Fund.
Cis-syn	0.446	0.038	0.224	−0.130	−0.191	0.997
Trans-syn	0.362	0.030	0.229	−0.059	−0.144	0.755
Cis-anti	0.300	0.069	0.116	−0.102	0.248	0.868
Trans-anti	0.306	0.037	0.184	−0.050	−0.064	0.716
Cytosine						
Cis-syn	0.412	0.034	0.196	−0.139	—	—
Trans-syn	0.309	0.032	0.169	−0.051	—	—
Cis-anti	0.251	0.069	0.072	−0.100	—	—
Trans-anti	0.265	0.044	0.102	−0.043	—	—

The delocalization and interaction energies are given in units of β and $\Delta E^{\neq}$ in eV.

The energies obtained by the three methods present a displacement, introduced in the theory of perturbations by repulsive terms resulting from the interaction of orbitals π doubly occupied of both molecules, owing to the overlapping, and in the Hückel extended method, by the repulsion of σ orbitals. It is interesting to note that relative results between the different conformations in the photodimerization and the dimerization, with both molecules in a fundamental state, remain practically constant. One must note that the results of the extended Hückel method come given in the eV and not in β units. The *cis-syn* conformation is the most favoured. The inclusion of all valence electrons does not modify this result. However, when using a monoelectronic operator, one does not bear in mind the electrostatic interaction between both molecules. An exploratory calculation with the distribution of the π charges gives a high repulsion in the *cis-syn* conformation, in such a way that the electrostatic factor comes in to compensate the energy differences between the different conformations. To discuss the conformation kinetically most favoured, it appears to be indespensable to introduce a bielectronic

Hamiltonian operator, as has already been done for reactions with both molecules in the fundamental state [47, 48]. It is necessary to give only a provisional value to the results obtained with a monoelectronic operator.

Experimentally, the *cis-syn* conformation presents a double rotation around the 5–6 bonds, and the line which unites the midpoints of the 5–6 bonds. One tries to study whether the improving reaction initiate these rotations which present the geometry of the final product. The theory of perturbations is inadequate to study the rotation around the line which unites the two mid-points of the 5–6 bonds, by excluding the interaction between 3 or more centers in the approximation used.

One has evaluated the delocalization energy by turning the upper base to the distance of 2.8 Å. In the transition state of the photodimerization of the uracil, there was an energetic minimum presented at 9° of rotation, 0.004 β underneath the starting position. On the other hand, the electrostatic repulsion diminishes with the rotation. Because of this, it appears that one can be sure that the rotation around the line which unites the mid-points of the 5–6 bonds is already initiated in the transition state. A similar result, but with a higher rotation of 28°, had been discovered earlier by another method [38].

The angle formed by the two planes of the bases at coming together can be studied by constructing the interaction energy surface in function of the distance and the angle of the turn. The analytical expression of the interaction energy in function to the overlap integrals between the orbitals $2p_z$ of both molecules permits a rapid evaluation of the points of the surface. Table 3 contains said equations for the *cis-syn* and *cis-anti* conformations of uracil, of the thymine and the cytosine. S_{rp} indicates the overlap integral between the $2p_z$ orbitals of the centre r of one molecule, and the centre p of another. The numeration which has been adopted for the different centres is to be found in Fig. 2. The expressions for the second order terms have been truncated, owing to their length. So for example, the equation for the uracil *cis-syn* contains 44 terms.

Table 3. Equations of interaction energy in function of overlap integrals

Uracil

$$\textit{Cis-syn}\,\Delta E = 1{,}060\,S_{11} + 0{,}109\,S_{22} + 0{,}198\,S_{33} + 0{,}282\,S_{44} + 1{,}915\,S_{55} + 1{,}668\,S_{66} + 0{,}470\,S_{77} + 0{,}299\,S_{88} - 5{,}692\,S_{11}^2 - 1{,}421\,S_{11}S_{22} - \cdots$$

$$\textit{Cis-anti}\,\Delta E = -0{,}440\,S_{14} + 0{,}131\,S_{23} + 0{,}131\,S_{32} - 0{,}440\,S_{41} + 1{,}402\,S_{56} + 1{,}402\,S_{65} + 0{,}117\,S_{14}^2 - 1{,}176\,S_{14}S_{23} - \cdots$$

Thymine

$$\textit{Cis-syn}\,\Delta E = 1{,}034\,S_{11} + 0{,}104\,S_{22} + 0{,}203\,S_{33} + 0{,}320\,S_{44} + 1{,}891\,S_{55} + 1{,}663\,S_{66} + 0{,}489\,S_{77} + 0{,}288\,S_{88} + 0{,}001\,S_{99} + 0{,}006\,S_{1010} - 5{,}688\,S_{11}^2 - 1{,}426\,S_{11}S_{22} - \cdots$$

$$\textit{Cis-anti}\,\Delta E = -0{,}462\,S_{14} + 0{,}125\,S_{23} + 0{,}125\,S_{32} - 0{,}462\,S_{47} + 1{,}441\,S_{56} + 1{,}441\,S_{65} + 0{,}083\,S_{14}^2 - 1{,}181\,S_{14}S_{23} - \cdots$$

Cytosine

$$\textit{Cis-syn}\,\Delta E = 1{,}027\,S_{11} + 0{,}050\,S_{22} + 0{,}530\,S_{33} + 0{,}735\,S_{44} + 1{,}290\,S_{55} + 1{,}362\,S_{66} + 0{,}659\,S_{77} + 0{,}347\,S_{88} - 5{,}466\,S_{11}^2 - 1{,}191\,S_{11}S_{22} - \cdots$$

$$\textit{Cis-anti}\,\Delta E = -0{,}689\,S_{14} + 0{,}124\,S_{23} + 0{,}124\,S_{32} - 0{,}689\,S_{41} + 0{,}798\,S_{56} + 0{,}798\,S_{65} + 0{,}481\,S_{11}^2 - 1{,}286\,S_{14}S_{23} - \cdots$$

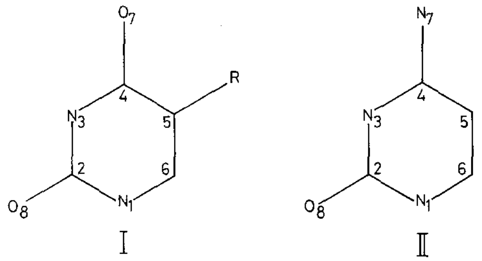

Fig. 2. The numeration of centres. I. R = H uracil, R = CH_3 thymine; II. cytosine

At great distances, the terms of the first order will be predominant. Among these, those which affect the 5 and 6 centres have the highest coefficients, owing to the high coefficients of the frontier orbitals in these positions. The Figs. 3, 4, and 5 present the interaction energy surfaces for the uracil *cis-syn*, the uracil *cis-anti* and the cytosine *cis-syn*. The most favourable path of reaction will be perpendicular to the isoenergetic lines, in the direction of the higher interaction energies. This path is the one which corresponds to the maximum stabilisation. In the Figs. 3, 4, and 5, a dotted line has been drawn for the most favourable path of reaction, starting from a position parallel to both molecules at great distances. The molecules come together in parallel until the distance of 2.8 Å. After this distance, the approximation is accompanied by a rapid turn of both molecules around the 5–6 bonds. A rapid decrease in the interaction energy impedes a parallel approximation after the distance indicated. If one had used the SCF

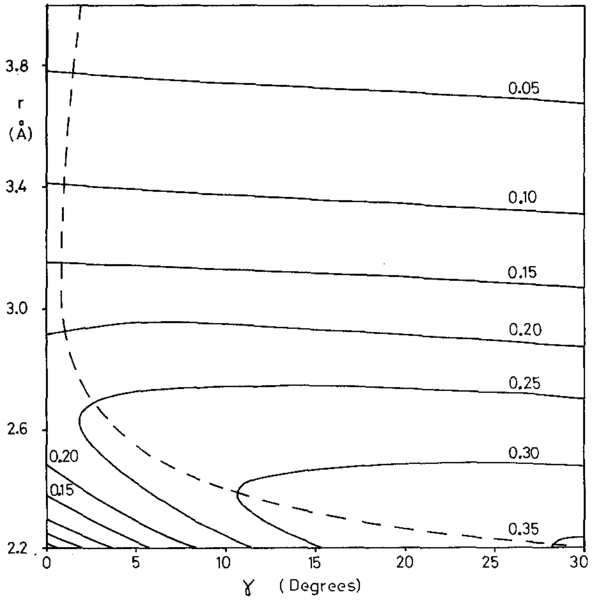

Fig. 3. Interaction energy surface between two *cis-syn* uracils, r distance between 5–6 bonds of both molecules, γ angle formed by the two planes of the bases

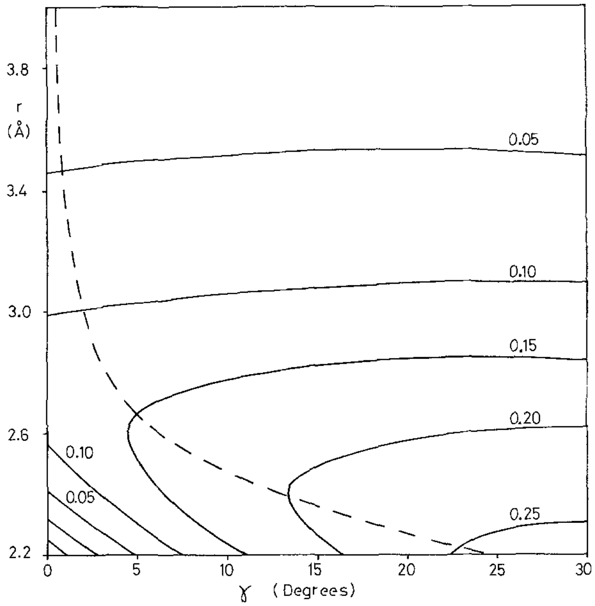

Fig. 4. Interaction energy surface between two *cis-anti* uracils, r and γ as in Fig. 3

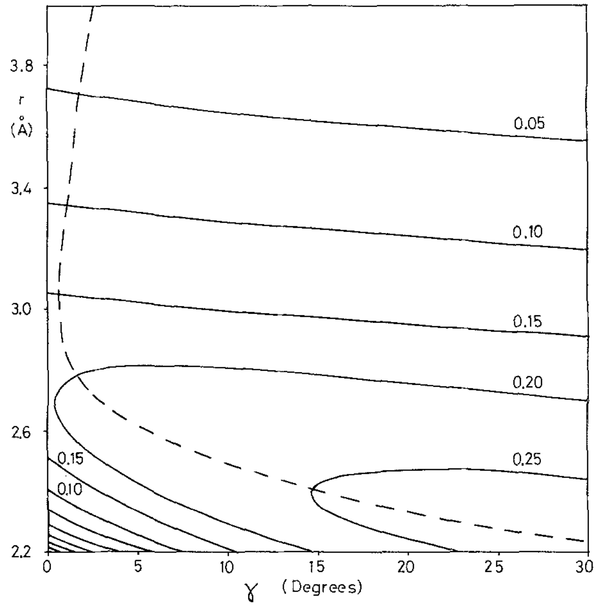

Fig. 5. Interaction energy surface between two *cis-syn* cytosines, r and γ as in Fig. 3

atomic orbitals instead of the Slater orbitals to evaluate the overlap integrals, one would have found a greater distance. Even though the interaction surface represents only one of the components of the potential surface, the path of reaction found, seems to be qualitively correct, since other factors, such as the electrostatic repulsion, confirm it.

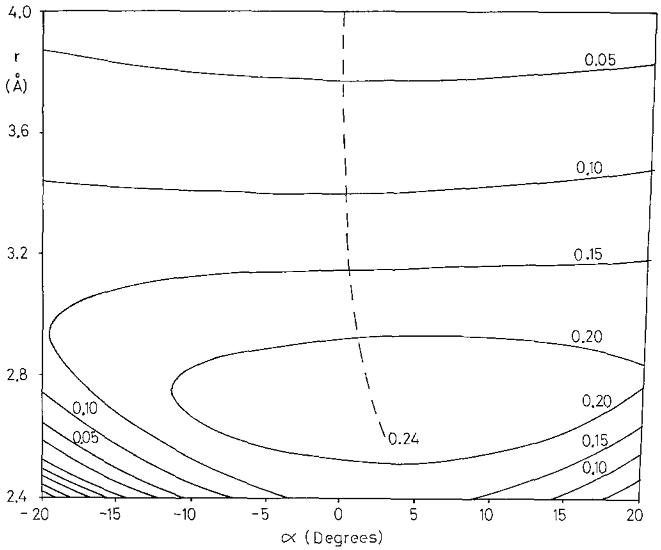

Fig. 6. Interaction energy surface between two *cis-syn* uracils, *r* distance between the midpoints of 5–6 bonds of both molecules, α angle formed by the two 5–6 bonds of the bases

The equations of the interaction energy permit the study also of the concerted character of the process of photodimerization. For the *anti* conformations it is undoubtable that the new bonds are formed at the same velocity, since there is an axis of symmetry in existance, but for the *syn* conformations, this is open to question. Fig. 6 presents the interaction energy surface in function of the distance between the two midpoints of the 5–6 bonds, and the angle formed by these two bonds. For positive values, the 6–6 bond is formed more rapidly, and vice-versa, the 5–5 bond for negative values. It can be seen that if both 5–6 bonds are parallel at great distances, they remain in parallel in the course of the reaction. Even if both bonds for a certain angle at great distances they will follow on adapting a parallel conformation as the reaction advances, since the no-null angles appear to have a high diminishing of the interaction energy, as the distance between the two bases decreases. The concerted character of the reaction seems, therefore, to be confirmed by this theoretic study.

Acknowledgement. The author is grateful to the European Molecular Biology Organization for financial support to visit Quantum Chemistry Group of Uppsala University where some calculations were carried out.

References

1. McLaren, A. D., Shugar, D.: Photochemistry of proteins and nucleic acids, Oxford: Pergamon 1964.
2. Setlow, J. K.: Current Topics Rad. Res. **2**, 195 (1966).
3. Smith, K. C.: Radiation Res. Suppl. C, 54 (1966).
4. Setlow, R. B.: Progress in Nucleic Acid Res. and Mol. Biol. **8**, 257 (1968).
5. Burr, J. G.: Advances Photochemistry **6**, 193 (1968).
6. Beukers, R., Berends, W.: Biochim. Biophys. Acta **41**, 550 (1960).

7. Swensen, P.A., Setlow, R.B.: Photochem. Photobiol. **2**, 419 (1963).
8. Brown, I.H., Freeman, K.B., Johns, H.E.: J. molecular Biol. **15**, 640 (1966).
9. Smith, K.C.: Photochem. Photobiol. **2**, 503 (1963).
10. Freeman, K.B., Hariharan, P.V., Johns, H.E.: J. molecular Biol. **13**, 833 (1965).
11. Ben-Ishai, R., Ben-Hur, E., Hornfeld, Y.: Israel J. Chem. **6**, 769 (1968).
12. Setlow, R.B., Carrier, W.L.: J. molecular Biol. **17**, 237 (1966).
13. Jordan, F., Pullman, B.: Theoret. chim. Acta (Berl.) **10**, 423 (1968).
14. Weinblum, D., Johns, H.E.: Biochim. Biophys. Acta **114**, 450 (1966).
15. Fisher, G.J., Johns, H.E.: Photochem. Photobiol. **11**, 429 (1970).
16. Johns, H.E., Helleincer, C.W., Pearson, M.L., LeBlanc, J.C.: J. molecular Biol. **9**, 503 (1964).
17. Blackburn, G.M., Davies, R.J.H.: Tetrahedron Letters **37**, 4471 (1966).
18. Camerman, N., Nyburg, S.C., Weinblum, D.: Tetrahedron Letters **42**, 4127 (1967).
19. Einstein, J.R., Hosszu, J.L., Longworth, J.W., Rahn, R.O., Wei, C.H.: Chem. Comm. **20**, 1063 (1967).
20. Adman, E., Gordon, M.P., Jensen, L.H.: Chem. Comm. **21**, 1019 (1968).
21. Camerman, N., Camerman, A.: Science **160**, 1451 (1968).
22. Karle, I.L.: Acta crystallogr. B **25**, 2119 (1969).
23. Camerman, N., Camerman, A.: J. Amer. chem. Soc. **92**, 2523 (1970).
24. Corey, E.J.: J. Amer. chem. Soc. **86**, 5570 (1964).
25. Lamola, A.A.: Photochem. Photobiol. **7**, 619 (1968).
26. — Eisinger, J.: Proc. nat. Acad. Sci. USA **59**, 46 (1968).
27. Whillans, D.W., Johns, H.E.: Photochem. Photobiol. **9**, 323 (1969).
28. Wagner, P., Bucheck, D.: J. Amer. chem. Soc. **92**, 181 (1970).
29. Eisinger, J., Schulman, R.G.: Proc. nat. Acad. Sci. USA **58**, 895 (1967).
30. Sutherland, B.M., Sutherland, J.C.: Biophys. J. **9**, 292 (1969).
31. Mantione, M.J., Pullman, B.: Biochim. Biophys. Acta **91**, 387 (1964).
32. Pullman, B., Mantione, M.J.: Biochim. Biophys. Acta **95**, 668 (1965).
33. Danilov, V.I.: Photochem. Photobiol. **6**, 233 (1967).
34. Denis, A., Pullman, A.: Theoret. chim. Acta (Berl.) **7**, 110 (1967).
35. Kuprievich, V.: Int. J. quant. Chemistry **1**, 561 (1967).
36. Pullman, B.: Photochem. Photobiol. **7**, 525 (1968).
37. Nagata, C., Imamura, Y., Tagashira, M., Kodama, M., Fukuda, N.: J. theoret. Biology **9**, 357 (1965).
38. Sayre, R., Harlos, J., Rein, R.: Molecular orbital studies in chemical pharmacology, ed. by L.B. Kier, p. 207. N.Y.: Springer-Verlag 1970.
39. Camerman, N., Nyburg, S.: Theoret. chim. Acta (Berl.) **13**, 162 (1969).
40. Bertrán, J., Chalvet, O., Daudel, R., McKillop, T.F., Schmid, G.H.: Tetrahedron **26**, 339 (1970).
41. — Schmid, G.H.: Tetrahedron **27**, 5191 (1971).
42. Mulliken, R.S.: J. Chim. physique **46**, 497, 675 (1949).
43. Roothaan, C.J.: J. chem. Physics **19**, 1445 (1951).
44. Hoffmann, T.A., Ladik, J.: Advances chem. Physics **7**, 84 (1964).
45. Salem, L.: J. Amer. chem. Soc. **90**, 543, 553 (1968).
46. Hoffman, R.: J. chem. Physics **39**, 1397 (1963).
47. Devaquet, A., Salem, L.: J. Amer. chem. Soc. **91**, 3793 (1969).
48. Sustmann, R., Binsch, G.: Molecular Physics **20**, 1, 9 (1971).

Dr. J. Bertran
Instituto de Biología Fundamental
Universidad Autónoma de Barcelona
Avd. San Antonio M^a Claret 171
Barcelona – 13, Spain