JOC Article

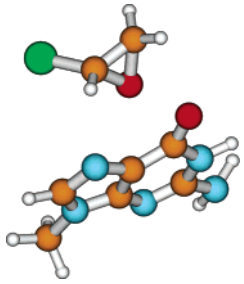
Chemical Reactivity as a Tool to Study Carcinogenicity: Reaction between Chloroethylene Oxide and Guanine

Urban Bren,[†] Mateja Zupan,[†] F. Peter Guengerich,[‡] and Janez Mavri*,[†]

National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia, and Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

janez@kihp2.cmm.ki.si

Received January 18, 2006



Chloroethylene oxide, an ultimate carcinogen of vinyl chloride, reacts with DNA giving rise to 7-(2-oxyethyl)guanine adduct in a nearly quantitative yield. This reaction represents an initial step of carcinogenesis associated with vinyl chloride. From experimental data for this reaction we calculated the second-order rate constant of 0.049 s^{-1} M^{-1}, which corresponds to the activation free energy of 19.5 kcal/mol. We also performed a series of medium high ab initio and density functional theory simulations. Effects of hydration were considered in the framework of the Langevine dipoles solvation model and the solvent reaction field method of Tomasi and co-workers. In silico calculated activation free energies are in a good agreement with the experimental value. This fact presents strong evidence in favor of the validity of the proposed reaction mechanism and points to the applicability of quantum-chemical methods to studies of other reactions associated with carcinogenesis. Insignificant stereoselectivity of the studied reaction was also predicted.

1. Introduction

Vinyl chloride (VC) is involved in the etiology of liver hemangiosarcoma and possibly other tumors of industrial workers.[1] Similar tumors can be induced in animal models thus making the VC a classic exogenous chemical carcinogen. Its neurotoxicity, immunotoxicity, hepatotoxicity, embryotoxicity, teratogenicity, and cardiovascular effects have also been reported.[2]

VC has been manufactured in large quantities for further processing, primarily to poly(vinyl chloride) (PVC). World production rose from 16 million tons in year 1980 to 27 million tons in year 1998. PVC has become indispensable in the building sector, automobile industry, agriculture, medical care, packaging, and electrical appliances. Small quantities of VC are also found in groundwater as a degradation product of chlorinated hydrocarbons and in tobacco smoke.[2]

Upon inhalative or oral uptake of VC, its oxidation by cytochrome P450 in the presence of oxygen and NADPH as cofactors gives rise to chloroethylene oxide (CEO, also known as 2-chlorooxirane), a highly electrophilic, short-lived epoxide. This is confirmed by studies in vitro[3] and in vivo.[4,5] This reaction takes place in human liver with P4502E1 as the major catalyst.[6]

* Author to whom correspondence should be addressed.
[†] National Institute of Chemistry.
[‡] Vanderbilt University School of Medicine.

(1) Forman, D.; Bennet, B.; Stafford, J.; Doll, R. *Br. J. Ind. Med.* **1985**, *42*, 750.
(2) http://www.inchem.org/documents/ehc/ehc/ehc215.htm Vinyl Chloride (EHC 215, 1999).
(3) Guengerich, F. P.; Crawford, W. M., Jr.; Watanabe, P. G. *Biochemistry* **1979**, *18*, 5177.
(4) Guengerich, F. P.; Mason, P. S.; Scott, W. J.; Fox, T. R.; Watanabe, P. G. *Cancer Res.* **1981**, *41*, 4391.
(5) Guengerich, F. P.; Watanabe, P. G. *Biochem. Pharmacol.* **1979**, *28*, 589.
(6) Guengerich, F. P.; Kim, D.-H.; Iwasaki, M. *Chem. Res. Toxicol.* **1991**, *4*, 168.

10.1021/jo060098l CCC: $33.50
Published on Web 05/02/2006

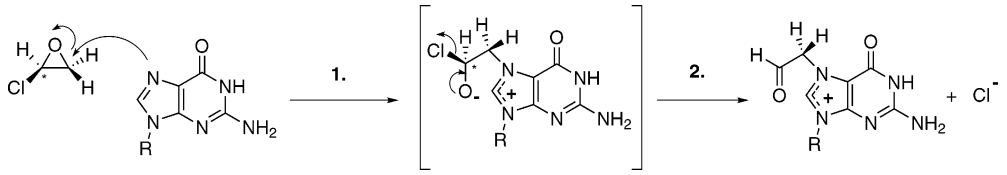


FIGURE 1. The proposed mechanism of the reaction between chloroethylene oxide and guanine giving rise to 7-(2-oxyethyl)guanine.

CEO is a genotoxic carcinogen that is capable of alkylating proteins and nucleic acid bases and shows similar toxicity/mutagenic profiles to its parent compound VC.[7] VC in this case represents a procarcinogen, while CEO is its ultimate carcinogen. This alkylation is followed by other reactions, of which depurination (leading to gene mutations and chromosomal aberrations) is a typical example. The major alkylation site is the N7 position of guanine (Gua), though other alkylation sites are reported. This reaction gives 7-(2-oxyethyl)guanine in a nearly quantitative yield.[8,9] The proposed reaction mechanism is depicted in Figure 1.

In this article we calculated the second-order rate constant (and the corresponding activation free energy) of the reaction between CEO and Gua from the available experimental data. In addition, we calculated the activation free energy of this reaction at several ab initio, density functional theory (DFT), and semiempirical molecular orbital (MO) levels. We incorporated the solvation effects by using the solvent reaction field method of Tomasi and co-workers[10] and the Langevin dipoles method of Florian and Warshel.[11] Moreover, the solution effects in conjunction with the semiempirical MO methods are studied at the AM1-SM1 and PM3-SM3 levels.

The organization of the article is as follows. In section 2 we present the calculation of the second-order rate constant (and the corresponding activation free energy) of the reaction between CEO and Gua from the experimental data. The applied computational methods are described in section 3. In section 4 we present the in silico calculated activation free energies of the reaction between CEO and Gua and compare them to the experimental value. Conclusions can be found in section 5.

2. Calculation of the Rate Constant from the Experimental Data

The rate constant of the reaction between CEO and Gua giving rise to 7-(2-oxyethyl)guanine has not been directly measured due to the instability of epoxide in aqueous solution and the lack of a clear signal for the reaction product. However, we were able to calculate it from the available experimental data.[3,9] The reaction of CEO with DNA is competitive with hydrolysis. We can neglect the other DNA adducts since 7-(2-oxyethyl)guanine represents more than 95% of the total DNA products[9] and write the following reaction scheme:

$$\mathrm{CEO} + \mathrm{H_2O} \xrightarrow{k_a} \text{glycolaldehyde}$$

$$\mathrm{CEO} + \mathrm{Gua} \xrightarrow{k_b} \text{7-(2-oxyethyl)guanine}$$

The corresponding rate equation for CEO is

$$\frac{\mathrm{d[CEO]}}{\mathrm{d}t} = -k_a[\mathrm{H_2O}][\mathrm{CEO}] - k_b[\mathrm{Gua}][\mathrm{CEO}] \qquad (1)$$

where [ ] indicates molar concentration and k_a and k_b represent corresponding second-order rate constants. This differential equation *cannot* be solved analytically.

Water represents the solvent and is therefore in great excess. Thus, we can treat the hydrolysis as a pseudo-first-order reaction and write

$$\frac{\mathrm{d[CEO]}}{\mathrm{d}t} = -k_a'[\mathrm{CEO}] - k_b[\mathrm{Gua}][\mathrm{CEO}] \qquad (2)$$

where $k_a' = k_a[\mathrm{H_2O}]$ represents the pseudo-first-order rate constant of hydrolysis, which was measured to be 0.012 s^{-1} (corresponding half-life was 60 s).[3]

In a yield experiment 10 mg of DNA (corresponding to 8.1 μmol of Gua) was reacted with 50 μmol of CEO in 2 mL of water to obtain 0.8 μmol of 7-(2-oxyethyl)guanine.[9] The remaining 49.2 μmol of CEO presumably reacted with water to yield glycolaldehyde.[3] From these data we can conclude that less than 10% of the total Gua reacted with CEO and consequently treat the reaction between CEO and Gua as a pseudo-first-order reaction by writing:

$$\frac{\mathrm{d[CEO]}}{\mathrm{d}t} = -k_a'[\mathrm{CEO}] - k_b'[\mathrm{CEO}] = -(k_a' + k_b')[\mathrm{CEO}] \qquad (3)$$

where $k_b' = k_b[\mathrm{Gua}]$ represents the corresponding pseudo-first-order rate constant.

In the case of two competing (pseudo) first-order reactions the quotient of their (pseudo) first-order rate constants equals the quotient of the amounts of their respective products:

$$\frac{k_a'}{k_b'} = \frac{[\text{glycolaldehyde}]}{[\text{7-(2-oxyethyl)guanine}]} \qquad (4)$$

This kinetic equation nicely accounts for the experimentally observed fact that, in this concentration range, the yield of 7-(2-oxyethyl)guanine was directly proportional to the amount of CEO used. The portion of 7-(2-oxyethyl)guanine among the products is constant, while the overall amount of products directly corresponds to the amount of CEO used. In addition, eq 4 allows for the determination of the pseudo-first-order rate constant of the reaction between CEO and Gua ($k_b' = 1.9 \times 10^{-4}$ s^{-1}). Using the definition $k_b' = k_b[\mathrm{Gua}]$, where [Gua] corresponds to the logarithmic average of Gua concentration ($[\mathrm{Gua}] = ([\mathrm{Gua}]^S - [\mathrm{Gua}]^F)/\ln([\mathrm{Gua}]^S/[\mathrm{Gua}]^F)$, where S refers to starting and F to final concentrations) we obtain the second-

(7) Zajdela, F.; Croisy, A.; Barbin, A.; Melaveille, C.; Tomatis, L.; Bartsch, H. *Cancer Res.* **1980**, *40*, 352.
(8) Scherer, E.; Van der Laken, C. J.; Gwinner, L. M.; Laib, R. J.; Emmelot, P. *Carcinogenesis* **1981**, *2*, 671.
(9) Guengerich, F. P. *Chem. Res. Toxicol.* **1992**, *5*, 2.
(10) Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1981**, *55*, 117.
(11) Florian, J.; Warshel A. *J. Phys. Chem. B* **1997**, *101*, 5583.

order rate constant of the reaction between CEO and Gua (k_b = 0.049 s^{-1} M^{-1}).

This section bears an important message. A common misconception would be that one needs a 10-fold excess of Gua over CEO to treat the reaction as a pseudo-first-order one. In the presented experiment we have quite the opposite regime: a 6-fold excess of CEO over Gua. Yet, without the use of the described simplifications, which not only yield an analytical solution to this kinetic problem but also bring a deeper understanding of the experimental results, we obtained a numerical solution of the system of differential equations that gives a second-order rate constant that differs from the analytical value by less than 0.003%. Therefore, to treat the reaction as a pseudo-first-order reaction, one only needs 90% preservation of one of the reactants. This is often the case when there is a competition for the second reactant with the solvent (as in the case of hydrolysis), because the second reactant is more likely to react with the solvent due to its 1000-fold excess.

The transition state theory of Eyring (based on the assumption that reactants and transition states are in chemical equilibrium) relates the rate constant of the reaction to the activation free energy ($\Delta G^{\ddagger}$) of the same reaction:

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{k_B T}\right) \quad (5)$$

where k_B represents the Boltzmann constant, h the Planck constant, and T the thermodynamic temperature. Using this equation we can calculate the experimental activation free energy of the reaction between CEO and Gua ($\Delta G_b^{\ddagger}$ = 19.5 kcal/mol). The validity of the transition state theory in biocatalysis was proved experimentally by the development of catalytic antibodies and theoretically by the success of the empirical valence bond (EVB) theory.[12−14] Recently, Chandler and co-workers performed a nice study of peptide isomerization in aqueous solution.[15] They treated the transition state as an ensemble of structures harvested on different transition paths rather than a single structure.[16] It remains a challenge to apply such methodology for treatment of biochemical reactions in solution.

3. Computational Methods

All calculations were performed at the National Institute of Chemistry in Ljubljana on the CROW clusters[17,18] consisting of about 270 Linux-based PCs running AMD Athlon processors at 700 MHz or AMD Opteron processors at 1.6 GHz. To obtain the Born−Oppenheimer (BO) hypersurface (and consequently the activation energy of the reaction between CEO and Gua) we applied a series of ab initio, DFT, and semiempirical MO calculations encoded in the Gaussian03 package.[19] For the reactants a full geometry optimization was performed. The transition state structure was located with the Berny algorithm. The difference between energies of the transition state and the reactants is the activation energy. Moreover, we performed vibrational analysis in the harmonic approximation and obtained only real frequencies for the reactants and a single imaginary frequency for the transition state at all levels of theory.

Calculation of the Born−Oppenheimer surface for chemical reactions is not a trivial task. It is generally accepted that one needs relatively flexible basis sets and adequate treatment of the electron correlation. The ab initio calculations were performed on the Hartree−Fock (HF) level of theory in conjunction with the 6-31G-(d), 6-31+G(d), and 6-311++G(d,p) basis sets and on the MP2 (Møller−Plesset perturbation method of the second order) level of theory, using the 6-31G(d) and 6-31+G(d) basis sets. In addition, we considered the DFT method B3LYP that has the Becke's three-parameter hybrid gradient corrected exchange functional[20] combined with the gradient-corrected correlation functional of Lee, Yang, and Parr.[21] Basis sets 6-31G(d), 6-31+G(d), and 6-311++G(d,p) were used. We are aware of the significant empirical character of the DFT methods, but they do to some extent include the electron correlation. Finally, we also applied the semiempirical MO methods AM1 and PM3. These two methods were used because of their low CPU cost, which facilitates their application in QM/MM methods and thermal averaging.

Solvation free energies of reactants and transition state were calculated with two methods, the solvent reaction field (SCRF) of Tomasi and co-workers[10] and the Langevin dipoles model (LD) parametrized by Florian and Warshel.[11] The SCRF method encoded in the Gaussian 03 package was applied at all ab initio and DFT levels. Merz−Kollman partial atomic charges calculated at all ab initio and DFT levels with the Gaussian03 package served as an input for the LD model built in the ChemSol program.[22] The AM1-SM1 and PM3-SM3 calculations were performed by the AMSOL-5.4.1 program of Truhlar and co-workers.[23]

4. Results and Discussion

The calculated activation energies, zero-point energies, and imaginary frequencies of the transition state structures are collected in Table 1. Table 2 shows the activation energies for the reaction between CEO and imidazole. Solvation free energies obtained by the SCRF method are presented in Table 3. Hydration free energies calculated with the LD model are shown in Table 4. The AM1-SM1 and PM3-SM3 based solvation free energies are collected in Table 5. Tables 3−5 also include activation free energies calculated as:

$$\Delta G^{\ddagger} = \Delta E^{\ddagger} + \Delta \mathrm{ZPE} + \Delta\Delta G_{\mathrm{hydr}} \quad (6)$$

where $\Delta E^{\ddagger}$ represents the activation energy, ΔZPE denotes the

(12) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; J. Wiley: Weinheim, Germany, 1997.
(13) Florian, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B* **2000**, *104*, 10092.
(14) Warshel, A.; Strajbl, M.; Villa, J.; Florian, J. *Biochemistry* **2000**, *39*, 14728.
(15) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci.* **2000**, *97*, 5877.
(16) Dellago, C.; Bolhuis, P. G.; Chandler D. *J. Chem. Phys.* **1999**, *110*, 6617.
(17) Borštnik, U.; Hodošček, M.; Janežič, D. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 359.
(18) Borštnik, U.; Janežič, D. *J. Chem. Inf. Model*. **2005**, *45*, 1600.
(19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Lahram, M. A.; Peng, C. Y.; Nanayakkara, A.; Challachombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzales, C.; Pople, J. A. *Gaussian 03*, Revision C02; Gaussian Inc.: Pittsburgh, PA, 2003.
(20) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(21) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

TABLE 1. Activation Energies Calculated for the Reaction between Chloroethylene Oxide and Guanine with Different Methods

method	$\Delta E^{\ddagger}$ [kcal/mol][a]	ZPE^{TS} [kcal/mol][b]	ZPE^{R} [kcal/mol][c]	ΔZPE [kcal/mol][d]	ω^{TS} [i cm^{-1}][e]	ω^{R} [cm^{-1}][f]	d^{TS} [Å][g]	d^{R} [Å][h]
AM1	45.03	121.92	123.49	−1.57	768	7.4	1.95	4.25
PM3	36.64	116.90	117.53	−0.63	842	6.7	1.98	4.44
HF/6-31G(d)	33.44	131.16	132.08	−0.92	621	8.0	2.10	4.35
HF/6-31+G(d)	32.78	130.84	131.76	−0.92	599	6.4	2.08	4.39
HF/6-311++G(d,p)	32.85	129.70	130.71	−1.01	597	6.9	2.09	4.39
B3LYP/6-31G(d)	19.78	121.58	122.15	−0.57	452	10.2	2.09	4.10
B3LYP/6-31+G(d)	19.31	121.21	121.81	−0.60	443	5.1	2.10	4.24
B3LYP/6-311++G(d,p)	18.83	120.36	121.02	−0.66	439	6.4	2.11	4.21
MP2/6-31G(d)	31.97	123.23	123.87	−0.64	592	20.0	1.97	4.35
MP2/6-31+G(d)	33.89	122.34	123.07	−0.73	612	12.9	1.94	4.40

[a] Classical activation energy. [b] Zero point vibrational energy of the transition state. [c] Zero point vibrational energy of the reactants. [d] Zero point energy of the transition state minus zero point energy of the reactants. [e] Imaginary frequency value corresponding to the transition state. [f] Lowest frequency value corresponding to the reactant structure. [g] Distance between the N7 atom of guanine and the nonchiral carbon of chloroethylene oxide for the transition state. [h] Distance between the N7 atom of guanine and the nonchiral carbon of chloroethylene oxide for the reactant structure.

TABLE 2. Activation Energies Calculated for the Reaction between Chloroethylene Oxide and Imidazole with Different Methods

basis set	$E^{\ddagger}_{MP2}$ [kcal/mol][a]	$E^{\ddagger}_{CCSD(T)}$ [kcal/mol][b]
6-31G(d)	23.65	23.06
6-31+G(d)	21.97	21.48
6-311++G(d,p)	23.12	22.33

[a] Activation energy obtained at the MP2 level of theory. [b] Activation energy obtained at the CCSD(T) level of theory.

TABLE 3. Activation Free Energies Calculated for the Reaction between Chloroethylene Oxide and Guanine with the Solvent Reaction Field (SCRF) Method ($\Delta G^{\ddagger}_{exp}$ = 19.5 kcal/mol)

method	ΔG^{SCRF}_{hydr} TS [kcal/mol][a]	ΔG^{SCRF}_{hydr} R [kcal/mol][b]	$\Delta\Delta G^{SCRF}_{hydr}$ [kcal/mol][c]	$\Delta G^{\ddagger}_{SCRF}$ [kcal/mol][d]
HF/6-31G(d)	−27.50	−14.29	−13.21	19.31
HF/6-31+G(d)	−31.04	−16.28	−14.76	17.10
HF/6-311++G(d,p)	−30.38	−15.80	−14.58	17.26
B3LYP/6-31G(d)	−21.91	−11.76	−10.15	9.06
B3LYP/6-31+G(d)	−25.97	−14.72	−11.25	7.46
B3LYP/6-311++G(d,p)	−25.31	−14.41	−10.90	7.27
MP2/6-31G(d)	−31.37	−15.91	−15.46	15.87
MP2/6-31+G(d)	−36.68	−18.80	−17.88	15.28

[a] Hydration free energy of the transition state obtained with the SCRF method. [b] Hydration free energy of the reactants calculated with the SCRF method. [c] Hydration free energy of the transition state minus hydration free energy of the reactants. [d] Activation free energy.

TABLE 4. Activation Free Energies Calculated for the Reaction between Chloroethylene Oxide and Guanine with the Langevin Dipoles (LD) Method ($\Delta G^{\ddagger}_{exp}$ = 19.5 kcal/mol)

method	ΔG^{LD}_{hydr} TS [kcal/mol][a]	ΔG^{LD}_{hydr} R [kcal/mol][b]	$\Delta\Delta G^{LD}_{hydr}$ [kcal/mol][c]	$\Delta G^{\ddagger}_{LD}$ [kcal/mol][d]
HF/6-31G(d)	−28.62	−20.79	−7.83	24.69
HF/6-31+G(d)	−30.40	−21.51	−8.89	22.97
HF/6-311++G(d,p)	−29.98	−21.01	−8.97	22.87
B3LYP/6-31G(d)	−24.47	−18.99	−5.48	13.73
B3LYP/6-31+G(d)	−26.39	−20.00	−6.39	12.32
B3LYP/6-311++G(d,p)	−25.75	−19.94	−5.81	12.36
MP2/6-31G(d)	−31.01	−18.65	−12.36	18.97
MP2/6-31+G(d)	−31.61	−18.94	−12.67	20.49

[a] Hydration free energy of the transition state calculated with the LD method. [b] Hydration free energy of the reactants calculated with the LD method. [c] Hydration free energy of the transition state minus hydration free energy of the reactants. [d] Activation free energy.

TABLE 5. Activation Free Energies Calculated for the Reaction between Chloroethylene Oxide and Guanine with the AM1-SM1 and PM3-SM3 Methods ($\Delta G^{\ddagger}_{exp}$ = 19.5 kcal/mol)

method	ΔG^{TS}_{hydr} [kcal/mol][a]	ΔG^{R}_{hydr} [kcal/mol][b]	$\Delta\Delta G_{hydr}$ [kcal/mol][c]	$\Delta G^{\ddagger}$ [kcal/mol][d]
AM1-SM1	−36.09	−25.54	−10.55	32.91
PM3-SM3	−41.85	−30.69	−11.16	24.85

[a] Hydration free energy of the transition state calculated with the AM1-SM1 and PM3-SM3 methods. [b] Hydration free energy of the reactants calculated with the AM1-SM1 and PM3-SM3 methods. [c] Hydration free energy of the transition state minus hydration free energy of the reactants. [d] Activation free energy.

relative zero-point energy of the transition state and the reactants, and $\Delta\Delta G_{hydr}$ stands for the corresponding relative hydration free energy. The entropic contribution is frequently calculated as the sum of translational, rotational, and vibrational entropies calculated with the ideal gas, rigid rotor, and harmonic oscillator approximations.[24] However, it has been argued that the use of these approximations often leads to a dramatic overestimation of the entropy term for reactions in solution, which could in fact be neglected.[25] The most important reason for the observed overestimation lies in the fact that the harmonic approximation underestimates the entropy contribution from the low-frequency modes that are more abundant in larger solutes. Moreover, the solvation free energy calculated by implicit solvation models neglects the changes in the solute entropy upon the transfer from the gas phase to the solution.

Our in silico calculations focus only on the first part of the reaction depicted in Figure 1 as this S_N2 substitution represents the rate limiting step of the whole reaction. It leads to the formation of the unstable intermediate with tetrahedral coordination of the chiral carbon atom. The activation free energy is defined as a free energy difference between the transition state and the reactants. To obtain the activation free energy of this first step we therefore need only to consider its reactants (CEO

(22) Florian, J.; Aqvist, J.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 11524.
(23) Hawkins, G. D.; Lynch, G. C.; Giesen, D. J.; Rossi, I.; Storer, J. W.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *AMSOL* version 5.4.1; University of Minnesota.
(24) Kuhn, B.; Kollman, P. A. *J. Am. Chem. Soc.* **2000**, *122*, 2586.
(25) Strajbl, M.; Florian, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 4471.

and Gua) and its transition state, i.e., the saddle point (characterized by a single imaginary frequency) on the potential energy surface connecting the reactants and the unstable intermediate.

In the second part of the same reaction the elimination of the chlorine ion from the intermediate takes place yielding 7-(2-oxyethyl)guanine. The activation free energy of this reaction step must be smaller than that of the corresponding elimination in the basic hydrolysis of methyl acetate (experimentally measured 7.4 kcal/mol[26]), because chlorine ion represents a far better leaving group than methoxide.[27] This elimination cannot be rate determining because the experimental activation free energy of the reaction between CEO and Gua reaches 19.5 kcal/mol.

From results of ab initio HF calculations of the activation energy collected in Table 1 it is evident that the convergence in terms of basis set size was reached. The addition of diffuse and polarization functions on heavy atoms (6-31+G(d) basis set) is crucial for obtaining the converged barrier in terms of the basis set size. The predicted reaction barrier at the HF level lies between 32 and 34 kcal/mol. Almost identical activation energies were obtained at the ab initio MP2 level. These calculations were CPU demanding due to the large size of the studied system. For example, the transition state structure search at the MP2/6-31+G(d) level took over one month of CPU time on a Linux-based PC running two AMD Opteron processors at 1.6 GHz. DFT calculations drastically reduce the reaction barrier, but semiempirical MO method AM1 increases it. By using the semiempirical MO method PM3, which was demonstrated to yield reasonable energetics for the reaction catalyzed by xylose isomerase,[28] an activation energy similar to the HF level was obtained.

The zero-point vibrational energy correction of the reaction barrier is almost negligible (Table 1). In addition, it should be noted that the DFT- and MP2-calculated BO surfaces are shallower than the HF-calculated ones. This fact is reflected in the values of the zero-point vibrational energies. A single imaginary vibrational frequency was obtained for the transition state structure at all theory levels. Values of corresponding imaginary frequencies are also collected in Table 1. To check whether the correct transition structure was found, we performed a visualization of the vibration mode of this imaginary frequency using the MOLDEN software[29] because it should correspond to the reaction coordinate of the first step of the reaction mechanism depicted in Figure 1. At all theory levels the vibration mode of this imaginary frequency coincided with the formation of a chemical bond between the N7 atom of Gua and the nonchiral carbon of CEO and with the cleavage of the chemical bond connecting this nonchiral CEO carbon atom to the epoxide oxygen, thus confirming the allocation of the correct transition state structure. The structures of the reactants and the transition state calculated at the MP2/6-31+G(d) level of theory are presented in Figure 2.

The elimination step (denoted by 2 in Figure 1) was studied at the HF/6-31G(d) and B3LYP/6-31G(d) levels of theory. Respective transition state structures served as a starting point.

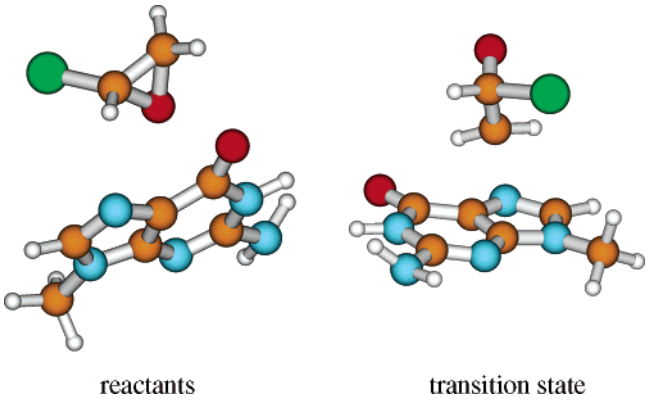


FIGURE 2. The structures of the reactants and the transition state of the chemical reaction between chloroethylene oxide and guanine calculated by using the MP2/6-31+G(d) level of theory. Oxygen is depicted in red, carbon in orange, nitrogen in blue, chlorine in green, and hydrogen in white.

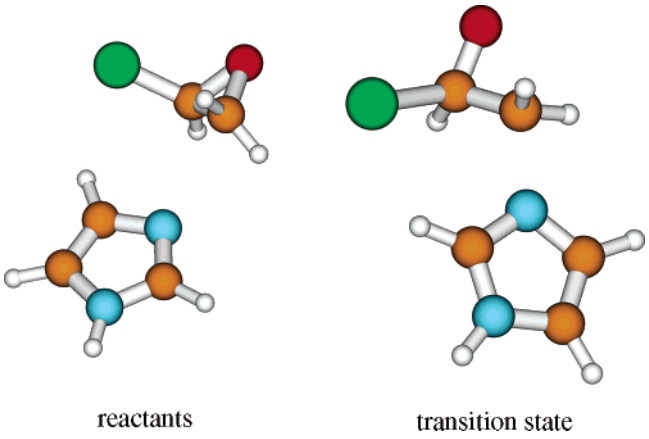


FIGURE 3. The structures of the reactants and the transition state of the chemical reaction between chloroethylene oxide and imidazole calculated by using the MP2/6-31G(d) level of theory. Oxygen is depicted in red, carbon in orange, nitrogen in blue, chlorine in green, and hydrogen in white.

Then, the distance between the N7 atom of Gua and the nonchiral carbon of CEO was shortened by 0.1 Å and a full geometry optimization was performed. In both simulations the product state characterized by the dissociated chlorine ion resulted. In this way the reactive system was moved through the proposed intermediate directly to the product valley thus indirectly indicating an insignificant energy barrier for the elimination step of the reaction.

To prove that good agreement between the calculated and experimental reaction barrier is not associated with coincidental cancellation of errors we truncated guanine to imidazole that allows for advanced application of post Hartree−Fock methodology. First, we performed a full geometry optimization and transition state search at the MP2/6-31G(d) level of theory. The resulting reactant and transition state structures for the reaction between CEO and imidazole are depicted in Figure 3. For these geometries subsequent single point calculations at the CCSD(T) level of theory with 6-31G(d), 6-31+G(d), and 6-311++G(d,p) basis sets were performed. The results are presented in Table 2. All in all, activation energies for the reaction between CEO and imidazole at the MP2 level compare favorably with the ones obtained at the CCSD(T) level of theory. The

(26) Gillan, C. J.; Knipe, A. C.; Watts, W. E. *Tetrahedron Lett.* **1981**, *22*, 597.
(27) Isaacs, N. S. *Physical Organic Chemistry*; Longman Scientific & Technical: Harlow, England, 1987.
(28) Garcia-Viloca, M.; Alhambra, C.; Truhlar, D. G.; Gao, J. *J. Comput. Chem.* **2003**, *24*, 177.
(29) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123.

convergence of reaction barriers is therefore achieved already at the MP2 level. Numerical comparison of the presented activation energies with the ones calculated for the reaction between CEO and Gua is, however, meaningless since it compares two different reacting systems.

In Table 3 we collected hydration free energies calculated by the SCRF method of Tomasi and co-workers. The solvent accelerates the reaction because the transition state is better solvated than the reactants. This finding reflects the formation of the zwitterionic intermediate in the first step of the reaction depicted in Figure 1. Reduction of the reaction barrier in terms of the hydration free energies is the smallest for DFT methods and the largest for MP2 methods. All in all, the SCRF model seems to underestimate the activation free energies. Very good agreement with the experimental activation free energy of 19.5 kcal/mol was obtained only at the HF/6-31G(d) level. This level of theory nicely accounts for the polarization effects of the solvent, which is why it was used for development of partial atomic charges in the construction of the AMBER force field.[30]

All presented results were obtained for the *R* stereoisomer of CEO. To check for the stereoselectivity of the reaction between CEO and Gua we recalculated its activation free energy using the SCRF model at the HF/6-31G(d) level of theory for the *S* stereoisomer of CEO. The reaction barriers were 19.20 and 19.31 kcal/mol for the *S* and *R* stereoisomers, respectively, thus making the stereoselectivity of the reaction insignificant. This fact can be easily understood in view of the nearly planar character of Gua. The experimental work involved what is presumed to be a racemate, prepared by photochemical chlorination of ethylene oxide.

Table 4 shows hydration free energies obtained by the LD solvation model of Florian and Warshel. Reduction of the reaction barrier in terms of the hydration free energies is smaller relative to the corresponding SCRF values. Again, the largest reduction is obtained for MP2 and the smallest for DFT methods. MP2 calculations in conjunction with the LD solvation model give activation free energies that are in very good agreement with the experimental value of 19.5 kcal/mol. HF reaction barriers are moderately overestimated and B3LYP reaction barriers are significantly underestimated.

All in all, the B3LYP functional regardless of the applied solvation model systematically underestimates the activation free energy of the reaction between CEO and Gua. It appears that for the studied system the LD solvation model outperforms the SCRF method, though there are no experimental data available on hydration free energies of the studied species. The discrepancy between the solvation free energies obtained by both models could be rationalized by the fact that there were no epoxy species nor zwitterions used in their parametrization sets.[11,31]

Table 5 presents the results of AM1-SM1 and PM3-SM3 methods. The AM1-SM1 level of theory significantly overestimates the activation free energy of the reaction between CEO and Gua. The reaction barrier obtained at the PM3-SM3 level of theory is slightly overestimated, but presents a good compromise in terms of the required CPU time and the quality of the results.

Disagreement between the experimental and calculated free energies could also be explained by considering only a part of DNA (Gua) and not treating water and counterions in atomic detail. In our study we applied only some of the available quantum-chemical methods. There is still room for the use of QM/MM methodology with all-atom representation of the polar environment and application of thermal averaging.[32−34] In addition, it would be a challenge to test novel DFT functionals[35] or to reparametrize the semiempirical methods as reported by Truhlar and co-workers.[36]

5. Conclusion

In this article we present a study of a chemical reaction between chloroethylene oxide, the ultimate carcinogen formed from vinyl chloride, and Gua giving rise to 7-(2-oxyethyl)-guanine, the major adduct formed. First we calculated the activation free energy of this reaction from the experimental data. In addition, we applied quantum-chemical calculations to the rate-limiting step of the proposed reaction mechanism. We demonstrated that the MP2 level of theory, which takes into account a large part of the electron correlation, in conjunction with flexible basis sets and LD solvation model gives a good agreement with the experimental activation free energy. This agreement presents strong evidence in favor of the validity of the proposed reaction mechanism. It also points to the applicability of quantum-chemical methods to studies of other reactions associated with carcinogenesis, for which the activation free energies (and the corresponding rate constants) or reaction mechanisms have not yet been experimentally determined. Finally, we predicted insignificant stereoselectivity of the studied reaction.

It is well established that the N7 atom of Gua is the major site of alkylation by ultimate carcinogens of the epoxy type.[37−39] A possible explanation lies in the fact that this nitrogen atom is not involved in the hydrogen bonding pattern with cytosine and no conformational changes of DNA during the alkylation are therefore needed. To verify this assumption we intend to calculate activation free energies for reactions of CEO with other nucleophilic sites on DNA and elaborate our findings in future studies.

Carcinogenesis is a complex pathological process, in which normal cells become neoplastic. In many of the experimental cases this process is associated with chemical modifications of DNA involving a series of chemical reactions. It is therefore a major challenge to understand and model these processes.[40−44] However, the computer modeling of reactions between DNA

(30) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(31) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210.
(32) Vayner, G.; Houk, K. N.; Jorgensen, W. L.; Brauman, J. I. *J. Am. Chem. Soc.* **2004**, *126*, 9054.
(33) Carloni, P.; Alber, F.; Eds. *Quantum Medicinal Chemistry*; J. Wiley: Weinheim, Germany, 2003.
(34) Spiegel, K.; Rothlisberger, U.; Carloni, P. *J. Phys. Chem. B* **2004**, *108*, 2699.
(35) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 6624.
(36) Pu, J.; Ma, S.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2005**, *109*, 8551.
(37) Kolman, A.; Chovanec, M.; Osterman Golkar, S. *Mutat. Res./Rev. Mutat. Res.* **2002**, *512*, 173.
(38) Their, R.; Bolt, H. M. *Crit. Rev. Toxicol.* **2000**, *30*, 595.
(39) Farmer, P. B.; Shuker, D. E. G. *Mutat. Res./Fundam. Mol. Mech. Mutagen.* **1999**, *424*, 275.
(40) Politzer, P.; Trefonas, P., III; Politzer, I. R.; Elfman, B. *Ann. N.Y. Acad. Sci.* **1981**, *367*, 478.
(41) Politzer, P.; Bar-Adon, R.; Zilles, B. A. *IARC Sci. Publ.* **1986**, *70*, 37.
(42) Scribner, J. D. *Natl. Cancer Inst. Monogr.* **1981**, *58*, 173.
(43) Ford, G. P.; Scribner, J. D. *Chem. Res. Toxicol.* **1990**, *3*, 219.
(44) Sayer, J. M.; Chadha, A.; Agarwal, S. K.; Yeh, H. J. C.; Yagi, H.; Jerina, D. M. *J. Org. Chem.* **1991**, *56*, 20.

and ultimate carcinogens has been overlooked. It is only now that computational studies of this highly relevant class of reactions are finding their way into scientific literature.[45−47] We believe that such calculations can contribute to our understanding, prevention, and treatment of cancer.

(45) Kranjc, A.; Mavri, J. *J. Phys. Chem. A*, published online Apr 7, http://dx.doi.org/10.1021/jp055092z.

(46) Huetz, Ph.; Kamarulzaman, E. E.; Wahab, H. A.; Mavri, J. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 310.

(47) Huetz, Ph.; Mavaddat, N.; Mavri, J. *J. Chem. Inf. Model*. **2005**, *45*, 1564.

Acknowledgment. The authors would like to thank Janez Cerkovnik and Ana Bergant from University of Ljubljana and Jan Florian from Loyola University Chicago for many helpful discussions. We are grateful to Jernej Stare from Los Alamos National Laboratory for performing the coupled cluster calculations. Financial support from the Slovenian Ministry of Science and Education through grant P1-0012 is gratefully acknowledged, along with United States Public Health Service grants R01 ES10546 and P30 ES00267.

JO060098L