Activation barriers for methylation of DNA bases by dimethyl sulfate

Daniel R. Eichler, George A. Papadantonakis *

Department of Chemistry, University of Illinois at Chicago, 845 W. Taylor St., Room 4500 SES, Chicago, IL 60607, United States

1. Introduction

Dimethyl sulfate (DMS) can be found in the environment in its gas phase and in airborne particulate matter that primarily originates from coal combustion [1]. The proposed use of methanol as an alternative fuel could increase exposure to significant levels of DMS with consequences affecting quality of life [2]. Hydrolysis of DMS and the subsequent methylation of DNA in cells is responsible for systemic toxic effects principally on the nervous system, heart, liver, kidneys and possible carcinogenicity [3].

DMS is a monofunctional strong alkylating agent and reacts rapidly with DNA at room temperature and penetrates intact cells [4,5]. DMS methylates predominately nitrogen sites, such as the N7 of guanine at the major groove and N3 of adenine at the minor groove in double-stranded DNA. In single-stranded DNA, DMS methylates the N7 site of guanine; the N1, N3 and N7 sites of adenine; and the N3 site of cytosine [4,5]. DMS both in vivo and in vitro can be used to identify the guanines that interact with protein in the major groove and adenes that interact with protein in the minor groove [7]. DMS can methylate the O6 site of guanine and cause direct mispairing. This is relevant because the O6 site of guanine is significant in mutagenesis [8]. Also, cases of lung cancer, bronchial carcinoma and choroidal melanoma were reported in men exposed occupationally to DMS [9,10]. Lastly, the EPA has classified DMS as a Group B2, probable human carcinogen [11].

In addition, the impact of the O6-methylation of guanosine on the relaxation mechanism of guanine monomers was reported in an experimental femtosecond laser study where the absorption spectrum is red-shifted and this resulted in a 40-fold decrease of the excited-state decay which increases the probability of radiation induced damage of cellular DNA [12].

In light of these reports, further investigation of the methylation of DNA bases by DMS becomes imperative. The methylation of DNA bases by DMS occurs through a bimolecular S_n2 nucleophilic substitution mechanism [13,14] and the experimental selectivity [4] observed in different sites is a strong indication that activation barriers exist. In theoretical investigations of DNA alkylation reactions, many studies have focused on the reactivity of methane diazonium ion with guanine [15,16]. These studies have shown that activation barriers are influenced by local ionization energy and steric interference. It was also reported that activation energies are significantly higher at the O6 than at the N7 sites and are slightly higher in solution than in the gas phase [16].

In a study investigating the steric retardation of S_n2 reactions in the gas phase and solution considering the reaction of chloride with ethyl and neopentyl chlorides and their α-cyano derivatives employing B3LYP, CBS-QB3 and PDDG/PM3 levels of theory for the gas phase and QM/MM Monte Carlo simulations for different solvents such as DMSO, methanol and water, as well as with the polarizable continuum model, CPCM, it was reported that steric effects raise the activation energy by 6 kcal/mol for the tert-butyl group relative to methyl for both cases and solvents cause a large increase in the activation energies for these reactions [17]. In an investigation of the S_n2 reaction between chloride and chloramine in dimethyl ether solution employing quantum mechanical calculations at the 6-311+G^*/MP2 and 6-311+G (2d, p)/MP2 level for the gas phase and classical force field to describe the solvent, it was described that the solvent made the ion-dipole complex well more shallow by 6.4 kcal/mol and raised the barrier for the reaction in solution to 15.0 kcal/mol [18]. In another study employing the VENUS chemical dynamics program interfaced with the NWChem electronic structure program, it was conveyed that the
dynamics of the monohydrated $\text{S}_2\text{O}_2$ reaction between fluoride ion and methyl iodide is quite complex and that the influence of solvent molecules varies essentially with collision activation [19]. In the same study, it was also stated that when more water molecules are involved, steric effects will raise the activation barrier.

However, in another investigation on the Menshutkin $\text{S}_2\text{O}_2$ reaction in gas phase and inside a carbon nanotube, it was observed that the activation energy and reaction endothermcity inside the nanotube was reduced significantly relative to the gas phase. This reduction was attributed to the separation of charge along the reaction coordinate [20]. In a theoretical investigation on the Menshutkin reaction of methyl chloride with ammonia, it was revealed that solvation causes the activation barrier to decrease relative to the gas phase [21]. The question of steric effects versus solvation effects on $\text{S}_2\text{O}_2$ reactions is very old and computational chemistry provides the necessary tools to dissect the problem [22]. The primary sites and the percent methylation in experiments obtained from reactions with double-stranded DNA from salmon sperm, calf thymus, salmon testes, rat liver and brain, human fibroblasts, and HeLa and V79 cells with DMS are the N7 of guanine (74%), the N3 of adenine (18%) and N3 of cytosine (~2%) [4].

The percent methylation at the O6 site of guanine is 0.2% [4] but this site is the main site of mutagensis in mammalian DNA and it is in competition with the N7 site [23]. The lack of direct experimental data for the activation barriers of the methylation reaction makes the theoretical evaluation of the activation barriers significant.

The goal of the present theoretical study is to investigate the factors (steric interference or charge separation) that influence the activation barriers for the reaction of DMS with the primary sites of methylation in double-stranded DNA in vitro; namely, guanine at the N7 and O6 sites, as well as adenine and cytosine at the N3 sites [4]. The present investigation employs the density functional theory (DFT) for gas-phase reactions and the conductor-like polarizable continuum model (C-PCM) [24] for reactions in water utilizing feedback from electrostatic potential maps.

2. Methods

All calculations were performed using the Spartan ‘16 Parallel Suite for Microsoft Windows 7, Professional 64-bit edition [25] on an Intel Xeon E3-1240 v3 processor utilizing 32 GB of RAM. Standard geometries of free DNA bases were generated using the Molecule Builder of Spartan 16 and were fully optimized employing the Density Functional Theory (DFT) using the dispersion-corrected density functional M06-2X with the 6-31+G* and the B3LYP-D3 with the 6-311+G (2df, 2p) basis sets. The zero-point-energy, (ZPE), corrections for the activation barriers are also included.

Both the 6-31+G* and the 6-311+G (2df, 2p) basis sets include polarization functions on all atoms and diffuse functions on heavy atoms and the later includes more of such functions than the former. Diffuse functions on heavy atoms are important for this project because the bases have lone pairs and polarization functions on all atoms, which affects the geometry and the energetics of the reaction profiles. Therefore, the basis sets selected constitutes an appropriate treatment for the present study.

Solvation energy calculations were performed employing the conductor-like polarizable continuum model (C-PCM) [24]. The implicit solvent is water which has a dielectric constant of 78.3. Electrostatic potential maps were computed using the embedded module in Spartan ‘16 at the M06-2X/6-31+G* level of theory and the $\text{S}_2\text{O}_2$ transition states of the methylation reactions were monitored by employing the Energy Profile module in Spartan ‘16.

3. Results and discussion

Table 1 contains theoretical results for the gas-phase methylation of selected sites of DNA bases obtained from DFT/ M06-2X/6-31+G* calculations and in water using the C-PCM solvation model. DMS methylates double-stranded DNA predominantly at the N7 site of guanine and at the N3 site of adenine and cytosine [4,6]. The methylation reaction at the O6 site of guanine is also considered.

Table 1 lists the activation energies relative to the reactants and products, the distances for forming and breaking bonds and the angles (attack angles) between the reaction center of the nucleic acid base, the leaving group, and the carbon of one of the methyl groups of DMS. The results show that the activation barriers in water are lower than that at the gas phase except for the reaction at the O6 site of guanine. The $\text{S}_2\text{O}_2$ reaction mechanism for the reaction of DMS with guanine at N7 is shown below.

Table 2 lists the activation energies and the transition-state geometries of the DMS methylation reactions employing the B3LYP-D3/6-311+G (2df, 2p) level of theory. It also lists the imaginary frequency calculations and the zero-point-energy correction of the activation energy. The imaginary frequency of the transition state and the normal mode of vibration that corresponds to the formation of N7 (Gua)-C3 (DMS) bond and breaking of C3 (DMS)-O4 (DMS) bond is (i220 cm$^{-1}$); for the formation of O6 (Gua) - C1 (DMS) bond and breaking of C1 (DMS)-O4 (DMS) bond is (i302 cm$^{-1}$); for the formation of N3 (Ade) – C1 (DMS) and breaking of C1 (DMS)-O3 (DMS) bond is (i458 cm$^{-1}$); and for the formation of N3 (Cyt) – C1 (DMS) and breaking of C1 (DMS)-O4 (DMS) bond is (i481 cm$^{-1}$). The gas-phase activation barriers obtained with the B3LYP-D3/6-311+G (2df, 2p) level of theory and with the zero-point-energy correction are lower in absolute value than the ones obtained by the M06-2X/6-31+G* level. However, the trend of the activation energies is the same for both levels of theory.

The use of the 6-31+G* basis set with the B3LYP functional provided results for the threshold photoionization energies of nucleotides similar to those of the more computationally expensive Dunning type basis sets and in agreement with experimental data [16] and the same level of theory was used in the theoretical evaluation of the activation barriers of DNA bases with methane diazonium ions [15]. Lastly, in a review paper on the assessment of the performance of DFT for several atomic and molecular properties it

<table>
<thead>
<tr>
<th>Site</th>
<th>Gas phase geometry and $E_a$</th>
<th>Geometry and $E_a$ in water</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r_{form}$ ($\text{Å}$)</td>
<td>$r_{break}$ ($\text{Å}$)</td>
</tr>
<tr>
<td>Guanine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N7</td>
<td>1.814</td>
<td>1.499</td>
</tr>
<tr>
<td>O6</td>
<td>1.918</td>
<td>2.546</td>
</tr>
<tr>
<td>Adenine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>2.023</td>
<td>1.673</td>
</tr>
<tr>
<td>Cytosine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td>2.023</td>
<td>1.699</td>
</tr>
</tbody>
</table>
was reported that the 6-31+G* basis set provided accurate results for vibrational frequencies, ionization potentials and for barrier heights for reactions with radical and singlet transition states [26]. The selection of the M02-2X/6-31+G* method for the evaluation of the activation barriers in water was based merely on grounds of computational expense; the results obtained show that the lowest activation energy (16.83 kcal/mol) is for the methylation reaction at the N7 site of guanine which is associated with 74% methylation and the highest activation energy is for the methylation at the O6 site (79.94 kcal/mol) which is associated with 0.2% methylation as it was reported [4].

The electrostatic potential maps employed in this study were computed using DFT/M06-2X/6-31+G* level of theory and they show the electrostatic potential on a surface of electron density. Colors near red represent large negative values of the potential, while colors near blue represent large positive values and other colors (orange, yellow and green) represent intermediate values of the potential. In addition, red areas show that charge is localized in that area, and green and blue areas show delocalization of charge over the entire molecule.

3.1. Activation barriers for guanine methylation by DMS at N7 and O6 sites

The initial step of the methylation reaction in the gas phase and in water at the N7 site shown in Fig. 1(a) and (d) respectively, indicates that guanine has a neutral potential (green areas) delocalized with small zones of positive potential (blue areas) around the hydrogens at the N9, C8, N1 and N2 sites. The initial state of the reaction at the N7 reaction center shows zones of negative

### Table 2

<table>
<thead>
<tr>
<th>Site</th>
<th>r_{form} (Å)</th>
<th>r_{break} (Å)</th>
<th>Attack angle (deg)</th>
<th>E_a (kcal/mol)</th>
<th>Imaginary frequency (cm⁻¹)</th>
<th>E_a (ZPE) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guanine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N7</td>
<td>1.814</td>
<td>1.499</td>
<td>83.71</td>
<td>68.93</td>
<td>220</td>
<td>68.49</td>
</tr>
<tr>
<td>O6</td>
<td>1.918</td>
<td>2.546</td>
<td>157.8</td>
<td>19.92</td>
<td>302</td>
<td>21.52</td>
</tr>
<tr>
<td>Adenine</td>
<td>2.023</td>
<td>1.673</td>
<td>174.56</td>
<td>24.72</td>
<td>458</td>
<td>24.60</td>
</tr>
<tr>
<td>Cytosine</td>
<td>2.023</td>
<td>1.699</td>
<td>174.56</td>
<td>21.59</td>
<td>481</td>
<td>21.22</td>
</tr>
</tbody>
</table>

![Guanine (N7) methylation by dimethyl sulfate](image)
potential (red areas) at the N7 and O6 sites suggesting that these are the most reactive sites. The N3 site also shows a slightly negative potential suggesting some reactivity (1%) [4]. DMS shows neutral potential at the two methyl groups and a negative potential at the sulfate core.

A closer look at the electrostatic maps that correspond to the transition state at the N7 reaction center in gas phase, Fig. 1(b), and in water, Fig. 1(e), clearly indicates the beginning of charge separation buildup between the methylated base and the leaving group. Previously neutral potential on guanine has become more positive and the potential at the O6 and N3 sites has also become more positive. Furthermore, the leaving group shows negative potential around the sulfate and is stabilized by resonance. Data listed in Table 1 reveals that the length of the bond formed between the methyl group and the N7 atom is longer in water, 2.127 Å compared to 1.814 Å in the gas phase; and the bond broken is shorter in the gas phase, 1.499 Å compared to 1.604 Å in water. The angle of attack between the sulfate, the methyl group, and guanine in the gas phase is very unfavorable, 83.71°, and it becomes more linear in water, 177.99°. The linearity of the attack angle is a main characteristic of SN2 reactions.

At the end state as shown in Fig. 1(c), the reaction geometry in the gas phase indicates that the methyl group has formed a bond at the N7 position, and the leaving group has successfully left, and is stabilized by resonance. 7-Methylguanine has large positive potential delocalized over the entire molecule at all sites except for the N3 and O6 sites and the positive potential is more predominant in water Fig. 1(f). At these sites, slightly negative values of the potential are observed and they suggest further reactivity. The leaving group, MeSO4, shows large negative potential and it is further stabilized by resonance. The end state shows large charge separation along the reaction coordinate which is significantly greater than that at the transition state in both gas-phase and in water. The energy of the activation barrier in solution is reduced from 71.51 kcal/mol to 16.83 kcal/mol. This is a clear indication that solvation enhances charge separation along the reaction coordinate and makes the N7 site of guanine the most reactive site among all of the nucleic acid bases with methylation rate of 74% [4,15,16,27].
In guanine, the reactivity at the O6 site is in competition with the N7 site and it is accepted that the O6 site is the main site of mutagenesis in mammalian DNA [8,12]. Fig. 2(a) and (d) shows the geometry in the gas phase and in water respectively of the initial state of the methylation reaction in which a hydrogen bond is formed between the N1 and N2 hydrogen atoms of guanine and the O2 oxygen atom of the sulfate group. Initially, there were two hydrogen bonds between the N1 and N2 hydrogen atoms of guanine and the sulfate O2 oxygen and this is still observed in the gas phase transition state Fig. 2(b), but at the transition state in water, Fig. 2(e), there is only one hydrogen bond interaction at the N1 hydrogen atom of guanine. This dual interaction occupies the oxygen orbitals of the sulfate leaving group and results in less charge separation seen in the gas phase.

At the end state in the gas phase, as shown in Fig. 2(c), the methyl group has formed a bond at the O6 site, and the leaving group has successfully left and is stabilized by resonance. O6-Methylguanine exhibits positive potential delocalized over the entire molecule except for the N3 and N7 sites where there is a slightly negative potential. The latter suggests further reactivity at these sites. The leaving group shows a large negative potential and is stabilized via resonance, but there is still a hydrogen interaction at the N2 hydrogen of guanine with the O2 of the sulfate, as shown in Fig. 2(c). This suggests that, MeSO$_4^-$ is a poor leaving group in the gas phase. At the end state, charge separation occurs but not adequately enough to prevail over the steric effects that dominate the reaction at the O6 site of guanine. The angle of

**Panel A**

(a) Initial Step (gas)  
(b) Transition State (gas)  
(c) N3 Me-Adenine (gas)

**Panel B**

(d) Initial Step in Water  
(e) Transition State in Water  
(f) N3 Me-Adenine in Water

**Fig. 3.** Panel A shows the electrostatic potential maps, clipped in the plane, generated by using the DFT/M06-2X functional with the 6-31+G* basis set for the reaction of dimethyl sulfate with adenine at the N3 site in the gas phase; initially (a), the transition state (b) and the end state (c). Panel B shows the electrostatic maps for the reaction in water; initially (d), the transition state (e) and the end state (f). Hydrogen bonding occurs between the O1 atom on the sulfate and the hydrogen atom on the N9 atom of adenine in (a), (b), (d), (e) and (f).
attack between the sulfate, the methyl group, and guanine in the gas phase is an unfavorable 157.80° and deviates further from linearity in water to 85.74°. The activation barrier energy is increased from 25.04 kcal/mol in the gas phase to 79.94 kcal/mol in solution because of steric interference effects.

3.2. Activation barriers for adenine methylation by DMS at the N3 site

The initial geometry of the methylation reaction at the N3 site of adenine in both gas phase and in water are shown in Fig. 3 (a) and (d) respectively. The N7 and N1 sites have small localized negative potential and therefore exhibit reactivity. The transition state of the reaction in the gas phase, Fig. 3(b), shows some charge separation between the leaving group and adenine. There is a negative potential localized around the sulfate on the leaving group, and on adenine, the slightly negative potentials around N7 and N3 have become more positive. There is a strong hydrogen bond present between the hydrogen at the N9 of adenine and the oxygen O1 of the sulfate and this hydrogen bond creates a worse angle of attack in the gas phase, 174.56°, than what is observed in water, 176.34°. The transition state in solution, Fig. 3(e), has a neutral potential delocalized around the adenine molecule; however, the negative potential localized around the N1 and N7 atoms becomes less negative. Furthermore, there is still a large negative potential around the sulfate leaving group. In solution, the angle of attack becomes 176.34°, and this indicates that solvent effects weaken the hydrogen bond interactions and the attack angle comes closer to linearity. The hydrogen bond length between the hydrogen of N9 atom of adenine and the O1 atom of sulfate is 1.908 Å in the gas phase and it weakens in water, lengthening to 2.068 Å. At the end of the reaction in both gas phase, Fig. 3(c) and in water, Fig. 3(f), there is a large positive potential located around the adenine and a large negative potential around the methyl sulfate leaving group. Finally at the end state of the N3 reaction in both gas-phase and in water, large charge separation is observable and the activation barrier decreases from 24.53 kcal/mol in the gas-phase to 17.59 kcal/mol in water.

3.3. Activation barriers for cytosine methylation by DMS at the N3 site

Fig. 4(a) shows that cytosine initially has a neutral potential in both gas phase and in water, Fig. 4(d), and is delocalized over the molecule. There are a few areas of positive potential localized around the perimeter hydrogens and some negative potential localized around the O2 and N3 active sites. DMS has neutral potential delocalized over the entire molecule and some localized negative potential around the oxygens of the sulfate. At the transition state for the reaction at the N3 site in the gas phase, Fig. 4(b),
there is a delocalized neutral potential around the molecule but the initial large negative potential around the O2 site becomes less negative. There is a hydrogen bond formed at the N4 hydrogen with the O1 atom of sulfate and it creates a less favorable SN2 attack angle, 174.54°, than what is observed at the transition state attack angle in water, 175.82°, in Fig. 4(e), and the hydrogen bond is disrupted. The activation barrier lowers from 23.07 kcal/mol in the gas phase to 17.92 kcal/mol in solution. Solvation also increases the length of the bond formed in the transition state from 2.023 Å in the gas phase, to 2.127 Å in water, and it decreases the breaking bond distance from 1.699 Å to 1.600 Å; also, the leaving group is stabilized upon solvation by resonance. At the end state of the reaction in the gas phase, as shown in Fig. 4(c), a large charge separation buildup along the reaction coordinate is significantly greater than the buildup at the transition state. Similar result is obtained for the reaction in water.

4. Conclusions

1. Gas-phase activation barriers with zero-point-correction obtained by DFT using the B3LYP-D3 functional with the 6-311+G (2df, 2p) basis set are lower in magnitude than the values obtained by using the M06-2X functional with the 6-31+G level. However, both sets of results follow the same trend in energetics.

2. The methylation at the N7 site of guanine has the lowest activation barrier in water, 16.83 kcal/mol, and, therefore it is the most reactive site to be methylated, in agreement with the percentage methylation reported from experimental data.

3. Charge separation between the nucleic acid base and the leaving group results in reduced hydrogen bond interactions and lowers the activation barriers in water for all nitrogen sites of the bases in study except for the O6 site of guanine which is influenced by steric interference.

Acknowledgments

Daniel R. Eichler is grateful for the financial support provided by the University of Illinois at Chicago Liberal Arts and Science Undergraduate Research Initiative (LASURI) and the Department of Chemistry. GAP acknowledges stimulating discussions with Professor Donald J. Wink of the UIC Chemistry Department.

References