Computational Study of the Deamination Reaction of Cytosine with H_2O and OH^-

Mansour H. Almatarneh, Christopher G. Flinn, and Raymond A. Poirier*

Department of Chemistry, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X7

W. Andrzej Sokalski

Molecular Modelling Laboratory, Institute of Physical and Theoretical Chemistry I-30, Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland

Received: April 13, 2006; In Final Form: May 16, 2006

The mechanism for the deamination reaction of cytosine with H_2O and OH^- to produce uracil was investigated using ab initio calculations. Optimized geometries of reactants, transition states, intermediates, and products were determined at RHF/6-31G(d), MP2/6-31G(d), and B3LYP/6-31G(d) levels and for anions at the B3LYP/6-31+G(d) level. Single-point energies were also determined at B3LYP/6-31+G(d), MP2/GTMP2Large, and G3MP2 levels of theory. Thermodynamic properties (ΔE, ΔH, and ΔG), activation energies, enthalpies, and free energies of activation were calculated for each reaction pathway that was investigated. Intrinsic reaction coordinate analysis was performed to characterize the transition states on the potential energy surface. Two pathways for deamination with H_2O were found, a five-step mechanism (pathway A) and a two-step mechanism (pathway B). The activation energy for the rate-determining steps, the formation of the tetrahedral intermediate for pathway A and the formation of the uracil tautomer for pathway B, are 221.3 and 260.3 kJ/mol, respectively, at the G3MP2 level of theory. The deamination reaction by either pathway is therefore unlikely because of the high barriers that are involved. Two pathways for deamination with OH^- were also found, and both of them are five-step mechanisms. Pathways C and D produce an initial tetrahedral intermediate by adding H_2O to deprotonated cytosine which then undergoes three conformational changes. The final intermediate dissociates to product via a 1−3 proton shift. Deamination with OH^-, through pathway C, resulted in the lowest activation energy, 148.0 kJ/mol, at the G3MP2 level of theory.

1. Introduction

Cytosine (Cyt), one of the pyrimidine bases, occurs naturally in many nucleic acids, DNA, and RNA and is the most unstable of the DNA bases. It is chemically bound to a sugar moiety and interacts with other nucleic acid bases via hydrogen bonds, most frequently with guanine.[1] Cytosine is also a parent compound of various modified nucleosides and nucleotides. Some of them occur naturally, and others are the product of chemical reactions of nucleic acids with various mutagenic agents.[2] Among nucleic acid bases, cytosine is the most alkaline in aqueous solution ($pK_a = 4.6$). This property plays an important role in many biochemical processes.[3] Cytosine, its nucleosides and nucleotides, and many of its derivatives have been extensively studied experimentally,[4−18] as well as computationally, in both the gas phase[17−42] and the aqueous phase.[43−49]

The nucleic acid bases have tremendous versatility in the formation of hydrogen bond complexes because of the presence of numerous hydrogen bond donor and acceptor groups. These interactions are responsible for maintaining the genetic code.[3] Tautomerism is fundamentally important to the structure and functioning of nucleic acid bases. Interest in this area is due to the fact that the formation of rare tautomers can induce alterations in the normal base pairing, leading to spontaneous mutations in the genetic code.[50] Hence, a great deal of research has been carried out on the tautomerism of nucleic acid bases using both experimental and theoretical approaches. Conversion of one tautomer to another is generally the result of proton transfer reactions whose activation energy barriers may control the formation of higher-energy tautomeric forms (rare tautomers). Purine and pyrimidine bases exist in several different tautomeric forms which differ from each other mainly in the position of one of the hydrogens which may be bound to the exocyclic nitrogen or oxygen atom or one of the ring nitrogen atoms. It has been suggested that conversion between tautomers can occur as a result of a simultaneous double-proton transfer, but such conversions are energetically rather unfavorable.[36] Numerous computational studies have discussed the tautomerism of the cytosine molecule[18,34−38,45−49] which have provided a reliable picture of the relative stability of its tautomers, both in the gas phase and in solution. Using infrared spectrometry, it has been found that, in Ar and N_2 matrices at 15 K, cytosine exists as a mixture of the "normal" amino-oxo (a-o), the canonical form found in DNA, and the "rare" aminohydroxy (a-h) tautomers.[6] In addition to the canonical form and the a-h tautomer, there are imine tautomers. Since the tautomers are very close in energy, their relative stabilities are very sensitive to the level of theory.[35,36] It appears that the canonical form is the local minimum at almost all theoretical levels.[38] DFT calculations favor slightly the a-o form.[35,51] The a-h tautomer has been predicted to predominate in the gas phase using a coupled cluster.[49] A microwave investigation has identified three of the six possible tautomers.[17]

The relative energies of the cytosine−water conformers have been computed at the MP2/6-31G(d) level of theory, and the results have shown that the cytosine (keto)−water complex is

10.1021/jp062300u CCC: $33.50
Published on Web 06/15/2006

more stable than the cytosine (enol)−water complex by 4.5 kJ/mol.[39] It has also recently been shown that the a-o tautomeric form of the cytosine−water complex is stable in both the ground and excited states.[52] By comparison with experimental data,[6,17] it was shown that kinetic barriers in the gas phase are generally high (126−155 kJ/mol) which affects the tautomer population at ≤500 K.[47] It must be mentioned that the interaction between the canonical form of cytosine and one and two water molecules has been investigated in several theoretical works.[37−39,47,48] Theoretical studies have shown that interaction with water molecules changes the relative energies of some or all tautomers, the canonical tautomer being better hydrated than the other tautomers.[53] Florian et al.[18] studied the relative stabilities of tautomers of protonated cytosine in the gas phase and in a polar solvent at the ab initio HF, MP2, and polarizable continuum approximations. In addition, the infrared and Raman frequencies and intensities for cytosine and its two protonated forms were calculated.

Metal ions can also affect the relative stability of tautomers of the pyrimidine base cytosine. There have been several studies of the interaction of cytosine and its tautomers with metals.[32,40−42] Metalation can play a role in the formation of rare tautomers of nucleic acid bases and also affect the ability of the nucleobase to be protonated or deprotonated.[42] More recently, Prado et al.[54] studied the Mg^{2+}−cytosine complex interaction at RHF, MP2, DFT, and CCSD(T) levels with the 6-31+G(d) basis set.

Interactions of other species such as hydrogen peroxide and hydroxyl radical have also been the object of recent studies. For example, authors have studied the interaction between the a-o and *cis* a-h tautomers of cytosine with one hydrogen peroxide molecule. The binding energies and harmonic vibrational frequencies for the complexes were calculated using the B3LYP/6-31++G(d,p) level of theory.[21] DFT calculations have shown that the addition of an OH radical to the C_5 and C_6 sites is thermodynamically and kinetically more favorable than addition to the N_3 and C_2 sites.[20]

Spontaneous mutations (changes to the nucleotide sequences of DNA) can arise as a result of chemical changes to individual bases in DNA. One such chemical change is the conversion of cytosine to uracil which is classified as a deamination reaction. In general, deamination refers to the loss of an amino group from a tetrahedral carbon with conversion to a carbonyl functional group. All the DNA bases, except thymine and uracil, have amino groups, but only deamination of cytosine gives a base found in DNA and RNA. Hydrolytic deamination of cytosine yields uracil, as shown in Figure 1.

Uracil (Ura) is found in RNA and can only base pair with adenine.[10] In principle, six tautomeric forms of uracil are possible,[35] and like cytosine, numerous studies on all tautomers of uracil have been reported.[27,35,36,39,45,55] If uracil is found in DNA, it poses a very serious problem. The cell, however, has a specific enzyme to remove it from DNA, called DNA uracil-*N*-glycosylase. The uracil formed by cytosine deamination is potentially mutagenic, changing the coding information during DNA replication and RNA transcription, resulting in altered base pairs in the genome.[13] Hydrolytic deamination is known to depend on pH and temperature.[9−12] It has been established that deamination of the DNA base cytosine is an extremely rare event under normal physiological conditions (40−100 deaminations in the human genome per day at pH 7.4), although the rate of deamination can be significantly increased in the presence of various reagents such as NO, HNO_2, and bisulfite. For example, bisulfite-induced deamination involving acid-catalyzed hydrolysis[14] has been shown to noticeably accelerate the rate of cytosine

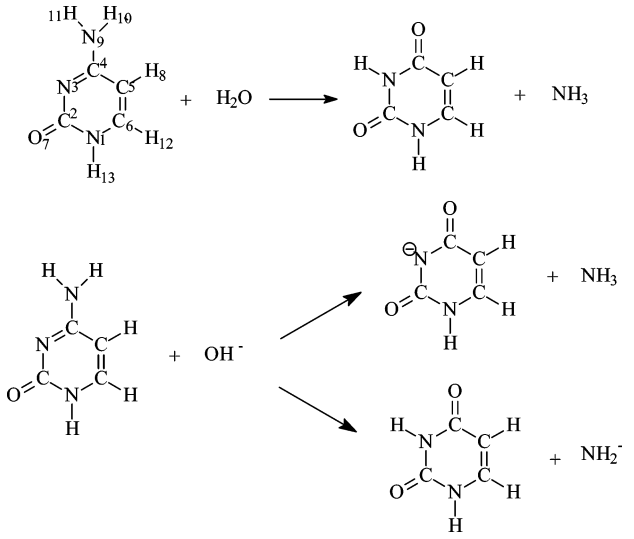

Figure 1. Deamination of cytosine with H_2O, deamination of cytosine with OH^- with two possible products, and atom labeling in the amino-oxo tautomer of cytosine.

deamination. Other pathways that were studied include a diazotization pathway[56] and base-catalyzed[10] and acid-catalyzed[8] deaminations. Federico et al.[16] were able to determine the rate constant of cytosine deamination for single- and double-stranded DNA under physiologically relevant conditions at 37 °C and pH 7.4 by a sensitive genetic assay. Their measured rate constants for single- and double-stranded DNA are 1×10^{-10} and 7×10^{-13} s^{-1}, respectively, with an activation energy of 117 ± 4 kJ/mol. This value agrees well with the value of 121 kJ/mol obtained over a 25 °C temperature range by Lindahl and Nyberg.[57]

Duncan et al.[58] pointed out that C_5 methylation of cytosine increases the likelihood of spontaneous mutations. They showed that the rate of deamination of 5-methylcytosine (m^5Cyt) is 3−4 times faster than that for cytosine. As a result, m^5Cyt residues are considered to represent hot spots for spontaneous transition mutations.[58] The deamination of m^5Cyt, which is involved in the regulation of gene expression, yields thymine which is naturally found in DNA. Replication of these deamination products will produce a C·G → T·A transition mutation.[13,16]

Yao et al.[4] studied the hydrolysis of cytosine to uracil by yeast cytosine deaminase, a zinc metalloenzyme. In particular, they studied the catalysis of the deamination of the prodrug 5-fluorocytosine to form the anticancer drug 5-fluorouracil. Sponer et al.[5] studied the metal-mediated deamination of 1-methylcytosine and 1,5-dimethylcytosine with a cationic complex of Pt^{II} both experimentally and using DFT calculations. They also studied the deamination of cytosine with OH^- using the PCM model to account for solvation effects. This study represents the only computational study of the hydrolytic deamination reaction of cytosine which has been reported.[5] However, their reported activation energy barrier (213.4 kJ/mol) is not very close to the experimentally accepted value (117 ± 4 kJ/mol).[16,55]

This paper represents a detailed computational study of the hydrolytic deamination of the DNA base cytosine with OH^- and H_2O and follows a similar study of the deamination of the related compound, formamidine (*E* isomer) with OH^-, H_2O, and H_3O^+.[59] The deamination of the structurally related *E*(*trans*) isomer of formamidine, as predicted, provided reaction pathways which proved to be very similar in terms of transition states, intermediates, and energetics to the reaction pathways of cytosine. As can be seen in Figure 2 below, the *E* isomer of

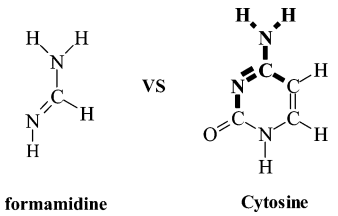

Figure 2. Comparison of formamidine (*E* isomer) and cytosine structure.

formamidine forms part of the cytosine structure, notably in the region of cytosine where deamination of cytosine to uracil can take place.

2. Computational Method

The MUNgauss[60] computational program was used for most of the geometry optimizations at the HF/6-31G(d) level for reactants, products, and transition state structures. All other calculations were performed with Gaussian 03.[61] The geometries of all reactants, transition states, intermediates, and products were fully optimized at the restricted HF, second-order Møller−Plesset (MP2) and B3LYP levels of theory using the 6-31G(d) basis set. Energies have also been calculated using the Gaussian-n theory, G3MP2. From our previous work,[62] we found that the activation energies and the heats of reaction calculated using Gaussian-n theories (G1, G2, G2MP2, G3, G3MP2, G3B3, and G3MP2B3) all agreed to within 10 kJ/mol, which is within the reported error of the Gaussian-n theories in the literature. On the basis of these results, we chose the G3MP2 level of theory which is the least computationally expensive method for providing reliable energetics. Single-point calculations at the B3LYP/6-31+G(d) and MP2/GTMP2Large levels of theory are also reported. All structures have been vibrationally characterized, i.e., checked for the absence of imaginary frequencies in the minima and for the presence of only one imaginary frequency in the transition states. The complete reaction pathways for all the mechanisms discussed in this paper have been verified using intrinsic reaction coordinate (IRC) analysis of all transition states. For each transition state, we optimized the structures at the last IRC points to positively identify the reactant and product to which each is connected.

3. Results and Discussion

Our previous results from the deamination reaction of formamidine[59] gave us a useful starting point for studying the mechanism of the deamination reaction of cytosine by providing good initial guesses for transition state structures that exist along the reaction pathways.

3.1. Deamination of Cytosine with H_2O in Pathways A and B. Computational studies predict that, in aqueous solution or the gas phase, cytosine protonation should occur at the N_3 position (see Figure 1).[18] Computational studies have also shown that the interaction with water changes the relative energies of the other tautomers, the canonical tautomer being better hydrated than the other tautomers.[53] Deamination of cytosine with H_2O can follow one of two possible pathways designated as pathway A and pathway B. These pathways are closely related to those found for the deamination of formamidine.[59] The geometries for the reactant, intermediates, transition states, and product involved in pathway A are shown in Figure 3. The relative energies of reactants, intermediates, transition states, and products for pathway A are shown in Figure 4.

Pathway A is a five-step mechanism. Initially, a hydrogen-bonded complex of cytosine with one water molecule forms

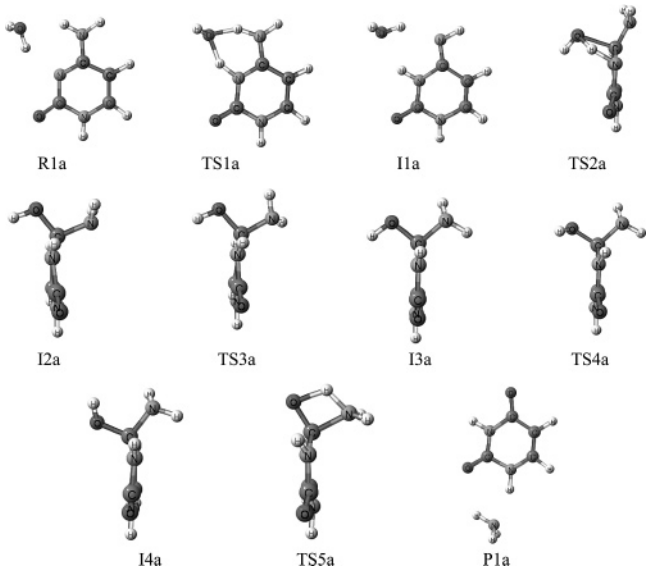

Figure 3. Deamination of cytosine with one water molecule to form the Ura−NH_3 complex (pathway A).

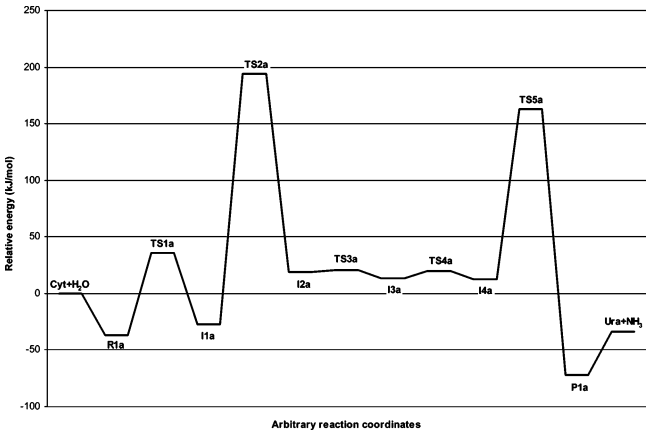

Figure 4. Reaction pathway for the deamination of cytosine with one water molecule at the G3MP2 level of theory (pathway A).

the imine−oxo tautomer of cytosine (**I1a**) with an activation energy of 72.6 kJ/mol at the G3MP2 level of theory. This is followed by nucleophilic attack by the water molecule on the C_4 atom and proton transfer from H_2O to the exocyclic imine nitrogen (sp^2 nitrogen) of cytosine to form a tetrahedral intermediate (**I2a**). This is followed by two conformational changes of the initial intermediate to give two new intermediates (**I3a** and **I4a**, in Figure 3), finally resulting in an intramolecular 1−3 proton transfer of the hydroxyl hydrogen to the amino group to yield a uracil−ammonia complex (**P1a**).

The geometries for the reactant, intermediates, transition states, and product involved in pathway B are shown in Figure 5. The relative energies of reactants, intermediates, transition states, and products for pathway B are shown in Figure 6. In pathway B, the water molecule attacks the C_4 atom with simultaneous proton transfer from H_2O to the amino group (sp^3 nitrogen) of cytosine producing the hydroxy-oxo tautomer of uracil and ammonia, **I1b**. A 1−3 proton shift of the hydroxy hydrogen to the sp^2 nitrogen of the tautomer forms the deamination product, the uracil−ammonia complex, as shown in Figure 5. Both complete reaction pathways have been verified using IRC analysis.

The mechanism for the deamination of cytosine with H_2O is very similar to that of the deamination of formamidine with

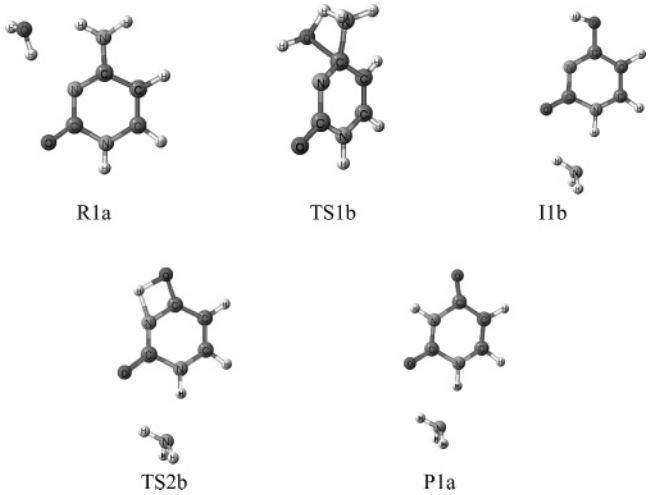

Figure 5. Deamination of cytosine with one water molecule to form the Ura−NH_3 complex (pathway B).

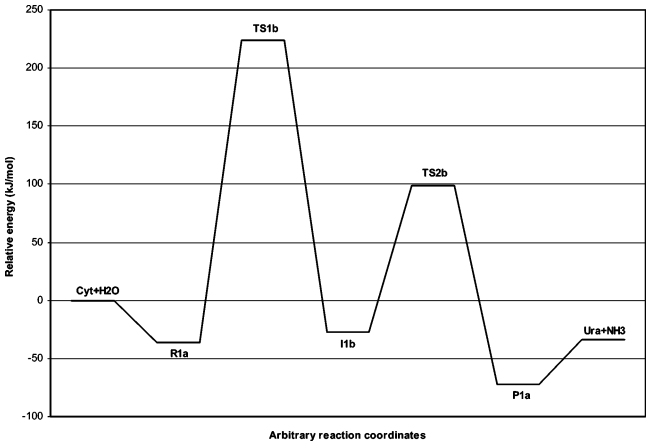

Figure 6. Reaction pathway for the deamination of cytosine with one water molecule at the G3MP2 level of theory (pathway B).

TABLE 1: Thermodynamic Properties for the Deamination of Cytosine with One Water Molecule in Kilojoules per Mole at 298.15 K

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	G3MP2
ΔE^a	−72.3	−62.3	−64.9	−35.3 (−36.2)[c]
ΔH^a	−69.1	−58.8	−62.2	−35.3
ΔG^a	−70.9	−60.5	−63.6	−37.0
ΔE^b	−75.0	−64.8	−70.7	−34.1 (−36.7)[c]
ΔH^b	−71.0	−60.6	−66.7	−35.5
ΔG^b	−70.0	−60.7	−66.1	−34.2

[a] For the Cyt−H_2O complex → Ura−NH_3 complex step. [b] For the separated species (Cyt + H_2O → Ura + NH_3). [c] The values in parentheses are for the MP2/GTMP2Large level of theory.

H_2O. The most notable differences were for pathway A in which the initial tetrahedral intermediate undergoes several conformational changes and the initial intermediate is formed from a tautomer of cytosine in the first step of the reaction. Pathway B is very similar for both cytosine and formamidine.

The thermodynamic properties for the deamination of cytosine with H_2O are listed in Table 1. The deamination reaction of cytosine is found to be exothermic and exergonic at all levels of theory for both separated species and for complex to complex, with negligible differences between them.

The activation energies, enthalpies of activation, and free energies of activation for the deamination of cytosine with H_2O at HF, MP2, and B3LYP levels of theory using the 6-31G(d) basis set and the G3MP2 level of theory for both pathways A and B are listed in Tables 2 and 3, respectively. The MP2 and B3LYP results are in reasonable agreement with G3MP2 results, differing by no more than 22 kJ/mol. The MP2/GTMP2Large results, for both the thermochemistry (Table 1) and barriers (Tables 2 and 3), are in excellent agreement with the G3MP2 values, differing by no more than 7 kJ/mol.

TABLE 2: Activation Energies, Enthalpies of Activation, and Free Energies of Activation for the Deamination of Cytosine with One Water Molecule in Kilojoules per Mole at 298.15 K (pathway A)

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	G3MP2
ΔE_{act}, **TS1**	138.8	85.3	73.5	72.6 (76.6)[a]
$\Delta H^{\ddagger}$, **TS1**	119.3	64.0	53.4	67.8
$\Delta G^{\ddagger}$, **TS1**	130.5	72.5	61.6	79.6
ΔE_{act}, **TS2**	271.3	240.2	223.2	221.3 (223.7)[a]
$\Delta H^{\ddagger}$, **TS2**	264.2	226.2	209.8	218.8
$\Delta G^{\ddagger}$, **TS2**	270.1	233.1	213.9	225.0
ΔE_{act}, **TS3**	5.0	5.7	1.8	1.7 (2.7)[a]
$\Delta H^{\ddagger}$, **TS3**	2.3	3.1	−0.7	0.5
$\Delta G^{\ddagger}$, **TS3**	4.1	4.1	1.7	2.5
ΔE_{act}, **TS4**	8.8	10.6	10.0	6.6 (9.2)[a]
$\Delta H^{\ddagger}$, **TS4**	5.6	7.1	6.8	5.4
$\Delta G^{\ddagger}$, **TS4**	7.8	8.8	9.5	7.9
ΔE_{act}, **TS5**	210.6	157.8	145.9	150.1 (156.7)[a]
$\Delta H^{\ddagger}$, **TS5**	194.9	141.8	129.3	149.4
$\Delta G^{\ddagger}$, **TS5**	196.0	141.1	128.9	150.6

[a] The values in parentheses are for the MP2/GTMP2Large level of theory.

TABLE 3: Activation Energies, Enthalpies of Activation, and Free Energies of Activation for the Deamination of Cytosine with One Water Molecule in Kilojoules per Mole at 298.15 K (pathway B)

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	G3MP2
ΔE_{act}, **TS1**	338.8	258.2	260.1	260.3 (258.3)[a]
$\Delta H^{\ddagger}$, **TS1**	329.3	246.5	249.0	256.2
$\Delta G^{\ddagger}$, **TS1**	339.6	254.8	256.7	267.1
ΔE_{act}, **TS2**	186.6	131.8	129.7	125.7 (132.4)[a]
$\Delta H^{\ddagger}$, **TS2**	172.0	118.4	116.2	125.2
$\Delta G^{\ddagger}$, **TS2**	172.3	118.7	116.4	125.7

[a] The values in parentheses are for the MP2/GTMP2Large level of theory.

An activation energy of 117 ± 4 kJ/mol was reported experimentally for this reaction.[16,57] In this study, the activation energies for the rate-determining step for pathways A and B are 221.3 and 260.3 kJ/mol, respectively, at the G3MP2 level of theory. These values are high compared with the experimental value and may explain why this reaction is an extremely rare event under neutral conditions. The activation energies for step 3 (**TS3a**) and step 4 (**TS4a**), which represent the formation of different conformers, are 1.7 and 6.6 kJ/mol at G3MP2, respectively. Pathway A is a more likely mechanism for the deamination with H_2O compared to pathway B due to a significantly lower activation energy (221.3 and 260.3 kJ/mol, respectively) for the rate-determining step. These activation energies are only slightly higher than those found for formamidine, 212.7 kJ/mol for pathway A and 249.2 kJ/mol for pathway B (at the G2 level of theory).[59] It can be seen from Tables 2 and 3 that the activation energies for this reaction are much higher than the experimentally accepted value, and hence, we can also conclude that neither pathway is likely to account for cytosine deamination in DNA. Although the barriers for the deamination of cytosine with H_2O are high compared to the experimental value, this mechanism may be important

TABLE 4: Thermodynamic Properties for the Deamination of Cytosine with OH^- in Kilojoules per Mole at 298.15 K

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	B3LYP/ 6-31+G(d)	G3MP2
			Cyt^--H_2O complex → $Ura^-(N_3)-NH_3$ complex[b]		
ΔE	−53.4	−40.3	−45.5	−33.5 (−35.0)[a]	−18.9 (−15.1)[c]
ΔH	−53.1	−37.9	−43.7	−31.2	−17.1
ΔG	−58.5	−40.8	−48.7	−37.0	−22.6
			$Cyt + OH^-$ → $Ura^-(N_3) + NH_3$[b]		
ΔE	−328.2	−339.6	−347.8	−229.9 (−229.8)[a]	−223.0 (−224.9)[c]
ΔH	−325.2	−337.6	−346.0	−229.3	−223.3
ΔG	−329.4	−344.5	−351.1	−233.8	−227.1
			$Cyt + OH^-$ → $Ura + NH_2^-$		
ΔE	−11.6	−2.9	−21.0	21.2 (23.4)[a]	18.5 (25.7)[c]
ΔH	−15.5	−6.7	−24.9	18.0	18.3
ΔG	−18.5	−10.8	−28.3	15.3	15.7
			$Cyt-H_2O$ complex → $Ura^-(N_1)-NH_3$ complex[b]		
ΔE	−115.4	−97.3	−103.5	−87.8 (−89.1)[a]	−65.5 (−64.3)[c]
ΔH	−114.1	−93.0	−100.2	−84.1	−64.1
ΔG	−119.1	−94.4	−104.5	−88.3	−69.2
			$Cyt + OH^-$ → $Ura^-(N_1) + NH_3$[b]		
ΔE	−395.7	−402.8	−412.0	−288.5 (−288.3)[a]	−272.6 (−278.0)[c]
ΔH	−391.5	−398.7	−408.4	−286.4	−273.7
ΔG	−393.8	−402.1	−411.2	−288.6	−275.5

[a] The values in parentheses are for the B3LYP/6-31+G(d)//B3LYP/6-31G(d) level of theory. [b] $Ura^-(N_1)$ and $Ura^-(N_3)$ for deprotonation at N_1 and N_3, respectively. [c] The values in parentheses are for the MP2/GTMP2Large level of theory.

for the catalyzed reaction or for deamination of cytosine in DNA or RNA.

3.2. Deamination of Cytosine with OH^-. Since this reaction takes place under physiological conditions (pH 7.4), we have included a detailed study of possible mechanisms for the deamination of cytosine with OH^-. Our results show that in this case, deamination of cytosine could occur via one of two possible pathways, designated as pathways C and D.

The thermodynamic properties for the deamination of cytosine with OH^- are listed in Table 4. Two possibilities for the separated products, uracil anion and ammonia, and uracil and amide anion (azanide), were considered as shown in Figure 1. Deamination of cytosine with OH^- to form a uracil anion is found to be highly exothermic and exergonic at all levels of theory for the separated species and less so for the formation of the Ura^--NH_3 complex. Formation of Ura and NH_2^- is found to be slightly exothermic and exergonic at HF, MP2, and B3LYP levels of theory and to be endothermic and endergonic at B3LYP/6-31+G(d) and G3MP2 levels of theory. This is due to the highly resonance-stabilized uracil anion which has a highly delocalized negative charge, versus the highly localized negative charge of NH_2^- (an extremely strong base). For the reaction involving separate species, the addition of diffuse functions is shown here to be essential. The B3LYP/6-31+G(d)//B3LYP/6-31G(d) and B3LYP/6-31+G(d)//B3LYP/6-31+G(d) results are in excellent agreement.

3.3. Deamination of Cytosine with OH^- (Pathway C). The geometries for the reactant, intermediates, transition states, and product involved in pathway C are shown in Figure 7. The relative energies of reactants, intermediates, transition states, and products for pathway C are shown in Figure 8.

Pathway C is a five-step mechanism. Deamination of cytosine by OH^- initially involves deprotonation at the amino group (H_{10}) of cytosine, as shown in Figure 7, to form a cytosine anion−H_2O complex. This complex is very stable due to delocalization of the negative charge. Despite extensive attempts, we were not successful in optimizing the geometry of a cytosine−OH^- complex due to the ease of deprotonation of cytosine by OH^-. In the next step, H_2O adds to C_4, simultaneously adding hydrogen to the imine nitrogen resulting in a tetrahedral intermediate anion (**I1**). This is followed by the formation of three different conformers (**I2**, **I3c**, and **I4c**). These structures are very similar, differing mainly in the torsion for H_{15} ($\angle H_{15}O_8C_4N_3$) and the $H_{15}-N_9$ bond length (Figure 7). The reactant, intermediates, transition states, and product illustrated in Figures 7 and 8 correspond to attack of OH^- on the *si* face (see Figure 9) of cytosine. The mirror of this mechanism also exists for attack of OH^- on the *re* face of cytosine, where the intermediates and transition state structures are enantiomers. In the last step, deamination occurs by intramolecular proton transfer from the hydroxyl group to the amino group to yield a hydrogen-bonded complex of the uracil anion (formal negative charge at N_3) and ammonia. The only other study of cytosine deamination with OH^- was reported by Sponer and co-workers.[5] Their results have shown that the activation energy for this system is 213.4 kJ/mol at the B3LYP/6-31G(d) level of theory using the PCM model. The potential energy surface for this system shows that the intermediate and the second transition state have the same energy relative to the reactant (Cyt^--H_2O) complex. The two transition states identified by the authors are not connected in the reaction mechanism as revealed by IRC

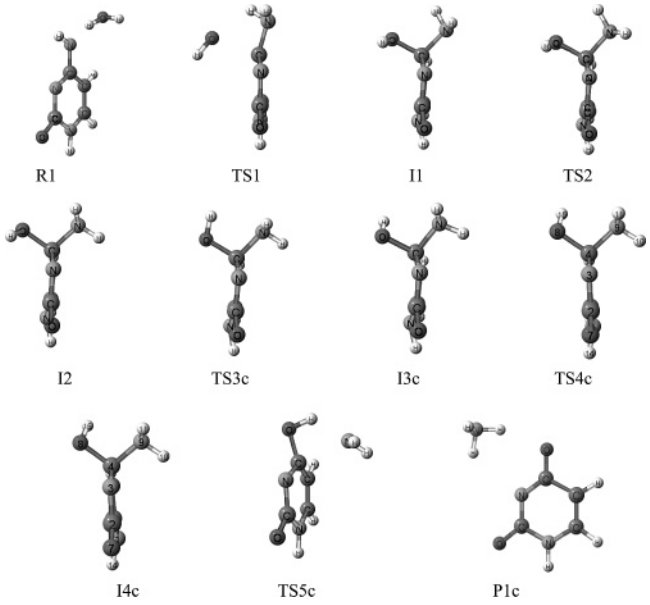

Figure 7. Deamination of cytosine with OH^- on the *si* face (see Figure 9) to form the Ura anion−NH_3 complex (pathway C).

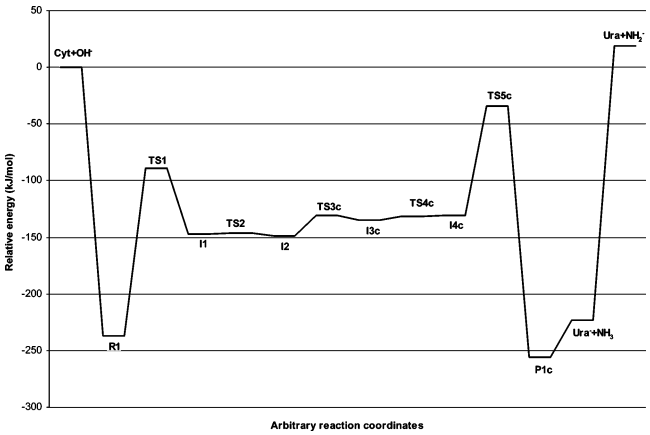

Figure 8. Reaction pathway for the deamination of cytosine with OH^- at the G3MP2 level of theory (pathway C).

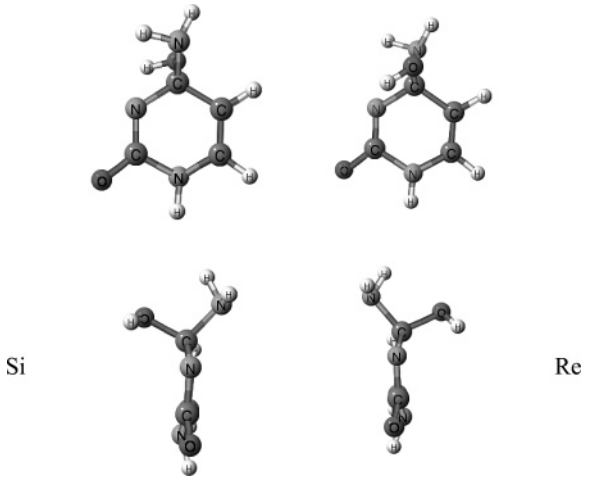

Figure 9. Two enantiomers of the intermediate structure (**I1**) corresponding to attack of OH^- on the *si* face and on the *re* face of the cytosine ring.

TABLE 5: Activation Energies, Enthalpies of Activation, and Free Energies of Activation for the Deamination of Cytosine with OH^- in Kilojoules per Mole at 298.15 K (pathway C)

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	G3MP2
ΔE_{act}, **TS1**	199.5	188.9	177.3	148.0 (153.8)[a]
$\Delta H^{\ddagger}$, **TS1**	194.6	184.5	172.5	145.5
$\Delta G^{\ddagger}$, **TS1**	204.6	195.8	180.3	155.7
ΔE_{act}, **TS2**	1.7	2.5	0.9	1.0 (1.8)[a]
$\Delta H^{\ddagger}$, **TS2**	−0.5	0.1	−1.4	−0.7
$\Delta G^{\ddagger}$, **TS2**	2.4	2.5	2.0	2.6
ΔE_{act}, **TS3c**	26.1	24.7	21.5	17.5 (19.9)[a]
$\Delta H^{\ddagger}$, **TS3c**	23.0	21.6	18.4	16.3
$\Delta G^{\ddagger}$, **TS3c**	25.0	23.4	19.9	18.6
ΔE_{act}, **TS4c**	3.0	6.1	8.2	2.7 (2.0)[a]
$\Delta H^{\ddagger}$, **TS4c**	−0.1	2.9	4.9	1.2
$\Delta G^{\ddagger}$, **TS4c**	3.2	5.0	3.8	4.8
ΔE_{act}, **TS5c**	138.0	126.5	92.7	97.0 (108.5)[a]
$\Delta H^{\ddagger}$, **TS5c**	125.0	110.2	79.9	97.8
$\Delta G^{\ddagger}$, **TS5c**	124.4	106.2	77.2	97.2

[a] The values in parentheses are for the MP2/GTMP2Large level of theory.

analysis. One of the transition states belongs to pathway C for attack of OH^- on the *re* face of cytosine and the other to pathway C for attack of OH^- on the *si* face of cytosine. For the last step of their mechanism, they reported an activation energy of 73.6 kJ/mol at the B3LYP/6-31G(d) level of theory compared to our value of 97.0 kJ/mol at the G3MP2 level of theory.

The activation energies, enthalpies of activation, and free energies of activation for the deamination of cytosine with OH^- at the HF, MP2, and B3LYP levels of theory using the 6-31G(d) basis set and the G3MP2 level of theory for pathway C are listed in Table 5. An activation energy of 213.4 kJ/mol [5] was reported for the deamination of cytosine with OH^- at the B3LYP/6-31G(d) level of theory using the PCM solvation model, which is far from the experimental value. In this study, the activation energies for the rate-determining step are 199.5, 188.9, 177.3, and 148.0 kJ/mol at the HF, MP2, B3LYP, and G3MP2 levels of theory, respectively. The G3MP2 result is fairly close to the experimental value. The small difference could be due in part to the environment of cytosine in DNA and RNA, where it is bonded to a sugar moiety and is hydrogen bonded to other nucleic acid bases.

3.4. Deamination of Cytosine with OH^- (Pathway D). The structures of the reactant, intermediates, transition states, and product involved in pathway D are shown in Figure 10. The relative energies of reactants, intermediates, transition states, and products for pathway D are shown in Figure 11. The complete reaction pathway was determined with the help of IRC analysis. Pathway D is a five-step mechanism. The first two steps are the same as those in pathway C. After the intermediate, **I2**, the reaction bifurcates into two different pathways, pathway C via transition state **TS3c** and pathway D via transition state **TS3d**. In step 4 of pathway D, intermediate **I3d** is converted, by elimination of NH_3, to a uracil anion tautomer−NH_3 complex via **TS4d**, where the C_4-NH_2 bond breaks and the NH_2 group abstracts the hydrogen atom (H_{13}) from N_1. In the final step, a 1−3 proton shift from the hydroxy group to the sp^2 nitrogen of the uracil tautomer results in formation of a uracil anion with the negative charge localized at N_1.

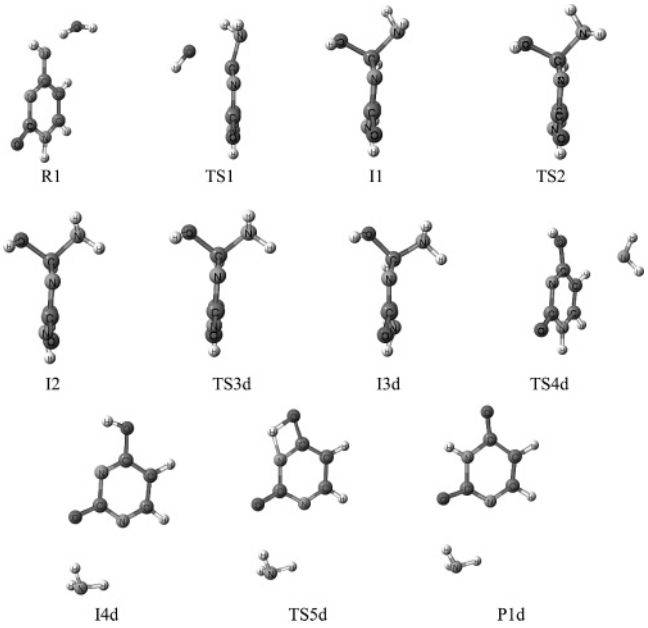

Figure 10. Deamination of cytosine with OH^- to form the Ura anion−NH_3 complex (pathway D).

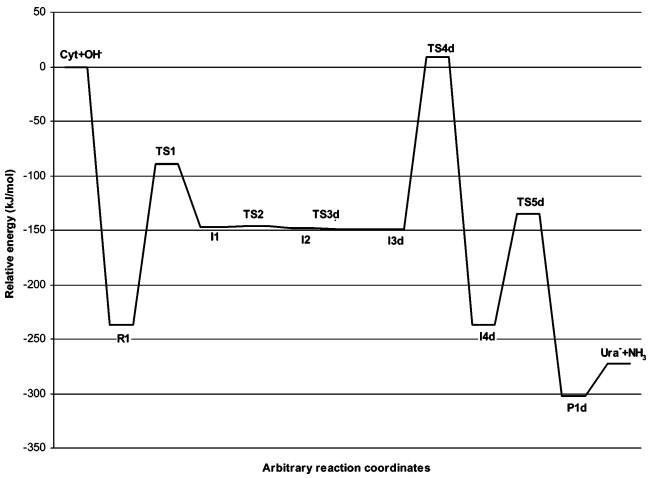

Figure 11. Reaction pathway for the deamination of cytosine with OH^- at the G3MP2 level of theory (pathway D).

The activation energies, enthalpies of activation, and free energies of activation for the deamination of cytosine with OH^- at the HF, MP2, and B3LYP levels of theory using the 6-31G(d) basis set and the G3MP2 level of theory for pathway D are listed in Table 6. The activation energies for the rate-determining step are 218.6, 233.3, 170.4, and 158.0 kJ/mol at the HF, MP2,

TABLE 6: Activation Energies, Enthalpies of Activation, and Free Energies of Activation for the Deamination of Cytosine with OH^- in Kilojoules per Mole at 298 K (pathway D)

	HF/ 6-31G(d)	MP2/ 6-31G(d)	B3LYP/ 6-31G(d)	G3MP2
ΔE_{act}, **TS1**	199.5	188.9	177.3	148.0 (153.8)[a]
$\Delta H^{\ddagger}$, **TS1**	194.6	184.5	172.5	145.5
$\Delta G^{\ddagger}$, **TS1**	204.6	195.8	180.3	155.7
ΔE_{act}, **TS2**	1.7	2.5	0.9	1.0 (1.8)[a]
$\Delta H^{\ddagger}$, **TS2**	−0.5	0.1	−1.4	−0.7
$\Delta G^{\ddagger}$, **TS2**	2.4	2.5	2.0	2.6
ΔE_{act}, **TS3d**	0.4	2.4	0.5	−0.3 (0.4)[a]
$\Delta H^{\ddagger}$, **TS3d**	−2.5	−0.8	−2.3	−1.9
$\Delta G^{\ddagger}$, **TS3d**	1.1	1.3	1.2	1.9
ΔE_{act}, **TS4d**	218.6	233.3	170.4	158.0 (182.3)[a]
$\Delta H^{\ddagger}$, **TS4d**	204.1	218.9	157.5	161.4
$\Delta G^{\ddagger}$, **TS4d**	196.2	211.2	151.6	153.2
ΔE_{act}, **TS5d**	157.6	107.7	103.6	101.9 (108.9)[a]
$\Delta H^{\ddagger}$, **TS5d**	143.5	94.8	90.7	101.4
$\Delta G^{\ddagger}$, **TS5d**	143.6	95.2	91.3	101.7

[a] The values in parentheses are for the MP2/GTMP2Large level of theory.

B3LYP, and G3MP2 levels of theory, respectively. Again, as for the reaction with water, the MP2/GTMP2Large results, for both the thermochemistry (Table 4) and barriers (Tables 5 and 6), are in good agreement with the G3MP2 values. Except for **TS4d**, where G3MP2 and MP2/GTMP2Large results differ by 24 kJ/mol, the results are all within 12 kJ/mol.

In summary, the more likely mechanism for the deamination of cytosine is pathway C with an activation energy of 148.0 kJ/mol, which is reasonably close to the experimental value, whereas pathway D is less likely to be the mechanism for this reaction due to the high barriers.

4. Conclusions

The mechanism for the deamination reaction of cytosine with H_2O and OH^- to produce uracil was investigated using ab initio calculations. Optimized geometries were determined at the RHF/6-31G(d), MP2/6-31G(d), and B3LYP/6-31G(d) levels of theory and at the B3LYP/6-31+G(d) level for the anions. Thermodynamic properties (ΔE, ΔH, and ΔG), activation energies, enthalpies, and free energies of activation were calculated for each reaction pathway that was investigated. Intrinsic reaction coordinate IRC analysis was carried out for all transition state structures to obtain the complete reaction pathway. The MP2/GTMP2Large result is overall in better agreement with the G3MP2 result, which may be an acceptable level of theory for larger systems. Two pathways for deamination with H_2O were found, a five-step mechanism (pathway A) and a two-step mechanism (pathway B). The deamination reaction by either pathway is unlikely because of the high barriers that are involved but may be a viable mechanism for catalyzed reactions, i.e., with bisulfite. Also, two pathways for deamination with OH^- were found, five-step mechanisms for both pathways C and D. Deamination with OH^- yields a tetrahedral intermediate with an activation energy much lower than that for the reaction with H_2O. The more likely mechanism of deamination of cytosine is pathway C with an activation energy of 148.0 kJ/mol, which is reasonably close to the experimental value, where the difference may be due to environmental factors in DNA.

Acknowledgment. We thank the Natural Sciences and Engineering Council of Canada (NSERC) for financial support. Also, we gratefully acknowledge the Memorial University of Newfoundland Advanced Computation and Visualization Centre (CVC) and the Atlantic Computational Excellence Network (ACEnet) for computer time.

Supporting Information Available: Full geometries and energies of all structures reported for all pathways at all levels of theory that were used. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Neidle, S. *Oxford Handbook of Nucleic Acid Structure*; Oxford University Press Inc.: New York, 1999.
(2) Burgur, A. *A Guide to the Chemical Basis of Drug Design*; Wiley: New York, 1983.
(3) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983.
(4) Yao, L.; Li, Y.; Wu, Y.; Liu, A.; Yan, H. *Biochemistry* **2005**, *44*, 5940−5947.
(5) Sponer, J. E.; Miguel, P. J.; Rodriguez-Santiago, L.; Erxleben, A.; Krumm, M.; Sodupe, M.; Sponer, J.; Lippert, B. *Angew. Chem., Int. Ed.* **2004**, *43*, 5396−5399.
(6) Person, W. B.; Szczepaniak, K.; Szczesniak, M.; Kwiatkowski, J. S.; Hernandez, L.; Czerminski, R. *THEOCHEM* **1989**, *194*, 239.
(7) Shapiro, R.; Servis, R. E.; Wecher, M. *J. Am. Chem. Soc.* **1970**, *92*, 422.
(8) Shapiro, R.; Klein, R. *Biochemistry* **1967**, *6*, 3576.
(9) Brown, D.; Phillips, J. H. *J. Mol. Biol.* **1965**, *11*, 663−671.
(10) Notari, R. E.; Chin, M. L.; Cardoni, A. *J. Pharm. Sci.* **1970**, *59*, 28.
(11) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *98*, 6338.
(12) Shapiro, R.; Klein, R. *Biochemistry* **1966**, *5*, 2358.
(13) Peng, W.; Shaw, B. R. *Biochemistry* **1996**, *35*, 10172−10181.
(14) Chen, H.; Shaw, B. R. *Biochemistry* **1994**, *33*, 4121−4129.
(15) Bobek, M.; Cheng, Y. C.; Bloch, A. *J. Med. Chem.* **1978**, *21*, 597.
(16) Frederico, L. A.; Kunkel, T. A.; Shaw, B. R. *Biochemistry* **1990**, *29*, 2532−2537.
(17) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1989**, *111*, 2308−2310.
(18) Florián, J.; Baumruk, V.; Leszczyński, J. *J. Phys. Chem.* **1996**, *100*, 5578−5589.
(19) Šponer, J.; Hobza, P. *J. Phys. Chem.* **1994**, *98*, 3161−3164.
(20) Ji, Y. J.; Xia, Y. Y.; Zhao, M. W.; Huang, B. D.; Li, F. *THEOCHEM* **2005**, *723*, 123−129.
(21) Wysokinski, R.; Bienko, D. C.; Michalska, D.; Zeegers-Huyskens, T. *Chem. Phys.* **2005**, *315*, 17−26.
(22) Florián, J.; Leszczyński, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010−1017.
(23) Šponer, J.; Leszczyński, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965−1974.
(24) Hobza, P.; Šponer, J.; Polšek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792−798.
(25) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBulat, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319−8330.
(26) Smets, J.; Jalbout, A. F.; Adamowicz, L. *Chem. Phys. Lett.* **2001**, *342*, 342−346.
(27) Hou, X.-J.; Nguyen, M. T. *Chem. Phys.* **2005**, *310*, 1−9.
(28) Fülscher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 2089−2095.
(29) Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L. *Photochem. Photobiol.* **1994**, *59*, 125−129.
(30) Kwiatkowski, J. S.; Leszczyñski, J. *J. Phys. Chem.* **1996**, *100*, 941−953.
(31) Danilov, V. I.; Les, A.; Alderfer, J. L. *Pol. J. Chem.* **2001**, *75*, 1039−1049.
(32) Monajjemi, M.; Ghiasi, R.; Ketabi, S.; Passdar, H.; Mollaamin, F. *J. Chem. Res.* **2004**, 11−18.
(33) Zhanpeisov, N. U.; Leszczyñski, J. *J. Phys. Chem. B* **1998**, *102*, 9109−9118.
(34) Šponer, J.; Leszczyñski, J.; Hobza, P. *J. Comput. Chem.* **1996**, *17*, 841−850.
(35) Estrin, D. A.; Paglieri, L.; Corongiu, G. *J. Phys. Chem.* **1994**, *98*, 5653−5660.
(36) Scanlan, M. J.; Hillier, I. H. *J. Am. Chem. Soc.* **1984**, *106*, 3737−3745.
(37) Chandra, A. K.; Nguyen, M. T.; Zeegers-Huyskens, T. *J. Mol. Struct.* **2000**, *519*, 1−11.
(38) Chandra, A. K.; Michalska, D.; Wysokinsky, R.; Zeegers-Huyskens, T. *J. Phys. Chem. A* **2004**, *108*, 9593−9600.
(39) Broo, A.; Holmén, A. *J. Phys. Chem. A* **1997**, *101*, 3589−3600.

(40) Monajjemi, M.; Ghiasi, R.; Abedi, A. *Theor. Inorg. Chem.* **2005**, *50*, 435−441.
(41) Burda, J.; Šponer, J.; Leszczyñski, J.; Hobza, P. *J. Phys. Chem. B* **1997**, *101*, 9670−9677.
(42) Šponer, J.; Burda, J. V.; Sabat, M.; Leszczyñski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951−5957.
(43) Leœ, A.; Adamowicz, L.; Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 4001−4005.
(44) Alemán, C. *Chem. Phys.* **2000**, *253*, 13−19.
(45) Civcir, P. Ü. *THEOCHEM* **2000**, *532*, 157−169.
(46) Gould, I. R.; Green, D. V. S.; Young, P.; Hillier, I. H. *J. Org. Chem.* **1992**, *57*, 4434−4437.
(47) Morpurgo, S.; Bossa, M.; Morpurgo, G. O. *Adv. Quantum Chem.* **2000**, *36*, 169.
(48) Sambrano, J. R.; Souza, A. R.; Queralt, J. J.; Andrés, J. *Chem. Phys. Lett.* **2000**, *317*, 437−443.
(49) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811−6821.
(50) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285.
(51) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.
(52) Sobolewski, A. L.; Adamowicz, L. *J. Chem. Phys.* **1995**, *102*, 5708.
(53) Fogarasi, G.; Szalay, P. G. *Chem. Phys. Lett.* **2002**, *356*, 383.
(54) Prado, M. A. S.; Garcia, E.; Martins, J. B. L. *Chem. Phys. Lett.* **2006**, *418*, 264−267.
(55) Leszczyñski, J. *J. Phys. Chem.* **1992**, *96*, 1649−1653.
(56) Glaser, R.; Rayat, S.; Lewis, M.; Son, M.-S.; Meyer, S. *J. Am. Chem. Soc.* **1999**, *121*, 6108−6119.
(57) Lindahl, T.; Nyberg, B. *Biochemistry* **1974**, *13*, 3405−3410.
(58) Duncan, B. K.; Miller, J. H. *Nature* **1980**, *287*, 560−561.
(59) Flinn, C.; Poirier, R. A.; Sokalski, W. A. *J. Phys. Chem. A* **2003**, *107*, 11174−11181.
(60) (a) Poirier, R. A. *MUNgauss* (Fortran 90 version); Chemistry Department, Memorial University of Newfoundland, St. John's, Canada. With contributions from Brooker, M.; Bungay, S. D.; El-Sherbiny, A.; Gosse, T.; Keefe, D.; Pye, C. C.; Reid, D.; Shaw, M.; Wang, Y.; Xidos, J. (b) Colonna, F.; Jolly, L.-H.; Poirier, R. A.; Angyan, J.; Jansen, G. *Comput. Phys. Commun.* **1994**, *81*, 293−317.
(61) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T., Jr.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.05; Gaussian, Inc.: Pittsburgh, PA, 2003.
(62) Almatarneh, M. H.; Flinn, C. G.; Poirier, R. A. *Can. J. Chem.* **2005**, *83*, 2082−2090.