J Mol Model (2013) 19:2993–3005
DOI 10.1007/s00894-013-1813-0

ORIGINAL PAPER

Comparison of gas phase intrinsic properties of cytosine and thymine nucleobases with their O-alkyl adducts: different hydrogen bonding preferences for thymine versus O-alkyl thymine

Zahra Aliakbar Tehrani · Alireza Fattahi

Received: 15 January 2012 / Accepted: 3 March 2013 / Published online: 7 April 2013

Abstract In recent years, there has been increasing interest in damaged DNA and RNA nucleobases. These damaged nucleobases can cause DNA mutation, resulting in various diseases such as cancer. Alkylating agents are mutagenic and carcinogenic in a variety of prokaryotic and eukaryotic organisms. The present study employs density functional theory (DFT/B3LYP) with the 6-311++G(d,p) basis set to investigate the effect of chemical damage in O-alkyl pyrimidines such as O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine. We compared the intrinsic properties, such as proton affinities, gas phase acidities, equilibrium tautomerization and nucleobase pair's hydrogen bonding properties, of these molecules with those in the normal nucleobases thymine and cytosine. The results are of interest for chemical reasons and also possibly for biological purposes since biological media can be quite non-polar. Furthermore, we found that N1-H of O^4-methylthymine is less acidic than N1-H of thymine, suggesting that alkyl DNA glycosylase enzyme cannot discriminate this damaged nucleobase from a normal thymine nucleobase. This result indicates that the conjugated base anion of O^4-methylthymine would be a worse leaving group and O^4-methylthymine is repaired in genome by demethylation rather than enzyme-catalyzed excision at N1.

Keywords Alkylation damage · Tautomer · O^4-methylthymine · Atom in molecule analysis · Demethylation

Z. Aliakbar Tehrani · A. Fattahi (✉)
Department of Chemistry, Sharif University of Technology,
PO Box 11365-9516, Tehran, Iran
e-mail: Fattahi@sharif.ir

Introduction

DNA is the molecule chosen by nature to store the information required to build organisms. These organisms in turn serve to replicate that information. The chemical structure of DNA was proposed in the seminal article of Watson and Crick in 1953 [1] and consists of two double helices of nucleotides linked by stabilizing hydrogen bonds formed between adenine-thymine (A:T) and guanine-cytosine (G:C) base pairs (Fig. 1). The hydrogen bond formation of a Watson-Crick type base pair is fundamental to molecular recognition in the duplex formation of nucleic acids. It is also essential for the transmission of genetic information [2]. Because of their biological activity, most substituted (or modified) nucleic bases have been frequently studied experimentally [3].

It was initially thought that DNA must be incredibly stable to maintain the integrity of the information, but it has been shown that DNA is in fact a dynamic molecule that is constantly damaged. In recent years, there has been increasing interest in DNA damage. Such damage may cause DNA mutation, resulting in various diseases such as cancer [4–8]. Many experimental and theoretical efforts have been directed at investigating the possible consequences of DNA cleavage, the mechanisms of DNA damage and the corresponding structural changes of base pairs, but still there are processes that are not well understood.

Damage can be caused by exogenous sources such as man-made mutagenic substances and naturally occurring agents including sunlight and dietary mutagens, and endogenous sources such as reactive oxygen species (ROS) formed during cellular metabolism. Moreover, DNA molecules are made up of many nucleophilic centers at which DNA damage can occur. The chemical species that react with these electron-rich atoms are typically electrophilic and form adducts where the

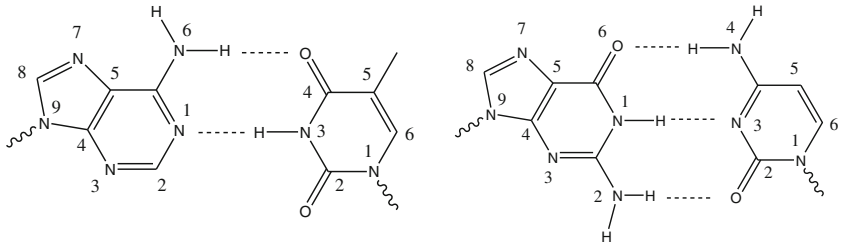


Fig. 1 Structures of adenine-thymine (*A-T*) and guanine-cytosine (*G-C*) Watson-Crick base pairs

electrophile is linked by one or more covalent bonds to the base, deoxyribose sugar or phosphate of the DNA. DNA damaging agents either act directly on target sites on DNA or are metabolically activated by enzymatic means.

DNA damage can result in nucleobase loss, nucleobase dimerization, alkylation, deamination, and oxidation as well as single- or double-strand breakage, leading to permanent changes in the information encoded by the DNA. Alterations in base sequence can also arise as a result of replication and recombination. Without maintenance, the information encoded by DNA would be altered so dramatically that the organism could not thrive. Nature has therefore devised a solution to this problem: DNA repair. A number of DNA repair systems have evolved, including direct damage reversal, nucleotide excision repair (NER), mismatch repair, base excision repair, and recombinational repair.

DNA alkylating agents are mutagenic, carcinogenic, or both in a variety of prokaryotic and eukaryotic organisms. Although the precise molecular mechanisms underlying mutagenesis by these agents are unknown, there is strong evidence suggesting that base-substitution mutations arise via the formation of alkyl-DNA adducts that direct the misincorporation of nucleotides during DNA replication [9]. These agents invade DNA by covalently binding to nucleophilic sites on the nucleophilic molecule leading to DNA-alkyl adducts. These agents modify almost all of the heteroatoms in DNA that could be potentially alkylated in the double helix. The preferred sites of alkylation in duplex DNA depend strongly on the nature of the alkylating agent. For example, in the case of diethylsulfate, the preferred sites of reaction follow the order: N7 of guanine>> P-O > N3 of adenine>> N1 of adenine ~ N7 of adenine ~ N3 of guanine ~ N3 of cytosine >>O6 of guanine [9, 10]. The preferences observed at these sites have been rationalized commonly in terms of hard-soft reactivity principles [11]. Hard alkylating agents (defined by small size, positive charge, and low polarizability) such as diazonium ions display increased reactivity with hard oxygen nucleophiles in DNA [9, 10, 12–14]. On the other hand, soft (large, uncharged, and polarizable) alkylating agents like dialkylsulfates favor reactions at the softer nitrogen centers in DNA.

Alkylating agents such as N-methyl-N′-nitro-N-nitrosoguanidine and N-methyl-N-nitrosourea (MNU) form reaction products at DNA oxygen and nitrogen atoms [15]. *O*-Alkylation is less common but more damaging; for example, O^2 alkylation of deoxycytidine, deoxyuridine and thymidine increased the rate of *N*-glycoside hydrolysis 10^4 fold [16]. Among oxygen adducts formed by these agents are O^6-methylguanine and O^4-methylthymine, which are widely believed to be the major sites of alkylation mutagenesis. For both DNA adducts, the mutational mechanism is thought to involve passive mispairing with incorrect nucleotides during replication [17]. MNU and related methylating agents produce G:C to A:T transitions almost exclusively [18, 19], and this change is also the dominant mutation induced by synthetic O^6-methylguanine in DNA in vitro and in vivo [20–22].

O^4-methylthymine arising from alkylation of thymine is highly mutagenic. Its mutagenicity arises from a A:T base pair to G:C base pair transition that occurs as a result of different hydrogen bonding preferences for normal thymine versus damaged O^4-methylthymine [20, 21, 23–26]. In DNA replication, thymine nucleobase pairs preferentially with adenine. However, once thymine is alkylated to form O^4-methylthymine, the preferred hydrogen bonding pattern is to guanine rather than to adenine. The result is that a DNA sequence that originally contained an A:T base pair will become an O^4-methylthymine-guanine base pair. When the strand containing guanine replicates, the end result will be an G:C base pair; this is A:T to G:C transition (Fig. 1). Because the exact sequence of DNA is necessary for proper life function (coding for proteins, signaling), the mutation to O^4-methylthymine is known to be highly carcinogenic [19, 26].

Many theoretical studies have been carried out on the interaction energies of base pairs between natural nucleic acid bases, but few systematic studies on damaged base pairs have been reported. Theoretical methods provide a powerful tool to aid in the interpretation of experimental data and to describe the effects of chemical modifications, including their effect on the local electronic structure properties of individual damaged DNA nucleobases, base pairs, and their more global effect on the helical base stack. These factors will influence their biologically relevant chemical reactivity.

In the present work,[1] we provide a comprehensive theoretical examination of the gas phase thermochemical

[1] Presented at the Spring 2010 meeting of the ACS Division of Physical Chemistry, Multiscale Nanomaterials, Polymer & Bimolecular Dynamics

properties of naturally occurring pyrimidine nucleobase cytosine and thymine as well as damaged nucleobases such as O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine (Fig. 2) by employing density functional theory (DFT) (B3LYP) with the 6-311++G(d,p) basis set. The goal of this study was to provide insight into the electronic properties, hydrogen bonding pattern, proton affinity, gas phase acidity and equilibrium tautomerization of these molecules. We discuss the interesting biological implications of these results.

Computational methods

Initial searches for minima on the potential energy surface of tautomers of O-alkylated nucleobases and/or their deprotonated/protonated forms at the relative energy range of 10.0 kcalmol^{-1} were carried out using the MMFF force field in Spartan software [27]. The most stable conformers were optimized by the DFT method using Becke3 (B3) exchange [28] and Lee, Yang, and Parr (LYP) correlation [29] potentials, in connection with the 6–311++G (d,p) basis set. Energy minimizations followed by harmonic vibrational calculations were performed at this level of theory. The absence of imaginary frequencies proved that energy-minimized structures correspond well to the local minima of the energy landscape (local minima were verified by establishing that the matrix of energy second derivatives has only positive eigenvalues). Bulk solvation effects on the stability order of tautomers were included in the series of single-point energy calculations on the optimized structures obtained from gas phase, through the integral equation formalism of the polarized model (IEF-PCM) [30]. The dielectric constant ε=78.4 was employed to model aqueous solution.

Natural bond orbital (NBO) analysis [31, 32] was performed at DFT-B3LYP/6-311++G (d,p) level. This procedure examines all possible interaction between filled (donor) Lewis-type NBOs and empty (acceptor) non-Lewis NBOs and estimating their energetic importance by the second-order perturbation interaction energy. For each donor NBO (i) and acceptor NBO (j), the stabilization energy $E^{(2)}$ associated with delocalization i/j is estimated as

$$E^{(2)} = \Delta E_{ij} = q_i \frac{F(i,j)^2}{\varepsilon_j - \varepsilon_i}$$

where, q_i is the donor orbital occupancy, ε_i; ε_j are diagonal elements (orbital energies) and F(i,j) is the off-diagonal NBO Fock matrix element.

Furthermore, electron densities $\rho(r)$ and Laplacian of electron densities $\nabla^2\rho(r)$ of various hydrogen bonds in non-classical and/or classical base pairs at bond critical points (BCPs) were calculated at DFT-B3LYP/6-311++G (d,p) level using Bader's theory of atoms in molecule (AIM) [33, 34]. AIM is very useful tool for analyzing hydrogen bonds, with a large electronic density at hydrogen BCP and positive value of $\nabla^2\rho(r)$ indicating a strong hydrogen bond [35, 36]. The topological analysis was performed using the AIM2000 program [37].

Results and discussion

Tautomerization of selected O-alkylated pyrimidine nucleobases

O-alkylated pyrimidine nucleobases

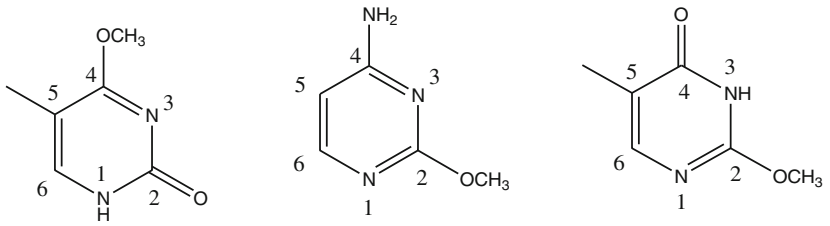
Tautomerism is a well-known phenomenon occurring in nucleic acid bases [38–57] in which a proton transfers from the heterocyclic ring nitrogen to an exocyclic oxo- or imino- group, leading to the formation of either an −OH or an $-NH_2$ functionality. The structure of the nucleobase is such that several tautomers are often possible and O-alkylated pyrimidine nucleobases are no exception. Structures of all possible tautomers of O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine are shown in Fig. 3. Furthermore, B3LYP/6-311++G (d,p) relative energies (ΔE, kcal mol^{-1}), relative free energies (ΔG, kcal mol^{-1}), relative stability in aqueous solution (ΔE_{aq}, kcal mol^{-1}) and dipole moments (μ in Debye) of selected O-alkylated pyrimidine nucleobases are collected in Table 1.

O^4-methylthymine

O^4-methylthymine has three possible tautomeric forms (structures **1a**, **2a** and **3a** in Fig. 3). The computation revealed energy disparity of 0.0 to 17.5 kcalmol^{-1} for this molecule. At B3LYP/6-311++G (d,p) level, we found that tautomer **1a** (where the proton resides on N1 atom) is the most stable tautomer of O^4-methylthymine. The next most stable **2a** tautomer (proton resides on O2 atom) is 2.7 kcal mol^{-1} less stable than **1a**. Results of calculations revealed that the **3a** tautomer, where the proton resides on the N3 atom, is 11.8 kcalmol^{-1} less stable than **1a**. The energy difference between tautomers **1b** and **3a** is large enough to consider that, in gas phase populations, the **3a** tautomer is negligible. To explore the relative stability order of cytosine and thymine nuceloabses as well as their O-alkylated adducts, we used the IEF-PCM model. As can seen from Table 2, the relative stability order of O^4-methylthymine tautomers in aqueous solution is consistent with the stability order of this molecule in gas phase.

O^2-methylcytosine

The structures of three most stable tautomers of O^2-methylcytosine are shown in Fig. 3. As shown in Table 1,



O^4-methyl thymine (**1**) O^2-methyl cytosine (**2**) O^2-methyl thymine (**3**)

Fig. 2 Chemical structures and atom numbering of the O-alkylated pyrimidine nucleobases considered in this study

the energy gap between tautomers of this molecule is higher than those of O^2-methylthymine and O^4-methylthymine molecules. Based on DFT calculations, structure **1b** was found to be the most stable tautomer of O^2-methylcytosine. The **2b** tautomer (protons reside on N3 and N4 atoms) is the next most stable tautomer, lying 18.5 kcalmol^{-1} higher in energy than **1b**. **2b** is followed by the **3b** structure (protons residing on N1 and N4 atoms), which is 0.6 kcalmol^{-1} less stable than **2b**. The energy difference between the **1b** and **2b**/**3b** tautomer is large enough to consider that, at least in the gas phase, populations of these tautomeric forms are negligible. As seen from Table 2, at B3LYP/6-311++G (d,p) level of theory in aqueous solution the most stable tautomer of O^2-methylcytosine molecule is **1b** tautomer followed by **2b** tautomer (lying 15.1 kcalmol^{-1} higher in energy than **1b**) and **3b** tautomer (lying 16.8 kcalmol^{-1} higher in energy than **1b**).

O^2-methylthymine

As shown in Fig. 3, O^2-methylthymine has three possible tautomeric forms as shown in Fig. 3. Results of calculation reveal that the canonical **1c** tautomer (where proton resides on N3 atom) is the most stable tautomer of O^2-methylthymine. Two other tautomers, **2c** and **3c** (where protons reside on O4 and N1 atoms, respectively), are 9.8 and 17.5 kcalmol^{-1} less stable than the **1c** tautomer (see Table 1 for more details). The energy difference between the **1c** and **3c** tautomers is large enough to consider that, in the gas phase, populations of **3c** tautomer are negligible. As seen from Table 2, at B3LYP/6-311++G (d,p) level of theory in aqueous solution, the most stable tautomer of O^2-methylthymine molecule is **1c** followed by **2c** and **3c**. This stability

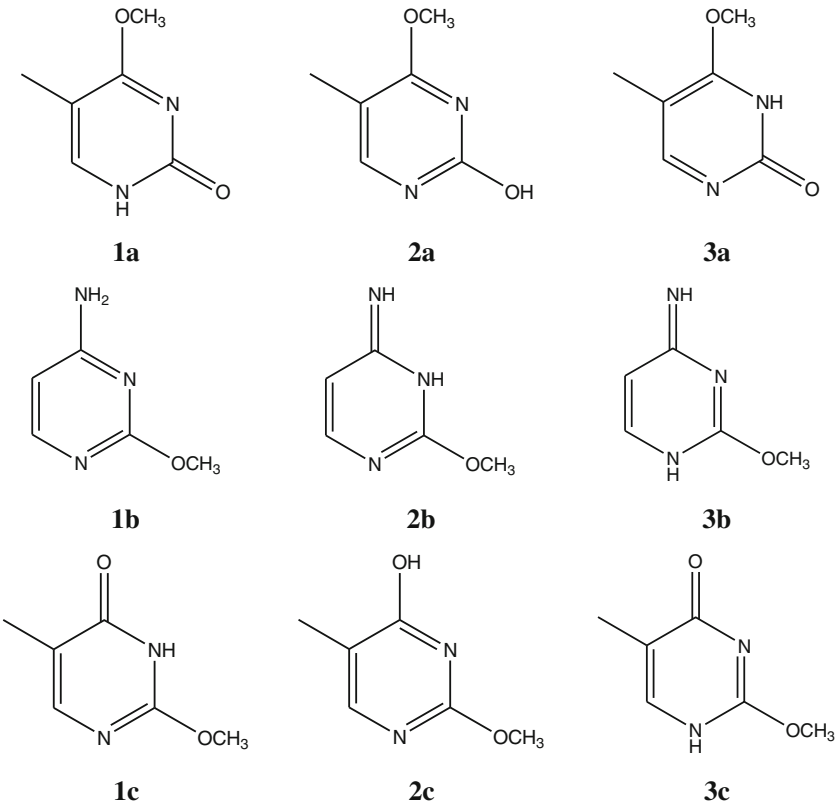


Fig. 3 Structures of possible tautomers of selected O-alkylated pyrimidine nucleobases calculated at B3LYP/6-311++G (d,p) level of theory

Table 1 Density functional theory (DFT)-B3LYP/6-311++G (d,p) calculated thermodynamic parameters for tautomers of selected O-alkylated pyrimidine nucleobases

System	Tautomer	μ (Debye)	ΔE_{298} (kcal mol^{-1})	ΔG_{298} (kcal mol^{-1})	ΔE_{aq}(kcal mol^{-1})
O^4-methylthymine	**1a**	3.63	0.0	0.0	0.0
	2a	4.08	2.7	2.9	6.7
	3a	9.19	11.8	11.2	9.5
O^2-methylcytosine	**1b**	4.33	0.0	0.0	0.0
	2b	1.76	18.5	18.5	15.1
	3b	5.44	19.1	18.8	16.8
O^2-methylthymine	**1c**	5.12	0.0	0.0	0.0
	2c	3.69	9.8	9.9	8.4
	3c	7.35	17.5	17.0	9.0

order is consistent with stability order of this molecule in gas phase.

Cytosine nucleobase

There have been numerous computational studies on the tautomeric equilibrium of the nucleobase cytosine. On the basis of these theoretical studies, the lowest energy conformers of this nucleobase in gas phase have been identified unambiguously. The optimized structures of the six most stable cytosine tautomers [54, 56, 58–60] determined from theoretical calculations performed at the B3LYP method in conjunction with the 6-311++G (d,p) basis set are depicted in Fig. 4.

The total energies, relative energies, relative free energies and dipole moment values of tautomers obtained at B3LYP/6-311++G (d,p) level are given in Table 2. The keto-amino form (**C1**) is the "canonical" structure of cytosine found in DNA and RNA. Indeed, X-ray diffraction (XRD) [61, 62] and neutron diffraction [63] studies find that this is the only tautomeric form of cytosine nucleobase. In contrast, resonance enhanced multiphoton ionization (REMPI) experiments [45] find that both keto-amino (**C1**) and enol-amino (**C2**) tautomers coexist in the gas phase. Mixtures of C1 and C3 as well as the keto-imino ($\mathbf{C_5}$) tautomers have also been observed in infra-red (IR) matrix isolation studies [64]. Similarly, molecular beam microwave (MW) spectroscopy studies indicate that **C1**, **C2** and **C5** tautomers coexist in the gas phase [65]. It should be noted that three of the six low-energy tautomers of isolated cytosine, i.e., **C2**, **C3**, and **C4**, are not accessible in DNA and RNA because the ribose would not migrate.

In agreement with most previous studies, our results reveal that the **C1** tautomer is the most stable tautomer in gas phase in solid cytosine. The **C2** tautomer was calculated to be less stable than **C1** by 1.2 kcalmol^{-1}. The **C2** tautomer is the next most stable gas phase tautomer, lying 2.1 kcalmol^{-1} higher in energy than the **C1** tautomer. The **C6** tautomer is found to be the next most stable cytosine tautomer, lying 6.9 kcalmol^{-1} higher in energy than the **C1** tautomer. Tautomers **C4** and **C5** are found to be least stable tautomers, lying 14.2 and 22.4 kcal mol^{-1} higher in energy than the **C1** tautomer. The energy difference between **C1** and **C4**/**C5** and **C6** tautomers is large enough to consider that populations of these tautomers are negligible in the gas phase.

In summary, cytosine tautomers have the following stability sequence predicted at B3LYP/6-311++G(d,p) level of theory: **C1** > **C2** > **C3** > **C4** > **C5** > **C6**. Moreover, as seen from Table 2, comparison of the relative stability order of

Table 2 B3LYP/6-311++G relative energies (ΔE, kcal mol^{-1}), relative free energies (ΔG, kcalmol^{-1}) relative stability in aqueous solution (ΔE_{aq}, kcalmol^{-1}), and dipole moments (μ in Debye) at 298 k for different cytosine and thymine tautomers

Tautomer species	ΔE_{298}	ΔG_{298}	μ (Debye)	ΔE_{aq}(kcal mol^{-1})
C1	0.0	0.0	6.75	0.0
C2	1.2	1.4	3.47	4.0
C3	1.9	2.3	4.94	5.1
C4	6.9	7.0	8.25	6.9
C5	14.2	14.6	1.84	19.5
C6	22.4	22.5	5.65	22.6
T1	0.0	0.0	8.34	0.0
T2	3.6	3.9	6.23	6.5
T3	11.7	12.2	4.58	15.0
T4	15.2	15.7	3.57	16.2
T5	18.4	18.6	2.91	17.5
T6	19.7	19.9	1.77	18.3

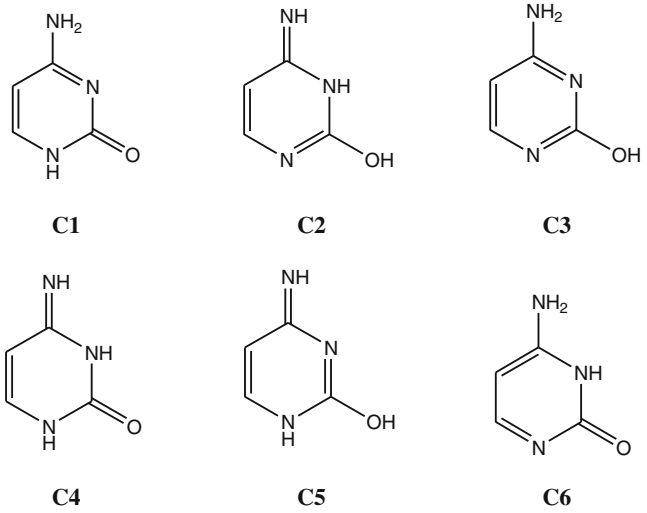


Fig. 4 Chemical structures of cytosine tautomers investigated in this study

uracil tautomers in aqueous solution with results calculated in the gas phase indicates that the solvent has different effects on the stability of these tautomers. The existence of solvent increases the stability of the **C4** tautomer and decreases the stability of the **C3** tautomer. However, the **C1** tautomer is the most stable in solvent as well as gas phase (cf. Table 2) followed by **C2**,**C4**, **C3**, **C5** and **C6**, respectively. The latter are located at 4.0, 5.1, 6.9, 19.5, and 22.6 kcal mol^{-1} above the **C1** tautomer, respectively.

Thymine nucleobase

According to previous reports [46, 66, 67], thymine nucleobase has six expected tautomers (shown in Fig. 5). DFT-B3LYP/6-311++G (d,p) relative energies (ΔE, kcal mol^{-1}), relative free energies (ΔG, kcal mol^{-1}) and dipole moments (μ in Debye) for these tautomers are given in Table 2; the canonical **T1** tautomer of thymine (protons resides on N1 and N3 atoms) is the most stable tautomer [24, 68, 69]. It is calculated to be about 3.6 kcal mol^{-1} more stable than the next nearest tautomer (i.e., **T2** tautomer in which protons reside on O2 and N3 atoms). The next most stable tautomers are **T3** (protons reside on N1 and O2 atoms) and **T4** (protons reside on O2 and O4 atoms), which are 11.7 and 15.2 kcal mol^{-1} less stable than the canonical tautomer **T1**. The remaining tautomers are all quite high in energy. On the basis of aqueous solution and gas phase calculations tautomer **T1** is found to be the most stable tautomeric form of the thymine nucleobase (cf. Table 2). **T1** is followed by **T2**, **T3**, **T4**, **T5** and **T6** tautomers, in that order.

The energy difference between **T1** and **T3**, **T4**, **T5** and **T6** tautomers is large enough to consider that populations of these tautomers are negligible in the gas phase.

Comparison of normal nucleobase properties with those of O-alkylated analogues

Damaged DNA nucleobases differ in structure and properties from normal nucleobases and, therefore, interfere with gene replication and expression, leading to cell death, aging, and carcinogenesis. Furthermore, hydrogen bonding modulates recognition of DNA and RNA nucleobases. The interaction energy between two complementary nucleobases that are held together by NH···O and NH···N hydrogen bonds depends on the intrinsic basicity of the acceptor atoms as well as on the acidity of the NH donor groups. Elucidating the intrinsic reactivity of these molecules can improve understanding of key biosynthetic mechanisms for which those nucleobases are substrates.

The first step toward understanding how normal nucleobases differ from damaged ones is to characterize the naturally occurring normal compounds. Herein, we undertook a comprehensive examination of the gas phase thermochemical properties of the naturally occurring nucleobases cytosine and thymine, as well as their O-alkylated analogues (i.e., O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine). We measured acidities of multiple sites as well as proton affinities of these molecules. $\Delta H_{acidity}$ were calculated as the enthalpy changes of the following process:

$$\mathbf{HB} \rightarrow \mathbf{H}^+ + \mathbf{B}^-$$
$$\Delta \mathbf{H}^{298} = \Delta \mathbf{E}^{298} + \Delta(\mathbf{PV}) = \Delta \mathbf{E}^{298} + \Delta \mathbf{n_g RT}$$
$$\Delta \mathbf{H}^{298} = \mathbf{E}\ (\mathbf{B}^-)^{298} - \mathbf{E}(\mathbf{HB})^{298} + (\mathbf{5/2})RT$$

where E^{298} represents calculated energy including thermal vibrational corrections and HB and B^- represents selected acidic sites and their conjugated bases in O-alkylated adducts. The (5/2) RT term includes translation energy of proton and Δ(PV) term. The computational acidity value (black values, $\Delta H_{acidity}$ in kcal mol^{-1}) and proton affinities (red values, PA in kcal mol^{-1}) for multiple sites of most stable tautomers of cytosine, thymine, O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine are shown in Fig. 6.

Cytosine nucleobase

The canonical form of cytosine nucleobase has three potentially acidic amino groups: N1-H, N4-H_a and N4-H_b protons (see Fig. 4 for atom numbering). For this keto tautomer, the most acidic site was reported to be N1, followed by N4-H_b and N4-H_a, with gas phase acidities of 345.1, 347.9 and 353.8 kcal mol^{-1}, respectively (see Fig. 6 for more details). Gas phase acidity of N4-H_a is about 5 kcal mol^{-1} higher than that of N4-H_b due to the repulsion between the electron lone pairs centered at N3 and at N4. It is worth mentioning that gas phase acidity values for N1-H and N4-H_b sites are in excellent

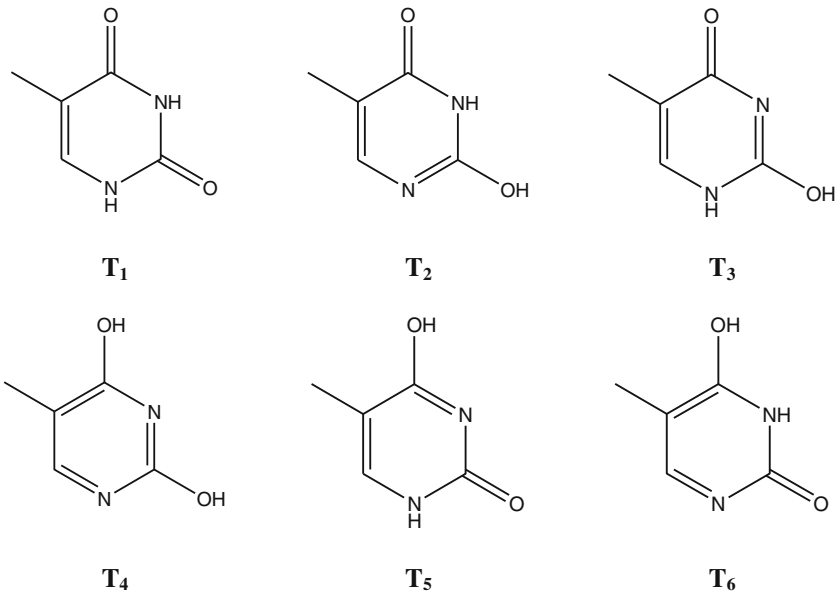


Fig. 5 Chemical structures of thymine tautomers investigated in this study

agreement with experimental values of 342±3 and 352±4 kcalmol^{-1} measured using the bracketing method [69]. Moreover, the acidity value of the most acidic site of cytosine measured in this study is in agreement with that previously measured by Chen and Chen (340±2 kcalmol^{-1}) using a variety of methods such as derivation from DMSO pK_a values, electron impact, and negative chemical ionization mass spectrometry [70, 71]. On the other hand, the canonical **C1** tautomer of cytosine nucleobase has two sites that are close in basicity: O2 atom (PA=219.2 kcalmol^{-1}) and N3 atom (PA=226.0 kcalmol^{-1}) (see Fig. 6 for more details). These values are in agreement with previously measured values [72, 73].

Thymine nucleobase

The acidity of the canonical tautomer of thymine nucleobase at the most acidic site, N1, is calculated to be $\Delta H_{acidity}$=336.2 kcalmol^{-1}, which is also in agreement with a previously measured value of 333.0±2 kcalmol^{-1} using different methods [74, 75]. The N3-H site is less acidic with $\Delta H_{acidity}$=342.5 kcalmol^{-1} (see Fig. 6 for more details). Proton affinity measurements yield a value of PA=119.2 kcalmol^{-1} and PA=201.3 kcalmol^{-1} for the O2 and O4 basic sites, respectively, which agrees with previous measurements (PA=210.5 kcalmol^{-1}) [76–80].

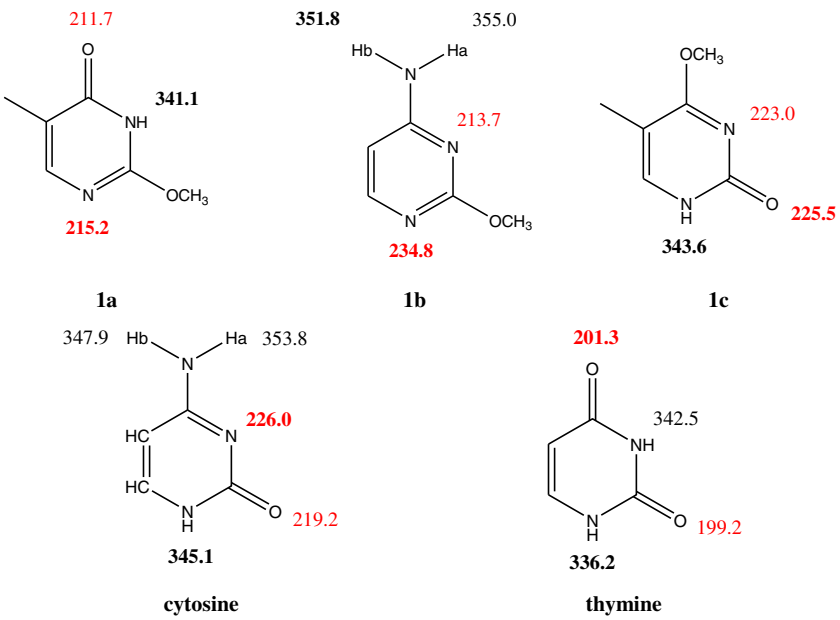


Fig. 6 Calculated acidities ($\Delta H_{acidity}$) and proton affinities (PA in kcal mol^{-1}) of the most stable tautomers of cytosine, thymine, O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine at B3LYP/6-311++G (d,p) level. The most acidic and basic sites are in *bold*

O^4-methylthymine

The most stable tautomer of O^4-methylthymine, **1a**, is predicted to have gas phase acidity value of 343.6 kcalmol^{-1}, for deprotonation of the N1 site because the N3 site is blocked by an imino bond (see Fig. 6). NBO analysis shows that, in anions generated from deprotonation of this molecule, a strong hyperconjugation interaction exists between lone pairs in anionic centers and the antibonding orbitals of their neighboring atoms. Generally, the more (less) acidic positions are those that generate the most (least) stable anionic species. Values of $E^{(2)}$ for orbital interactions of N1 anionic center of the **1a** tautomer generated during deprotonation are: $LP_{N1} \rightarrow BD^*_{C2\text{-}N3}$ (11.3 kcalmol^{-1}), $LP_{N1} \rightarrow BD^*_{C2\text{-}O2}$ (5.8 kcalmol^{-1}), $LP_{N1} \rightarrow BD^*_{C5\text{-}C6}$ (7.3 kcalmol^{-1}) and $LP_{N1} \rightarrow BD^*_{C6\text{-}H6}$ (2.8 kcalmol^{-1}), respectively.

O^2-methylcytosine

The gas phase acidities for protons of the most stable tautomeric forms of O^2-methylcytosine, **1b**, calculated at B3LYP/6-311++G (d,p) level of theory are shown in Fig. 6. The most acidic site of this tautomer is N4-H_b with a $\Delta H_{\text{acidity}}$ value of 351.8 kcalmol^{-1}. The N4-H_a is less acidic at 355.0 kcalmol^{-1}. As shown by NBO analysis, the negative charge at N4-H_b deprotonated form of **1b** tautomer is stabilized by $LP_{N4} \rightarrow BD^*_{N3\text{-}C4}$ (13.0 kcalmol^{-1}), $LP_{N4} \rightarrow BD^*_{C2\text{-}N3}$ (0.5 kcalmol^{-1}) and $LP_{N4} \rightarrow BD^*_{C4\text{-}C5}$ (0.5 kcalmol^{-1}). In N4-H_a, the deprotonated structure of this molecule $LP_{N4} \rightarrow BD^*_{N3\text{-}C4}$ (1.8 kcalmol^{-1}) and $LP_{N4} \rightarrow BD^*_{C4\text{-}C5}$ (11.4 kcalmol^{-1}) interactions play an important role in stabilization of negative charge on the N4 atom. Furthermore, the results of calculations revealed that the **1b** tautomer of O^2-methylcytosine has two positions that could accept a proton: the most basic site N1 with proton affinity of 234.8 kcalmol^{-1} and the N3 site (the next most basic position with a PA value of 213.7 kcalmol^{-1}).

O^2-methylthymine

In the **1c** tautomer of O^2-methylthymine the most acidic site is the N3-H proton, with a $\Delta H_{\text{acidity}}$ value of 340.1 kcalmol^{-1} (see Fig. 6). For this anion, the value of second perturbation energies of the LP → BD* types of interactions are: $LP_{N3} \rightarrow BD^*_{C2\text{-}O2}$ (2.5 kcalmol^{-1}), $LP_{N3} \rightarrow BD^*_{C2\text{-}N3}$ (10.2 kcalmol^{-1}), $LP_{N3} \rightarrow BD^*_{C5\text{-}C6}$ (9.7 kcalmol^{-1}) and $LP_{N3} \rightarrow BD^*_{C6\text{-}H6}$ (3.6 kcalmol^{-1}), respectively. Moreover, the results of calculation revealed that the most basic site of this molecule is N1 atom, with a proton affinity value of 215.2 kcalmol^{-1}. The O4 is the next basic position with a proton affinity value of 211.7 kcalmol^{-1}.

Hydrogen bonding interaction and base pair properties

As mentioned in the Introduction, the O-alkylated adducts of cytosine and thymine nucleobases that arise from exposure to certain chemicals are highly mutagenic. Their mutagenicity arises from base pair transition that occurs as a result of the different hydrogen bonding preferences of normal nucleobase versus damaged ones. On the other hand, because proton affinities and acidities are related to hydrogen bonding acceptor and donor ability, we can use these values to try to understand the hydrogen bonding properties of O-alkylated adducts. The optimized structures of A-T (adenine-thymine), O^4-methylthymine-A and O^4-methylthymine-G base pairs as well as their respective hydrogen bond lengths are shown in Fig. 7.

Furthermore, the stability of base pairs is a result of many contributions [74] including hydrogen bonds (H-bonds)—the main binding factor in base pairs. The equilibrium distance between the proton donor and the proton acceptor atom in base pairs is generally controlled by ΔE, with a stronger hydrogen bond associated with a shorter length. On the other hand, Guerra et al. [75] investigated the fact that the nature of the H-bond in DNA and RNA base pairs is controlled by charge transfer and resonance assistance. BP86/TZ2P analysis of adenine-thymine (A-T) and guanine-cytosine (G-C) disproved that H-bonding in DNA base pairs is a predominantly electrostatic phenomenon. Instead, it has a substantial charge-transfer character caused by donor acceptor orbital interactions between O and N lone pairs and $\sigma^*_{\text{N-H}}$ acceptor orbitals. On the other hand, according to the relationship between hydrogen bond energy and the potential energy density at the BCP, the hydrogen bond energies $\mathbf{E_{H\cdots X}}$ can be calculated using the following equation [81–83]:

$$\mathbf{E_{H\cdots X} = 1/2\ V(r)\ V(r) = 1/4\nabla^2\rho(r) - 2G(r)}$$

Optimized geometry parameters, electron density ρ(*r*) in (a.u), Laplacian of electron density$\nabla^2\rho(r)$ in (a.u), H-bond energy $E_{H\cdots X}$ (in kcal mol^{-1}), selected NBO charge, and the H-bond stabilization energy $E^{(2)}$ (in kcal mol^{-1}) calculated at B3LYP/6-311++G (d,p) level for every H-bond in A-T, O^4-methylthymine-A and O^4-methylthymine-G base pairs are given in Table 3. As seen in Fig. 7, the A-T base pair has two H-bonds: from N6-H in adenine to O4 atom in O^4-methylthymine (1.924 Å), and from N3-H in thymine to N1 atom of adenine (1.841 Å). The N6-H···O4 H-bond has a $E^{(2)}$ energy value of 11.7 kcalmol^{-1}, an electron density ρ(*r*) of 0.021 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.120. The N3-H···N1 H-bond in this base pair has a $E^{(2)}$ energy value of 24.5 kcalmol^{-1}, an electron density ρ(*r*) of 0.027 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.170. These

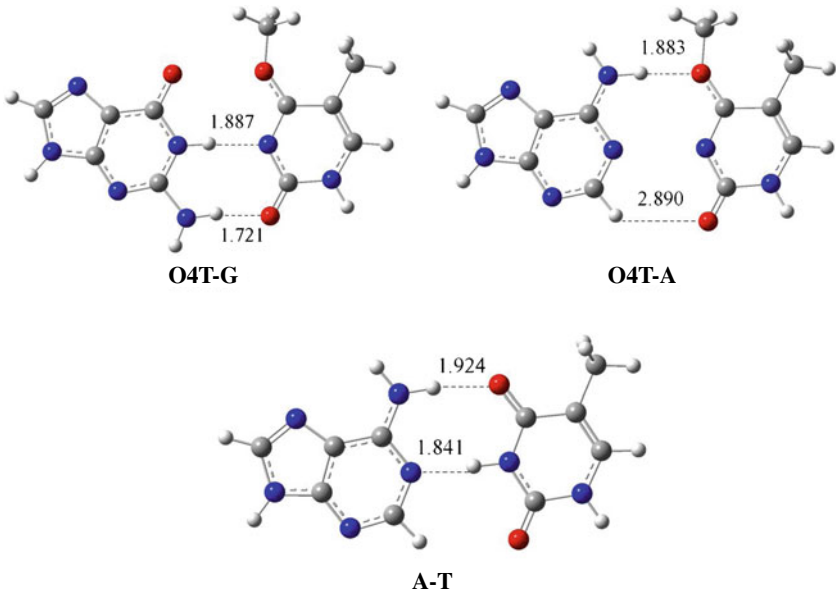


Fig. 7 Optimized structures of A-T (adenine-thymine), O^4-methylthymine-A and O^4-methylthymine-G base pairs obtained at B3LYP/6-311++G (d,p) level of theory. Bond lengths are in Ångstroms

short and strong H-bonds contribute to the stability of the A-T base pair.

The O^4-methylthymine-A base pair also has two H-bonds: from N6-H in adenine to O4 atom in thymine (1.883 Å), and from C2-H in adenine to O2 atom in thymine (2.890 Å). From Fig. 7, it can be seen that the H-bond pattern in O^4-methylthymine-A base pair differs from that of the normal A-T base pair. The most distinct changes are the O4···H-N6 H-bond in the O^4-methylthymine-A base pair, which decreased by 0.041, and the N1···H-N3 H-bond in the A-T base pair which disappeared in the O^4-methylthymine-A base pair during alkylation of the O4 atom. As seen from Table 3, the O4···H-N1 H-bond in the O^4- methylthymine-A base pair has an $E^{(2)}$ energy value of 8.5 kcalmol^{-1}, an electron density $\rho(r)$ of 0.0017 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.047. The O2···H-C2 H-bond in this base pair has an $E^{(2)}$ energy value of 1.4 kcalmol^{-1}, an electron density $\rho(r)$ of 0.004 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.015.

The O^4-methylthymine-G base pair is marked by two H-bonds as shown in Fig. 7: one between the N1-H atom of guanine and the N3 atom of O^4-methylthymine with a distance of 1.887 Å, and the other between the N2-H atom of guanine to the O2 atom of O^4-methylthymine with a distance of 1.721 Å. As shown in Table 3, the N3···H-N1 H-bond in this base pair has an $E^{(2)}$ energy value of 19.6 kcalmol^{-1}, an electron density $\rho(r)$ of 0.016 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.079. The O2···H-N2 H-bond has an $E^{(2)}$ energy value of 18.9 kcalmol^{-1}, an electron density $\rho(r)$ of 0.018 and Laplacian of electron density $\nabla^2\rho(r)$ of 0.101 (see Table 3).

As expected for close-shell interactions, all Laplacians of electron densities are positive, which indicate a depletion of electron density from the inter-atomic surface towards the interacting nuclei. Values of $\rho(r)$ and $\nabla^2\rho(r)$ for H-bond interactions in the A-T base pair are notably higher than those of base pairs of O-alkylated adducts (see Table 3). As a rule, increasing values of $\rho(r)$ and $\nabla^2\rho(r)$ at H-BCPs

Table 3 Optimized geometry parameters, electron density $\rho(r)$ in (a.u), Laplacian of electron density$\nabla^2\rho(r)$ in (a.u), H-bond energy $E_{H...X}$ (in kcalmol^{-1}), selected natural bond orbital (NBO) charge, and the H-bond stabilization energy $E^{(2)}$ (in kcalmol^{-1}) calculated at B3LYP/6-311++G (d,p) level for A-T, O^4-methylthymine-G and O^4-methylthymine-A

Base pair	Hydrogen bond	$\rho(r)$	$\nabla^2\rho(r)$	$E_{H...X}$	d (Å)	NBO charge	$E^{(2)}$
A-T	O4…H-N6	0.021	0.120	7.5	1.924	−0.653, 0.435, −0.763	11.7
	N1…H-N3	0.027	0.170	11.9	1.841	−0.618, 0.460, −0.663	24.5
O^4-methylthymine -G	O2…H-N2	0.018	0.101	6.1	1.721	−0.654, 0.401, −0.786	18.9
	N1…H-N3	0.016	0.079	4.6	1.887	−0.655, 0.453, −0.616	19.4
O^4-methylthymine -A	O4…H-N1	0.017	0.047	9.0	1.883	−0.551, 0.412, −0.784	8.5
	O2…H-C2	0.004	0.015	0.8	2.890	−0.624, 0.191, −0.293	1.4

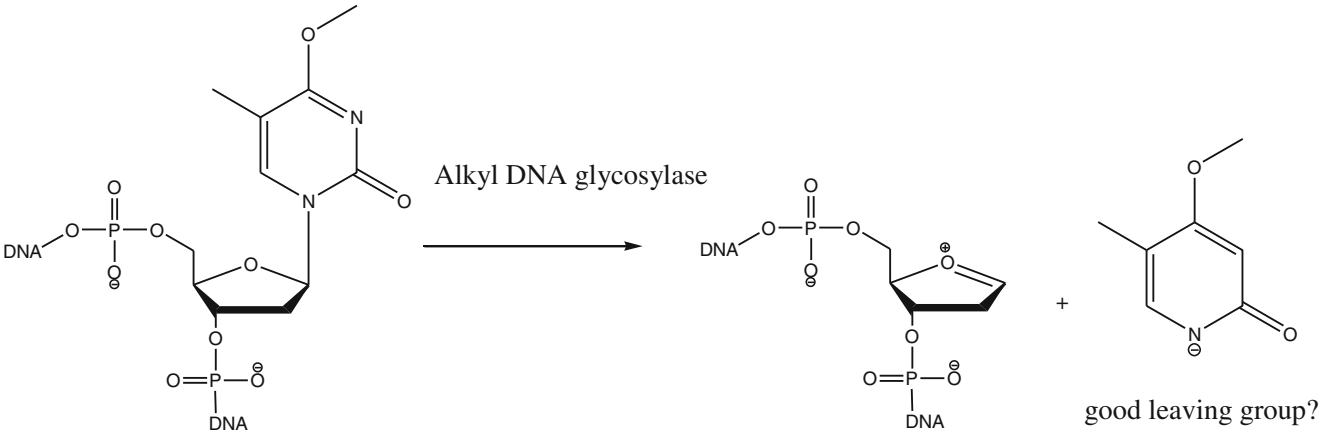


Fig. 8 Schematic illustration of the anionic cleavage mechanism used for DNA repair by the alkyl DNA glycosylase enzyme

results in increasing H-bond strength. Comparison of hydrogen bonds energies ($E_{H\cdots X}$) with $\rho(r)$ values of H-bonds in the A-T base pair with those of O-alkylated base pairs confirms this prediction.

On the other hand, as seen from Fig. 6, alkylation of thymine nucleobase at the O4 atom increases the basicity of the O2 atom considerably from 199.2 kcalmol^{-1} to 225.5 kcalmol^{-1}. Therefore, any H-bond to this site would be fairly strong, and the ability of this site to accept protons from hydrogen donor sites in base pairs increases. For example, as seen from Table 3, the O4···H-N6 H-bond in the A-T base pair has an H-bond ($E_{H\cdots X}$) value of 7.5 kcalmol^{-1}, while in the O^4-methylthymine-A base pair, $E_{H\cdots X}$ increases to 8.9 kcalmol^{-1}. On the other hand, alkylation of the thymine nucleobase at the O4 site increases the $\Delta H_{acidity}$ of the N1-H site considerably from 336.2 kcalmol^{-1} to 343.6 kcalmol^{-1} (i.e., its deprotonation reaction becomes more endothermic). Therefore, any hydrogen bond to this site would be fairly weak and thus the ability of this proton to act as a hydrogen bond donor decreases. Furthermore, the N3 atom in the thymine nucleobase changes from an acidic site to a basic site in O^4-methylthymine as result of alkylation, leading to the hydrogen bond pattern shown in Fig. 7. The N1···H-N3 hydrogen bond in the A-T base pair has an $E_{H\cdots X}$ of 11.9 kcalmol^{-1}. However, $E_{H\cdots X}$ for the same hydrogen bond in the O^4-methylthymine-A base pair is 3.2 kcalmol^{-1} (Table 3). In general, hydrogen bond energies in the A-T base pair are stronger than those in the O^4-methylthymine-A base pair (see $E_{H\cdots X}$ values in Table 3). On the other hand, hydrogen bonds energies in the O^4-methylthymine-G base pair are stronger than those of the O^4-methylthymine-A base pair. These results confirm the A-T to G-C transition that occurs as a result of different hydrogen bonding preference for normal thymine versus damaged O^4-methylthymine.

Biological applications

An interesting aspect of cytosine and thymine nucelobases is how their O-alkylated damaged nucleobases are repaired in DNA. DNA is inevitably damaged and, as mentioned in the Introduction, nature has devised various ways to repair damaged nucleobases. The main (preferred) category of repair is the base excision repair (BER) pathway, which uses the glycosylase family of enzymes. The initial step in BER is the removal of a nucleobase rather than a nucleotide. The enzymes involved in BER pathways are DNA glycosylases, which recognize damaged bases and excise them by hydrolyzing the N-glycosidic bond between the nucleobase and the sugar moiety [84, 85], leading to apurinic or apyrimidinic sites.

As shown in Fig. 8, alkyl DNA glycosylase uses an anionic cleavage mechanism for DNA repair. This mechanism is also proposed for a related enzyme, thymine DNA glycosylase (TDG), which cleaves mutated pyrimidine nucleobases from DNA. This mechanism, and the relative amounts of purine and pyrimidine nucleosides and glycosidic bond hydrolysis are highly dependent on the relative acidities of various functional groups on the leaving group nucleobases. This dependence reflects the profound activating effect of nucleobase protonation, which makes the leaving group nucleobase more electron deficient, thereby accommodating the increased electron density that develops during glycosidic bond cleavage. In terms of this mechanism, one can imagine cleavage with a

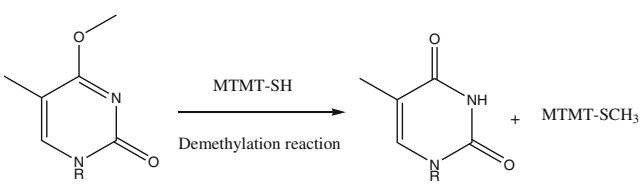


Fig. 9 Schematic illustration for O^4-methylthymine repairing by a demethylation mechanism

"deprotonated damaged nucleobase" as the leaving group (Fig. 8), or protonation of the damaged nucleobase first to facilitate cleavage.

In terms of this mechanism, as we demonstrated recently [86], the ease of excision should be related to the leaving group property of the deprotonated damaged nucleobase versus that of the normal nucleobase (Fig. 8). The alkyl DNA glycosylase enzyme discriminates the damaged deprotonated nucleobase from the normal nucleobase because the former is a better leaving group. Does the enzyme find the O^4-methylthymine nucleobase a better leaving group than thymine nucleobase? The $\Delta H_{acidity}$ of N1-H of O^4-methylthymine (**1a** tautomer) is 343.6 kcalmol^{-1}, which is 7.4 kcalmol^{-1} less acidic than thymine nucleobase (**T1** tautomer with $\Delta H_{acidity}$=336.2 kcalmol^{-1}; Fig. 6). These results suggest that the thymine DNA glycosylase enzyme not cleave the N-glycosylic bond of O^4-methylthymine. Because O^4-methylthymine has a much higher value of $\Delta H_{acidity}$ than thymine, it is not prone to anionic cleavage. Consequently, Nature has presumably devised an alternative method to repair of the O^4-methylthymine adduct. For example, demethylation via methylthymine methyltransferase (MTMT-SH; Fig. 9) could occur.

Conclusions

In the present paper, we performed a detailed DFT study in conjunction with 6-311++G (d,p) atomic basis set to probe the intrinsic properties, such as acidity, proton affinity, equilibrium tautomerization, and base pairing properties, of some O-alkylated DNA adducts (i.e., O^4-methylthymine, O^2-methylcytosine and O^2-methylthymine). In particular, we investigated how these O-alkylated damaged nucleobases differ from normal cytosine and thymine nucleobases. Base pair hydrogen bond energies in adenine and thymine normal nucleobase and their O-alkylated adducts were investigated by means of NBO and AIM. The results showed that O-alkylated adducts have a different hydrogen bonding pattern from normal nucleobases in the DNA strand. On the other hand, O-alkylated adducts have fewer tendencies to form hydrogen bonds in DNA strands because the hydrogen bond donor/acceptor ability of these molecules depends on their proton affinity and acidity properties. Furthermore, results of calculation have shown that the properties of normal versus O-alkylated damaged nucleobases lends insight into the mechanism whereby damaged nucleobases are cleaved. Comparison of the acidic properties of these DNA O-alkylated adducts to those of normal cytosine and thymine nucleobases supports the theory that the alkyl DNA glycosylase enzyme does not cleave O-alkylated damaged nucleobases as anions, and that these molecules would be a worse leaving group than normal nucleobases and be repaired in the genome by demethylation rather than by enzyme-catalyzed excision.

Acknowledgment Support from Sharif University of Technology is gratefully acknowledged.

References

1. Watson JD, Crick FHC (1953) A structure for deoxyribose nucleic acid. Nature 171:737–738
2. Brenner S, Jacob F, Meselson M (1961) An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature 190:576–581
3. Saenger W (1984) Principles of nucleic acid structure. Springer, New York, pp 113–150
4. Demple B, Harrison L (1994) Repair of oxidative damage to DNA: enzymology and biology. Annu Rev Biochem 63:915–948
5. Loft S, Poulsen HE (1996) Cancer risk and oxidative DNA damage in man. J Mol Med 74:297–312
6. Nakabeppu Y, Sakumi K, Sakamoto K, Tsuchimoto D, Tsuzuki T, Nakatsu Y (2006) Mutagenesis and carcinogenesis caused by the oxidation of nucleic acids. Biol Chem 387:373–379
7. Limoli CL, Kaplan MI, Phillips JW, Adair GM, Morgan WF (1997) Differential induction of chromosomal instability by DNA strand-breaking agents. Cancer Res 57:4048–4056
8. Palmer CM, Serafini DM, Schellhorn HE (1997) Near ultraviolet radiation (UVA and UVB) causes a formamidopyrimidine glycosylase-dependent increase in G to T Transversions. Photochem Photobiol 65:543–549
9. Singer B, Grunberger D (1983) Molecular biology of mutagens and carcinogens. Plenum, New York, pp 55–79
10. Beranek DT (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat Res 231:11–30
11. Lo TL (1975) Hard soft acids bases (HSAB) principle in organic chemistry. Chem Rev 75:1–20
12. Moschel RC, Hudgins WR, Dipple A (1979) Selectivity in nucleoside alkylation and aralkylation in relation to chemical carcinogenesis. J Org Chem 44:3324–3328
13. Loechler EL (1994) A violation of the swain-scott principle, and not SN_1 versus SN_2Reaction mechanisms. Explains why carcinogenic alkylating agents can form different proportions of adducts at oxygen versus nitrogen in DNA. Chem Res Toxicol 7:277–280
14. Lu X, Heilman JM, Blans P, Fishbein JC (2005) The Structure of DNA dictates purine atom site selectivity in alkylation by primary diazonium ions. Chem Res Toxicol 18:1462–1470
15. Lawley P, Bartsch D, Tomatis L (1976) Screening tests in chemical carcinogenesis. IARC Scientific Publications, Lyon, France, pp 181–208
16. Singer B, Kroger M, Carrano M (1978) O^2- and O^4-Alkyl pyrimidine nucleosides: stability of the glycosyl bond and of the alkyl group as a function of pH. Biochemistry 17:1246–1250
17. Loveless A (1969) Possible relevance of O^6-Alkylation of deoxyguanine to the mutagenicity and carcinogenicity of nitrosamines and nitrosamides. Nature 223:205–207
18. Coulondre C, Miller JH (1977) Mutagenic specificity in the Lacl gene of Escherichia coli. J Mol Biol 117:577–606
19. Richardson KK, Richardson FC, Swenberg JA, Skopek TR (1987) DNA base changes and alkylation following in vivo exposure of Escherichia coli to N-methyl-N-nitrosourea or N-ethyl-N-nitrosourea. Proc Natl Acad Sci USA 84:344–348

20. Abbott PJ, Saffhill R (1979) DNA synthesis with methylated Poly(dC-dG) templates. Evidence for a competitive nature to miscoding by O′-methylguanine. Biochim Biophys Acta 562:51–61
21. Beranek DT, Heflich RH, Kodell RL, Morris SM, Casciano DA (1983) Correlation between specific DNA-methylation products and mutation induction at the HGPRT locus in Chinese hamster ovary cells. Mutat Res 110:171–180
22. Loechler EL, Green CL, Essigmann JM (1984) In vivo mutagenesis by O^6-methylguanine built into a unique site in a viral genome. Proc Natl Acad Sci USA 81:6271–6275
23. Singer B, Sagi J, Kusmierek JT (1983) The vinyl chloride-derived nucleoside, N^2,3-ethenoguanosine, is a highly efficient mutagen in transcription. Proc Natl Acad Sci USA 80:4884–4888
24. Chandra AK, Nguyen MT, Uchimaru T, Zeegers-Huyskens T (1999) Protonation and deprotonation enthalpies of guanine and adenine and implications for the structure and energy of their complexes with water: comparison with uracil, thymine, and cytosine. J Phys Chem A103:8853–8860
25. Richardson KK, Richardson FC, Swenberg JA, Skopek TR (1986) DNA base changes in alkylation due to methylnitrosourea (MNU) or ethylnitrosourea (ENU) in *E. coli*. Proc Am Assoc Cancer Res 27:95–95
26. Singer B, Spengler S, Bodell WJ (1981) Tissue-dependent enzyme-mediated repair or removal of O-ethyl pyrimidines and ethyl purines in carcinogen-treated rats. Carcinogenesis 2:1069–1073
27. Spartan '06V102' Wavefunction, Irvine, CA
28. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652
29. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti conelation energy formula into a functional of the electron density. Phys Rev B 37:785–789
30. Tomasi J, Persico M (1994) Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent. Chem Rev 94:2027–2094
31. Carpenter JE, Weinhold F (1988) Analysis of the geometry of the hydroxymethyl radical by "different hybrids for different spins" natural bond orbital procedure. J Mol Struct 169:41–62
32. Foster JP, Weinhold F (1980) Natural hybrid orbitals. J Am Chem Soc 102:7211–7218
33. Alkorta I, Elguero J, Foces-Foces C (1996) Dihydrogen bonds (A–HH–B). Chem Commun 14:1633–1634
34. Alkorta I, Rozas I, Elguero J (1998) Bond length-electron density relationships: from covalent bonds to hydrogen bond interactions. Struct Chem 9:243–247
35. Bader RFW (1991) A quantum theory of molecular structure and its applications. Chem Rev 91:893–928
36. Bader RFW (1998) A bond path: a universal indicator of bonded interactions. J Phys Chem A 102:7314–7323
37. Bader RFW (2002) AIM2000 program package, Ver. 2.0. McMaster University, Hamilton
38. Kosenkov D, Kholod Y, Gorb L, Shishkin O, Hovorun DM, Mons M, Leszczynski J (2009) Ab initio kinetic simulation of gas-phase experiments: tautomerization of cytosine and guanine. J Phys Chem B 113:6140–6150
39. Mons M, Piuzzi F, Dimicoli I, Gorb L, Leszczynski J (2006) Picosecond X-ray absorption spectroscopy of a photoinduced Iron(II) spin crossover reaction in solution. J Phys Chem A 110:38–44
40. Gorb L, Kaczmare A, Gorb A, Sadle AJ, Leszczynski J (2005) Thermodynamics and kinetics of intramolecular proton transfer in guanine. Post Hartree–Fock Study. J Phys Chem B 109:13770–13776
41. Katritzky AR, Karelson M (1991) AM1 calculations of reaction field effects on the tautomeric equilibria of nucleic acid pyrimidine and purine bases and their 1-Methyl analogs. J Am Chem Soc 113:1561–1566
42. Ha TK, Gunthard HH (1993) Quantum chemical investigation of the structure and stability of all geometric isomers and conformers of all tautomeric forms of thymine. J Am Chem Soc 115:11939–11950
43. Ha TK, Keller HJ, Gunde R, Gunthard HH (1999) Energy increment method based on quantum chemical results: a general recipe for approximative prediction of isomerization and tautomerization energies of pyrimidine and purine nucleic acid bases and related compounds. J Phys Chem A 103:6612–6623
44. Nir E, Janzen C, Imhof P, Kleinermanns K, De Vries MS (2001) Guanine tautomerism revealed by UV–UV and IR–UV hole burning spectroscopy. J Chem Phys 115:4604–4611
45. Nir E, Muller M, Grace LI, De Vries MS (2002) REMPI spectroscopy of cytosine. Chem Phys Lett 355:59–64
46. Rejnek J, Hanus M, Labelac M, Ryjacek F, Hobza P (2005) Correlated ab initio Study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Part 4. Uracil and thymine. Phys Chem Chem Phys 7:2006–2017
47. Plutzer C, Kleinermanns K (2002) Tautomers and electronic states of jet-cooled adenine investigated by double resonance spectroscopy. Phys Chem Chem Phys 4:4877–4882
48. Shukla MK, Leszczynski J (2006) Spectral origins and ionization potentials of guanine tautomers: theoretical elucidation of experimental findings. Chem Phys Lett 429:261–265
49. Kosenkov D, Gorb L, Shishkin O, Sponer V, Leszczynski J (2008) Tautomeric equilibrium, stability, and hydrogen bonding in 2′-Deoxyguanosine monophosphate complexed with Mg^{2+}. J Phys Chem B 112:150–157
50. Mons M, Dimicoli I, Piuzzi F, Tardivel B, Elhanine M (2002) Tautomerism of the DNA base guanine and its methylated derivatives as studied by gas-phase infrared and ultraviolet spectroscopy. J Phys Chem A 106:5088–5094
51. Trygubenko A, Bogdan TV, Rueda M, Orozco M, Luque FJ, Sponer J, Slavıcek P, Hobza P (2002) Hydrated hydronium: a cluster model of the solvated electron? Phys Chem Chem Phys 4:4192
52. Kryachko EK, Nguyen MT, Zeegers-Huyskens T (2001) Theoretical study of tautomeric forms of uracil. 1. Relative order of stabilities and their relation to proton affinities and deprotonation enthalpies. J Phys Chem A 105:1288–1295
53. Kryachko ES, Nguyen MT, Zeegers-Huyskens T (2001) Theoretical study of uracil tautomers. 2. Interaction with water. J Phys Chem A 105:1934–1943
54. Fogarasi G (2002) Relative stabilities of three low-energy tautomers of cytosine: a coupled cluster electron correlation study. J Phys Chem A 106:1381–1390
55. Piacenza M, Grimme S (2004) Systematic quantum chemical study of DNA-base tautomers. J Comput Chem 25:83–99
56. Yang Z, Rodgers MT (2004) Theoretical studies of the unimolecular and bimolecular tautomerization of cytosine. Phys Chem Chem Phys 6:2749–2757
57. Tsolakidis A, Kaxiras E (2005) A TDDFT study of the optical response of DNA bases, base pairs, and their tautomers in the gas phase. J Phys Chem A 109:2373–2380
58. Civcir PU (2000) A theoretical study of tautomerism of cytosine, thymine, uracil and their 1-methyl analogues in the gas and aqueous phases using AM1 and PM3. J Mol Struct (THEOCHEM) 532:157–169
59. Fogarasi G, Szalay PG (2002) The interaction between cytosine tautomers and water: An MP2 and coupled cluster electron correlation study. Chem Physic Lett 356:383–390
60. Liu M, Tingting L, Amegayibor FS, Cardoso DS, Fu Y, Lee JK (2008) Gas-phase thermochemical properties of pyrimidine nucleobases. J Org Chem 73:9283–9291
61. Barker DL, Marsh RE (1964) The crystal structure of cytosine. Acta Crystallogr 17:1581–1587

62. McClure RJ, Craven BM (1973) New investigation of cytosine and its monohydrate. Acta Crystallogr Sect B 29:1234–1238
63. Weber HP, Craven BM, McMullan RK (1980) The structure of deuterated cytosine monohydrate at 82 K by neutron diffraction. Acta Crystallogr Sect B 36:645–649
64. Szczesniak M, Szczepaniak K, Kwiatkowski JS, Kubulat K, Person WB (1988) Matrix isolation infrared studies of nucleic acid constituents. 5. Experimental matrix-isolation and theoretical ab initio SCF molecular orbital studies of the infrared spectra of cytosine monomers. J Am Chem Soc 110:8319–8330
65. Brown RD, Godfrey PD, McNaughton D, Pierlot AP (1989) Tautomers of cytosine by microwave spectroscopy. J Am Chem Soc 111:2308–2310
66. Morsy MA, Al-Somali AM, Suwaiyan A (1999) Fluorescence of thymine tautomers at room temperature in aqueous solutions. J Phys Chem B 103:11205–11210
67. Fan J, Shang Zh, Liang J, Liu X, Jin H (2010) Systematic theoretical investigations on the tautomers of thymine in gas phase and solution. J Mol Struct (Theochem) 939:106–111
68. Huang Y, Kenttamaa H (2003) Theoretical estimations of the 298 K gas-phase acidities of the pyrimidine-based nucleobases uracil, thymine, and cytosine. J Phys Chem A 107:4893–4897
69. Podolyan Y, Gorb L, Leszczynski J (2000) Protonation of nucleic acid bases. a comprehensive Post-Hartree–Fock study of the energetics and proton affinities. J Phys Chem A 104:7346–7352
70. Chen ECM, Chen ES (2000) Negative ion mass spectra, electron affinities, gas phase acidities, bond dissociation energies, and negative ion states of cytosine and thymine. J Phys Chem B 104:7835–7844
71. Chen ECM, Herder C, Chen ES (2006) The experimental and theoretical gas phase acidities of adenine, guanine, cytosine, uracil, thymine and halouracils. J Mol Struct 798:126–133
72. Russo N, Toscano M, Grand A, Jolibois F (1998) Protonation of thymine, cytosine, adenine, and guanine dna nucleic acid bases: theoretical investigation into the framework of density functional theory. J Comput Chem 19:989–1000
73. Colominas C, Luque FJ, Orozko M (1996) Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes. J Am Chem Soc 118:6811–6821
74. Sponer J, Leszczynski J, Hobza P (1996) Structures and energies of hydrogen-bonded DNA base pairs. A Nonempirical Study with inclusion of electron correlation. J Phys Chem 100:1965–1974
75. Guerra CF, Bickelhaupt FM, Snijders JG, Baerends EJ (1999) The Nature of the hydrogen bond in dna base pairs: the role of charge transfer and resonance assistance. Chem Eur J 5:3581–3594
76. Meot-Ner M (1979) Ion thermochemistry of low-volatility compounds in the gas phase. Intrinsic basicities and hydrogen-bonded dimers of nitrogen heterocyclics and nucleic bases. J Am Chem Soc 101:2396–2403
77. Greco F, Liguori A, Sindona G, Uccella N (1990) Gas-phase proton affinity of deoxyribonucleosides and related nucleobases by fast atom bombardment tandem mass spectrometry. J Am Chem Soc 112:9092–9096
78. NIST Chemistry WebBook, Linstrom PJ, Mallard WG (eds) (2005) NIST standard reference database number 69. National Institute of Standards and Technology, Gaithersburg, http://webbook.nist.gov
79. Szczesniak M, Leszczynski J, Person WB (1992) Identification of the Imino-oxo form of 1-Methylcytosine. J Am Chem Soc 114:2731–2733
80. Wilson MS, McCloskey JA (1975) Chemical ionization mass spectrometry of nucleosides. Mechanisms of ion formation and estimations of proton affinity. J Am Chem Soc 97:3436–3444
81. Yurenko YP, Zhurakivsky RO, Samijlenko SP, Ghomi M, Hovorun DM (2007) The whole of Intramolecular H-bonding in the Isolated DNA Nucleoside Thymidine. AIM Electron Density Topological Study. Chem Phys Lett 447:140–146
82. Espinosa E, Molins E, Lecomte C (1998) Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. Chem Phys Lett 285:170–173
83. Yurenko YP, Zhurakivsky RO, Samijlenko SP, Hovorun DM (2011) Intramolecular CH···O Hydrogen bonds in the AI and BI DNA-like conformers of canonical nucleosides and their Watson-Crick Pairs. Quantum chemical and AIM analysis. Biomol Struct Dyn 1:51–56
84. Parikh SS, Walcher G, Jones GD, Slupphaug G, Krokan HE, Blackburn GM, Tainer JA (2000) Uracil-DNA glycosylase–DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. Proc Natl Acad Sci USA 97:5083–5088
85. Berti PJ, McCann JAB (2006) Toward a detailed understanding of base excision repair enzymes: transition state and mechanistic analyses of N-Glycoside hydrolysis and N-Glycoside transfer. Chem Rev 106:506–555
86. Tehrani AZ, Torabifard H, Fattahi A (2012) Thermochemical properties of some vinyl chloride-induced DNA lesions: detailed view from NBO & AIM analysis. Struct Chem 23:1987–2001. doi:10.1007/s11224-012-0026-y