PCCP

Accepted Manuscript

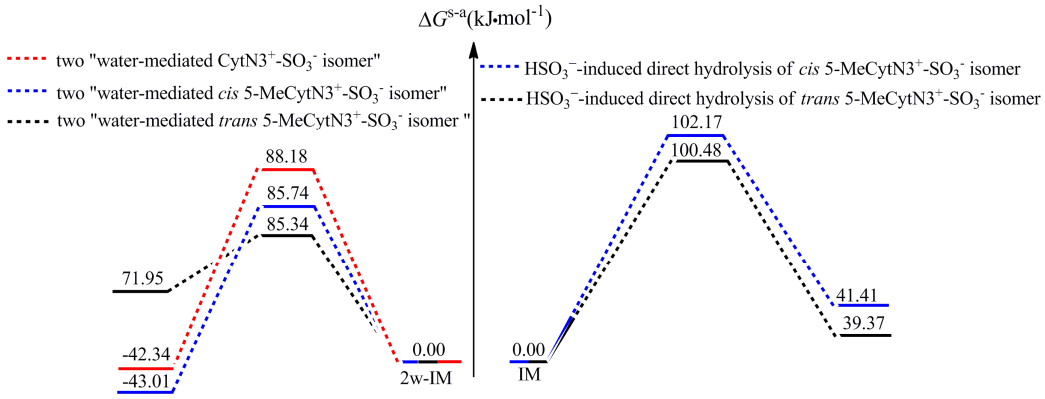
ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Table of contents

Is the Equal Contribution of *Cis* and *Trans* Protonated 5-Methylcytosine-SO_3^- Isomers to Thymine-SO_3^- under Bisulfite Conditions? A Theoretical Perspective

Lingxia Jin, Lu Wang, Caiying Zhang, Wenliang Wang, Suotian Min, and Daodao Hu



Activation free energies ($\Delta G^{s\text{-}a\neq}$) of the rate-limiting step for the direct hydrolytic paths are obviously reduced by extra two water molecules, and the $\Delta G^{s\text{-}a\neq}$ of the rate-determining step for *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers paths exhibit no obvious difference. Meanwhile, the $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for both 5-MeCytN3$^+$-SO_3^- isomers paths are very close to that of the theoretical date for CytN3$^+$-SO_3^-.

# Is the Equal Contribution of *Cis* and *Trans* Protonated 5-Methylcytosine-SO_3^- Isomers to Thymine-SO_3^- under Bisulfite Conditions? A Theoretical Perspective†

Lingxia Jin[a,b] Lu Wang[a] Caiying Zhang[a] Wenliang Wang[a,*] Suotian Min[b] Daodao Hu[a]

[a]*Key Laboratory for Macromolecular Science of Shaanxi Province, School of Chemistry and Chemical Engineering, Shaanxi Normal University, Xi'an 710062*

[b]*School of Chemical and Environmental Sciences, Shaanxi University of Technology, Hanzhong 723001*

ABSTRACT

Cytosine (Cyt) can be converted to 5-methylcytosine (5-MeCyt) in CpG sequences of DNA. The conventional bisulfite sequencing can discriminate Cyt from 5-MeCyt, whereas the inappropriate conversion of 5-MeCyt to thymine and a failure conversion of Cyt to uracil always occurred when Cyt and 5-MeCyt is treated by bisulfite, which would lead to the erroneous estimates of DNA methylation densities. Here the direct hydrolytic deamination of *cis* (paths A-C) and *trans* (paths A'-C') 5-MeCytN3$^+$-SO_3^- isomers with bisulfite have been explored at the MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) level. The activation free energies ($\Delta G^{s\text{-}a\neq}$) of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers paths exhibit no obvious difference, implying both isomers may make an equal contribution to the hydrolytic deamination of 5-MeCyt under bisulfite conditions. This result has high expectations for the experimental scientists to explore new methods to avoid the formation of the deaminated reactants (5-MeCytN3$^+$-SO_3^-). Meanwhile, the HSO_3^--induced direct hydrolytic deamination of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers paths A and A' are represented and have been further explored in the presence of two water molecules. The result is that the contribution of two water molecules make the HSO_3^--induced direct hydrolytic deamination of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers paths A and A' become favourable. In

*Corresponding author: Wenliang Wang Tel: +86-29-81530815, Fax: +86-81530727

E-mail: wlwang@snnu.edu.cn (W. L. Wang).

†Electronic supplementary information (ESI) available: The corresponding geometrical structures of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers in the aqueous phase are listed in Fig. S1. The important bond lengths of all stationary points for direct deamination paths of *cis* (paths A-C) and *trans* 5-MeCytN3$^+$-SO_3^- (paths A'-C') isomers and two water-mediated paths A and A' in the aqueous phase are listed in Figs.S2-S9. The relative energies for two water-mediated path of CytN3$^+$-SO_3^- both in the gas and aqueous phases are listed in Table S1. The relative energies for one water-mediated paths A and A' both in the gas and aqueous phases are listed in Tables S2 and S3.

addition, the $\Delta G^{s\text{-}a\neq}$ (85.74~85.34 kJ·mol^{-1}) of the rate-limiting step of two water-mediated paths A and A' are very close to that of the theoretical date for CytN3^{+}-SO$_3^{-}$ (88.18 kJ·mol^{-1}), implicating that the free barrier gap is very small between Cyt and 5-MeCyt in bisulfite conditions. That further suggests that the bisulfite sequencing technology may be easily influenced by external environment.

1. Introduction

Analysis of methylation states in genomic DNA has provided insights into biological phenomena as disparate as human disease and atypical floral morphologies. [1-6] In eukaryotic genomes, DNA methylation usually involves addition of a methyl group to the 5-carbon of cytosine (Cyt), yielding 5-methylcytosine (5-MeCyt). Because both Cyt and 5-MeCyt are complementary to guanine, conventional sequencing does not distinguish between them. Of the various methods, bisulfite genomic sequencing is a method most typically used to assess DNA methylation states. [7-10] Treatment of DNA with bisulfite converts Cyt into uracil by deamination, while 5-MeCyt remains unaltered.[7,11] However, the experiments demonstrated that the inappropriate conversion of 5-MeCyt to thymine and a failure conversion of Cyt to uracil always occurred, [12] and both types of conversion would lead to the erroneous estimates of DNA methylation densities. Therefore, it is significantly important to understand these inappropriate conversion mechanisms for improving the reliability of the bisulfite sequencing technology.

As for 5-MeCyt, the asymmetric carbon at position 5 exists from the viewpoint of stereochemistry. Shiragami et al. [13] demonstrated that two diastereomers of 5,6-dihydrothymine-6-sulfonate were formed by the deamination of 5-MeCyt with bisulfite. Meanwhile, our previously calculated results also manifested *cis* and *trans* 5-MeCyt-SO$_3^{-}$ isomers obtained in the addition reaction of bisulfite with 5-MeCyt (Fig. S1), [14] thus the deamination of 5-MeCyt-SO$_3^{-}$ may proceed by two possible types of routes, raising that the question of whether there is competition between them treated with bisulfite.

Based on experiments[15] and our previous theoretical analysis[14], the speed of the HSO$_3^{-}$ group addition to 5-MeCyt is significantly faster in acidic medium than in neutral condition, and the key step of deamination proceeds principally through the protonated adduct. Moreover, recent theoretical works[16-19] also indicate that the activation free energy is likely to be further lowered in

acidic medium. Therefore, the acidic media may tremendously accelerate the deamination of *cis* and *trans* 5-MeCyt-SO_3^- isomers (5-MeCytN3$^+$-SO_3^-). In addition, Hayatsu et al. have discovered the rate of CytN3$^+$-SO_3^- to uracil conversion proportionately is dependent on bisulfite concentrations,[20,21] and our previous theoretical work[22] has also demonstrated the presence of the HSO_3^- group significantly contributes to facilitating the direct hydrolytic process of CytN3$^+$-SO_3^-, suggesting that the HSO_3^- group may also greatly accelerate the direct hydrolytic deamination of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers.

As mentioned above, the direct hydrolytic deamination of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers with the HSO_3^- group are firstly explored from a theoretical perspective to clarify which of isomers can easily deaminate to thymine in the presence of the HSO_3^- group. Meanwhile, combined the recent theoretical works [23-28] and our previous reports[29], the participation of the extra water molecules is likely to further change the activation free energy. Therefore, the HSO_3^--induced direct hydrolytic deamination of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers may be affected by the presence of the additional water. Finally, the hydrolytic deamination mechanisms and activation free energies of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers have been compared with those of CytN3$^+$-SO_3^- in the participation of HSO_3^- group. The aim of this study is to make clear the mechanism of the inappropriate-conversion event and to find out the essential differences between Cyt and 5-MeCyt under bisulfite conditions, which may be the theoretical aid for the experimental scientists.

2. Computational methods

All the calculations were performed using the Gaussian 09 package.[30] The diffuse function has better performance in predicting the geometries of anionic systems, so the density functional method B3LYP[31] with a standard 6-311++G(d,p) basis set[32] was applied to the gas phase calculations and additionally the polarizable continuum model (PCM)[33] with dielectric constant 78.39 of the solvent for the aqueous solution. Frequency analysis has also been computed at the same level of theory to verify whether the obtained structures are transition structures or local minima. Intrinsic reaction coordinate (IRC)[34] calculations have been carried out from each transition state to ensure that the obtained transition state connected the appropriate reactants and products. From our previous work[22], it has been found that the activation free energy calculated

using the MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) level[35] (hereafter named MP2// B3LYP method) is in close proximity to the previous experimental data, thus the single point energies of the species have been further refined at the MP2// B3LYP method. Noted that the changes in free energies and the free energy barriers in solution with PCM model based on the optimized geometries in the gas and aqueous phases are donated $\Delta G^{s\text{-}g}$, $\Delta G^{s\text{-}a}$, $\Delta G^{s\text{-}a\neq}$ and $\Delta G^{s\text{-}g\neq}$, respectively.

3. Results and discussion

3.1 Stationary point structures and energetics in the gas phase

Based upon pathways proposed in our previous work[22] for the HSO_3^- -induced direct hydrolytic deamination of CytN3$^+$-SO_3^-, three possible reaction pathway for the direct hydrolytic deamination of each 5-MeCytN3$^+$-SO_3^- isomer in the presence of HSO_3^- group were taken into account in our work, and three possible reaction pathway were obtained respectively by two isomers. Thus, our interest has been focused on the direct hydrolytic deamination routes of *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomers in the following discussion.

3.1.1 HSO_3^--induced direct hydrolytic deamination of *cis* 5-MeCytN3$^+$-SO_3^- isomer

Indentify with CytN3$^+$-SO_3^- paths, the $HSO_3^-\cdots H_2O$ group is approaching to the left, above, and right side of *cis* 5-MeCytN3$^+$-SO_3^- isomer, denoted as paths A, B, and C, respectively. The optimized geometries and corresponding potential energy profile are shown in Figs.1-4. The relative energies (ΔE^g), the relative free energies (ΔG^g, $\Delta G^{s\text{-}g}$, and $\Delta G^{s\text{-}a}$), the activation free energies ($\Delta G^{g\neq}$, $\Delta G^{s\text{-}g\neq}$, and $\Delta G^{s\text{-}a\neq}$) of the different stationary points are presented in Table 1.

Path A As presented in Fig.1, with an approaching the HSO_3^- group toward the left side of *cis* 5-MeCytN3$^+$-SO_3^- isomer (close to H2 atom), the *cis* 5-MeCytN3$^+$-SO_3^- isomer forms a complex (A-RC) with one HSO_3^- group and one H_2O, in which three hydrogen bonds exist, which is 211.09 kJ·mol^{-1} more stable than the sum of monomers. With the approaching of the $HSO_3^-\cdots H_2O$ group to C4 atom and the elongation of O5-H6 bond, in A-TS1, the two eight-centered structures are formed by the C4-O5-H6-O1-S1-O2-H2-N4 and C4-O5-H6-O1-S1-O2-H1-N3, respectively, in which one hydrogen transfer and one hydroxyl oxygen addition, namely H6 from O5 to O1 atom and O5 addition to C4 atom, occur synchronously. The $\Delta G^{g\neq}$ associated with this step is 112.10 kJ·mol^{-1}. After the nucleophilic addition O5 to C4 atom, the C4-N4 bond has to be broken in order

to deaminate. For this to occur (starting from A-IM1), the hydroxyl hydrogen must be adjusted to make O4-H4 bond weaken, and then easily transfer to N4 atom. Indentify with $CytN3^+$-SO_3^- paths, this path, besides the formation of multiply hydrogen bonded transition states, involves the hydrogen transfers from H_2SO_3 to the amino group via an eight-membered transition state A-TS3 with effectively a barrierless process for producing the intermediate A-IM3. And then, the H5 from hydroxyl oxygen transfers to O2 site of the HSO_3^- group with simultaneous the NH_3 released (A-TS4), yielding the hydrogen-bonded complex A-P (*cis* T···NH_3···H_2SO_3).

Path B As shown in Fig. 2, introducing the HSO_3^- group directly approaching to the above side of *cis* 5-$MeCytN3^+$-SO_3^- isomer (close to H2 and H3), this reveals a trend that is similar to the previous case. This path exhibits a prereactive complex with a strong hydrogen bond having a binding energy of 182.74 kJ·mol^{-1}. The transition state B-TS1 is similar to the one for path A, with the HSO_3^- group aiding in the addition of the OH group to C4 site, and the $\Delta G^{g\neq}$ obtained in this step is 100.76 kJ·mol^{-1}. Identical with path A, path B also involves hydrogen transfer from H_2SO_3 to N4 site as well as the hydrogen from hydroxyl oxygen transfer to oxygen site of the HSO_3^- group. The $\Delta G^{g\neq}$ of this two steps are both small as expected and do not have significant effects on the mechanism of this reaction.

Path C As shown in Fig. 3, with nucleophilic attack by the HSO_3^-···H_2O group close to the right moiety of *cis* 5-$MeCytN3^+$-SO_3^- isomer, results in the formation of the intermediate anion C-RC, in which lies 197.76 kJ·mol^{-1} below the reactants. Dissimilar with A-RC and B-RC, the geometry of C-RC is far from the initial structure design, and the HSO_3^-···H_2O group is always close to the left moiety of *cis* 5-$MeCytN3^+$-SO_3^- isomer, which may result from the hydrophobic interactions of the methyl group. Indentify with A-TS1 and B-TS1, the HSO_3^- group also acts as charge dispersion in C-TS1, abstracting the hydrogen from water to O1 atom of the HSO_3^- group with the $\Delta G^{g\neq}$ of 98.67 kJ·mol^{-1}. Similar with the two previous cases, the subsequent steps of path C also involves the creation of O-H single bonds, the rotation of single bond, and the cleavage of C-N bond. Unlike with paths A and B, path C involves the hydrogen transfer from the hydroxyl oxygen to NH_2, which is the rate-determining step associated with the $\Delta G^{g\neq}$ of 124.56 kJ·mol^{-1}. The $\Delta G^{g\neq}$ for path C is much higher than those of paths A and B, suggesting that the path C in the gas phase is less possible to occur.

As depicted in Figs. 1-3, these reaction mechanisms mainly take place in two stages: (a) with

the $HSO_3^-\cdots H_2O$ group gradually approaching the C4 atom, the H6 transfer and O5 addition occur synchronously. The HSO_3^- group acts as charge dispersion in this stage. And correspondingly, the O1-H6 and O5-C4 bonds for three paths are formed in this stage. This process is taken as the charge dispersion stage. (b) The hydrogen transfers play a key role in the second stage. The hydrogen transfer from H_2SO_3 to N4 site or the hydrogen transfer from OH group to N4 site, leading to the formation NH_3 and C4-N5 bond partially rupturing. And then, the hydrogen transfer to oxygen site of the HSO_3^- group, results in H_2SO_3 formed and C4-N5 bond completely ruptured. This process is named as the hydrogen transfer stage. As seen from Fig. 4, in this stage, the $\Delta G^{g\neq}$ of the hydrogen transfer from H_2SO_3 to the amino group is very small, whereas for the hydrogen transfer from OH group to N4 site, the $\Delta G^{g\neq}$ is very high. As for paths A and B, the rate-determining step is the first stage involved the $HSO_3^-\cdots H_2O$ group nucleophilic attack C4 atom, whereas for path C, the second stage involved the hydrogen transfer from the hydroxyl oxygen to NH_2 is the rate-determining step. Similarly, the same reaction mechanisms also exist in *trans* 5-MeCytN3$^+$-SO_3^- isomer paths.

3.1.2 HSO_3^--induced direct hydrolytic deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer

Similar with *cis* 5-MeCytN3$^+$-SO_3^- isomer, three paths for the direct hydrolysis deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer with the HSO_3^- group also exist, namely paths A', B', and C', respectively. Note that the conformational changes involved $\Delta G^{g\neq}$ are still small and have no significant effects on the *trans* 5-MeCytN3$^+$-SO_3^- isomer mechanisms. Thus, the rate-limiting steps of these paths are only discussed in the subsequent context. The relative energies (ΔE^g), the relative free energies (ΔG^g, $\Delta G^{s\text{-}g}$, and $\Delta G^{s\text{-}a}$), the activation free energies ($\Delta G^{g\neq}$, $\Delta G^{s\text{-}g\neq}$, and $\Delta G^{s\text{-}a\neq}$) of the different stationary points are presented in Table 2. The optimized geometries and corresponding potential energy profile are shown in Figs. 5-8.

As shown in Figs. 5, 6, and 8, the rate-determining step of *trans* 5-MeCytN3$^+$-SO_3^- isomer paths A' and B' also involves one hydrogen transfer and one hydroxyl oxygen addition with the $\Delta G^{g\neq}$ of 99.11 kJ·mol^{-1}, 95.67 kJ·mol^{-1}, respectively. As shown in Figs. 7 and 8, similar with path C, the rate-determining step for path C' involves the hydrogen transfer from OH group to N4 site associated with $\Delta G^{g\neq}$ of 126.24 kJ·mol^{-1}, suggesting that path C' in the gas phase is significant to the disadvantage of this hydrolysis process. As seen from Figs. 4 and 8, the $\Delta G^{g\neq}$ of the rate-determining step for paths A, B, A', and B' are much lower than those of paths C and C'. Then,

it is of great interest whether the $\Delta G^{g\neq}$ of paths A, B, A', and B' are further reduced by the contribution of the extra water molecules. However, the geometry optimization for some important transition structures in the aqueous phase along path B' failed, so it is difficult to compare the direct hydrolysis with extra water-mediated hydrolysis by the contribution of the bulk water. Hence, the effect of the additional water molecules on paths A and A' will be investigated.

3.1.3 Effect of the extra water molecules on paths A and A'

In this section, the extra water molecules mediated the HSO_3^--induced direct hydrolysis deamination (hereafter named "water-mediated mechanisms") is clarified. In addition, two additional water molecules were added in active sites for the HSO_3^--induced direct hydrolysis deamination of $CytN3^+$-SO_3^-, where the calculated activation free energies($\Delta G^{s\text{-}a\neq}$=88.18 kJ·mol^{-1} in Table S1) is in very good agreement with the experimental values ($\Delta G^{s\text{-}a\neq}$=89.87 kJ·mol^{-1}). [36] Therefore, two extra water molecules were added in active sites for the original reaction (the HSO_3^--induced direct hydrolysis deamination of *cis* and *trans* 5-$MeCytN3^+$-SO_3^- isomer). Meanwhile, the effect of one extra water molecule for the original reaction have also been calculated, and the results show that the $\Delta G^{g\neq}$ of paths A and A' are to some extent affected by one additional water molecule. The relevant energies of these paths are listed in Tables S2 and S3. According to our previously reports, [29] the names of stationary structures in "water-mediated mechanisms" are further modified with the prefix 1w or 2w, which implies the participation of one or two extra water molecules. Note that the conformational changes involved activation free energies are still small and have no significant effects on the "water-mediated mechanisms". Thus, the rate-limiting steps of water-mediated paths A (Fig. 9) and A' (Fig. 10) are only discussed in the follow-up context.

As shown in Figs. 4 and 11, two "water-mediated mechanism" (2w-path A), relative to the HSO_3^--induced direct hydrolysis (path A), shows a less complex process, but the rate-limiting step still is the first step, which associated with the $\Delta G^{g\neq}$ is 15.40 kJ·mol^{-1} lower than that in path A, suggesting that 2w-path A in the gas phase is significant to the advantage of this hydrolysis process. This implies that two extra water molecules have an obviously positive catalytic effect on reducing the free energy barrier of the rate-limiting step of path A.

On the contrary, the $\Delta G^{g\neq}$ of the rate-determining step for path A' is slightly increased by the presence of two additional water molecules (Figs. 8 and 11, Table 3). The $\Delta G^{g\neq}$ of the

rate-determining step for 2w-path A' is 100.56 kJ·mol^{-1}, which amounts to an increase by 1.45 kJ·mol^{-1} relative to that for path A', indicating that two additional water molecules have almost no effect on the rate-limiting step of path A'.

As mentioned above, the $\Delta G^{g\neq}$ of paths A and A' are to some degree affected by two extra water molecules. Then, it is of great interest whether the $\Delta G^{g\neq}$ of these paths are further flattened by the contribution of the bulk water.

3.2 Stationary point structures and energetics in the aqueous phase

In modeling solvation effects, an important question arises as to whether computational results are sensitive to the presence of the bulk water. The solvation effect on the HSO_3^--induced direct hydrolytic deamination and "water-mediated mechanism" has been discussed. It is noteworthy to mention here that the conformational changes in the aqueous phase involved activation free energies are still small and have no significant effects on the mechanism of these reactions. Thus, our interest has been focused on the rate-determining steps of all paths in the follow-up context. The important bond lengths of all stationary points for all paths in the aqueous phase are also shown in Figs. S2-S9. The influence of solvation on the activation free energies can be explained by the evolution of the dipole moments for the rate-limiting steps of all paths (see Table 4).

3.2.1 Solvation effect on the HSO_3^--induced direct hydrolytic deamination

As for path B (Figs. 2 and S3), the stationary structures in the aqueous phase are similar with those in the gas phase, and for path A' (Figs. 5 and S5), both the optimized geometries of A'-RC and A'-TS1 have changed from the gas phase to the aqueous phase. Combined with Tables 1 and 2, the free energy barrier of rate-determining step for path B is slightly decreases from 100.76 to 100.42 kJ·mol^{-1}, whereas the free energy barrier of rate-determining step for path A' is slightly increases from 99.11 to 100.48 kJ·mol^{-1}. This reveals that water has no significant impact on the free energy barrier of these paths and is comparatively negligible. By this time there is almost no difference between $\Delta G^{s\text{-}a\neq}$ and $\Delta G^{s\text{-}g\neq}$.

Compared Fig. 1 with Fig. S2, only the optimized geometry in A-TS1 along path A has changed in the aqueous phase, leading to the difference in free energy barrier between in the aqueous and gas phase. Indentify with path A, the solvent effects obviously change the geometries of C'-TS3 (Figs. 7 and S7). The free energy barriers of rate-determining step for paths A and C' are

significantly reduced from 112.10, 126.24 to 102.17, 108.60 kJ·mol^{-1}, respectively, implying that water has noticeably effect on the activation free energy of these paths and is not comparatively negligible. At this point, there is a difference between $\Delta G^{s\text{-}a\neq}$ and $\Delta G^{s\text{-}g\neq}$. However, as shown in Tables 1 and 2, although there is a little difference between $\Delta G^{s\text{-}a\neq}$ and $\Delta G^{s\text{-}g\neq}$, the variational trends of both methods are the same.

As for paths C and B', the important structures for the rate-determining step are not located in the aqueous phase, and the geometries of these stationary points may or may not change from the gas phase to the aqueous phase. Consequently, the activation free energy of these paths may also be affected from the gas phase to the aqueous phase. However, as mentioned above, the variational trends of $\Delta G^{s\text{-}a\neq}$ and $\Delta G^{s\text{-}g\neq}$ of rate-determining step for these paths are the same, thus instead of $\Delta G^{s\text{-}a\neq}$, the $\Delta G^{s\text{-}g\neq}$ of rate-determining step for paths C and B' is approximately taken as a criterion of the reaction activity. As seen from Tables 1 and 2, it is obviously in aqueous solution that paths A and B (paths A' and B') are much more favorable in the HSO_3^--induced direct hydrolytic deamination of *cis* (*trans*) 5-MeCytN3$^+$-SO_3^- isomer.

For a good understanding of the effect of the solvation on the free energy barriers, the evolution of dipole moments for the rate-limiting steps are discussed. As seen from Table 4, for path A, the activation free energy ($\Delta G^{s\text{-}a\neq}$) of the rate-determining step in water is 9.93 kJ·mol^{-1} lower than that in the gas phase, attributed to that the dipole moment of A-TS1 (μ=8.7 debye) is much more than that of A-RC (μ=6.3 debye), and the transition state A-TS1 is stabilized in water by solvation. Indentify with path A, the dipole moment (μ=10.2 debye for C'-TS3) is larger than that of C'-IM2 (μ=9.2 debye), and then the solvation of water on C'-TS3 is stronger than that on C'-IM2. This illustrates that water has noticeably effect on the activation free energies of these paths. On the contrary, for path B, the dipole moment (*u*=7.7 debye for B-TS1), relative to B-RC (u=7.6 debye), has a very small change, indicating that solvation is comparatively negligible. And for path A', the dipole moment (μ=8.4 debye for A'-TS1) is less than that of A'-RC (μ=8.9 debye). As a result, the solvent water destabilizes the transition state A'-TS1, leading to the increase of free energy barrier by 1.37 kJ·mol^{-1} as compared to that in the gas phase.

As seen from Figs. 4 and 8, the $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for paths A, B, A', and C' are 102.17, 100.42, 100.48, and 108.60 kJ·mol^{-1}, respectively. It is obviously in aqueous solution that the $\Delta G^{s\text{-}a\neq}$ of hydrolysis deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer exhibits no obvious

difference with *cis* 5-MeCytN3$^+$-SO_3^- isomer in the presence of bisulfite.

3.2.2 Solvation effect on the "water-mediated mechanism"

Similarity with paths A and C', the solvation has obviously effect on the water-mediated paths A and A'. For two "water-mediated mechanism" (2w-paths A and A'), the activation free energies of the rate-limiting step are tremendously reduced from 96.70, 100.56 kJ·mol^{-1} to 85.74, 85.34 kJ·mol^{-1}. Meanwhile, the rate-limiting step of 2w-path A involved in the first step in the gas phase is changed into the second step in water, which associated with the $\Delta G^{s\text{-}a\neq}$ is 85.14 kJ·mol^{-1}, illustrating that water has noticeably effect on this path.

For 2w-paths A and A', the dipole moments (μ=16.2 debye for 2w-A-TS1 and 15.3 debye for 2w-A'-TS1) are much more than their intermediates (μ=12.3 debye for 2w-A-RC and 12.6 debye for 2w-A'-RC). As a result, the solvent water stabilizes the transition states 2w-A-TS1 and 2w-A'-TS1, leading to the decrease of free energy barriers with the range of 10.96 ~15.22 kJ·mol^{-1} as compared to those in the gas phase.

As seen from Table 3, the $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for two water-mediated paths A and A' are 85.74 and 85.34 kJ·mol^{-1}, respectively, showing that there is almost no difference in $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for the "water-mediated mechanism" between *cis* and *trans* 5-MeCytN3$^+$-SO_3^- isomer in the presence of the bulk water, suggesting both 5-MeCytN3$^+$-SO_3^- isomers may be contributed to the hydrolytic deamination of 5-MeCyt under bisulfite conditions, which is in agreement with previous experimental observation[20,21] that the deaminated products of 5-MeCyt with bisulfite were a mixture of two diastereomers.

3.3 Relevance of the results

The results of the present study suggest that both isomers may be contributed to the formation of thymine-SO_3^-. To obtain a more complete knowledge of the effect of the different 5-MeCytN3$^+$-SO_3^- isomers on the hydrolytic deamination of 5-MeCyt under bisulfite conditions, it is necessary to compare the reaction rate of the different 5-MeCytN3$^+$-SO_3^- isomers paths. Take two water-mediated paths A and A' as an example, the reaction rates of the rate-determining step of 2w-paths A and A' can be written as

$$v_{2w\text{-}A}=k_{2w\text{-}A}[2H_2O][H_2O][HSO_3^-][cis\ \text{isomer}] \quad (1)$$

$$v_{2w\text{-}A'}=k_{2w\text{-}A'}[2H_2O][H_2O][HSO_3^-][trans\ \text{isomer}] \quad (2)$$

and the relative rate is

$$\frac{v_{2w-A}}{v_{2w-A'}} = \frac{k_{2w-A}}{k_{2w-A'}} \frac{[2H_2O][H_2O][HSO_3^-][cis\ \text{isomer}]}{[2H_2O][H_2O][HSO_3^-][trans\ \text{isomer}]} = \frac{k_{2w-A}}{k_{2w-A'}} \frac{[cis\ \text{isomer}]}{[trans\ \text{isomer}]} \quad (3)$$

Thus, the different 5-MeCytN3$^+$-SO$_3^-$ isomers and the 2H$_2$O···HSO$_3^-$···H$_2$O group as reactants in the deamination reaction depend not only on the rate constants but also on the relative concentration of the two isomers. The rate constant of the rate-determining step is inferred from the equation given by the activated complex theory[37]

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

Where the free energy barriers almost have no difference for both 2w-paths (A and A'), thus $k_{2w\text{-}A}$ is approximately equal to $k_{2w\text{-}A'}$ and the equation (3) is changed as

$$\frac{v_{2w-A}}{v_{2w-A'}} = \frac{[cis\ \text{isomer}]}{[trans\ \text{isomer}]} \quad (5)$$

Our calculations predict that the relative rate is the ratio of the concentration of the different 5-MeCytN3$^+$-SO$_3^-$ isomers. When the ratio of the concentration of both isomers is 1, both isomers make an equal contribution to the hydrolytic deamination of 5-MeCyt under bisulfite conditions.

4. Conclusions

Two distinct groups of the direct hydrolytic deamination mechanisms of 5-MeCytN3$^+$-SO$_3^-$ have been explored in the presence of the HSO$_3^-$ group, *cis* 5-MeCytN3$^+$-SO$_3^-$ isomer reaction (paths A-C), *trans* 5-MeCytN3$^+$-SO$_3^-$ isomer reaction (paths A'-C'), respectively. Meanwhile, paths A and A' are represented and have been further investigated in the participation of two extra water in this work.

(1) The $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for the direct hydrolytic deamination paths A, B, A', and C' are 102.17, 100.42, 100.48, and 108.60 kJ·mol^{-1}, respectively. It is clear that the $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step exhibits no obvious difference in the direct hydrolytic deamination paths between *cis* and *trans* 5-MeCytN3$^+$-SO$_3^-$ isomers, suggesting that both isomers may make an equal contribution to the hydrolytic deamination of 5-MeCyt in the participation of bisulfite. Meanwhile, the results also show that the real impact of the reactions only depends on the relative concentrations of the both isomers.

(2) The $\Delta G^{s\text{-}a\neq}$ of the rate-limiting step for two water-mediated paths A and A' are 85.74 and

85.34 kJ·mol^{-1}, respectively. It is obviously that the participation of two water molecules obviously reduce the $\Delta G^{s\text{-}a\neq}$ of the rate-limiting for paths A and A'. For two water-mediated paths A and A', the $\Delta G^{s\text{-}a\neq}$ (85.34~85.74 kJ·mol^{-1}) of the rate-limiting step for *cis* and *trans* 5-MeCytN3^{+}-SO$_3^{-}$ isomers paths are in very close proximity to that of the theoretical date for CytN3^{+}-SO$_3^{-}$ (88.18 kJ·mol^{-1}), implying that the difference of Cyt and 5-MeCyt is very small from the view of the free energy barrier. This further suggests that the bisulfite sequencing technology may have strong dependency on the environmental factors.

5. Biochemical implications

As seen from Tables. 1-3, for 5-MeCyt, there is little difference in reaction mechanisms and activation free energies under bisulfite conditions compared with theoretical data of Cyt, making the bisulfite genomic sequencing be easily influenced by external environment, which is not benefit for the reliability and selectivity of this method. Thus, to improve the reliability and selectivity of this method, it is crucial to find the appropriate solutions for widening of the free energy barrier gap between Cyt and 5-MeCyt in the presence of bisulfite. To achieve this, some measures should be taken in aspect of the internal and external factors. According to the experimental results,[38,39] the saturation of the C5-C6 double bond is critical for efficient deamination, thus it is firstly important to find the suitable substituent for the C5-C6 bond to widen the free energy barrier gap between Cyt and 5-MeCyt. In addition, as it is well known, the appropriate solvent may enhance the selectivity of the method, thus selection of appropriate solvent for DNA sample is also an important factor. Finally, the adjustment of other parameters, such as the bisulfite concentration, pH value, and temperature, may be able to improve the reliability of the bisulfite sequencing technology. This work not only may be an important mechanism for revealing the inappropriate-conversion event but by extension is also likely to be a significant basis for improving the efficiency and reliability of the bisulfite genomic sequencing.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No:21173139), National Training Programs of Innovation and Entrepreneurship for Undergraduates (No:201210781070), Shaanxi Innovative Team of Key Science and Technology (2013KCT-17),

and the Fundamental Research Funds for the Central Universities (GK: 201101004).

References

(1) R. Sto¨ger, P. Kubicka, C. G. Liu, T. Kafri, A. Razin, H. Cedar and D. P. Barlow, *Cell*, 1993, **73**, 61.
(2) J. L. Swain, T. A. Stewart, P. Leder, *Cell*, 1987, **50**, 719.
(3) A. P. Feinberg and B.Vogelstein, *Semin. Surg. Oncol*., 1987, **3**, 149.
(4) C. D. Laird, E. Jaffe, G. Karpen, M. Lamb and R. Nelson, *Trends Genet*., 1987, **3**, 274.
(5) S. E. Jacobsen and E. M.Meyerowitz, *Science*, 1997, **277**, 1100.
(6) P. Cubas, C. Vincent and E. Coen, *Nature*, 1999, **401**, 157.
(7) M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy and C. L. Paul, *Proc. Natl. Acad. Sci. USA*., 1992, **89**, 1827.
(8) S. J. Clark, J. Harrison and M. Frommer, *Nat. Genet*.,1995, **10**, 20.
(9) R. Stöger, T. M. Kajimura, W. T. Brown and C. D. Laird, *Hum. Mol. Genet*., 1997, **6**, 1791.
(10) D. P. Genereux, B. E. Miner, C. T. Bergstrom and C. D. Laird, *Proc. Natl Acad. Sci. USA*., 2005, **102**, 5802.
(11) S. J. Clark, J. Harrison, C. L. Paul and M. Frommer, *Nucleic Acids Res*. 1994, **22**, 2990.
(12) D. P. Genereux, W. C. Johnson, A. F. Burden, R. Stöger and C. D. Laird, *Nucleic Acids Res*., 2008**, 36**, e150.
(13) M. Shiragami, S. Iida, I. Kudo and H. Hayatsu, *Chem Pharm Bull*., 1975, **23**, 3027.
(14) L. X. Jin, W. L. Wang, D. D Hu and S. T. Min, *J. Phys. Chem. B*., 2013, **117**, 3.
(15) R. Shapiro, V. DiFate and M. Welcher, *J. Am. Chem. Soc.*, 1974, **96**, 906.
(16) B. Wang and Z. Cao, *J. Phys. Chem. A*, 2010, **114**, 12918.
(17) Z. q. Chen and Y. Xue, *J. Phys. Chem. B*, 2010, **114**, 12641.
(18) V. Labet, C. Morell, J. Cadet, L. A. Eriksson and A. Grand, *J. Phys. Chem. A* 2009**, 113**, 2524.
(19) F. C. Meng and H. Wang, J. *Theor Chem Acc.*, 2010**, 127**, 561.
(20) H. Hayatsu, *Proc. Jpn. Acad. Ser. B.*, 2008, **84**, 321.
(21) H. Hayatsu, *Genes and Environment.*, 2006, **28**, 1.
(22) L. X. Jin, W. L. Wang, D. D. Hu and J. Lü, *Phys. Chem. Chem. Phys*. 2013, **15**, 9034.
(23) D. Wei, B. L. Lei, M. S.Tang and C. G. Zhan. *J. Am. Chem. Soc*., 2012, **134**, 10436.
(24) (a) Z. Q. Chen, C. H. Zhang, C. K. Kimand and Xue, Y. *Phys. Chem. Chem. Phys*., 2011, **13**, 6471.(b) V. Labet, A. Grand, C. Morell, J. Cadet and L. A. Eriksson, *Theor Chem Account*., 2008, **120**, 429.
(25) K. M. Uddin and R. A. Poirier, *J. Phys. Chem. B*., 2011, **115**, 9151.
(26) K. M. Uddin, M. H. Almatarneh, D. M. Shaw and R. A. Poirier, *J. Phys. Chem. A*, 2011, **115,** 2065.
(27) (a) M. H. Almatarneh, C. G. Flinn and R. A. Poirier, *J. Chem. Inf. Model.*, 2008, **48,** 831. (b) V. Labet, C. Morell, T. Douki, J. Cadet, L. A. Eriksson and A. Grand, *J. Phys. Chem. A*, 2010, **114**, 1826.
(28) J. L. Przybylski and S. D. Wetmore, *J. Phys. Chem. B*, 2009, **113**, 6533.
(29) L. X. Jin, W. L. Wang, D. D. Hu and J. Lü, *Phys. Chem. Chem. Phys*., 2014, **16**, 3573.
(30) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. J. Zheng, L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A.Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O.

Farkas, J. B. Foresman, J.V. Ortiz, J. Cioslowski,; D. J. Fox, Gaussian 09, Revision A.02. Gaussian, Inc., Wallingford, CT, 2009.

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

(32) A. D. McLean and G. S. Chandler, *J Chem Phys*., 198**0**, **72**, 5639.

(33) E. M. B. Cancès and J. Tomasi, *J. Chem. Phys*., 1997, **107**, 3032.

(34) W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257.

(35) M. S. Gordon, *Chem. Phys. Lett.*, 1980, **76**, 163.

(36) M. Sono, Y.Wataya and H. Hayatsu, *J. Am. Chem. Soc.,* 1973, **95**, 4745.

(37) H. Eyring, *Chem. Rev*., 1935, **17**, 65.

(38) B. Duncan and J. Miller, *Nature*, 1980, **287**, 560.

(39) M. Wyszynski, S. Gabbara, S. Ashok and A. Bhagwat, *Proc. Natl.Acad. Sci. U.S.A.*, 1994, **91**, 1574.

Table content:

Table 1 Relative Energies [a] (kJ·mol^{-1}) for the Direct Hydrolytic Deamination of *Cis* 5-MeCytN3^{+}-SO_3^{-} Isomer with the HSO_3^{-} Group (paths A-C) in the Gas and Aqueous Phases

System	MP2//B3LYP [b]			PCM[c]			
	ΔE^{g}	ΔG^{g}	$\Delta G^{g\neq}$	$\Delta G^{s\text{-}g}$	$\Delta G^{s\text{-}g\neq}$	$\Delta G^{s\text{-}a}$	$\Delta G^{s\text{-}a\neq}$
Path A							
R[d]	0.00	0.00		0.00		0.00	
A-RC	-286.33	-211.09		25.75		-5.47	
A-TS1	-193.11	-98.99		131.21		96.70	
A-IM1	-238.18	-145.20		74.80		35.94	
A-TS2	-232.94	-137.95		84.87		45.27	
A-IM2	-234.75	-141.21		81.85		40.93	
A-TS3	-240.03	-147.02		71.69		--	
A-IM3	-236.59	-146.11		62.03		17.62	
A-IM4	-284.57	-193.09		47.50		-2.18	
A-TS4	-229.55	-139.07		108.15		67.20	
A-P	-269.26	-187.22		58.42		13.75	
A-RC→A-IM1			112.10		105.46		102.17
A-IM1→A-IM2			7.25		10.07		9.33
A-IM2→A-IM3			-5.81		-10.16		--
A-IM4→A-P			54.02		60.65		69.38
Path B							
R	0.00	0.00		0.00		0.00	
B-RC	-263.67	-182.74		48.89		3.26	
B-TS1	-172.77	-81.98		144.25		103.68	
B-IM1	-236.29	-146.11		81.02		40.89	
B-TS2	-222.35	-130.10		85.24		42.62	
B-IM2	-230.86	-139.90		85.30		40.69	
B-TS3	-226.25	-132.40		94.77		52.49	
B-IM3	-225.33	-133.10		93.32		53.37	
B-TS4	-224.91	-131.35		90.19		46.23	
B-IM4	-227.66	-136.21		78.00		21.01	
B-IM5	-288.13	-196.97		45.20		0.88	
B-TS5	-270.59	-171.69		75.38		40.00	
B-P	-263.20	-167.58		79.65		44.88	
B-RC→B-IM1			100.76		95.36		100.42
B-IM1→B-IM2			16.01		4.22		1.73
B-IM2→B-IM3			7.50		9.47		11.80
B-IM3→B-IM4			1.75		-3.13		-7.14
B-IM4→B-P			25.28		30.18		39.12
Path C							
R	0.00	0.00		0.00		0.00	
C-RC	-279.79	-197.76		36.92		-9.34	
C-TS1	-189.59	-99.09		129.90		90.79	
C-IM1	-209.20	-123.56		99.09		62.14	
C-TS2	-203.65	-115.96		110.32		70.80	
C-IM2	-215.17	-126.02		98.54		--	

C-TS3	-215.02	-122.48	103.31	--	
C-IM3	-215.17	-126.01	98.56	63.93	
C-TS4	-93.08	-1.45	215.84	--	
C-IM4	-223.19	-130.89	73.26	29.73	
C-IM5	-304.74	-206.17	29.45	7.67	
C-TS5	-254.60	-154.58	91.61	55.35	
C-P	-319.89	-233.26	8.31	-35.62	
C-RC→C-IM1			98.67	92.98	100.13
C-IM1→C-IM2			7.60	11.23	8.66
C-IM2→C-IM3			3.54	4.77	--
C-IM3→C-IM4			124.56	117.28	--
C-IM5→C-P			51.59	62.16	47.68

[a] ΔE^{g}, ΔG^{g}, and $\Delta G^{g\neq}$ are the relative energy, relative free energy, and activation free energy in the gas phase, respectively.; $\Delta G^{s\text{-}g}$ and $\Delta G^{s\text{-}g\neq}$ are the relative free energy and activation free energy with PCM model based on the optimized geometries in the gas phase respectively.; $\Delta G^{s\text{-}a}$ and $\Delta G^{s\text{-}a\neq}$ are the relative free energy and activation free energy with PCM model based on the optimized geometries in the aqueous phase, respectively. [b] MP2/6-311++G(3df,3pd)//B3LYP/ 6-311++G(d,p) level. [c] MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) with PCM model. [d] denotes *cis* 5-MeCytN3^{+}-SO_3^{-} isomer +H_2O+HSO_3^{-}.

Table 2 Relative Energies [a] (kJ·mol^{-1}) for the Direct Hydrolytic Deamination of *Trans* 5-MeCytN3^{+}-SO$_3^{-}$ Isomer with the HSO$_3^{-}$ Group (Paths A'-C') in the Gas and Aqueous Phases

System	MP2//B3LYP [b]			PCM[c]			
	ΔE^{g}	ΔG^{g}	$\Delta G^{g\neq}$	$\Delta G^{s\text{-}g}$	$\Delta G^{s\text{-}g\neq}$	$\Delta G^{s\text{-}a}$	$\Delta G^{s\text{-}a\neq}$
path A'							
R'[d]	0.00	0.00		0.00		0.00	
A'-RC	-280.72	-197.19		8.42		-0.78	
A'-TS1	-193.17	-98.08		115.84		99.70	
A'-IM1	-243.13	-146.84		60.66		38.59	
A'-TS2	-235.79	-138.93		67.05		51.42	
A'-IM2	-236.61	-140.93		65.11		44.94	
A'-TS3	-242.19	-148.33		74.87		33.71	
A'-IM3	-238.70	-144.89		41.30		25.79	
A'-IM4	-293.98	-198.64		51.77		9.15	
A'-TS4	-230.59	-139.54		117.14		73.09	
A'-P	-286.03	-202.90		47.79		6.19	
A'-RC→A'-IM1			99.11		107.42		100.48
A'-IM1→A'-IM2			7.91		6.39		12.83
A'-IM2→A'-IM3			-7.40		9.76		-11.23
A'-IM4→A'-P			59.10		65.37		63.94
Path B'							
R'	0.00	0.00		0.00		0.00	
B'-RC	-262.65	-180.18		55.86		-0.39	
B'-TS1	-174.03	-84.51		148.96		--	
B'-IM1	-239.60	-146.83		91.03		48.42	
B'-TS2	-224.69	-129.24		95.97		--	
B'-IM2	-234.04	-140.78		93.83		40.35	
B'-TS3	-229.01	-130.46		105.97		61.44	
B'-IM3	-228.06	-134.20		101.43		59.42	
B'-TS4	-226.70	-132.18		96.04		--	
B'-IM4	-227.52	-132.90		86.97		41.92	
B'-IM5	-306.82	-209.83		39.42		-0.47	
B'-TS5	-275.30	-173.91		81.52		19.99	
B'-P	-268.95	-171.20		83.70		49.61	
B'-RC→B'-IM1			95.67		93.10		--
B'-IM1→B'-IM2			17.59		4.94		--
B'-IM2→B'-IM3			10.32		12.14		21.09
B'-IM3→B'-IM4			2.02		-5.39		--
B'-IM4→B'-P			35.92		42.10		20.46
Path C'							
R'	0.00	0.00		0.00		0.00	
C'-RC	-276.82	-184.46		30.17		0.89	
C'-TS1	-205.78	-107.25		107.17		79.96	
C'-IM1	-257.18	-162.58		62.37		40.29	
C'-TS2	-254.90	-159.32		67.89		47.44	
C'-IM2	-266.20	-170.14		53.11		33.68	
C'-TS3	-144.31	-43.90		170.88		142.28	

C'-P	-299.48	-215.20		26.82	0.72	
C'-RC→C'-IM1			77.21	77.00		79.07
C'-IM1→C'-IM2			3.26	5.52		7.15
C'-IM2→C'-P			126.24	117.77		108.60

[a] ΔE^{g}, ΔG^{g}, and $\Delta G^{g\neq}$ are the relative energy, relative free energy, and activation free energy in the gas phase, respectively.; $\Delta G^{s\text{-}g}$ and $\Delta G^{s\text{-}g\neq}$ are the relative free energy and activation free energy with PCM model based on the optimized geometries in the gas phase respectively.; $\Delta G^{s\text{-}a}$ and $\Delta G^{s\text{-}a\neq}$ are the relative free energy and activation free energy with PCM model based on the optimized geometries in the aqueous phase, respectively. [b] MP2/6-311++G(3df,3pd)//B3LYP/ 6-311++G(d,p) level. [c] MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) with PCM model. [d] denotes *trans* 5-MeCytN3^{+}-SO_3^{-} isomer +H_2O+HSO_3^{-}.

Table 3 Relative Energies [a] (in kJ·mol^{-1}) for Two Water-mediated Paths A (2w-path A) and A' (2w-path A') Both in the Gas and Aqueous Phases

System	MP2//B3LYP [b]			PCM[c]	
	ΔE^{g}	ΔG^{g}	$\Delta G^{g\neq}$	$\Delta G^{s\text{-}a}$	$\Delta G^{s\text{-}a\neq}$
2w-path A					
R[d]+2H_2O	0.00	0.00		0.00	
2w-A-RC	-349.70	-187.35		26.86	
2w-A-TS1	-258.18	-90.65		112.00	
2w-A-IM1	-258.66	-94.19		95.11	
2w-A-TS2	-186.02	-17.54		180.85	
2w-A-IM2	-318.47	-151.48		52.10	
2w-A-TS3	-329.39	-157.26		92.84	
2w-A-P	-370.16	-212.27		47.60	
2w-A-RC→2w-A-IM1			96.70		85.14
2w-A-IM1→2w-A-IM2			76.65		85.74
2w-A-IM1→2w-A-P			-5.78		40.74
2w-path A'					
R[e]+2H_2O	0.00	0.00		0.00	
2w-A'-RC	-351.27	-189.32		26.84	
2w-A'-TS1	-260.63	-88.76		112.18	
2w-A'-IM1	-262.55	-96.04		98.79	
2w-A'-TS2	-192.40	-22.89		181.74	
2w-A'-IM2	-319.11	-150.74		57.49	
2w-A'-TS3	-331.92	-160.53		88.33	
2w-A'-P	-353.36	-201.52		56.36	
2w-A'-RC→2w-A'-IM1			100.56		85.34
2w-A'-IM1→2w-A'-IM2			73.15		82.95
2w-A'-IM1→2w-A'-P			-9.79		30.84

[a] ΔE^{g}, ΔG^{g}, and $\Delta G^{g\neq}$ are the relative energy, relative free energy, and activation free energy in the gas phase, respectively.; $\Delta G^{s\text{-}a}$ and $\Delta G^{s\text{-}a\neq}$ are the relative free energy and activation free energy with PCM model based on the optimized geometries in the aqueous phase, respectively. [b] MP2/6-311++G(3df,3pd)//B3LYP/ 6-311++G(d,p) level. [c] MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) with PCM model. [d] denotes *cis* 5-MeCytN3^{+}-SO_3^{-} isomer +H_2O+HSO_3^{-}. [e] denotes *trans* 5-MeCytN3^{+}-SO_3^{-} isomer +H_2O+HSO_3^{-}.

Table 4 The Evolution of the Dipole Moments (μ, in Debye) of the Rate-limiting Step for the HSO_3^--induced Direct Hydrolytic Deamination of *Cis* (Paths A-C) and *Trans* 5-MeCytN3$^+$-SO_3^- Isomers (Paths A′-C′) and Two Water-mediated Paths A (2w-path A) and A' (2w-path A')

HSO_3^--induced direct hydrolytic deamination of *cis* 5-MeCytN3$^+$-SO_3^- isomers					
path A	μ	path B	μ	path C	μ
A-RC	6.3	B-RC	7.6	C-IM3	11.8
A-TS1	8.7	B-TS1	7.7	C-TS4	--
HSO_3^--induced direct hydrolytic deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomers					
path A'	μ	path B'	μ	path C'	μ
A'-RC	8.9	B'-RC	7.9	C'-IM2	9.2
A'-TS1	8.4	B'-TS1	--	C'-TS3	10.2
Two water-mediated hydrolytic deamination					
2w-Path A	μ	2w-Path A	μ	2w-Path A'	μ
2w-A-RC	12.3	2w-A-IM1	17.7	2w-A'-RC	12.6
2w-A-TS1	16.2	2w-A-TS2	12.8	2w-A'-TS1	15.3

Figure Captions:

Fig. 1 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *cis* 5-MeCytN3$^+$-SO_3^- isomer (path A-the HSO_3^- group toward the left side of *cis* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 2 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *cis* 5-MeCytN3$^+$-SO_3^- isomer (path B-the HSO_3^- group toward the above side of *cis* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 3 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *cis* 5-MeCytN3$^+$-SO_3^- isomer (path C-the HSO_3^- group toward the right side of *cis* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 4 The potential energy surfaces (ΔG in kJ·mol^{-1}) along *cis* 5-MeCytN3$^+$-SO_3^- isomer paths A (b), B (c), and C (d) both in the gas (ΔG^g in black line) and aqueous phases ($\Delta G^{s\text{-}a}$ in red line) are at MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) level of theory. R denotes *cis* 5-MeCytN3$^+$-SO_3^- isomer +H_2O+HSO_3^-.

Fig. 5 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer (path A'-the HSO_3^- group toward the left side of *trans* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 6 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer (path B'-the HSO_3^- group toward the above side of *trans* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 7 Optimized structures (bond distances in Å) in the gas phase for the HSO_3^- -induced direct hydrolytic deamination of *trans* 5-MeCytN3$^+$-SO_3^- isomer (path C'-the HSO_3^- group toward the right side of *trans* 5-MeCytN3$^+$-SO_3^- isomer) at the B3LYP/6-311++G(d, p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 8 The potential energy surfaces (ΔG in kJ·mol^{-1}) along *trans* 5-MeCytN3$^+$-SO_3^- isomer paths A' (a'), B' (b'), and C' (c') both in the gas (ΔG^g in black line) and aqueous phases ($\Delta G^{s\text{-}a}$ in red line) are at MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) level of theory. R' denotes *trans* 5-MeCytN3$^+$-SO_3^- isomer +H_2O+HSO_3^-.

Fig. 9 Optimized stationary structures (bond distances in Å) in the gas phase for two water-mediated path A (2w-path A) are at B3LYP/6-311++G(d,p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 10 Optimized stationary structures (bond distances in Å) in the gas phase for two water-mediated path A' (2w-path A') are at B3LYP/6-311++G(d,p) level. Both the bonded and H-bond species in chemical structural formulas are denoted in solid and dotted lines, respectively. 5-MeCytN3$^+$-SO_3^- and Thymine-SO_3^- in chemical structural formulas are marked M and T, respectively.

Fig. 11 The potential energy surfaces (ΔG in kJ·mol^{-1}) along two water-mediated paths A(a) and A' (b) both in the gas (ΔG^g in black line) and aqueous phases ($\Delta G^{s\text{-}a}$ in red line) are at

MP2/6-311++G(3df,3pd)//B3LYP/6-311++G(d,p) level. R denotes *cis* 5-MeCytN3^{+}-SO_3^{-} isomer + H_2O + HSO_3^{-}; R' denotes *trans* 5-MeCytN3^{+}-SO_3^{-} isomer + H_2O + HSO_3^{-}.

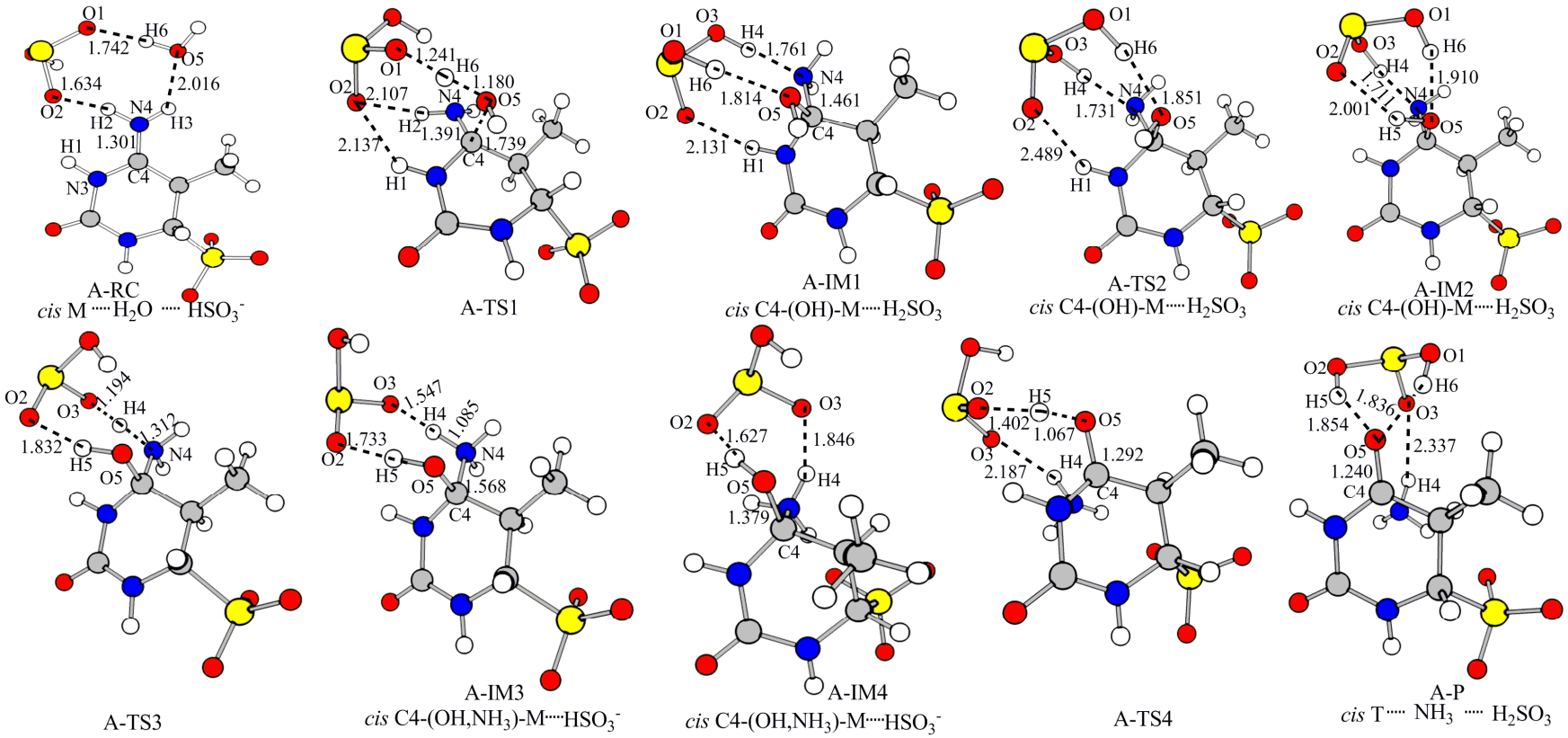


Fig. 1

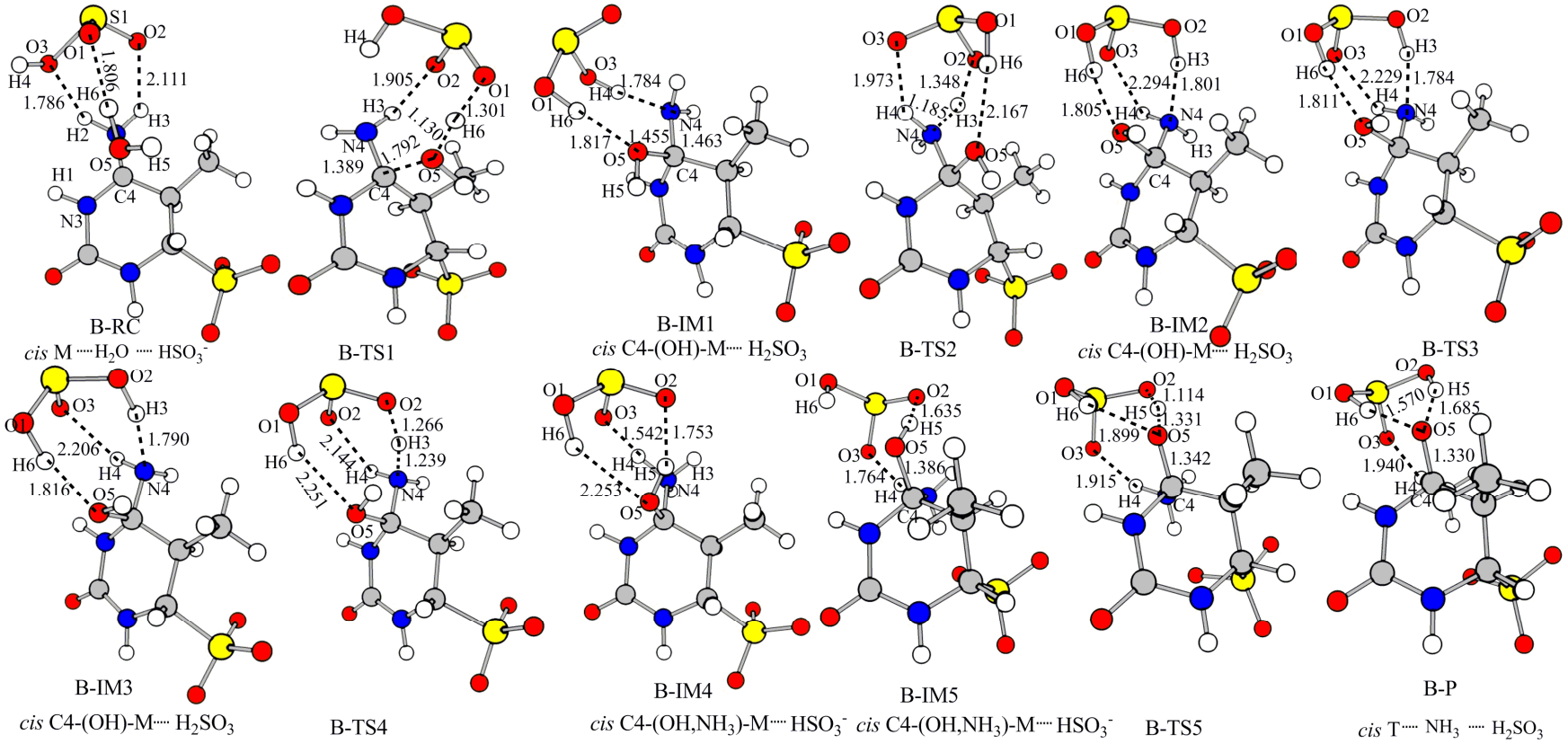


Fig. 2

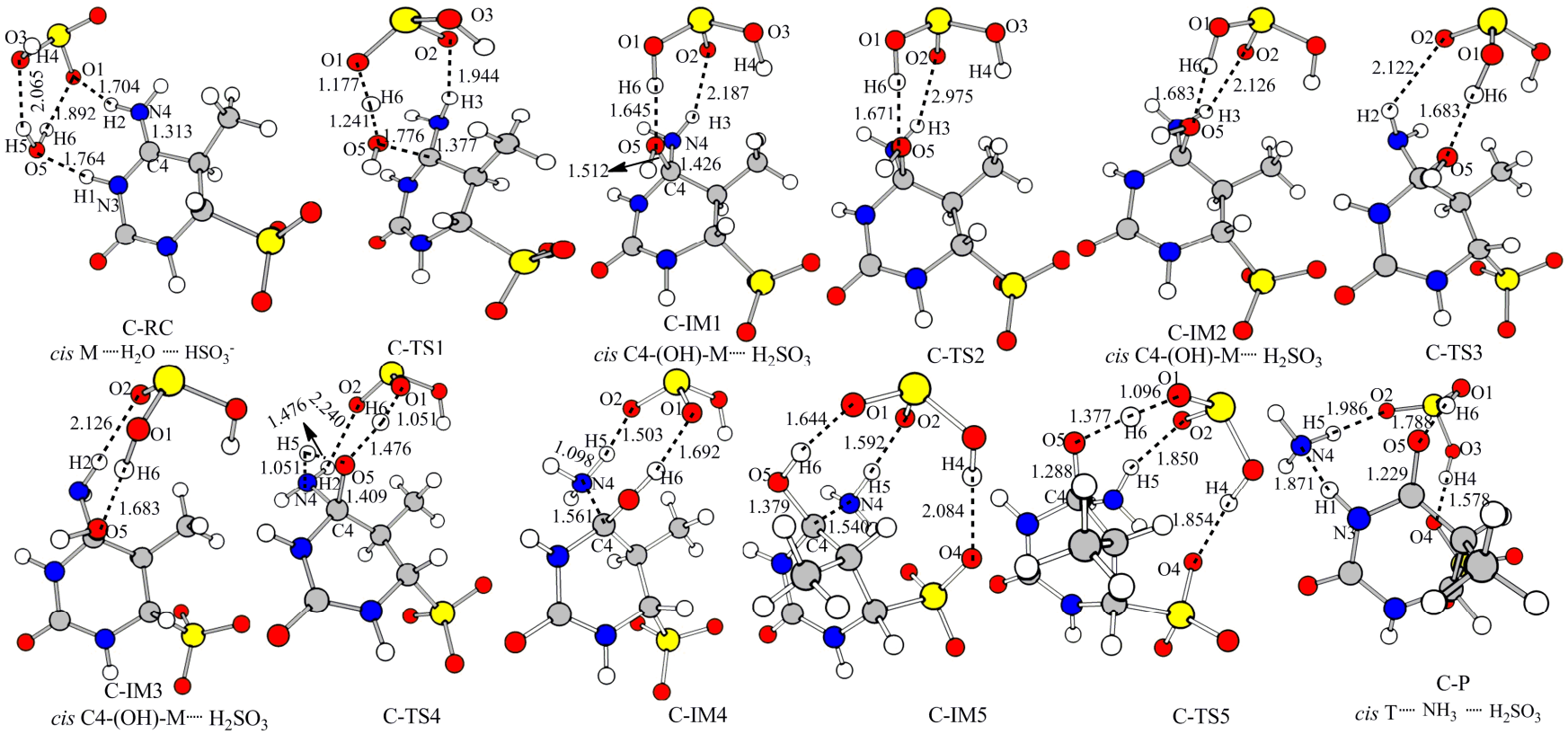


Fig. 3

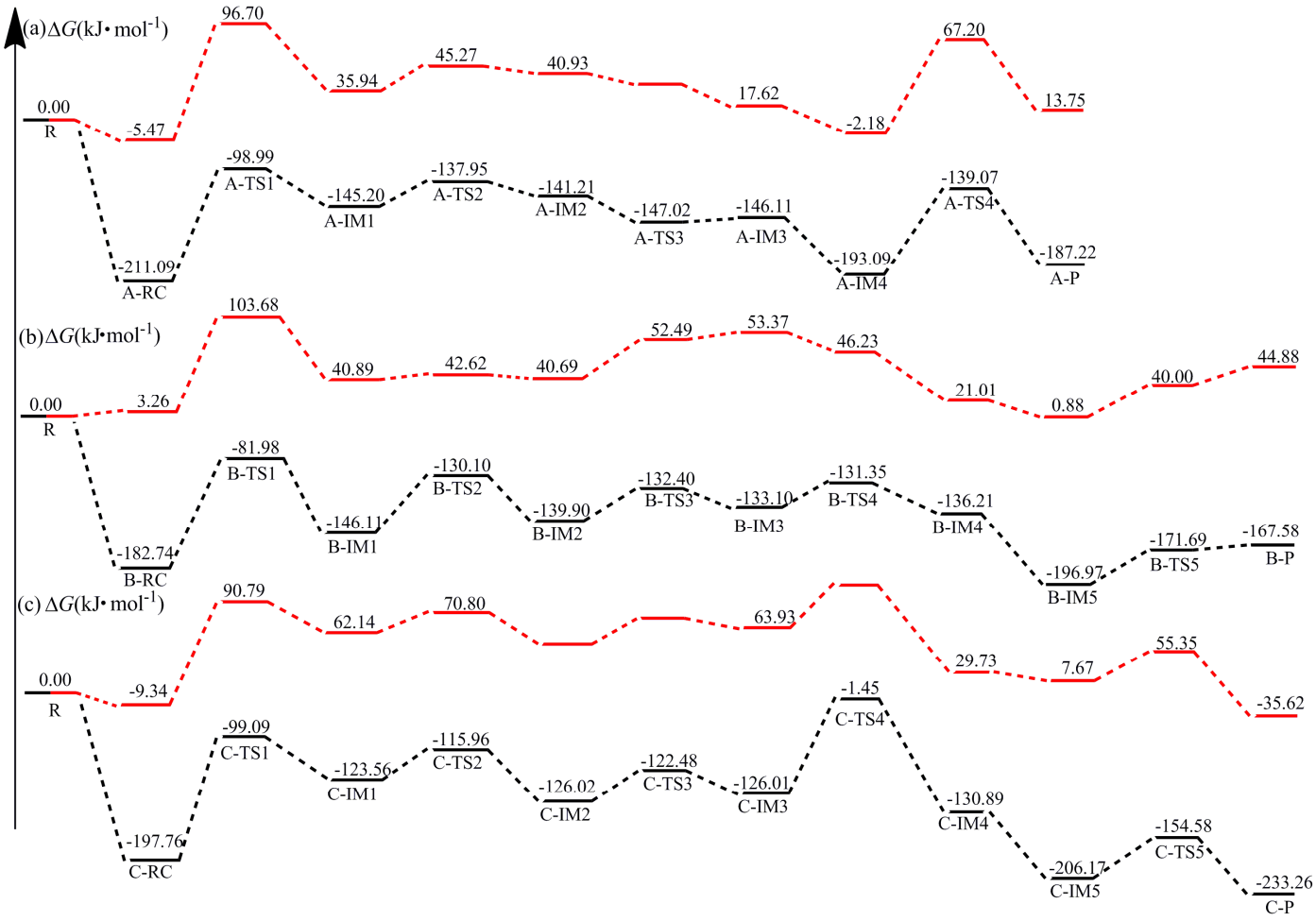


Fig. 4

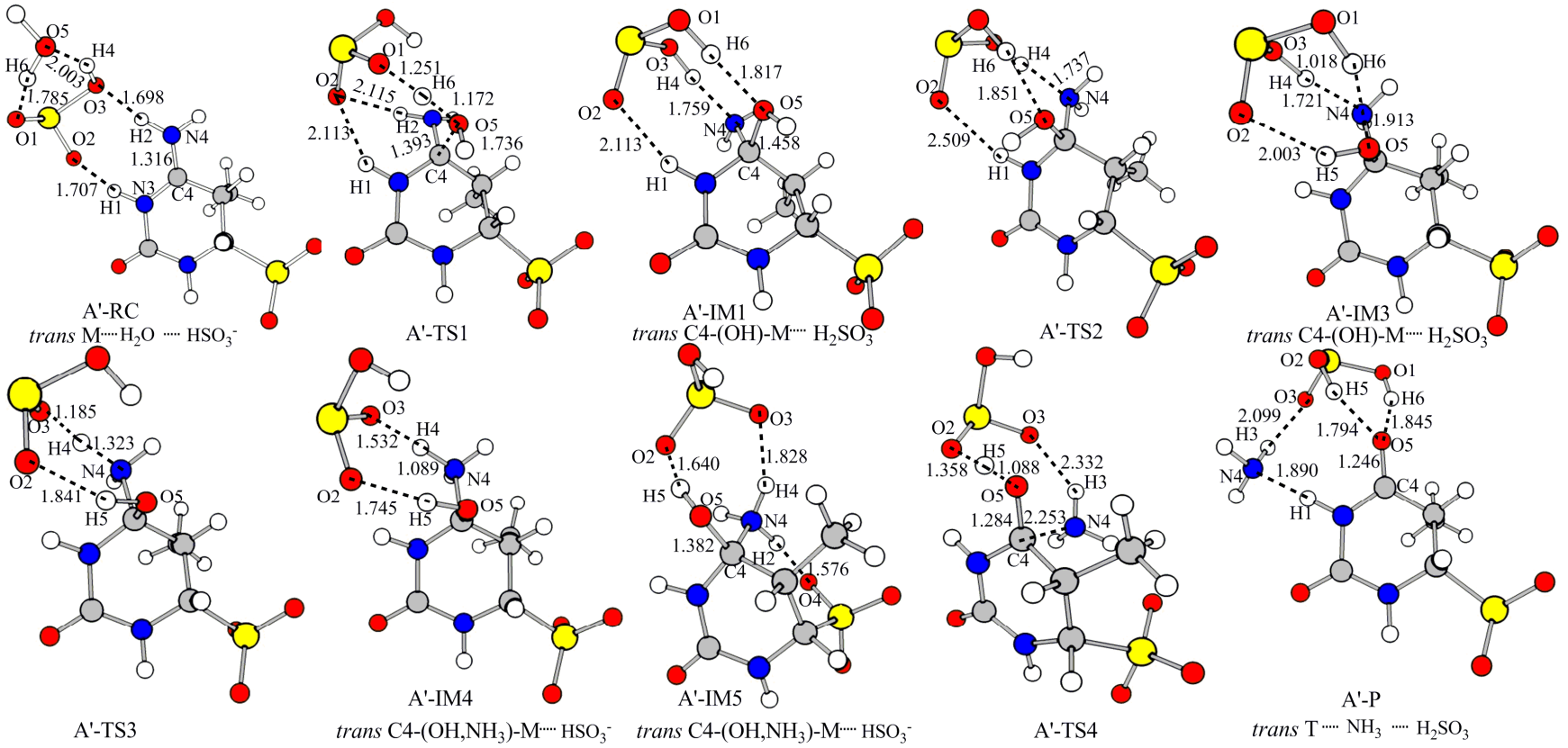


Fig. 5

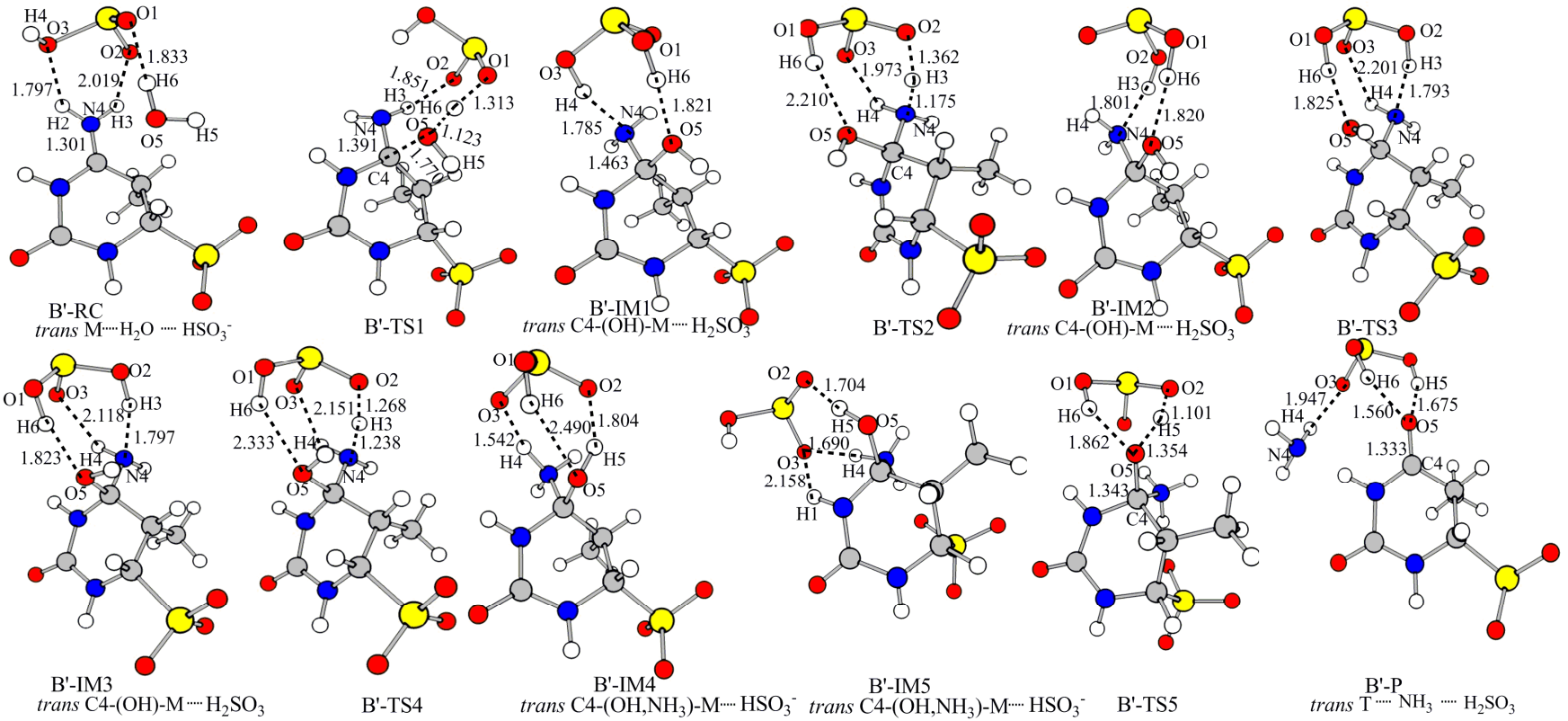


Fig. 6

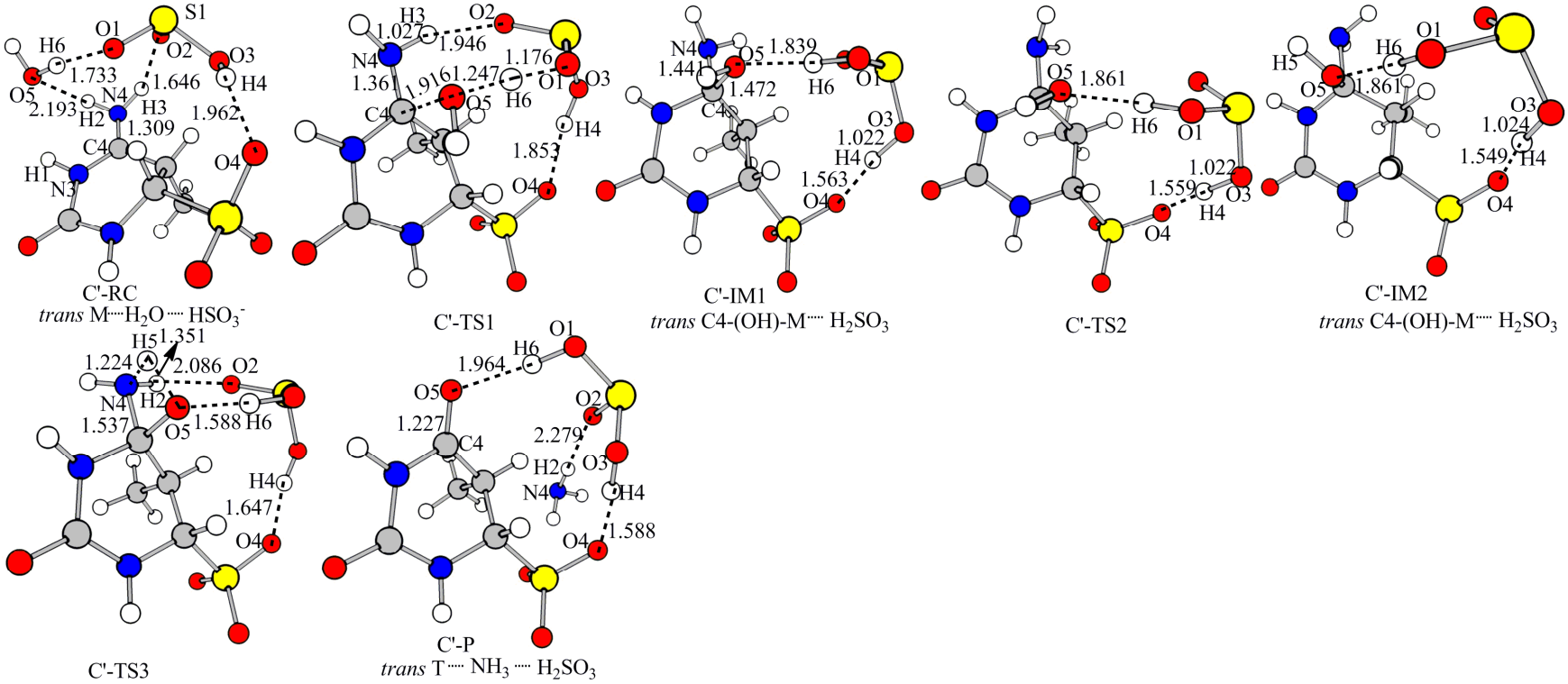


Fig. 7

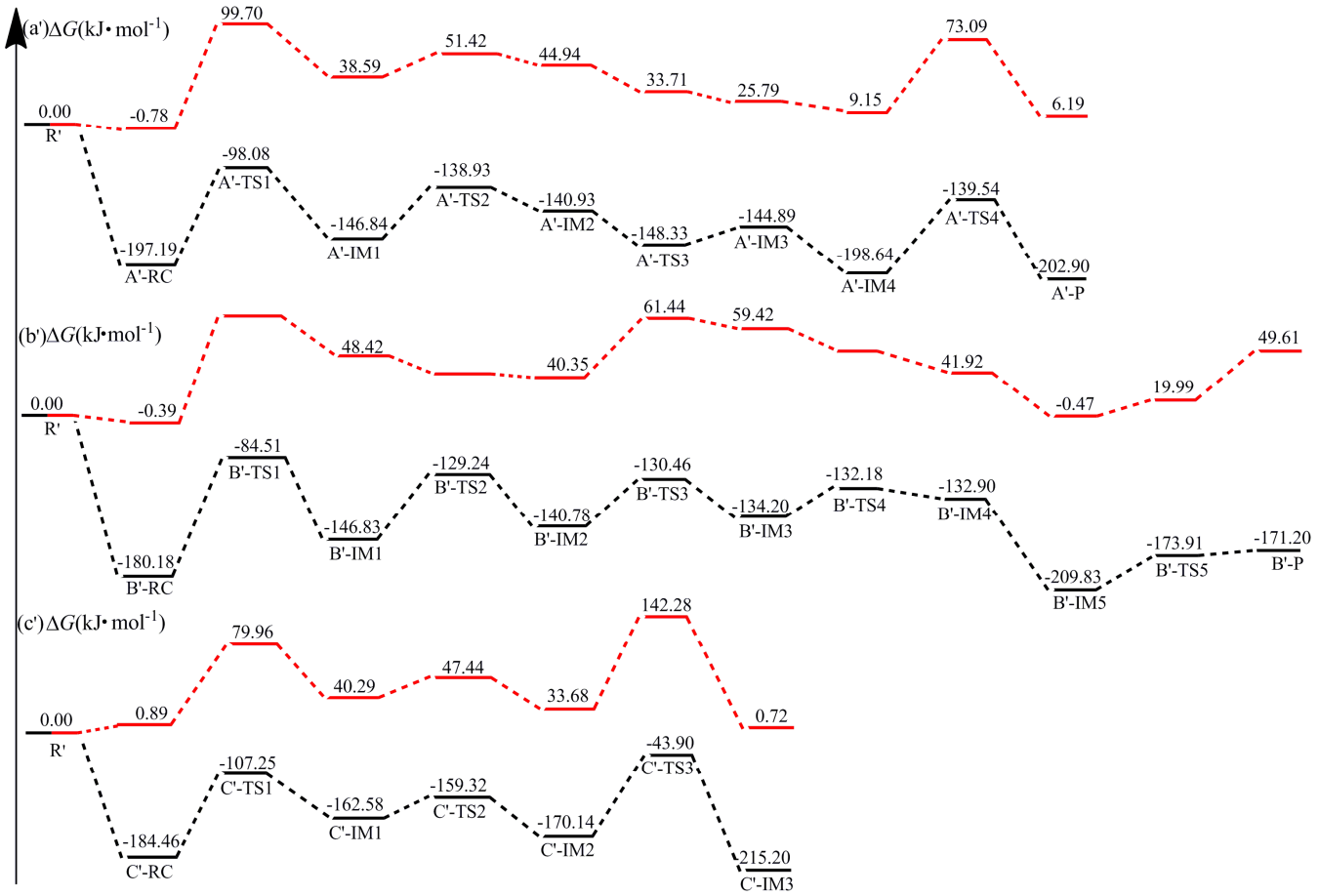


Fig. 8

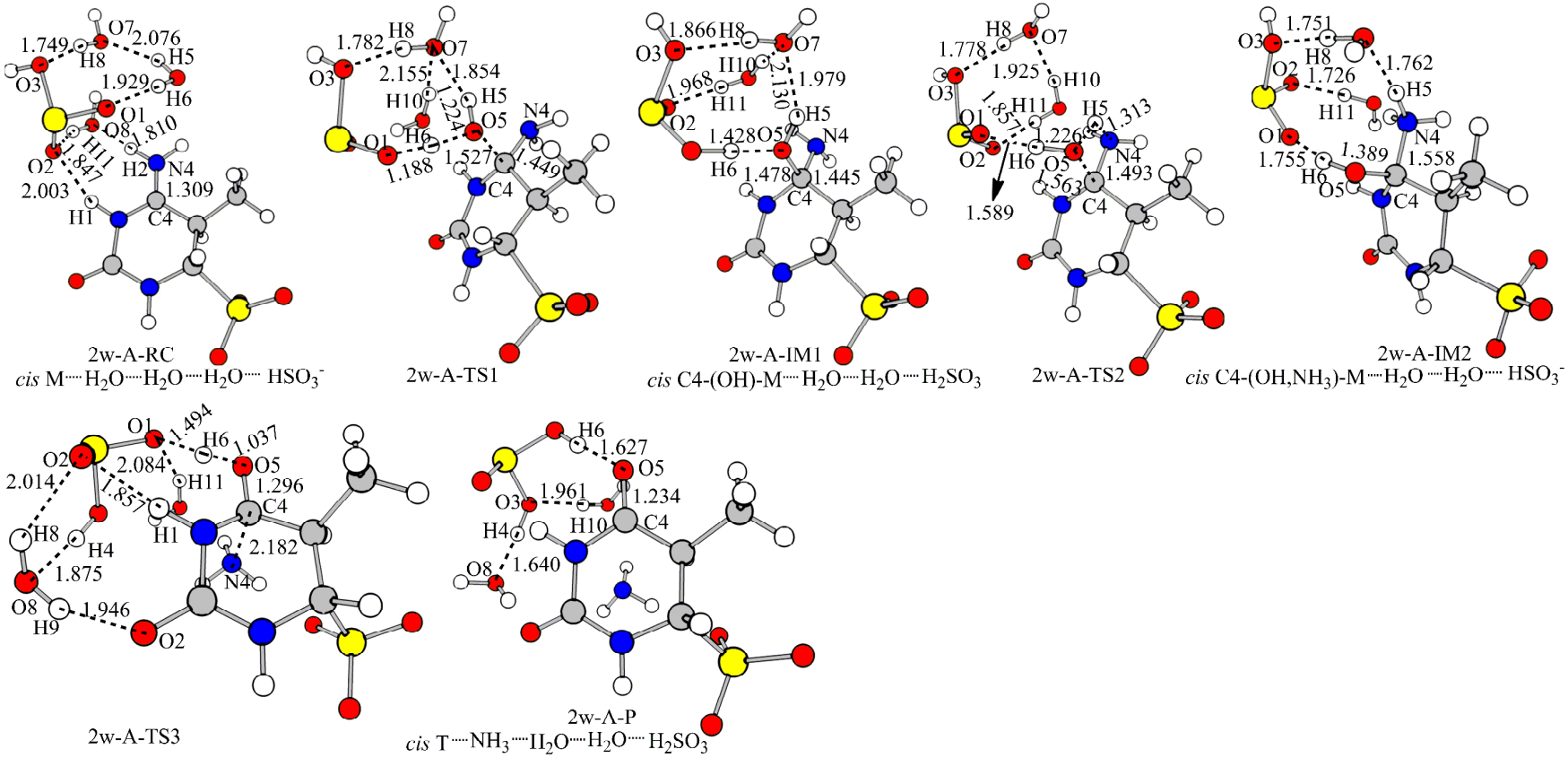


Fig. 9

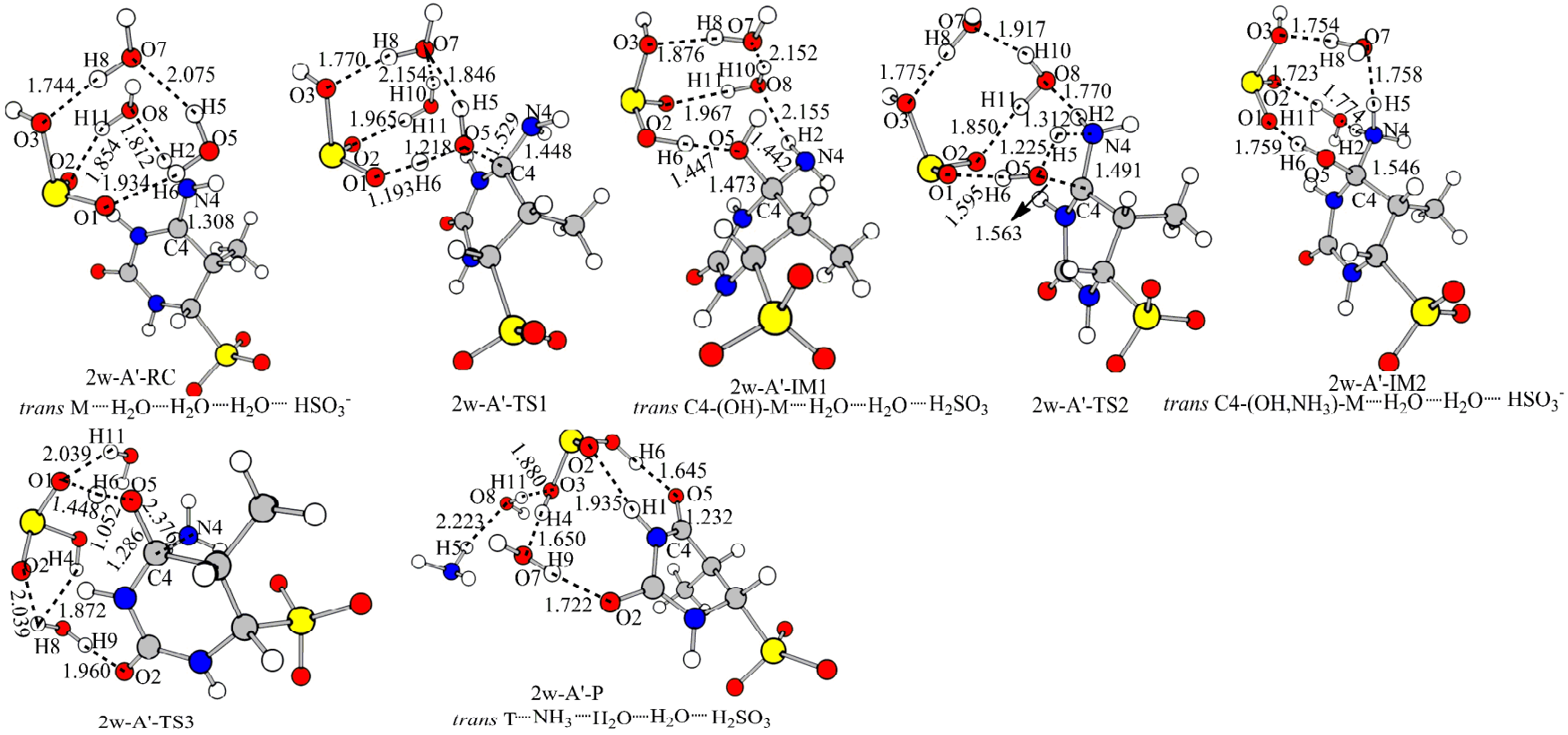


Fig. 10

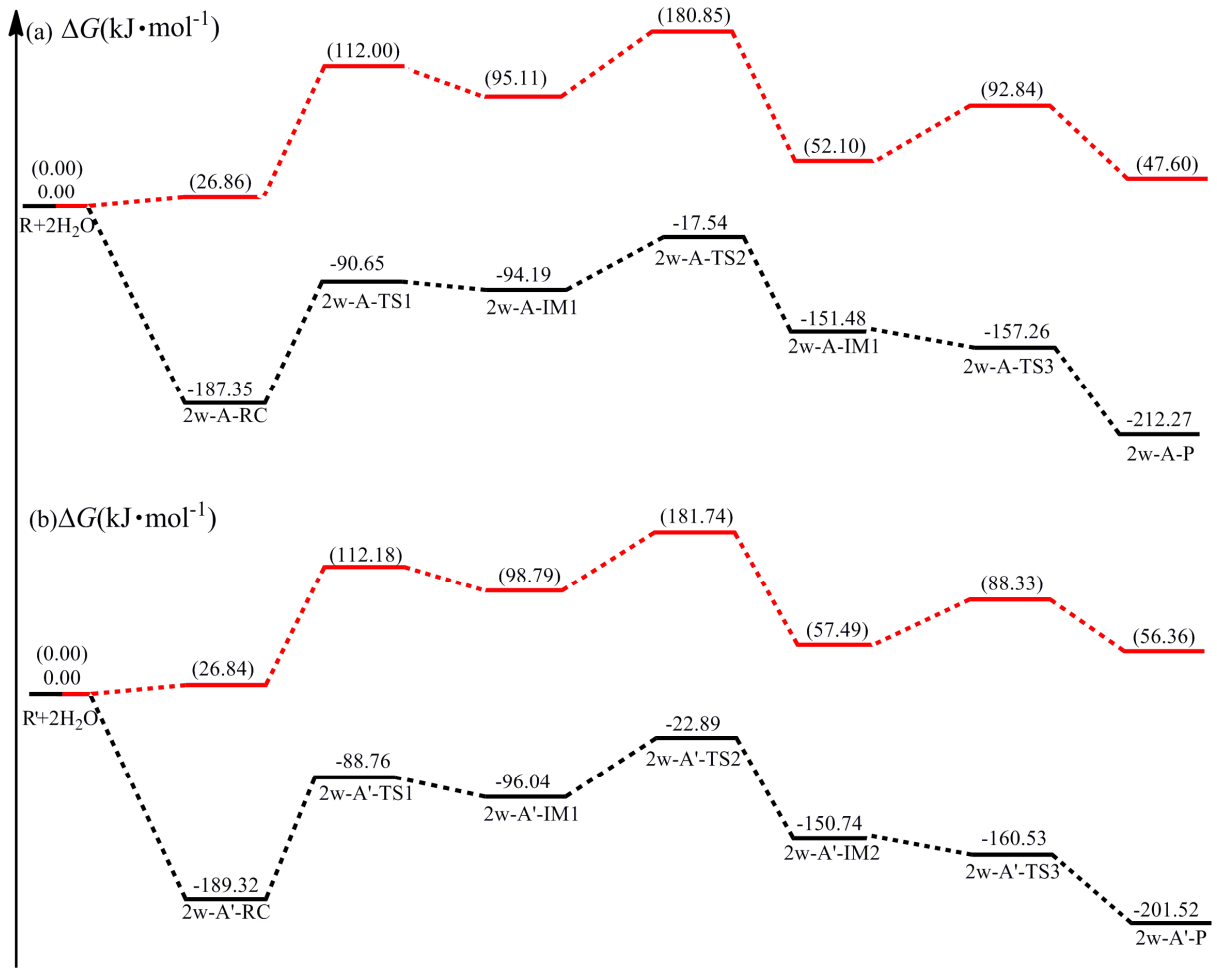


Fig. 11