

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 6471–6483

www.rsc.org/pccp

PAPER

Quantum mechanics study and Monte Carlo simulation on the hydrolytic deamination of 5-methylcytosine glycol[†]

Ze Qin Chen,^{a,b} Cheng Hua Zhang,^a Chan Kyung Kim^c and Ying Xue*^a

Received 5th December 2010, Accepted 27th January 2011

DOI: 10.1039/c0cp02783a

The efficient formation of 5-methylcytosine glycol (mCg) and its facile deamination to thymine glycol (Tg) may account for the prevalent C → T transition mutation found at methylated CpG site (mCpG) in human p53 gene, a hallmark for many types of human tumors. In this work, the hydrolytic deamination of mCg was investigated at the MP2 and B3LYP levels of theory using the 6-311G(d,p) basis set. In the gas phase, three pathways were explored, paths A–C, and it indicates that the direct deamination of mCg with H₂O by either pathway is unlikely because of the high activation free energies involved in the rate-determining steps, the formation of the tetrahedral intermediate for paths A and B as well as the formation of the Tg tautomer for path C. In aqueous solution, the role of the water molecules in the deamination of mCg with H₂O was analyzed in two separate parts: the direct participation of one water molecule in the reaction pathway, called the water-assisted mechanism; and the complementary participation of the aqueous solvation. The water-assisted mechanism was carried out for mCg and the cluster of two water molecules by quantum mechanical calculations in the gas phase. This indicates that the presence of the auxiliary water molecule significantly contributes to decreasing all the activation free energies. The bulk solution effect on the water-assisted mechanism was included by free energy perturbation implemented on Monte Carlo simulations, which is found to be substantial and decisive in the deamination mechanism of mCg. In this case, the water-assisted path A is the most plausible mechanism reported for the deamination of mCg, where the calculated activation free energy (22.6 kcal mol⁻¹ at B3LYP level of theory) agrees well with the experimentally determined activation free energy (24.8 kcal mol⁻¹). The main striking results of the present DFT computational study which is in agreement with previous experimental data is the higher rate of deamination displayed by mCg residues with respect to 5-methylcytosine (mC) bases, which supports that the deamination of mCg contributes significantly to the C → T transition mutation at mCpG dinucleotide site.

1. Introduction

Mutation in the p53 tumor suppressor gene is a hallmark of human cancers. Six major mutational hotspots in p53, *i.e.* codons 175, 213, 245, 248, 273 and 282, contain CpG sites.¹ In mammalian cells, the cytosine (C) residues at CpGs are methylated at the C5 carbon to form 5-methylcytosine (mC),

which accounts for ~4% of all dC residues in humans.^{2,3} Although CpGs are under-represented by 5-fold of their expected frequency in mammalian DNA,⁴ methylated CpGs (mCpGs) are the sites for more than 30 percent of all known disease related point mutations.^{5–7} The most common mutation at CpGs observed is C → T transition mutation.

Numerous hypotheses have been proposed to account for the ubiquitously observed C → T transition mutation at CpGs, most of which emphasize the importance of methylation of C residues. Methylation increases the rate of hydrolytic deamination and also increases the reactivity of neighboring guanines to electrophiles.^{1,8–10} Detailed studies on hydrolytic deamination of mC and C, either as free nucleotides or when inserted into DNA fragments, have been performed at neutral physiological pHs. It has been indicated that, as a general trend, the deamination rate of mC is about five-fold higher than that of C in both nucleoside 5'-monophosphates¹¹ and single-stranded DNA^{12–15} at 37 °C. The rate constant for the

^aCollege of Chemistry, Key Laboratory of Green Chemistry and Technology in Ministry of Education, Sichuan University, Chengdu 610064, People's Republic of China.
E-mail: yxue@scu.edu.cn; Tel: 86-28-85418330

^bLab of Applied Chemistry and Pollution Control Technology, College of Chemistry and Chemical Engineering, China West Normal University, Nanchong 637002, People's Republic of China

^cDepartment of Chemistry, Inha University, Incheon 402-751, Korea
† Electronic supplementary information (ESI) available: Listings of optimized Cartesian coordinates and energies of all stationary points along the potential energy profile in hartree at the B3LYP/6-311G(d,p) and MP2/6-311G(d,p) levels of theory. See DOI: 10.1039/c0cp02783a

deamination of mC in duplex DNA at 37 °C is $5.8 \times 10^{-13} \text{ s}^{-1}$, which is twice as fast as the corresponding deamination of C.⁹ However, it should be noted that these rate constants are at least 3 orders of magnitude lower than those estimated for mC and C in nucleotides and single-stranded DNA, manifesting a major protecting effect of stacking and base-pairing on hydrolytic deamination of mC and C in double-stranded DNA. These rates, albeit low, are sufficient to explain the observed frequency of mutation at sites containing mC and emphasize the importance of hydrolytic deamination as one of the major sources of human mutations.¹⁶

Since the deamination of mC proceeds at a low rate and since there are multiple repair systems that operate on T/G mismatches derived from deaminated mC,^{17,18} it has been questioned if the deamination of mC is the only or even the prevailing reaction that leads to CpG transition. Recently, another mechanism accounting for the common C → T transition mutation at CpGs has been confirmed by Cao *et al.*,¹⁹ which involves the formation of 5-methyl-5,6-dihydroxy-5,6-dihydrocytosine (or 5-methylcytosine glycol, mCg; Scheme 1) and its subsequent deamination to 5,6-dihydroxy-5,6-dihydrothymine (or thymine glycol, Tg; Scheme 1). They found out that the deamination of mCg is about 7 orders of magnitude faster than the spontaneous deamination of mC in duplex DNA. Consequently, the high susceptibility to oxidation and the significant deamination rate of oxidized mC may contribute to the high mutagenic rate observed at mCpG doublets in genomic DNA.

Reactive oxygen species (ROS) generated by either aerobic metabolism or exogenous oxidizing agents constitute an important endogenous source of DNA damage, which is associated with aging and a variety of human diseases, such as cancer and neurodegeneration.^{20,21} The main species responsible for the ROS-induced DNA damage appears to be the hydroxyl radical generated by a Fenton-type reaction involving the reduction of H₂O₂ by transition metal ions (*e.g.* Fe²⁺ and Cu⁺).²² The most likely pathways of mCg formation involve the addition of hydroxyl radical to C6 atom of mC giving rise

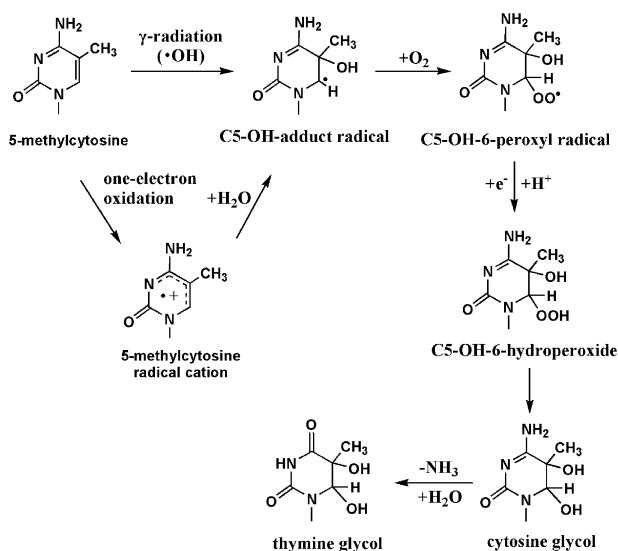
to transient C6-OH-adduct radical of mC that then reacts with oxygen at a diffusion-controlled rate (shown in Scheme 1).²³ The resulting pyrimidine peroxy radical is expected to be reduced and then protonated to yield the related hydroperoxide, which is likely to be converted by reduction of the peroxidic bond into mCg. Another possible mechanism is one-electron oxidation of mC that leads to the generation of a related radical cation, which predominantly yields C5-OH-adduct radical of mC through predominant hydration and then the other process are the same as discussed above.²⁴ Tg, as the deamination product of mCg,²⁵ although primarily a replication blocking lesion, can be bypassed by translesion synthesis DNA polymerases including yeast DNA polymerases ζ and η as well as human DNA polymerase κ , and all three polymerases preferentially insert a dAMP opposite the lesion,^{26–28} resulting in a GC → AT transition mutation.

On the theoretical side, early studies have investigated the deamination mechanisms of C,^{29–33} mC,³⁴ and 5,6-dihydrocytosine (5-DHC).³⁵ However, it remains unexplored how methylation affects the deamination of the ring-saturated glycol derivative. To pursue more understanding of the biological role of mCg, we focus our work on the deamination of *cis*-(5S,6S)-mCg. This system provides an interesting case study because its deaminated product, *cis*-(5R,6S)-Tg, is the most stable among the four stereoisomers, no matter when it located at nucleoside, or in duplex.³⁶ The present study, to the best of our knowledge, is for the first time aimed at using the computational model of mCg to investigate its hydrolytic deamination mechanisms. For the gas phase reactions, the MP2 and B3LYP methods have been used in our calculations. In our previous work, we adopted the quantum chemical molecular orbital method and Monte Carlo (MC) simulation with the free energy perturbation (FEP) technique to study effects of solvents on some reaction systems.³⁷ In this study, the effect of solvent water is studied using the Monte Carlo free energy perturbation method.

2. Computational details

2.1 Quantum mechanics calculations

The model of *cis*-(5S,6S)-mCg used here was established using the X-ray crystallographic structure of *cis*-(5R,6S)-Tg by replacing the oxygen atom at C4 atom with amino group and deleting a hydrogen atom on the N3 atom³⁸ with subsequent full optimization. All the computations were performed with the Gaussian 03 suite of programs.³⁹ The geometries of all reactant complex (RC), transition state (TS), intermediate (IM), and product complex (PC) in the gas phase were fully optimized at MP2 and B3LYP levels of theory using the 6-311G(d,p) basis set. This combination of basis set and methods has been widely utilized to study the deamination mechanisms of the related compounds, *i.e.* C,^{33,40} mC,³⁴ and 5-DHC.³⁵ A frequency analysis was performed for each stationary point in order to ensure that all minima have no imaginary frequencies on the potential-energy surface and that transition states have a single imaginary frequency. The complete reaction pathways for all the mechanisms discussed in this paper have been verified using intrinsic reaction



Scheme 1 ROS-induced DNA Lesions of mC and the Deamination of mCg.

coordinate (IRC) analysis⁴¹ of all transition states. In this way, the geometries along the minimum energy path (MEP) were also given. For all points along MEP, atomic charges were obtained *via* the CHELPG⁴² procedure at the B3LYP/6-311G(d,p) level.

2.2 Monte Carlo simulation

To test the solvent effect on the deamination of mCg in aqueous solution through the water-assisted mechanism, the MC simulations with free energy perturbation method (MC-FEP) were performed to determine the free energies of solvation using standard procedure for Metropolis sampling technique applied for liquid system of rigid molecules in the NPT ensemble.⁴³ In the present model, a pre-equilibrated box containing 400 solvent water molecules was used for each free energy perturbation simulation at a pressure of 1 atm and a temperature of 298.15 K. The molecules were kept rigid during the simulation. For all solute molecules, the geometries were obtained in IRC calculations at the B3LYP/6-311G(d,p) level in the gas phase; the potential interaction was described by the Lennard-Jones parameters of the OPLS force field⁴⁴ in BOSS program database, which scaled linearly in going from their reactant to product values, and the CHELPG charges were calculated for each system at the B3LYP/6-311G(d,p) level in the gas phase and were assigned to all atoms in the solute as partial charges. Perturbations were carried out with use of double-wide sampling scheme.⁴⁵ The geometry parameters determining volume changes and solute rotations and translations were adjusted to have an acceptance level *ca.* 40%. The variation of the Gibbs free energy between two adjacent states *i* and *j* in solution is calculated *via* Zwanzig perturbation expression⁴⁶ as follows:

$$\Delta G_{ij} = G_j - G_i = -kT \ln\{\exp[-(E_j - E_i)/kT]\} \quad (1)$$

where *k* is the Boltzman constant, *E_i* is the total energy of the point *i*, and *T* is the absolute temperature. Each simulation involved 10^6 configurations of equilibration and 2×10^6 configurations of averaging. All of the simulations were executed using the BOSS program.⁴⁷

The intermolecular interaction is described by a potential energy function consisting of Coulomb and Lennard-Jones (LJ) terms between the atom *i* in the molecule *a* and atom *j* in the molecule *b*, which are separated by a distance *r_{ij}*

$$\Delta E_{ab} = \sum \sum (q_i q_j e^2 / r_{ij} + 4 \epsilon_{ij} [(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6]} \quad (2)$$

where the crossing terms were obtained using the geometric combining rules:

$$\epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2}, \sigma_{ij} = (\sigma_i \sigma_j)^{1/2} \quad (3)$$

In this study, the four-site TIP4P model^{48–50} was used for the solvent water. 395 water molecules in an approximate $20 \times 20 \times 30 \text{ \AA}^3$ box were used for the water-assisted deamination of mCg.

3. Results and discussion

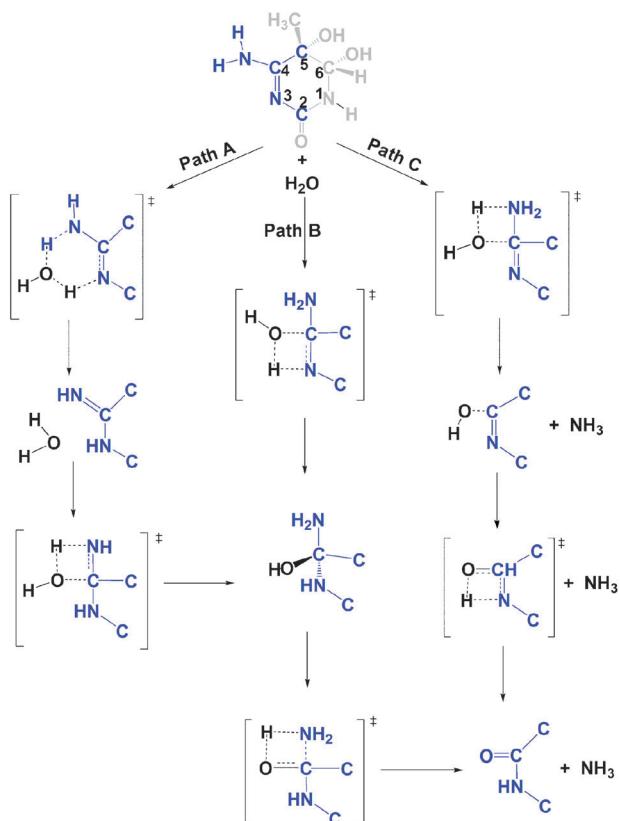
First, it should be pointed out that, the geometries on the PES of the title reaction located at the B3LYP/6-311G(d,p) are very close to those located at the MP2/6-311G(d,p) level, whose

corresponding structure parameters are presented in Fig. s1–s3 in the Supporting Information.[†] Therefore, our geometry discussions here mainly focus on the B3LYP/6-311G(d,p) level of theory. Based upon pathways proposed in the literature for the deamination of unmodified and modified cytosine,^{29–35} three possible reaction pathways for the hydrolytic deamination of mCg were taken into account in our work, denoted as paths A, B, and C (shown in Scheme 2). As a matter of fact, paths A and B, proposed in the deamination of C in early studies,^{29–31} both involve the formation of a tetrahedral intermediate, then the intermediate adjusts its geometry by rotating some groups to offer a favorable structure to proceed the final intramolecular 1–3 proton shift yielding the deamination product, the Tg–ammonia complex. This is the fist time that the two mechanisms were considered together. For path C, two processes are concerned: the initial cleavage of C4–N4 bond to form the hydroxyl-oxo tautomer of Tg and ammonia and the subsequent 1–3 proton shift of the hydroxyl hydrogen to the sp² nitrogen producing the Tg–ammonia complex. The changes in Gibbs free energies (ΔG) and electronic energies for the direct deamination mechanism of paths A, B, and C are presented in Table 1.

3.1 Stationary point structures and energetics in the gas phase

3.1.1 Direct deamination mechanisms

Path A. The optimized structures of reactant complex, intermediates, transition states, and product complex for path A in the gas phase are depicted in Fig. 1. Path A is a five-step mechanism. Starting from the hydrogen-bonded reactant



Scheme 2 Proposed deamination mechanisms of mCg with H₂O.

Table 1 Changes of Gibbs free energies (ΔG) and electronic energies (ΔE) for the direct deamination mechanisms of mCg in the gas phase (in kcal mol⁻¹)

	B3LYP/6-311G(d,p)		MP2/6-311G(d,p)	
	$\Delta G(\text{gas})$	$\Delta E(\text{gas})$	$\Delta G(\text{gas})$	$\Delta E(\text{gas})$
Path A				
RE + H ₂ O → A-RC	-3.1	-12.0	-3.0	-11.8
A-RC → A-TS1	13.3	12.3	13.6	12.5
A-RC → A-IM1	1.6	2.6	-0.4	0.2
A-IM1 → A-TS2	47.3	45.4	52.5	50.9
A-IM1 → A-IM2	8.2	5.9	4.3	2.2
A-RC → A-IM2	9.8	8.5	3.9	2.5
A-IM2 → A-TS3	0.4	0.1	0.5	0.2
A-IM2 → A-IM3	-1.6	-1.8	-1.7	-1.9
A-RC → A-IM3	8.2	6.7	2.2	0.6
A-IM3 → A-TS4	0.9	0.7	1.0	0.9
A-IM3 → A-IM4	-2.4	-2.4	-2.4	-2.3
A-RC → A-IM4	5.8	4.3	-0.1	-1.7
A-IM4 → A-TS5	27.3	27.6	28.9	29.3
A-IM4 → A-PC	-23.7	-21.8	-16.4	-14.4
A-RC → A-PC	-17.9	-17.5	-16.5	-16.2
Path B				
RE + H ₂ O → B-RC	-0.5	-8.3	-0.7	-9.1
B-RC → B-TS1	45.4	43.5	47.5	45.9
B-RC → B-IM1	-0.2	-2.9	-5.8	-7.9
B-IM1 → B-TS2	8.5	8.7	9.3	9.5
B-IM1 → B-IM2	4.7	5.1	5.6	5.8
B-RC → B-IM2	4.5	2.2	-0.2	-2.0
B-IM2 → B-TS3	30.7	30.7	32.3	32.7
B-IM2 → B-PC	-21.4	-19.2	-15.1	-13.0
B-RC → B-PC	-16.8	-17.0	-15.3	-15.0
Path C				
RE + H ₂ O → C-RC	3.0	-5.1	1.5	-6.1
C-RC → C-TS1	50.5	48.7	50.1	48.1
C-RC → C-IM1	-2.7	-2.8	-0.6	-1.5
C-IM1 → C-TS2	24.9	24.4	24.8	24.7
C-IM1 → C-PC	-22.1	-22.4	-21.0	-21.0
C-RC → C-PC	-24.8	-25.3	-21.7	-22.5

complex (A-RC), the hydrogen atom H_{w1} involved in the strong hydrogen bond with N3 atom is transferred to the nitrogen atom *via* the transition state A-TS1 with the corresponding activation free energies of 13.3 and 13.6 kcal mol⁻¹ at the B3LYP/6-311G(d,p) and MP2/6-311G(d,p) levels, respectively (hereafter named B3LYP and MP2 levels, respectively). In A-TS1, a six-membered ring is formed by carbon C4, nitrogen N4, hydrogen H1, oxygen O_{w1}, hydrogen H_{w1}, and nitrogen N3. Two proton transfers occur synchronously, namely, H1 atom from N4 to O_{w1} atom and H_{w1} from O_{w1} to N3 atom. The intermediate A-IM1 is hence best described as a hydrogen-bonded complex between the imine-oxo tautomer of mCg and the water molecule.

Following the hydrogen-atom transfer, a nucleophilic addition of the hydroxyl group occurs on C4 atom with simultaneous proton transfer from H₂O to the exocyclic imine nitrogen (sp² nitrogen) *via* a four-membered transition state A-TS2. This step is associated with the activation free energies of 47.3 and 52.5 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively, and leads to the formation of the intermediate A-IM2. In A-IM2, the C4 atom has a tetrahedral structure, and the distances of C4–N4, C4–N3, and C4–C5 are stretched to 1.433, 1.456, and 1.564 Å, respectively, longer than those in the A-IM1 (1.272 Å for C4–N4, 1.383 Å for C4–N3, and 1.532 Å for C4–C5), suggesting that the hybridization of C4 atom

converts from sp² in the A-IM1 to sp³ and the C4=N4 double bond turns into a single bond.

To achieve deamination, the N4 amine group must dissociate through the transfer of H_{w2} atom from the hydroxyl group on C4 atom, which makes it essential to adjust the hydroxyl hydrogen toward the N4 atom. For this to occur (starting from A-IM2), two low barrier conformational changes connect intermediates A-IM2 and A-IM4. These intermediate conformers (A-IM2, A-IM3, A-IM4) are very similar, differing mainly in the torsion and the angles of the two functional groups (–NH₂ and –OH). The activation free energies of A-TS3 and A-TS4 are 0.4 and 0.9 kcal mol⁻¹, respectively, at the B3LYP level, which correspond to 0.5 and 1.0 kcal mol⁻¹, respectively, at the MP2 level. These activation free energies for these conformational changes are small as expected and do not have significant effects on the mechanism of this reaction.

In the subsequent step along path A, the N4 amine group dissociates from the molecule to yield the product complex (A-PC). The activation free energy associated with the step is 27.3 kcal mol⁻¹ at the B3LYP level and 28.9 kcal mol⁻¹ at the MP2 level. The transition state A-TS5 is a four-membered structure, which mainly corresponds to the transfer of H_{w2} from the hydroxyl group to the amino group. In A-TS5, the distance of C4–N4 is 1.586 Å, 0.131 Å longer than that in A-IM4 (1.455 Å), indicating that the C4–N4 bond has been partially broken in A-TS5. In the product complex A-PC, the two fragments, Tg and NH₃, are interacting through two hydrogen bonds.

Path B. The optimized structures of reactant complex, intermediates, transition states, and product complex for path B are shown in Fig. 2. In path B, an incoming H₂O at first forms a reactant complex B-RC with mCg through the electrostatic interaction and then adds to the N3=C4 double bond to form the tetrahedral intermediate B-IM1 *via* the transition state B-TS1. B-TS1 is a four-membered ring structure, which is mainly associated with the transfer of H_{w1} atom from O_{w1} to N3 atom and the addition of O_{w1} to C4 atom. Similar to A-IM2, the intermediate B-IM1 also has tetrahedral structure and the hybridization of C4 atom changes from sp² in B-RC into sp³. The activation free energies concerned in this step are 45.4 and 47.5 kcal mol⁻¹ at the B3LYP and MP2 levels of theory, respectively, in the gas phase.

For path B, the remainder processes are similar to those of path A but only one conformational change exists in the potential energy surface. The activation free energies for this conformational process are 8.5 and 9.3 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively. Starting from B-IM2, it is apparent that N4 atom must become protonated to facilitate the breakage of C4–N4 bond. The intramolecular 1–3 proton transfer from hydroxyl group to the amino group is associated with the transition state B-TS3 and with the activation free energy of 30.7 kcal mol⁻¹ at the B3LYP level and 32.3 kcal mol⁻¹ at the MP2 levels in the gas phase.

Path C. The optimized structures of the reactant complex, intermediates, transition states, and product complex for path C are shown in Fig. 3. Path C is a two-step process mechanism. Initially, the water molecule attacks the C4 atom with

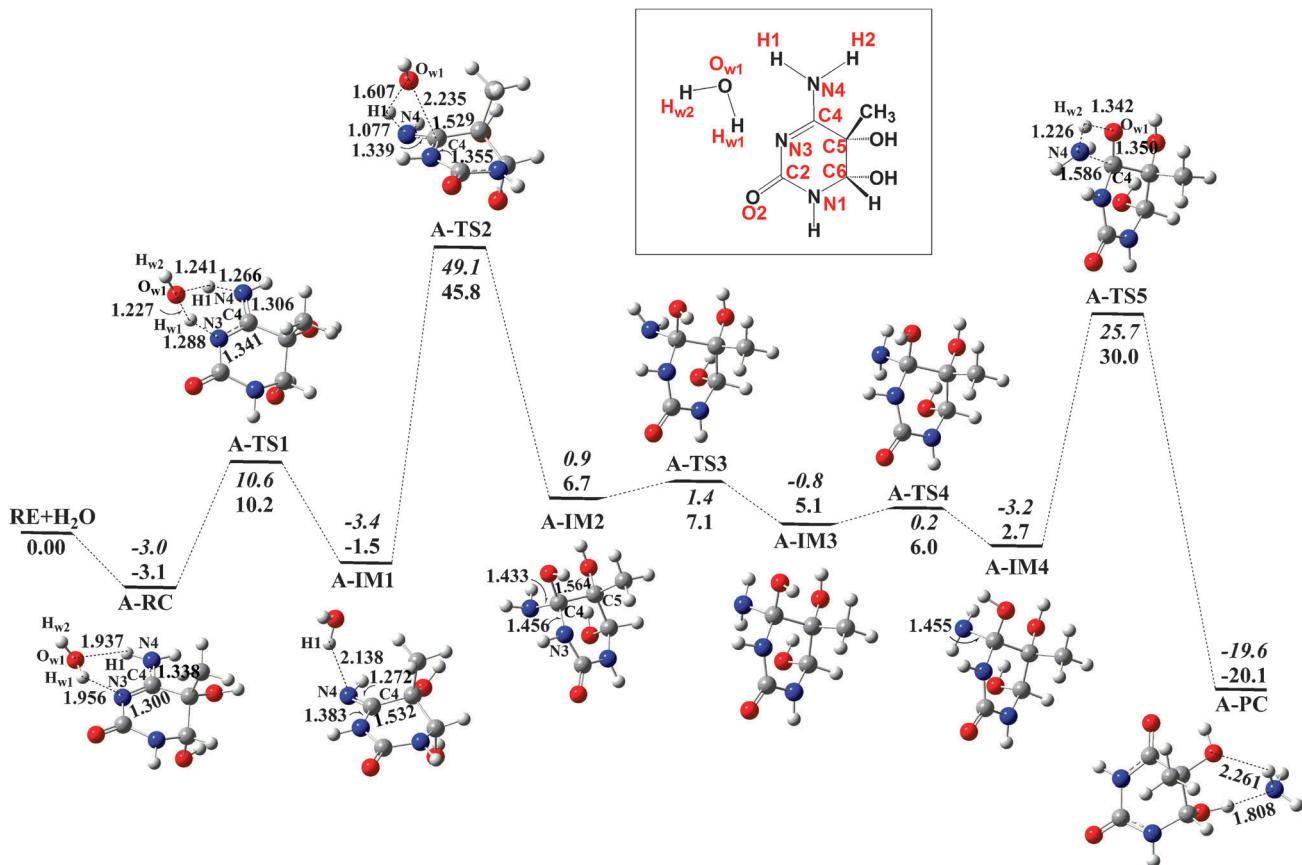


Fig. 1 Potential energy surface (in kcal mol⁻¹) and optimized structures (in Å) for the direct deamination mechanism of *cis*-(5*S*,6*S*)-mCg (path A) in the gas phase (top values in italic, ΔG at the MP2/6-311G(d,p) level; bottom values, ΔG at the B3LYP/6-311G(d,p) level. Definitions are the same in the following Figures). Inset: atomic labeling used in the current study.

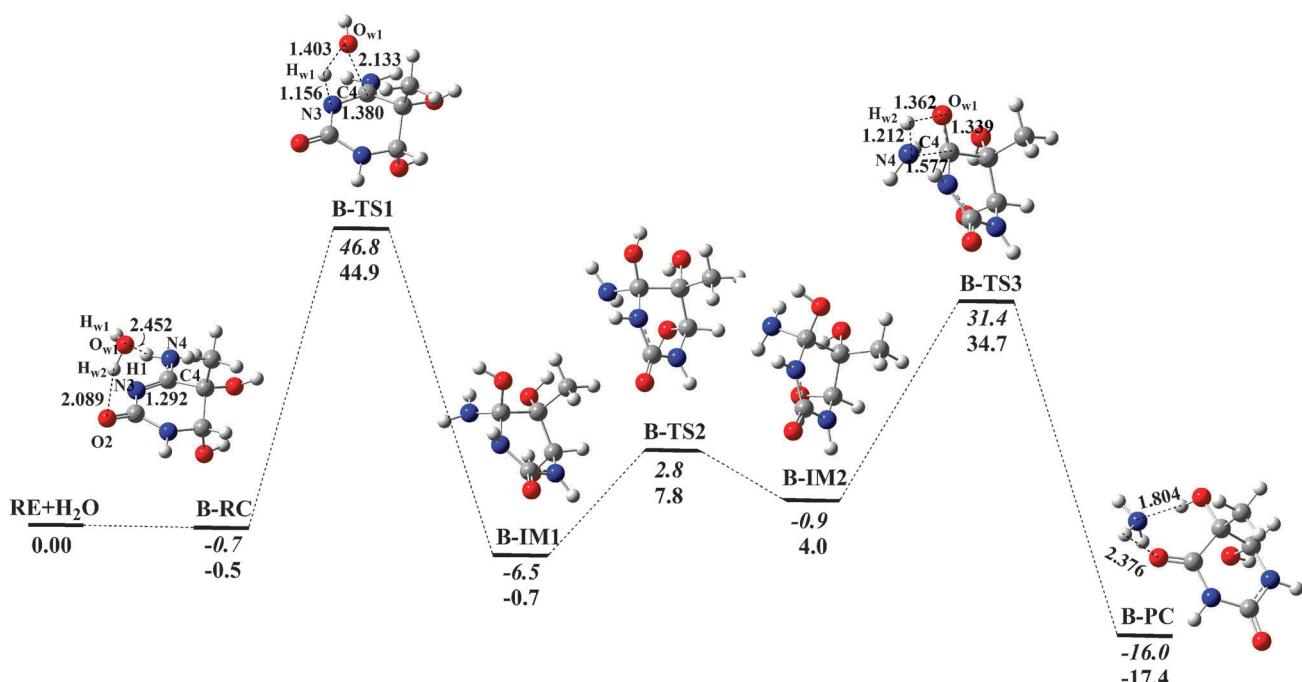


Fig. 2 Potential energy surface (in kcal mol⁻¹) and optimized structures (in Å) for the direct deamination mechanism of *cis*-(5*S*,6*S*)-mCg (path B) in the gas phase.

simultaneous proton transfer from H₂O to the amino group (sp³ nitrogen) of mCg producing the intermediate C-IM1, a hydroxyl-oxo tautomer of thymine glycol and ammonia. This step occurs via a four-membered transition state C-TS1 with the activation free energies of 53.5 and 51.6 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively. In C-TS1, the distance of C4–N4 is 1.510 Å, longer than that of C-RE (1.351 Å), indicating that the C4–N4 bond has been partially cleaved in C-TS1. After surmounting transition state C-TS1, the intermediate C-IM is formed, in which, the released NH₃ keeps bonded with the hydroxyl-oxo tautomer of thymine glycol through a strong hydrogen bond.

Following step 1, thymine glycol is produced from C-IM1 by the migration of the hydroxyl hydrogen to the N3 atom. The transition state B-TS2 is a four-membered ring structure, in which the transferred hydrogen atom, H_{w2}, is almost equidistant between N3 (1.356 Å) and O_{w1} (1.306 Å) atoms. The activation free energies associated with this unique step are 24.9 and 24.8 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively.

The potential energy surfaces for paths A, B, and C are shown in Fig. 1–3, respectively. As depicted in Fig. 1–3, the deamination of mCg is exothermic at both levels of theory. Each rate-determining step, the formation of the tetrahedral intermediate for paths A and B and the formation of the thymine glycol tautomer for path C, involves so high activation free energies that it is unlikely for the deamination of mCg to occur by either pathway. The concerned high activation free energies mainly arise from the four-centered transition states for the rate-determining steps. In such cyclic transition states, the orbits required for the bond dissociation and formation are deformed so much that a large amount of deformation energy is substantially needed. Then, it is of great interest whether the activation free energies of paths A, B, and C are reduced by

the contribution of an extra water molecule because it is well-known that water can act not only as a solvent but as a catalyst where it can donate or accept a proton to promote long-range proton transfer. Then, the water-mediated mechanisms with one extra water molecule were investigated in each pathway.

3.1.2 Water-assisted deamination mechanisms. In this section, the water-assisted deamination mechanisms are clarified, where one extra water molecule serves as a bridge to help the transfer of the hydrogen atom in the reaction. In order to conveniently state the water-assisted mechanisms, the names of stationary structures in water-mediated mechanisms are further modified with the prefix w, which implies the participation of an extra water molecule. Due to no significant effects on the deamination of mCg, the conformational changes (A-step 3, A-step 4, and B-step 2) involved in the water-assisted mechanisms have not been discussed in detail any more. The changes in Gibbs free energies (ΔG) and electronic energies (ΔE) for the water-assisted paths A, B, and C are presented in Table 2.

Path A. The optimized geometries and important bond lengths for the stationary points are shown in Fig. 4. As presented in Fig. 4, attributed to the presence of the one-water-bridge, transition state wA-TS1 is expanded to eight-centered structure, in which three proton transfers, namely H1 from N4 to O_{w1} atom, H_{w1} from O_{w1} to O_{w2} atom as well as H_{w3} from O_{w2} to N3 atom, occur synchronously. The activation free energies associated with this step are 12.5 and 13.6 kcal at the B3LYP and MP2 levels, respectively, which are very close to those of the transition state A-TS1 in the direct deamination mechanism. This indicates that expanding the transition state structure from six-membered to eight-membered ring can not further lower the free energy barrier in the gas phase and the six-membered ring transition state A-TS1 is the most favorable reaction route for this proton transfer process.

For wA-TS2, a six-membered ring is formed by carbon C4, oxygen O_{w1}, hydrogen H1, oxygen O_{w2}, hydrogen H_{w1}, and nitrogen N4. Carbon C4 shows a pronounced tetrahedral character. In this step, a single bond is created between C4 and O_{w1} atoms, a proton (H1) is transferred between the two water molecules, and a second one (H_{w1}) is transferred from O_{w2} to N4 atom. The two proton transfers are asynchronous because the H_{w1} proton transfer has proceeded further than the H1 one. The activation free energies associated with this step are 31.8 and 37.2 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively, which are indeed lower than those associated with A-TS2 in the direct deamination mechanism by about 15 kcal mol⁻¹. After the nucleophilic addition of one water molecule to C4 with the assistance of a second water molecule, the C4–N4 bond has to be broken in order to allow deamination of the substrate. In step 5, with one extra water molecule bridging the H_{w2} and N4 atoms, the located transition state wA-TS5 is also a six-membered ring structure, whose imaginary frequency is mainly associated with the transfer of H_{w2} atom from O_{w1} to O_{w2} atom and the transfer of H1 atom from O_{w2} to N4 atom, leading to the cleavage of

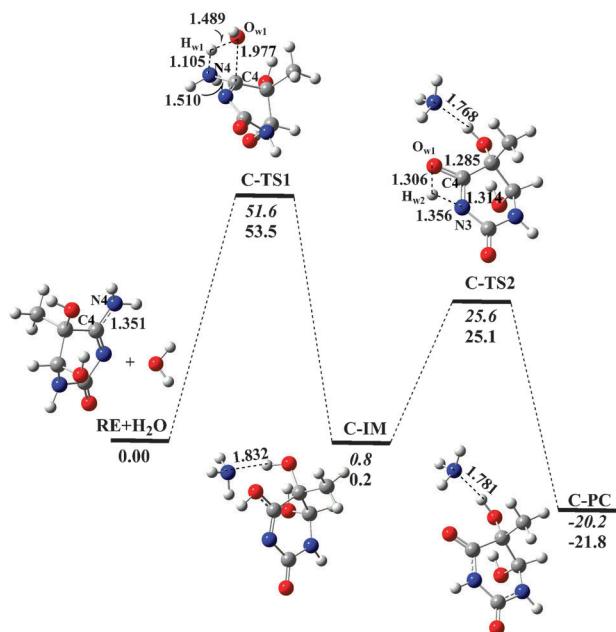


Fig. 3 Potential energy surface (in kcal mol⁻¹) and optimized structures (in Å) for the direct deamination mechanism of *cis*-(5S,6S)-mCg (path C) in the gas phase.

Table 2 Changes of Gibbs free energies (ΔG) and electronic energies (ΔE) for the water-assisted deamination mechanisms of *cis*-(5S,6S)-mCg in the gas phase and aqueous solution (in kcal mol⁻¹)

	B3LYP/6-311G(d,p)		MP2/6-311G(d,p)		Monte Carlo $\Delta(\Delta G_{\text{sol}})$
	$\Delta G(\text{gas})$	$\Delta E(\text{gas})$	$\Delta G(\text{gas})$	$\Delta E(\text{gas})$	
Path A					
RE + 2H ₂ O → wA-RC	-7.2	-24.7	-6.4	-23.7	—
wA-RC → wA-TS1	12.5	10.9	13.6	12.0	1.0 ^a
wA-RC → wA-IM1	3.3	3.0	0.1	-0.5	2.6 ^a
wA-IM1 → wA-TS2	31.8	29.9	37.2	35.4	-9.2
wA-IM1 → wA-IM2	3.8	2.2	-0.5	-2.0	-1.4
wA-RC → wA-IM2	7.1	5.2	-0.4	-2.5	—
wA-IM2 → wA-TS3	11.1	11.7	12.1	12.7	—
wA-IM2 → wA-IM3	5.1	5.3	5.6	5.5	—
wA-RC → wA-IM3	12.2	10.5	5.2	3.1	—
wA-IM3 → wA-TS4	0.9	0.7	1.2	1.1	—
wA-IM3 → wA-IM4	-4.8	-5.1	-4.9	-5.0	—
wA-RC → wA-IM4	7.4	5.4	0.3	-1.9	—
wA-IM4 → wA-TS5	14.1	13.6	16.5	16.0	-5.5
wA-IM4 → wA-PC	-20.1	-18.1	-12.9	-10.8	0.7
wA-RC → wA-PC	-12.7	-12.7	-12.7	-12.7	—
Path B					
RE + 2H ₂ O → wB-RC	-2.6	-19.9	-3.1	-20.5	—
wB-RC → wB-TS1	32.1	29.8	38.4	35.9	-6.4
wB-RC → wB-IM1	2.3	0.3	-4.1	-6.0	-2.2
wB-IM1 → wB-TS2	9.7	10.3	10.7	11.6	—
wB-IM1 → wB-IM2	2.6	2.5	3.7	3.5	—
wB-RC → wB-IM2	4.8	2.7	-0.4	-2.5	—
wB-IM2 → wB-TS3	16.2	15.6	18.7	18.3	-8.3
wB-IM2 → wB-PC	-22.3	-20.5	-15.6	-13.6	-1.7
wB-RC → wB-PC	-17.5	-17.8	-16.0	-16.1	—
Path C					
RE + 2H ₂ O → wC-RC	3.8	-13.6	1.8	-15.8	—
wC-RC → wC-TS1	43.6	42.0	44.8	43.4	-0.5
wC-RC → wC-IM1	-10.3	-10.1	-7.0	-6.6	7.6
wC-IM1 → wC-TS2	5.1	4.1	6.0	5.0	-0.1
wC-IM1 → wC-PC	-15.5	-15.4	-15.0	-14.7	1.3
wC-RC → wC-PC	-25.8	-25.4	-22.0	-21.3	—

^a The values are for the six-membered transition state A-TS1 in the direct deamination reaction.

C4–N4 bond. As shown in Table 2, this elementary step involves relatively low values of activation free energy of 14.1 at the B3LYP level and 16.5 kcal mol⁻¹ at the MP2 level, respectively, which are about 13 kcal mol⁻¹ lower than those of A-TS5. This suggests that the presence of one-water-bridge is significantly beneficial to this proton transfer process.

Path B. The optimized geometries and important bond lengths for the stationary points are shown in Fig. 5. As shown in Fig. 5, both of the transition states, wB-TS1 and wB-TS3, have converted from the four-membered ring structure in B-TS1 and B-TS3 into six-membered ring structure due to the one-water-bridge. As in the case of the A-step 2, the first step of path B also involves two proton transfers and the creation of one C–O single bond. In this case, one proton (H_{w1}) is transferred from O_{w1} to O_{w2} atom, a second proton (H_{w3}) is transferred from O_{w2} to N3 atom, and a single bond is created between carbon C4 and oxygen O_{w1}. The activation free energies associated with this step drop to 32.1 and 38.4 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively, as the result of a decrease by about 8–13 kcal mol⁻¹ with respect to those of B-TS1. Comparing with A-step 5, the final proton transfer process of path B has similar structural feature and energetics. The values of activation free energy concerned in this step are 16.2 and 18.7 kcal mol⁻¹ at the B3LYP and MP2

levels, respectively, which are 14 kcal mol⁻¹ lower than those of B-TS3 but close to those of wA-TS5.

Path C. The optimized geometries and important bond lengths for the stationary points are shown in Fig. 6. As depicted in Fig. 6, a six-membered ring structure transition state wC-TS1 is formed in the cleavage process of C4–N4 bond due to the assistance of one extra water molecule. In wC-TS1, two proton transfers occur synchronously, *i.e.*, H_{w1} from O_{w1} to O_{w2} atom and H_{w3} from O_{w2} to N4 atom. The activation free energy of wC-TS1 are 47.4 and 46.6 kcal mol⁻¹ at the B3LYP and MP2 levels, respectively, which were decreased by about 6 kcal mol⁻¹ compared to those of C-TS1. In the subsequent proton transfer process, the transition state wC-TS2 is expanded to a six-membered ring structure, which mainly corresponds to the transfer of the proton H_{w2} from O_{w1} to O_{w2} atom and H_{w1} from O_{w2} to N3 atom. The activation free energies associated with this step were found to be 5.1 and 6.0 kcal mol⁻¹ as inferred from B3LYP and MP2 level calculations, respectively. This represents a decrease of about 19 kcal mol⁻¹ with respect to those of C-TS2. This result shows that the proton transfer process can be significantly facilitated by the extra water molecule.

The potential energy surfaces for the water-assisted paths A, B, and C are shown in Fig. 4–6, respectively. Compared to

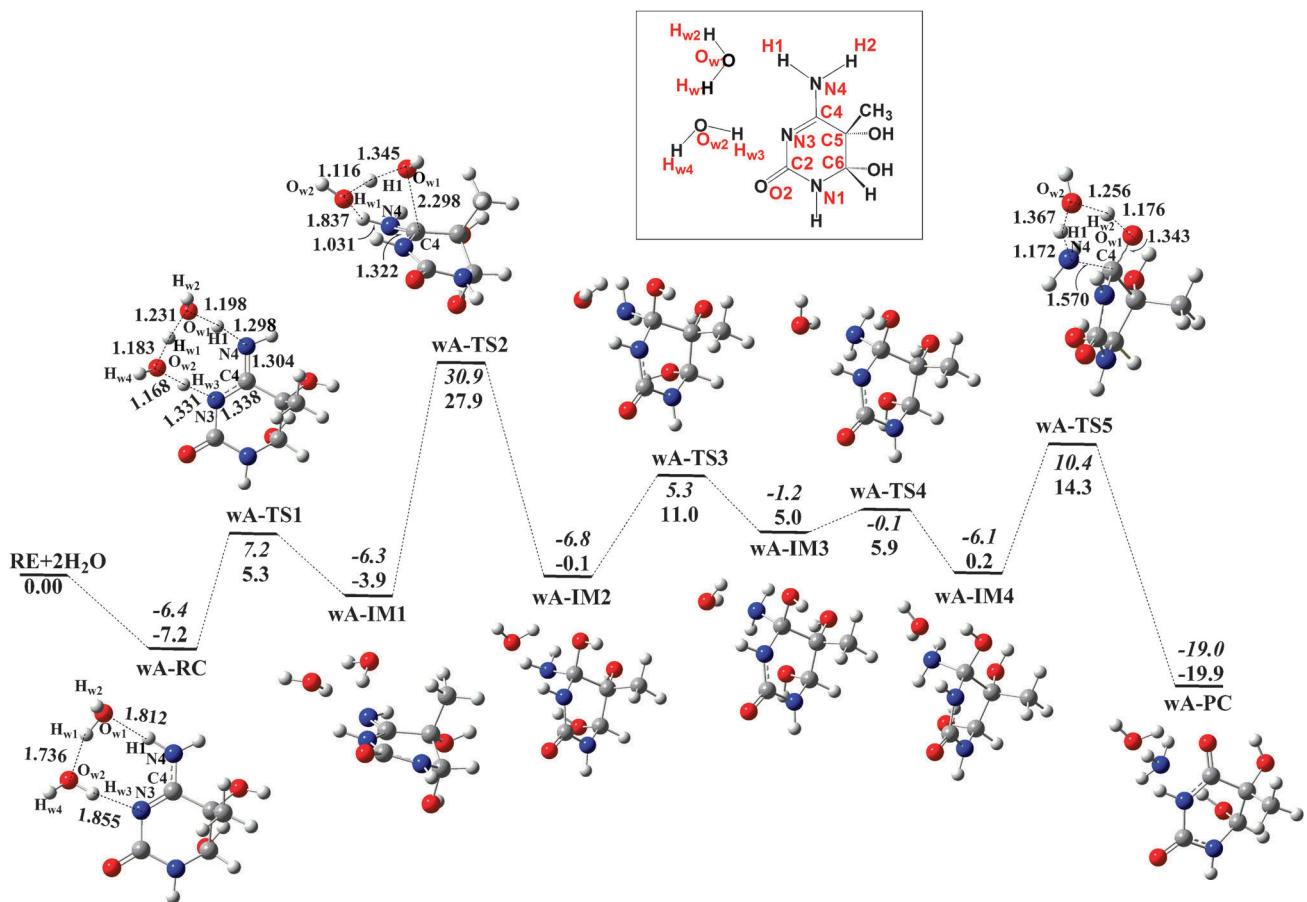


Fig. 4 Potential energy surface (in kcal mol^{-1}) and optimized structures (in \AA) for the water-assisted deamination mechanism of *cis*-(5*S*,6*S*)-mCg (path A) in the gas phase. Inset: atomic labeling used in the current study.

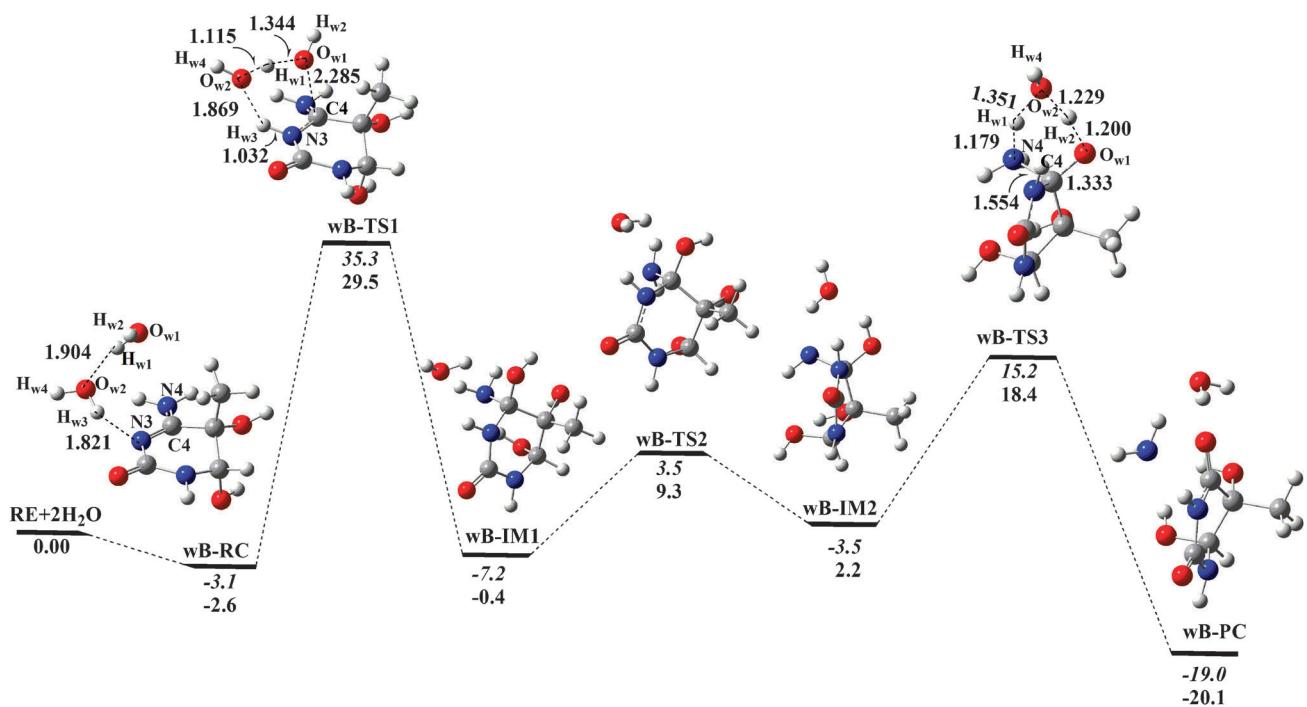


Fig. 5 Potential energy surface (in kcal mol^{-1}) and optimized structures (in \AA) for the water-assisted deamination mechanism of *cis*-(5*S*,6*S*)-mCg (path B) in the gas phase.

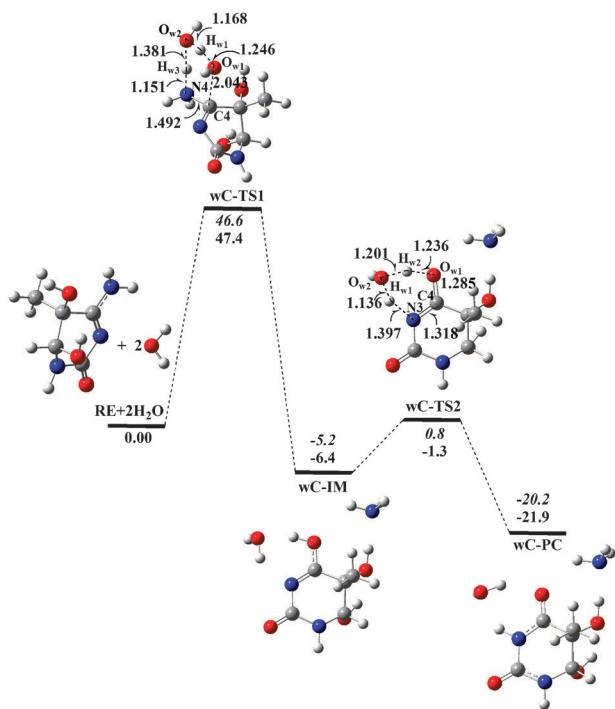


Fig. 6 Potential energy surface (in kcal mol^{-1}) and optimized structures (in \AA) for the water-assisted deamination mechanism of *cis*-(5S,6S)-mCg (path C) in the gas phase.

the direct deamination mechanisms, the potential energy surfaces of the water-assisted mechanisms become smoother with the contribution of the extra water molecule. The transition state in each step is stabilized. It is obvious that the deamination reaction is greatly enhanced by the participation of the extra water molecule. Moreover, this work also shows that mCg has a lower activation free energy of 2 kcal mol^{-1} than mC ($34.0 \text{ kcal mol}^{-1}$) obtained by Labet *et al.*³⁴ based on the B3LYP/6-311G(d,p) level and water-assisted path B, suggesting that the dihydroxylation of mC leads to an increase in the deamination rate. From a biochemical point of view, one of the main striking results of the present DFT computational study which is in agreement with previous experimental data is the higher rate of deamination displayed by mCg residues with respect to mC bases.¹⁹ This provides a relevant molecular basis for the occurrence of mutagenic hot spots at mCpG islands in eukaryotic genes despite the existence of efficient repair of G:T mismatches by dedicated and specific thymine DNA *N*-glycosylases in the DNA of mammalian cells.

3.2 Solvent effects by Monte Carlo simulation

In this section, the solvent effect on the water-assisted deamination of *cis*-(5S,6S)-mCg in aqueous solution was evaluated with Monte Carlo simulation (MC) and free energy perturbation methods. As stated above, owing to no important effects on the deamination mechanism, the conformational changes were not taken into account in this study. So, only seven steps were investigated, *i.e.*, A-step 1, A-step 2, A-step 5, B-step 1, B-step 3, C-step 1, and C-step 2. For A-step 1, the proton transfer process *via* the six-membered transition state A-TS1 is investigated in this study because of its advantage over the

eight-membered ring structure wA-TS1. The changes in free energies of solvation for the activation and reaction procedures of the water-assisted deamination mechanism are presented in Table 2. Fig. 7 depicts the relative potential energy profiles $E_{\text{MEP}}(\text{s})$ along the minimum energy path (MEP) in the gas phase and in aqueous solution. As shown in Fig. 7, the aqueous environment stabilizes the transition states, wA-TS2, wB-TS1, and wC-TS1 by -9.2 , -6.4 , and $-0.5 \text{ kcal mol}^{-1}$, respectively. Our results indicate that the water bulk effect significantly contributes to decreasing the activation free energies of the rate-determining steps of paths A and B and increasing the reaction velocity. Combining B3LYP/6-311G(d,p) calculation with Monte Carlo simulation, the activation free energies of the rate-determining steps for paths A, B, and C are 22.6 , 25.7 , and $43.0 \text{ kcal mol}^{-1}$, respectively, in aqueous solution. It is clear that both of the water-assisted mechanisms (paths A and B) are much more favorable than path C in aqueous solution and have high possibility to occur. Moreover, path A is the most plausible mechanism reported for the deamination of mCg, where the calculated activation free energy agrees well with the experimentally determined activation free energy ($24.8 \text{ kcal mol}^{-1}$),⁴⁸ inferred from the equation given by the activated complex theory⁵¹

$$k = \frac{k_B T}{h} \exp\left(\frac{-\Delta G^\ddagger}{RT}\right) \quad (4)$$

In addition, for paths A and B, the water environment has similar effect on the rate-determining steps but decreases the former more the activation free energy by $2.81 \text{ kcal mol}^{-1}$ than that of the latter. This indicates that, compared to the gas phase, path A has more advantage than path B in the aqueous solution. To account for this, the solute–water energy pair distributions (EPD) for the two rate-determining steps, A-Step 2 and B-Step 1, in the aqueous solution were plotted in Fig. 8. The plots give the number of solvent molecules on the y -axis that interact with the solute and the interaction energy on the x -axis. The bands at low-energy result from the hydrogen-bonded solvent molecules and the spike centered at 0 kcal mol^{-1} comes from the weak interactions between the solute and many distant solvent molecules in outer shells. For A-step 2, as shown in Fig. 8a, hydrogen bonding in water is reflected in the left most regions with energies more attractive than *ca.* $-2.5 \text{ kcal mol}^{-1}$. The hydrogen-bonded energy band for wA-IM1 in water covers the range from -12.5 to $-2.5 \text{ kcal mol}^{-1}$. Integrations of the bands from -12.5 to $-2.5 \text{ kcal mol}^{-1}$ yield 9.8 , 10.9 , and 9.9 water molecules for wA-IM1, wA-TS2, and wA-IM2, respectively. For B-path 1, as presented in Fig. 8b, the situation is similar but the hydrogen-bonded energy band for wB-RC in water ranges from -9.0 to $-3.5 \text{ kcal mol}^{-1}$. Integration of the bands from -9.0 to $-3.5 \text{ kcal mol}^{-1}$ forms 6.2 , 7.2 , and 6.3 water molecules for wB-RC, wB-TS1, and wB-IM1, respectively. Both two transition states, wA-TS2 and wB-TS1, show greater numbers of solvent molecules than their corresponding intermediates or reaction complexes, thus suggesting that in aqueous solution the transition states interact more strongly with water solvent and the solvent plays an important role in the increase of the stabilization of these

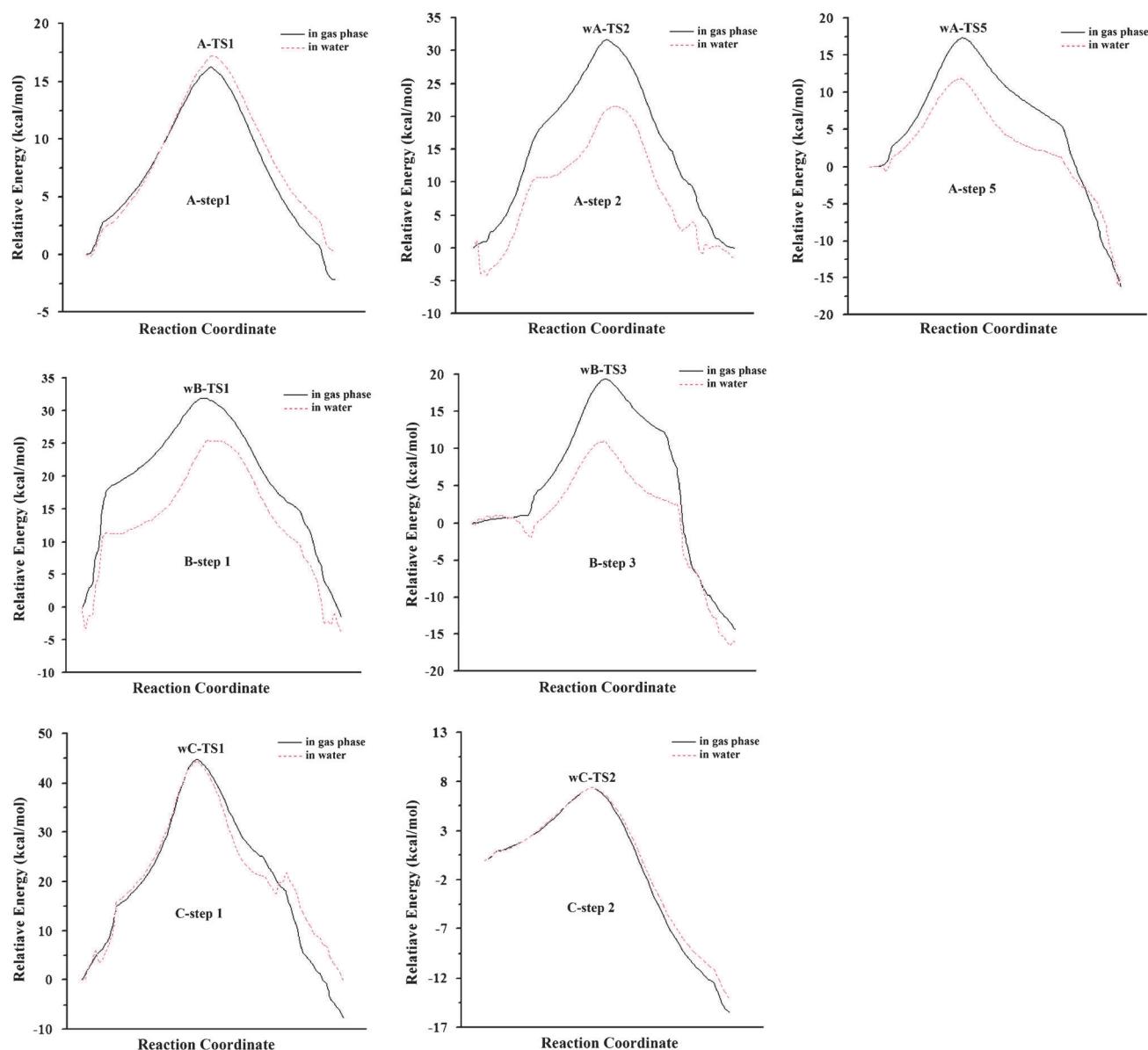


Fig. 7 Potential energy profiles for the water-assisted deamination mechanisms of *cis*-(5*S*,6*S*)-mCg in gas phase and water solution.

transition states, which results in the decreases of the two activation free energies. Furthermore, for A-step 2, the number of the solvent molecules in wA-TS2 is 1.1 larger than that of wA-IM1, whereas for B-step 1, the difference in the solvent molecules number between wB-TS1 and wB-RC is 1.0, smaller than that of A-step2, which, therefore, leads to a less reduction of the activation free energy of B-step 1 in water.

3.3 Effects of methylation and stereochemistry on the deamination rate

As some earlier studies have reported, the methylation of C increases the rate of hydrolytic deamination. Hydrolytic deamination of mC is 3.5 times faster than that of C.^{1,8–10} Then, Another important issue concerns the effect of the C5 methylation on the deamination efficiency of 5,6-saturated pyrimidine glycols. To gain more understanding about the

methylation and stereochemistry on the deamination rates of cytosine glycol components, further investigations on the deaminations of *cis*-(5*S*,6*S*) diastereomer of cytosine glycol (Cg) and *cis*-(5*R*,6*R*)-mCg in the gas phase have also been carried out based on the water-assisted path A, which, as stated above, is the most optimum pathway for the deamination of *cis*-(5*S*,6*S*)-mCg both in the gas phase and aqueous solution. The corresponding optimized structures of stationary points for *cis*-(5*S*,6*S*)-Cg and *cis*-(5*R*,6*R*)-mCg are shown in Fig. 9 and 10, respectively. As shown in Fig. 9 and 10, all the stationary points in the deamination processes of *cis*-(5*S*,6*S*)-Cg and *cis*-(5*R*,6*R*)-mCg have similar structural features with *cis*-(5*S*,6*S*)-mCg. Then, energetic aspects of the deamination reaction were considered. As the initial product of cytosine oxidation, Cg is unstable and can easily convert into stable uracil glycol through deamination reaction.^{52,53} As shown in Fig. 9, the rate-determining step for the deamination of

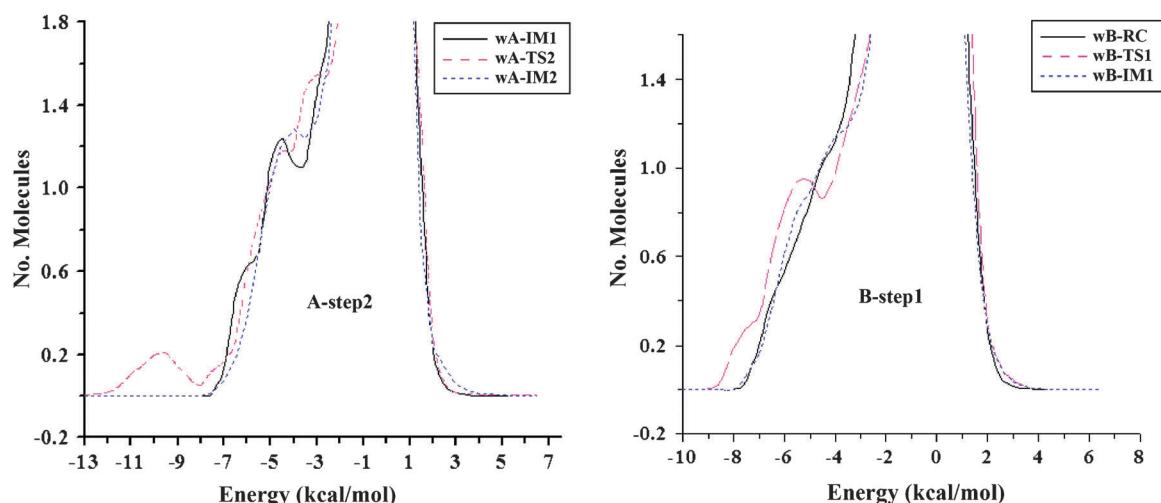


Fig. 8 Pair-energy distributions of the solute-water interaction for the rate-determining steps of the water-assisted paths A (a) and B (b) of *cis*-(5S,6S)-mCg in water.

cis-(5S,6S)-Cg still involves the formation of the tetrahedral intermediate with the activation free energies of 30.1 and 35.7 kcal mol⁻¹ at the B3LYP and MP2 level, respectively, which are about 2 kcal mol⁻¹ lower than those of *cis*-(5S,6S)-mCg. This indicates that the presence of a methyl group at C5 leads to a decrease in the deamination rate of the ring-saturated glycol, which may result from the steric hindrance of the methyl group.

As the other *cis* diastereomer of mCg, *cis*-(5R,6R) diastereomer has the same rate-determining step as *cis*-(5S,6S)-mCg and *cis*-(5S,6S)-Cg have done, whose activation free energy is 32.5 kcal mol⁻¹ at the B3LYP theory of level and 38.1 kcal mol⁻¹ at the MP2 level, higher than those of *cis*-(5S,6S)-mCg by 1 kcal mol⁻¹ or so. This suggests that the 2'-deoxyribonucleoside derivative of *cis*-(5S,6S)-mCg exhibits a slightly higher

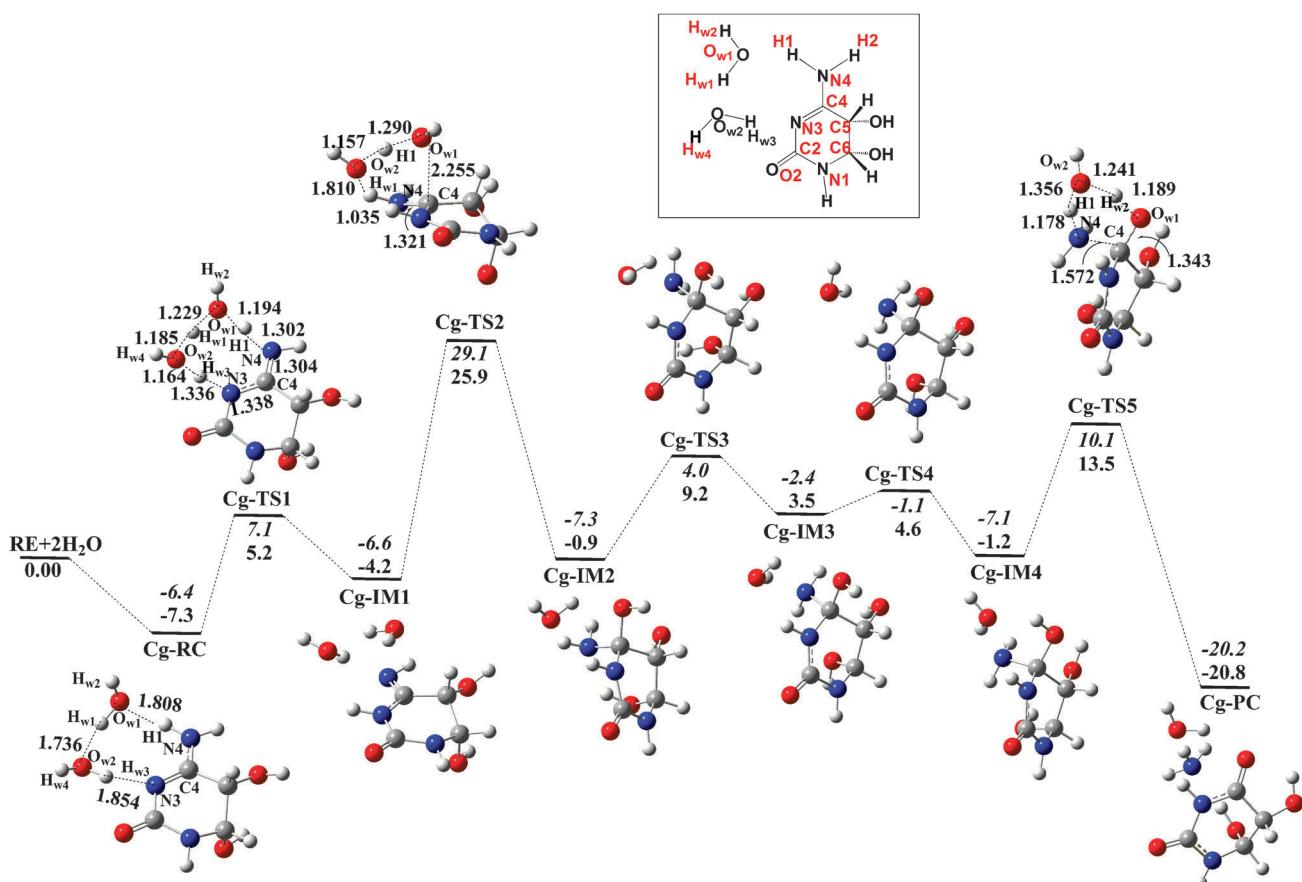


Fig. 9 Potential energy surface (in kcal mol⁻¹) and optimized structures (in Å) for the water-assisted deamination mechanism of *cis*-(5S,6S)-Cg (path A) in the gas phase. Inset: atomic labeling used in the current study.

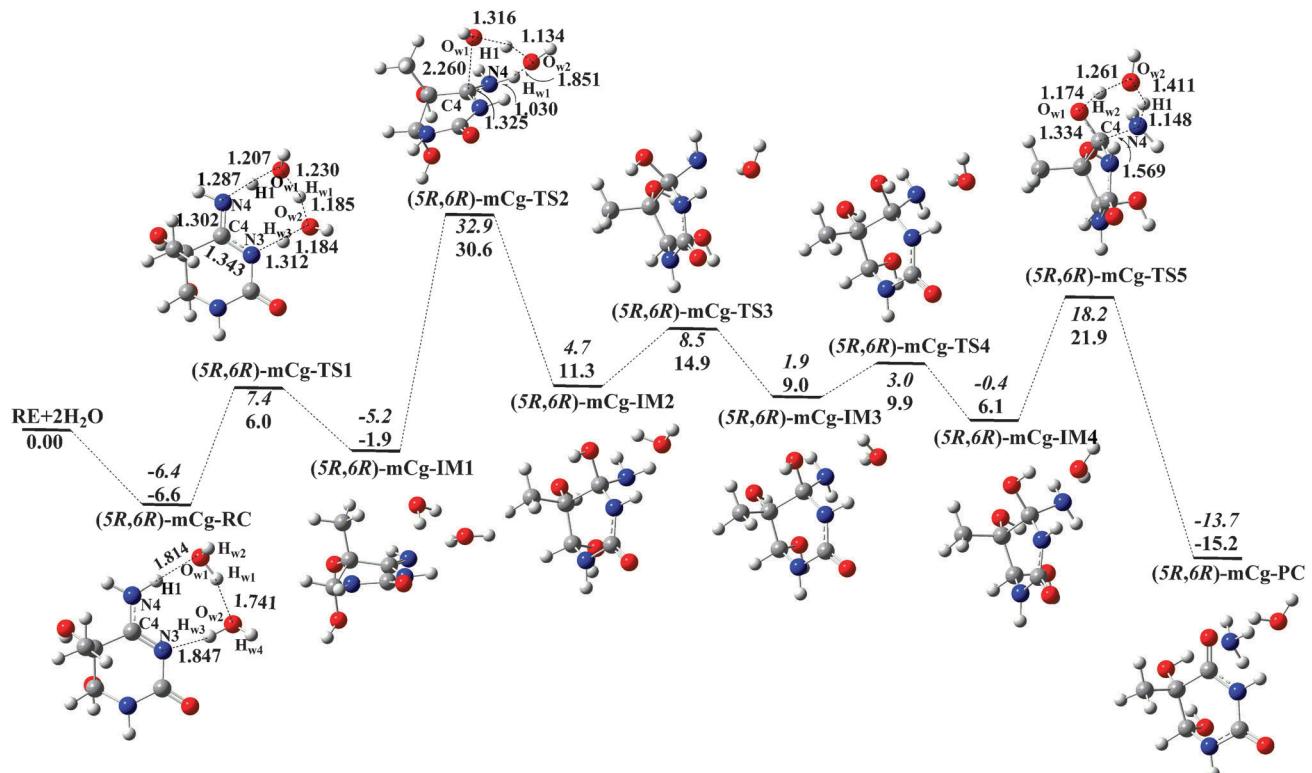


Fig. 10 Potential energy surface (in kcal mol⁻¹) and optimized structures (in Å) for the water-assisted deamination mechanism of *cis*-(5R,6R)-mCg (path A) in the gas phase.

deamination rate than the *cis*-(5R,6R) one. In other words, the former is more prone to deamination than the latter, which may be attributed to an ordinary transition state in *cis*-(5R,6R)-mCg as the result of immobilization of water molecules, a process which is correlated with a decrease in the reaction rate.⁵⁴ All the results can reappear the experimental data well.^{25,55}

4. Conclusions

The formation and deamination of 5-methylcytosine glycol may contribute significantly to the C → T transition mutation at mCpG dinucleotide site. This paper reports the first detailed study of possible mechanisms for the deamination of mCg with H₂O in the gas phase and in aqueous solution using *ab initio* quantum mechanics and free energy perturbation implemented in Monte Carlo simulations. Optimized geometries were determined at the B3LYP/6-311G(d,p) and MP2/6-311G(d,p) levels of theory. Three pathways for deamination of mCg with H₂O were found. In each pathway, two distinct reaction mechanisms were considered in this work, a direct deamination mechanism and a water-assisted mechanism. It is clear that the direct participation of an auxiliary water molecule in the reaction is crucial for obtaining a realistic barrier and the bulk solvent water molecules significantly contribute to decreasing the activation free energy and accelerate the reaction velocity. Our computational results in kinetics and thermodynamics clearly manifest that the water-assisted paths A and B both have high possibility to occur in aqueous solution and the water-assisted path A is the most plausible mechanism reported for the deamination of mCg,

which agrees well with the available experimental conclusion. Further investigations about the effects of C5 methylation, 5,6-dihydroxylation and stereochemistry on the rate of deamination of cytosine derivatives in nucleosides indicate that the dihydroxylation of mC at C5-C6 leads to a significant increase in deamination susceptibility, whereas methylation of Cg at C5 site slows down the rate of deamination. It is also found that *cis*-(5S,6S)-mCg is more prone to deamination than the (5R,6R) diastereomer. Our results agree well with the experimental data and support that the deamination of mCg may contribute significantly to the C → T transition mutation at mCpG dinucleotide site.

Acknowledgements

This project has been supported by the National Natural Science Foundation of China (Grant Nos. 20773089 and 20835003) and National Basic Research Program of China (973 Program)(2011CB201202).

References and notes

- 1 G. P. Pfeifer, *Mutat. Res.*, 2000, **450**, 155.
- 2 M. Ehrlich, M. A. Gamasa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune and C. Gehrke, *Nucleic Acids Res.*, 1982, **10**, 2709.
- 3 A. Razin and A. D. Riggs, *Science*, 1980, **210**, 604.
- 4 P. Grippo, M. Iaccarino, E. Parisi and E. Scarano, *J. Mol. Biol.*, 1968, **36**, 195.
- 5 D. W. Yandell, *N. Engl. J. Med.*, 1989, **321**, 1689.
- 6 D. N. Cooper and Y. Youssoufian, *Hum. Genet.*, 1988, **78**, 151.
- 7 J. M. Nigro, *Nature*, 1989, **342**, 705.

- 8 M. F. Denissenko, J. X. Chen, M. S. Tang and G. P. Pfeifer, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 3893.
- 9 J. C. Shen, W. M. III Rideout and P. A. Jones, *Nucleic Acids Res.*, 1994, **22**, 972.
- 10 D. H. Lee, T. R. O'Connor and G. P. Pfeifer, *Nucleic Acids Res.*, 2002, **30**, 3566.
- 11 F. T. Liu and N. C. Yang, *Biochemistry*, 1978, **17**, 4865.
- 12 A. Ziegler, D. J. Leffel, S. Kunala, H. W. Sharma, P. E. Shapiro, A. E. Bale and D. E. Brash, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 4216.
- 13 N. Jiang and J. S. Taylor, *Biochemistry*, 1993, **32**, 472.
- 14 W. Peng and B. R. Shaw, *Biochemistry*, 1996, **35**, 10172.
- 15 L. A. Frederico, T. A. Kunkel and B. R. Shaw, *Biochemistry*, 1990, **29**, 2532.
- 16 W. M. III Rideout, G. A. Coetzee, A. F. Olumi and P. A. Jones, *Science*, 1990, **249**, 1288.
- 17 B. Hendrich, U. N. Hardeland, H. H. Jiricny and J. Bird, *A. Nature*, 1999, **401**, 301.
- 18 P. Neddermann, P. Gallinari, T. Lettieri, D. Schmid, O. Truong, J. J. Hsuan, K. Wiebauer and J. J. Jiricny, *J. Biol. Chem.*, 1996, **271**, 12767.
- 19 H. C. Cao, Y. Jiang and Y. S. Wang, *Nucleic Acids Res.*, 2009, **37**, 6635.
- 20 M. S. Cooke, M. D. Evans, M. Dizdaroglu and J. Lunec, *FASEB J.*, 2003, **17**, 1195.
- 21 T. Finkel and N. J. Holbrook, *Nature*, 2000, **408**, 239.
- 22 E. S. Henle and S. J. Linn, *J. Biol. Chem.*, 1997, **272**, 19095.
- 23 S. J. Zuo, R. J. Boorstein and G. E. Teebor, *Nucleic Acids Res.*, 1995, **23**, 3239.
- 24 J. R. Wagner and J. Cadet, *Acc. Chem. Res.*, 2010, **43**, 564.
- 25 C. Bienvenu and J. Cadet, *J. Org. Chem.*, 1996, **61**, 2632.
- 26 R. E. Johnson, S. L. Yu, S. Prakash and L. Prakash, *Genes Dev.*, 2003, **17**, 77.
- 27 P. L. Fischhaber, V. L. Gerlach, W. J. Feaver, Z. Hatahet, S. S. Wallace and E. C. Friedberg, *J. Biol. Chem.*, 2002, **277**, 37604.
- 28 R. Kusumoto, C. Masutani, S. Iwai and F. Hanaoka, *Biochemistry*, 2002, **41**, 6090.
- 29 M. H. Almatarneh, C. G. Flinn, R. A. Poirier and W. A. Sokalski, *J. Phys. Chem. A*, 2006, **110**, 8227.
- 30 M. H. Almatarneh, C. G. Flinn and R. A. Poirier, *J. Chem. Inf. Model.*, 2008, **48**, 831.
- 31 T. Matsubara, M. Ishikura and M. Aida, *J. Chem. Inf. Model.*, 2006, **46**, 1276.
- 32 V. Labet, A. Grand, C. Morell, J. Cadet and L. A. Eriksson, *Theor. Chem. Acc.*, 2008, **120**, 429.
- 33 V. Labet, A. Grand, J. Cadet and L. A. Eriksson, *ChemPhysChem*, 2008, **9**, 1195.
- 34 V. Labet, C. Morell, J. Cadet, L. A. Eriksson and A. Grand, *J. Phys. Chem. A*, 2009, **113**, 2524.
- 35 V. Labet, C. Morell, J. Cadet, T. Douki, J. Cadet, L. A. Eriksson and A. Grand, *J. Phys. Chem. A*, 2010, **114**, 1826.
- 36 (a) M. J. Lustig, J. Cadet, R. J. Boorstein and G. W. Teebor, *Nucleic Acids Res.*, 1992, **20**, 4839; (b) K. L. Brown, T. Adams, V. P. Jasti, A. K. Basu and M. P. Stone, *J. Am. Chem. Soc.*, 2008, **130**, 11701.
- 37 (a) Y. Xue and C. K. Kim, *J. Phys. Chem. A*, 2003, **107**, 7945; (b) Y. Xue, C. K. Kim, Y. Guo, D. Q. Xie and G. S. Yan, *J. Comput. Chem.*, 2005, **26**, 994; (c) Y. Wu, Y. Xue, D. Q. Xie, C. K. Kim and G. S. Yan, *J. Phys. Chem. B*, 2007, **111**, 2357; (d) X. F. Xia, C. H. Zhang, Y. Xue, C. K. Kim and G. S. Yan, *J. Chem. Theory Comput.*, 2008, **4**, 1643.
- 38 (a) F. E. Hruska, R. Sebastian, A. Grand, L. Voituriez and J. Cadet, *Can. J. Chem.*, 1987, **65**, 2618; (b) Z. Q. Chen, C. H. Zhang and Y. Xue, *J. Phys. Chem. B*, 2009, **113**, 10409.
- 39 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision D.01)*, Gaussian, Inc., Wallingford, CT, 2004.
- 40 V. Labet, C. Morell, A. Grand and A. Toro-Labbe, *J. Phys. Chem. A*, 2008, **112**, 11487.
- 41 K. Fukui, *J. Phys. Chem.*, 1970, **74**, 4161.
- 42 C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361.
- 43 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1989.
- 44 W. J. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
- 45 W. J. Jorgensen and C. Ravimohan, *J. Chem. Phys.*, 1985, **83**, 3050.
- 46 R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420.
- 47 W. L. Jorgensen and J. Tirado-Rives, *J. Comput. Chem.*, 2005, **26**, 1689.
- 48 G. A. Kaminski and W. J. Jorgensen, *J. Phys. Chem. B*, 1998, **102**, 1787.
- 49 E. M. Duffy, D. L. Severance and W. J. Jorgensen, *J. Am. Chem. Soc.*, 1992, **114**, 7535.
- 50 W. J. Jorgensen, *J. Phys. Chem.*, 1986, **90**, 1276.
- 51 H. Eyring, *Chem. Rev.*, 1935, **17**, 65.
- 52 S. Tremblay and J. R. Wagner, *Nucleic Acids Res.*, 2008, **36**, 284.
- 53 S. Tremblay, T. Douki, J. Cadet and J. R. Wagner, *J. Biol. Chem.*, 1999, **274**, 20833.
- 54 A. J. S. C. Vieira and S. Steenken, *J. Phys. Chem.*, 1987, **91**, 4138.
- 55 T. Douki and J. Cadet, *Biochemistry*, 1994, **33**, 11942.