Hydrolytic Deamination of 5-Methylcytosine in Protic Medium—A Theoretical Study

Vanessa Labet,[†] Christophe Morell,[†] Jean Cadet,[†] Leif A. Eriksson,[‡] and André Grand*[,†]

Laboratoire "Lésions des Acides Nucléiques", INAC/SCIB—UMR-E n°3 CEA-UJF, CEA Grenoble, 17 rue des Martyrs, F-38 054 Grenoble cedex 9, France, and School of Science and Technology, and Örebro Life Science Center, Örebro University, 701 82 Örebro, Sweden

Received: October 8, 2008; Revised Manuscript Received: January 5, 2009

The mechanism for the deamination reaction of 5-methylcytosine with H_2O in protic medium was investigated using DFT calculations at the B3LYP/6-311G(d,p) level of theory. Two pathways were found. Pathway 5mA is a two-step mechanism where the N3-protonated 5-MeCyt undergoes a nucleophilic attack to carbon C4 by a water dimer before the elimination of an ammonium cation. Pathway 5mB is a three-step mechanism where neutral 5-MeCyt is directly attacked by a water dimer. The resulting intermediate is then protonated to allow the elimination of an ammonium cation. Both pathways lead to the formation of thymine in interaction with an ammonium cation and a water molecule. Pathway 5mA can explain the spontaneous deamination of 5-MeCyt in protic medium at acidic pH, whereas pathway 5mB is more representative of the deamination in protic medium at neutral pH. The nucleophilic addition of the water dimer is rate-determining in both pathways and is associated with an activation free energy in aqueous solution of 137.4 kJ/mol for pathway 5mA and 134.1 kJ/mol for pathway 5mB. This latter value is in agreement with the experimental observation that 5-MeCyt deaminates four- to fivefold faster than Cyt at neutral pH. Both electrostatic and electron-transfer contributions appear to have significant importance. In vacuum, the former one dominates when the substrate is positively charged and the latter one when it is neutral.

1. Introduction

Enzymatic methylation of cytosine (Cyt) residues giving rise to 5-methylcytosine (5-MeCyt) occurs mostly at CpG sequences of DNA, with a frequency of about 5% in human cells.[1,2] This constitutes a critical cellular event that plays a significant role in gene regulation[3] and defense against viral infection.[4,5] In addition, hypo- and hypermethylation reactions of DNA have been suggested to be involved in the initiation and progression steps of human cancers.[6−8] Spontaneous deamination of 5-methylcytosine in cellular DNA leads to the formation of thymine, which in turn forms a highly mutagenic G:T mismatch with the opposite guanine base.[9,10] The Vsr DNA endonuclease and several glycosylases, which have been identified in prokaryotic and/or eukaryotic cells,[11,12] are able to efficiently prevent the formation of C → T transition by repairing G:T mismatch to predominantly G:C base pairs.[13]

Detailed studies on hydrolytic deamination of 5-MeCyt and Cyt, either as free nucleotides or when inserted into DNA fragments, have been performed at neutral physiological pH's. The deamination mechanism of 5-MeCyt in neutral aqueous solutions is likely to involve protonation at N3 in the initial step, accompanied by an addition−elimination reaction leading to the hydrolytic loss of the exocyclic 4-amino group as also proposed for Cyt.[14] It has been shown as a general trend that the deamination rate of 5-MeCyt is about fivefold higher than that of Cyt in both nucleoside 5′-monophosphates[15] and single-stranded DNA[16−18] at 37 °C. The rate constant for spontaneous hydrolytic deamination of 5-MeCyt in double-stranded DNA at 37 °C and neutral pH was determined to be 5.8×10^{-13} s^{-1} using a sensitive genetic assay, twofold higher than that of cytosine.[19] It can be noted that these values are at least 3 orders of magnitude lower than those estimated for 5-MeCyt and Cyt in nucleotides and single-stranded DNA, showing a major protecting effect of stacking and base-pairing on hydrolytic deamination of 5-MeCyt and Cyt in double-stranded DNA. These rates, albeit low, are able to explain the mutation frequencies observed, as these are even lower due to the efficient repair of G:T mismatch. It was recently proposed that enzymatic deamination of 5-MeCyt mediated by DNA deaminases of Aid/Apobec could take place in pluripotent tissues,[20] and that these appear to be more efficient than the hydrolytic reaction.[20]

Saturation of the 5,6-double bond as the result of photohydration,[21] •OH-mediated 5,6-dihydroxylation,[22] or dimerization including formation of cyclobutane pyrimidine dimers and pyrimidine (6−4) pyrimidone photoproducts[23,24] leads to a significant enhancement of the deamination rate of the 5-MeCyt residue. However 5,6-dihydroderivatives of 5-MeCyt were found to display increased resistance toward hydrolytic deamination as compared to related Cyt compounds,[22,23] whereas an opposite trend is observed for the normal unsaturated bases.

The present study is aimed at gaining insight into the mechanism of hydrolytic deamination of 5-MeCyt using computational modeling and at trying to rationalize the fact that 5-MeCyt is more sensitive than Cyt toward spontaneous deamination. This constitutes a follow-up theoretical investigation of spontaneous deamination of Cyt[25] that was achieved at the PCM-corrected B3LYP/6-311G(d,p) level of theory, thereby allowing a comparison between the two pyrimidine aminobases. In the previous study,[25] two pathways (pathway A and pathway B) were proposed to explain the cytosine deamination in protic medium (Figure 1). The same two pathways will be investigated in the case of 5-MeCyt (pathway 5mA and 5mB) to allow the comparison in terms of energetic barriers. After section 2 where the different computational methods used in this theoretical study

* To whom correspondence should be addressed. Phone: +33 438 783 917. Fax: +33 438 785 090. E-mail: andre.grand@cea.fr.
[†] CEA Grenoble.
[‡] Örebro University.

10.1021/jp808902j CCC: $40.75
Published on Web 02/16/2009

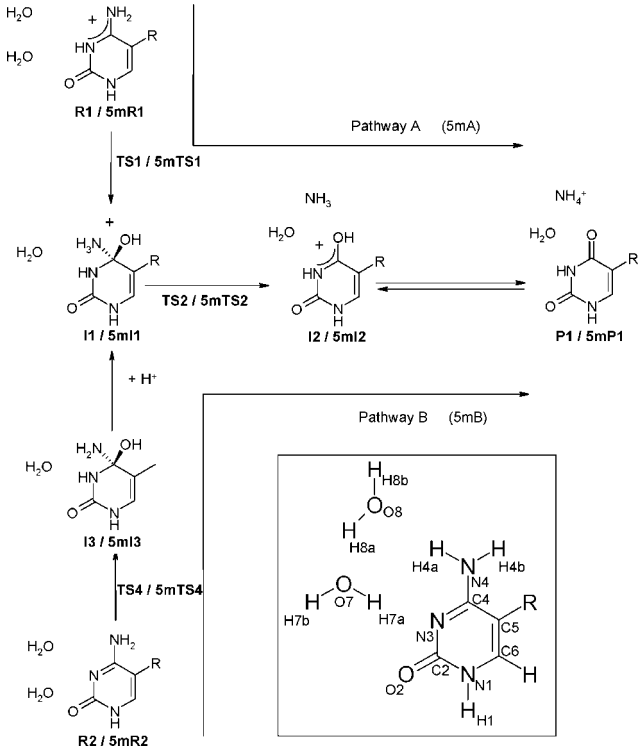

Figure 1. Reaction pathways studied in this work for the spontaneous deamination of Cyt (R = H, top labels) and 5-MeCyt (R = CH_3, bottom labels) in protic medium. Inset: atomic labeling used in the current study.

are specified, section 3 is devoted to the presentation and discussion of the results obtained for the reaction mechanism of 5-MeCyt deamination in protic medium. After a brief description of the energetics of pathways 5mA and 5mB (subsection 3.1) including a comparison with those of pathways A and B, the geometries involved in all cases are assessed in details (subsection 3.2). Subsection 3.3 focuses on a comparison of the electronic structure of the two pyrimidine aminobases and their N3-protonated forms in order to rationalize the results described in the two previous subsections. For this purpose, different properties such as partial charges on the nuclei, orbital energies, and reactivity indices derived from conceptual DFT are discussed. The theoretical background on these reactivity indices and similar is introduced in subsection 2.2.

2. Methodology

All the calculations were performed using the Gaussian 03 package[26] at the B3LYP/6-311G(d,p)[27−31] level.

2.1. Computational Investigation of Reaction Pathways. ***2.1.1. General Considerations.*** As a first step, the geometries of molecular systems of interest were optimized in vacuum. After full optimization, the different stationary points were characterized as either minima or transition states by computing the vibrational frequencies within the harmonic approximation at the same level of theory. Thermal data were extracted in order to obtain the different thermodynamic functions of reaction and the corresponding activation parameters at 298.15 K and 1 atm.

Energy calculations were performed on the optimized geometries including a polarized continuum model in the integral equation formalism (IEF-PCM; hereafter named PCM)[32−34] with dielectric constant $\epsilon = 78.39$ in order to simulate the bulk effect of an aqueous environment. If this value is clearly not appropriate to simulate the quite apolar environment inside

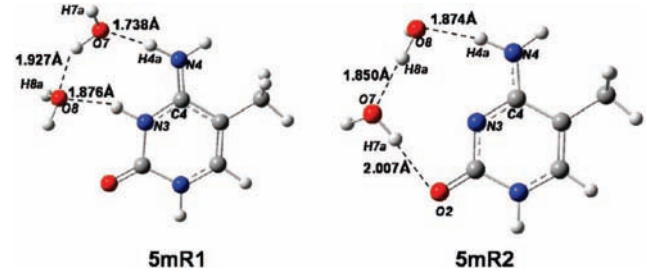

Figure 2. Geometries of the reactant complexes of pathways 5mA and 5mB, optimized in vacuum without constraints.

duplex DNA, it is adapted to reproduce the one of nucleoside 5′-monophosphates in aqueous solvent, for which was experimentally observed a difference of reactivity between cytosine and 5-methylcytosine toward deamination.[15] The cavity was built using the united atom topologic model applied on the atomic radii of the UFF force field, with an average area of 0.2 Å^2 for the tesserae generated on each sphere. The default cavity was modified by adding individual spheres on all hydrogen atoms linked to nitrogen and oxygen atoms, using the keyword SPHEREONH.

Reactions paths in vacuum were characterized at the same level of theory by performing intrinsic reaction coordinate (IRC) calculations. This was done from each gas phase optimized transition structure, to ensure that they are connected to the appropriate reactants and products.

2.1.2. The Reactive Forms **5mR1** ***and*** **5mR2.** For the reactant complexes **5mR1** and **5mR2** which involve electrostatic interactions between 5-MeCyt (or its N3-protonated form) and two water molecules, unconstrained geometry optimizations render hydrogen bond stabilized molecular complexes where the oxygen atoms of the two water molecules and 5-MeCyt (or its N3-protonated form) are coplanar, as shown in Figure 2.

Since both pathways 5mA and 5mB are initiated by a nucleophilic addition of a water molecule on carbon C4, these geometries for **5mR1** and **5mR2** are not realistic. Indeed, the attacking water molecule has to come from a direction perpendicular to the plane defined by the pyrimidine ring. Consequently, the optimization of the reactant complexes was achieved by imposing constraints in order to impose a correct relative orientation of the three molecules forming the complexes. This orientation was determined by studying the reaction force profiles along the IRC of the first elementary step for the two pathways, in the case of Cyt (Figure 3). The concept of reaction force that was introduced a few years ago by Toro-Labbé[35] is defined as

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

where ξ is the intrinsic reaction coordinate that represents the force acting on the system to bring the reactant(s) into the product(s). By definition, the reaction force vanishes at the three stationary points characterizing an elementary step, i.e., the reactant, the transition state, and the product. From Figures 3c and 3d, it can be seen that the reaction force is almost equal to zero in a region located before the transition states **TS1** and **TS4** but after **R1** and **R2** (indicated by a dashed vertical line in each graph). The relative orientations of the two water molecules at these particular points were then chosen for the constrained geometry optimization of the reactive complexes **5mR1** and **5mR2**. To distinguish them from the ones displayed in Figure 2, the notation **5mR1′** and **5mR2′** is used. Geometries for the constrained reactants are shown in Figure 4. It can be noted

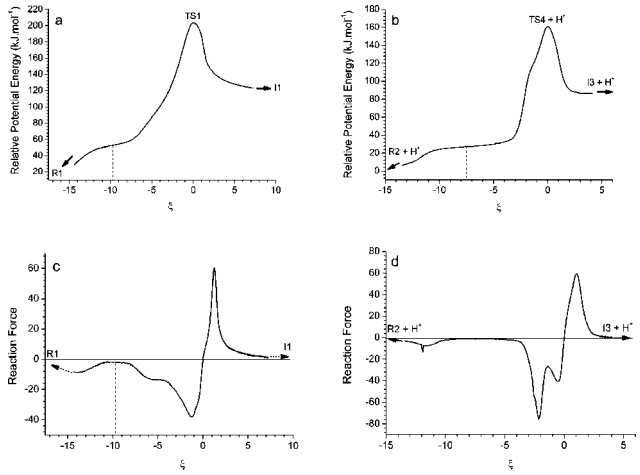

Figure 3. Potential energy (a and b) and reaction force (c and d) profiles along the **R1**−**TS1**−**I1** (a and c) and **R2**−**TS4**−**I3** (b and d) elementary steps in vacuum.

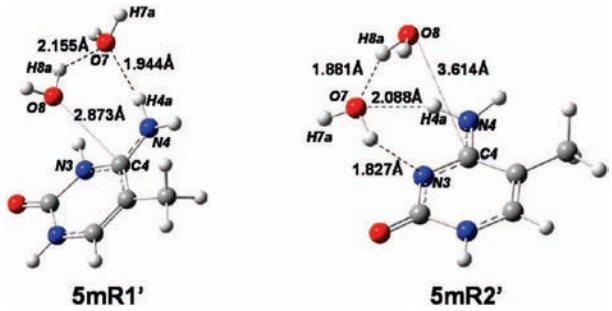

Figure 4. Optimized constrained geometries of the new reactant complexes in vacuum for pathways 5mA and 5mB.

that, as expected, the water molecule that will attack carbon C4 is not in the plane defined by the pyrimidine ring but well situated to add to C4 perpendicularly to the ring plane. In the following, the energies of the different stationary points involved in pathways 5mA and 5mB are calculated with respect to **5mR1′** and **5mR2′**.

2.2. Evaluation of Reactivity Indices. Several reactivity indices derived from conceptual DFT[36−38] and allowing a measure of electrophilicity of molecular systems[39,40] were used in this work to rationalize the difference of reactivity between cytosine and 5-methylcytosine and their N3-protonated forms toward their hydrolytic deamination. They were calculated from geometries optimized on one hand in vacuum and on the other hand in an aqueous solvent modelized as described in section 2.1. Some of the indices used in this work are global; others are local. Moreover, some are more accurate to describe electrostatic control (or charge control), whereas others are more adapted to describe an electron-transfer control (or frontier-orbital control).[41−44]

2.2.1. Definition and Computation of the Global Reactivity Indices Used in This Work. *2.2.1.a. Chemical Potential (μ).* Within the conceptual DFT, the chemical potential μ has been defined as[36,37,45]

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v(\vec{r})} \tag{2}$$

It measures the escaping tendency of the electronic cloud from equilibrium.[46] Consequently, a good electrophilic reactant must have a low chemical potential μ.

2.2.1.b. Absolute Hardness (η). The absolute hardness η, which can be viewed as the resistance of a molecular system to a charge transfer,[47−49] is defined as[36,37,45]

$$\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} \tag{3}$$

The smaller the absolute hardness, the bigger the electrophilicity of a molecular system.

2.2.1.c. Global Electrophilicity Index (ω). The global electrophilicity index ω is defined as[50,51]

$$\omega = \frac{\mu^2}{2\eta} \tag{4}$$

By analogy with the equation of power in classical electricity ($W = V^2/R$ with W the power, V the potential difference, and R the resistance), ω can be considered as the measure of the electrophilic power of a molecular system. Assuming that the chemical potential is often negative, a large electrophilic power is consistent with a low (very negative) chemical potential and a small absolute hardness, as already mentioned.

2.2.1.d. Computational Evaluation. μ, η, and ω were calculated using approximate expressions deriving from the use of a finite difference approximation and the Janak's theorem[52,53]

$$\mu \approx \frac{1}{2}(\varepsilon_{\mathrm{LUMO}} + \varepsilon_{\mathrm{HOMO}}) \tag{5}$$

$$\eta \approx (\varepsilon_{\mathrm{LUMO}} - \varepsilon_{\mathrm{HOMO}}) \tag{6}$$

$$\omega \approx \frac{1}{8}\frac{(\varepsilon_{\mathrm{LUMO}} + \varepsilon_{\mathrm{HOMO}})^2}{(\varepsilon_{\mathrm{LUMO}} - \varepsilon_{\mathrm{HOMO}})} \tag{7}$$

where $\varepsilon_{\mathrm{HOMO}}$ and $\varepsilon_{\mathrm{LUMO}}$ are the energies of the highest occupied (HOMO) and the lowest unoccupied (LUMO) molecular orbitals, respectively.

2.2.2. Definition and Computation of the Local Reactivity Indices Used in This Work. All the previous indices provide information on the reactivity of a molecular system as a whole. However, they are not relevant for assessing differences in reactivity within the system. For this, local indices were used: the net atomic charges to describe reactivity under electrostatic control, and three other indices more accurate to describe reactivity under electron-transfer control.

2.2.2.a. Net Atomic Charges. The net atomic charges obtained from the electrostatic potential provide a good description of the way in which the electron density distribution of a molecular system interacts with other molecules.

2.2.2.b. Local Electrophilicity Index ($\omega^+(\vec{r})$). The local electrophilicity index $\omega^+(\vec{r})$, which is a local version of the global electrophilicity index ω, is defined as[54]

$$\omega^+(\vec{r}) = \omega f^+(\vec{r}) \tag{8}$$

where $f^+(\vec{r})$ is the Fukui function used when the system undergoes a nucleophilic attack[55,56]

$$f^+(\vec{r}) = \left[\frac{\partial \rho(\vec{r})}{\partial N}\right]^+_v \tag{9}$$

with ρ the electronic density and N the number of electrons in the molecular system.

$f^+(\vec{r})$ measures the intramolecular reactivity at site $\vec{r}$ toward a nucleophilic reagent. The bigger $f^+(\vec{r})$, the more reactive the site will be with respect to other sites of the same molecule toward a nucleophilic attack. Since ω measures the global electrophilicity power, $\omega^+(\vec{r})$ takes into account both global electrophilicity and local intramolecular reactivity.

2.2.2.c. Excess Electrophilicity ($\Delta\omega(\vec{r})$). The excess electrophilicity $\Delta\omega(\vec{r})$ is defined as[57]

$$\Delta\omega(\vec{r}) = \omega\Delta f(\vec{r}) \tag{10}$$

where $\Delta f(\vec{r})$ is the dual descriptor,[58,59] a third-order response function[60] defined as

$$\Delta f(\vec{r}) = \left(\frac{\partial^2\rho(\vec{r})}{\partial N^2}\right)_v \tag{11}$$

If $\Delta f(\vec{r}) > 0$, the site is more electrophilic than nucleophilic, and the reverse if $\Delta f(\vec{r}) < 0$. $\Delta\omega(\vec{r})$, like $\omega^+(\vec{r})$, takes into account both global electrophilicity and local intramolecular reactivity and is consequently a good intermolecular descriptor. Since ω is always positive, the more positive $\Delta\omega(\vec{r})$, the more favored the site will be to undergo nucleophilic attack, whereas the more negative it is, the more favored the site will be to undergo electrophilic attack.

2.2.2.d. Another Local Reactivity Inex Derived from the Equivalent of the Dual Descriptor in the Grand Canonical Ensemble. Another reactivity indicator based on the dual descriptor $\Delta f(\vec{r})$ was introduced in a recent article[61] to allow gaining information on the relative reactivity of molecules with different sizes. This indicator is the equivalent of the $\Delta f(\vec{r})$ function in the grand canonical ensemble and can be expressed as

$$s^{(2)}(\vec{r}) = \frac{\Delta f(\vec{r})}{\eta^2} - \frac{f(\vec{r})}{\eta^3}\left\{\left(\frac{\partial\eta}{\partial N}\right)_{v(\vec{r})}\right\} \tag{12}$$

This descriptor, contrary to $\Delta\omega(\vec{r})$ and $\omega^+(\vec{r})$, is size extensive. Consequently, it scales correctly the philicity with respect to system size and can be used as an appropriate tool to compare the relative reactivity of Cyt and 5-MeCyt.

The more positive $s^{(2)}(\vec{r})$, the more favored a nucleophilic attack will be on that site. Assuming that the second term of the right-hand side of eq 12 is negligible with respect to the first one, the values of $\Delta f(\vec{r})/\eta^2$ were used in our study to compare the electrophilicity of Cyt and 5-MeCyt and the one of their N3-protonated forms.

2.2.2.e. Computational Evaluation. The partial charges on the nuclei have been computed with the ChelpG[62] method. $\omega^+(\vec{r})$, $\Delta\omega(\vec{r})$, and $\Delta f(\vec{r})/\eta^2$ were evaluated through their condensed-to-atom version where the atom of interest is carbon C4. To calculate these, the following formulas were used

$$\omega^+_{C4} = \omega f^+_{C4} \tag{13}$$

$$\Delta\omega_{C4} = \omega\Delta f_{C4} \tag{14}$$

$$f^+_{C4} = p_{C4}(N+1) - p_{C4}(N) \tag{15}$$

$$\Delta f_{C4} = p_{C4}(N+1) + p_{C4}(N-1) - 2p_{C4}(N) \tag{16}$$

where $p_{C4}(N+1)$, $p_{C4}(N-1)$, and $p_{C4}(N)$ are, respectively, the electronic population of carbon C4 when an electron is added to the molecular system, when an electron is removed, and in the normal system.

3. Results and Discussion

3.1. Energetics of Pathways 5mA and 5mB. Pathway 5mA is initiated by the nucleophilic addition of a first water molecule to carbon C4 of N3-protonated 5-MeCyt with the assistance of a second water molecule. This leads to the formation of the tetrahedral cationic intermediate **5mI1** via transition state **5mTS1**. The C4−N4 bond is then broken via transition state **5mTS2**, also including a simultaneous proton transfer from the hydroxyl group at C4 to NH_3, thus forming thymine interacting with one water molecule and an ammonium cation. Pathway 5mB can be described by a first nucleophilic addition of a water molecule to carbon C4 of neutral 5-MeCyt with the assistance of a second water molecule. This gives rise to the neutral tetrahedral intermediate **5mI3** via **5mTS4**. The exocyclic amino group of **5mI3** is protonated through an intermolecular proton transfer to form **5mI1** as in pathway 5mA. Pathway 5mB is subsequently obtained in a similar way as pathway 5mA. In Figure 5 we display the evolution of the free energy along the two pathways, in aqueous solvent.

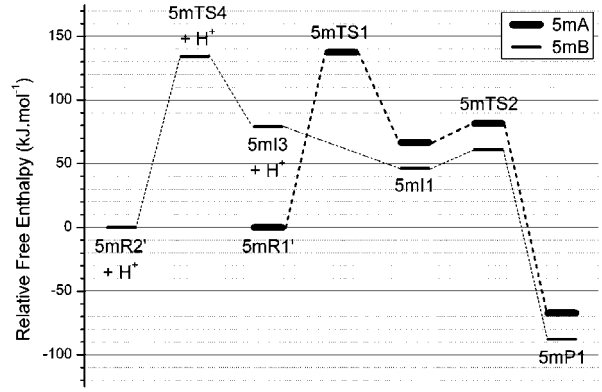

Figure 5. Relative free energies along pathways 5mA and 5mB in aqueous solvent. Values are expressed in kJ·mol^{-1}.

It is worth noting that no intermediate **5mI2** with H_2O, NH_3, and CytOH (analogous to that found for cytosine) could be optimized. All attempts gave **5mP1**, i.e., with the C4OH proton transferred to the leaving ammonia, including also optimization of the last point of the IRC calculation from **5mTS2**. This suggests that **5mI2** does not exist and that **5mI1** proceeds directly to **5mP1** via **5mTS2**. If existing, **5mI2** would be a van der Waals complex between a protonated form of thymine, an ammonia molecule, and a water molecule. The instability in vacuum of such complexes has already been noticed in our theoretical study dealing with the deamination of the radical cation of cytosine.[63]

From Figure 5 it can be concluded that deamination of 5-MeCyt follows the same pathways as cytosine in protic medium. For both pathways the first step, the nucleophilic addition to carbon C4, is the rate-determining step. The activation free energies are comparable even if pathway 5mB appears to be a little bit more efficient than pathway 5mA in aqueous solution.

In Table 1 we report the relative energies in vacuum, the free energies in vacuum, and the free energies in aqueous solvent, calculated for the different stationary points of the two pathways. The related values of Cyt and its N3-protonated form are also provided in order to allow a comparison. Perhaps it is worth noting that the activation free energies of pathways A and B in aqueous solvent ($\Delta G^{\ddagger}_{aq,TS1}$ = 136.0 kJ/mol and $\Delta G^{\ddagger}_{aq,TS4}$ = 138.5 kJ/mol) are quite close to the activation energy determined experimentally for the spontaneous deamination of cytidine 5′-monophosphate at 37 °C and pH = 7.8 (E_a = 121 kJ/mol).[15]

The main observations made in the case of Cyt are also valid for 5-MeCyt. In the two pathways 5mA and 5mB, the nucleophilic addition to carbon C4 is the rate-determining step, both in vacuum and in aqueous solution. In vacuum, the nucleophilic addition to carbon C4 of 5-MeCyt is easier than to the N3-protonated form ($\Delta G^{\ddagger}_{vacuum,5mTS4}$ = 136.3 kJ/mol and $\Delta G^{\ddagger}_{vacuum,5mTS1}$ = 169.7 kJ/mol). This is also true in aqueous solution ($\Delta G^{\ddagger}_{aq,5mTS1}$ = 137.4 kJ/mol and $\Delta G^{\ddagger}_{aq,5mTS4}$ = 134.1 kJ/mol), whereas the trend was reversed in the case of Cyt and

TABLE 1: Relative Energies in Vacuum (ΔE_{vac}), Free Energies in Vacuum (ΔG_{vac}), and Free Energies in Aqueous Solution (ΔG_{aq}) for the Stationary Points of Pathways 5mA, 5mB, A, and B, at $T = 298.15$ K and $P = 1$ atm[a]

pathway	system	ΔE_{vac}	ΔG_{vac}	ΔG_{aq}
5mA	**5mR1′**	0.0	0.0	0.0
(A)	(**R1′**)	(0.0)	(0.0)	(0.0)
5mA	**5mTS1**	150.6	169.7	137.4
(A)	(**TS1**)	(145.1)	(163.8)	(136.0)
5mA	**5mI1**	84.2	93.1	66.7
(A)	(**I1**)	(77.9)	(87.9)	(64.2)
5mA	**5mTS2**	89.2	96.8	81.5
(A)	(**TS2**)	(84.9)	(93.9)	(82.3)
5mA	**5mI2**	–	–	–
(A)	(**I2**)	(36.8)	(20.0)	(35.0)
5mA	**5mP1**	−39.7	−44.1	−67.3
(A)	(**P1**)	(−43.2)	(−47.8)	(−65.2)
5mB	**5mR2′** + H^+	0.0	0.0	0.0
(B)	(**R2′** + H^+)	(0.0)	(0.0)	(0.0)
5mB	**5mTS4** + H^+	115.3	136.3	134.1
(B)	(**TS4** + H^+)	(121.4)	(142.9)	(138.5)
5mB	**5mI3** + H^+	62.4	75.7	79.2
(B)	(**I3**+ H^+)	(62.1)	(74.9)	(79.3)
5mB	**5mI1**	−908.5	−867.0	46.4
(B)	(**I1**)	(−908.7)	(−865.6)	(44.0)
5mB	**5mTS2**	−903.4	−863.3	61.2
(B)	(**TS2**)	(−901.6)	(−859.6)	(62.1)
5mB	**5mI2**	–	–	–
(B)	(**I2**)	(−949.7)	(−933.5)	(14.8)
5mB	**5mP1**	−1032.4	−1004.3	−87.6
(B)	(**P1**)	(−1029.7)	(−1001.3)	(−85.4)

[a] All values are expressed in kJ/mol.

its N3-protonated form ($\Delta G^{\ddagger}_{aq,TS1} = 136.0$ kJ/mol and $\Delta G^{\ddagger}_{aq,TS4} = 138.5$ kJ/mol).

It appears that according to pathway 5mA, deamination of 5-MeCyt is more difficult than that of Cyt in both vacuum ($\Delta G^{\ddagger}_{vacuum,5mTS1} = 169.7$ kJ/mol and $\Delta G^{\ddagger}_{vacuum,TS1} = 163.8$ kJ/mol) and aqueous solvent ($\Delta G^{\ddagger}_{aq,5mTS1} = 137.4$ kJ/mol and $\Delta G^{\ddagger}_{aq,TS1} = 136.0$ kJ/mol). The opposite is seen in pathways 5mB ($\Delta G^{\ddagger}_{vacuum,5mTS4} = 136.3$ kJ/mol and $\Delta G^{\ddagger}_{vacuum,TS4} = 142.9$ kJ/mol; $\Delta G^{\ddagger}_{aq,5mTS4} = 134.1$ kJ/mol and $\Delta G^{\ddagger}_{aq,TS4} = 138.5$ kJ/mol). As already mentioned in the Introduction, experimental results[15] have shown that the deamination rate of 5-MeCyt is about fivefold higher than that of Cyt in related nucleoside 5′-monophosphates at 37 °C and pH = 7.8. In DNA, it is quite difficult to know the pK_A of Cyt and 5-MeCyt[64] and consequently to determine the predominant form of the nucleobases between neutral and N3-protonated. Nevertheless, in nucleoside 5′-monophosphates, it is quite clear that the pK_A of Cyt is about 4.2 whereas that of 5-MeCyt is around 4.6. Consequently, in the case of nucleoside 5′-monophosphates, at pH = 7.8, the proportion of N3-protonated Cyt or 5-MeCyt will be small, and hence pathways A and 5mA have a much smaller probability to occur than pathways B and 5mB where a water molecule is added to the neutral Cyt and 5-MeCyt. The activation free energy associated with the 5-methylcytosine **5mR2′**−**5mTS4**−**5mI3** step is in aqueous solvent about 4.4 kJ/mol smaller than that associated with the cytosine **R2′**−**TS4**−**I3** step. This difference in the energetic barriers is in agreement with the difference in the rate constants inferred from the equation given by the activated complex theory[65]

$$k = \frac{k_B T}{h} \exp\left(\frac{-\Delta G^{\ddagger}}{RT}\right) \qquad (17)$$

Solvation effects make pathways 5mA and A easier, whereas it has essentially no effect on pathways 5mB and B. Consequently, in aqueous solution, pathway A appears to be more favorable than pathway B ($\Delta G^{\ddagger}_{aq,TS1} = 136.0$ kJ/mol and $\Delta G^{\ddagger}_{aq,TS4} = 138.5$ kJ/mol) whereas pathway 5mB remains more favorable than pathway 5mA ($\Delta G^{\ddagger}_{aq,5mTS1} = 137.4$ kJ/mol and $\Delta G^{\ddagger}_{aq,5mTS4} = 134.1$ kJ/mol). It would be interesting to explore experimentally whether deamination of 5-MeCyt is faster at acidic pH than at neutral pH, as observed for Cyt. Indeed, whereas our results are supportive of the existence of acidic catalysis for the deamination of Cyt, they are predictive of a lack of the latter process in the case of 5-MeCyt.

TABLE 2: Dipole Moments (μ, in D) of the Stationary Points Involved in Pathways 5mA, 5mB, A, and B in Vacuum

system	μ	system	μ
5mR1′	5.0	**R1′**	5.2
5mTS1	7.1	**TS1**	7.4
5mI1	6.1	**I1**	6.4
5mTS2	5.8	**TS2**	6.0
5mI2	–	**I2**	3.6
5mP1	10.8	**P1**	11.3
5mR2′	5.1	**R2′**	4.8
5mTS4	5.0	**TS4**	4.9
5mI3	3.9	**I3**	4.0

The influence of solvation on the energetic barriers can be explained by the evolution of the dipole moments along the different pathways (see Table 2). **TS1** and **5mTS1** are more polar than **R1′** and **5mR1′**, respectively. This can explain why pathways A and 5mA are associated with lower barriers in aqueous solvent than in vacuum. On the contrary, **TS4** and **5mTS4** have dipole moments similar to those of **R2′** and **5mR2′**, and thus solvation has essentially no effect on the energetic barriers associated with the first step of pathways B and 5mB.

It may be noted that **TS2** and **5mTS2** have smaller dipole moments than **I1** and **5mI1,** respectively. Hence, the removal of NH_3 is more difficult in aqueous solvent than in vacuum, as seen in Table 1.

For a better understanding of the effect of C5-methylation on the energetic barriers, the geometries of the stationary points involved in the two kinds of pathways have been studied in more detail. Emphasis was placed on the geometries of the transition states involved in the first, rate-determining, steps.

3.2. Geometries of the Stationary Points. ***3.2.1. Comparison of the Transition States Involved in the Rate-Determining Steps of Pathway A and Pathway 5mA.*** In the first step of pathways 5mA and A, a single bond is created between carbon C4 of the N3-protonated 5-methylcytosine and oxygen O8 of water, a proton (H8a) is transferred between the two water molecules, and a second one (H7a) is transferrd from O7 of the second water molecule to N4 (Figure 6).

Figure 6 shows the optimized geometries of the transition states and the products associated with this step, i.e.,**TS1** and **I1** of pathway A, and **5mTS1** and **5mI1** of pathway 5mA, respectively.

The two transition states have very similar geometries. A six-membered ring is formed by carbon C4, oxygen O8, hydrogen H8a, oxygen O7, hydrogen H7a, and nitrogen N4. Carbon C4 shows a pronounced tetrahedral character. The two proton transfers are asynchronous, and H8a proton transfer from O8 to O7 has proceeded further than a H7a one from O7 to N4, giving more of a "H_3O^+" character at O7.

When the substrate is C5-methylated, the distance between carbon C4 and oxygen O8 is longer by 0.010 Å. Consequently, the C4−O8 bond formation is a bit less advanced in the C5-methylated transition state than in its nonmethylated counterpart.

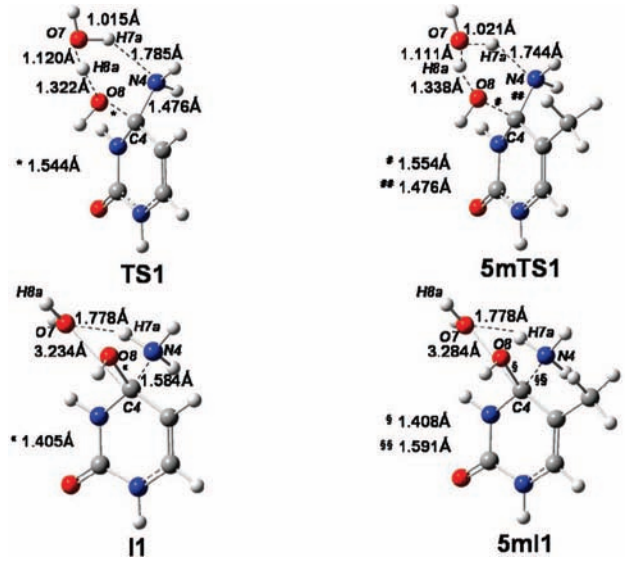

Figure 6. Optimized geometries of **TS1**, **5mTS1**, **I1**, and **5mI1** in vacuum.

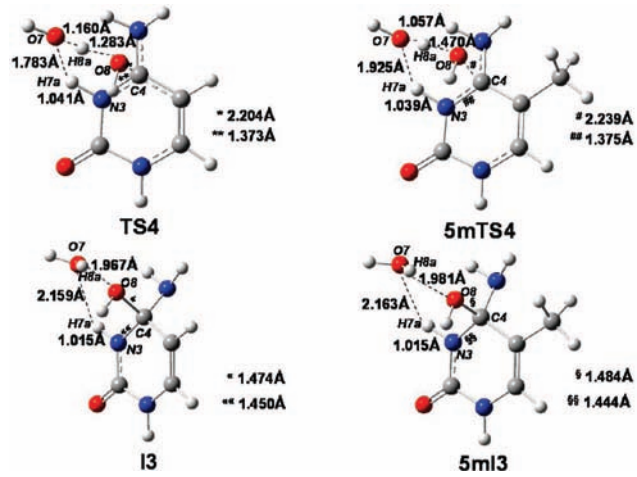

Figure 7. Optimized geometries of **TS4**, **5mTS4**, **I3**, and **5mI3** in vacuum.

On the contrary, the H8a proton is transferred further when C5 is methylated. The H7a proton transfer has progressed slightly more when cytosine is C5-methylated (d_{H7a-O7} = 1.015 Å in **TS1** and d_{H7a-O7} = 1.021 Å in **5mTS1**). Hence, it appears that the C5-methylation delays the C4−O8 bond creation and promotes the two proton transfers. A consequence is that the nucleophilic agent becomes more resemblant of a hydroxyl anion in the case of the 5-MeCyt. IRC calculations from **TS1** and **5mTS1** show that the C4−O8 bond creation begins before the two proton transfers. Consequently, the concerted mechanism associated with this elementary step is more synchronous when C5 is methylated. Intuitively, it can be assumed that the corresponding energetic barrier is narrower but higher. The values reported in Table 1 shows that, effectively, the free energy barrier associated with **5mTS1** is indeed higher than the one associated with **TS1** (by 5.9 kJ/mol in vacuum and by 1.4 kJ/mol in aqueous solvent) and the associated imaginary frequency of **5mTS1** is about 8 cm^{-1} higher than that of **TS1** (ν_{5mTS1} = 457.0 icm^{-1}; ν_{TS1} = 449.4 icm^{-1}). It is well documented that the imaginary frequency is related to the barrier width: the higher the frequency, the narrower the energetic barrier. This is all consistent with the fact that the deamination of the N3-protonated 5-MeCyt is more synchronous than that of the N3-protonated Cyt.

3.2.2. Comparison of the Transition States Involved in the Rate-Determining Steps of Pathway B and Pathway 5mB. As in the case of pathways A and 5mA, the first step of pathways B and 5mB also involves two proton transfers and the creation of one C−O single bond. In this case, one proton (H7a) is transferred from O7 to N3, a second proton (H8a) is transferred from O8 to O7, and a single bond is created between carbon C4 of the cytosine derivative and oxygen O8. In Figure 7 we present the optimized geometries of the transition states and the products associated with this elementary step, in the case of both Cyt and 5-MeCyt, i.e., **TS4** and **5mTS4**, and **I3** and **5mI3**.

The geometries of **TS4** and **5mTS4** are very similar, as were those of **TS1** and **5mTS1**. Both cases involve six-center transition states (N3, H7a, O7, H8a, O8, and C4). The H7a proton transfer from O7 to N3 appears much later than that for H8a. The tetrahedral character of carbon C4 in **TS4** and **5mTS4** is less pronounced than in **TS1** and **5mTS1**. Moreover, the distances between carbon C4 and oxygen O8 are much longer than those involved in **TS1** and **5mTS1**. The IRC calculation from **TS4** shows that when the substrate is neutral, the two proton transfers begin before the C4−O8 bond creation, whereas when the substrate is N3-protonated, it is the C4−O8 bond creation which induces the two proton transfers. This is in agreement with the fact that the N3-protonation increases the aminobase electrophilic power. When the substrate is unprotonated, carbon C4 is not electrophilic enough to undergo a nucleophilic attack; instead N3 needs to first become protonated. On the contrary, when the substrate is already N3-protonated, carbon C4 is sufficiently electrophilic to directly undergo nucleophilic addition.

In **5mTS4** the distance between carbon C4 and oxygen O8 is longer than in **TS4** (d_{C4-O8} = 2.204 Å in **TS4**, d_{C4-O8} = 2.239 Å in **5mTS4**). Consequently, the C4−O8 single bond creation is less advanced when the substrate is C5-methylated, in comparison with Cyt. The H7a proton transfer from O7 to N3 is almost finished at the transition states since the distance between H7a and N3 is 1.041 Å in **TS4** and 1.039 Å in **5mTS4**. The H8a proton transfer from O8 to O7 is much more advanced when C5 is methylated than when it is not; the distance between H8a and O8 is 1.283 Å in **TS4** and 1.410 Å in **5mTS4**. As already observed in the previous section, the C5-methylation delays the C4−O8 bond creation and brings forward the proton transfers. As previously noted, along the elementary step, the reactions move toward the creation of a better nucleophilic agent (closer to a "hydroxyl anion") on carbon C4. However, in the case of pathway 5mB, since the C4−O8 bond creation is the last of the three events, C5-methylation makes the H8a proton transfer and the C4−O8 bond creation less synchronous. Contrary to what was suggested in the case of pathways A, which appear to be more synchronous in the case of 5-MeCyt, in pathways B the mechanism is made less synchronous, and thus the C5-methylation must render a wider but lower energetic barrier to cross. Values that are reported in Table 1 confirm that the barrier is less high for 5-MeCyt (by 6.6 kJ/mol in vacuum and 4.4 kJ/mol in aqueous solvent). The imaginary frequency associated with **5mTS4** is 71.6 cm^{-1} smaller than that associated with **TS4** (ν_{5mTS4} = 633.1 icm^{-1}; ν_{TS1} = 704.7 icm^{-1}), which is consistent with the barrier being wider.

3.2.3. Comparison of the Transition States Involved in the NH_3 Departure. After the nucleophilic addition of one water molecule to C4 with the assistance of a second water molecule, the C4−N4 bond has to be broken in order to allow deamination of the substrates. As shown in Table 1, this elementary step

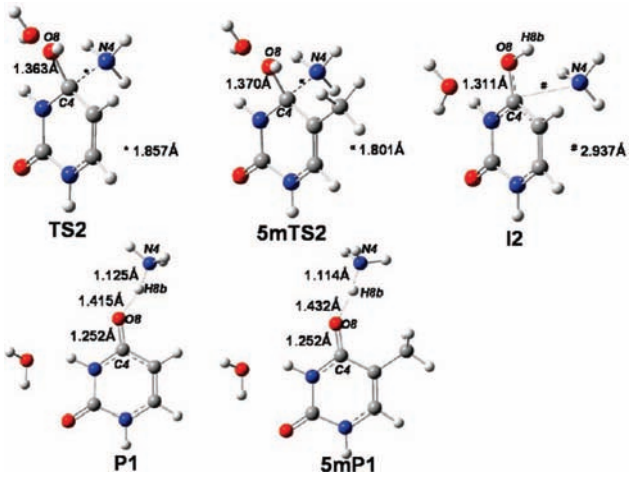

Figure 8. Optimized geometries of **TS2**, **5mTS2**, **I2**, **P1**, and **5mP1** in vacuum.

involves a very low free energy barrier ($\Delta G^{\ddagger}_{vacuum,5mTS2} = 3.7$ kJ/mol and $\Delta G^{\ddagger}_{vacuum,TS2} = 6.0$ kJ/mol; $\Delta G^{\ddagger}_{aq,5mTS2} = 14.8$ kJ/mol and $\Delta G^{\ddagger}_{aq,TS2} = 18.1$ kJ/mol), and it can be assumed that the bond breaking occurs more or less automatically. The fact that the step is more difficult in aqueous solvent than in vacuum is attributed to the dipole moments of **TS2** and **5mTS2,** which are smaller than those of **I1** and **5mI1**, respectively (see Table 2).

In Figure 8 we show the transition sates associated with this C4−N4 bond breakage and their products, i.e., **TS2**, **5mTS2**, **I2**, **P1**, and **5mP1**. The C5-methylation tends to reduce the distance between the two atoms at the transition state by about 0.05 Å, probably because of the steric repulsion between the C5-methyl group and the NH_3 substituent.

3.3. Comparison of the Electronic Properties of Cyt, 5-MeCyt, and Their N3-Protonated Forms. Since the rate-determining step in the spontaneous deamination of Cyt and 5-MeCyt is the nucleophilic addition to carbon C4, it may be assumed that the difference of reactivity between the two pyrimidine bases is accounted for by a difference in their electrophilic power.

3.3.1. Global Reactivity Indices. In the frontier molecular orbital theory, along a nucleophilic addition which is under frontier-orbital control, the electrophilic reagent is known to react with its lowest unoccupied molecular orbital (LUMO) and the nucleophile with its highest occupied molecular orbital (HOMO). The closer the two frontier orbitals, the easier the reaction will be. Assuming that the LUMO of the electrophile is higher than the HOMO of the nucleophile, the smaller the gap between these, the easier the reaction will be. In Table 3 we report the LUMO and HOMO energies for all the cytosine derivatives studied in this section, both in vacuum (section a) and in aqueous solvent (section b).

C5-methylation induces a very small destabilization (0.1−0.4 eV) of the HOMO but has essentially no effect on the LUMO energy. The LUMO energy hence does not appear to explain the difference in hydrolysis rate between Cyt and 5-MeCyt.

N3-protonation leads to a stabilization of the two frontier orbitals. This stabilization is very significant (about 5 eV) when the molecules are in vacuum but is more moderate when they are in aqueous solution (about 1 eV). In the N3-protonated forms of Cyt and 5-MeCyt, since there is the same number of electrons but more nuclei than in the nonprotonated forms, the repulsion between the electrons is less important and the attraction between electrons and nuclei more stabilizing. A consequence should be that N3-protonated Cyt and N3-protonated 5-MeCyt are more prone to undergo a nucleophilic attack than Cyt and 5-MeCyt, respectively. Yet, in vacuum, pathways 5mA and A are more difficult than pathways 5mB and B, respectively (see Table 1). Thus, the frontier molecular orbital theory approach appears to fail.

TABLE 3: Kohn−Sham Frontier Orbital Energies (ϵ_{LUMO} and ϵ_{HOMO}), Chemical Potential (μ), Absolute Hardness (η), and Global Electrophilicity Index (ω) of Cyt, 5-MeCyt, N3-Protonated Cyt, N3-Protonated 5-MeCyt, and a Water Molecule in Vacuum (a) and in Aqueous Solvent (b)[a]

molecular system	ϵ_{LUMO}	ϵ_{HOMO}	μ	η	ω
	a				
Cyt	−1.03	−6.39	−3.71	5.36	1.29
5-MeCyt	−0.95	−6.17	−3.56	5.22	1.22
N3-protonated Cyt	−6.64	−11.86	−9.25	5.22	8.19
N3-protonated 5-MeCyt	−6.45	−11.42	−8.94	4.98	8.02
H_2O	1.60	−8.02	−3.21	9.63	0.53
	b				
Cyt	−0.95	−6.53	−3.74	5.58	1.26
5-MeCyt	−0.93	−6.34	−3.63	5.41	1.22
N3-protonated Cyt	−2.18	−7.53	−4.86	5.36	2.20
N3-protonated 5-MeCyt	−2.18	−7.34	−4.76	5.17	2.19
H_2O	1.14	−8.30	−3.58	9.44	0.68

[a] All values are expressed in eV.

In conceptual density functional theory, one reasons in terms of chemical potential and absolute hardness instead of frontier orbital energies. The modifications in the frontier orbital energies mentioned above are correlated with modifications in the values of chemical potential, absolute hardness, and global electrophilicity index, as inferred from the values reported in Table 3.

C5-methylation tends to increase the chemical potential by about 0.2−0.3 eV in vacuum and 0.1 eV in aqueous solvent. Simultaneously, the absolute hardness decreases somewhat. These trends are a consequence of the HOMO destabilization caused by the C5-methylation and result in a smaller value of the global electrophilicity index for these compounds. Indeed, although the two C5-methylated forms have a smaller absolute hardness than their counterparts, and consequently a smaller resistance to an electronic transfer, they show a higher chemical potential and appear to be less susceptible to gain electrons. This can be explained by the electron-donating character of the methyl group, which makes the ring systems of 5-MeCyt and N3-protonated 5-MeCyt more electron rich than Cyt and N3-protonated Cyt.

N3-protonation lowers the chemical potential by about 5.5 eV in vacuum and 1.1 eV in aqueous solvent. This effect is much more important than that induced by C5-methylation. In addition, the absolute hardness decreases by about 0.2 eV. Consequently, the N3-protonated forms will as expected have a higher electrophilic power than their neutral counterparts.

It is worth noting that C5-methylation and N3-protonation have opposite effects on the chemical potential and similar effects on the absolute hardness. The consequence is that among Cyt, 5-MeCyt, and their N3-protonated forms, 5-MeCyt is the system whose chemical potential is the highest but the closest to that of a water molecule, whereas N3-protonated Cyt is the lowest but the furthest from water. Moreover, Cyt is the hardest of the four and N3-protonated 5-MeCyt the softest. Chemical potential differences drive electron transfer, but when two systems with very different chemical potential interact, the reaction proceeds under charge control. Thus, pathways 5mA and A are more likely to be under charge control than pathways

TABLE 4: Partial Charges on C4 (q_{C4}), Local Electrophilicity Index (ω^+_{C4}, eV), Excess Electrophilicity ($\Delta\omega_{C4}$, eV), and Index Derived from the Grand Canonical Dual Descriptor ($\Delta f_{C4}/\eta^2$, $10^{-2}\cdot eV^{-2}$) Condensed to Carbon C4 for Cyt, 5-MeCyt, N3-Protonated Cyt, and N3-Protonated 5-MeCyt, in Vacuum (a) and in Aqueous Solvent (b)

molecular system	q_{C4}	ω^+_{C4}	$\Delta\omega_{C4}$	$\Delta f_{C4}/\eta^2$
	a			
Cyt	0.91	0.27	0.27	0.71
5-MeCyt	0.74	0.28	0.28	0.84
N3-protonated Cyt	0.85	2.57	2.64	1.18
N3-protonated 5-MeCyt	0.67	2.46	2.65	1.33
	b			
Cyt	1.07	0.34	0.32	0.81
5-MeCyt	0.87	0.39	0.34	0.95
N3-protonated Cyt	0.90	0.72	0.67	1.06
N3-protonated 5-MeCyt	0.68	0.77	0.71	1.22

5mB and B, which are probably under electron-transfer control. Moreover, electrons must be hard to transfer between two systems with very different chemical potential levels. This could explain why in vacuum pathways A and 5mA are more difficult than pathways B and 5mB, respectively.

Solvation has no effect on the chemical potential of the neutral molecules, but raises it by more than 4 eV for the cationic forms. Moreover, for all systems, solvation tends to increase the absolute hardness by about 0.2 eV. This leads to the result that the electrophilic power of the neutral molecules is nearly the same in vacuum as in aqueous solvent, whereas that of the cationic molecules is considerably reduced in aqueous solvent. This is consistent with the fact that electric charges are more stabilized in aqueous solvent than in vacuum. Since the chemical potential of N3-protonated Cyt and N3-protonated 5-MeCyt is considerably increased in aqueous solution, it is closer to that of a water molecule. Consequently, in aqueous solution, pathways A and 5mA may proceed no longer under charge control, but under electron transfer control. This could explain the fact that pathways A and 5mA are easier in aqueous solution than in vacuum.

It is interesting to note that in the case of the deamination reaction, the global electrophilicity index does not appear to be a good indicator of reactivity. This is due to the fact that depending on the substrate the reaction can proceed under charge control or under electron-transfer control.

3.3.2. Local Reactivity Indices. In Table 4 we report the electrostatic partial charges on C4 in vacuum and in aqueous solvent for cytosine, 5-methylcytosine, and their N3-protonated forms, which should be good indicators to describe an electrostatic control.

Surprisingly, N3-protonation decreases the positive charge held by carbon C4 by 0.06−0.07 e in vacuum and 0.17−0.19 e in aqueous solution. As for C5-methylation, it also induces a decrease of the positive charge held by carbon C4, both when the aminobases are N3-protonated (−(0.18−0.22) e) or not (−(0.17−0.20) e). Consequently, according to a charge control model, both C5-methylation and N3-protonation make nucleophilic attack more difficult.

In Table 4 we also report the values of the different reactivity descriptors derived from the Fukui function described in subsection 2.2.2.b.

The values of ω^+_{C4}, $\Delta\omega_{C4}$, and $\Delta f_{C4}/\eta^2$ all indicate that carbon C4 of 5-MeCyt and N3-protonated 5-MeCyt are more electrophilic than that of Cyt and N3-protonated Cyt. As for N3-protonated Cyt and N3-protonated 5-MeCyt, their C4 carbon appears to have a better electrophilic power than that of their neutral counterparts. $\Delta f_{C4}/\eta^2$ is probably among these three indicators the one which is the most accurate in providing information concerning an electron-transfer control. Under frontier-orbital control, both C5-methylation and N3-protonation make it easier to undergo nucleophilic attack. This is exactly the reverse for a charge control.

Because of its chemical potential level in vacuum, the deamination of N3-protonated Cyt is probably under charge control. The C5-methylation, by decreasing the positive charge held by C4, renders the deamination reaction more difficult. That is in agreement with the fact that pathway 5mA is less favored than pathway A (see Table 1). On the contrary, the deamination of Cyt is probably under frontier-orbital control. The C5-methylation, by increasing the value of $\Delta f_{C4}/\eta^2$, makes the nucleophilic attack to C4 easier, explaining why pathway 5mB is favored with respect to pathway B (see Table 1).

4. Conclusion

In this work, two pathways proposed in a recent article[25] for the spontaneous deamination of Cyt in protic medium have been investigated in the case of 5-MeCyt. In pathway 5mA, the N3-protonated form of 5-MeCyt undergoes nucleophilic addition to carbon C4 by a water molecule with the assistance of a second water. In pathway 5mB, the neutral 5-MeCyt is instead attacked. The two pathways appear feasible for 5-MeCyt and have comparable activation parameters. Since N3-protonated 5-MeCyt molecules are predominant over neutral 5-MeCyt ones at acidic pH, spontaneous deamination under these conditions may be accounted for by pathway 5mA. In contrast since there are more 5-MeCyt molecules than the N3-protonated ones at neutral pH, deamination is better rationalized in term of pathway 5mB.

This work shows that pathways 5mB and B are in agreement with the experimental observation that at pH = 7.4 5-MeCyt deaminates four- to fivefold faster than Cyt. It may be also be predicted from the comparison of pathways 5mA and A that at acidic pH deamination of 5-MeCyt is slower than that of Cyt. Moreover, pathway 5mB appears to be more operative than pathway 5mA, which would mean that contrary to cytosine, there would be no acid catalysis effect for spontaneous deamination of 5-MeCyt. It would be interesting to challenge this hypothesis experimentally.

From the evaluation of the electrostatic partial charges on the nuclei, and other local descriptors such as ω^+_{C4}, $\Delta\omega_{C4}$, and $\Delta f_{C4}/\eta^2$, it seems that the C5-methylation decreases the positive charge on carbon C4 but increases its electrophilic power. The relative rates of reaction of pathways A and 5mA and B and 5mB seem to indicate that in the deamination reaction both electrostatic and electron-transfer contributions are important. In vacuum, the former dominates when the cytosine derivative is N3-protonated and the latter one when it is not. It seems to be due largely to the chemical potential level. These results have to be taken with precaution. Indeed, the functional B3LYP used in this study systematically underestimates reaction barriers and tends to favor electron delocalization over charge localization.[66] As a consequence, if a mechanism is charge-controlled, it is possible that it appears orbital-controlled.

The comparison of the transitions states involved in the rate-determining steps shows that the two proton transfers involved in the nucleophilic addition steps are more advanced and the C−O bond creation delayed when the substrates are C5-methylated. The consequence is the formation of a

nucleophilic agent more resemblant of a hydroxyl anion. This phenomenon has opposite consequences in pathways 5mA and 5mB. In pathway 5mA, where the proton transfers start after the C−O bond creation, the mechanism is more synchronous and the energetic barrier becomes narrower and higher. In pathway 5mB, where the proton transfers begin before the C−O bond creation, the mechanism is more asynchronous and the energetic barrier thus wider and lower.

From a biological point of view, one of the main striking results of the present DFT computational study which is in agreement with previous experimental data is the higher rate of deamination displayed by the relatively minor 5-MeCyt residues with respect to Cyt bases.[16−19] This provides a relevant molecular basis for the occurrence of mutagenic hot spots at CpG islands in eukaryotic genes despite the existence of efficient repair of G:T mismatches by dedicated and specific thymine DNA *N*-glycosylases in the DNA of mammalian cells. The C to T transitions thus generated in human genome have been shown to represent about 20% of all base substitutions.[67] Deficiency in the repair activity of MBD4, one of the two proteins involved in the removal of G:T mismatches in mice, has been shown to increase by threefold the number of deaminated 5-MeCyt sites that was correlated with a higher incidence of tumorigenesis.[68] Photomodification,[69] chlorination, or radical oxidation[70] of 5-MeCyt has been suggested to lead to heritable changes in methylation patterns that may interfere with normal epigenic control and favor human cancer promotion.

Acknowledgment. The Swedish Research Council (VR.) and the Faculty of Science and Technology at Örebro University are gratefully acknowledged for financial support (L.A.E.).

Supporting Information Available: Molecular modeling coordinates of all stationary points involved in pathways A, 5mA, B, and 5mB. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Barlow, D. P. *Science* **1993**, *260*, 309–310.
(2) Verdine, G. L. *Cell* **1994**, *76*, 197–200.
(3) Cedar, H. *Cell* **1988**, *53*, 3–4.
(4) Groudine, M.; Eisenman, R.; Weintraub, H. *Nature* **1981**, *292*, 311–317.
(5) Doerfler, W. *Biol. Chem. Hoppe-Seyler* **1991**, *372*, 557–564.
(6) Laird, P. W.; Jaenisch, R. *Hum. Mol. Genet.* **1994**, *3*, 1487–1495.
(7) Venitt, S. *Environ. Health Perspect.* **1996**, *104*, 633–637.
(8) Jones, P. A.; Gonzalgo, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2103–2105.
(9) Lieb, M. *Genetics* **1991**, *128*, 23–27.
(10) Pfeifer, G. P. *Curr. Top. Microbiol. Immunol.* **2006**, *301*, 259–281.
(11) Kow, Y. W. *Free Radical Biol. Med.* **2002**, *33*, 886–893.
(12) Cortázar, D.; Kunz, C.; Saito, Y.; Steinacher, R.; Schär, P. *DNA Repair* **2007**, *6*, 489–504.
(13) Brown, T. C.; Jiricny, J. *Cell* **1988**, *54*, 705–711.
(14) Shapiro, R.; Klein, R. S. *Biochemistry* **1966**, *5*, 2358–2362.
(15) Lindhal, T.; Nyberg, B. *Biochemistry* **1974**, *13*, 3405–3410.
(16) Wang, R. Y.; Kuo, K. C.; Gehrke, C. W.; Huang, L. H.; Ehrlich, M. *Biochim. Biophys. Acta* **1982**, *697*, 371–377.
(17) Ehrlich, M.; Norris, K. F.; Wang, R. Y.-H.; Kuo, K. C.; Gehrke, C. W. *Biosci. Rep.* **1986**, *6*, 387–393.
(18) Ehrlich, M.; Zhang, X.-Y.; Inamdar, N. M. *Mutat. Res.* **1990**, *238*, 277–286.
(19) Shen, J. C.; Rideout, W. M., III.; Jones, P. A. *Nucleic Acids Res.* **1994**, *22*, 972–976.
(20) Morgan, H. D.; Dean, W.; Coker, H. A.; Reik, W.; Petersen-Mahrt, S. K. *J. Biol. Chem.* **2004**, *279*, 52353–52360.
(21) Privat, E.; Sowers, L. C. *Chem. Res. Toxicol.* **1996**, *9*, 745–750.
(22) Bienvenu, C.; Cadet, J. *J. Org. Chem.* **1996**, *61*, 2632–2637.
(23) Douki, T.; Cadet, J. *Biochemistry* **1994**, *33*, 11942–11950.
(24) Lee, D.-H.; Pfeifer, G. P. *J. Biol. Chem.* **2003**, *278*, 10314–10321.
(25) Labet, V.; Grand, A.; Morell, C.; Cadet, J.; Eriksson, L. A. *Theor. Chem. Acc.* **2008**, *120*, 429–435.
(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskortz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chem, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03* , Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(27) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(28) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785–789.
(29) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.
(30) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
(31) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265–3269.
(32) Cancès, M. T.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
(33) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151–5158.
(34) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.
(35) Toro-Labbé, A. *J. Phys. Chem. A* **1999**, *103*, 4398–4403.
(36) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
(37) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793–1874.
(38) Chermette, H. *J. Comput. Chem.* **1999**, *20*, 129–154.
(39) Chattaraj, P. K.; Sarkar, U.; Roy, D. R. *Chem. Rev.* **2006**, *106*, 2065–2091.
(40) Ayers, P. W.; Anderson, J. S. M.; Bartolotti, L. J. *Int. J. Quantum Chem.* **2005**, *101*, 520–534.
(41) Klopman, G. *J. Am. Chem. Soc.* **1968**, *90*, 223–234.
(42) Melin, J.; Aparicio, F.; Subramanian, V.; Galvan, M.; Chattaraj, P. K. *J. Phys. Chem.* **2004**, *108*, 2487–2491.
(43) Ayers, P. W.; Parr, R. G.; Pearson, R. G. *J. Chem. Phys.* **2006**, *124*, 194107.
(44) Anderson, J. S. M.; Melin, J.; Ayers, P. W. *J. Chem. Theory Comput.* **2007**, *3*, 358–374.
(45) Parr, R. G.; Yang, W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701–728.
(46) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, *68*, 3801–3807.
(47) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512–7516.
(48) Pearson, R. G. *J. Am. Chem. Soc.* **1985**, *107*, 6801–6806.
(49) Pearson, R. G. *J. Chem. Educ.* **1987**, *64*, 561–567.
(50) Maynard, A. T.; Huang, M.; Rice, W. G.; Covel, D. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11578–11583.
(51) Parr, R. G.; Szentpaly, L. V.; Liu, S. *J. Am. Chem. Soc.* **1999**, *121*, 1922–1924.
(52) Janak, J. F. *Phys. Rev. B* **1978**, *18*, 7165–8.
(53) Cohen, A. J.; Mori, San.; Yang, W. T. *Phys. Rev. B* **2008**, *77*, 115–123.
(54) Chattaraj, P. K.; Maiti, B.; Sarkar, U. *J. Phys. Chem. A* **2003**, *107*, 4973–4975.
(55) Parr, R. G.; Yang, W. *J. Am. Chem. Soc.* **1984**, *106*, 4049–4050.
(56) Ayers, P. W.; Levy, M. *Theor. Chem. Acc.* **2000**, *103*, 353–360.
(57) Padmanabhan, J.; Parthasarathi, R.; Elango, M.; Subramanian, V.; Krishnamoorthy, B. S.; Gutierrez-Oliva, S.; Toro-Labbé, A.; Roy, D. R.; Chattaraj, P. K. *J. Phys. Chem. A* **2007**, *111*, 9130–9138.
(58) Morell, C.; Grand, A.; Toro-Labbé, A. *J. Phys. Chem. A* **2005**, *109*, 205–212.
(59) Morell, C.; Grand, A.; Toro-Labbé, A. *Chem. Phys. Lett.* **2006**, *425*, 342–346.
(60) Geerlings, P.; De Proft, F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3028–3042.
(61) Ayers, P. W.; Morell, C.; De Proft, F.; Geerlings, P. *Chem. Eur. J.* **2007**, *13*, 8240–8247.
(62) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

(63) Labet, V.; Grand, A.; Cadet, J.; Eriksson, L. A. *Chem. Phys. Chem.* **2008**, *9*, 1195–1203.
(64) Robinson, H.; Wang, A. H.-J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5224–5228.
(65) Eyring, H. *Chem. Rev.* **1935**, *17*, 65–77.
(66) Sousa, F. P.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439–10452.
(67) Krawczak, M.; Ball, E. V.; Cooper, D. N. *Am. J. Hum. Genet.* **1998**, *63*, 474–488.
(68) Millar, C. B.; Guy, J.; Sansom, O. J.; Selfridge, J.; MacDougall, E.; Hendrich, B.; Keightley, P. D.; Bishop, S. M.; Clark, A. R.; Bird, A. *Science* **2002**, *297*, 403–405.
(69) Pfeifer, G. P.; You, Y.-H.; Besaratinia, A. *Mutat. Res.* **2005**, *571*, 19–31.
(70) Valinluck, V.; Sowers, L. C. *Cancer Res.* **2007**, *67*, 5583–5586.

JP808902J