Electronic Spectra of the H and OH Adducts of Cytosine

M. Krauss*

Center for Advanced Research in Biotechnology, National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850

R. Osman†

Department of Physiology and Biophysics, Mount Sinai School of Medicine, City University of New York, New York, New York 10029

Received: September 19, 1996; In Final Form: March 24, 1997⊗

Theoretical assignments of the spectra of the H and OH adduct radicals of cytosine are obtained with ab initio multiconfiguration self-consistent-field calculations. The C6OH isomer is substantially more stable than the C5OH. The C6H isomer is slightly more stable than C5H, but both are much less stable than the N3H isomer. The C5 and C6 adducts for both H and OH are found to support transitions in entirely different spectral regions. Visible absorption transitions are found for the C6 adducts in which the open shell orbital is delocalized by conjugating to a neighboring double bond or system of bonds. In C5 adducts, the open-shell orbital remains localized giving rise to absorption transitions in the UV. The calculations show that the very broad absorption observed at 440 nm is the convolution of two transitions for the C6OH isomer and that the observed peak at 340 nm has contributions from both the C5OH and C6OH isomers but is dominated by the absorption of C5OH. The similarity of the absorption pattern of the cytosine and uracil adducts suggest a general behavior in the spectroscopy of pyrimidine adducts. Base radical calculations were reported at the 8th International Congress of Quantum Chemistry, Prague, June 20, 1994.

1. Introduction

Radiation-induced radicals formed in aqueous solution are known to react with nucleic acid bases and cause DNA damage.[1] Understanding the overall mechanism of DNA damage requires an accounting of the time dependence of base−radical formation. Identification of these radical species is generally attempted with transient absorption spectroscopy. However, little is known of the electronic properties of these radicals and the relative stabilities of the various isomers. We have already shown that the assignment of uracil adducts produced by reactive radicals are confused by significant differences in the electronic structure between isomers.[2] Identifying the isomer is an important object of the spectral assignment since the site of the unpaired electron is relevant to the mechanism of strand scission in radiation-damaged DNA.[3] Transient absorption spectra have been observed for OH adducts of cytosine,[4] but the preference for OH addition across the C5−C6 double bond was not determined spectroscopically. Adding the hydroxyl radical to the C5−C6 bond of cytosine leads to two isomers, C6OH and C5OH. These spectra were also reported earlier without isomeric assignment.[5,6] The addition of OH radicals to the C5 position in cytosine has been deduced to be favored.[1,4,7] Assignments of a C5OH adduct has been made from conductomeric measurements[7] and by indirect considerations drawn from electron spin resonance (EPR) ESR measurements.[8] Even though the reactive preference for the direct attack of the OH radical to pyrimidines is C5[1,7,9,10] calculations indicate that the C6OH isomers of uracil and thymine are more stable than the C5OH adducts.[2,11] The relative stability of the C6OH isomer is supported by the observation that the C5OH radical of thymine is not formed under thermodynamically controlled conditions, i.e., in the reaction of water with the radical cation.[12]

There is little information on the formation of hydrogen adducts at room temperature. However, the cytosine anion or its reversibly N3-protonated adduct is obtained in high probability in γ-irradiated DNA at low temperature.[13] Protonation at C6 is also reported,[13] but the yields are low and decrease rapidly as the temperature increases above 200 K. Simulation of the low-temperature EPR spectrum in aqueous glasses at low temperature shows that it can be assigned to both the C5 and C6 H adducts in a ratio of about 60% to 40%.[14] No electronic spectra have been reported for the H adducts.

The spectra of the OH radical adducts to cytosine show a prominent, broad absorption in the visible with a peak around 440 nm. When open shell orbitals can overlap and couple to relatively low-lying π or conjugated orbitals, low-energy excited electronic states result with the possibility of visible absorption. In the C6OH adduct of uracil, the site of the open-shell on C5 is next to the carbonyl bond. The open-shell electron can delocalize into the carbonyl π orbital, and a single transition is calculated in the visible for this isomer,[2] while the absorption for the C5OH isomer is in the UV because no such delocalization can take place. Analogously, the delocalization of the open-shell orbital to the C−C π bond in ethenyl peroxyl radical produces a visible transition for an unsaturated peroxyl while in all saturated peroxyl radicals the lowest transition is in the far UV.[15] Since the open-shell orbital is localized at C6 in the C5OH radical, our expectation is that the observed visible spectra should be assigned to the C6OH rather than the C5OH adduct. Electronically these isomers are very different as reflected in their redox properties for both uracil and cytosine.[16,17] Analysis of the π orbital structure of cytosine also suggests that there should be more than one visible transition, since there is a conjugated pair of π bonds in cytosine in contrast to uracil where the coupling is only to the exocyclic π orbital of the carbonyl bond. Cytosine provides another test of the hypothesis that the delocalization of the open-shell orbital

† This work was supported in part by the U.S. Public Health Service Grant CA 63317 to R. Osman.

⊗ Abstract published in *Advance ACS Abstracts,* May 1, 1997.

S1089-5639(96)02862-9 CCC: $14.00

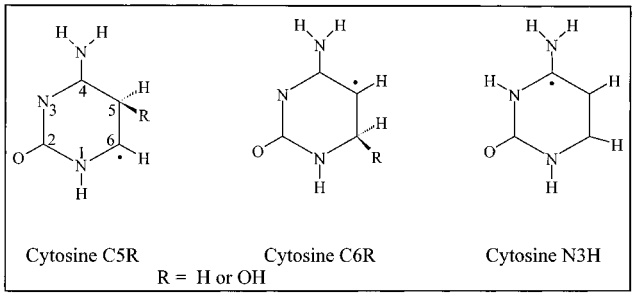


Figure 1. Schematic description of H and OH radical adducts.

supports the visible transitions in pyrimidine radicals, as well as other unsaturated radicals.

A hydrogen atom can be added to cytosine at the three sites depicted in Figure 1. The isomer formed by adding a H to C6 is denoted C6H. Three isomers are formed: C6H, C5H, and N3H. In this note we will focus on the spectroscopy of the adducts of H and OH to cytosine to expand our understanding of the electronic structure of the radical adducts and their excited states. The isomeric assignments were investigated for both H and OH adducts. The qualitative difference in predicted peak absorptions between the C5 and C6 H or OH adducts is so great that assignment of the spectra of the OH adduct is possible.

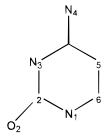
2. Method

The complete active space multiconfiguration self-consistent-field (CAS-MCSCF) method is applied here to the cytosine radicals. MCSCF calculations of the cytosine molecule have been reported[18] for a wide range of excited states. Where dynamical correlation plays an important role, the MCSCF solution must be augmented by a calculation of the dynamical electron correlation. However, explicit consideration of dynamical correlation is found to be less important for the lowest valence states[19,20] since excitations remained within the π orbital space. The differential electron correlation between the π and π^* orbitals is calculated to be small. The accuracy of the representation of the relevant excited states depends on the extent to which the active valence orbitals model the bond breaking that accompanies the excitation. These states are dominated by either transferring the open-shell orbital among those in the Hartree−Fock (HF) representation of the ground state or by the excitation to the lowest unoccupied orbital that correlates the bond or bonds being broken. It is assumed that the remaining dynamical correlation is constant and approximately cancels in the calculation of transition energies.

Energy gradient optimized ground-state structures in Table 1 are determined at the self-consistent-field (ROHF) level for all model systems using the GAMESS codes.[20] The heavy atom bond distances of the radicals are compared to those for neutral cytosine obtained by an RHF optimization. All coordinates are available upon request. The MC-SCF calculations used the CEP-31G basis for only the valence electrons and their concomitant effective core potentials[21] to replace the K-shell electron cores. For the OH adduct to C6, where the spectral predictions are compared to experiment, d polarization functions were added to all the heavy atoms.

The number of active orbitals was limited by requiring the occupation number of the smallest natural orbital (NO) always to exceed that of the next NO by at least a factor of 4 and the neglected NO to have a magnitude less than 0.03. Only two additional valence π or a″-type orbitals were used with thirteen doubly occupied and one open-shell orbitals of the Hartree−Fock configuration (13D1A2V). Orbitals are considered a″-type in C_1 symmetry cases where ring atoms are essentially in the *xy* plane and the orbitals are dominated by p_z functions. Antibonding π^* orbitals correlate the π-type orbitals. The excitation energies in Table 2 were obtained by averaging the density matrices of the ground and excited states by weighting the ground state with a coefficient of 0.5 equal to the sum of all equivalently weighted excited states considered in the calculation. This ensures that the orbital set used for both the ground and excited states is determined in the average MCSCF to be appropriate for both states and does not introduce a bias toward either the ground or excited state. Since both the C6OH and C5OH isomers are calculated to have transitions in the near UV that could be assigned to the experimental peak around 340 nm, the oscillator strength in the dipole approximation was calculated for the first three transitions of C6OH and the first two transitions of C5OH.

TABLE 1. Comparison of Bond Distances between Radical Adducts and Cytosine



						cytosine	
bond	N3H	C6H	C5H	C6OH	C5OH	calcd	exptl[a]
N1-C2	1.368	1.390	1.399	1.403	1.402	1.419	1.399
C2-O2	1.246	1.237	1.234	1.233	1.232	1.238	1.237
C2-N3	1.390	1.409	1.407	1.403	1.406	1.387	1.356
N3-C4	1.412	1.305	1.300	1.306	1.300	1.352	1.334
C4-N4	1.392	1.368	1.364	1.364	1.358	1.313	1.337
C4-N5	1.367	1.475	1.527	1.481	1.537	1.413	1.426
C5-C6	1.451	1.509	1.517	1.507	1.507	1.341	1.337
C6-N1	1.431	1.464	1.431	1.444	1.411	1.362	1.364

[a] Taylor, R.; Kennard, O. *J. Mol. Struct.* **1982**, *78*, 1.

TABLE 2. Cytosine Adduct Spectra (nm) and Dipole Moments (D)

state	1	2	3	4	5[a]
C6H		460.2	428.8	396.5	306.7
	7.30	5.21	6.13		
C5H		316.1	299.2		
	7.84	3.20			
N3H		390.8			
	4.39				
C6OH[b]		461.7	440.4	391.1	281.0
	5.76	3.43	4.97	3.13	
		425.2	417.1	353.0	274.5[c]
	5.63	4.39			
C5OH		340.6	315.5		
	4.61	1.61	3.06		

[a] State 1 is the ground state with the absorption wavelength for excited states in nm in the first row and dipole moments in the second row. [b] Experimental absorption peak at 430 nm, Hissung, A.; von Sontag, C.; *Z. Naturforsch.* **1978**, *88b*, 321. [c] Excitation wavelengths and dipole moments calculated with polarized basis set; all other values obtained with DZ basis.

Absorption shifts in water are estimated from a classical formula,[22]

$$\Delta E_{xe} = -\mu_x(\mu_e - \mu_x)a^{-3}[2(f(D) - g(n)]$$

where $f(D) = (D-1)/(2D+1)$ with D being the dielectric constant, g the refractive index, and x and e representing the ground and excited states, respectively. The classical molecular radius, a, was chosen to be 7*b*. Dipole moments of the ground and excited states were obtained by solving each state separately

in an MCSCF calculation. Only those states required for the qualitative analysis of the spectra were converged separately. The density of close lying states can make convergence difficult even when input vectors are chosen as the orbitals obtained from the average density solution.

3. Results and Discussion

The bond distances between the heavy atoms in Table 1 reflect the electronic structure of the radicals. The single, double, and conjugated bonds can be identified by comparison with the distances for neutral cytosine. Adding an H or OH radical across the C5−C6 double bond converts it into a single bond. The addition also disrupts the weak bond conjugation between the N3−C4 and C5−C6 bonds leading to a substantial increase in length for the C4−C5 bond. For the C6 adducts, this bond shows a smaller increase relative to the C5 adducts in which the open-shell orbital cannot be conjugated to the N3−C4 double bond. The C6−N1 bond distance, isolated from the localized open shell on the C6 adduct shows the opposite complementary behavior. Adding the H to N3 localizes the spin density on C6 with the double bond shifting to C4−C5 as can be seen from the corresponding changes in the bond lengths. This still places the open shell in a conjugating position with a double bond. Thus, the C5 adducts are expected to behave like alkyl radicals and support transitions in the far UV while the C6 and N3 adducts should have low-energy excited states.

At the MCSCF/CEP-31G level the C6OH isomer is 19 kJ/mol more stable than C5OH in vacuo while N3H is 24 and 26 kJ/mol more stable than C6H and C5H, respectively. Only 2 kJ/mol separate the in vacuo energies of C6H and C5H. Both OH adducts are polar, but the dipole moment of C6OH is somewhat larger as seen in Table 2 so that the OH adducts' energy difference will increase in water. These results are in good agreement with experimental findings that the C6OH adduct obtained from hydration of the radical cation[12] is more stable. Thermodynamic stability is distinct from the kinetically controlled addition of OH radicals where a preference for the C5OH adduct is observed.[5]

The N3H isomer is calculated to be most stable with C5H and C6H much higher in energy. In water, the C6H and C5H isomers are stabilized more than N3H, and all three isomers will get closer in energy. The accuracy of the relative energies is not easy to gauge. Higher level calculations would be required to provide more reliable relative energies of the isomers. The relative stability of the N3H adduct over the C5 or C6 adducts is also clearly observed.[13] However, the relative similarity in product distribution of the C5H and C6H adducts[14] cannot be used to support their comparable energies since the hydrogen additions are probably also kinetically controlled. The relative energetics among the OH and H adduct isomers are in agreement with previous calculations for uracil[2] and thymine.[11]

Three transitions are predicted within the visible or near UV range for the C6OH adduct (see Table 2). MCSCF solutions were obtained for the first three excited states at the DZ level to obtain consistent dipole moments for the estimation of the solvent effects. Substantial blue shifts are calculated in water for the first two transitions to 450 and 430 nm. These are very near the broad absorption peak maximum observed experimentally.[4] At the same ground state geometry, the MCSCF calculations with the polarized basis set predict in vacuo transitions at 425 and 417 nm, but the blue shift for the first excited state is to 421 nm. The shift for the second transition cannot be calculated because the third state cannot be converged because of the close proximity of the second and third states. On the basis of the DZ moments, an even smaller shift is expected for the transition to the second excited state. The shifts are smaller since the dipole moments for the ground and excited states are much closer. Thus, two transitions are predicted close to the peak of the observed spectra for either the DZ or DZd basis. Dipole oscillator strengths for the first two transitions in C6OH are calculated to be 0.0012 and 0.0387 so that the absorption to the second excited state dominates. The OH adducts are nonplanar, but the electronic character of the excited state wave functions retains the approximate symmetry of a planar molecule. The first excited state is an A′-like state while the second excited state has an approximate A″ symmetry. This assignment is supported by the symmetry of respective states in the planar C6H adduct. A large oscillator strength is expected and is found for a transition in cytosine from the A″-like ground state to the second excited state with a similar symmetry. Thus, the absorption in the visible is a convolution of the two transitions with the major intensity accounted for by the second transition to the A″-like excited state.

The transition to the third excited state could be assigned to the observed 340 nm peak. With the DZd basis, the MCSCF transition energy for this state is 353 nm. The oscillator strength calculated for this transition is only 0.0005, thus only a small contribution is possible to the observed peak. However, transitions to the first two excited states in C5OH can also be assigned to the 340 nm transition with a reasonable overlap between predicted and observed wavelengths. The first C5OH transition shifts in water to 320 from 341 nm, since the dipole moment for the first excited state is much reduced from that of the ground state. This transition contributes weakly to the absorption because the oscillator strength is 0.0003. The second excited state for C5OH is calculated to absorb at 316 nm with a shift in water to 306 nm but with a substantial oscillator strength of 0.0081. This transition is identified as the dominant contributor to the observed peak at 340 nm. The predicted absorption wavelengths for the dominant transitions of C6OH and C5OH are both several thousand wavenumbers (˜0.25 eV) to the blue of the observed transitions in water. This is comparable to both the direction and magnitude of the error of MCSCF predicted transition energies relative to observed absorptions in cytosine.[18]

Theoretical predictions for the H adducts in Table 2 show that the C6H and C5H isomers behave similarly to the comparable OH adducts. The transitions for C6H are predicted to be remarkably similar to C6OH and would be hard to distinguish from OH adducts expected to form in greater concentration. Since the N3H adduct is relatively the most stable, this should be observable when conditions are favorable for themodynamic equilibrium in the formation. It is clear that it is not observed under the experimental conditions of Hissung and von Sonntag.[4]

The theoretical transition energies and oscillator strengths show that the OH adduct isomers contribute differentially to the overall absorption. The C5OH adduct dominates the absorption peak at 340 nm while the C6OH contributes largely to the peak at 440 nm. It is difficult to determine which absorption is stronger since they are both broad and overlapping transitions. We will assume the areas are comparable. Based on the different redox properties of the isomers, 87% of the OH adducts have been deduced to be C5OH.[7] The ratio of the absorption transition probability values for the dominant transitions at 340 nm (C5OH) and 440 nm (C6OH) is calculated to be 0.15, which is fortuitously close to the observed ratio of the number density for the C6OH to C5OH adducts. The absorption coefficient is equal to the product of the transition energy, transition probability, and the number density of the absorbing

molecule. Since the ratio of the transition energies is about 1.3 and within the uncertainty in estimating the areas of the two transitions, it appears that the weaker transition calculated for the C5OH radical is compensated by the larger number density of this adduct. The predicted spectra and intensities lead to assignment of the dominant radical production to the C5OH adduct in agreement with kinetic measurements.[7] There is little experimental evidence on the relative thermodynamic stability of the cytosine isomers. Theoretical energies predict the C6OH isomer is most stable. But kinetic factors have determined the preference for C5OH in the direct addition of the radical. Nontheless, the absorption peak at 440nm is entirely due to the C6OH isomer. Since the separate absorption peaks are assigned to different isomers, it is possible to follow differences in the chemistry of the isomers by time resolved spectroscopy.

References and Notes

(1) von Sonntag, C. *Chemical Basis of Radiation Biology*, Taylor and Francis: London, 1987.

(2) Krauss, M.; Osman, R. *J. Phys. Chem.* **1993**, *97*, 13515.

(3) Hildenbrand, K.; Schulte-Frohlinde, D. *Int. J. Radiat. Biol.* **1989**, *55*, 725.

(4) Hissung, A.; von Sonntag, C. *Z. Naturforsch.* **1978**, *88b*, 321.

(5) Myers, L. S., Jr.; Hollis, M. L.; Theard, L. M.; Peterson, F. C.; Warnick, A. *J. Am. Chem. Soc.* **1970**, *92*, 2875.

(6) Hayon, E.; Simic, M. *J. Am. Chem. Soc.* **1973**, *95*, 1029.

(7) Hazra, D. K.; Steenken, S. *J. Am. Chem. Soc.* **1983**, *105*, 4380.

(8) Joshi, A.; Rustgi, S.; Riesz, P. *Int. J. Radiat. Biol.* **1976**, *30*, 151.

(9) Fujita, S.; Steenken, S. *J. Am. Chem. Soc.* **1981**, *103*, 2540.

(10) Das, S.; Deeble, D. J.; von Sonntag, C. *Z. Naturforsch.* **1985**, *40c*, 292.

(11) Osman, R.; Miaskiewicz, K.; Weinstein, H. In *Physical and Chemical Mechanisms in Molecular Radiation Biology*; Glass, W. A., Varma, M. N., Eds.; Plenum Press: New York, 1991; pp 423−452.

(12) von Sonntag, C. In *Physical and Chemical Mechanisms in Molecular Radiation Biology*; Glass, W. A., Varma, M. N., Eds.; Plenum Press: New York, 1991; pp 287−321.

(13) Wang, W.; Sevilla, M. D. *Rad. Res.* **1994**, *138*, 9 and references therein.

(14) Ohlmann, J.; Huttermann, J. *Int. J. Radiat. Biol.* **1993**, *63*, 427.

(15) Krauss, M.; Osman, R. *J. Phys. Chem.* **1995**, *99*, 11387.

(16) Deeble, D. J.; Das, S.; von Sonntag, C. *J. Phys. Chem.* **1985**, *89*, 5784.

(17) Steenken, S.; Jagannadham, V. *J. Am. Chem. Soc.* **1985**, *107*, 6818.

(18) Matos, J. M. O.; Roos, B. O. *J. Am. Chem. Soc.* **1988**, *110*, 7664.

(19) Krauss, M.; Garmer, D. R. *J. Phys. Chem.* **1993**, *97*, 831.

(20) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(21) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026.

(22) Suppan, P. *J. Photochem. Photobiol., A* **1990**, *50*, 293.