Theor Chem Acc (2012) 131:1187
DOI 10.1007/s00214-012-1187-x

REGULAR ARTICLE

# Comparison of the mechanism of deamination of 5,6-dihydro-5-methylcytosine with other cytosine derivatives

**André Grand · Jean Cadet · Leif A. Eriksson · Vanessa Labet · Nelly L. Jorge · Maria L. Schreiber · Thierry Douki · Christophe Morell**

Received: 8 July 2011 / Accepted: 8 December 2011 / Published online: 4 April 2012


**Abstract** The mechanism of deamination of 5,6-dihydro-5-methylcytosine has been investigated theoretically and compared to those of other cytosine derivatives. The main goal is to understand the effect of C5-methylation and C5–C6 saturation upon the deamination rate. It is found that C5–C6 saturation tends to increase the local electrophilicity of the cytosine derivative on carbon C4. It is also concluded that C5-methylation displays an opposite effect on saturated versus unsaturated systems: on unsaturated systems, C5-methylation tends to increase the local electrophilicity on C4, while it reduces the local electrophilicity on C4 for saturated ones.

Dedicated to Professor Vincenzo Barone and published as part of the special collection of articles celebrating his 60th birthday.

A. Grand · J. Cadet · T. Douki · C. Morell (✉)
Laboratoire Lésions des Acides Nucléiques,
INAC/SCIB-UMR-E no. 3 CEA-UJF, CEA Grenoble,
17 rue des Martyrs, 38054 Grenoble cedex 9, France
e-mail: christophe.morell@ujf-grenoble.fr

L. A. Eriksson
Department of Chemistry and Molecular Biology,
University of Gothenburg, 412 96 Gothenburg, Sweden

V. Labet
UPMC Univ Paris 06, UMR 7075, Laboratoire de Dynamique,
Interactions et Réactivité (LADIR), 75005 Paris, France

V. Labet
CNRS, UMR 7075, Laboratoire de Dynamique,
Interactions et Réactivité (LADIR), 75005 Paris, France

N. L. Jorge · M. L. Schreiber
Laboratorio de Investigaciones en Tecnología del Medio Ambiente, FACENA-UNNE, Av. Libertad 5640,
3400 Corrientes, Argentina

**Keywords** Deamination · 5,6-Dihydro-5-methylcytosine · DNA damage · Mutation · DFT · Reactivity indices · Dual descriptor

## 1 Introduction

Damage to DNA is a major deleterious process in cells that can lead to lethality or modification of the genetic information. The latter process is of particular importance because it may represent through mutation the initial step toward malignant tumor formation. DNA damage arises as a result of exposure to a variety of exogenous agents such as ionizing and UV radiations, chemicals and alkylating agents. In addition, endogenous processes like mitochondrial respiration and inflammation lead to the release of genotoxic compounds including oxygen and nitrogen reactive species. This withstanding, water remains the most efficient endogenous damaging compound; spontaneous hydrolytic reactions of the *N*-glycosidic bonds are at the origin of the formation of abasic sites upon depurination and of the conversion of cytosine (**Cyt**) into uracil. In the latter reaction, water substitutes the exocyclic amino group of cytosine, giving rise to a tautomer of uracil. The biological relevance of deamination in genome stability is indirectly demonstrated by the very high uracil DNA *N*-glycosylase activity in cells.

5-Methylcytosine (**5mCyt**) also undergoes deamination. This minor base represents up to 5% of cytosine bases in human cells and plays a major role in gene regulation and defense against viruses [1, 2]. It is produced enzymatically from cytosine at CpG sites by DNA methyltransferases, with *S*-adenosyl-methionine as the methyl donor. 5-Methylcytosine is more sensitive than cytosine toward deamination. The rate constants for deamination in double-stranded DNA have been reported to be $5.8 \times 10^{-13}$ $s^{-1}$ (**5mCyt**) and $2.6 \times 10^{-13}$ $s^{-1}$ (**Cyt**) at 37 °C and physiological pH,

respectively [3]. The deamination product of 5mCyt is thymine, which is excised from DNA by a specific repair glycosylase when thymine is paired with guanine [4]. However, mutational events at 5-methylated cytosines remain among the most frequent ones [5].

Saturation of the C5–C6 double bond of cytosine is another chemical modification that drastically affects the deamination rate. Values in the range of $10^{-6}$ $s^{-1}$ were reported for 5,6-dihydrocytosine under various experimental conditions [6, 7]. Accelerated deamination is also observed in 5,6-saturated cytosine derivatives including oxidation products [8–10], photohydrates [11, 12], and cyclobutane pyrimidic dimeric photoproducts [13–17]. Interestingly, methylation at position 5 in such compounds decreases the deamination rate [18–21] in contrast to what is observed for the 5,6-unsaturated parent cytosine molecule [4].

In recent works, we have applied quantum chemistry tools to better describe the elementary steps involved in the deamination of cytosine [22, 23], 5-methylcytosine [24] and 5,6-dihydrocytosine (**5dhCyt**) [25]. The main conclusion of these studies is that more than one water molecule is needed to fully explain the observed rate constants. In each case, the assisted nucleophilic addition of one water molecule to C4 (see Fig. 1) of the cytosine derivative with the assistance of a second water molecule is the rate-determining step. It appears that the imino tautomer of cytosine can be viewed as the active form of the nucleobase for this reaction. For cytosine and 5-methylcytosine, the imino tautomer is formed in a concerted way during the nucleophilic addition, whereas for 5,6-dihydrocytosine it has to be generated prior to the addition step. We herein extend these studies to 5,6-dihydro-5-methylcytosine (**5mdhCyt**) with the purpose of exploring the mechanistic reasons for the opposite effect of methylation on deamination of cytosine and its 5,6-saturated derivative. The numbering of the atoms within all compounds is given in Fig. 1.

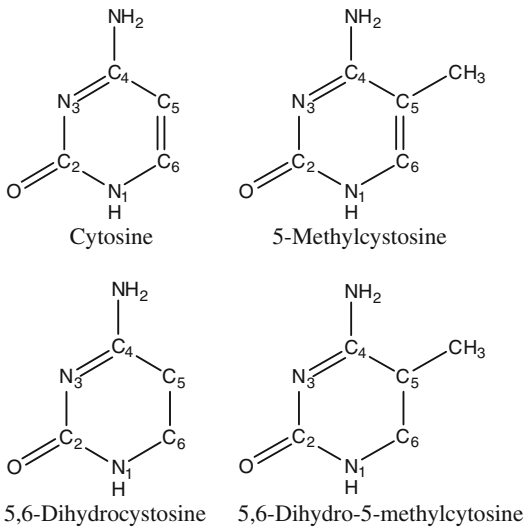


**Fig. 1** Atom numbering of cytosine and cytosine derivatives

## 2 Computational details

### 2.1 Quantum chemical calculations

In order to enable a straightforward comparison with data previously reported for hydrolytic deamination of cytosine, 5-methylcytosine, and 5,6-dihydrocytosine, the same quantum methodology was adopted as in those studies [22–25]. All calculations were thus performed using the Gaussian 03 package [26] at the B3LYP/6-311G(d,p) level of theory [27–31]. As a first step, molecular systems of interest were fully optimized in vacuum. The stationary points were then characterized as minima or transition states by computing the vibrational frequencies within the harmonic approximation at the same level of theory, and thermal data were extracted to obtain thermodynamic functions of reaction and corresponding activation parameters at 298.15 K and 1 atm. The B3LYP functional is known to predict with moderate accuracy the barrier heights of H-transfers. Previous work [22–25] has however shown that, although B3LYP underestimates barrier heights in comparison with single-point MP2 calculations using the same basis set, the general energy trends are not affected.

Single-point energy calculations were subsequently performed on the optimized geometries, including the integral equation formalism polarized continuum model (IEF-PCM, hereafter referred to as PCM [32–34]), with dielectric constant $\varepsilon = 78.39$ to simulate the bulk effects of an aqueous environment, and to enable straightforward comparison with earlier work. The cavity was built using the united atom topologic model applied to 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 to all hydrogen atoms linked to nitrogen and oxygen atoms, using the keyword SPHEREONH.

Reaction paths in vacuum were characterized at the same level of theory through intrinsic reaction coordinate (IRC) calculations from the gas-phase-optimized transition structures to ensure that they connected the appropriate reactants and products.

### 2.2 Reactivity indices

In order to rationalize the difference in reactivity between the investigated systems, a range of global and local reactivity indices as derived from conceptual DFT [35–40] were computed for the imine tautomers of cytosine, 5-methylcytosine, and their 5,6-dihydrogenated counterparts. Calculations were performed on the optimized structures in *vacuum* and in bulk solvent. The different indices evaluated are as follows (full details and mathematical expressions can found in for instance Ref. [25]).

#### 2.2.1 Global reactivity indices

These parameters, which give information on the reactivity of the molecular system as a whole, include the *chemical potential* [41] $\mu$, giving a measure of the escaping tendency of the electrons; the *absolute hardness* [42, 43] $\eta$, which may be viewed as the resistance toward charge transfer; and the *global electrophilicity index* [44, 45] $\omega$, which can be considered as a measure of the electrophilicity of the molecular system. From a computational perspective, these are evaluated from frontier orbital energies (HOMO, LUMO) together with expressions based on finite differences approximations and Koopmans' theorem [46, 47].

#### 2.2.2 Local reactivity indices

In order to address differences in reactivity within a particular system, we instead need to explore local features. The first is the *net atomic charges*, in particular of carbon C4 at which the deamination reaction takes place and which gives a measure of the reactivity under electrostatic control. Three other local indices evaluated provide information about reactivity under electron-transfer control. These include the *local electrophilicity index* [48] $\omega^{+}(\mathbf{r})$ which is the product of the global electrophilicity index $\omega$ and the Fukui function [49–51] $f^{+}(\mathbf{r})$; the *excess electrophilicity* [52] $\Delta\omega(\mathbf{r})$ defined as the product of the *dual descriptor* [53–56] $\Delta f(\mathbf{r})$ and the global electrophilicity index $\omega$ and finally another local reactivity index $s^{(2)}(\mathbf{r}) = \frac{\Delta f(\mathbf{r})}{\eta^2}$ derived from the *dual descriptor within the grand canonical ensemble* [57] *(GCDD)*. The latter descriptor is size extensive and thus more suitable to compare molecules with different number of electrons. In its reduced form used in this paper, it can be calculated as the product of the dual descriptor and the inverse of the square of the global hardness. The more positive the *GCDD* is the more willing the site is to undergo nucleophilic attack, whereas a large negative number indicates a strong tendency to undergo electrophilic attack. The partial charges were evaluated using the ChelpG method [58], and the dual descriptors are based on comparisons of the electronic populations on carbon C4 upon adding or subtracting an electron to or from the normal system. For full details, we refer to Ref. [25].

## 3 Results and discussion

### 3.1 Geometries of the starting materials

5,6-Dihydro-5-methylcytosine possesses an asymmetric carbon, namely the carbon bearing the methyl group. As a consequence, two enantiomers can exist. It is obvious that the results observed are the same for both enantiomers.

We furthermore note that saturation of the C5=C6 bond produces a puckering of the ring system, and thus two orientations of the C5 methyl group, axial and equatorial, are possible, which will influence the energetics and interactions of the incoming water/OH addition at C4. Both cases are considered herein.

Cytosine and cytosine derivatives can furthermore exist under two tautomeric forms. The classical amino form that possess an exocyclic amino group and an imino form characterized by a double bond between the carbon C4 and the exocyclic nitrogen. Quite interestingly, for the 5,6-dihydrogenated forms, the axial conformer is the most stable one for the amino form, while it is the equatorial for the imino tautomer. All conformers do, however, lie very close in energy; the range is about 1.5 kJ/mol.

### 3.2 Reaction pathways—geometries and energies

#### 3.2.1 Description of the mechanism studied

The reaction mechanism for spontaneous deamination of 5,6-dihydro-5-methylcytosine in protic medium is outlined schematically in Fig. 2. Following hydrogenation across the C5=C6 double bond (not shown), the reaction starts with a water-assisted tautomerization of the common amine form (**5mdhR**) to the corresponding imine (**5mdhIR1**). This generates a reactive C4=N4 double bond (atomic numbering given in Fig. 2), across which water is added to yield **5mdhIR2**. The exocyclic amine group is protonated from the medium (**5mdhIR3**), followed by $NH_3$ release from C4 and abstraction of the OH proton from the just added C4–OH group to give the **5mdhP** product (uracil tautomer) and an ammonium cation. Of the two water molecules included in the cluster calculations, one is involved in the proton shuttle mechanisms thus stabilizing hydrogen bonding, whereas the other actively adds across the C4–N4 bond as OH and H. The above labels for the different stationary structures will be used throughout, with the addition **eq** or **ax** for the conformers with C5 methyl group in equatorial or axial orientation, respectively.

#### 3.2.2 Formation of the active form of the nucleobase

In Fig. 3a, the main structures for the stationary points along the initial tautomerization reaction for the equatorial conformer are shown, and in Fig. 3b the corresponding structures for the axial form are displayed. In the reactant complex **5mdhReq**, "water 7" is positioned with hydrogen bonding between O7 and H4a (2.096 Å) and between H7a and N3 (1.815 Å) to readily facilitate the proton shift from N4 to N3. The two protons transfer concertedly, with the

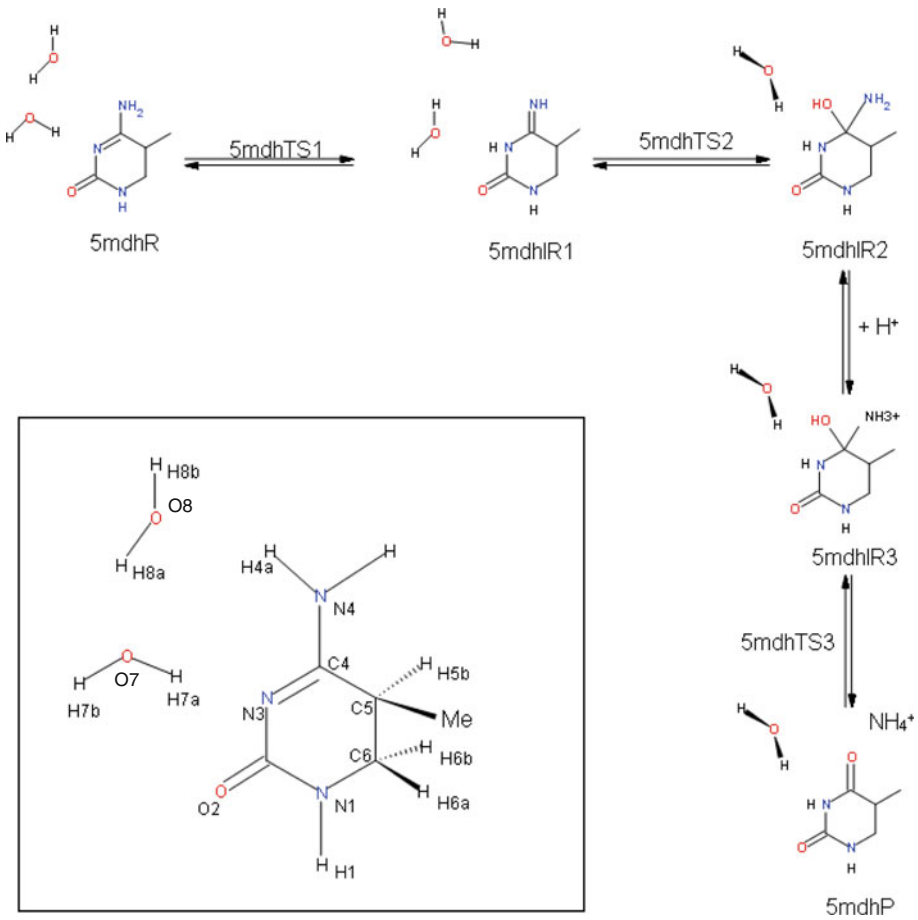


**Fig. 2** Reaction pathways involved in the spontaneous deamination of 5-methyl-5,6-dihydrocytosine in protic medium. *Insert* atomic labeling used throughout the study

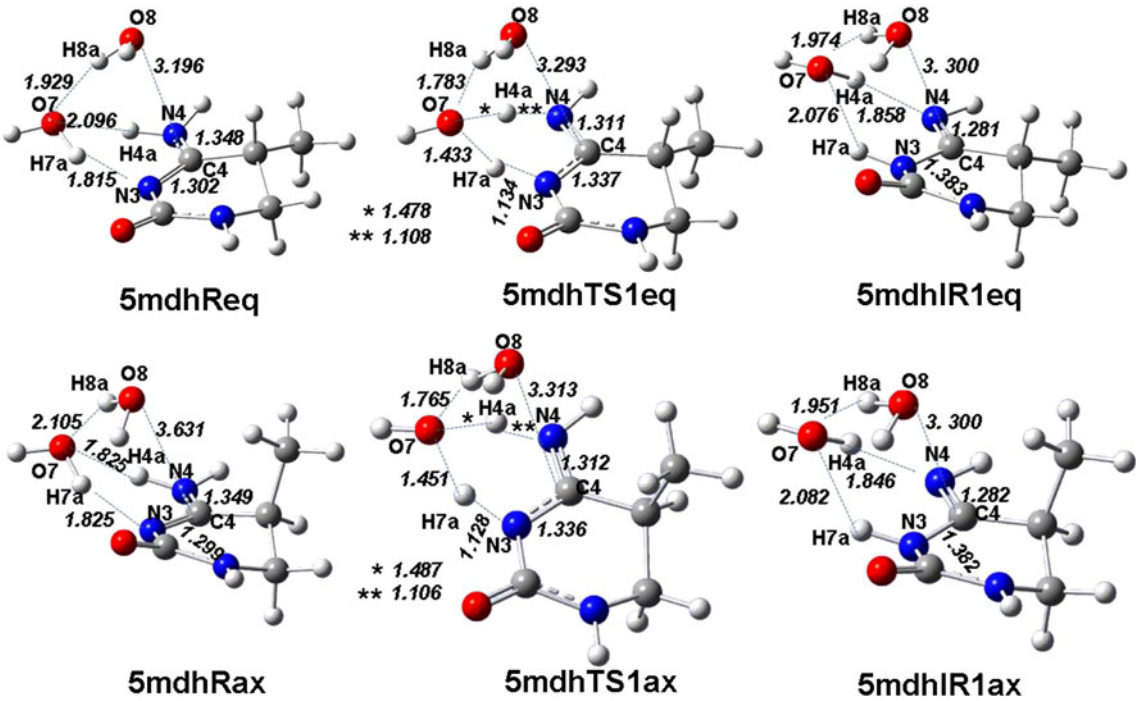


**Fig. 3** **a** Optimized geometries of **5mdhReq**, **5mdhTS1eq**, and **5mdhIR1eq**; **b** **5mdhRax**, **5mdhTS1ax**, and **5mdhIR1ax** Distances in Å

**Table 1** Relative energies in vacuum, free energies in vacuum, and relative free energies in aqueous solvent for the stationary points of the deamination pathway of 5-methyl-5,6-dihydrocytosine

| System | $\Delta E_{vac}$ | $\Delta G_{vac}$ | $\Delta G_{aq}$ |
|---|---|---|---|
| a) Equatorial position | | | |
| 5mdhR + $H^+$ | 0.0 | 0.0 | 0.0 |
| 5mdhTS1 + $H^+$ | 44.0 | 48.5 | 50.1 |
| 5mdhIR1 + $H^+$ | −11.2 | −10.8 | −9.2 |
| 5mdhIR1 constrained + $H^+$ | −11.1 | −9.8 | −7.9 |
| 5mdhTS2 + $H^+$ | 105.3 | 115.5 | 118.5 |
| 5mdhIR2 + $H^+$ | 15.9 | 24.6 | 16.7 |
| 5mdhIR3 | −928.9 | −918.6 | −18.4 |
| 5mdhTS3 | −916.3 | −907.4 | 12.9 |
| 5mdhP | −1,035.8 | −1,039.8 | −150.5 |
| b) Axial position | | | |
| 5mdhR + $H^+$ | 0.0 | 0.0 | 0.0 |
| 5mdhTS1 + $H^+$ | 48.2 | 50.4 | 57.7 |
| 5mdhIR1 + $H^+$ | −21.9 | −22.4 | −17.0 |
| 5mdhIR1 constrained + $H^+$ | −6.2 | −4.4 | −1.5 |
| 5mdhTS2 + $H^+$ | 115.8 | 124.3 | 127.0 |
| 5mdhIR2 + $H^+$ | 25.8 | 31.4 | 26.0 |
| 5mdhIR3 | −925.1 | −922.3 | −36.8 |
| 5mdhTS3 | −914.8 | −909.1 | 3.2 |
| 5mdhP | −1,056.1 | −1,058.3 | −152.8 |

All values in kJ/mo

H4a transfer representing an early TS and the H7a transfer a late one. The reaction free energy is 50.1 kJ/mol in aqueous solution, and the resulting **5mdhIR1eq** lies 9.2 kJ/mol below the reactant (Table 1a). For the axial form, the overall geometric features of the tautomerization reaction are highly similar to those of the equatorial conformer, except that the two water molecules are slightly (∼0.02 Å) further out. Energetically, the reaction free energy is higher by 7.6 kJ/mol than for the equatorial system, and the **5mdhIR1ax** tautomer is 17 kJ/mol more stable than the reactant (Table 1b).

During the reaction, the C4–N4 bond shortens from 1.35 to 1.28 Å, and the N3–C4 bond lengthens from 1.30 to 1.38 Å. We also note that due to the hydrogenation of the C5–C6 bond, the conjugation of the ring system is broken and the molecule becomes highly non-planar. It is worth noticing that for 5,6-dihydro-5-methylcytosine the imino form is more stable than the amino form. The same trend is observed for the 5,6-dihydrocytosine [25]. One can hypothesize that it is a general tendency that when the C5–C6 bond is saturated, the imino tautomer is more stable than the amino tautomer. The overstabilization of the imino tautomer versus the amino tautomer is likely due to the presence of a urea moiety in the imino form. Urea fragments are very stable entities since the lone pair of each nitrogen interacts with the vicinal π* (C=O) bond. A NBO [59] analysis confirms this hypothesis with a quite strong stabilizing energy between nitrogen N3 and the carbonyl π* (C=O); about 48 kJ/mol.

The effects of free energy corrections relative to electronic energies in vacuum, as well as the effect of inclusion of the solvent, can also be seen from the data reported in Table 1. Introducing thermal corrections in the equatorial system increases the barrier by 4.5 kJ/mol, and adding the solvent effects raises this by another 1.6 kJ/mol. For the axial conformer, the corresponding increases are 2.2 and 7.3 kJ/mol, respectively. Similarly, the relative energy of the products becomes slightly more positive, following the same trends as for the barriers, with the exception of the thermal free energy correction to **5mhdIR1ax**, which lowers the energy by 0.5 kJ/mol.

### 3.2.3 *Preparation for the nucleophilic addition of water*

In order for the subsequent water addition to occur, the attacking oxygen (O8) must approach C4 in a perpendicular trajectory with respect to the molecular plane. In the current systems, the **5mCyt** ring is already distorted due to the hydrogenation at C5 and C6, whereas in the case of non-hydrogenated cytosine and 5-methylcytosine, the ring system remains planar and the two water molecules align with stable hydrogen bonds in that same plane. To overcome the unphysical situation of an in-plane water required to attack carbon C4 perpendicularly to the plane, the concept of reaction forces was employed herein as well as in our previous studies on deamination reactions.

The reaction force, introduced by Toro-Labbé [60], is defined as the negative derivative of the energy with respect to the intrinsic reaction coordinate. It represents the force acting on the system to bring the reactants into the products. The reaction force vanishes at stationary points, and it has been shown (see, e.g., Ref. [25]) that the reaction force becomes very close to zero in a region between reactant and TS. The coordinates at this point are then used in a constrained geometry optimization of a more physically realistic reactant complex. In the current context, the structural differences between the tautomerization product **5mdhIR1eq** and the reactive complex (Fig. 4) turned out to be relatively small due to the non-planarity of the cytosine derivative. This is also reflected in the close-to-zero electronic energy difference in vacuum between the two, and just over 1 kJ/mol difference in free energy in aqueous solution (Table 1a). The main difference involves a shortening of the C4–O8 distance from 3.300 to 3.181 Å. The situation is quite different for the axial system. Due to the repulsion between the axially oriented methyl group and the incoming water, the energy of the constrained system is raised by 16–18 kJ/mol. The energetics of the

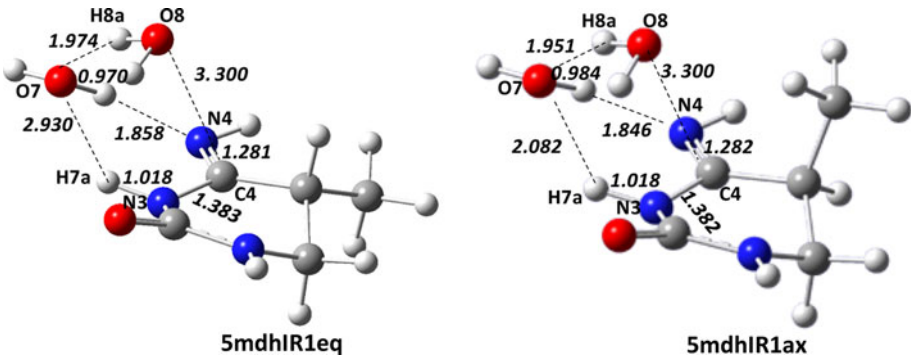


**Fig. 4** Optimized constrained geometries of the **5mdhIR1** reactant complexes, in vacuum. Distances in Å

subsequent points on the reaction energy surfaces (Fig. 5a (**eq**), b (**ax**)) are shifted to be relative to the energy of the corresponding reactive complexes.

### *3.2.4 Nucleophilic addition of water to C4*

Water addition has been shown in all our previous deamination reaction studies to be the rate-determining step. This is the case also in the current system, where the TS **5mdhTS2eq** lies 126.4 kJ/mol above the **5mdhIR1eq** reactant complex (Fig. 6). A free energy barrier of 128.5 kJ/mol is observed for the corresponding **5ndhTS2ax** system (Fig. 6). The geometries of both the TS's and the resulting adducts (**5mdhTS2**, **5mdhIR2**; Fig. 6) are very similar to those of the corresponding 5,6-dihydrocytosine (**dhTS2**, **dhIR2**; Fig. 6)—the differences in bond distances throughout being within 0.01 Å between the corresponding stationary points for the equatorial conformer, and slightly larger differences (0.02–0.03 Å) for the distances to the two waters in the axial system. The TS forms a 6-membered ring (C4–N4–H4a–O7–H8a–O8) and thus involves the back-transfer of H4a that was originally transferred to O7 of water in the tautomerization reaction, concomitantly with the transfer of H8a to O7, as O8 binds to C4. The resulting adduct forms three hydrogen bonds to the O7 water molecule (O7–H7a, H8a–O8, and O7–H4a), all in the range 2.0–2.3 Å. In the case of the equatorial conformer, the inclusion of thermal corrections raises the barriers by 7–10 kJ/mol, whereas solvent effects are less pronounced, by 2–3 kJ/mol. For the axial system, inclusion of thermal effects raises the energy of the products by 6–9 kJ/mol, whereas inclusion of solvent effects provides product stabilization by a similar amount.

### *3.2.5 Protonation of the amino group and deamination*

Water addition is followed by protonation of the N4 amine by the medium, which stabilizes the complex by 33–35 kJ/mol compared to the energy of **5mdhIR2** + $H^+$ [61]. Deamination of the protonated **5mdhIR3** leads to simultaneous

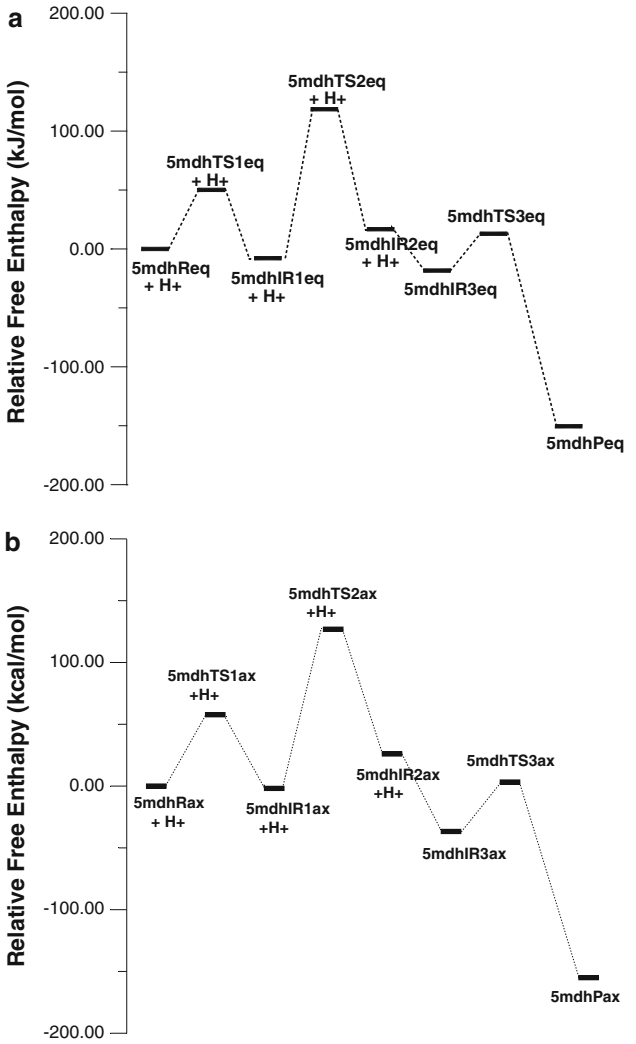


**Fig. 5** **a** Relative free energies (kJ/mol) along the studied deamination pathway in aqueous solution for equatorial isomer; **b** As (**a**) for axial isomer

deprotonation by the leaving ammonia, of the just added O8 (Fig. 7). The barrier to deamination is 31 kJ/mol for the equatorial conformer, and the final product **5mdhPeq** is located 132 kJ/mol below the protonated reactant **5mdhIR3eq**. In the case of the axial conformer, the barrier is

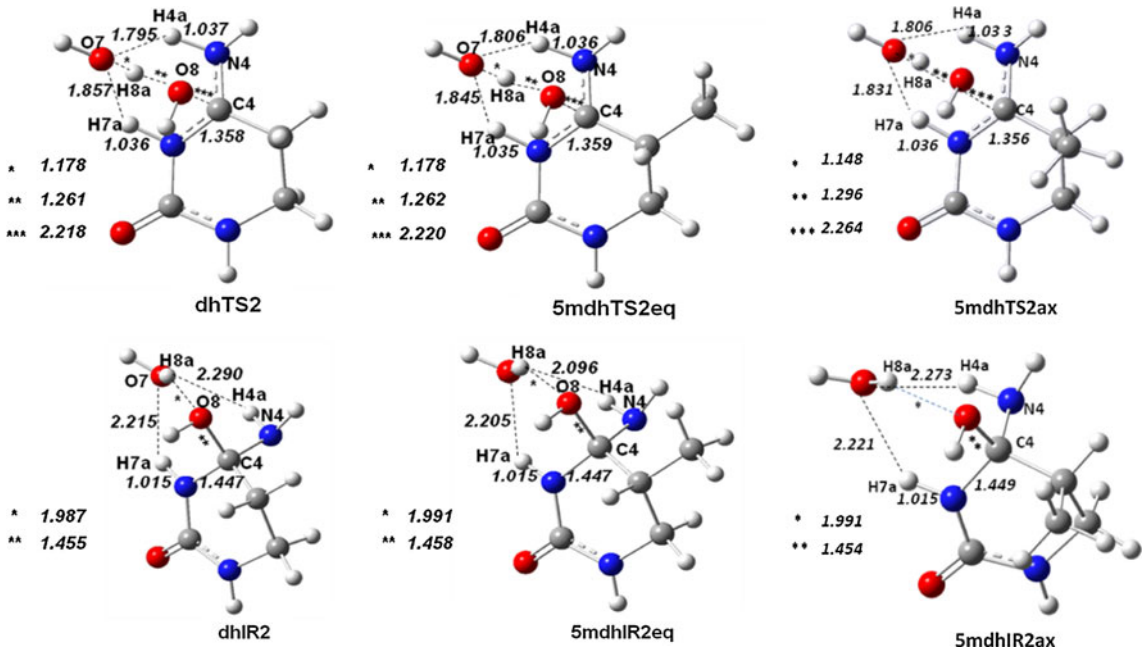


**Fig. 6** Optimized geometries of IR2 and TS2 for **dhCyt**, **5mdhCyteq,** and **5mdhCytax**. Distances in Å

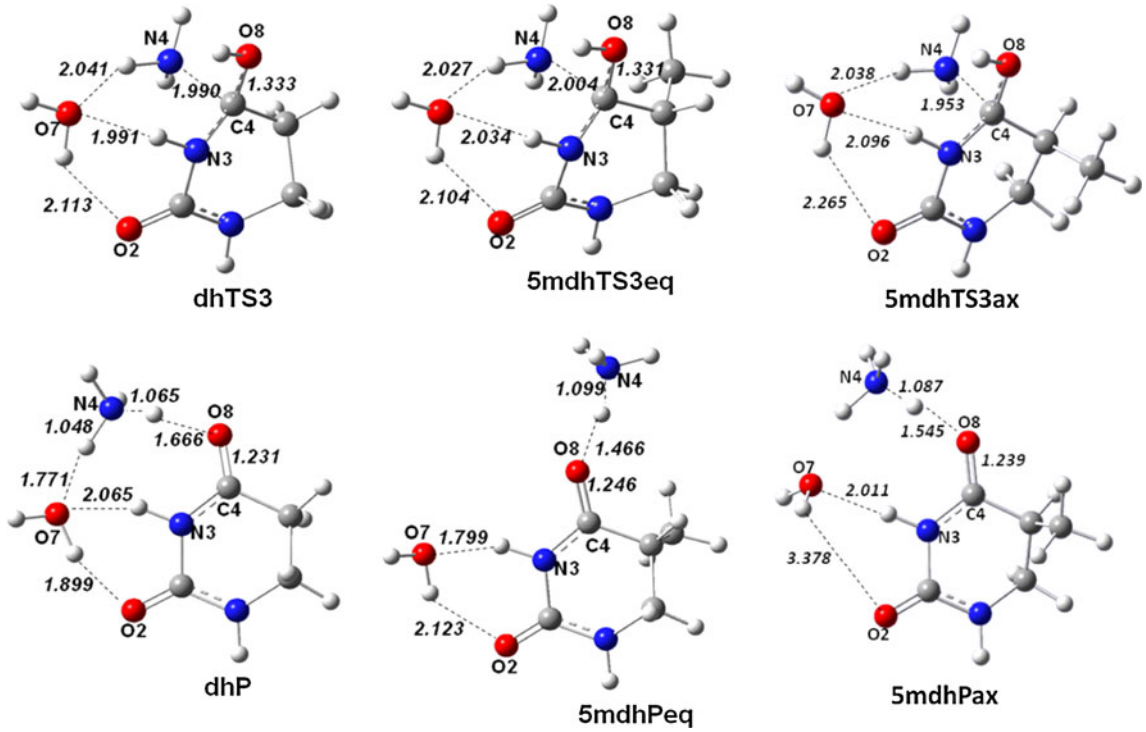


**Fig. 7** Optimized geometries TS3 and product for **dhCyt**, **5mdhCyteq,** and **5mdhCytax**. Distances in Å

considerably higher, 40 kJ/mol, and the reaction less exergonic ($\Delta\Delta G_{aq} = 116.0$ kJ/mol). The difference between the two conformers is largely due to the higher relative energy of the **5mdhIR3ax** complex over the corresponding equatorial one. The transition states for deamination are again very similar to the 5,6-dihydrocytosine counterpart, see Fig. 7.

### 3.3 Energetics

The computed energetics of the rate-determining steps for **Cyt**, **5mCyt**, **dhCyt,** and **5mdhCyt**, following similar pathways as outlined herein, are 138, 134, 110, and 128 kJ/mol, respectively. As seen from the energetics above,

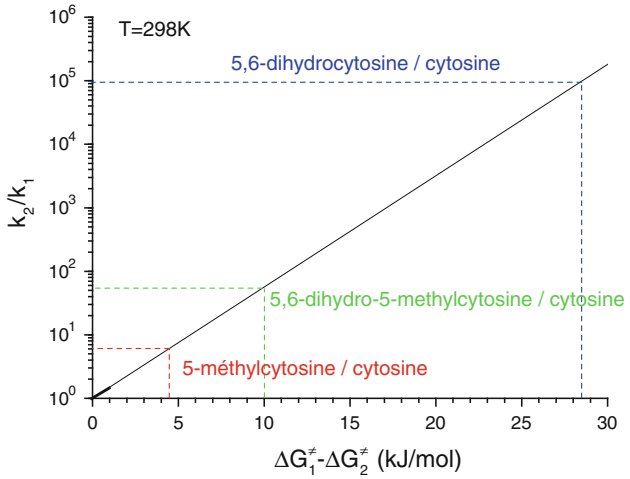


**Fig. 8** Evolution of the deamination rate ratio

saturation of the double bond in cytosine gives a significantly increased rate of deamination, whereas for the **5mCyt** counterparts the effects are smaller but still significant. The trend agrees with that duly experimentally noted [16–19] and, interestingly, the effect of the methylation is opposite in cytosine and 5,6-dihydrocytosine.

From the data below and using the Eyring's law, one can estimate a deamination rate ratio between 5,6-dihydro-5methylcytosine and cytosine. Using Eyring's Law, the reaction rate ratio can be assessed using the following relation:

$$\frac{k_2}{k_1} \approx \exp\left(\frac{\Delta G_1^{\neq} - \Delta G_2^{\neq}}{\mathrm{RT}}\right)$$

As the deamination rate ratio between 5,6-dihydrocytosine and cytosine is about $\frac{k_{56\mathrm{dhcyt}}}{k_{\mathrm{cyt}}} \approx 10^5$ ($\Delta\Delta G = 138–110 = 28$ kJ/mol) and the one between 5-methylcytosine and cytosine is about $\frac{k_{5\mathrm{mcyt}}}{k_{\mathrm{cyt}}} \approx 5$ ($\Delta\Delta\mathrm{G} = 138–134 = 4$ kJ/mol), one can assess that the deamination rate ratio between 5,6-dihydro-5-methylcytosine and cytosine would be about $\frac{k_{56\mathrm{dh5mcyt}}}{k_{\mathrm{cyt}}} \approx 50$ ($\Delta\Delta G = 138–128 = 10$ kJ/mol); cf. Fig. 8. Experimental value obtained at 25 °C for the deamination rate of cytosine is 2.1 $10^{-10}$ $s^{-1}$. The estimated value for 5,6-dihydro-5-methylcytosine, based on the current analysis, is therefore 1.2 $10^{-8}$ $s^{-1}$. Unfortunately, to our best knowledge, no accurate experimental determination of such value has been carried out so far.

## 3.4 Reactivity indices

### *3.4.1 Global reactivity indexes*

In Table 2a and b, we report the frontier orbital energies (HOMO and LUMO) for **Cyt**, **5mCyt**, **dhCyt,** and the axial and equatorial forms of **5mdhCyt**. Since for each of these cytosine derivatives the imine tautomer has been shown to be the actual active form of the nucleobases for the rate-determining step, be it formed prior to the rate-determining step as in the **dhCyt** and **5mdhCyt** cases, or in a concerted way with it as in the **Cyt** and **5mCyt** cases, we therefore focus our analyses on these. From the orbital energies, we have also computed global reactivity indices such as the chemical potential $\mu$, the absolute hardness $\eta$, and the global electrophilicity index $\omega$. In nucleophilic reactions under frontier orbital control, the reactivity of the electrophile is, according to frontier molecular orbital theory (FMOT), driven by its LUMO, whereas that of the nucleophile is by its HOMO. The closer the LUMO of the electrophile and the HOMO of the nucleophile lie in energy, the easier the reaction will be. In the following

**Table 2** Kohn-Sham frontier orbital energies ($\varepsilon_{\mathrm{LUMO}}$ and $\varepsilon_{\mathrm{HOMO}}$), chemical potentials ($\mu$), absolute hardnesses ($\eta$), and global electrophilicity indices ($\omega$), in vacuum (a) and in aqueous solvent (b)

| Molecular system | $\varepsilon_{\mathrm{LUMO}}$ | $\varepsilon_{\mathrm{HOMO}}$ | $\mu$ | $\eta$ | $\omega$ |
|---|---|---|---|---|---|
| a) Vacuum | | | | | |
| Cytosine imine tautomer | −1.09 | −6.50 | −3.80 | 5.41 | 1.33 |
| 5,6-Dihydrocytosine imine tautomer | −0.16 | −7.02 | −3.59 | 6.86 | 0.94 |
| 5-Methylcytosine imine tautomer | −0.98 | −6.36 | −3.67 | 5.38 | 1.25 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer(equatorial) | −0.19 | −7.00 | −3.59 | 6.82 | 0.95 |
| 5-Methylcytosine-5,6-dihydrocytosine imine tautomer(axial) | −0.20 | −7.00 | −3.60 | 6.81 | 0.95 |
| b) Aqueous solvent | | | | | |
| Cytosine imine tautomer | −0.87 | −6.37 | −3.62 | 5.50 | 1.19 |
| 5,6-Dihydrocytosine imine tautomer | −0.22 | −7.10 | −3.66 | 6.88 | 0.97 |
| 5-Methylcytosine imine tautomer | −0.74 | −6.16 | −3.45 | 5.42 | 1.10 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer(equatorial) | −0.11 | −6.96 | −3.53 | 6.85 | 0.91 |
| 5-Methylcytosine-5,6-dihydrocytosine imine tautomer(axial) | −0.22 | −7.03 | −3.63 | 6.80 | 0.97 |

All values are in eV

discussion, the effects of methylation and saturation of cytosine on the frontier orbital energies are discussed and analyzed.

*3.4.1.1 Frontier orbital energies* Methylation of cytosine has a mild effect on the frontier orbital energies. The energies of both frontier orbitals increase by 0.12–0.14 eV. For 5,6-dihydrocytosine, methylation has a weaker effect on the frontier orbital energies than for cytosine. Just like for cytosine, the frontier orbital energies increase, but only by 0.02 eV. The conformation does not significantly influence the frontier energies. The equatorial and axial conformers display identical values within a range of 0.02 eV. The same trends are found when solvation is added. Two comments can be drawn from these observations. First, the effect of methylation is much higher when the methyl group is directly connected to an atom of a double bond than when the bond is saturated. Second, the general tendency is that methylation increases the frontier orbital energies. As a consequence, following the FMOT, methylation renders both systems saturated and unsaturated less electrophilic. The deamination rates should therefore be higher for cytosine or 5,6-dihydrocytosine compared to their methylated counterparts. This is in agreement with experimental results for 5,6-dihydro-5-methylcytosine versus 5,6-dihydrocytosine but does not apply for 5-methylcytosine versus cytosine.

Saturation of the C5–C6 bond, on the other hand, has a dramatic effect on the LUMO energies which are raised (destabilized) by up to 0.9 eV relative to their unsaturated counterparts. The HOMO levels, on the other hand, become somewhat stabilized through saturation. As a result, unsaturated forms appear softer than their saturated counterparts. Solvation has a similar effect for the saturated systems and the unsaturated ones, in that it leaves the HOMO–LUMO gap essentially intact. According to FMOT, the deamination rate of the saturated system should be lower than for their unsaturated counterpart, which is in total contradiction with experiments. The same contradiction arises from the global conceptual DFT descriptors.

*3.4.1.2 Conceptual DFT indices* Turning to the global conceptual DFT indices of chemical potential, absolute hardness, and electrophilicity, we note that although the saturated compounds overall display less negative values for the chemical potentials *in vacuum*, the data when including bulk solvation become more similar. The saturated compounds have significantly larger hardness than the unsaturated ones. This is a mathematical consequence of the larger HOMO–LUMO gaps seen for these systems relative to the unsaturated ones and can be rationalized by that the system lacks one $\pi$-bond, which implies a reduction in the polarizability of the system. The larger the absolute hardness is, the bigger the resistance to charge transfer should be, which is also reflected in the smaller electrophilicity indices, since the chemical potentials are very similar. C5-methylation appears to have a relatively modest effect on these. As noted above, the two conformers behave identically in gas phase, whereas in aqueous solution, differences between the two conformers can be noted [62]. Comparing again with the saturated 5,6-dihydrocytosine system, the equatorial conformer of 5,6-dihydro-5-methylcytosine imine deviates more in terms of chemical potential, whereas for the absolute hardness the axial form deviates more from the values seen for **dhCyt**. The global electrophilicities, however, are affected to a much less degree.

Since the next reaction step is water addition across the C4–N4 bond, it is of interest to compare the data with the corresponding values for $H_2O$. At the same level of theory, the chemical potential, absolute hardness, and global electrophilicity index in aqueous solution are −3.58, 9.44, and 0.68 eV, respectively [25]. We note that the data for the 5,6-saturated derivatives of the imine tautomers are considerably closer to these values than ones for the unsaturated systems are.

From the global reactivity indices, we conclude that methylation at C5 has essentially the same effect on the saturated and unsaturated cytosine derivatives. The explanation as to why C5-methylation of the saturated forms of cytosine reduces the deamination rate while that of the unsaturated forms raises it, thus has to be found from other considerations.

### 3.4.2 Local reactivity indexes

In Table 3, we focus on the local reactivity indices at C4 for the same systems, *in vacuum* and in aqueous solution, the rationale being that addition of water oxygen to C4 is the rate-determining step in all these deamination reactions. First of all, we note that the electrostatic charges on C4 are higher, by 0.07–0.16 *e*, in aqueous solution than *in vacuum*. The same applies to the different local reactivity indices derived from conceptual DFT, in all cases.

*3.4.2.1 Partial charges* C5–C6 saturation of both **Cyt** and **5mCyt** leads to a reduced local charge on C4. This indicates that electrostatic effects cannot be at the root of the observed phenomenon. On the other hand, saturation of the C5–C6 increases the values of local DFT-based reactivity indices, such as the grand canonical dual descriptor. This indicates that C4 becomes more electrophilic in the saturated derivatives. This rationalizes the much reduced barrier for this reaction in **dhCyt** and **5mdhCyt** over **Cyt** and **5mCyt,** respectively, and in turn explains the much increased reaction rate as seen experimentally for the saturated derivatives of cytosine.

**Table 3** Electrostatic charges on carbon C4 $q_{C4}$, local electrophilicity ($\omega^{+}_{C4}$, in eV), excess electrophilicity ($\Delta\omega_{C4}$, in eV), and grand canonical dual descriptor ($\frac{\Delta f_{C4}}{\eta^2}$, in $10^{-2}$ $eV^{-2}$) condensed to carbon C4, in vacuum (a) and in aqueous solvent (b)

| Molecular system | $q_{C4}$ | $\omega^{+}_{C4}$ | $\Delta\omega_{C4}$ | $\frac{\Delta f_{C4}}{\eta^2}$ |
|---|---|---|---|---|
| a) Vacuum | | | | |
| Cytosine imine tautomer | 0.73 | 0.11 | 0.13 | 0.33 |
| 5,6-Dihydrocytosine imine tautomer | 0.65 | 0.26 | 0.30 | 0.67 |
| 5-Methylcytosine imine tautomer | 0.63 | 0.10 | 0.15 | 0.40 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer (equatorial) | 0.56 | 0.20 | 0.24 | 0.55 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer (axial) | 0.57 | 0.21 | 0.25 | 0.58 |
| b) Aqueous solvent | | | | |
| Cytosine imine tautomer | 0.81 | 0.21 | 0.18 | 0.49 |
| 5,6-Dihydrocytosine imine tautomer | 0.73 | 0.49 | 0.49 | 1.05 |
| 5-Methylcytosine imine tautomer | 0.70 | 0.22 | 0.22 | 0.68 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer (equatorial) | 0.63 | 0.42 | 0.43 | 1.01 |
| 5-Methyl-5,6-dihydrocytosine imine tautomer (axial) | 0.65 | 0.44 | 0.45 | 1.00 |

*3.4.2.2 Fukui functions-based indices* For the 5-methylated nucleobases, the local reactivity indices increase upon saturation of the C5–C6 bond. The trends are again more pronounced in bulk solvent. Interestingly, C5-methylation of cytosine induces an increase in $\Delta\omega_{C4}$ and $(\Delta f_{C4})/(\eta^2)$ values, both *in vacuum* and in aqueous solvent, while methylation of 5,6-dihydrocytosine results in a decrease in $\Delta\omega_{C4}$ and $(\Delta f_{C4})/(\eta^2)$ values. For the unsaturated system, the increase in $(\Delta f_{C4})/(\eta^2)$ results from a lowering of the conjugation between C5=C6 double bond and C4=N3 double bond. Indeed, the substitution of C5 by a methyl group induces a polarization of the C5=C6 electron cloud toward C6. Thus, in the case of a reaction under charge transfer control, C5-methlyation is expected to have opposite effect on the local reactivity of carbon C4 of saturated and unsaturated forms of cytosine. This is in agreement with the fact that experimentally methylation at position 5 in 5,6-saturated cytosine derivatives decreases the deamination rate in contrast to what is observed with the 5,6-unsaturated cytosine derivatives. However, the range of the deamination rate decrease cannot totally be explained by a relatively small decrease in the reactivity index values. It is therefore likely that besides these electronic effects, the steric hindrance of the methyl group affects the approach of the water molecules. The methyl group probably acts as a shield for at least one side of the ring system, especially when it is in axial position. As both axial and equatorial positions lie very close in energy, the odds that the methyl group is in an axial orientation during the process are considerable.

## 4 Conclusions

We have in the current study investigated the deamination of 5,6-dihydro-5-methylcytosine. The investigation has been carried out in close comparison with deamination of cytosine, 5-methylcytosine, and 5,6-dihydrocytosine. It appears that the deamination mechanism is very similar to that of cytosine and cytosine derivatives. Just like for cytosine and cytosine derivatives, the imine tautomer is the reactive form for the deamination process. But in contrast to cytosine and 5-methylcytosine, the imine tautomer is produced before the deamination and constitutes the more stable form for the saturated systems. Experimentally, the deamination rates is lower for 5,6-dihydro-5-methylcytosine than for 5,6-dihydrocytosine. This fact has been rationalized herein through quantum chemical calculations and conceptual DFT-based descriptors. It appears that besides the electronic effects of the methyl group on the carbon C4 of cytosine, steric hindrance caused by the methyl group, especially when it is in axial position, shields the approach of the reacting water molecules.

**Acknowledgments** The Faculty of Science at the University of Gothenburg is gratefully acknowledged for financial support (LAE). This research has beneficiated of the support of Aviesan ITMO cancer within the "Plan Cancer 2009–2013" and the application of the action 3.3.

## References

1. Attwood JT, Yung RL, Richardson BC (2002) Cell Molr Life Sci 59:241
2. Ramsahoye BH, Davies CS, Mills KI (1996) Blood Rev 10:249
3. Shen J-C, Rideout WM III, Jones PA (1994) Nucleics Acids Res 22:972
4. Cortazar D, Kunz C, Saito Y, Steinacher R, Schar P (2007) DNA Repair(Amst) 6:489
5. Pfeifer GP (2006) Current topics in microbiology and immunology. In: DNA methylation: basic mechanisms, vol 301, Part IV. Springer, pp 259–281
6. Slae S, Shapiro R (1978) J Org Chem 43:1721
7. Green M, Cohen SS (1957) J Biol Chem 228:601
8. Decarroz C, Wagner JR, van Lier JE, Murali KC, Riesz P, Cadet J (1986) Int J Radiat Biol 50:491
9. Tremblay S, Douki T, Cadet J, Wagner R (1999) J Biol Chem 274:20833

10. Wagner JR, Hu CC, Ames BN (1992) Proc Natl Acad Sci USA 89:3380–3384
11. Boorstein RJ, Hilbert TP, Cunningham RP, Teebor GW (1990) Biochemistry 29:10455–10460
12. Douki T, Vadesne-Bauer G, Cadet J (2002) Photochem Photobiol Sci 1:565
13. Freeman KB, Hariharan PV, Johns HE (1965) J Mol Biol 13:833–848
14. Lemaire DGE, Ruzsicska BP (1993) Photochem Photobiol 57:755
15. Lemaire DGE, Ruzsicska BP (1993) Biochemistry 32:2525
16. Douki T, Cadet J (1992) J Photochem Photobiol B Biol 15:199–213
17. Liu F-T, Yang NC (1978) Biochemistry 17:4865
18. Bienvenu C, Cadet J (1996) J Org Chem 61:2632
19. Cannistraro VJ, Taylor JS (2009) J Mol Biol 392:1145
20. Celewicz L, Mayer M, Shetlar MD (2005) Photochem Photobiol 81:404
21. Douki T, Cadet J (1994) Biochemistry 33:11942
22. Labet V, Grand A, Morell C, Cadet J, Eriksson LA (2008) Theor Chem Acc 120:429
23. Labet V, Morell C, Grand A, Toro-Labbe A (2008) J Phys Chem A 112:11487
24. Labet V, Morell C, Cadet J, Eriksson LA, Grand A (2009) J Phys Chem A 113:2524
25. Labet V, Morell C, Douki T, Cadet J, Eriksson LA, Grand A (2010) J Phys Chem A 114:1826
26. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) *Gaussian 09*, Revision A.02. Gaussian, Inc, Wallingford, CT
27. Becke AD (1993) J Chem Phys 98:5648–5652
28. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
29. McLean AD, Chandler GS (1980) J Chem Phys 72:5639
30. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) J Chem Phys 72:650
31. Frisch MJ, Pople JA, Binkley JS (1984) J Chem Phys 80:3265
32. Cancès MT, Mennucci B, Tomasi J (1997) J Chem Phys 107:3032
33. Mennucci B, Tomasi J (1997) J Chem Phys 106:5151–5158
34. Cossi M, Scalmani G, Rega N, Barone V (2002) J Chem Phys 117:43–54
35. Parr RG, Yang W (1989) Density functional theory of atoms and molecules. Oxford University Press, New York
36. Geerlings P, De Proft F, Langenaeker W (2003) Chem Rev 103:1793–1874
37. Chattaraj PK, Sarkar U, Roy DR (2006) Chem Rev 106:2065
38. Ayers PW, Anderson JSM, Bartolotti LJ (2005) Int J Quantum Chem 101:524
39. Chermette H (1999) J Comp Chem 20:129
40. Parr RG, Yang W (1995) Annu Rev Phys Chem 46:701
41. Parr RG, Donnelly RA, Levy M, Palke WE (1978) J Chem Phys 68:3801
42. Parr RG, Pearson RG (1983) J Am Chem Soc 105:7512
43. Mineva T, Sicilia E, Russo N (1998) J Am Chem Soc 120:9053
44. Parr RG, Szentpaly Lv, Liu SB (1922) J Am Chem Soc 1999:121
45. Morell C, Labet V, Grand A, Chermette H (2009) Phys Chem Chem Phys 11:3417
46. Koopmans TA (1934) Physica 1:104
47. Parr RG, Zhou Z (1993) Acc Chem Res 26:256
48. Chattaraj PK, Maiti B, Sarkar U (2003) J Phys Chem A 107:4973
49. Parr RG, Yang W (1984) J Am Chem Soc 106:4049
50. Ayers PW, Levy M (2000) Theor Chem Acc 103:353
51. Mineva T, Parvanov V, Petrov I, Neshev N, Russo N (1959) J Phys Chem A 2001:105
52. Padmanabhan J, Parthasarathi R, Elango M, Subramanian V, Krishnamoorthy BS, Gutierrez-Oliva S, Toro-Labbé A, Roy DR, Chattaraj PK (2007) J Phys Chem A 111:9130
53. Morell C, Grand A, Toro-Labbé A (2005) J Phys Chem A 109:205
54. Morell C, Grand A, Toro-Labbé A (2006) Chem Phys Lett 425:342
55. Morell C, Ayers PW, Grand A, Chermette H (2011) Phys Chem Chem Phys 13:9601
56. Cardenas C, Rabi N, Ayers P, Morell C, Jaramillo P, Fuentealba P (2009) J Chem Phys A 113:8860
57. Ayers PW, Morell C, De Proft F, Geerlings P (2007) Chem Eur J 13:8240
58. Breneman CM, Wiberg KB (1990) J Comp Chem 11:361
59. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88:899
60. Toro-Labbé A (1999) J Phys Chem A 103:4398
61. Llano J, Eriksson LA (2002) J Chem Phys 117:10193
62. De Luca G, Sicilia E, Russo N, Mineva T (2002) J Am Chem Soc 124:1494