Photoaddition of water to uracil: A model study of loss of absorbance by pyrimidine adducts using geometry optimization of ground and excited states

P C MISHRA and R D TEWARI

Department of Physics, Banaras Hindu University, Varanasi 221 005, India

Abstract. Geometries of the ground, first triplet and first singlet excited states of a hydrohydroxide of uracil have been optimized using the SINDO 1 SCF molecular orbital method in conjunction with configuration interaction. The first triplet and first singlet transitions of the molecule are found to be of $(\pi\text{-}\sigma^*)$ type which shows their orbitally forbidden character. Thus an explanation as to why the adducts formed by the addition of H_2O, hydroxylamine, bisulphite etc at the CC double bonds of pyrimidines do not absorb in the characteristic 260 nm region has been provided.

1. Introduction

Molecules of hydroxylamine, certain bisulphites, water, etc add across the C_5C_6 double bonds of pyrimidines and in certain cases it has been observed to be of great biological significance (Brooks and Newbold 1980; Hayatsu 1976). The reaction depends on the conditions of pH, temperature and exposure to appropriate ultraviolet radiation. Further, these depend considerably on the pyrimidine base under study. It is established that consequent to the formation of an adduct, the pyrimidine bases do not absorb in the characteristic ultraviolet region of 260 nm, though this radiation is known to enhance efficiency of the addition reaction (Smith and Hanawalt 1969). It has also been suggested that decrease of absorption in this region may be used to monitor the amount of adduct formed (Hayatsu 1976). However, its theoretical basis is not understood. The present work was performed to explain this basis and to show that the decrease of absorption can genuinely serve as a criterion to monitor the addition reaction.

2. Method of calculations

A hydrohydroxide of uracil was taken as a representative of the various adducts formed by addition at the CC double bonds of pyrimidines. The SINDO 1 SCF molecular orbital method was used to calculate wave functions (Nanda and Jug 1980; Jug and Nanda 1980) and the Newton-Raphson procedure was used to search each of the

ground and excited state potential surface minima separately. Ground state as well as the first triplet and first singlet excited states of the adduct were studied by optimizing all the internal coordinates. Three lowest energy excitations $25 \to 26$, $25 \to 27$ and $24 \to 26$, where the highest occupied molecular orbital is numbered 25 and this orbital numbering has been done in the increasing order of energy, were included in the configuration interaction (CI) treatment. The necessary first and second derivatives of energy with respect to internal coordinates were calculated numerically. The diagonal elements of the Hessian matrix (second derivatives of energy) were examined in order to ensure that the extrema searched were minima. Excitation energies were obtained as differences of total energies of optimised ground and excited states. A H_2O molecule was considered near the CC double bond of uracil. The starting distance between the centre of the CC double bond and oxygen of H_2O was 3 Å. Then the separation of the H and OH fragments of H_2O was increased and the whole system optimized. In this process, the H and OH fragments of H_2O got attached across the CC double bond of uracil: OH at C_5 and H at C_6 (see figure 1). Similarly, addition of H at C_5 and OH at C_6 could be effected. Both these situations are expected to lead to similar results. Excited state optimization was performed beginning with the SCF molecular orbitals of the ground state. A statistics as to the performance of the SINDO 1 method in comparison to other methods is available elsewhere (Nanda and Jug 1980; Jug and Nanda 1980). On account of its satisfactory performance not only for the ground state but also for excited states of a variety of molecules and reactions (Jug 1980; Mishra and Jug 1982; Müller-Remmers *et al* 1984), this method was chosen for the present study.

3. Results and discussion

The following results were obtained for the adduct studied. It is desirable to compare the nature of electronic excitations of uracil with those of the uracil-H_2O adduct to understand their relationships and differences. The highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) of uracil can be numbered as 21 and 22 respectively. The orbital just above LUMO (from the energy point of view) would be numbered 23 while that just below HOMO would be numbered 20. The orbitals 20, 21, 22 and 23 are of types σ (non-bonding), π, σ^* and π^* respectively in the equilibrium ground state geometry of uracil. The vertical transitions T_1 and S_1 of uracil arise from the excitations $21 \to 23$ and $21 \to 22$ respectively. Thus they are of types $(\pi\text{-}\pi^*)$ and $(\pi\text{-}\sigma^*)$ respectively. T_2 arises from the $21 \to 22$ excitation and is of the type $(\pi\text{-}\sigma^*)$. T_3, S_2 and S_3 involve both the $20 \to 22$ and $21 \to 23$ excitations. Thus S_1 and T_2 as vertical
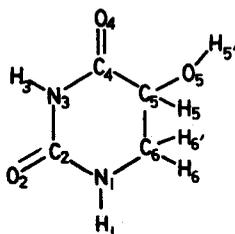

Figure 1. Atomic numbering scheme in the uracil-water adduct.

transitions have orbitally forbidden character whereas T_1, S_2 and S_3 have orbitally allowed character. On geometry optimization the orbitals acquire some mixed character. Further, in going to the equilibrium geometries of the excited states T_1 and S_1, the orbitals 22 and 23 cross. Also in the equilibrium geometry of T_1 the corresponding transition arises from the $21 \rightarrow 22$ excitation and thus retains its orbitally allowed nature. In the equilibrium geometry of S_1, the singlet transition gets a large contribution from the $21 \rightarrow 22$ excitation and a small contribution from the $21 \rightarrow 23$ excitation. Thus the nature of S_1 changes from one of forbidden (π-σ^*) type to one of largely allowed (π-π^*) type due to the excited state geometry relaxation.

The nature of the first singlet transition of uracil has continued to be a mystery, despite several experimental and theoretical studies. It is mainly because the transition is experimentally observed to be fairly intense and the theoretical methods like CNDO-S/CI (Dal Bene and Jaffé 1968) which are designed for only equilibrium to equilibrium transitions, cannot account for change of nature of transitions on geometrical relaxation. It is only recently that a forbidden character for the first singlet transition of uracil has been suggested on the basis of experimental work though no clear explanation for the forbiddenness of the transition could be given by the authors. A possibility was suggested that this transition might be of (n-π^*) type and overlaps with another (π-π^*) transition (Becker and Kogan 1980; Salet *et al* 1979). The SINDO 1 calculation does not only explain the partial forbidden character of the transition but it also clarifies why its observed intensity is fairly high (Mishra and Jug 1982) as explained above.

To assign the transitions in the uracil-water adduct we must examine the nature of its molecular orbitals and CI wave functions. Since the molecule has significant deviation from planarity, the orbitals such as π and σ cannot be classified exactly. One may devise different ways to obtain appropriate indices for approximate classifications of orbitals in such a situation. We have used here the following simple criterion for the purpose. One would not expect major changes in the classification by improving upon the criterion. In a fully planar molecule with the molecular plane coinciding with xy-plane, p_z orbitals would be rigorously classified as π-orbitals. Let the sum of squares of p_z atomic orbital coefficients, multiplied by 100, be written as $\chi(\pi)$. In a fully planar molecule $\chi(\pi)$, due to normalization, would be equal to 100. One may call $\chi(\pi)$ percentage π-character. Small and large values of $\chi(\pi)$ would respectively represent σ and π molecular orbitals. It may be noted there the above definition of $\chi(\pi)$ does not include hyperconjugation effects and hence it would generally tend to somewhat underestimate the π-characters of molecular orbitals in molecules having partial deviation from planarity. The values of $\chi(\pi)$ for the molecular orbitals involve in the transitions T_1 and S_1, in the respective equilibrium geometries, are given in table 1. We find that the HOMO (number 25) in both the cases is of π type while the other orbitals (numbers 24, 26 and 27) are of σ or σ^* types.

CI wavefunctions presented in table 2 show that the first triplet and first singlet transitions of the adduct, in the corresponding excited state equilibrium geometries, arise mainly from $25 \rightarrow 26$ orbital excitation. In both the cases they get appreciable contributions from the $25 \rightarrow 27$ excitation also. Thus these two transitions should be classified as (π-σ^*). The second triplet and second singlet transitions of the adduct have (σ-σ^*) character as they arise prominently from $24 \rightarrow 26$ orbital excitation. Thus while the first triplet and first singlet transitions of the adduct would be orbitally forbidden, the second triplet and second singlet transitions would be allowed.

Table 1. Assignments of orbitals.

Orbital no.	$\chi(\pi)^a$	Approx. assignment
(A) T_1 optimization		
24	2·3	σ
25	74·7	π
26	3·9	σ^*
27	3.8	σ^*
(B) S_1 optimization		
24	4·4	σ
25	68·4	π
26	5·1	σ^*
27	4·9	σ^*

^{a} percentage π-character (see text).

Table 2. Assignment of electronic transitions.

State[a]	Excitation wavelengths[b] (nm)	Main CI wave-function[c]	Assignment
(A) T_1 optimization			
T_1	317·2	0·67 Ψ (25 → 26)	$\pi\text{-}\sigma^*$
	(383·3)	+0·33 Ψ (25 → 27)	
S_1	235·7	0·71 Ψ (25 → 26)	$\pi\text{-}\sigma^*$
	(257·1)	+0·29 Ψ (25 → 27)	
T_2	184·8	0·99 Ψ (24 → 26)	$\sigma\text{-}\sigma^*$
S_2	160·8	0·95 Ψ (24 → 26)	$\sigma\text{-}\sigma^*$
	(221·5)		
(B) S_1 optimization			
T_1	303·3	0·73 Ψ (25 → 26)	$\pi\text{-}\sigma^*$
	(383·3)	+0·27 Ψ (25 → 27)	
S_1	234·0	0·76 Ψ (25 → 26)	$\pi\text{-}\sigma^*$
	(257·1)	+0·24 Ψ (25 → 27)	
T_2	179·5	1·0 Ψ (24 → 26)	$\sigma\text{-}\sigma^*$
S_2	158·3	0·99 Ψ (24 → 26)	$\sigma\text{-}\sigma^*$
	(221·5)		

^{a} T_n and S_n represent nth triplet and nth singlet excited states respectively; ^{b} Values given in parentheses correspond to planar uracil (Mishra and Jug 1982); ^{c} Contributions smaller than 5 % are not given.

Excitation wavelengths evaluated with respect to the optimized ground state and assignments of transitions in the adduct are presented in table 2. In this table, the results given under (A) were obtained when the first triplet excited state was optimized whereas those under (B) were obtained when the first singlet excited state was optimized. Excitation wavelengths of the states which were not optimized were obtained as vertical separations with respect to the corresponding optimized excited state. It is noted that the results presented under (A) and (B) are in reasonable agreement. Excitation

wavelengths of uracil (Mishra and Jug 1982) are compared with those of the adduct in table 2. In making this comparison the following two points are to be noted: (i) comparison would be proper only between the corresponding optimized values. Thus while the excitation wavelengths of T_1 of uracil and of the adduct may be reasonably compared under category A, those of S_1 may be compared under category B. Excitation wavelengths of the adduct for other states would be appreciably shorter as these are not optimized. Since the calculated excitation energies are very small fractions ($\sim 0.2\%$) of the total energies as the differences of which these are obtained, accuracies in the excitation wavelengths are not expected to be high. Thus although table 2 shows that excitation wavelengths of the adduct would be blue-shifted with respect to those of uracil, as expected from the more "saturated" nature of the adduct, it would be inappropriate to base the assignments of excited states mainly on the calculated excitation wavelengths. In this situation, the nature and correlation of orbitals appears to be a better guide to assign the transitions.

The highest occupied molecular orbital is localized upto about 7 % on the oxygen of the OH group in the ground state of the adduct. In the excited states, this molecular orbital loses even this small amount of localization on the OH group, but it develops about 10 % localization on the other hydrogen atom coming from H_2O and attached to the other carbon atom of uracil. No other molecular orbital involved in the transitions is localized even to this extent on the fragments of H_2O. It shows that the electronic transitions of the adduct mainly belong to uracil and they are only perturbed by the fragments of H_2O. Consequently, the transitions T_1 and S_1 of the adduct can be correlated with the corresponding transitions of uracil. However, perturbation by the addition of H_2O causes change of nature of the transitions T_1 and S_1 from largely (π-π^*) type to (π-σ^*) type in the excited state equilibrium geometries and thus the transitions acquire orbitally forbidden character. This seems to be the reason why the ultraviolet absorption around 260 nm is not observed in the uracil-H_2O adduct as shown in figure 2 as well as in the other products formed by the addition of bisulphites, hydroxylamine etc at the CC double bonds of pyrimidines (Budowsky 1976; Smith and Hanawalt 1969).

Introduction of non-planarity following the addition at the CC double bonds of pyrimidines has been suggested earlier. It is proposed that the hydroxylamine adducts have half-chair conformation (Budowsky 1976). The present results of geometry
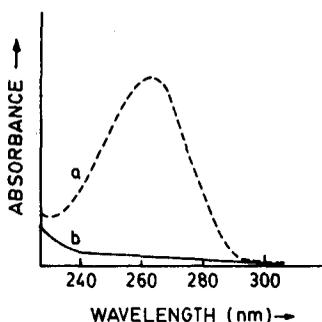

Figure 2. UV absorption of uridylic acid. a. before reaction; b. after reaction (Smith and Hanawalt 1969).

Table 3. Values of internal coordinates in the addition region (Å and degrees)[†].

Coordinate	Ground state	Triplet excited state	Singlet excited state
C_4C_5	1·621	1·612	1·630
	(1·547)	(1·532)	(1·560)
C_5C_6	1·616	1·600	1·575
	(1·372)	(1·561)	(1·423)
C_1N_6	1·472	1·465	1·441
	(1·427)	(1·394)	(1·418)
C_5O'	1·411	1·415	1·414
$C_4C_5C_6$	115·4	115·5	115·2
	(119·8)	(115·8)	(116·8)
$C_5C_6N_1$	117·0	117·4	117·4
	(119·8)	(116·5)	(118·5)

[†] Values given in parentheses correspond to uracil (Mishra and Jug 1982).

optimization, the important ones of which are given in table 3 and are discussed below, do not support this view. The C_4C_5 and C_5C_6 bond lengths are appreciably increased in comparison to uracil following the adduct formation. Most noticeable is the increase in the C_5C_6 bond length in the ground state of the adduct. Similarly, the bond angles $C_4C_5C_6$ and $C_5C_6N_1$ are appreciably reduced in comparison to uracil in the ground state of the adduct; in the excited states the effect of adduct formation on these angles is very small.

We may discuss non-planarity with respect to the plane of the atoms N_3, C_4 and C_5 of the ring. The other atoms of the ring generally have deviation of 0·1 to 0·2 Å from the reference plane. However, as expected, the exocyclic atoms attached to C_5 and C_6 have large deviations from the reference plane. H_5 and H_6 lie on the same side of the plane and the displacements for H_5 are 0·98, 1·0 and 1·04 Å while those of H_6 are 0·67, 0·34 and 0·32 Å in the ground state, triplet excited state and singlet excited state respectively. The fragments H and OH of H_2O lie on the other side of the reference plane. Let us label the oxygen and hydrogen of H_2O which are attached respectively to C_5 and C_6 of uracil as O_5 and $H_{6'}$ while the hydrogen of the OH group may be labelled as $H_{5'}$. The displacements of O_5, $H_{6'}$ and $H_{5'}$ from the reference plane respectively are as follows: ground states: $-1·03$, $-1·08$ and $-1·12$ Å, triplet excited state $-0·99$, $-1·22$ and $-1·06$ Å, singlet excited state: $-0·90$, $-1·28$ and $-0·99$ Å. Thus each fragment of H_2O bound to the adduct is at a distance of about 1 Å from the reference plane.

Intramolecular hydrogen bonding between the oxygen O_4 and a hydrogen of the NHOH group has been invoked to explain stability of the hydroxylamine adduct (Budowsky 1976). It appears desirable to comment on this aspect in the present case also. The distances between the hydrogen atom of the OH group and O_4 are all close to 2·7 Å in the ground state, the triplet excited state and the singlet excited state of the adduct. Considering the net charges, one gets electrostatic interaction energies between these two atoms as nearly 4·9 kcal/mole in all the three states. It shows the existence of an intramolecular hydrogen bonding between O_4 and the hydrogen of the OH group.

4. Conclusion

An important conclusion of the present work is that the first singlet transitions of adducts formed by addition at the CC double bonds of pyrimidines are orbitally forbidden whereas these are largely allowed in the pyrimidines themselves. Consequently, the adduct formation can be quantitatively monitored with the help of decrease of absorption in the characteristic 260 nm region.

References

Becker R S and Kogan G 1980 *Photochem. Photobiol.* **31** 5
Brookes P and Newbold R F 1980 in *Carcinogenesis: fundamental mechanisms and environmental effects* (eds) B Pullman, P O P Ts's and H Gelboin (Boston: Reidel Press) p. 81
Budowsky E I 1976 *Progr. Nucl. Acid Res. Mol. Biol.* **16** 125
Del Bene J and Jaffé H H 1968 *J. Chem. Phys.* **48** 1807
Hayatsu H 1976 *Progr. Nucl. Acid Res. Mol. Biol.* **16** 75
Jug K 1980 *Theor. Chim. Acta* **54** 263
Jug K and Nanda D N 1980 *Theor. Chim. Acta* **57** 107, 131
Müller-Remmers P, Mishra P C and Jug K 1984 *J. Am. Chem. Soc.* **106** 2538
Mishra P C and Jug K 1982 *Theor. Chim. Acta* **61** 559
Nanda D N and Jug K 1980 *Theor. Chim. Acta* **57** 95
Salet C, Benasasson R and Becker R S 1979 *Photochem. Photobiol.* **30** 325
Smith K C and Hanawalt P C 1969 *Molecular photobiology, inactivation and recovery* (New York: Academic Press) p. 59