THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Effects of Protonation and C5 Methylation on the Electrophilic Addition Reaction of Cytosine: A Computational Study

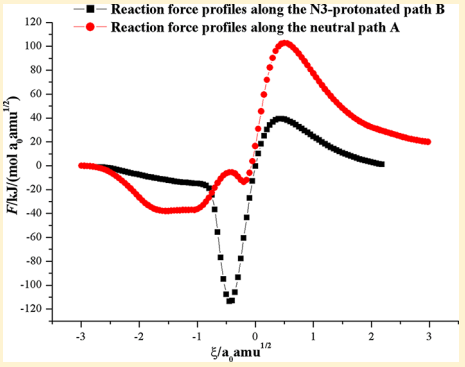
Lingxia Jin,[†,‡] Wenliang Wang,*[,†] Daodao Hu,[†] and Suotian Min[‡]

[†]Key Laboratory for Macromolecular Science of Shaanxi Province, School of Chemistry and Chemical Engineering, Shaanxi Normal University, Xi′an 710062, China

[‡]School of Chemical and Environmental Sciences, Shaanxi University of Technology, Hanzhong 723001, China

Supporting Information

ABSTRACT: The mechanism for the effects of protonation and C5 methylation on the electrophilic addition reaction of Cyt has been explored by means of CBS-QB3 and CBS-QB3/PCM methods. In the gas phase, three paths, two protonated paths (N3 and O2 protonated paths B and C) as well as one neutral path (path A), were mainly discussed, and the calculated results indicate that the reaction of the HSO_3^- group with neutral Cyt is unlikely because of its high activation free energy, whereas O2-protonated path (path C) is the most likely to occur. In the aqueous phase, path B is the most feasible mechanism to account for the fact that the activation free energy of path B decreases compared with the corresponding path in the gas phase, whereas those of paths A and C increase. The main striking results are that the HSO_3^- group directly interacts with the C5═C6 bond rather than the N3═C4 bond and that the C5 methylation, compared with Cyt, by decreasing values of global electrophilicity index manifests that C5 methylation forms are less electrophilic power as well as by decreasing values of NPA charges on C5 site of the intermediates make the trend of addition reaction weaken, which is in agreement with the experimental observation that the rate of 5-MeCyt reaction is approximately 2 orders of magnitude slower than that of Cyt in the presence of bisulfite. Apart from cis and trans isomers, the rare third isomer where both the CH_3 and SO_3 occupy axial positions has been first found in the reactions of neutral and protonated 5-MeCyt with the HSO_3^- group. Furthermore, the transformation of the third isomer from the cis isomer can occur easily.



1. INTRODUCTION

Cytosine (Cyt), one of the pyrimidine bases, occurs naturally in many nucleic acids, DNA, and RNA. It is also a parent compound of various modified nucleosides and nucleotides. Some of them occur naturally, and others are the product of chemical reactions of nucleic acids with various mutagenic agents. In eukaryotic genomes, enzymatic methylation of Cyt residues giving rise to 5-methylcytosine (5-MeCyt) occurs mostly at CpG sequences of DNA and it is thought to be involved in the regulation of gene expression and gene silencing.[1−3] Moreover, 5-MeCyt is believed to cause about one-third of all transition mutations and is responsible for human genetic diseases, cancer in particular.[4−7] To provide insights into these special biological phenomena, it has become clinically important to know the methylation status at specific sites in genomic DNA. However, conventional sequencing does not distinguish 5-MeCyt from Cyt, because both Cyt and 5-MeCyt are complementary to guanine. For now, the bisulfite genomic sequencing is believed to be the most common and classical approach for determining of the position of 5-MeCyt in a given DNA.[8−11]

Bisulfite treatment is expected to deaminate Cyt to uracil, and to leave 5-MeCyt unaltered.[8,12] Unfortunately, the experiments demonstrated[13] that unmethylated Cyt treated with bisulfite was either converted to uracil or failed to be converted and remained as Cyt, and the 5-MeCyt either did not undergo conversion or was inappropriately converted to thymine. Both types of conversion would lead to the erroneous estimates of methylation densities. To avoid these errors, Shiraishi and Hayatsu[14,15] introduced the LowMT (low molarity/temperature) and HighMT (high molarity/temperature) bisulfite-conversion protocols, but its inappropriate-conversion events remains unresolved. Therefore, to improve the bisulfite sequencing technology, it is important to understand the conversion mechanism and determine the best reaction condition.

Bisulfite reacts with pyrimidine nucleosides and undergoes addition to the 5,6 double bond to form pyrimidine-5,6-dihydro-6-sulfonates, which have been demonstrated to be important intermediates in causing deamination of Cyt to produce uracil,[8,12−17] particularly in acidic conditions. Furthermore, on the one hand many experiments demonstrated[18,19] that in the nonenzymatic condition the reaction called as spontaneous processes could take place readily but was pH-

Received: May 3, 2012
Revised: December 3, 2012
Published: December 5, 2012

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

dependent in that the acidic would certainly promote the reaction; on the other hand, the acidic-catalyzed process is likely to further lower the activation energy in recent theoretical works.[20−23] Therefore, the acidic media may tremendously accelerate the bisulfite addition and this case will be a subject of concern in this article.

In acidic media, there are three plausible protonation sites for Cyt, namely, N3, O2, and N4, respectively. It is certainly expected that the precedent protonation will strikingly enhance the electrophilicity of the C6, which can pave a more facile way for the subsequent nucleophilic attack of the bisulfite. Based on previous reports[24] and our calculated results, six possible protonated forms of Cyt exist; thus, the bisulfite addition reaction may proceed by six possible routes. But which of the six protonated routes will be preferred? For a long time, it is widely accepted that the bisulfite mediated deamination is likely to involve N3 protonation in the initial step. However, in the acidic aqueous phase, the O2-protonated Cyt is present in small amounts.[25] These intriguing phenomena inspire us to question whether the bisulfite also reacts with O2-protonated Cyt. In addition, it has been shown as a general trend that the deamination rate of 5-MeCyt is about 5-fold higher than that of Cyt in both nucleoside 5′-monophosphates and single-stranded DNA at physiological condition, pH 7.4,[26−29] whereas the rate of Cyt reaction is approximately 2 orders of magnitude faster than that of 5-MeCyt in the presence of bisulfite.[8,12−17] Such discrepancy may be also interesting for improving the bisulfite sequencing technology. Furthermore, the aromatic ring of Cyt possesses two double bonds, one for C5═C6 and the other for N3═C4, and the bisulfite may directly interact with the C5═C6 and N3═C4 bonds in theory. Why have only the products of bisulfite addition to the C5═C6 bond been reported rather than the N3═C4 bond?

Based on the information mentioned above, we will perform a detailed computational study on possible mechanisms for the effects of the protonation and C5 methylation on the electrophilic addition reaction of Cyt. Thus, herein three aspects are concerned as follows: first, the different sites of the proton-catalyzed process with the corresponding noncatalytic process are compared; second, the difference of the reaction trend in Cyt + $HSO_3^-$and 5-MeCyt + HSO_3^- reactions is taken into account under neutral and acidic conditions where the HSO_3^- group as a nucleophile was examined because it plays an important role in distinguishing 5-MeCyt from Cyt;[8,12−17] last, which of the double bonds would be preferred is investigated.

Our calculations point out that the proton-catalyzed process in reactions of Cyt + HSO_3^- and 5-MeCyt + HSO_3^- is more favorable than their noncatalytic process in both the gas and aqueous phases. Regarding the effects of the proton position on the reactivity, the obtained results show that in the gas phase the O2-protonated path is the most likely to occur, whereas in the aqueous phase the N3-protonated path is the most feasible mechanism. Compared with Cyt, the trend of the C5 methylation reaction with the HSO_3^- group, in neutral and acidic conditions, tends to be decreased. With regard to the choice of the addition sites, the HSO_3^- group directly interacts with the C5═C6 bond rather than the N3═C4 bond. Overall, our calculated results are in agreement with previous experimental phenomena.

2. CALCULATION METHODS AND COMPUTATIONAL INVESTIGATION OF REACTION PATHS

All calculations have been performed using the *Gaussian 03* package.[30] All stationary points including reactant complexes, transition states, and products have been optimized using density functional method B3LYP at the 6-311G(d,p) level.[31,32] Frequency analysis has been performed on all stationary points to verify whether the obtained structures are transition structures or local minima. Intrinsic reaction coordinate (IRC)[33] calculations have been carried out at the same level of theory from each transition state to ensure that the obtained transition states connected the appropriate reactants and products. To obtain more reliable energies, the CBS-QB3 composite approach[34] was performed based on the B3LYP/6-311G(d,p) geometries. This highly accurate method is a five-step method starting with a CBSB7 geometry and frequency calculation, followed by CCSD(T), MP4SDQ, and MP2 single-point calculations as well as a CBS extrapolation. Besides, the obtained stationary points have been further optimized using polarized continuum model (PCM)[35] at CBS-QB3 method, with dielectric constant 78.39 to simulate the aqueous environment.

2.1. Evaluation of Reactivity. *2.1.1. Dynamic Evaluation: Reaction Force F(ξ) Analysis.* In order to focus on the comparison of the mechanisms of the Cyt and 5-MeCyt and their protonated forms reactions with the HSO_3^- group, the reaction force $F(\xi)$ has been introduced by the negative of the derivative of the potential energy $E(\xi)$ with respect to the reaction coordinate ξ[36]

$$F(\xi) = -\frac{dE}{d\xi} \tag{1}$$

As the potential energy, the $F(\xi)$ presents a universal form along the reaction coordinate. For an elementary step it exhibits a minimum between the reactant and the transition state and a maximum between the transition state and the product. These particular points define the transition state region and correspond to transitions between different stages of the elementary step.[37−40]

2.1.2. Static Evaluation: Local Philicity Index $\omega^{\alpha}(\vec{r})$ and Natural Population Analysis. a. Local Philicity Index $\omega^{\alpha}(\vec{r})$. The local philicity index is defined as[41]

$$\omega^{\alpha}(\vec{r}) = \omega f^{\alpha}(\vec{r}) \tag{2}$$

α = + and − refer to nucleophilic and electrophilic attacks, respectively. Related condensed-to-atom variants for the atomic site k can be written as

$$\omega^{\alpha}(\vec{r}) = \omega f_k^{\alpha}(\vec{r}) \tag{3}$$

Where ω is global electrophilicity index,[42] which can be considered as the measure of the electrophilic power of a molecular system. $f_k^a(\vec{r})$ is a condensed form of Fukui functions of an atom k in a molecule with N electrons measuring the intramolecular reactivity at site r and is calculated using the finite difference approximation proposed by Yang and Mortier:[43]

$$f_k^+(\vec{r}) = \rho_k(N+1)(\vec{r}) - \rho_k(N)(\vec{r})$$

for nucleophilic attack (4)

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

$$f_k^-(\vec{r}) = \rho_k(N)(\vec{r}) - \rho_k(N-1)(\vec{r})$$

for electrophilic attack (5)

The ρ_k $(N + 1)$, $\rho_k(N - 1)$, and $\rho_k(N)$ are the electron population for the atom k in the anionic, cationic, and neutral states of a system, respectively. To compute $f_k^a(\vec{r})$, the ChelpG procedure[44] has been used.

b. Natural Population Analysis. The natural population analysis (NPA) charges have been used to analyze the difference of the reaction trend in the reaction of protonated Cyt and 5-MeCyt with HSO_3^- group. They were calculated at the CBS-QB3 method,[34] adopting natural population analysis.[45,46]

3. RESULTS AND DISCUSSION

The relative energies (ΔE^g), the relative free energies (ΔG^g and ΔG^s), and the activation free energies ($\Delta G^{g\neq}$ and $\Delta G^{s\neq}$) of the different stationary points for three paths (A-C and A′∼C′) both in the gas and aqueous phases are given in Tables 1 and 2,

Table 1. Energy Information (kJ·mol^{-1}) of Paths A (Neutral Cyt Path), B (N3-Protonated Cyt Path), and C (O2-Protonated Cyt Path) in the Gas and Aqueous Phases[a]

system	in the gas phase			in the aqueous phase	
	ΔE^g	ΔG^g	$\Delta G^{g\neq}$	ΔG^s	$\Delta G^{s\neq}$
		path A			
Ra (HSO_3^-+X_1)	0.00	0.00		0.00	
IM1a	−76.13	−39.06		53.16	
TSIM1a/P1a	−3.49	53.89		166.22	
P1a	−126.33	−71.16		12.01	
IM1a→P1a			92.95		113.06
		path B			
Rb (HSO_3^-+X_2)	0.00	0.00		0.00	
IM1b	−438.02	−387.97		81.92	
TSIM1b/P1b	−377.03	−321.26		113.16	
P1b	−417.51	−363.50		6.56	
IM1b→P1b			66.71		31.24
		path C			
Rc (HSO_3^-+X_3)	0.00	0.00		0.00	
IM1c	−395.27	−346.21		24.16	
TSIM1c/P1c	−358.90	−302.64		128.36	
P1c	−424.72	−370.46		−16.82	
IM1c→P1c			43.57		104.20

[a] ΔE^g, relative energies in the gas phase; ΔG^g, relative free energies in the gas phase; $\Delta G^{g\neq}$, activation free energies in the gas phase; ΔG^s, relative free energies in the aqueous phase; $\Delta G^{s\neq}$, activation free energies in the aqueous phase. X_1, neutral Cyt; X_2, protonation @ N3 oxo-amino Cyt; X_3, protonation @ O2 *trans*-hydroxy-amino Cyt.

respectively. The values of global electrophilicity index (ω), local philicity index for Cyt and its derivatives, and the NPA charges of C5 site (ρ_{C5}) of reaction complexes for three paths both in the gas and aqueous phases are listed in Table 3. The relevant information of other protonated paths is given in Tables S1 and S2 (Supporting Information). Meanwhile, the relative free energies of different protonated tautomers and their isomerization reactions are presented in Table S3. The information on the addition of HSO_3^- group to N3═C4 bond is also listed in Tables S4 and S5. The optimized structures, the potential energy surfaces for three paths of the reactions of

Table 2. Energy Information (kJ·mol^{-1}) of Paths A′ (Neutral 5-MeCyt Path), B′ (N3-Protonated 5-MeCyt Path), and C′ (O2-Protonated 5-MeCyt Path) in the Gas and Aqueous Phases[a]

system	in the gas phase			in the aqueous phase	
	ΔE^g	ΔG^g	$\Delta G^{g\neq}$	ΔG^s	$\Delta G^{s\neq}$
		path A′			
Ra′ (HSO_3^-+X_1′)	0.00	0.00		0.00	
IM1a′	−85.85	−48.95		55.21	
TSIM1a′/P1a′	4.77	62.37		175.02	
P1a′	−121.25	−65.14		24.23	
TSP1a′/P3a′	−108.71	−47.21		34.29	
P3a′	−119.71	−63.75		−11.28	
IM2a′	−93.00	−53.36		36.38	
TSIM2a′/P2a′	82.61	145.41		272.68	
P2a′	−115.02	−58.95		4.00	
IM1a′→P1a′			111.32		119.81
P1a′→P3a′			17.93		10.06
Ra′→P2a′			198.82		236.30
		path B′			
Rb′ (HSO_3^-+X_2′)	0.00	0.00		0.00	
IM1b′	−424.97	−376.04		85.06	
TSIM1b′/P1b′	−367.09	−309.98		119.20	
P1b′	−412.14	−356.17		16.94	
TSP1b′/P3b′	−401.30	−341.57		29.31	
P3b′	−442.60	−385.25		3.19	
IM2b′	−417.30	−370.77		91.56	
TSIM2b′/P2b′	−283.09	−224.89		194.72	
P2b′	−397.47	−343.20		22.05	
IM1b′→P1b′			66.06		34.14
P1b′→P3b′			14.60		12.37
Rb′→P2b′			145.88		103.16
		path C′			
Rc′ (HSO_3^-+X_3′)	0.00	0.00		0.00	
IM1c′	−380.35	−332.12		44.99	
TSIM1c′/P1c′	−345.04	−288.03		138.34	
P1c′	−414.71	−358.51		23.11	
TSP1c′/P3c′	−403.72	−343.64		33.41	
P3c′	−421.84	−366.51		11.92	
IM2c′	−369.47	−323.88		7.91	
TSIM2c′/P2c′	−243.44	−185.99		244.49	
P2c′	−398.63	−343.67		28.98	
IM1c′→P1c′			44.09		93.35
P1c′→P3c′			14.87		10.30
Rc′→P2c			137.89		226.58

[a] ΔE^g, relative energies in the gas phase; ΔG^g, relative free energies in the gas phase; $\Delta G^{g\neq}$, activation free energies in the gas phase; ΔG^s, relative free energies in the aqueous phase; $\Delta G^{s\neq}$, activation free energies in the aqueous phase. X_1′, neutral 5-MeCyt; X_2′, protonation @ N3 oxo-amino 5-MeCyt; X_3′, protonation @ O2 *trans*-hydroxy-amino 5-MeCyt.

neutral and protonated Cyt (A∼C) and C5 methylation (A′∼C′) with HSO_3^- group, and the evolution of the reaction force along reaction complexes, transition states, and products for the rate-determining step are shown in Figures 1−6. The evolution of the reaction force of other protonated paths (D∼G and D′∼G′) and the corresponding structure coordinates of all paths (A∼G and A′∼G′) are presented in Figures S1−S3.

3.1. Electrophilic Reaction Mechanisms in the Gas Phase. According to previous reports[24] and our calculated results, six protonated paths as well as one neutral path have

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

Table 3. Kohn-Sham Frontier Orbital Energies (ε_{HOMO} and ε_{LUMO}), Global Electrophilicity Index (ω), Local Electrophilicity Index on C6 (ω^{+}_{C6}), Local Nucleophilicity Index on C5 (ω^{-}_{C5}) of Cyt and Derivatives, and NPA Charge on C5 (ρ_{C5}) of the Reaction Complexes of Three Paths Both in the Gas (a) and Aqueous Phases (b)

species	$\varepsilon_{LUMO}{}^{a}$	$\varepsilon_{HOMO}{}^{a}$	ω^{a}	$\omega^{+}_{C6}{}^{a}$	$\omega^{-}_{C5}{}^{a}$	species	ρ_{C5}
			a				
Cyt	−1.05	−6.41	1.30	0.86	0.17	IM1a	−0.4832
5-MeCyt	−0.97	−6.20	1.23	0.80	0.03	IM1b	−0.4643
N3-protonated Cyt(CytN3$^+$)	−6.65	−11.87	8.21	4.16	0.19	IM1c	−0.5133
N3-protonated 5-MeCyt(5-MeCytN3$^+$)	−6.47	−11.45	8.06	3.94	1.06	IM1a′	−0.2842
O2-protonated Cyt(Cyt2t$^+$)	−6.16	−11.93	7.09	3.56	0.40	IM1b′	−0.2674
O2-protonated 5-MeCyt (5-MeCyt2t$^+$)	−5.96	−11.55	6.86	3.14	1.32	IM1c′	−0.3034
			b				
Cyt	−0.98	−6.57	1.27	0.84	0.16	IM1a	−0.4755
5-MeCyt	−0.94	−6.34	1.23	0.75	0.05	IM1b	−0.5064
N3-protonated Cyt(CytN3$^+$)	−2.17	−7.55	2.20	1.17	0.08	IM1c	−0.4566
N3-protonated 5-MeCyt(5-MeCytN3$^+$)	−2.14	−7.29	2.16	1.17	0.11	IM1a′	−0.2554
O2-protonated Cyt(Cyt2t$^+$)	−1.86	−7.64	1.95	1.04	0.05	IM1b′	−0.3037
O2-protonated 5-MeCyt (5-MeCyt2t$^+$)	−1.82	−7.44	1.91	0.97	0.14	IM1c′	−0.1984

[a]All values are expressed in eV; $\omega \approx 1/8(\varepsilon_{LUMO} + \varepsilon_{HOMO})^2/(\varepsilon_{LUMO} - \varepsilon_{HOMO})$;[48,49] ε_{HOMO} and ε_{LUMO} are the energies of the highest occupied (HOMO) and the lowest unoccupied (LUMO) molecular orbitals, respectively.

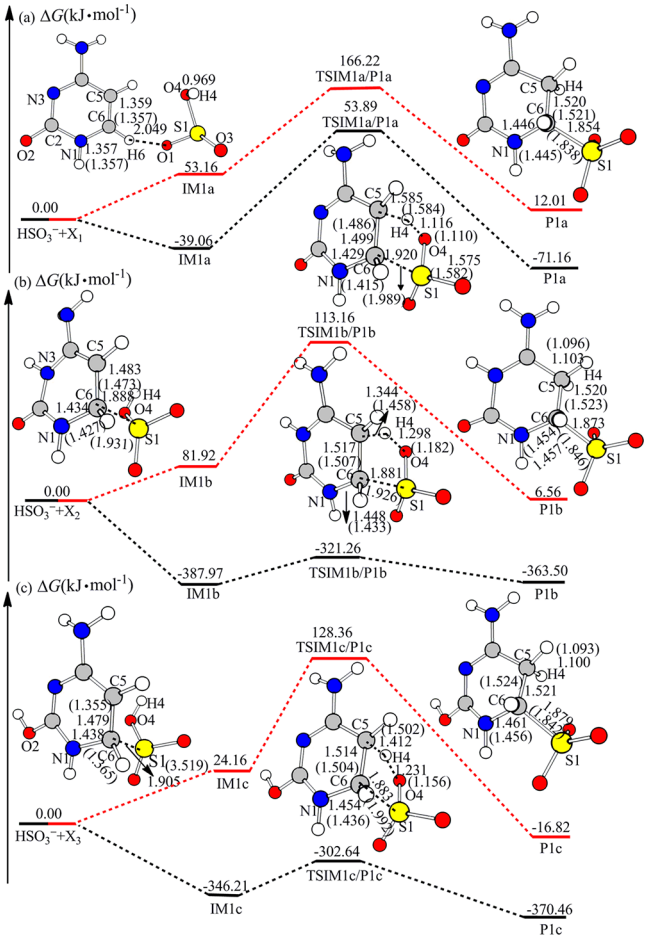


Figure 1. Potential energy surfaces of paths A (a), B (b), and C (c) in the gas (blank line) and aqueous phases (red line). The total free energy of reactants respectively in the gas and aqueous phases is set to zero. Relative free energies and optimized stationary structures are at the CBS-QB3 composite approach. Bond distances (Å) both in the gas phase and in aqueous phases (in brackets) are indicated. X_1 denotes neutral Cyt; X_2 denotes protonation @ N3 oxo-amino Cyt; X_3 denotes protonation @ O2 *trans*-hydroxy-amino Cyt.

been found. Path A refers to the reaction of neutral Cyt with HSO_3^- group. Paths B and G are protonation @ N3 oxo-amino (CytN3$^+$) and N4 oxo-amino (CytN4$^+$) reactions, respectively. Paths C and D are protonation @ O2 *trans*-hydroxy-amino (Cyt2t$^+$) and *cis*-hydroxy-amino (Cyt2c$^+$) reactions, respectively. Paths E and F are protonation @ O2 *trans*-hydroxy-imine (Cyt23t$^+$) and *cis*-hydroxy-imine (Cyt23c$^+$) reactions, respectively.

Among these protonated forms, the protonated forms of Cyt23t$^+$, Cyt23c$^+$, and CytN4$^+$ are obviously less stable than those of CytN3$^+$, Cyt2t$^+$, and Cyt2c$^+$ and are impossible to convert into CytN3$^+$, Cyt2t$^+$, and Cyt2c$^+$ due to the higher free energy barriers. Thus the reactions of these protonated forms with the HSO_3^- group are neglected in the following discussion. The relevant energies of these protonated paths are given in Table S3. Furthermore, for CytN3$^+$, Cyt2t$^+$, and Cyt2c$^+$, in the gas phase, the order of stability obtained is Cyt2t$^+$ > CytN3$^+$ > Cyt2c$^+$, with an energetic difference of 34.00 kJ·mol^{-1} between Cyt2t$^+$ and Cyt2c$^+$, and in the aqueous phase, the order of stability obtained is CytN3$^+$ > Cyt2t$^+$ > Cyt2c$^+$, with an energetic difference of 39.22 kJ·mol^{-1} between CytN3$^+$ and Cyt2c$^+$. According to the literature,[47] the relative energy of 40.00 kJ·mol^{-1} is a reasonable limit for the existence of the stable species, so CytN3$^+$, Cyt2t$^+$, and Cyt2c$^+$ can coexist. However, Cyt2c$^+$ is the least stable among the three tautomers and is easily converted into Cyt2t$^+$; thus, the relevant energies of the reaction of Cyt2c$^+$ with the HSO_3^- group are also listed in Table S3.

In addition, the reactions of the C5═C6 and N3═C4 bonds of Cyt with the HSO_3^- group have also been investigated. As for the reaction between C5═C6 bond and the HSO_3^- group, three paths discussed above have been evaluated, whereas one neutral path (path J) and one O2-protonated (Cyt2t$^+$) path (path H) have been investigated for the reaction of N3═C4 bond with HSO_3^-group. As seen from Tables S4 and S5, the activation free energy of path J is so high ($\Delta G^{g\neq}$ = 123.39 kJ·mol^{-1}) that it is unlikely to occur, and as for path H, the activation free energy of the reverse reaction ($\Delta G^{g\neq}$ = 2.08 kJ·mol^{-1}) is much lower than the forward ($\Delta G^{g\neq}$ = 27.99 kJ·mol^{-1}), indicating that the product is impossible to be formed under acidic condition. This can explain why the

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

adducts of HSO_3^-group addition to N3═C4 bond has not been found in experiments.[8,12−17] Therefore, only the addition reaction of HSO_3^- group to the C5═C6 bond of Cyt has been investigated in this paper.

3.1.1. Electrophilic Addition to Cyt Reactions. **Path A (neutral Cyt path).** As shown in Figure 1a, path A is a one-step mechanism. In path A, the HSO_3^-group first forms a hydrogen bond complex IM1a, where only one hydrogen bond between the H6 of Cyt and the O1 of the HSO_3^- group is formed. With the approaching of the HSO_3^- group, the H6···O1 bond has been cleaved and then adds to the C5═C6 double bond to form 5,-6-dihydro-6-sulfonate anion via transition state TSIM1a/P1a, which has a five-membered-ring structure that is mainly associated with the addition of S1 to the C6 site and the transfer of H4 from O4 to C5 atom. In TSIM1a/P1a, the O4−H4 bond is lengthened from 0.969 Å in IM1a to 1.116 Å, and the S1−O4 bond is shorted to 1.575 Å, indicating that the S1═O4 bond has been partially formed. The distance of C5−C6 is 1.499 Å, 0.140 Å longer than that in IM1a (1.359 Å), indicating that the C5═C6 bond has been partially broken in TSIM1a/P1a. The H4 transfer and S1 addition occur synchronously with $\Delta G^{g\neq}$ = 92.95 kJ·mol^{-1} and lead to the formation of P1a. In P1a, the C6 atom has a tetrahedral structure, and the distances of C6−C5, C6−N1, and C6−H6 are stretched to 1.520, 1.446, and 1.097 Å, respectively, longer than those in IM1a (1.359 Å for C6−C5, 1.357 Å for C6−N1, and 1.091 Å for C6−H1), suggesting that the hybridization of C6 atom converts from sp^2 in IM1a to sp^3 and the C5═C6 double bond turns into a single bond.

Path B (N3-protonated Cyt path). Different from path A, path B seen in Figure 1b is a two-step process. The initial reaction complex IM1b is formed barrierlessly via S1 atom of the HSO_3^- group attacking C6 atom, which is dissimilar with IM1a that is formed by the hydrogen bonds. In IM1b, the distance of the C5═C6 bond is stretched to 1.483 Å, that is fairly close to a C−C bond (1.530 Å). The distance of C6−S1 is 1.888 Å, which is approximate to the length of a normal C−S bond (1.818 Å). With the approaching of the HSO_3^- group, the distance of C5−C6 is elongated, and at the same time, C6−S1 is shortened with the proton transfer from O4 to C5 atom to form P1b. Similar to the formation of P1a, the step from IM1b to P1b involved a tetrahedral structure also occurs through a five-membered transition state TSIM1b/P1b ($\Delta G^{g\neq}$ = 66.71 kJ·mol^{-1}). In TSIM1b/P1b, the distance of C6−S1 is 1.881 Å, 0.007 Å shorter than that in IM1b, suggesting that the process of path B only involves the H4 transfer from O4 to C5 atom.

Path C (O2-protonated Cyt path) Similar to path B, path C shown in Figure 1c is also a two-step reaction. In IM1c, the distances of C5−C6 and C6−S1 are 1.479 and 1.905 Å, respectively, which is similar to that in IM1b. Analogous to TSIM1b/P1b, TSIM1c/P1c also undergoes the H4 transfer from O4 to C5 atom to form P1c with the $\Delta G^{g\neq}$ of 43.75 kJ·mol^{-1}. Compared with paths A and B, the $\Delta G^{g\neq}$ of path C is the lowest and the formation of P1c is more stable than others. Thus path C is favorable thermodynamically and kinetically. It is worth noting that the diastereomer does not exist in the products of the reactions of Cyt and its protonated forms with the HSO_3^- group due to the symmetric carbon at the C5 site. Therefore, such case has not been considered. On the contrary, as for C5 methylation, the diastereomers have been investigated.

3.1.2. Effects of Methylation and Stereochemistry on the Activation Free Energy. Considering the asymmetry of the C5 site, apart from P1a′, P1b′, and P1c′ obtained by a similar way as the corresponding paths of the reactions of Cyt and its protonated forms with the HSO_3^- group, the other diastereomers can also be located. P1a′ is 5-H cis to 6-SO_3^- (cis isomer), as well as P1b′ and P1c′ are 5-H cis to 6-SO_3 (cis isomer). For the three cis isomers above, both CH_3 and SO_3 occupy equatorial positions. Here, we take the formation of 5-H trans to 6-SO_3^- (trans isomer) as an example. This reaction begins with the hydrogen-bonded complex IM2a′ formed by the interaction of O1···H1 and O4···H6 (Figure 2). Starting

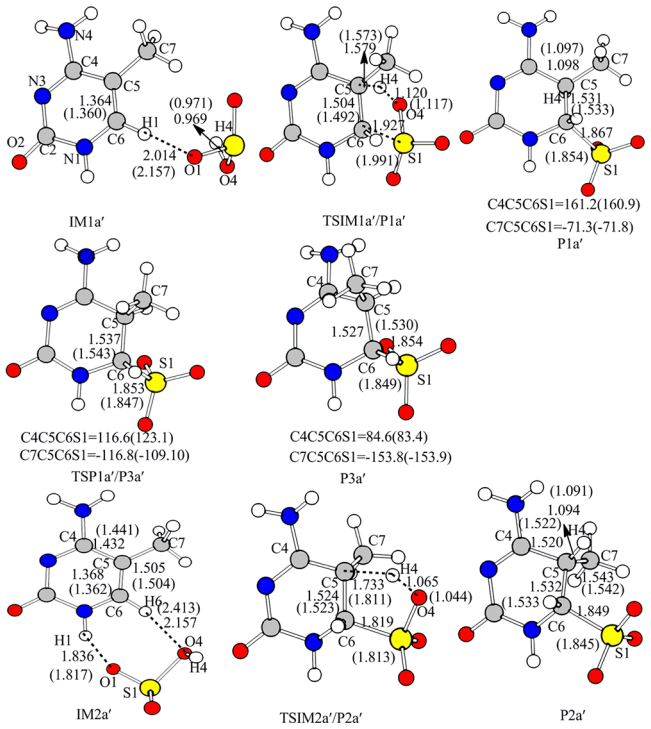


Figure 2. Optimized stationary structures of path A′ (neutral 5-MeCyt path) at the CBS-QB3 composite approach. Bond distances (Å) and dihedral angles (°) both in the gas and aqueous phases (in brackets) are indicated.

from IM2a′, the H4 atom is transferred to C5 atom via a five-membered transition state TSIM2a′/P2a′. This step is associated with the $\Delta G^{g\neq}$ = 198.82 kJ·mol^{-1} and leads to the formation of P2a′. In P2a′, the C5 atom has a tetrahedral structure, and the distances of C5−C7, C5−C4, and C5−C6 are stretched to 1.543, 1.520, and 1.532 Å, respectively, longer than those in IM2a′ (1.505 Å for C5−C7, 1.432 Å for C5−C4, and 1.368 Å for C5−C6), showing that the hybridization of C5 atom converts from sp^2 in IM2a′ to sp^3 and the C5═C6 double bond turns into a single bond. Hence, the 5-H trans to 6-SO_3^- (trans isomer) where the CH_3 occupies axial position and the SO_3 occupies equatorial position is obtained.

In addition to the paths mentioned above, the possible isomerization path based on the analysis of the two products is also considered, and the 5-H cis to 6-SO_3^- (third isomer) where both the CH_3 and SO_3 occupy axial positions is obtained. Starting from P1a′, there is a conformational change with low barrier ($\Delta G^{g\neq}$ = 17.93 kJ·mol^{-1}) leading to the formation of P3a′. P1a′ and P3a′ are very similar, which is mainly differing in the partial torsion angles of the pyrimidine cycle and the sites of the two functional groups (−CH_3 and −SO_3). Besides, the ΔG^g of P3a′ and P1a′ are −63.75 kJ·mol^{-1} and ΔG^g = −65.14

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

kJ·mol^{-1}, respectively, so the two adducts may coexist based on Mezey theory.[47]

As shown in Figures 3 and 4, in paths B′ (N3-protonated 5-MeCyt path) and C′ (O2-protonated 5-MeCyt path), the

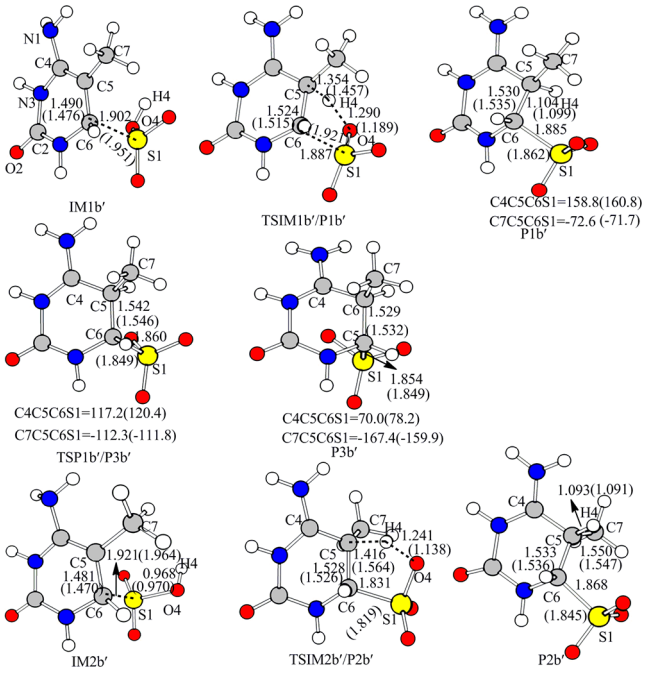


Figure 3. Optimized stationary structures along path B′ (N3-protonated 5-MeCyt path) at the CBS-QB3 composite approach. Bond distances (Å) and dihedral angles (°) both in the gas and aqueous phases (in brackets) are indicated.

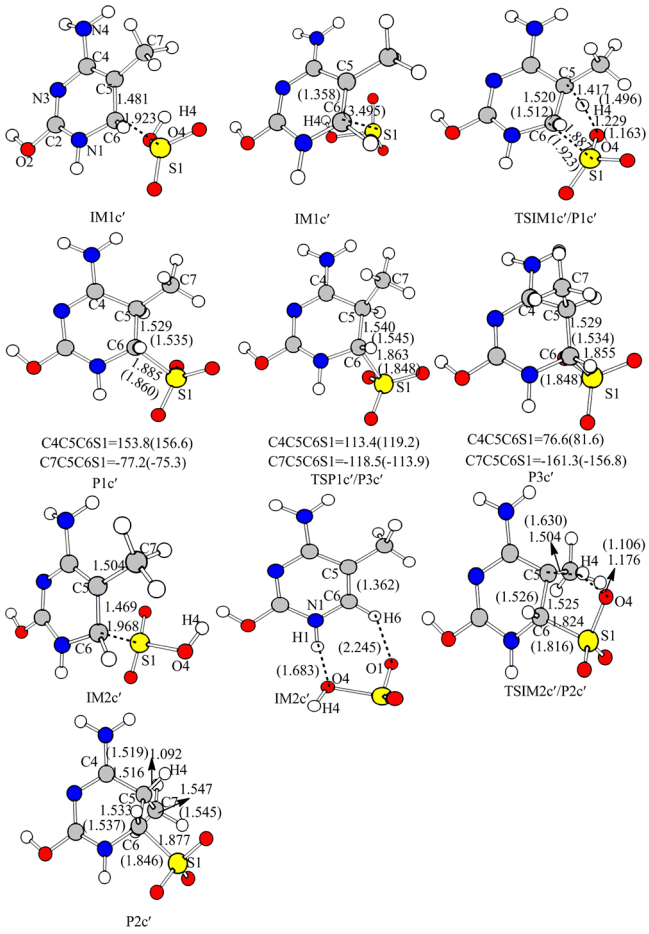


Figure 4. Optimized stationary structures along path C′ (O2-protonated 5-MeCyt path) at the CBS-QB3 composite approach. Bond distances (Å) and dihedral angles (°) both in the gas and aqueous phases (in brackets) are indicated.

neutral trans isomers P2b′ and P2c′ are also obtained, respectively. Unlike IM2a′ that is formed by the hydrogen bonds, IM2b′ and IM2c′ are formed by barrierlessly via S1 atom of the HSO_3^- group attacking C6 atom. Then, the H4 is also transferred from O4 to C5 atom leading to the formation of P2b′ and P2c′ with the activation free energies of 145.88 and 137.90 kJ·mol^{-1}, respectively. P3b′ and P3c′ where both the CH_3 and SO_3 occupy axial positions also exist only with the $\Delta G^{g\neq}$ of 14.60 and 14.87 kJ·mol^{-1}, respectively. Therefore, once P1b′ and P1c′ formed, they can quickly convert into P3b′ and P3c′.

The potential energy surfaces for the C5 methylation paths A′, B′, and C′ are shown in Figure 5. As for the rate-determining step of C5 methylation paths, the three paths for the formation of trans isomers and the path A′ for the formation of *cis*-isomer are impossible to occur due to so high activation free energies. As for the reaction of neutral 5-MeCyt with the HSO_3^- group, the $\Delta G^{g\neq}$ is increased by 18.37 kJ·mol^{-1} compared with the reaction of neutral Cyt with the HSO_3^- group, whereas there is a little difference in $\Delta G^{g\neq}$ between the reactions of protonated 5-MeCyt and protonated Cyt with the HSO_3^- group. Thus, for a good understanding of the difference of reaction trend between protonated Cyt and protonated 5-MeCyt, the global electrophilicity index ω, local philicity index, and NPA charge (ρ_{C5}) will be discussed in follow-up context. It is noteworthy that the cis isomers convert into third isomers with small $\Delta G^{g\neq}$, indicating there is no significant effect on the mechanism of these reactions. Thus, our interest has been focused on the formation mechanisms of cis and trans isomers.

3.2. Electrophilic Reaction Mechanisms in the Aqueous Phase. Herein we make a short discussion about the effect of water on the reaction mechanisms. The potential energy surfaces and optimized stationary structures in the aqueous phase are also involved in Figures 1−4. As seen from Figures 1−4, except for IM2a′ one hydrogen bond instead of two hydrogen bonds, and the O4 atom is far from H6 atom than that in the gas phase, the stationary structures of paths A and B (paths A′ and B′) in the aqueous phase are similar with those in the gas phase. As for paths C and C′, the optimized geometries in IM1c and IM1c′ change a lot in the aqueous phase. The HSO_3^-group is far from Cyt2t$^+$ and 5-MeCyt2t$^+$. The distances of C6···S1 in IM1c and IM1c′ are elongated to 3.519 and 3.495 Å, respectively, whereas the C5−C6 bond is shortened to 1.355 and 1.358, respectively. This indicates that the interaction has substantially been weakened between HSO_3^-group and Cyt2t$^+$ and 5-MeCyt2t$^+$. The same happens in IM2c′, and the position of the HSO_3^- group in IM2c′ is different than that in the gas phase. Two hydrogen bonds formed in the aqueous phase, one is N1−H1···O4, and the other is O1···H6−C6. The solvent effects obviously change the geometries of reaction complexes for paths C and C′, and this may be lead to the free energy barrier to be changed.

As seen in Figures 1 and 5, for path A (neutral Cyt path), the activation free energy barrier slightly increases from 92.95 to 113.06 kJ·mol^{-1}, and the activation free energy barrier of rate-

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

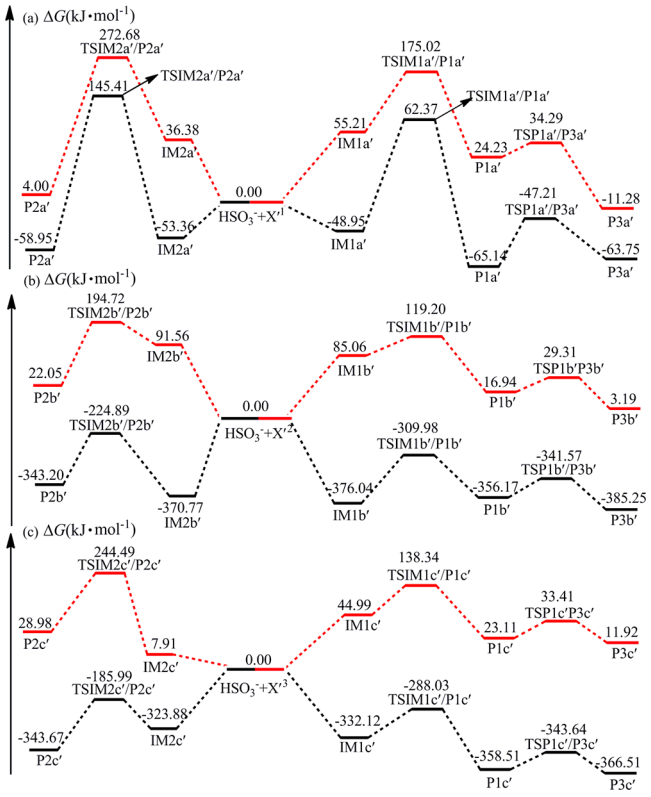


Figure 5. Potential energy surfaces of paths A′ (a), B′ (b), and C′ (c) in the gas phase (blank line) and aqueous phase (red line). The total free energy of reactants respectively in the gas and aqueous phases is set to be zero. Relative free energies are at the CBS-QB3 composite approach. X'_1 denotes neutral 5-MeCyt; X'_2 denotes protonation @ N3 oxo-amino 5-MeCyt; X'_3 denotes protonation @ O2 *trans*-hydroxy-amino 5-MeCyt.

determining step of path A′ (neutral 5-MeCyt path) also slightly increases from 111.32 to 119.81 kJ·mol^{-1}. This reveals that water has no significant impact on the activation free energy barrier of neutral counterparts. However, the activation free energies for the protonated paths B (N3-protonated Cyt path) and C (O2-protonated Cyt path) change obviously in water. The activation free energy for path B is significantly reduced from 66.71 to 31.24 kJ·mol^{-1}, whereas the path C is tremendously rose to 104.20 kJ·mol^{-1}. It is observed that the path B in the aqueous phase is energetically more favorable than that in the gas phase, whereas the path C in the aqueous phase is less favorable to occur than that in the gas phase. As for the reactions of protonated 5-MeCyt (paths B′ and C′), the activation free energies of cis isomers formation are changed by the range of 31.92−49.26 kJ·mol^{-1}. This illustrates that water has noticeably effect on the activation free energy barriers of protonated paths. It is worth noting that the activation free energies of the formation of trans isomers are changed from 103.16 to 236.30 kJ·mol^{-1}, and the reactions retain much more unfavorable in the aqueous phase.

For a good understanding of the solvation effect on the free energy barriers, the NPA charges along the three paths in the aqueous phase are compared with those in the gas phase. It is obvious in Table 3 that, in the aqueous phase, the NPA charge on C5 atom of IM1b is −0.5064 e, which has more negative charge than that in the gas phase (−0.4643 e for IM1b) to make C5 atom win more chances to react with the electrophile. On the contrary, for path C, the stronger nucleophilicity of the C5 atom in IM1c is significantly reduced with the NPA charge change from −0.5133 e (gas phase) to −0.4566 e (aqueous phase). This may explain that why path C in the aqueous phase is associated with poorer reaction trend than that in the gas phase. In addition, for path A, the NPA charge in IM1a has a very small change (∼0.08 e) in the aqueous phase, indicating that the solvation effect on this path is comparatively negligible.

3.3. Comparison of the Three Paths. *3.3.1. Dynamic Evaluation: Reaction Force F(ξ) Analysis.* As it is well-known, the $F(\xi)$ of an elementary step has a minimum and a maximum at the two inflection points of potential energy, and the reaction force vanishes at the transition state.[25−29] Figure 6a is the $F(\xi)$ evolution of path A (neutral Cyt path). From IM1a to P1a, the $F(\xi)$ goes through two local minima (ξ = −1.55, $F(\xi)$ = −38.04 kJ·mol^{-1} and ξ=-0.20, $F(\xi)$ = −13.59 kJ·mol^{-1}) and two local maxima (ξ = −0.40, $F(\xi)$ = −5.52 kJ·mol^{-1} and ξ = 0.51, $F(\xi)$ = 102.79 kJ·mol^{-1}). This confirms the fact that this elementary step is composed of two events and that it is their simultaneity that leads to only one energy barrier and a unique transition state.[38] It can be supposed that the second minimum is not as deep as the first one because of the superposition between the relaxation part of the first event and the activation part of the second one.

The $F(\xi)$ evolution of path B (N3-protonated Cyt path) is shown in Figure 6b. Unlike path A, for path B, the relaxation part is less deep and wider than the activation part. The $F(\xi)$ displays a minimum (ξ_1 = −0.45, $F(\xi)$ = −111.84 kJ·mol^{-1}) and a maximum (ξ_2 = 0.45, $F(\xi)$ = 35.75 kJ·mol^{-1}). These two extremes divide the reaction into three different regions: reactant region ($\xi_R \sim \xi_1$), transition state region ($\xi_1 \sim \xi_2$), and product region ($\xi_2 \sim \xi_P$). In the first region, the HSO_3^- group rapidly adjusts bond stretching and angle bending. The negative $F(\xi)$ increases rapidly, which reaches its greatest strength (minimum). At ξ_{TS}, the maximum of potential energy, the driving and the retarding forces balance exactly. Between ξ_1 and ξ_2, the transition state region is characterized as structural change by the movement of the H4 from O4 to C5 atom and S1 from the HSO_3^-group to C6 atom. Finally, in the product region ($\xi_2 \sim \xi_P$), the O4−H4 bond is broken and the newly H4−C5 and S1−C6 bonds are formed.

The $F(\xi)$ evolution of path C (O2-protonated Cyt path) is shown in Figure 6c. The $F(\xi)$ evolution of paths C and B look alike in terms of shape and intensity, indicating that the same reaction mechanism exists.

For path A, from the structural point of view, TSIM1a/P1a undergoes a big change compared with IM1a. According to the literature,[40] this suggests that, during the step (M1a-TSIM1a/P1a-P1a), the first process corresponds to the addition of HSO_3^- group to C6 site followed by the large structural distortions, whereas the second one refers to the H4 transfer from O4 to C5 atom. This has been confirmed in the above section by the examination of the evolution of structural parameters during the step (M1a-TSIM1a/P1a-P1a). However, for paths B and C, the transition states only involve the H4 transfer from O4 to C5 atom, and the changes in structures of TSIM1b/P1b and TSIM1c/P1c compared with their reaction complexes are reversed in the case of path A, indicating that the $F(\xi)$ of path A is different from that of paths B and C, which further lead to the local minimum and maximum.

As seen from Figure 6a in path A′ (neutral 5-MeCyt path), the $F(\xi)$ also has a local minimum and local maximum, but the intensity is bigger than that in path A, suggesting that the simultaneity of path A′ is worse than path A.[23,40] Moreover, the

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

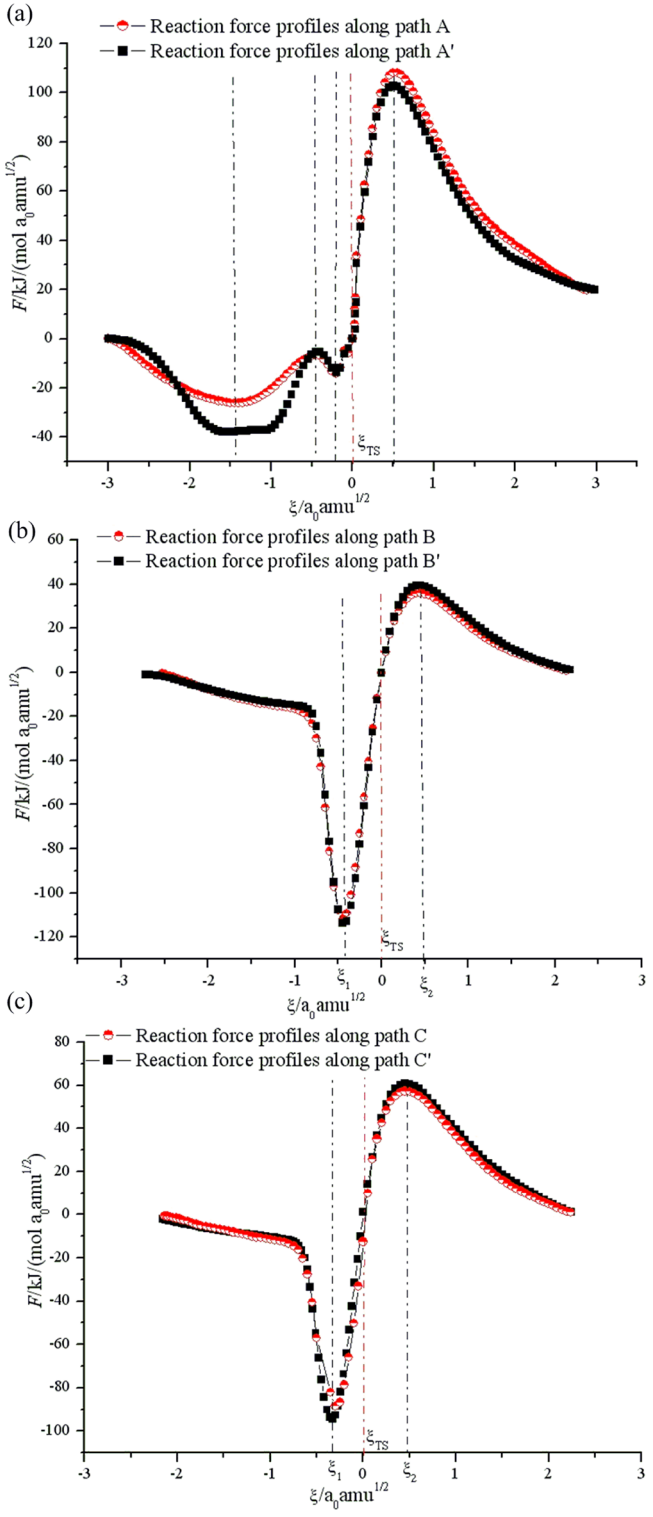


Figure 6. Reaction force profiles for the rate-determining step of path A∼C (A′∼C′) along the reaction coordinates. Paths A∼C (a∼c) denote the reaction of neutral and protonated Cyt with HSO_3^-group; paths A′∼C′ (a∼c) denote the reaction of neutral and protonated 5-MeCyt with the HSO_3^- group. The red dashed lines indicate the position of the transitions states and the black ones the limits of the different regions identified along the reaction coordinate.

product region of paths A′ and A look alike in terms of shape and intensity due to the similar products obtained. As seen from Figure 6b, both paths B′ and B look alike to some extent, indicating that the same reaction mechanism is involved in these two paths. Similar to Figure 6b, the consistent conclusion is also obtained as seen from Figure 6c.

3.3.2. Static Evaluation of Reactivity: Global Electrophilicity Index, Local Philicity Index, and NPA Analysis. In Table 3, the increasing value of ω_{C6}^{+} indicates C6 site of neutral Cyt and protonated Cyt are more electrophilic than that of neutral 5-MeCyt and protonated 5-MeCyt, and the protonated 5-MeCyt, by increasing the value of ω_{C5}^{-}, makes the electrophilic attack to C5 easier. Although the C6 site of protonated Cyt possesses the higher electrophilicity, the ω_{C5}^{-} value of protonated Cyt is much lower than protonated 5-MeCyt, thus it difficult to make a clear decision on the overall trend of these compounds. For this reason, the global electrophilicity index was introduced.

As shown in Table 3, it turns out that all protonated forms are characterized by significant electrophilicity values ranging from 1.91 to 8.21 eV, whereas the neutral forms are ranging from 1.23 to 1.30 eV. Therefore, according to the electrophilicity scale proposed by Domingo et al.,[50] the diene systems can be considered as strong electrophiles (ω>1.50 eV), whereas all the diene systems can be classified as moderate or marginal electrophiles (ω < 1.50 eV). Because the pyrimidine base has azadiene unit, the same criteria are taken to classify the electrophilic power in this paper. Note that the protonated forms act as strong electrophiles, whereas the neutral Cyt and 5-MetCyt behave as marginal electrophiles. In addition, the values of ω indicate that the neutral and protonated Cyt are more electrophilicity power than the corresponding C5 methylation forms. Consequently, the protonated forms make nucleophilic attack easier whereas the C5 methylation forms make nucleophilic attack more difficult. This can explain that why protonated paths are more favorable and the C5 methylation is less feasible. Nevertheless, in the gas phase, O2-protonated path is more feasible than N3-protonated path in both Cyt and 5-MeCyt systems (seen in Tables 1 and 2). Thus, the global electrophilicity index approach appears to fail.

As mentioned above, the NPA charge distribution from NBO method can be further analyzed instead of global electrophilicity and philicity indexes. On the basis of the reaction force $F(\xi)$ analysis, the rate-determining steps of proton-catalyzed processes in both Cyt and 5-MeCyt reactions with HSO_3^-group only involve the H4 transfer from O4 to C5 site where electrophilic attack occurs. As seen from Table 3, in the gas phase, the C5 sites in paths of protonated 5-MeCyt (−0.2674 e for IM1b′ and −0.3034 e for IM1c′) have less negative charge than those in paths of protonated Cyt (−0.4643 e for IM1b and −0.5133 e for IM1c). This indicates that, the reaction trend of protonated Cyt paths tends to be increased compared with protonated 5-MeCyt paths, and the more negative charge on C5 site of path C gains more opportunities to react with the electrophile.

Compared with in the gas phase (section b of Table 3), the C5 site in path B (−0.5064 e for IM1b) in the aqueous phase has more negative charge than that in path A (−0.4755 e for IM1a) and C (−0.4566 e for IM1c); thus, the C5 site in path B is easier to undergo electrophilic attack than other paths. Meanwhile, the negative charges of the C5 site of C5 methylation paths are much smaller than those in Cyt paths, which makes the reaction trend weaken. These results are in agreement with the experimental observations[8,12−17] that the HSO_3^- group adds to Cyt with a speed significantly faster than that for 5-MeCyt at acidic pH, especially for its N3 protonation.

dx.doi.org/10.1021/jp304282z | J. Phys. Chem. B 2013, 117, 3−12

4. CONCLUSIONS

The effects of protonation and C5 methylation on electrophilic addition of Cyt have been studied using the CBS-QB3 method. The solvent effects of water on the three paths have also been evaluated by the CBS-QB3/PCM method.

(1) The protonated paths are more feasible to occur with the lower $\Delta G^{g\neq}$. The $\Delta G^{g\neq}$ for paths B (N3-protonated Cyt path) and C (O2-protonated Cyt path) are 66.71 and 43.57 kJ·mol^{-1}, respectively, whereas the $\Delta G^{g\neq}$ for path A (neutral Cyt path) is 92.95 kJ·mol^{-1}. In addition, the increasing values of ω, indicate that protonated Cyt are more electrophilic than neutral Cyt, and the increasing values of ρ_{C5}, imply that the reaction trend of protonated Cyt paths tends to be increased. Overall, the reaction occurs more easily under acidic conditions than neutral conditions.

(2) The difference in $\Delta G^{g\neq}$ is small between protonated 5-MeCyt and protonated Cyt paths. The decreasing values of ω manifest that C5 methylation forms are less electrophilic power and the decreasing values of ρ_{C5} make the trend of addition reaction weaken.

(3) Three isomers are obtained in the paths of C5 methylation. Apart from the cis isomer where both CH_3 and SO_3 occupy equatorial positions and the trans isomer where the CH_3 occupies the axial position and the SO_3 occupies the equatorial position obtained, the rare third isomer where both the CH_3 and SO_3 occupy axial positions has been first found in this paper. The cis isomers can be converted into the third isomers easily, whereas the formation of trans isomers involved so high activation free energies are unlikely to occur.

(4) The solvent effects of water significantly contribute to changing the activation free energy. Both neutral path A and O2-protonated path C are impossible to occur in the aqueous phase, and the N3-protonated path B is the most favorable mechanism, which agrees well with the available experimental phenomena.

5. BIOCHEMICAL IMPLICATIONS

Experimental studies[8,12−17] and our computed results have verified that the reactions of Cyt and HSO_3^- as well as 5-MeCyt and HSO_3^- can be accelerated by the protonation and that the protonated site can be controlled conveniently via adjusting the pH value. Hence, the reaction is promoted by means of reducing pH, but the strong acid leading to the DNA chain cleavage causes the inefficient PCR performance of the treated DNA.[16] In addition, in the aqueous phase the anionic adduct formed by Cyt and HSO_3^- is more stable in acidic condition than in neutral condition. However, the exorbitant stability of the adduct disfavors to form thymine in the following bisulfite-mediated deamination process. Thus, to explore the best reaction conditions, the choice of appropriate isomer is critical in the bisulfite sequencing technology.

It should be pointed out that the optimum pH value should be ensured elementary in the bisulfite treatment of DNA and then adjusting other parameters, such as the bisulfite concentration and temperature. Overall, this work can explain the previous experiment results successfully and provide some implications for the bisulfite sequencing technology.

■ ASSOCIATED CONTENT

Supporting Information

Information of other protonated paths given in Tables S1 and S2. Relative free energies of different protonated tautomers and their isomerization reactions presented in Table S3. Information on the addition of HSO_3^-group to the N3═C4 bond listed in Tables S4 and S5. Evolution of the reaction force of other protonated paths (D∼G and D′∼G′) and the corresponding structure coordinates of all paths (A∼G and A′∼G′) presented in Figures S1−S3. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +86-29-81530815. Fax: +86-29-81530727. E-mail: wlwang@snnu.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 21173139), the Innovation Funds of Graduate Programs of the Shaanxi Normal University (No. 2012CXB017), and the Shaanxi Province Education Ministry Research Foundation (No. 09JK377).

■ REFERENCES

(1) Stöger, R.; Kubicka, P.; Liu, C. G.; Kafri, T.; Razin, A.; Cedar, H.; Barlow, D. P. *Cell* **1993**, *73*, 61−71.
(2) Swain, J. L.; Stewart, T. A.; Leder, P. *Cell* **1987**, *50*, 719−727.
(3) (a) Jones, P. A.; Takai, D. *Science* **2001**, *293*, 1068−1070. (b) Martienssen, R. A.; Colot, V. *Science* **2001**, *293*, 1070−1074. (c) Cogoni, C. *Annu. Rev. Microbiol.* **2001**, *55*, 381−406. (d) Paszkowski, J.; Whitham, S. *Curr. Opin. Plant Biol.* **2001**, *4*, 123−129.
(4) Feinberg, A. P.; Vogelstein, B. *Semin. Surg. Oncol.* **1987**, *3*, 149−151.
(5) Laird, C. D.; Jaffe, E.; Karpen, G.; Lamb, M.; Nelson, R. *Trends Genet.* **1987**, *3*, 274−281.
(6) Jacobsen, S. E.; Meyerowitz, E. M. *Science* **1997**, *277*, 1100−1103.
(7) (a) Esteller, M.; Garcia-Foncillas, J.; Andion, E.; Goodman, S. N.; Hidalgo, O. F.; Vanaclocha, V.; Baylin, S. B.; Herman, J. G. N. *Engl. J. Med.* **2000**, *343*, 1350−1354. (b) Robertson, K. D. *Oncogene* **2001**, *20*, 3139−3155.
(8) Frommer, M.; McDonald, L. E.; Millar, D. S.; Collis, C. M.; Watt, F.; Grigg, G. W.; Molloy, P. L.; Paul, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1827−1831.
(9) Clark, S. J.; Harrison, J.; Frommer, M. *Nat. Genet.* **1995**, *10*, 20−27.
(10) Stöger, R.; Kajimura, T. M.; Brown, W. T.; Laird, C. D. *Hum. Mol. Genet.* **1997**, *6*, 1791−1801.
(11) Genereux, D. P.; Miner, B. E; Bergstrom, C. T.; Laird, C. D. *Proc. Natl Acad. Sci. U.S.A.* **2005**, *102*, 5802−5807.
(12) Clark, S. J.; Harrison, J.; Paul, C. L.; Frommer, M. *Nucleic Acids Res.* **1994**, *22*, 2990−2997.
(13) Genereux, D. P.; Johnson, W. C.; Burden, A. F.; Stöger, R.; Laird, C. D. *Nucleic Acids Res.* **2008**, *36*, e150.
(14) Shiraishi, M.; Hayatsu, H. *DNA Res.* **2004**, *11*, 409−415.
(15) Hayatsu, H.; Tsuji, K.; Negishi, K. *Nucleic Acids Symp. Ser.* **2006**, *50*, 69−70.
(16) Hayatsu, H. *Genes Environ.* **2006**, *28*, 1−8.
(17) Huang, Y.; Pastor, W. A.; Shen, Y.; Tahiliani, M.; Liu, D. R.; Rao, A. *PLoS ONE* **2010**, *5*, e8888.
(18) Sedlàk, M.; Keder, R.; Skàla, P.; Hanusek, J. *J. Phys. Org. Chem.* **2005**, *18*, 743−750.
(19) Curran, T. C.; Farrar, C. R.; Niazy, O.; Williams, A. *J. Am. Chem. Soc.* **1980**, *102*, 6828−6837.
(20) Wang, B.; Cao, Z. *J. Phys. Chem. A* **2010**, *114*, 12918−12927.
(21) Chen, Z. q.; Xue, Y. *J. Phys. Chem. B* **2010**, *114*, 12641−12654.

(22) Labet, V.; Morell, C.; Cadet, J.; Eriksson, L. A.; Grand, A. *J. Phys. Chem. A* **2009**, *113*, 2524−2533.
(23) Meng, F. C.; Wang, H. J. *Theor. Chem. Acc.* **2010**, *127*, 561−571.
(24) Florián, J.; Baumruk, V.; Leszczyński, J. *J. Phys. Chem.* **1996**, *100*, 5578−5589.
(25) Purrello, R.; Molina, M.; Wang, Y.; Smulevich, G.; Fossella, J.; Fresco, J. R.; Spiro, T. G. *J. Am. Chem. Soc.* **1993**, *115*, 760−767.
(26) Lindhal, T.; Nyberg, B. *Biochemistry* **1974**, *13*, 3405−3410.
(27) Wang, R. Y.; Kuo, K. C.; Gehrke, C. W.; Huang, L. H.; Ehrlich, M. *Biochim. Biophys. Acta* **1982**, *697*, 371−377.
(28) Ehrlich, M.; Norris, K. F.; Wang, R. Y. H.; Kuo, K. C.; Gehrke, C. W. *Biosci. Rep.* **1986**, *6*, 387−393.
(29) Ehrlich, M.; Zhang, X. Y.; Inamdar, N. M. *Mutat. Res.* **1990**, *238*, 277−286.
(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C. et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(31) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(32) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639−5648.
(33) Gonzales, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154−2161.
(34) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822−2827.
(35) (a) Cancès, E. M. B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032−3041. (b) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151−5158.
(36) Toro-Labbé, A. *J. Phys. Chem. A* **1999**, *103*, 4398−4403.
(37) Herrera, B.; Toro-Labbé, A. *J. Phys. Chem. A* **2007**, *111*, 5921−5926.
(38) Bulat, F.; Toro-Labbé, A. *Chem. Phys. Lett.* **2002**, *354*, 508−517.
(39) Politzer, P.; Murray, J. S.; Lane, P.; Toro-Labbé, A. *Int. J. Quantum Chem.* **2007**, *107*, 2153−2157.
(40) Labet, V.; Morell, C.; Grand, A.; Toro-Labbe, A. *J. Phys. Chem. A* **2008**, *112*, 11487−11494.
(41) Chattaraj, P. K.; Maiti, B.; Sarkar, U. *J. Phys. Chem. A* **2003**, *107*, 4973−4975.
(42) (a) Parr, R. G.; Szentpály, L.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922−1924. (b) Maynard, A. T.; Huang., M.; Rice., W. G.; Covell., D. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11578−11583.
(43) (a) Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4049−4050. (b) Ayers, P. W.; Levy, M. *Theor. Chem. Acc.* **2000**, *103*, 353−360.
(44) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.
(45) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.
(46) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.
(47) Mezey, P. G.; Ladik, J. J.; Barry, M. *Theoret. Chim. Acta.* **1980**, *54*, 251−258.
(48) Janak, J. F. *Phys. Rev. B* **1978**, *18*, 7165−7168.
(49) Cohen, A. J.; Mori, S.; Yang, W. T. *Phys. Rev. B* **2008**, *77*, 115123−115126.
(50) Domingo, L. R.; Aurell, M. J. P.; Contreras, P. R. *J. Phys. Chem. A* **2002**, *106*, 6871−6875.