J|A|C|S
ARTICLES
Published on Web 02/26/2002

Theoretical Studies of the Cross-Linking Mechanisms between Cytosine and Tyrosine

Fuqiang Ban,[†] Maria J. Lundqvist,[‡] Russell J. Boyd,[†,*] and Leif A. Eriksson[§]

Contribution from the Department of Chemistry, Dalhousie University, Halifax, Nova Scotia B3H 4J3, Canada, Department of Biology and Chemical Engineering, Box 325, Mälardalen University, 631 05 Eskilstuna, Sweden, Department of Quantum Chemistry, Box 518, Uppsala University, 751 20 Uppsala, Sweden, and Department of Biochemistry, Box 576, Uppsala University, 751 23 Uppsala, Sweden

Received June 22, 2001

Abstract: DNA−protein cross-linking is one of the many DNA lesions mediated by hydroxyl radicals, the most damaging among the reactive oxygen species in biological systems. Density functional theory methods are employed to investigate the complex reaction mechanisms of the formation of cytosine−tyrosine cross-links as observed in γ-irradiated aqueous solutions of cytosine and tyrosine, as well as in γ-irradiated nucleohistone. The majority of the radical addition mechanisms considered are found to have significant barriers and therefore to be thermodynamically unfavorable for the formation of the initial cross-linked product. Our calculated reaction potential energy surfaces suggest that a feasible complete mechanism consists of radical combination forming the initial cross-linked product, a hydrogen shuffle within the initial cross-linked product, and an acid-catalyzed dehydration reaction. Water and hydrogen-bonding interactions are suggested to play a key role in catalyzing the hydrogen-transfer step of the reaction.

Introduction

Oxidative stress[1] mediated by reactive oxygen species (ROSs) is a disturbance in the oxidant−antioxidant balance in favor of the former, and contributes to numerous pathological conditions including cancer,[2−6] Alzheimer's disease[7,8] and Parkinson's disease.[9] The importance of radical chemistry in living systems is well recognized by biologists and chemists.[1] Large quantities of ROSs, such as singlet oxygen ($^1O_2$), hydrogen peroxide (H_2O_2), superoxide anion ($O_2^{\bullet-}$), hydroxyl radical ($^\bullet$OH), and hydroperoxyl radical (HOO$^\bullet$), can be generated from endogenous cellular metabolism and external perturbations (such as ionizing irradiation).[1,10] Of these, $^\bullet$OH is believed to be the most damaging.[11,12] Abstraction of hydrogen atoms by $^\bullet$OH from, or addition of $^\bullet$OH to, DNA bases, sugars, and amino acid residues of proteins produces huge numbers of secondary radicals. There can be up to 10^3 oxidative damaging events upon the DNA of each cell in the human body every day.[12] Extensive experimental studies have revealed that the interaction of the $^\bullet$OH radicals with DNA bases results in a diversity of adducts of purines and pyrimidines.[10,14,15] Tandem base lesions within isolated DNA are also mediated by $^\bullet$OH.[16,17]

The biological importance of oxidative damage to proteins has been considerably less well studied.[1] One reason for this is that research on protein radicals has focused primarily on their formation and functionality in enzymatic catalysis.[18] For instance, it is commonly known that tyrosyl radicals play important roles as reactive storage sites or metastable intermediates in a number of enzymes including ribonucleotide reductase[19] and photosystem II,[20,21] while little is known about the damage caused by tyrosyl radicals. However, recent research on proteins[22−27] has revealed that the actions of ROSs on proteins

* To whom correspondence should be addressed. E-mail: boyd@is.dal.ca.
† Dalhousie University.
‡ Mälardalen University and Department of Quantum Chemistry, Uppsala University.
§ Department of Biochemistry, Uppsala University.

(1) Halliwell, B.; Gutteridge, J. M. C. *Free Radicals in Biology and Medicine*; Oxford University Press Inc.: New York, 1999.
(2) Malins, D. C.; Polissar, N. L.; Gunselman, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2557.
(3) Ames, B. N.; Shigenaga, M. K.; Hagen, T. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7915.
(4) Loft, S.; Poulsen, H. E. *J. Mol. Med.* **1996**, *74*, 297.
(5) Totter, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1763.
(6) Ohshima, H.; Bartsch, H. *Mutat. Res.* **1994**, *305*, 253.
(7) Hsiao, K.; Chapman, P.; Nilsen, S.; Eckman, C.; Harigaya, Y.; Younkin, S.; Yang, F. S.; Cole, G. *Science* **1996**, *274*, 99.
(8) Selkoe, D. J. *Annu. Rev. Cell Biol.* **1994**, *10*, 373.
(9) Youdim, M. B. H.; Riederer, P. *Sci. Am.* **1997**, *276*, 52.
(10) von Sonntag, C. *The Chemical Basis of Radiation Biology*; Taylor & Francis: London, 1987.
(11) Chapman, J. D.; Reuvers, R. P.; Borsa, J.; Greenstock, C. L. *Radiat. Res.* **1973**, *56*, 291.
(12) de Lara, C. M.; Jenner, T. J.; Townsend, K. M. S.; Marsden, S. J.; O'Neill, P. *Radiat. Res.* **1995**, *144*, 43.
(13) Helbock, H. J.; Beckman, K. B.; Shigenaga, M. K.; Walter, P. B.; Woodall, A. A.; Yeo, H. C.; Ames, B. N. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 288.
(14) Dizdaroglu, M. *Mutat. Res.* **1992**, *275*, 331.
(15) Breen, A. P.; Murphy, J. A. *Free Radical Biol. Med.* **1995**, *18*, 1033.
(16) Bourdat, A. G.; Douki, T.; Frelon, S.; Gasparutto, D.; Cadet, J. *J. Am. Chem. Soc.* **2000**, *122*, 4549.
(17) Reid, T. M.; Loeb, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3904.
(18) (a) Stubbe, J.; van der Donk, W. A. *Chem. Rev.* **1998**, *98*, 705. (b) Stubbe, J. *Annu. Rev. Biochem.* **1989**, *58*, 257. (c) Stubbe, J. *Biochemistry* **1988**, *27*, 3593.
(19) Hoganson, C. W.; Sahlin, M.; Sjöberg, B.-M.; Babcock, G. T. *J. Am. Chem. Soc.* **1996**, *118*, 4672.
(20) Farrar, C. T.; Gerfen, G. J.; Griffin, R. G.; Force, D. A.; Britt, R. D. *J. Phys. Chem. B* **1997**, *101*, 6634.
(21) Dorlet, P.; Boussac, A.; Rutherford, A. W.; Un, S. *J. Phys. Chem. B* **1999**, *103*, 10945.

may lead to significant damage via hydrogen abstraction. In particular, the formation of tyrosyl radicals has been recognized as one typical feature of oxidative stress.[22,23] Further tyrosyl radical-mediated damage includes the formation of protein cross-links[28] via dityrosine and trityrosine, which possibly can serve as a marker for tyrosyl radical-mediated oxidative damage in atherosclerosis and other inflammatory conditions.

As DNA bases and proteins inside the nucleus are attacked by reactive oxygen species, the resulting protein radicals and DNA radicals may give rise to DNA−protein cross-links. In this respect, radiation biology[10] is well established to probe the oxidative effect of hydroxyl radicals. It has been shown that ionizing radiation produces DNA−protein cross-links in living cells and isolated chromatin,[29−31] which may cause serious problems in DNA replication. Even though great effort has been devoted to elucidating the mechanisms of DNA lesions and protein damage, the coupled reactions involving biological radical intermediates are poorly understood.

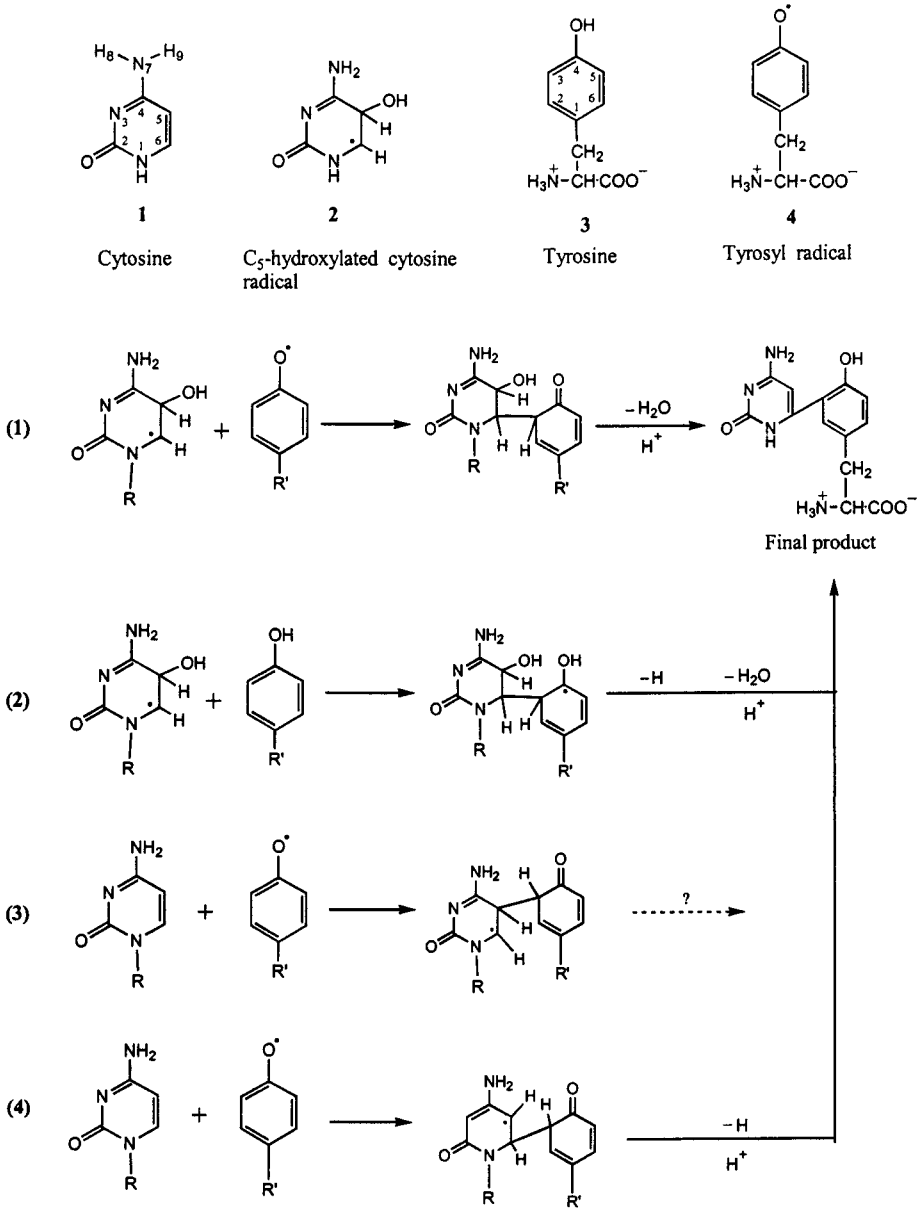
The cross-link structures[32−38] of model systems including pyrimidine bases and various amino acids in aqueous solution have been characterized via GC−MS techniques. In particlar, hydroxyl radical-induced cross-linking of cytosine with tyrosine,[35] thymine with tyrosine,[33] and thymine with lysine[37] has been studied in detail. For the cytosine + tyrosine or the thymine + tyrosine system, the molecular steps by which the •OH radical induces the DNA−protein cross-links are unclear. Several possible mechanisms based on the identified final products have been hypothesized.[33,35,38]

In this paper, we focus on the reaction mechanisms between cytosine and tyrosine upon irradiation. The formation of cytosine−tyrosine cross-links was observed in γ-irradiated aqueous solution of cytosine and tyrosine as well as in γ-irradiated nucleohistone.[35] The cross-link was proposed to involve formation of a covalent bond between C_6 of cytosine and C_3 of tyrosine (the numbering of the atoms of cytosine and tyrosine is shown in Scheme 1). It has been found[39] that the •OH adds to the C_5═C_6 bond of cytosine with a preference at C_5 (see **2** in Scheme 1) to the extent of 87%. In addition, the tyrosyl radical (**4**) in Scheme 1 has been found to be the major radical resulting from the addition of •OH to C_3 of tyrosine and subsequent water elimination.[10,40,41] Thus, the neutral cytosine, tyrosine, C_5-hydroxylated cytosine radical (hereafter denoted C(C_5OH)), and tyrosyl radical moieties are most likely the reactants that generate the Cyt−Tyr cross-link. Hence, the cross-linking mechanisms for the irradiated nucleohistone were proposed[35] as schematically shown in Scheme 1: (1) the combination of C(C_5OH) radical and tyrosyl radical, (2) the addition of C(C_5OH) radical to tyrosine, (3) the addition of the tyrosyl radical to C_5 of cytosine, and (4) the addition of the tyrosyl radical to C_6 of cytosine. The last two mechanisms were suggested to be the addition of **4** to the C_5═C_6 bond of cytosine.[35] Although only the radical combination mechanism (1) was suggested to account for the formation of the cross-link in γ-irradiated aqueous solution of cytosine and tyrosine, the radical addition mechanisms have been suggested as more likely in the case of nucleohistone reactions. In each of the mechanisms shown in Scheme 1, the final cross-link is formed through several steps of the reaction. As such, it is very important to examine the feasibility of each step of every reaction mechanism to obtain more insight into the likelihood of the complete reaction mechanisms.

Computationally, density functional theory (DFT) methods have been tested extensively against other accurate correlated methods for prediction of reliable geometries[42] and reaction barriers.[43] The reliability and efficiency of DFT methods offers an attractive theoretical approach to gaining insight into large biological systems.[44,45] In particular, the paramagnetic properties of numerous biological radicals derived from the DNA bases,[46a−d] sugars,[46e,47] and amino acids[48] have been extensively studied using density functional theory, and the reaction mechanisms of hydroxyl radical with imidazole, tyrosine, pyrimidine, and purine bases have been investigated using B3LYP, Hartree−Fock (HF), and MP2 methods.[49] The unrestricted Hartree−Fock (UHF) and MP2 methods suffer from severe spin contamination for radicals, whereas the B3LYP method is able to describe the potential energy surfaces, although it may have problems in locating transition structures in cases where the reaction barrier is very small.[49c]

(22) Diner, B. A.; Force, D. A.; Randall, D. W.; Britt, R. D. *Biochemistry* **1998**, *37*, 17931.
(23) Xiao, G. S.; Tsai, A. L.; Palmer, G.; Boyar, W. C.; Marshall, P. J.; Kulmacz, R. J. *Biochemistry* **1997**, *36*, 1836.
(24) Rauk, A.; Armstrong, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 4185.
(25) Shamovsky, I. L.; Riopelle, R. J.; Ross, G. M. *J. Phys. Chem. A* **2001**, *105*, 1061.
(26) Dean, R. T.; Fu, S. L.; Stocker, R.; Davies, M. J. *Biochem. J.* **1997**, *324*, 1.
(27) Liaw, S. H.; Villafranca, J. J.; Eisenberg, D. *Biochemistry* **1993**, *32*, 7999.
(28) Jacob, J. S.; Cistola, D. P.; Hsu, F. F.; Muzaffar, S.; Mueller, D. M.; Hazen, S. L.; Heinecke, J. W. *J. Biol. Chem.* **1996**, *271*, 19950.
(29) Oleinick, N. L.; Chiu, S.; Ramakrishnan, N.; Xue, L. *Br. J. Cancer* **1987**, *55*, 135.
(30) Mee, L. K.; Adelstein, S. J. *Int. J. Radiat. Biol.* **1979**, *36*, 359.
(31) Yamamoto, O. In *Aging, Carcinogenesis, and Radiation Biology*; Smith, K. C., Ed.; Plenum: New York, 1976; p 165.
(32) Gajewski, E.; Fuciarelli, A. F.; Dizdaroglu, M. *Int. J. Radiat. Biol.* **1988**, *54*, 445.
(33) Dizdaroglu, M.; Gajewski, E.; Reddy, P.; Margolis, S. A. *Biochemistry* **1989**, *28*, 3625. Margolis, S. A.; Coxon, B.; Gajewski, E.; Dizdaroglu, M. *Biochemistry* **1988**, *27*, 6353.
(34) Dizdaroglu, M.; Gajewski, E. *Cancer Res.* **1989**, *49*, 3463.
(35) Gajewski, E.; Dizdaroglu, M. *Biochemistry* **1990**, *29*, 977.
(36) Nackerdien, Z.; Rao, G.; Cacciuttolo, M. A.; Gajewski, E.; Dizdaroglu, M. *Biochemistry* **1991**, *30*, 4873.
(37) Morimoto, S.; Hatta, H.; Fujita, S.; Matsuyama, T.; Ueno, T.; Nishimoto, S. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 865.
(38) Charlton, T. S.; Ingelse, B. A.; Black, D. StC.; Craig, D. C.; Mason, K. E.; Duncan, M. W. *Free Radical Biol. Med.* **1999,** *27*, 254.
(39) Hazra, D. K.; Steenken, S. *J. Am. Chem. Soc.* **1983**, *105*, 4380.
(40) Solar, S.; Solar, W.; Getoff, N. *J. Phys. Chem.* **1984**, *88*, 2091.
(41) Land, E. J.; Ebert, M. *Trans. Faraday Soc.* **1967**, *63*, 1181.
(42) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.
(43) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936.
(44) Himo, F.; Siegbahn, P. E. M. *J. Phys. Chem. B* **2000**, *104*, 7502.
(45) Ban, F.; Gauld, J. W.; Boyd, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 7320.
(46) (a) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *J. Phys. Chem. B* **1998**, *102*, 5369. (b) Wetmore, S. D.; Himo, F.; Boyd, R. J.; Eriksson, L. A. *J. Phys. Chem. B* **1998**, *102*, 7484. (c) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *J. Phys. Chem. B* **1998**, *102*, 9332. (d) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *J. Phys. Chem. B* **1998**, *102*, 10602. (e) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *J. Phys. Chem. B* **1998**, *102*, 7674.
(47) Chatgilialoglu, C.; Ferreri, C.; Bazzanini, R.; Guerra, M.; Choi, S. Y.; Emanuel, C. J.; Horner, J. H.; Newcomb, M. *J. Am. Chem. Soc.* **2000**, *122*, 9525.
(48) (a) Ban, F.; Gauld, J. W.; Boyd, R. J. *J. Phys. Chem. A* **2000**, *104*, 5080. (b) Ban, F.; Gauld J. W.; Boyd, R. J. *J. Phys. Chem. A* **2000,** *104*, 8583. (c) Ban, F.; Wetmore, S. D.; Boyd, R. J. *J. Phys. Chem. A* **1999**, *103*, 4303. (d) Himo, F.; Gräslund, A.; Eriksson, L. A. *Biophys. J.* **1997**, *72*, 1556. (e) Himo, F.; Eriksson, L. A. *J. Phys. Chem. B* **1997**, *101*, 9811. (f) Himo, F.; Babcock, G. T.; Eriksson, L. A. *Chem. Phys. Lett.* **1999**, *313*, 374. (g) Himo, F.; Eriksson, L. A.; Blomberg, M. R. A.; Siegbahn, P. E. M. *Int. J. Quantum Chem.* **2000**, *76*, 714. (h) Himo, F. *Chem. Phys. Lett.* **2000**, *328*, 270.
(49) (a) Lundqvist, M. J.; Eriksson, L. A. *J. Phys. Chem. B* **2000**, *104*, 848. (b) Llano, J.; Eriksson, L. A. *J. Phys. Chem. B* **1999**, *103*, 5598. (c) Wetmore, S. D.; Boyd, R. J.; Llano, J.; Lundqvist, M. J.; Eriksson, L. A. In *Recent Advances in Density Functional Methods*; Barone, V., Bencini, A., Fantucci, P., Eds.; World Scientific: Singapore, Vol. 3, in press.

Scheme 1. Schematic Illustration of the Four Investigated Cross-Linking Mechanisms of Cytosine, Tyrosine, C_5-Hydroxylated Cytosyl Radical, and Tyrosyl Radical



In the present paper, density functional theory methods have been employed to investigate possible mechanisms of the Cyt−Tyr cross-linking. Since it has been reported that calculated structures and spin density distributions of phenoxyl radical and tyrosine side chain phenoxyl radical are essentially the same,[50] tyrosine and tyrosyl radical are modeled by phenol (**5**) and phenoxyl radical (**6**) (see Figure 1a) in the present study to reduce computational requirements.

Computational Methods

All geometry optimizations were performed with the B3LYP hybrid density functional in conjunction with the 6-31G(d,p) basis set using the GAUSSIAN 98 suite of programs.[51] The B3LYP functional is a combination of Becke's three-parameter hybrid exchange functional,[52] as implemented in GAUSSIAN 98,[53] and the Lee−Yang−Parr correlation functional.[54] Harmonic vibrational frequencies and zero-point vibrational energies (ZPVEs) were obtained at the same level of theory. Relative energies were obtained by performing single-point calculations at the B3LYP level in conjunction with the 6-311G(2df,p) basis set using the above optimized geometries and by including the zero-point vibrational energy, i.e., B3LYP/6-311G(2df,p)//B3LYP/6-31G(d,p) + ZPVE. The entropy contribution to the free energies at 298 K is derived from B3LYP/6-31G(d,p) frequency calculations.

(50) Qin, Y.; Wheeler, R. A. *J. Am. Chem. Soc*. **1995**, *117*, 6083.

(51) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *GAUSSIAN 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(52) Becke, A. D. *J. Chem. Phys*. **1993**, *98*, 1372.

(53) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem*. **1994**, *98*, 11623.

(54) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

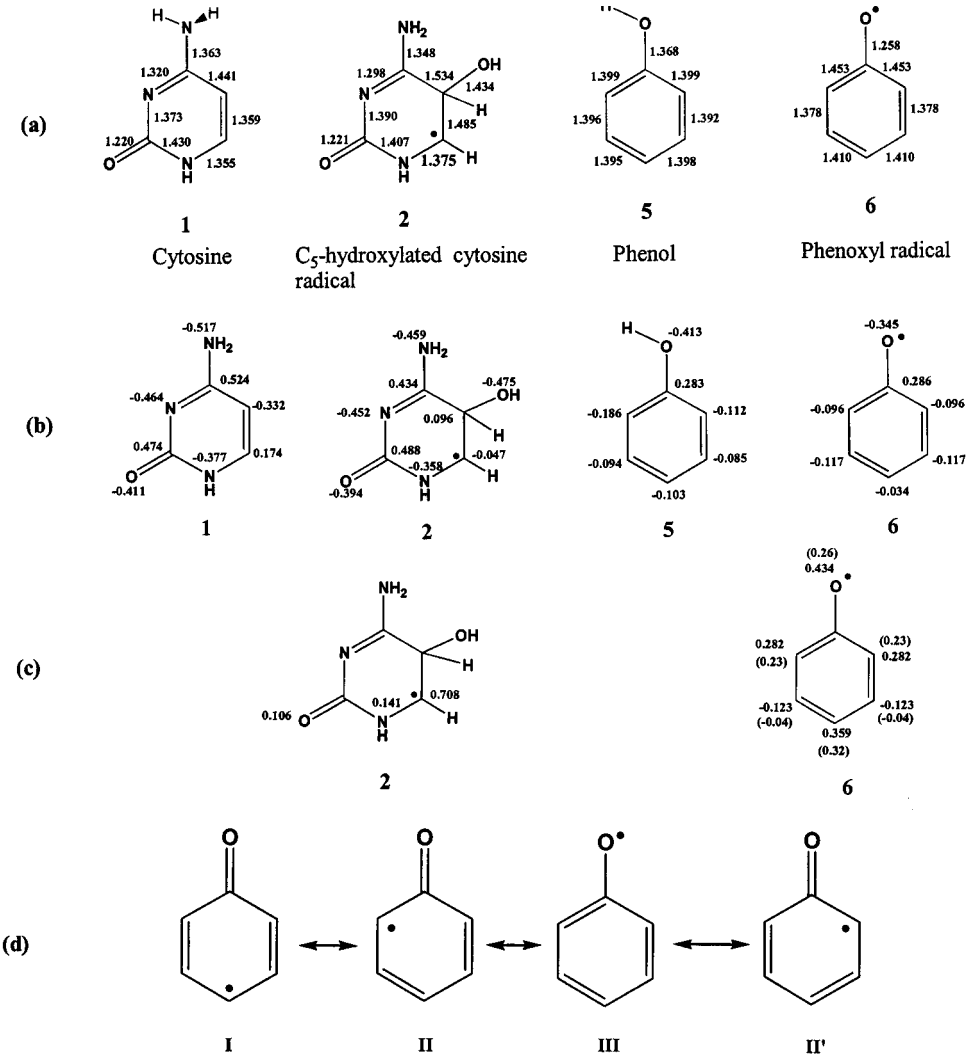


Figure 1. (a) Optimized structures (bond lengths) of cytosine (**1**), C_5-hydroxylated cytosyl radical (**2**), phenol (**5**), and phenoxyl radical (**6**). **(b)** Mulliken charges on the heavy atoms of cytosine, C_5-hydroxylated cytosyl radical, phenol, and phenxoyl radical. **(c)** Spin populations (only values numerically larger than 0.10 are shown) of C_5-hydroxylated cytosyl radical and phenxoyl radical. Values in parentheses are experimental estimates from ref 64. **(d)** Possible resonance structures of the phenoxyl radical.

The solvent effect on the cross-linking mechanisms was investigated by single-point calculations at the B3LYP/6-311G(2df,p) level on the optimized gas-phase structures using the polarizable dielectric model (DPCM)[55] as implemented in GAUSSIAN 98 with a dielectric constant of 78.39 for H_2O. For all open- and closed-shell systems, the unrestricted (UB3LYP) and restricted (RB3LYP) B3LYP procedures have been used, respectively. The symbols U and R have been neglected for simplicity. All energies are in kJ mol^{-1} and bond lengths in angstroms, unless otherwise specified. For the dimeric systems, we have used a slanted prime to denote atoms originally on the tyrosine (phenol) moiety.

Results and Discussion

Reactivity of the Reactants. The optimized structures of cytosine (**1**), C(C_5OH) radical (**2**), phenol (**5**), and phenoxyl radical (**6**) are shown schematically in Figure 1a. Mulliken charges on the heavy atoms of **1**, **2**, **5**, and **6** and atomic spin populations larger than 0.10 of radicals **2** and **5**, as calculated at the B3LYP/6-311G(2df,p)//B3LYP/6-31G(d,p) level of theory, are shown in parts b and c, respectively, of Figure 1.

The skeletal structure of cytosine is nearly planar. However, the amino group of cytosine is pyramidal by approximately 25° ($\angle(H_8N_7C_4H_9) = 154.7°$). Even though the cytosine ring shows a normal conjugated π-electron pattern, the C_5–C_6 bond (1.359 Å) has a localized double bond character that is much shorter than the C−C bond of benzene (1.396 Å).[56] Furthermore, the charge at C_5 is negative (−0.332), while at C_6 it is positive (+0.174). This particular charge distribution is a consequence of the position of the heteroatoms within the cytosine molecule. Since the addition of •OH to a polarizable double bond has remarkable selectivity toward an electron-rich center,[57] OH• addition to C_5 of cytosine dominates the reaction.[39]

The six-membered ring of **2** has a half-chair conformation with C_5 being the chair-back. The formation of the tetrahedral C_5 center in **2** weakens the N_1–C_6 and C_2–N_3 double bond character, but enhances the N_1−C_2 and N_3−C_4 double bonds. As a consequence, C_6 is pyramidal by 23° and becomes a localized radical center with a spin population of 0.708.

(55) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (b) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239. (c) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327.

(56) Cabana, A.; Bachand, J.; Giguere, J. *Can. J. Phys.* **1974**, *52*, 1949.

(57) (a) Steenken, S.; O'Neill, P. *J. Phys. Chem.* **1978**, *82*, 372. (b) Anbar, M.; Neta, P. *J. Phys. Chem.* **1966**, *70*, 2660.

5 is planar with typical aromatic C−C bond distances. The *ortho* C_3 and C_5 have larger negative Mulliken charges than the other carbons of the ring. This charge distribution of phenol explains the preference for the electrophilic addition to C_3/C_5 in the direct OH radical attack on tyrosine.[40,49a,58]

The structure of **6** has been carefully explored by various theoretical methods from AM1 to CASSCF/6-311G(2d,p) using a full π-active space of seven π-electrons correlated in seven π-orbitals.[59−61] It has been noted that the B3LYP/6-31G(d,p) geometry shows excellent agreement with the CASSCF/6-311G-(2d,p) geometry except that the C−O bond is longer by 0.030 Å. Due to the uncertainty associated with the restricted size of the active space, a too short C_4−O bond length of 1.228 Å was predicted at the CASSCF/6-311G(2d,p) level, essentially equal to the benzoquinone C_4−O bond length of 1.225 Å,[62] while the B3LYP/6-31G(d,p) level of theory was suggested to provide a more accurate structure for the phenoxyl radical.[59] The main geometrical changes from **5** to **6** are the shortening of the C_4−O, C_2−C_3, and C_5−C_6 bonds and the elongation of the C_1−C_2, C_3−C_4, C_4−C_5, and C_1−C_6 bonds. As noted previously,[60] the C_4−O bond length of 1.258 Å is much closer to the C═O bond length of 1.225 Å observed for benzoquinone[62] than to the calculated phenol C−O bond length of 1.368 Å or the value of 1.381 Å from electron diffraction measurements.[63]

We note that the calculated spin populations are in fair agreement with the experimental values, shown in parentheses in Figure 1c. The calculated spin population at the oxygen is approximately 0.434, while the experimental estimate is 0.26.[64] The large discrepancy between the values has also been noted earlier,[48d] and shown to arise primarily from an erroneous empirical value employed in the McConnell relation used to convert measured hyperfine parameters into spin populations. From the calculated spin populations and the experimental data, it is however clear that most of the unpaired electron is delocalized on the benzene ring. This agrees with the resonance structures of the radical, displayed in Figure 1d. Since the C_1 site is blocked in the tyrosyl radical (see **4** in Scheme 1), the equivalent C_3 or C_5 position is the most likely reactive site. This conclusion supports the proposal[35] that the cross-linking site is predominantly on the C_3 (C_5) carbon.

Initial Step of the Cross-Linking Mechanisms. The four mechanisms shown in Scheme 1 are considered as competitive reaction paths in the present study. The feasibility of the initial step of each reaction mechanism may determine its fate. Thus, we begin by investigating the potential energy surfaces for the initial step of each reaction mechanism.

The energy profile for the direct combination of **2** and **6** in forming the most stable conformer **1a** of the product is schematically shown in Figure 2a. All attempts to locate a transition structure (TS) for the formation of **1a,** including scans of the singlet and triplet energy surfaces, were unsuccessful. The radical combination forming the product **1a** is energetically favorable by 157.0 kJ mol^{-1}.

The energy profile for the addition of **2** to **5** in forming the most stable conformer of the product **2c** is schematically shown in Figure 2b. The C(C_5OH) radical initially interacts with the phenol, forming a hydrogen-bonded complex, **2a**, lying 10.1 kJ mol^{-1} lower in energy. The hydrogen-bonding distance between the −OH group of the C(C_5OH) radical and the oxygen of phenol is approximately 1.913 Å. The cross-linked product **2c** is formed via TS **2b** with an activation barrier of 74.6 kJ mol^{-1}. In **2b,** the distance between C_6 of the C(C_5OH) radical and C_3' of phenol is approximately 2.096 Å, and the −OH group of the cytosine moiety interacts with the oxygen of phenol at a shorter distance of 1.882 Å. Interestingly, despite the even stronger hydrogen bond (1.822 Å) in the product **2c** than in the addition complex **2a**, the product lies about 28 kJ mol^{-1} higher in energy than the isolated reactants.

The spin partitioning between the cytosine and the phenol moieties of **2a**, **2b**, and **2c** is shown in Table 1. The interaction between the two moieties is essentially electrostatic in **2a**, and the spin remains on the cytosine moiety. In the TS, approximately 0.428 of the spin is shifted to the phenol, whereas in the product, approximately 0.926 of the spin is localized on the phenol ring. In particular, the *meta* carbons C_2' and C_6' of the phenol possess approximately 0.886 of the spin, C_6' having the largest spin population of 0.498. Overall, the addition of the C(C_5OH) radical to phenol shifts the unpaired electron from the cytosine moiety to the *meta* carbon sites of the phenol ring. The highly localized double bond character of the C_1'═C_2' (1.366 Å) and C_4'═C_5' (1.368 Å) bonds suggests that the formation of the C_6−C_3' bond in **2c** deforms the π-electron conjugation of the phenol ring. As a consequence, the cross-linked product **2c** is higher in energy than the reactant system.

The energy profile for the addition of the phenoxyl radical to C_5 of cytosine to form the most stable product conformer **3c** is shown schematically in Figure 2c. Initially, the phenoxyl oxygen forms a hydrogen bond of 1.996 Å with the cytosine $-NH_2$ moiety. The resulting reactant complex **3a** lies 62.4 kJ mol^{-1} lower in energy than the isolated reactant system. The formation of the cross-linked product **3c** involves an activation barrier of 131.0 kJ mol^{-1} via TS **3b**. In **3b,** the cytosine C_5 approaches the C_3' of the phenoxyl radical at a distance of 1.862 Å. Product **3c** lies 66.3 kJ mol^{-1} higher in energy than the reactant system, and has a weak hydrogen bond (1.937 Å) that stabilizes the conformation.

The spin partitioning between the cytosine and the phenol moieties of **3a**, **3b**, and **3c** is shown in Table 1. The spin is essentially localized on the phenoxyl moiety in **3a**, while more than half (0.639) of the spin is delocalized onto the cytosine ring in the TS **3b**. The formation of the C_5–C_3' bond in the product results in the cytosine ring having a spin population of approximately 0.889, 0.698 of which is localized at C_6 of the cytosine moiety. The localized double bond character of C_1'═C_2' (1.346 Å) and C_5'═C_6' (1.356 Å) again indicates the loss of π-delocalization on the phenol ring. Due to the formation of a localized cytosine C_6 radical center, **3c** is higher in energy than the reactant system.

The energy profile for the addition of the phenoxyl radical to C_6 of cytosine to form the most stable conformer **4c** of the product is shown schematically in Figure 2d. The phenoxyl

(58) Raghavan, N. V.; Steenken, S. *J. Am. Chem. Soc.* **1980**, *102*, 3495.
(59) Wright, J. S.; Carpenter, D. C.; Mckay, D. J.; Ingold, K. U. *J. Am. Chem. Soc.* **1997**, *119*, 4245.
(60) Qin, Y.; Wheeler, R. A. *J. Chem. Phys.* **1995**, *102*, 1689−1698.
(61) Chipman, D. M.; Liu, R.; Zhou, X.; Pulay, P. *J. Chem. Phys.* **1994**, *100*, 5023.
(62) Hagen, K.; Hedberg, K. *J. Chem. Phys.* **1973**, *59*, 158.
(63) Portalone, G.; Schultz, G.; Domenicano, A.; Hargittai, I. *Chem. Phys. Lett.* **1992**, *197*, 482.
(64) Hulsebosch, R. J.; van den Brink, J. S.; Nieuwenhuis, S. A. M.; Gast, P.; Raap, J.; Lugtenburg, J.; Hoff, A. J. *J. Am. Chem. Soc.* **1997**, *119*, 8685.

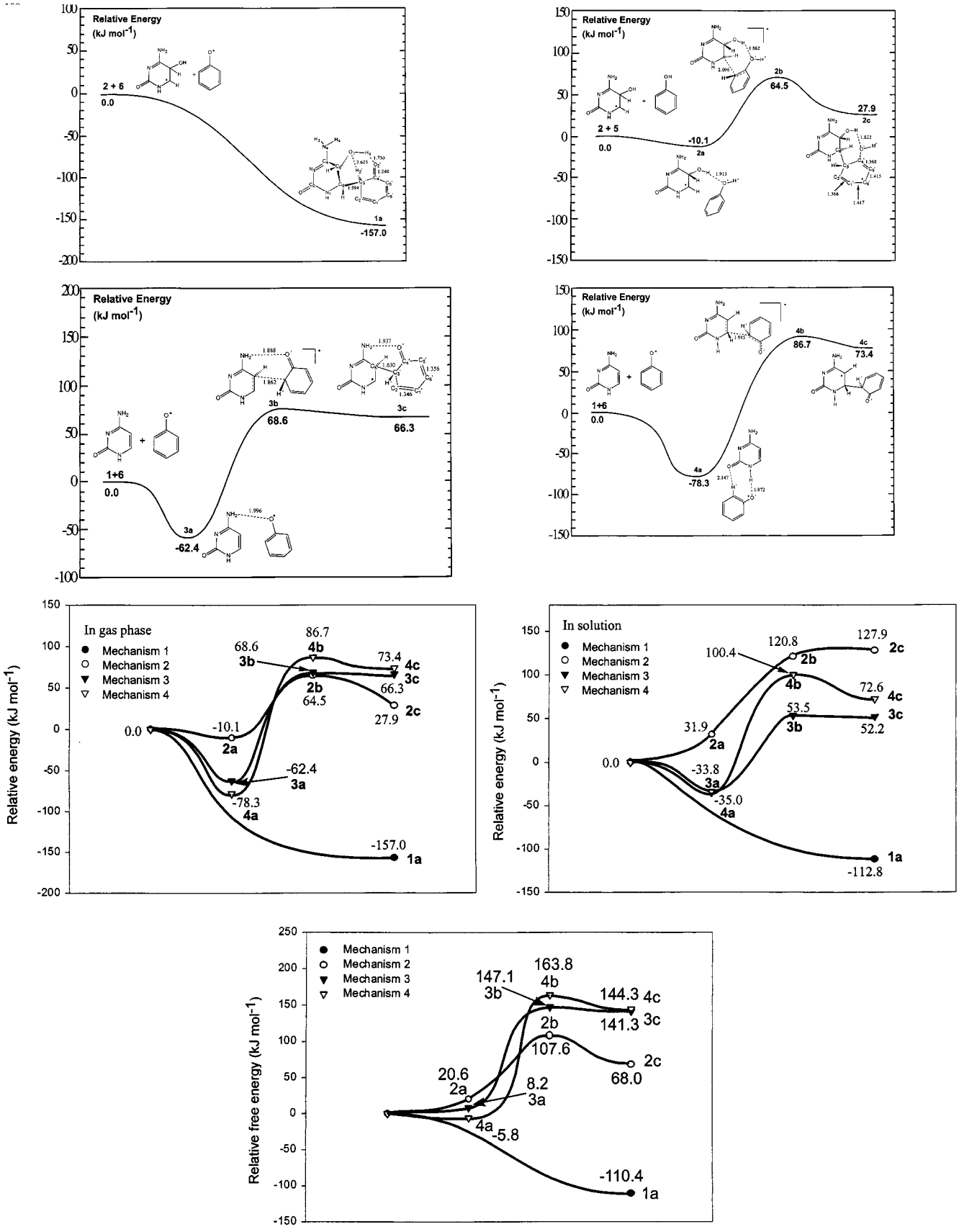


Figure 2. Calculated energy profiles for the initial reaction step of (a, top left**)** mechanism 1 (**2** + **6**), **(**b, top right**)** mechanism 2 (**2** + **5**), **(**c, second row left**)** mechanism 3 (**1** + **6**), addition to C_5 of cytosine, and **(**d, second row right**)** mechanism 4 (**1** + **6**), addition to C_6 of cytosine, **(**e, third row left**)** the summary of mechanisms 1, 2, 3, and 4 in the gas phase, and **(**f, third row right**)** the summary of mechanisms 1, 2, 3, and 4 showing the effects of the solvent. **(**g, bottom**)** Free energy profiles (gas phase) for the initial reaction steps at 298 K and 1 atm.

Table 1. Spin Partitioning between the Cytosine and Phenol Moieties in Compounds **2a**, **2b**, **2c**, **3a**, **3b**, **3c**, **4a**, **4b**, and **4c**

	spin population								
	2a	**2b**	**2c**	**3a**	**3b**	**3c**	**4a**	**4b**	**4c**
cytosine moiety	0.999	0.512	0.074	−0.002	0.639	0.889	−0.003	0.628	0.950
phenol moiety	0.001	0.428	0.926	1.002	0.361	0.111	1.003	0.372	0.050

radical and cytosine form an intermolecularly hydrogen bonded complex, **4a**, lying 78.3 kJ mol^{-1} lower in energy (Figure 2c). Product **4c** is formed via TS **4b** with a barrier of 165.0 kJ mol^{-1}. In **4b,** the C_6-C_3' distance is 1.915 Å. **4c** lies approximately 73.4 kJ mol^{-1} higher in energy than the reactants. Similar spin partitioning (Table 1) between the cytosine and phenoxyl moieties has been observed for **4a, 4b**, and **4c** as for **3a, 3b**, and **3c**. However, in **4c** the spin is now predominantly localized at the C_5 of cytosine.

We note that product **4c** is approximately 7.1 kJ mol^{-1} higher in energy than product **3c** (see Figure 2c), and more importantly, that TS **4b** is approximately 18.1 kJ mol^{-1} higher in energy than TS **3b**. Thus, the computed energy surfaces suggest that the addition of the tyrosyl radical to cytosine prefers the C_5 site over C_6, and it can be concluded that the addition is predominantly electrophilic in nature.

For comparison, the gas-phase potential energy surfaces of the initial steps are summarized in Figure 2e. As the original experiments were performed in aqueous solution, the effects of the solvent on the potential energy surface were investigated by single-point DPCM/B3LYP/6-311G(2df,p) calculations using a dielectric constant of 78.39 for water. The resulting energy profiles of the four mechanisms with solvation included are shown in Figure 2f. Solvation reduces the barriers of mechanisms 3 and 4 from 131.0 and 165.0 kJ mol^{-1} to 87.3 and 135.4 kJ mol^{-1}, respectively. In the case of mechanism 2, however, the energy cost required for the formation of **2c** increases from 74.6 kJ mol^{-1} in the gas phase to 127.9 kJ mol^{-1} with solvation. However, all the solvated products **1a**, **2c**, **3c**, and **4c** have significantly higher relative energies (−112.8, 127.9, 52.2, and 72.6 kJ mol^{-1}) than in the gas phase (−157.0, 27.9, 66.3, and 73.4 kJ mol^{-1}). According to the calculated energy profiles both in the gas phase and in solution, the initial steps of mechanisms 2, 3, and 4 have larger barriers than the initial step of mechanism 1, and are thermodynamically unfavorable. The energetically most favorable of these three is mechanism 3, which in the aqueous phase has a TS ca. 53 kJ mol^{-1} above the reactants. The calculated entropy (Table 2) changes of the initial steps of the four mechanisms are negative, and the free energy profiles (Figure 2g) of these reactions at 298 K suggest that the reactions are less favorable at 298 K than at 0 K. Of the four mechanisms proposed, the preferred mechanism for producing the cross-link between cytosine and tyrosine is the radical−radical addition. Thus, for the remainder of this paper we only consider the possible reaction steps that may lead to the final product from intermediate product **1a**.

Hydrogen Transfer. After the initial step of the radical−radical combination, three possible mechanisms for an internal hydrogen transfer from C_3' to O_2' within **1a** (Figure 2a) are considered: (A) direct hydrogen transfer from C_3' to O_2', (B) hydrogen transfer from C_3' to O_2' using a bridging water molecule, and (C) concerted hydrogen transfers from C_3' to O_1 and from O_1 to O_2' using a water molecule and the $-O_1H_1$ group

Table 2. Calculated Entropy (kJ mol^{-1} K^{-1}), Free Energy (kJ mol^{-1}), and Enthalpy (kJ mol^{-1}) Changes for Various Reaction Steps at 298 K and 1 atm

starting point	final point	ΔS_{298}°	ΔG_{298}°	ΔH_{298}°
2 + 6	**1a**	−0.202	−110.4	−170.5
2 + 5	**2a**	−0.101	20.6	−9.4
2 + 5	**2b**	−0.183	107.6	52.9
2 + 5	**2c**	−0.191	68.0	11.2
1 + 6	**3a**	−0.123	8.2	−28.5
1 + 6	**3b**	−0.186	147.1	91.6
1 + 6	**3c**	−0.176	141.3	88.7
1 + 6	**4a**	−0.141	−5.8	−47.9
1 + 6	**4b**	−0.170	163.8	113.1
1 + 6	**4c**	−0.156	144.3	97.9
1a	**1b**	−0.012	228.4	224.8
1a	**1c**	−0.007	−48.6	−50.7
1a + H_2O	**1d**	−0.128	10.8	−27.5
1a + H_2O	**1e**	−0.158	119.5	72.4
1a + H_2O	**1f**	−0.157	−74.6	−121.5
1a + H_2O	**1g**	−0.182	72.3	18.1
1c	**1h** + H_2O	0.169	−82.6	−32.1

of **1a** as bridges. The computed energy profiles for the three possible mechanisms are shown schematically in Figure 3a−c.

The calculated energy profile of mechanism A is schematically shown in Figure 3a. The hydrogen-transferred product **1c** can form via TS **1b** with an activation barrier of 227.3 kJ mol^{-1}. **1c** is 66.3 kJ mol^{-1} lower in energy than **1a**. We conclude that this pathway is unlikely due to the high barrier.

The structure of TS **1b** suggests that the high barrier is mainly due to strain in the four-membered ring $C_3'-H_2'-O_2'-C_4'$. As the reaction was carried out in aqueous solution, we speculate that water molecules may be specifically involved in the hydrogen-transfer reaction. Thus, mechanism B is proposed to investigate the role of a water molecule in the hydrogen transfer.

The computed energy profile of mechanism B is shown schematically in Figure 3b. Initially, **1a** interacts with a water molecule, forming hydrogen-bonded complex **1d** lying 24.7 kJ mol^{-1} lower in energy than the isolated reactants. Product complex **1f** is formed via a six-membered ring TS, **1e**. The activation barrier is approximately 109.0 kJ mol^{-1}. As expected it is much lower than the barrier 227.4 kJ mol^{-1} of mechanism A. **1f** lies 114.7 kJ mol^{-1} lower in energy. Isolation of **1c** + H_2O from **1f** costs 48.4 kJ mol^{-1}. In this particular situation the water molecule clearly acts as a catalyst for the hydrogen transfer. Thus, mechanism B is preferred over mechanism A.

We note that there is a strong hydrogen bond (1.730 Å) between O_1-H_1 and O_2' of **1a** (see Figure 2a). For this particular conformation, transfer of H_2' from C_3' to O_1 may induce a simultaneous transfer of H_1 from O_1 to O_2'. This possible mechanism is equivalent to a net hydrogen transfer from C_3' to O_2'. However, the distance between H_2' and O_1 is approximately 2.623 Å. Hence, it would be difficult for the direct transfer of H_2' from C_3' to O_1 to take place over such a distance. As we have learned, a water molecule may play a key role in catalyzing the hydrogen transfer in mechanism B. Hence, a water molecule may similarly catalyze the transfer of H_2' from C_3' to O_1. Thus,

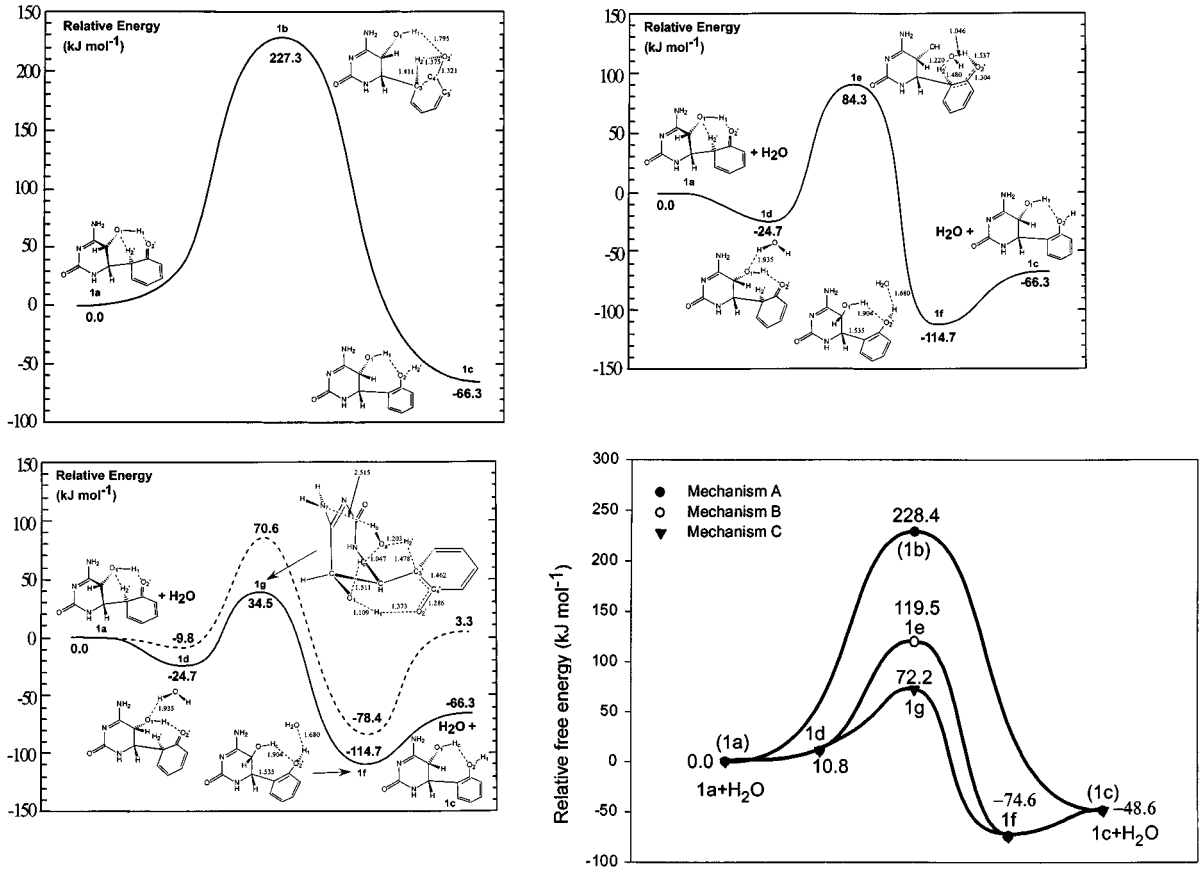


Figure 3. Calculated energy profiles of the hydrogen transfer from C_3' to O_2' (a, top left**)** for mechanism A, **(**b, top right**)** for mechanism B, and **(**c, bottom left**)** for mechanism C in the gas phase (solid line) and in aqueous solution (dashed line). **(**d, bottom right**)** Free energy profiles (gas phase) for the three hydrogen-transfer reactions at 298 K and 1 atm.

we investigated the feasibility of concerted hydrogen transfers along the chain $C_3'-O_{water}-O_1-O_2'$ of mechanism C.

The computed energy profile of mechanism C is shown in Figure 3c. As mentioned for mechanism B, **1a** and a water form the reactant complex **1d** lying 24.7 kJ mol^{-1} lower in energy. The formation of the product complex **1f** may proceed via TS **1g** with a barrier of 59.2 kJ mol^{-1}. The participation of a water molecule and the $-O_1H_1\cdots O_2'$ hydrogen bond in the reaction leads to the eight-membered ring structure $-C_3'---H_2'---O_a---H_c---O_1---H_1---O_2'---C_4'---$ **1g**. Because TS **1g** is 49.8 kJ mol^{-1}lower in energy than TS **1e**, we conclude that mechanism C is preferred over mechanism B.

Interestingly, TS **1g** possesses a hydrogen bond between H_b of the water and N_7 of the cytosine moiety, with a distance of 2.515 Å. Thus, the $H_b\cdots N_7$ hydrogen bond assists the concerted hydrogen transfers. The entropy changes at 298 K of the three hydrogen-transfer mechanisms are shown in Table 2. The computed free energy profiles of the reactions at 298 K (Figure 3d) suggest that the hydrogen transfer is less favorable at 298 K than at 0 K.

As the original experiments were performed in aqueous solution, solvent effects have been investigated by performing single-point DPCM/B3LYP/6-311G(2df,p) calculations. The barriers for mechanisms A, B, and C in solution are 279.1, 114.2, and 80.4 kJ mol^{-1}, respectively, which are higher than in the gas phase (227.3, 109.0, and 59.2 kJ mol^{-1}, respectively). We conclude that mechanism C is the preferred pathway both in the gas phase and in solution. Furthermore, the eight-membered ring structure of **1g** provides new insight for hydrogen-transfer mechanisms in biological systems. For simplicity, we only show the solvated energy profile of mechanism C (dashed curve in Figure 3c).

Acid-Catalyzed Dehydration of Intermediate 1c. Dehydration of **1c** leads to the final cross-linked product. Experimentally, 6 M HCl was added for the dehydration step.[35] Thus, the dehydration of **1c** should follow the standard reaction mechanism, involving protonation at O_1, loss of water at C_5, and deprotonation at C_6. The formation of the final product **1h** (see Figure 4) via dehydration of **1c** is energetically favored by 43.1 kJ mol^{-1}. The calculated entropy change at 298 K is 0.169 kJ mol^{-1} K^{-1} (Table 2). Thus, the free energy difference (Table 2) from 0 to 298 K is favorable for the dehydration.

Conclusions

The cross-linking reaction mechanisms of cytosine and tyrosine have been investigated by use of the B3LYP density functional theory method.

The direct combination of the C(C_5OH) radical and the phenoxyl radical is suggested to be the preferred mechanism in aqueous solution with a small or effectively no barrier, while all the initial steps of the radical addition mechanisms examined are found to be thermodynamically unfavorable.

Of the three possible mechanisms considered for the subsequent hydrogen transfer, mechanism C, by which a net hydrogen transfer from C_3' to O_2' of **1a** proceeds by a water-assisted chain of concerted hydrogen transfers in transition structure **1g**, is suggested to be the most favored.

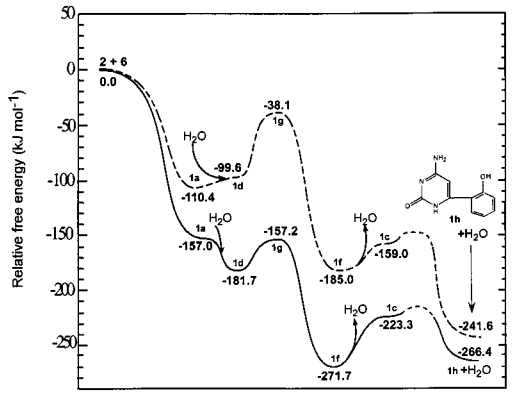


Figure 4. Free energy profiles at 0 K (solid line) and 298 K (dashed line) of the complete reaction mechanism in the gas phase at 1 atm.

It is furthermore found that bulk solvation has significant effects on the energetics of the cross-linking mechanisms. However, the preferred mechanism is suggested to be the same both in the gas phase and in solution.

In Figure 4, we show the calculated free energy profiles at 0 and 298 K of the most favorable energy surface for the overall mechanism leading to the final cross-linking product in an irradiated solution of cysotine and tyrosine. Note that in the segment of **1d** → **1f**, the water molecule is partaking explicitly in the complexes. The overall thermodynamics of the free energy profiles at the two temperatures are similar, but with differences of magnitude in the segment of **1d** → **1c**. Due to the entropy contribution to the free energy, temperature may be a key factor influencing the overall mechanism.

Acknowledgment. We gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC), the Killam Trusts, the KK Foundation, and the Swedish Research Council (VR) for financial support. F.B. thanks Drs. James W. Gauld and Zhenming Hu for helpful discussions. We thank the reviewers for valuable comments and suggestions.

Supporting Information Available: Archive entries of the B3LYP/6-31G(d,p) optimized structures. This material is available free of charge via the Internet at http://pub.acs.org.

JA011528M