THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Quantum Mechanical Calculations for the Misincorporation of Nucleotides Opposite Mutagenic 3,N^4-Ethenocytosine

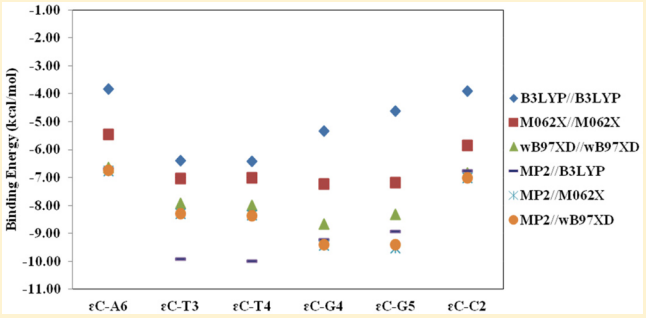
Venkatesan Srinivasadesikan,[†] Prabhat K. Sahu,*[,‡] and Shyi-Long Lee*[,†]

[†]Department of Chemistry and Biochemistry, National Chung Cheng University, Taiwan

[‡]Department of Chemistry, Bhubaneswar Institute of Technology, India

Supporting Information

ABSTRACT: The ubiquitous nature and persistence of exocyclic DNA adducts suggest their involvement as initiators of carcinogenesis. We have investigated the misincorporation properties of the exocyclic DNA adduct, 3,N^4-ethenocytosine, using DFT and DFT-D methods. Computational investigations have been carried out by using the B3LYP, M062X, and wB97XD methods with the 6-31+G* basis set to determine the hydrogen bonding strengths, binding energy, and physical parameters. The single point energy calculations have been carried out at MP2/6-311++G** on corresponding optimized geometries. The energies were compared among the 3,N^4-ethenocytosine adduct with DNA bases to find the most stable conformer. The solvent phase calculations have also been carried out using the CPCM model. The computed reaction enthalpy values provide computational insights to the earlier experimental observation in *in vitro*, *E.coli*, and mammalian cells of a high level of substitution mutation in which C → A transversion results from εC−T pairing [εC−T3 and εC−T4] in the adduct containing DNA sequence.



1. INTRODUCTION

Numerous occupational chemical carcinogens induce the formation of exocyclic DNA adducts.[1] Accumulating *in vitro* evidence as well as the *in vivo* study indicate that cyclic lesions contribute to the mutagenic and carcinogenic properties of a number of bifunctional chemical mutagens such as vinyl chloride and urethane.[2] The cellular DNA reacts with alkylating agents to produce 3,N^4-ethenocytosine (εC) and other exocyclic lesions. Moreover, the lipid peroxidation, reactive oxygen, and reactive nitrogen species also damage DNA and produce a variety of potentially genotoxic exocyclic lesions; such damage could lead to carcinogenesis.[3] Etheno adducts result from the formation of a new imidazole ring on nucleic acid bases. There are possible ethenobases: ethenoadenine (εA), ethenoguanine (εG), ethanoadenine, εC·H_2O, and εC, which have been reviewed by Barbin et al.[4] These exocyclic adducts were detected as endogenous lesions in DNA isolated from humans and untreated rat livers.[5] The mispairing properties of the lesions εC vary significantly, depending on the DNA polymerase (Pol α; incorporation of T, A in larger amount with C, G in smaller amount; Pol β; C, A; Pol δ; T in larger, A, G in smaller) used to catalyze the reaction with different quantity of nucleotide incorporation.[6] Speina et al.[7] have shown impaired repair of εA and εC in patients developing the AD (adenocarcinoma) type of lung tumor, the etiology of which is linked to inflammatory processes. The ethenocytosine has novel mutagenic properties that make it a valuable model for dissecting the mechanisms of induced mutagenesis.[8] The toxicity and mutagenic properties of adducts have been investigated using *in vitro* and *in vivo* systems, and the endogenous DNA adducts were highly genotoxic to *Escherichia coli* and human cells. The lesion has been found highly mutagenic in mammalian cells with the frequency of εC → A transversions almost twice that of the εC → T transitions.[9] The mutagenic potential of εC in *Escherichia coli* and simian kidney cells has been reported as εC·G to A·T transversions and T·A transitions.[9,10] Investigation has also been made on the mutagenic potential, and its structure−function relationships in DNA fragments have covalently modified bases.[11]

In our present study, we report the base pairing and mispairing specificity of 3,N^4-ethenocytosine at the molecular level. Quantum mechanical calculations for the misincorporation of nucleotides opposite mutagenic 3,N^4-ethenocytosine, based on different possible conformations, have been carried out. The present findings allow us to determine the potential role of base pairing and mispairing states, the most favorable conformation that exhibits unique mutagenic properties. Such a study will be helpful to lead the mechanism and base pairing specificity.

Received: July 21, 2012
Revised: August 13, 2012
Published: August 14, 2012

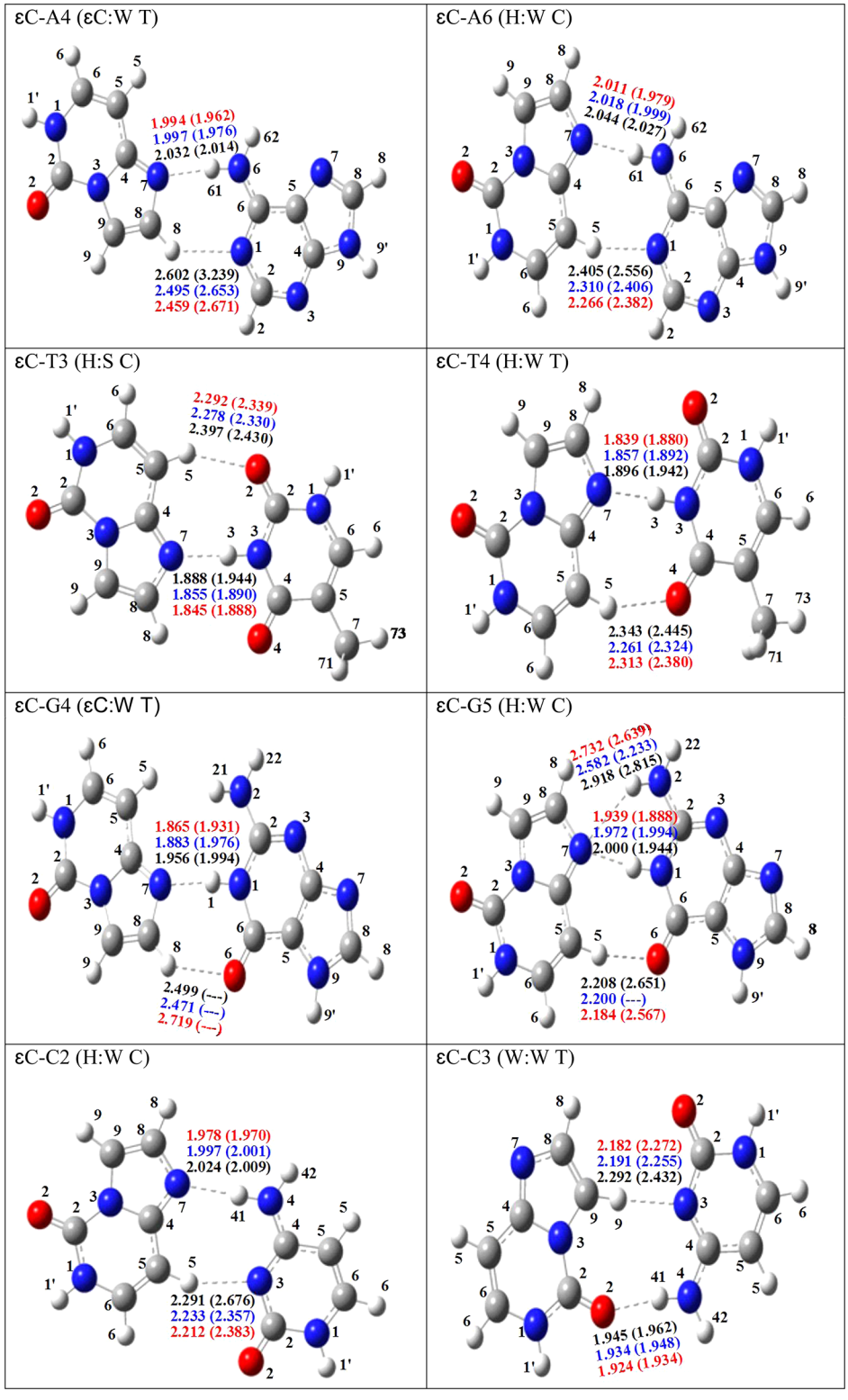


Figure 1. Most stable structures for 3,N^4-ethenocytosine adduct and DNA base complexes in the gas phase (aqueous phase) optimized using different DFT and DFT-D levels with the 6-31+G* basis set. Computed values at different levels: B3LYP, black; M062X, blue; wB97XD, red.

2. MODEL SYSTEM AND COMPUTATIONAL DETAILS

Toward determining the preference for the incorporation of nucleobases opposite to 3,N^4-ethenocytosine have been evaluated using the hydrogen bonding strengths and energetic parameters. Moriya et al. reported that the frequency of εC → A transversions is twice the amount of εC → T transitions in mammalian cells.[9] Drennan et al.[26] have shown the crystal structure of Δ79AAG bound to a DNA duplex containing εC:G. The DNA glycosylase activity has been investigated for the four nucleobases opposite to the εC, and its rates of excision have also been reported.[27] Thus, possibly four nucleobases can be incorporated opposite to the εC lesion. Despite the experimentally studied base pairing specificity[6,9,28]

dx.doi.org/10.1021/jp307239x | J. Phys. Chem. B 2012, 116, 11173−11179

Table 1. Hydrogen Bonding Strengths (kcal/mol) for 3,N^4-Ethenocytosine Adduct and DNA Base Complexes in the Gas Phase (ΔE_g) and Aqueous Phase (ΔE_{aq}) Calculated Using DFT Levels/6-311++G**//DFT Levels/6-31+G*[a]

	B3LYP//B3LYP		M062X//M062X		wB97XD//wB97XD	
system	ΔE_g	ΔE_{aq}	ΔE_g	ΔE_{aq}	ΔE_g	ΔE_{aq}
εC−A1	−6.44 (−6.03)	−3.22	−7.86 (−7.36)	−4.81	−8.70 (−8.21)	−5.53
εC−A2	−2.25 (−1.94)	−0.73	−3.76 (−3.28)	−2.29	−4.51 (−4.07)	−2.96
εC−A3	−4.02 (−3.71)	−1.15	−5.50 (−5.09)	−2.65	−6.12 (−5.72)	−3.19
εC−A4	−6.29 (−5.91)	−3.45	−7.83 (−7.37)	−4.98	−8.82 (−8.38)	−5.75
εC−A5	−1.62 (−1.38)	−0.22	−3.19 (−2.86)	−1.73	−3.67 (−3.36)	−2.10
εC−A6	−7.77 (−7.34)	−3.84	−9.27 (−8.75)	−5.45	−10.80 (−10.29)	−6.64
εC−A7	−2.36 (−2.06)	−0.49	−3.86 (−3.49)	−2.02	−4.74 (−4.39)	−2.70
εC−A8	−4.54 (−4.23)	−1.03	−6.30 (−5.89)	−2.63	−7.22 (−6.82)	−3.43
εC−T1	−8.03 (−7.53)	−5.22	−8.44 (−7.87)	−5.56	−8.71 (−8.12)	−6.25
εC−T2	−8.23 (−7.71)	−5.26	−8.53 (−7.94)	−5.59	−8.90 (−8.29)	−5.91
εC−T3	−10.80 (−10.20)	−6.40	−11.49 (−10.80)	−7.03	−12.60 (−11.93)	−7.92
εC−T4	−10.94 (−10.35)	−6.41	−11.59 (−10.88)	−7.02	−12.78 (−12.10)	−7.99
εC−T5	−5.12 (−4.75)	−2.02	−5.54 (−5.08)	−2.51	−6.40 (−5.93)	−3.12
εC−G1	−8.28 (−7.74)	−2.41	−9.52 (−8.85)	−3.87	−10.82 (−10.16)	−5.10
εC−G2	−5.27 (−4.93)	−0.28	−6.30 (−5.87)	−1.41	−7.19 (−6.76)	−2.13
εC−G3	−10.35 (−9.79)	−3.16	−11.67 (−10.98)	−4.56	−13.65 (−12.98)	−6.05
εC−G4	−9.52 (−8.98)	−5.34	−11.50 (−10.85)	−7.23	−12.95 (−12.30)	−8.66
εC−G5	−11.60 (−11.04)	−4.64	−13.77 (−13.11)	−7.18	−15.39 (−14.75)	−8.33
εC−G6	−10.74 (−10.20)	−4.56	−12.47 (−11.84)	−6.67	−13.32 (−12.70)	−7.29
εC−C1	−2.79 (−2.48)	0.02	−4.75 (−4.32)	−1.92	−4.76 (−4.37)	−1.78
εC−C2	−10.77 (−10.22)	−3.90	−12.84 (−12.19)	−5.86	−14.15 (−13.52)	−6.85
εC−C3	−8.32 (−7.81)	−3.28	−10.19 (−9.58)	−5.26	−10.86 (−10.25)	−5.75

[a]BSSE-corrected binding energies (kcal/mol) in parentheses.

and mode of incorporation for the 3,N^4-ethenocytosine adduct, molecular level information is not yet clear. Therefore, we have considered all the possible four nucleobases paired with mutagenic exocyclic DNA adduct 3,N^4-ethenocytosine for our model calculation to determine the misincorporation.

The geometrical and energetic parameters of ethenocytosine with the DNA bases were calculated using different density functional theory (DFT) methods with the Pople-type split valence basis set (6-31+G*). Several successful results, operating costs, and less computational time together attract DFT methods as a versatile alternative to the *ab initio* methods.[12−16] The aforementioned motivations together attract us to use the DFT methods for the evaluation of misincorporation properties of mutagenic 3,N^4-ethenocytosine. In our model system, the sugar (five-carbon pentose) and phosphate groups have been substituted by hydrogen atom and several possible interactions between εC adduct and DNA bases were considered. These replaced hydrogen atoms were excluded from the interaction, since it may lead to non-physiological base pairing.[17] Geometry optimizations were carried out using the Gaussian 09 programs.[18] In the density functional theory (DFT) framework B3LYP[19] (i.e., Becke's three-parameter hybrid functional, LYP correlation functional), the hybrid meta density functional M062X[20] and long-range with dispersion corrected wB97XD[21] density functional at the 6-31+G* level have been considered for computation. Single point energy calculations at the MP2[12,22]/6-311++G** levels on considered DFT and DFT-D optimized geometries have also been carried out to better estimate the hydrogen bonding strengths. The hydrogen bonding strength has been calculated as the energy difference between complex and monomers. The calculated hydrogen bonding energies were corrected for basis set superposition error (BSSE)[23] using the counterpoise method.[24] To estimate the solvent effect in the protein environment, the continuum model CPCM (polarizable conductor calculation model along with united-atom topological model, ε = 78.39)[25] that accounts for the solvent polarizability has been considered for the optimization, single point energy, and enthalpy computations at various levels.

3. GEOMETRIC AND ENERGETIC PARAMETERS FOR 3,N^4-ETHENOCYTOSINE−DNA COMPLEXES

The computations have been carried out for the different possible conformations of εC−DNA complexes (εC−A, εC−T, εC−G, εC−C). All the optimized structures in the gas phase and solvent phase (in brackets) with hydrogen bonding lengths at various levels (B3LYP, black; M062X, blue; wB97XD, red) are shown in Figure S1 in the Supporting Information, and energetically low lying conformations are shown in Figure 1 with the accepted nomenclature according to Leontis et al. and others[29] with the modification which is having the interacting edges of etheno group with the nucleobases. The hydrogen bonding lengths and bond angles at different DFT and DFT-D levels in the gas phase and aqueous phase have been listed in Table S1a−d in the Supporting Information. Biomacromolecules, DNA, RNA, and proteins all play a dominant role in our life. The hydrogen bond is the strong intermolecular force and plays a crucial role in the structure of biomolecules. All the biochemical processes are predominantly solution-based, with water serving as a ubiquitous solvent. Because of these reasons, our discussion is based on solvent phase computed results. Four different types of hydrogen bonds such as N−H···N, N−H···O, C−H···N, and C−H···O play a vital role in stabilizing the εC−DNA complexes. On the basis of the DFT and DFT-D levels, it can be observed that the N−H···N and N−H···O hydrogen bonds are shorter than the C−H···O and C−H···N hydrogen bonds. For the bond angle, ∠N-H...N and ∠N-H...O types are

dx.doi.org/10.1021/jp307239x | J. Phys. Chem. B 2012, 116, 11173−11179

Table 2. Hydrogen Bonding Strengths (kcal/mol) for 3,N^4-Ethenocytosine Adduct and DNA Base Complexes in the Gas Phase (ΔE_g) and Aqueous Phase (ΔE_{aq}) Calculated Using MP2 Levels/6-311++G**//DFT Levels/6-31+G*[a]

	MP2//B3LYP		MP2//M062X		MP2//wB97XD	
system	ΔE_g	ΔE_{aq}	ΔE_g	ΔE_{aq}	ΔE_g	ΔE_{aq}
εC−A1	−8.80 (−6.88)	−5.82	−8.82 (−6.80)	−5.87	−8.79 (−6.73)	−5.85
εC−A2	−5.39 (−3.81)	−3.92	−5.70 (−3.69)	−4.33	−5.68 (−3.65)	−4.09
εC−A3	−6.50 (−4.88)	−4.33	−6.57 (−4.79)	−4.34	−6.55 (−4.78)	−4.39
εC−A4	−8.86 (−7.20)	−5.75	−8.92 (−7.16)	−5.94	−8.90 (−7.13)	−5.93
εC−A5	−4.00 (−2.77)	−2.40	−4.08 (−2.69)	−2.65	−4.09 (−2.73)	−2.64
εC−A6	−10.67 (−8.73)	−6.70	−10.72 (−8.69)	−6.76	−10.67 (−8.60)	−6.74
εC−A7	−5.40 (−3.84)	−3.51	−5.18 (−3.59)	−3.76	−5.04 (−3.45)	−3.71
εC−A8	−7.57 (−5.98)	−3.78	−7.66 (−5.93)	−3.93	−7.64 (−5.90)	−3.90
εC−T1	−10.62 (−8.32)	−8.21	−9.03 (−6.67)	−6.46	−9.05 (−6.60)	−7.43
εC−T2	−10.78 (−8.44)	−8.25	−9.20 (−6.79)	−6.55	−9.20 (−6.73)	−6.55
εC−T3	−14.20 (−11.67)	−9.93	−12.70 (−10.06)	−8.30	−12.69 (−10.02)	−8.28
εC−T4	−14.33 (−11.76)	−9.99	−12.84 (−10.15)	−8.36	−12.86 (−10.15)	−8.36
εC−T5	−8.44 (−6.32)	−4.49	−6.87 (−4.56)	−4.03	−7.20 (−4.70)	−4.02
εC−G1	−10.45 (−8.07)	−5.02	−10.68 (−8.13)	−5.28	−10.58 (−7.98)	−6.04
εC−G2	−7.33 (−5.65)	−2.60	−7.44 (−5.63)	−2.73	−7.36 (−5.55)	−2.66
εC−G3	−12.95 (−10.53)	−5.84	−13.16 (−11.36)	−6.16	−13.07 (−10.45)	−6.34
εC−G4	−13.09 (−10.78)	−9.22	−13.37 (−10.85)	−9.43	−13.52 (−10.92)	−9.41
εC−G5	−15.00 (−12.44)	−8.93	−15.35 (−12.39)	−9.54	−15.31 (−12.43)	−9.40
εC−G6	−12.74 (−10.11)	−7.39	−12.79 (−9.95)	−7.47	−12.73 (−9.91)	−7.51
εC−C1	−5.28 (−3.86)	−1.88	−5.29 (−3.59)	−2.27	−5.36 (−3.78)	−2.40
εC−C2	−13.75 (−11.43)	−6.76	−13.82 (−11.34)	−7.01	−13.82 (−11.34)	−7.01
εC−C3	−10.83 (−8.55)	−6.09	−10.89 (−8.46)	−6.16	−10.87 (−8.43)	−6.14

[a]BSSE-corrected binding energies (kcal/mol) in parentheses.

linear than the ∠C-H...N and ∠C-H...O types. Our current results are in good agreement with the previous results.[30,31]

The computed total energy for several different conformations of the 3,N^4-ethenocytosine−DNA complexes is listed in Table S2a,b in the Supporting Information. Table 1 depicts the hydrogen bonding strengths in the gas phase (with BSSE corrections) and in the aqueous phase calculated using the DFT and DFT-D levels. Table 2 depicts the hydrogen bonding strengths in the gas phase (with BSSE corrections) and in the aqueous phase calculated using the MP2/6-311++G**//DFT and DFT-D/6-31+G* levels. Parts A and B of Figure 2 show the hydrogen bonding strengths for the most stable conformers among the 3,N^4-ethenocytosine−DNA complexes at the DFT, DFT-D, and MP2//DFT,DFT-D levels in the gas phase and in the aqueous phase, respectively.

3.1. 3,N^4-Ethenocytosine Adduct and Adenine (εC−A) Complexes. From Tables 1 and 2, the εC−A1, εC−A4, and εC−A6 have been observed to be the most stable structures in the gas and aqueous phases. Such most stable conformations are in the gas phase varying within 1 kcal/mol and in the aqueous phase varying within 0.4 kcal/mol among the εC−A complexes. The reason for the most stable conformers (εC−A1, εC−A4, and εC−A6) of the εC−A complexes, one among the two hydrogen bonds is having N−H···N or N−H···O type. Those hydrogen bonding strengths have been observed not to be elongated when shifted to the aqueous phase by the solvent polarization effects, due to the strong interactions. Interestingly, the hydrogen bonding lengths and bond angles have also been observed to deviate ∼0.1 Å and ∼1°, respectively, between the gas phase and the aqueous phase.

3.2. 3,N^4-Ethenocytosine Adduct and Thymine (εC−T) Complexes. From Tables 1 and 2, it can be observed that εC−T3 and εC−T4 are the most stable complexes among the εC−T1, εC−T2, εC−T3, εC−T4, and εC−T5 complexes. The most stable complexes of comparable binding energies were shown in Figure 2. Table 1 also reveals that the N−H···N (εC−T3, εC−T4) type hydrogen bond is shorter than the N−H···O (εC−T1, εC−T2) type, which reflects on the binding energy in the gas phase. When shifted to the aqueous phase, the qualitative trend is the same but quantitatively differs within 1 kcal/mol. Additionally, the weak force of attraction arises between O4···H8 (εC−T3) and O2···H8 (εC−T4), which might be the reason for the stability. Moreover, in εC−T5, the C−H···N type hydrogen bond has not been observed when shifted to the aqueous phase from the gas phase at the B3LYP level, exhibiting T-shaped structure as compared to planar structure (see Figure S2, Supporting Information).

3.3. 3,N^4-Ethenocytosine Adduct and Guanine (εC−G) Complexes. For εC−G complexes, εC−G4 and εC−G5 have been observed to be the most stable complexes (see Tables 1 and 2). The N−H···N type hydrogen bond has been observed for εC−G4 and εC−G5 complexes. Additionally, the C−H···O type hydrogen bond has been observed for εC−G4 and εC−G5 complexes in the gas phase. When shifted to the aqueous phase, such a hydrogen bond has been diminished for the εC−G4 conformation (Figure S2, Supporting Information) due to the weak interaction and solvent polarization effect. Notably, in the εC−G5 complex in the aqueous phase, the C−H···O type hydrogen bond has been diminished at the M062X level. Only two hydrogen bonds have been observed for the εC−G5 complex at the M062X level compared to three hydrogen bonds at other DFT levels. When we compare the earlier reported interaction energies of the G−C WC base pair by Sponer et al.[32] (27.5 kcal/mol) and Hobza et al.[33] (25.4 kcal/mol) with the corresponding adduct in our current investigation, it is found to be εC−G6 (13.32 kcal/mol) with the lowest binding energy (because the etheno group on cytosine blocked the hydrogen bonds). The geometrical

dx.doi.org/10.1021/jp307239x | J. Phys. Chem. B 2012, 116, 11173−11179

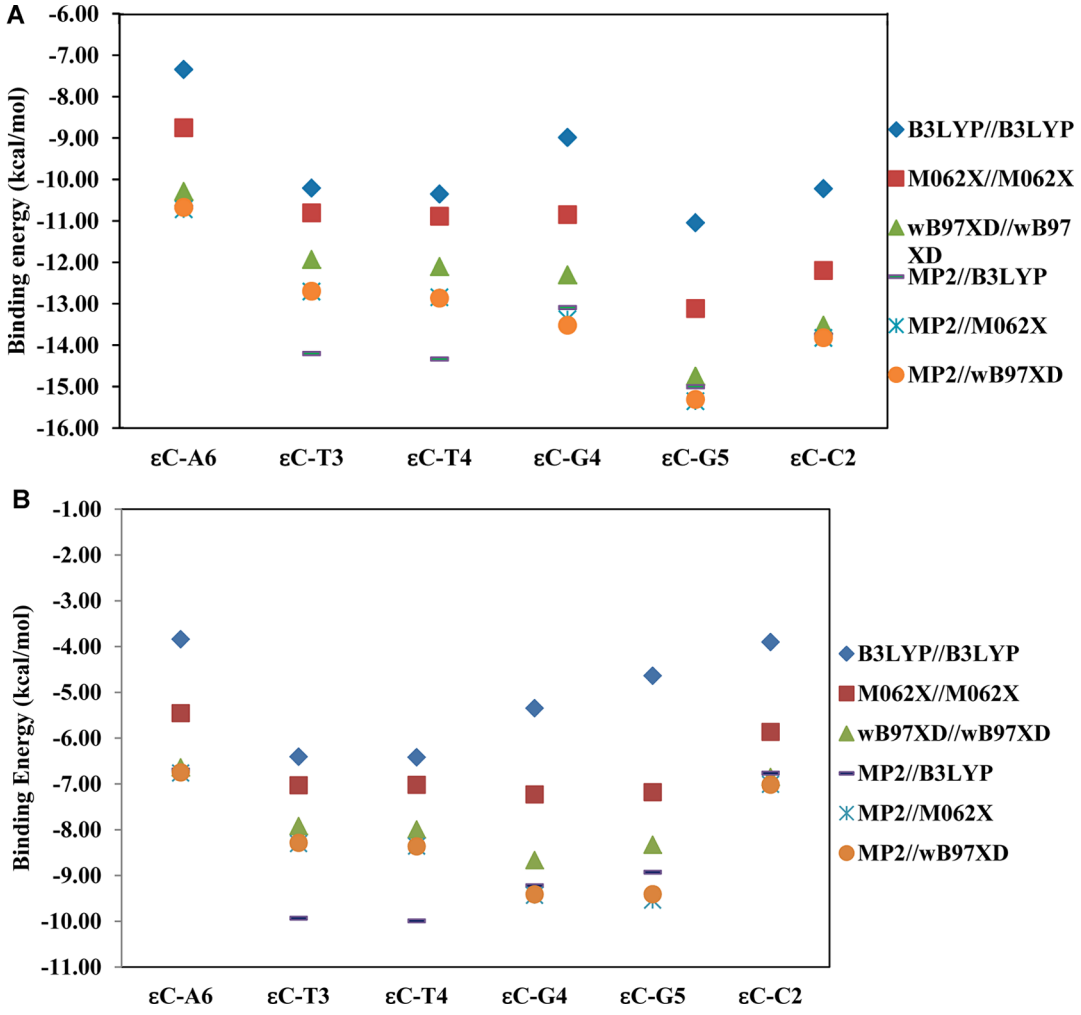


Figure 2. (A) Binding energy for the most stable complexes in 3,N^4-ethenocytosine−DNA base pair complexes in the gas phase at different DFT levels. (B) Binding energy for the most stable complexes in 3,N^4-ethenocytosine−DNA base pair complexes in the aqueous phase at different DFT levels.

comparison reveals that the donor NH_2 group in the GC base pair has been modified to the donor CH group in εC−G6 which interacts with O6 of guanine and acceptor N3 with H1 of guanine in the GC base pair has not shown any interaction in εC−G6. O2 of cytosine in the GC base pair interacts with H21 of guanine; such O2 of cytosine in εC−G6 has delivered two hydrogen bonds with H21 and H1 of guanine. These modifications due to etheno adduct formation on cytosine reflect on the binding energy. In our previous computation[30a] for the εG−DNA studies, it was revealed that the G−C base pair comparable adduct εG−C2 exhibits the lower binding energy due to the N^2,3-ethenoguanine adduct formation.

3.4. 3,N^4-Ethenocytosine Adduct and Cytosine (εC−C) Complexes. Figure 2 shows the binding energy for the most stable complexes of εC−C2 and εC−C3. It can also be observed from Tables 1 and 2. In the εC−C2 complex, C−H···N and N−H···N type hydrogen bonds have been observed. In the εC−C3 complex, C−H···N and N−H···O type hydrogen bonds have been observed (see Figure 1).

4. SUMMARY

The present study reports that B3LYP underestimates the hydrogen bonding strength as compared with the other DFT and DFT-D methods which are in good agreement with the literature.[34] The computed BSSE corrected hydrogen bonding strength at the DFT, DFT-D, and MP2 levels using a larger basis set on corresponding optimized geometry at the DFT and DFT-D levels are listed in Tables 1 and 2. From Tables 1 and 2, it can be observed that the energy difference is within 1 kcal/mol for different conformations. At the MP2//DFT,DFT-D levels (in gas phase) for the εC−DNA complexes, the hydrogen bonding strengths were found to be εC−G5 (15.3 kcal/mol) > εC−T4 (12.8 kcal/mol) ≈ εC−T3 (12.7 kcal/mol) > εC−C2 (13.8 kcal/mol) ≈ εC−C3 (10.8 kcal/mol) > εC−A6 (10.7 kcal/mol). The dispersion corrected wB97XD//wB97XD level performs very close to the MP2//DFT,DFT-D levels, the energy difference around 1 kcal/mol, with M062X the energy difference around 1.5 kcal/mol. However, the qualitative trend was found to be the same, εC−G5 ≈ εC−G4 > εC−C2 ≈ εC−C3 > εC−T4 ≈ εC−T3 > εC−A6 ≈ εC−A1 at the wB97XD, M062X, and B3LYP levels (see Table 1) in the gas phase. We can conclude from the above mentioned qualitative chronological order for the different conformations at different levels that εC−G5 is the most favorable potential candidate for the mispairing. However, other potential candidates are εC−C2, εC−T4, and εC−A6 which have close binding energy, and cannot be ignored.

dx.doi.org/10.1021/jp307239x | J. Phys. Chem. B 2012, 116, 11173−11179

On shifting to the aqueous phase, the stability trend at the MP2//DFT,DFT-D levels was found to be εC−T4 (8.3 kcal/mol) ∼ εC−T3 (8.3 kcal/mol) > εC−G5 (9.5 kcal/mol) ∼ εC−G4 (9.4 kcal/mol) > εC−C2 (∼7 kcal/mol) > εC−A6 (6.7 kcal/mol) ∼ εC−C3 (6.1 kcal/mol) ∼ εC−A4 (5.9 kcal/mol). The stability trend at the DFT//DFT,DFT-D levels was also qualitatively the same and found to be εC−T4 ∼ εC−T3 > εC−G4 ∼ εC−G5 > εC−C2 > εC−A6 > εC−A4. Among the most stable conformers in the same group, the energy difference was within 0.5 kcal/mol. Therefore, qualitatively, no significant difference with respect to the dispersion corrected DFT methods or the stability trend has been found in the most stable conformations. B3LYP quantitatively underestimates the hydrogen bonding strength, whereas the wB97XD method overestimates the hydrogen bonding strengths around 1 kcal/mol compared to M062X. Since we have noticed the performance of wB97XD is in close agreement with MP2 level in gas phase and aqueous phase quantitatively as well as qualitatively. It can be concluded that wB97XD is more reliable than M062X and other DFT methods. Such a conclusion was also drawn by Kanchana et al.[35] From the above-mentioned view and aqueous-phase discussion, we can conclude that 3,N^4-ethenocytosine adduct is the most favorable to pair with thymine [εC−T4 conformation] (*in vitro* and mammalian cells) as well as guanine [εC−G4 conformaion] at the MP2//B3LYP level, which is also in line with the experimental observation[6,9,31] of a high level of C → A transversion resulting from εC−T pairing upon replication in εC containing DNA helix. Also, one cannot ignore the C → C transition such as εC pairing with guanine in εC containing DNA helix. However, the other pairings εC−T3, εC−G5, εC−C2, and εC−A6 are close in binding energy and cannot be ignored. DNA damages with alteration of structure, that have not been repaired, cause errors during DNA synthesis, leading to mutations that can give rise to cancer. Thus, individuals with an inherited impairment in DNA repair capability are often at increased risk of cancer.[36]

Figure 3A presents the reaction enthalpy values in the gas phase for the εC−DNA complexes. The following stability trend can be observed: εC−G5 (13.5 kcal/mol) ∼ εC−G3 (11.9 kcal/mol) > εC−C2 (12.4 kcal/mol) > εC−T4 (11.11 kcal/mol) ∼ εC−T3 (10.86 kcal/mol) > εC−A6 (9.51 kcal/mol) in the gas phase at the B3LYP and DFT-D levels, an almost similar trend at both levels. On shifting to the aqueous phase (see Figure 3B), the trend is εC−T3 (6.8 kcal/mol) ∼ εC−T4 (6.4 kcal/mol) > εC−G4 (6.7 kcal/mol) ∼ εC−G5 (6.3 kcal/mol) > εC−C2 (5.8 kcal/mol) ∼ εC−C3 (4.7 kcal/mol) > εC−A6 (4.9 kcal/mol). The present findings based on enthalpy also show that thymine and guanine are the favorable candidates for the misincorporation opposite to 3,N^4-ethenocytosine which is in line with the experimental results[6,9,28] but contrary to the prediction of adenine and cytosine.

The present investigations provide underlying molecular mechanisms for the mutagenic 3,N^4-ethenocytosine adduct misincorporation with DNA bases. The computed reaction enthalpy values provide computational insights to the earlier experimental observation in *in vitro*,[6] *E. coli*,[2b,28] and mammalian cells[9] of a high level of substitution mutation in which C → A transversion results from εC−T pairing [εC−T3 and εC−T4] in the adduct containing DNA sequence. When guanine is replaced by thymine opposite to 3,N^4-ethenocytosine, the reaction enthalpy values are larger for thymine (6.84 kcal/mol) than guanine (6.78 kcal/mol) in the aqueous phase. Such aqueous phase results represent a point mutation occuring in the cell as a prime effect for carcinogenesis. It has been proposed that, along with εC−T3 and εC−T4, mispairs like εC−G4, εC−G5, εC−C2, and εC−A6 are also potent and cannot be ignored for further exploration.

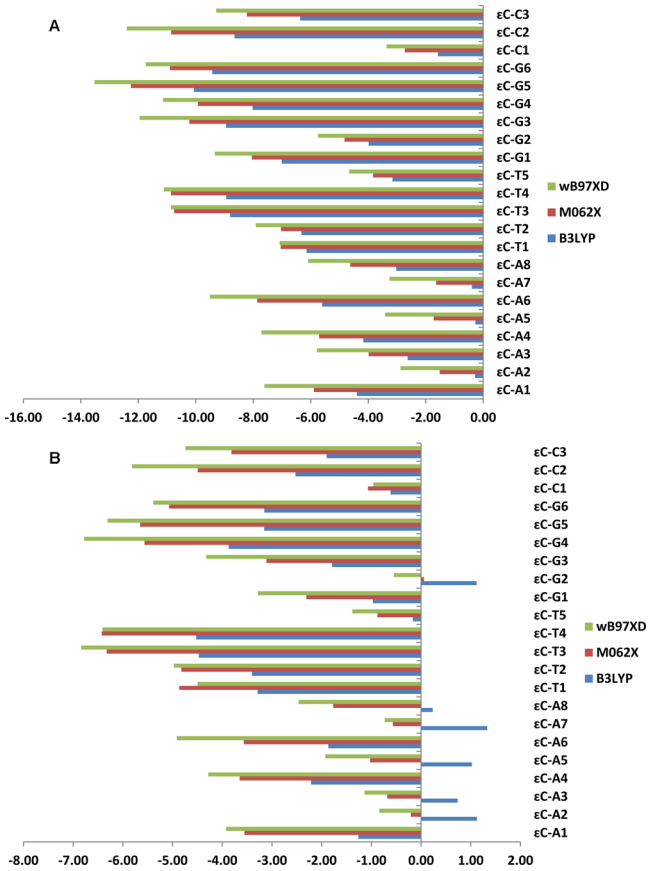


Figure 3. (A) Computed reaction enthalpy values (kcal/mol) for the 3,N^4-ethenocytosine adduct and DNA bases at different DFT levels in the gas phase. (B) Computed reaction enthalpy values (kcal/mol) for the 3,N^4-ethenocytosine adduct and DNA bases at different DFT levels in the solvent phase.

■ ASSOCIATED CONTENT

Supporting Information

Optimized hydrogen bond distances and bond angles for the several different conformations of 3,N^4-ethenocytosine adduct and DNA nucleotide complexes (Table S1a−d), computed total energy for the several different conformations of 3,N^4-ethenocytosine adduct and DNA bases (Table S2a,b), computed reaction enthalpy values for the 3,N^4-ethenocytosine adduct and DNA nucleotide complexes (Table S3a,b), optimized structures of 3,N^4-ethenocytosine adduct and DNA nucleotide complexes (Figures S1 and S2). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: prabhat.sahu@bit.edu.in (P.K.S.); chesll@ccu.edu.tw (S.-L.L.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the National Science Council (NSC) of Taiwan, and the computational resource is partially supported by National Center for High-Performance Computing (NCHC), Hsin-Chu, Taiwan.

■ REFERENCES

(1) Barbin, A.; Bartsch, H. *IARC Sci. Publ.* **1986**, *70*, 345−358.
(2) (a) Jacobsen, J. S.; Perkins, C. P.; Callahan, J. T.; Sambamurti, K.; Humayun, M. Z. *Genetics* **1989**, *121*, 213−222. (b) Palejwala, V. A.; Rzepka, R. W.; Humayun, M. Z. *Biochemistry* **1993**, *32*, 4112−4120. (c) Malaveille, C.; Bartsch, H.; Montesano, R.; Barbin, A.; Camus, A. M. *Biochem. Biophys. Res. Commun.* **1975**, *63*, 363−370.
(3) (a) Coussens, L. M.; Werb, Z. *Nature* **2002**, *420*, 860−867. (b) Wiseman, H.; Halliwell, B. *Biochem. J.* **1996**, *313*, 17−29.
(4) Barbin, A. *Mutat. Res.* **2000**, *462*, 55−69.
(5) (a) Nair, J.; Barbin, A.; Guichard, Y.; Bartsch, H. *Carcinogenesis* **1995**, *16*, 613−617. (b) Nath, R. G.; Chung, F. -L. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7491−7495.
(6) (a) Shibutani, S.; Suzuki, N.; Matsumoto, Y.; Grollman, A. P. *Biochemistry* **1996**, *35*, 14992−14998. (b) Simha, D.; Palejwala, V. A.; Humayun, M. Z. *Biochemistry* **1991**, *30*, 8727−8735.
(7) Speina, E.; Zielinska, M. *Cancer Res.* **2003**, *63*, 4351−4357.
(8) (a) Jacobsen, J. S.; Humayun, M. *Z. Biochem.* **1990**, *29*, 496−504. (b) Jacobsen, J. S.; Perkins, C. P.; Callahan, J. T.; Sambamurti, K.; Humayun, M. Z. *Genetics* **1989**, *121*, 213−222.
(9) Moriya, M.; Zhang, W.; Johnson, F.; Grollman, A. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11899−11903.
(10) Basu, A. K.; Wood, M. L.; Niedernhofer, L. J.; Ramos, L. A.; Essigmann, J. M. *Biochemistry* **1993**, *32*, 12793−12801.
(11) Friedberg, E. C.; Walker, G. C.; Siede, W. *DNA Repair and Mutagenesis*; ASM Press: Washington, DC, 1995.
(12) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503−506.
(13) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford Science: Oxford, U.K., 1989.
(14) Loos, P. F.; Assfeld, X.; Rivail, J. L. *Theor. Chem. Acc.* **2007**, *118*, 165−171.
(15) Magalhaes, A. L.; Madail, S. R. R. S.; Ramos, M. *J. Theor. Chem. Acc.* **2000**, *105*, 68−76.
(16) Fuqiang, B.; Kathryn, N. R.; James, W. G.; Russell, J. B. *Theor. Chem. Acc.* **2002**, *108*, 1−11.
(17) Bhattacharyya, D.; Koripella, S. C.; Mitra, A.; Rajendran, V. B.; Sinha, B. *J. Biosci.* **2007**, *32*, 809−825.
(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
(19) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100. (b) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652. (c) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554−8560. (d) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789. (e) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624−9631.
(20) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(21) Chai, J. D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(22) (a) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503−506. (b) Saebo, S.; Almloef, J. *Chem. Phys. Lett.* **1989**, *154*, 83−89. (c) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275−280. (d) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281−289.
(23) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873−1885.
(24) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.
(25) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995−2001.
(26) Lingaraju, G. M.; Davis, C. A.; Setser, J. W.; Samson, L. D.; Drennan, C. L. *J. Biol. Chem.* **2011**, *286*, 13205−13213.
(27) Gros, L.; Ishchenko, A. A.; Saparbaev, M. *Mutat. Res.* **2003**, *531*, 219−229.
(28) Palejwala, V. A.; Rzepka, R. W.; Simha, D.; Humayun, M. Z. *Biochemistry* **1993**, *32*, 4105−4111.
(29) (a) Leontis, N. B.; Stombaugh, J.; Westhof, E. *Nucleic Acids Res.* **2002**, *30*, 3497−3531. (b) Lemieux, S.; Major, F. *Nucleic Acids Res.* **2002**, *30*, 4250−4263. (c) Das, J.; Mukherjee, S.; Mitra, A.; Bhattacharyya, D. *J. Biomol. Struct. Dyn.* **2006**, *24*, 149−161. (d) Lu, X.-J.; Olson, W. K. *Nat. Protoc.* **2008**, *3*, 1213−1227.
(30) (a) Srinivasadesikan, V.; Sahu, P. K.; Lee, S.-L. *J. Phys. Chem. B* **2011**, *115*, 10537−10546. (b) Sahu, P. K.; Kuo, C.-W; Lee, S.-L. *J. Phys. Chem. B* **2007**, *111*, 2991−2998.
(31) Czyznikowska, Z.; Gora, R. W.; Zalesny, R.; Lipkowski, P.; Jarzembska, K. N.; Dominiak, P. M.; Leszczynski, J. *J. Phys. Chem. B* **2010**, *114*, 9629−9644.
(32) Sponer, J.; Jurecka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142−10151.
(33) Hobza, P.; Kabelac, M.; Sponer, J.; Mejzlik, P.; Vondrasek, J. *J. Comput. Chem.* **1997**, *18*, 1137−1150.
(34) Bren, U.; Zupan, M.; Guengerich, F. P.; Mavri, J. *J. Org. Chem.* **2006**, *71*, 4078−4084. and 2 Galesa, K.; Bren, U.; Kranjc, A.; Mavri, J. *J. Agric. Food Chem.* **2008**, *56*, 8720−8727. and 3 Bren, U.; Guengerich, F. P.; Mavri, J. *Chem. Res. Toxicol.* **2007**, *20*, 1134−1140.
(35) Thanthiriwatte, K. S.; Hohenstein, E. G.; Burns, L. A.; Sherrill, C. D. *J. Chem. Theory Comput.* **2011**, *7*, 88−96.
(36) (a) Bernstein, C; Bernstein, H; Payne, C. M.; Garewal, H. *Mutat. Res.* **2002**, *511*, 145−178. (b) Bernstein, H.; Byerly, H.; Hopf, F.; Michod, R. *Science* **1985**, *229*, 1277−1281.