Ab Initio Study of the Deamination of Formamidine

Christopher 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 & Theoretical Chemistry I-30, Wroclaw University of Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland

Received: June 12, 2003; In Final Form: September 15, 2003

The mechanism for the deamination of formamidine with OH^-, H_2O, and H_3O^+ was investigated using computational techniques. Optimized geometries for reactants, intermediates, products, and transition states were determined at the RHF/6-31G(d) and MP2/6-31G(d) levels of theory. Single-point energies were determined at G1 and G2 and using density functional theory (B3LYP) at both optimized geometries and the 6-311G(d,p), 6-311+G(d,p), 6-311++G(d,p) (with OH^- only), 6-311G(2df,p), and 6-311+G(3df,2p) basis sets. Frequencies were calculated for all structures studied at both geometries. Intrinsic reaction coordinate analysis was carried out for all transition states. Activation energies were calculated for each pathway investigated. Deamination with H_3O^+ involves the formation of the same tetrahedral intermediate from the addition of water to either possible protonated formamidine cation followed by a 1,3-proton shift and dissociation to products. Deamination is an unlikely event because of the high activation-energy barrier to the formation of the intermediate (242.4 kJ mol^{-1} for pathway A and 243.5 kJ mol^{-1} for pathway B using G2 data). Two pathways for deamination with H_2O were studied. Pathway A produced a tetrahedral intermediate by the addition of water followed by a 1,3-proton shift and dissociation to products, and pathway B involved the formation of the tautomer of the formamide product by reaction with water followed by a 1,3-proton shift to yield formamide. Deamination by either pathway is unlikely because of the high activation-energy barriers involved (212.7 kJ mol^{-1} for the formation of the tetrahedral intermediate of pathway A and 249.2 kJ mol^{-1} for the formation of the formamide tautomer of pathway B using G2 data). Deamination with OH^- yields a tetrahedral intermediate with a much lower activation-energy barrier than for the reaction with H_3O^+ and H_2O (58.2 kJ mol^{-1} using G2 data), but the deprotonation of formamidine by hydroxide appears to be the most likely process because of the resonance-stabilized deprotonated formamidine anion/water complex formed ($\Delta E = -163.2$ kJ mol^{-1}).

1. Introduction

Amidines are an important class of biologically active molecules. The amidine group consists of a N1−C−N2 framework with possible substituents R1, R2, R3, and R4 as shown in Figure 1. Several important properties of multiply substituted noncyclic amidines were investigated by Kosturkiewicz et al.[1] on the basis of X-ray crystallographic structures that they had solved as well as others reported in the literature. The properties and reactions of amidines are of general interest because of the structural relationship of some amidines to molecules of biological interest (e.g., penicillins[2]) and to the DNA bases cytosine, guanine, and adenine as well as their noted biological activity (e.g., benzamidine derivatives are inhibitors of serine proteases[3] and an amidine analogue of glucose is an inhibitor of glycosidases[4]). Cyclic amidines are known to be anthelmintic and serve as antibacterial agents.[5] Amidines have been reported to have antibiotic, antifungal, and anaesthetic properties.[6−8] Terfort et al.[9] have pointed out that the amidinium functionality is a recognition element in a self-replication system.

A study of the deamination of the more stable E isomer of the simplest amidine (formamidine) with OH^-, H_2O, and H_3O^+ to yield formamide as a product was undertaken as a precursor

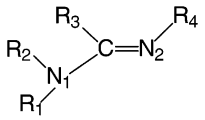

Figure 1. General structure of amidines.

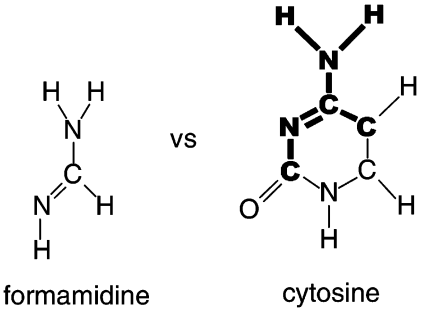

Figure 2. Comparison of formamidine (E isomer) and cytosine structures.

to a study of the deamination of the DNA base cytosine. Cytosine is the most unstable of the DNA bases, deaminating spontaneously to uracil with an activation energy of 117 ± 4 kJ mol^{-1}.[10] The deamination of 5-methylcytosine leads to thymine and G/C to T/A transitions dominating the spectra of spontaneous mutations for primates.[11] As can be seen in Figure 2,

* Corresponding author. E-mail: cgflinn@mun.ca. Tel: (709) 737-7911. Fax: (709) 737-3702.

10.1021/jp035677y CCC: $25.00 © 2003 American Chemical Society
Published on Web 11/20/2003

the E(trans) isomer of formamidine forms part of the cytosine structure, notably in the region of cytosine where the deamination of cytosine to uracil can take place.

DNA bases such as cytosine are bonded to the phosphate groups of the DNA framework and are hydrogen bonded to a matching base pair on the complementary strand of the double helix. The gas-phase results are significant because cytosine in DNA would have an environment more closely related to the gas phase than an aqueous solution. Dixon et al.[12] investigated solvent effects on the intramolecular conversion of cyclic amidines and imidates by means of the self-consistent reaction field (SCRF) procedure. The size of the solvation correction was small and stated that their major conclusions were unlikely to be changed by using a more accurate procedure. For these reasons, an investigation into solvent effects was not carried out.

The hydrolysis of cyclic amidines and cyclic amidinium ions[13] has been shown to result in two possible distinct reactions. Ring cleavage produces amino amides, and cleavage of the exocyclic C−N bond produces lactams. Much debate has occurred in the literature as to the presence or absence of Deslongschamps' hypothesis of stereoelectronic control[14] that is sometimes known as the "antiperiplanar lone-pair hypothesis",[15] which states that the preferential cleavage of a tetrahedral intermediate occurs when there are two lone pairs antiperiplanar to the leaving group. An alternate explanation proposed by Perrin et al.[16] concludes that the products observed from the hydrolyses of acetals, amides, and imidates can be explained on the basis of lactone instability, the instability of anti (E) isomers, or steric effects.

Tautomerization by intramolecular proton transfer involving formamidine and formamidine complexes with one to three water molecules was studied by Zhang et al.[17] using ab initio and density functional methods. Their results indicated the importance of hydrogen-bonded water molecules in lowering the barrier to tautomerization. Nguyen et al.[18] investigated double proton transfer in formamidine dimers using ab initio methods including single-point correlation corrections up to fourth order (MP4) many-body perturbation theory. The potential energy surface (PES) for double proton transfer involving formamidine dimers was studied by Lim et al.[19] using Hartree−Fock and density functional theory. The solvent effect on the PES was included in their study using the Onsager self-consistent reaction field model. Solvent effects on proton transfer involving formamidine hydrogen bonded to a water molecule were investigated by Nagaoka et al.[20] using the chemical reaction molecular dynamics simulation method. A recent systematic analysis of optimal catalytic fields for the tautomerization of a formamidine−formamide complex mimicking an A/T base pair by Dziekonski et al. revealed the dominant electrostatic nature of environmental effects facilitating double proton tranfer.[21]

The unimolecular decomposition mechanism of formamidine to give ammonia and hydrogen cyanide has been investigated by Andrés et al.[22] at the Hartree−Fock level of theory with different basis sets. A potential energy hypersurface was established, and the stationary points representing the reactants, products, and transition states were located. The first step was the tautomerization of the E(endo) to Z(exo) isomer followed by a 1,3 shift of the hydrogen bonded to the imino nitrogen to the amino nitrogen accompanied by the cleavage of the bond between the carbon and the amino nitrogen.

Tortajada[23] et al. investigated the structure of protonated formamidine isomers and the complexes of formamidine with several monovalent cations. They found that complexes with the less stable Z isomer of formamidine were found at the global minima versus those of the E isomer and complexes where the metal cation bridges both basic nitrogens.

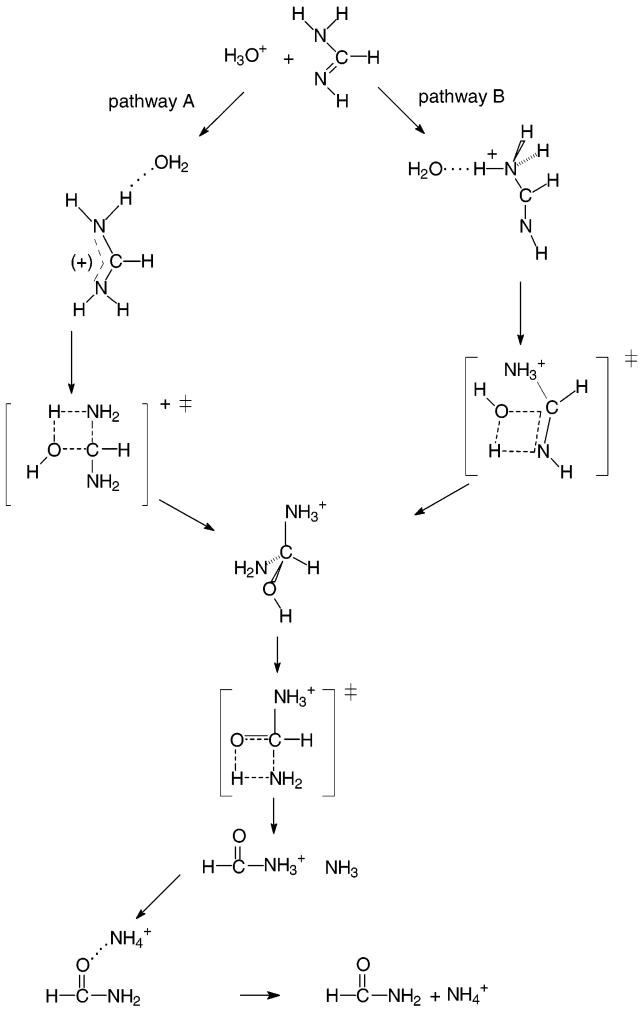

Figure 3. Deamination mechanisms for pathways A and B with H_3O^+.

No computational studies of the deamination of formamidine or any other related amidines have been reported. For this reason, we have performed a detailed study of possible mechanisms for the deamination of formamidine with OH^-, H_2O, and H_3O^+. The optimized formamidine transition states would likely be very useful in producing good initial guesses for the cytosine transition states. One of the main objectives of the formamidine study was to select the lowest level of theory/basis set that would give an acceptable level of confidence for the energetics of deamination reactions that involve many intermediates and transitions states including neutral, positively, and negatively charged species.

2. Method

Optimized geometries for reactants, transition states, intermediates, and products were determined at RHF/6-31G(d) and MP2(FULL)/6-31G(d). Single-point energies required for Gaussian G1 and G2 were obtained using the MP2(FULL)/6-31G(d) optimized geometry. B3LYP single-point calculations were also performed, using several different basis sets, at both the RHF/6-31G(d) and MP2(FULL)/6-31G(d) optimized geometries.

The MUNgauss[24] computational program was used for most of the geometry optimizations at RHF/6-31G(d) for reactants, products, intermediates, and transition-state structures. Geometry optimizations at MP2(FULL)/6-31G(d), frequency calculations,

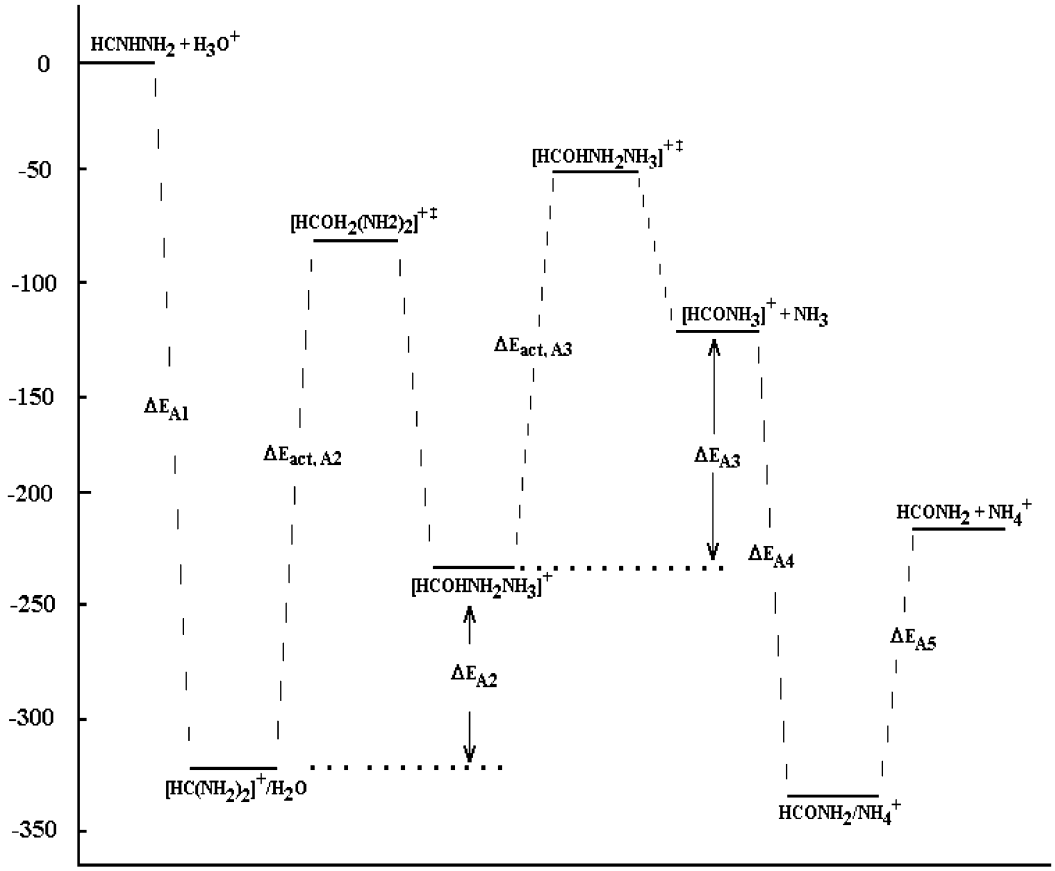

Figure 4. Deamination of formamidine with H_3O^+, pathway A. Relative energy in kJ mol^{-1} vs progress of reaction using G2 theory.

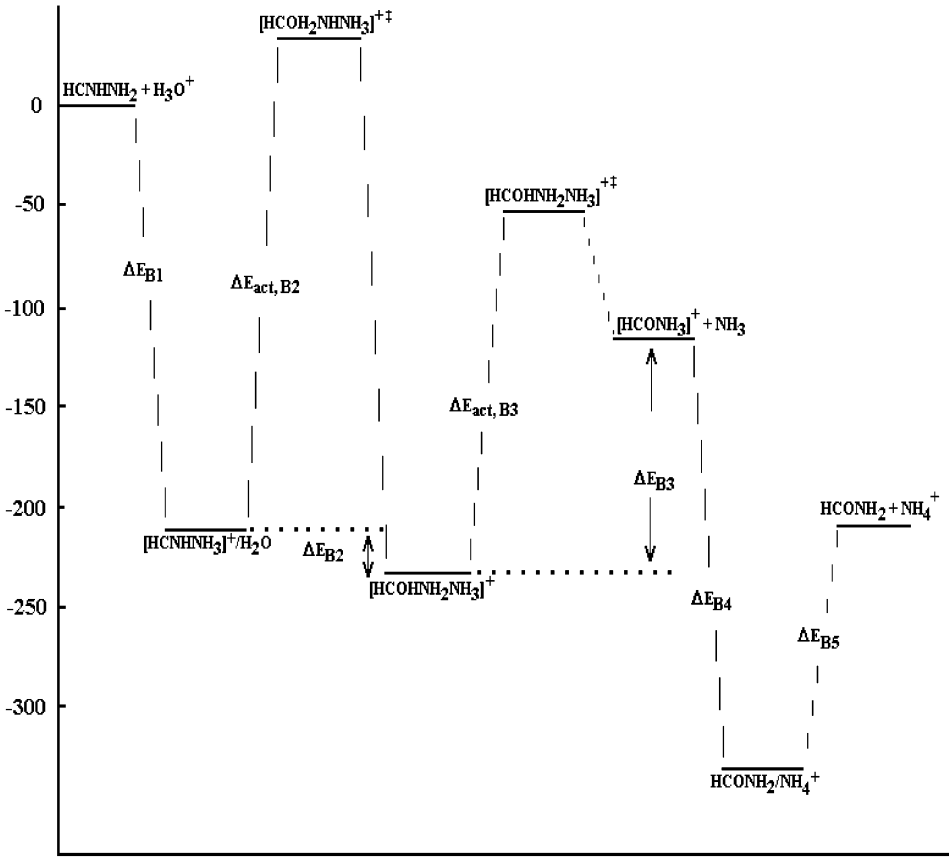

Figure 5. Deamination of formamidine with H_3O^+, pathway B. Relative energy in kJ mol^{-1} vs progress of reaction using G2 theory.

single-point energies for G1 and G2 theory, intrinsic reaction coordinate (IRC) analyses of all transition states, and all single-point density functional theory (DFT) calculations using B3LYP were performed with either Gaussian 92[25] or Gaussian 94.[26]

TABLE 1: Energetics as a Function of the Level of Theory and Basis Set for the Deamination of Formamidine with H_3O^+ in kJ mol^{-1}, Pathway A (Figures 3 and 4)[a]

theory/basis set	ΔE_{A1}	$\Delta E_{act,A2}$	ΔE_{A2}	$\Delta E_{act,A3}$	ΔE_{A3}	ΔE_{A4}	ΔE_{A5}
RHF/6-31G(d)[b]	−357.6	229.1	82.3	247.3	132.3	−233.3	**124.0**
MP2(FU)/6-31G(d)[c]	−344.9	**246.7**	72.7	**183.5**	130.0	−235.7	131.5
MP2/6-311G(d,p)[c]	−330.1	**241.9**	80.5	187.0	130.0	−224.3	131.6
MP2/6-311+G(d,p)[c]	**−325.7**	**248.8**	**84.8**	188.3	125.8	−220.9	**125.1**
MP2/6-311G(2df,p)[c]	−333.7	**238.9**	79.4	180.9	135.9	−233.8	133.7
MP2/6-311+G(3df,2p)[c]	−329.4	**243.3**	80.3	184.5	132.2	−227.2	**128.1**
MP4SDTQ/6-311G(d,p)[c]	−328.1	**241.9**	78.6	185.2	119.5	**−216.5**	130.3
MP4SDTQ/6-311+G(d,p)[c]	**−324.4**	**248.3**	82.4	186.5	**116.4**	**-213.7**	**124.0**
MP4SDTQ/6-311G(2df,p)[c]	−331.9	**240.4**	78.3	**179.8**	124.9	−226.2	132.1
QCISD(T)/6-311G(d,p)[c]	−329.6	**244.1**	79.8	189.8	125.1	**−217.9**	131.3
B3LYP/6-311G(d,p)[b]	−342.8	258.9	**99.3**	184.4	128.2	−242.2	135.2
B3LYP/6-311+G(d,p)[b]	−337.7	265.8	104.0	185.2	120.1	−236.0	**129.0**
B3LYP/6-311+G(3df,2p)[b]	−334.3	262.5	101.8	**183.2**	**116.6**	−232.1	**128.7**
B3LYP/6-311G(d,p)[c]	−345.8	259.9	100.6	**181.6**	128.0	−245.3	138.4
B3LYP/6-311+G(d,p)[c]	−340.2	266.7	104.9	**183.2**	120.1	−222.6	**115.2**
B3LYP/6-311+G(3df,2p)[c]	−336.6	262.4	102.5	**181.7**	117.0	−234.5	130.9
G1	−319.2	244.9	94.6	173.5	107.8	−214.3	120.0
G2	−319.2	242.4	91.1	175.7	108.2	−211.0	120.9
standard deviation	9.8	10.8	10.9	15.6	8.1	10.1	5.8
standard deviation (no RHF)	8.1	9.8	11.1	4.2	8.0	10.3	5.9

[a] Bold values are within 8.4 kJ mol^{-1} of the G2 values, not including G1. [b] Energies calculated at RHF/6-31G(d) optimized geometries. [c] Energies calculated at MP2(FU)/6-31G(d) optimized geometries.

TABLE 2: Energetics as a Function of the Level of Theory and Basis Set for the Deamination of Formamidine with H_3O^+ in kJ mol^{-1}, Pathway B (Figures 3 and 5)[a]

theory/basis set	ΔE_{B1}	$\Delta E_{act,B2}$	ΔE_{B2}	$\Delta E_{act,B3}$	ΔE_{B3}	ΔE_{B4}	ΔE_{B5}
RHF/6-31G(d)[b]	−230.1	313.9	−45.2	247.3	132.3	−233.3	**124.0**
MP2(FU)/6-31G(d)[c]	−238.5	**248.7**	−33.6	**183.5**	130.0	−235.7	131.5
MP2/6-311G(d,p)[c]	−226.1	**251.1**	−23.5	187.0	130.0	−224.3	131.6
MP2/6-311+G(d,p)[c]	−220.9	255.0	**−20.0**	188.3	125.8	−220.9	**125.1**
MP2/6-311G(2df,p)[c]	−220.7	**236.4**	−33.6	**180.9**	135.9	−233.8	133.7
MP2/6-311+G(3df,2p)[c]	**−215.6**	**239.2**	−33.5	184.5	132.2	−227.2	**128.1**
MP4SDTQ/6-311G(d,p)[c]	−229.2	255.4	**−20.3**	185.2	119.5	**−216.5**	130.3
MP4SDTQ/6-311+G(d,p)[c]	−224.7	258.9	**−17.4**	186.5	**116.4**	**−213.7**	**124.0**
MP4SDTQ/6-311G(2df,p)[c]	−223.6	**241.8**	−30.0	**179.8**	124.9	−226.2	132.1
QCISD(T)/6-311G(d,p)[c]	−230.1	260.4	**−19.7**	189.8	125.1	**−217.9**	131.3
B3LYP/6-311G(d,p)[b]	−220.1	**250.2**	−23.5	184.4	128.2	−242.2	135.2
B3LYP/6-311+G(d,p)[b]	**−215.7**	256.6	**−18.1**	185.2	120.1	−236.0	**129.0**
B3LYP/6-311+G(3df,2p)[b]	**−210.0**	**251.5**	**−23.0**	**183.2**	**116.6**	−232.1	**128.7**
B3LYP/6-311G(d,p)[c]	−222.8	**247.0**	**−22.4**	**181.6**	128.0	−245.3	138.4
B3LYP/6-311+G(d,p)[c]	**−217.5**	253.3	**−17.8**	**183.2**	120.1	−222.6	**115.2**
B3LYP/6-311+G(3df,2p)[c]	**−211.1**	**247.6**	**−23.0**	**181.7**	117.0	−234.5	130.9
G1	−209.8	244.6	−14.8	173.5	107.8	−214.3	120.0
G2	−209.9	243.5	−18.1	175.7	108.2	−211.0	120.9
standard deviation	8.1	16.6	7.9	15.6	8.1	10.1	5.8
standard deviation (no RHF)	8.1	6.9	6.1	4.2	8.0	10.3	5.9

[a] Bold values are within 8.4 kJ mol^{-1} of the G2 values, not including G1. [b] Energies calculated at RHF/6-31G(d) optimized geometries. [c] Energies calculated at MP2(FU)/6-31G(d) optimized geometries.

Basis sets used with DFT B3LYP calculations were 6-311G(d,p), 6-311+G(d,p), 6-311++G(d,p), and 6-311+G(3df,2p).

3. Results and Discussion

3.1. Choice of Theory/Basis Set. The relative energies obtained at the different levels of theory for each of the mechanisms investigated are given in Tables 1 to 5. Standard deviations for each calculated energy change are also included in Tables 1 to 5. The standard deviations range from 4.1 to 55.0 and 3.6 to 56.5 when the RHF/6-31G(d) results are excluded. In general, the standard deviations are within 10 kJ mol^{-1} except for the activation energies and for the reaction of formamidine with hydroxide, where much larger standard deviations are obtained because of the importance of diffuse functions. The effect of a second diffuse function on the reaction energetics was investigated for the deamination of formamidine with OH^-. It showed no significant gains versus the G1 and G2 results when compared to a single diffuse function. Only minimal gains resulted from using MP2(FULL)/6-31G(d) geometries and either G1 or G2 theory or a more complete basis set, 6-311+G(3df,2p), with DFT.

3.2. Deamination with H_3O^+. Two distinct cations can be formed by proton transfer from the hydronium ion, leading to two possible pathways for deamination. In this paper, the mechanisms are designated as pathway A (steps A1 to A5) and pathway B (steps B1 to B5). The structures of reactants, intermediates, transition states, and products involved in both mechanisms are shown in Figure 3. The relative energies of reactants, intermediates, transition states, and products for pathways A and B are shown in Figures 4 and 5, respectively. Internal-energy changes and activation-energy values using various basis sets at Hartree−Fock, MP2, MP4, G1, and G2 levels of theory for pathways A and B are given in Tables 1 and 2, respectively.

In pathway A, formamidine is protonated at its sp^2 (imino) nitrogen by the hydronium ion forming a hydrogen-bonded complex with the resulting water molecule. This process is very exothermic because of the significant delocalization of the positive charge. In pathway B, protonation occurs at the sp^3 nitrogen of the amino group, giving a different formamidine cation/H_2O complex. This complex is less stable than the former because there is little delocalization of the positive charge. In

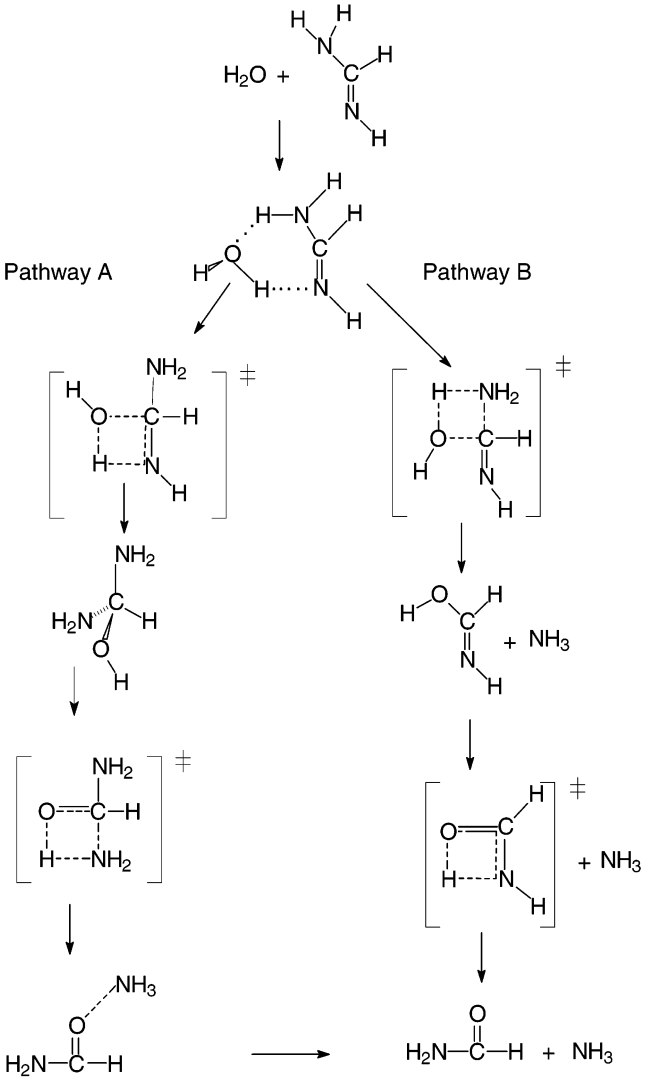

Figure 6. Deamination mechanisms for pathways A and B with H_2O.

each case, nucleophilic attack by the water molecule on the carbon atom results in the same intermediate cation. This was confirmed by intrinsic reaction coordinate (IRC) analysis of the two transition states. In pathway A, the transfer of hydrogen from water can occur to either NH_2 to form the intermediate. In pathway B, hydrogen is transferred to the imino nitrogen. This step is followed by an internal proton transfer from the hydroxyl group to the amino group, resulting in deamination. Protonated formamide and ammonia are the products. Proton transfer between protonated formamide and ammonia, giving a formamide/ammonium ion complex, is very exothermic because of the involvement of the nitrogen 2p orbital in the π bonding in planar formamide.

3.3. Deamination with H_2O. In this case, deamination can follow one of two possible pathways designated as pathway A (steps A1 to A4) and pathway B (steps B1 to B4). The structures of reactants, intermediates, transition states, and products involved in both mechanisms are shown in Figure 6. The relative energies of reactants, intermediates, transition states, and products for pathways A and B are shown in Figures 7 and 8, respectively. Internal-energy changes and activation-energy values using various basis sets at Hartree−Fock, MP2, MP4, G1, and G2 levels of theory for pathways A and B are given in Tables 3 and 4, respectively.

Formamidine initially forms a hydrogen-bonded complex with a water molecule. In pathway A, nucleophilic attack by the water molecule on the carbon atom and simultaneous proton transfer from H_2O to the sp^2 nitrogen of formamidine results in a tetrahedral intermediate. This is followed by an internal proton transfer from the hydroxyl group to either amino group, resulting in deamination to yield a formamide/ammonia complex. Both the intermediate and the transition state can be one of two isomeric forms due to the partial double-bond character of the C−N2 bond. The hydrogen atom bonded to N2 can be cis or trans to the hydrogen bonded to the carbon. In pathway B, nucleophilic attack by the water molecule on the carbon atom and simultaneous proton transfer from H_2O to the amino group of formamidine produce the formamide tautomer and ammonia. Internal proton transfer from the hydroxyl group to the sp^2 nitrogen of the tautomer results in the deamination product, formamide.

TABLE 3: Energetics as a Function of the Level of Theory and Basis Set for the Deamination of Formamidine with H_2O in kJ mol^{-1}, Pathway A (Figures 6 and 7)[a]

theory/basis set	ΔE_{A1}	$\Delta E_{act,A2}$	ΔE_{A2}	$\Delta E_{act,A3}$	ΔE_{A3}	ΔE_{A4}
RHF/6-31G(d)[b]	−42.0	267.0	−6.4	231.0	**−59.8**	**34.0**
MP2(FU)/6-31G(d)[c]	−58.3	234.7	7.9	177.4	**−66.4**	48.6
MP2/6-311G(d,p) [c]	−54.7	232.6	**12.7**	179.2	**−52.7**	45.0
MP2/6-311+G(d,p)[c]	−42.2	228.5	**13.9**	176.9	−47.1	36.3
MP2/6-311G(2df,p) [c]	−58.1	225.7	**13.2**	169.6	**−59.1**	46.3
MP2/6-311+G(3df,2p)[c]	−44.9	221.9	**13.8**	170.1	−50.5	36.0
MP4SDTQ/6-311G(d,p)[c]	−52.7	229.6	10.9	180.0	**−55.5**	43.9
MP4SDTQ/6-311+G(d,p)[c]	**−40.9**	224.3	11.2	177.5	−49.1	35.8
MP4SDTQ/6-311G(2df,p)[c]	−55.8	223.5	12.0	170.8	**−63.0**	45.0
QCISD(T)/6-311G(d,p)[c]	−52.1	232.8	10.4	183.7	−50.2	43.7
B3LYP/6-311G(d,p)[b]	−56.7	**216.6**	**26.2**	**165.9**	−74.4	43.8
B3LYP/6-311+G(d,p)[b]	−42.1	202.6	**28.5**	**165.7**	−69.1	**32.7**
B3LYP/6-311+G(3df,2p)[b]	**−38.9**	**205.6**	**28.7**	**164.0**	−69.5	**29.2**
B3LYP/6-311G(d,p)[c]	−59.2	**219.3**	**28.1**	**165.5**	−75.9	45.5
B3LYP/6-311+G(d,p)[c]	−43.0	**218.8**	**29.3**	**164.9**	−70.5	**33.5**
B3LYP/6-311+G(3df,2p)[c]	**−39.9**	**217.7**	**29.0**	**163.3**	−69.9	**29.6**
G1	−33.3	212.5	21.7	156.0	−62.3	27.5
G2	−32.7	212.7	21.1	158.9	−59.2	25.9
standard deviation	8.8	14.1	9.8	16.3	9.1	7.3
standard deviation (no RHF)	9.0	9.3	8.0	8.0	9.4	7.5

[a] Bold values are within 8.4 kJ mol^{-1} of the G2 values, not including G1. [b] Energies calculated at RHF/6-31G(d) optimized geometries. [c] Energies calculated at MP2(FU)/6-31G(d) optimized geometries.

TABLE 4: Energetics as a Function of the Level of Theory and Basis Set for the Deamination of Formamidine with H_2O in kJ mol^{-1}, Pathway B (Figures 6 and 8)[a]

theory/basis set	ΔE_{B1}	$\Delta E_{act,B2}$	ΔE_{B2}	$\Delta E_{act,B3}$	ΔE_{B3}
RHF/6-31G(d)[b]	−42.0	308.6	27.4	206.6	−59.6
MP2(FU)/6-31G(d)[c]	−58.3	**255.4**	47.6	**144.7**	**−57.5**
MP2/6-311G(d,p)[c]	−54.7	259.4	52.4	148.3	**−47.5**
MP2/6-311+G(d,p)[c]	−42.2	**251.8**	51.1	149.9	**−48.1**
MP2/6-311G(2df,p)[c]	−58.1	**252.5**	50.6	**140.4**	**−50.2**
MP2/6-311+G(3df,2p)[c]	−44.9	**245.8**	50.1	**143.4**	**−50.8**
MP4SDTQ/6-311G(d,p)[c]	−52.7	**257.1**	46.2	149.0	**−46.9**
MP4SDTQ/6-311+G(d,p)[c]	**−40.9**	**248.2**	**45.5**	151.0	**−47.6**
MP4SDTQ/6-311G(2df,p)[c]	−55.8	**251.0**	**43.9**	142.5	**−49.9**
QCISD(T)/6-311G(d,p)[c]	−52.1	264.1	46.7	155.2	**−42.9**
B3LYP/6-311G(d,p)[b]	−56.7	**252.1**	49.8	146.5	**−54.2**
B3LYP/6-311+G(d,p)[b]	−42.1	**245.5**	46.3	149.2	**−54.3**
B3LYP/6-311+G(3df,2p)[b]	**−38.9**	**247.2**	**41.0**	148.9	**−52.7**
B3LYP/6-311G(d,p)[c]	−59.2	**253.0**	51.7	**145.7**	**−54.0**
B3LYP/6-311+G(d,p)[c]	−43.0	**250.3**	47.0	148.1	**−54.7**
B3LYP/6-311+G(3df,2p)[c]	**−39.9**	**250.9**	**41.6**	147.7	**−53.0**
G1	−33.4	248.3	36.5	136.4	−49.5
G2	−32.8	249.2	37.4	137.8	−49.6
standard deviation	8.8	14.2	6.4	15.0	4.1
standard deviation (no RHF)	9.0	4.9	4.8	4.9	3.6

[a] Bold values are within 8.4 kJ mol^{-1} of the G2 values, not including G1. [b] Energies calculated at RHF/6-31G(d) optimized geometries. [c] Energies calculated at MP2(FU)/6-31G(d) optimized geometries.

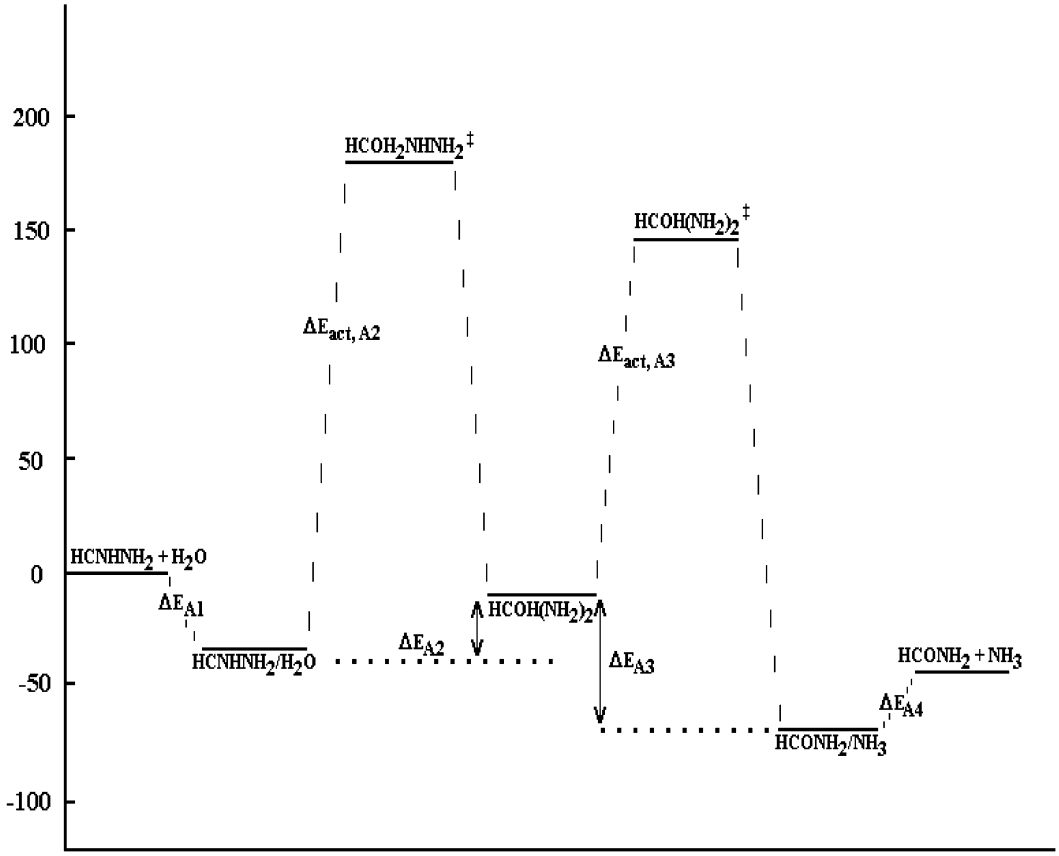

Figure 7. Deamination of formamidine with H_2O, pathway A. Relative energy in kJ mol^{-1} vs progress of reaction using G2 theory.

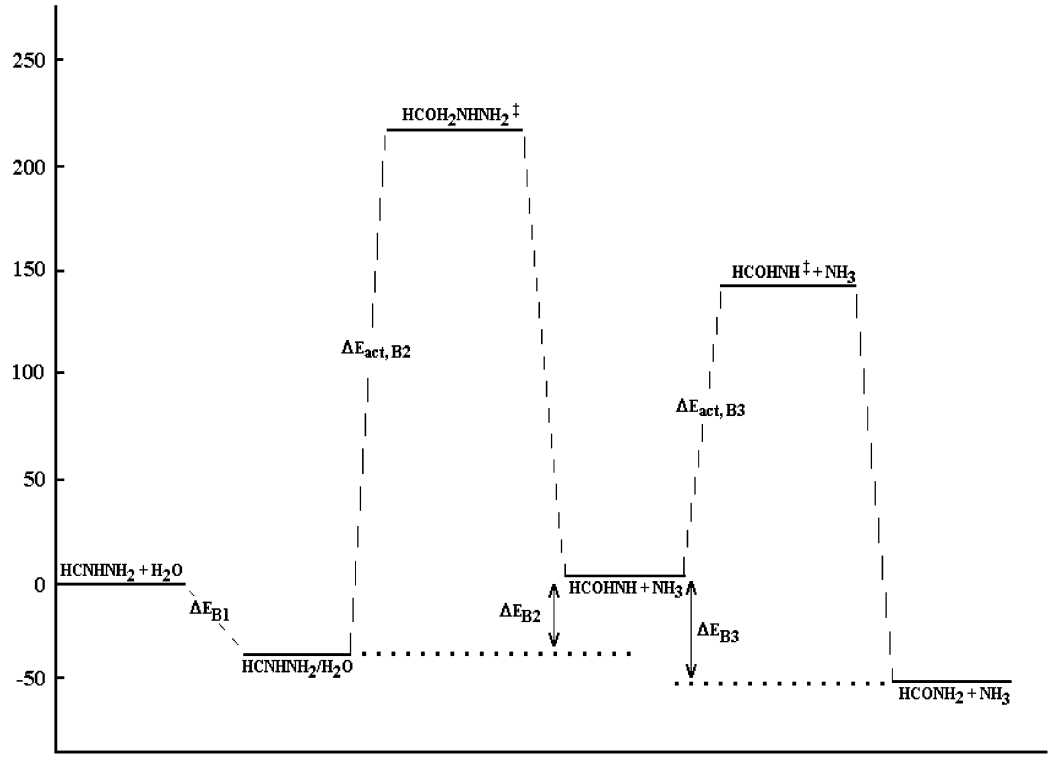

Figure 8. Deamination of formamidine with H_2O, pathway B. Relative energy in kJ mol^{-1} vs progress of reaction using G2 theory.

3.4. Deamination with OH^-. It is important to consider the deprotonation of formamidine by the hydroxide ion as well as deamination. Structures involved in deprotonation and deamination mechanisms are shown in Figure 9. The relative energies of reactants, intermediates, transition states, and products for deprotonation and deamination are shown in Figure 10. Internal-energy changes and activation-energy values using various basis sets at Hartree−Fock, MP2, MP4, G1, and G2 levels of theory for deprotonation and deamination are given in Table 5.

The deamination pathway with OH^- has a significantly lower activation-energy barrier than deamination with H_3O^+ or H_2O. However, the deprotonation of formamidine by the hydroxide

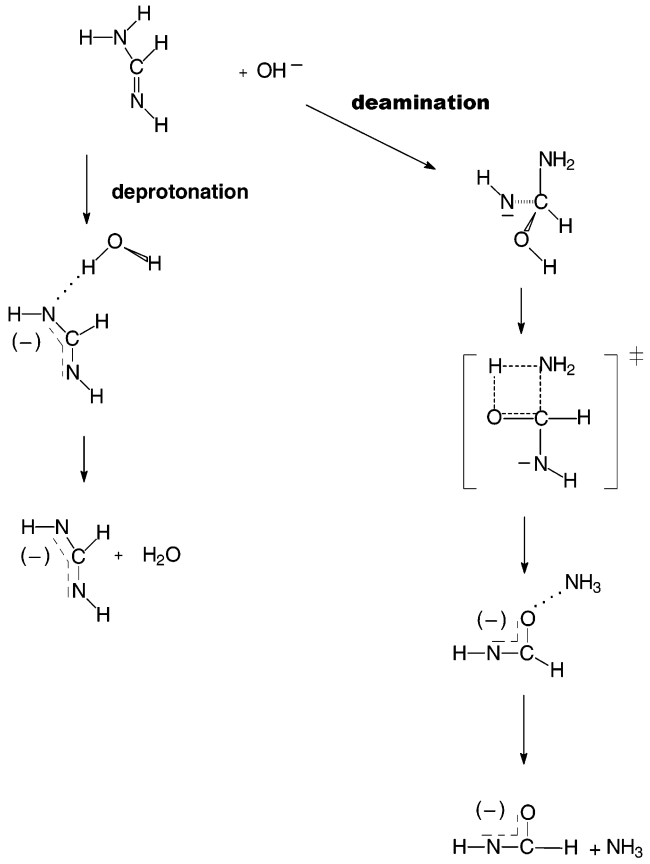

Figure 9. Deamination and deprotonation mechanisms with OH^-.

ion is very exothermic and hence much more favorable. This is because of the delocalization of the negative charge in the formamidine anion/H_2O complex formed. An initial formamidine/OH^- complex was not found because of the ease of deprotonation of formamidine by OH^-. The addition of OH^- to amidine resulted in a tetrahedral intermediate anion. Deamination occurs by an internal proton transfer from the hydroxyl group to the amino group to yield a hydrogen-bonded complex of the formamide anion and ammonia. Both the intermediate and the transition state can be one of two isomeric forms due to the partial double-bond character of the C−N2 bond. The hydrogen atom bonded to N2 can be cis or trans to the hydrogen atom bonded to C. The relative energies of reactants, intermediates, transition states, and products for deamination and deprotonation are shown in Figure 7.

TABLE 5: Energetics as a Function of the Level of Theory and Basis Set for the Deamination and Deprotonation of Formamidine with OH^- in kJ mol^{-1} (Figures 9 and 10)[a]

theory/basis set	ΔE_{A1}	$\Delta E_{act,A2}$	ΔE_{A2}	ΔE_{A3}	ΔE_{B1}	ΔE_{B2}
RHF/6-31G(d)[b]	−114.0	87.4	−172.4	43.7	−232.3	81.6
MP2(FU)/6-31G(d)[c]	−151.7	77.2	−163.1	57.7	−274.3	103.6
MP2/6-311G(d,p)[c]	−165.6	**74.1**	**−140.8**	56.1	−285.8	100.2
MP2/6-311+G(d,p)[c]	−67.7	75.8	−124.8	**39.4**	−179.0	83.1
MP2/6-311G(2df,p)[c]	−168.8	**70.9**	−147.0	55.8	−286.4	100.4
MP2/6-311+G(3df,2p)[c]	−70.0	76.8	−128.6	**40.9**	−175.4	81.0
MP4SDTQ/6-311G(d,p)[c]	−165.8	**70.2**	−146.7	56.2	−285.5	98.5
MP4SDTQ/6-311+G(d,p)[c]	−69.1	**72.9**	**−129.4**	**39.9**	−178.2	82.4
MP4SDTQ/6-311G(2df,p)[c]	−167.2	**66.4**	−154.0	55.5	−285.6	98.5
QCISD(T)/6-311G(d,p)[c]	−163.1	**71.5**	**−142.8**	55.4	−282.3	91.8
B3LYP/6-311G(d,p)[b]	−147.7	57.1	−159.1	54.2	−272.0	95.9
B3LYP/6-311+G(d,p)[b]	**−48.0**	**60.8**	**−145.1**	**37.0**	**−168.4**	80.7
B3LYP/6-311++G(d,p)[b]	**−48.9**	**60.7**	**−144.9**	**37.4**	**−168.5**	80.8
B3LYP/6-311+G(3df,2p)[b]	**−48.1**	**59.5**	−149.4	**35.9**	**−167.6**	**76.0**
B3LYP/6-311G(d,p)[c]	−151.6	57.0	−157.9	55.9	−274.7	97.7
B3LYP/6-311+G(d,p)[c]	**−48.8**	**59.7**	**−145.7**	**37.0**	**−168.7**	80.0
B3LYP/6-311+G(3df,2p)[c]	**−49.0**	58.6	−150.3	36.1	−171.1	**78.6**
G1	−53.3	63.1	−140.0	32.3	−167.5	73.0
G2	−52.4	67.2	−137.6	34.1	−163.2	70.6
standard deviation	52.8	8.5	11.9	9.6	55.0	10.5
standard deviation (no RHF)	54.2	7.2	10.4	9.8	56.5	10.7

[a] Bold values are within 8.4 kJ mol^{-1} of the G2 values, not including G1. [b] Energies calculated at RHF/6-31G(d) optimized geometries. [c] Energies calculated at MP2(FU)/6-31G(d) optimized geometries.

4. Conclusions

The stability of formamidine toward deamination is apparent from this investigation. The deprotonation of formamidine by OH^- is very exothermic because the anion formed by the loss of a proton is highly resonance-stabilized. The formation of the

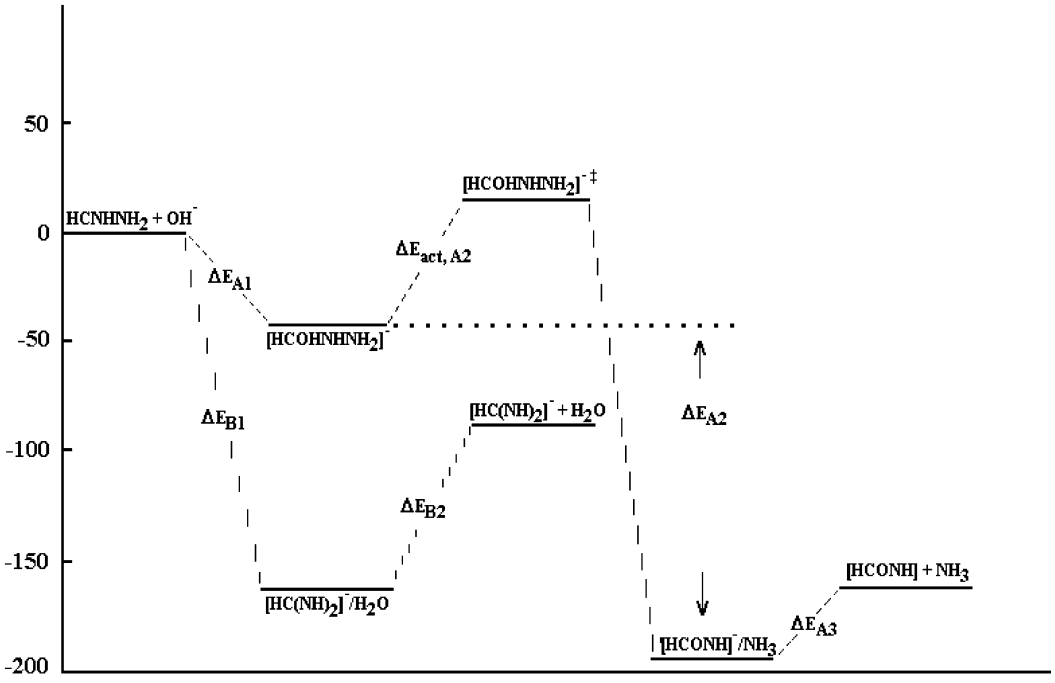

Figure 10. Deamination and deprotonation of formamidine with OH^-. Relative energy in kJ mol^{-1} vs progress of reaction using G2 theory.

tetrahedral intermediate anion is a much less exothermic process and is less likely to occur than deprotonation. There is a high activation-energy barrier for the addition of H_2O to the neutral or protonated formamidine/H_2O complexes and a similar high activation-energy barrier for the subsequent 1,3-proton shift that leads to the decomposition of the tetrahedral intermediate. This is the case for all deamination pathways investigated including the reaction with H_2O (closest to the physiological pH experienced by cytosine in living systems) as well as with H_3O^+. It is also a well-known fact that the deamination of the DNA base cytosine is an extremely rare event under normal physiological conditions (40−100 deaminations in a human genome per day[27]). This process, which is dependent on the protonation state of cytosine, seems to be the leading cause of mutations. It is also related to the appearance of so-called cytosine mutation "hotspots",[28] the origin of which is still a matter of controversy. The underlying purpose of this investigation, to be able to use the results from the deamination of formamidine as the basis to study the deamination of cytosine, promises to be a worthwhile endeavor, given the common structural features of both molecules and their high degree of stability toward deamination. No single level of theory/basis set agreed consistently with the G1 or G2 levels. An examination of the results suggests that B3LYP/6-311+G(d,p) was the most consistent overall. The use of MP2(full)/6-31G(d) geometries does not show any significant advantage compared to using RHF/6-31G(d) geometries for single point calculations.

Acknowledgment. We thank the Natural Sciences and Engineering Council of Canada (NSERC) for financial support and gratefully acknowledge the Memorial University of Newfoundland Advanced Computation and Visualization Centre for computer time.

Supporting Information Available: Calculations for the deamination of formamidine. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Kosturkiewicz, Z.; Ciszak, E.; Tykarska, E. *Acta Crystallogr., Sect. B* **1992**, *48*, 471.

(2) Page, M. I.; Webster, P.; Ghosez L. *J. Chem. Soc., Perkin Trans. 2* **1990**, *805*, 813.

(3) Brandstetter, H.; Turk, D.; Heoffken, H. W.; Gosse, D.; Stürzbecher, J. S.; Martin, P. D.; Edwards, B. F. P.; Bode, W. *J. Mol. Biol.* **1992**, *226*, 1085.

(4) Tong, M. K.; Papandreou, G.; Ganem, B. *J. Am Chem. Soc*. **1990**, *112*, 6137.

(5) Kourany-LeFoll, E.; Pais, M.; Sévenet, T.; Guittet, E.; Montagnac, A.; Fontaine, C.; Guénard, D.; Adeline, M. T. *J. Org. Chem*. **1992**, *57*, 3832.

(6) Zielinski, T. J.; Poirier, R.A.; Peterson, M. R.; Csizmadia, I. G. *J. Comput. Chem*. **1983**, *4*, 419.

(7) Steinman, U.; Estler, C. J.; Pann, O. *Drug Dev. Res.* **1986**, *7*, 153.

(8) Foussard-Blaupin, O.; Quevauviller A. *Ann. Pharm. Fr*. **1982**, *40*, 231.

(9) Terfort, A.; von Kiedrowski G. *Angew. Chem., Int. Ed Engl.* **1992**, *31*, 654.

(10) Frederico, L. A.; Kunkel, T. A.; Shaw, B. R. *Biochemistry* **1990**, *29*, 2532.

(11) Gojobori, T.; Li, W.H.; Graur, D. *J. Mol. Evol*. **1982**, *18*, 360.

(12) Dixon, W. J.; Hibbert, F.; Overill, R. E. *J. Chem. Soc., Perkin Trans. 2*. **1998**, *N.3*, 635.

(13) Perrin, C. L.; Thoburn, J. D. *J. Am. Chem. Soc.* **1993**, *115*, 3140.

(14) Deslongschamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon Press: Oxford, England, 1983.

(15) Kirby, A. J. *Acc. Chem. Res*. **1994**, *17*, 305.

(16) Perrin, C. L.; Arrhenius, G. M. L. *J. Am Chem. Soc*. **1982**, *104*, 2839.

(17) Zhang, Q.; Bell, R.; Truong, T. N. *J. Phys. Chem*. **1995**, *99*, 592.

(18) Nguyen, K. A.; Gordon, M. S.; Truhlar, D. G. *J. Am Chem. Soc*. **1991**, *113*, 1596.

(19) Lim, J.-H.; Lee, E. K.; Kim, Y. *J. Phys. Chem*. **1997**, *101*, 2233.

(20) Nagaoka, M.; Okuno, Y.; Yamabe, T. *J. Phys. Chem*. **1994**, *98*, 12506.

(21) Dziekonski, P.; Sokalski, W. A.; Podolyan, Y.; Leszczynski, J. *Chem. Phys. Lett*. **2003**, *367*, 367.

(22) Andrés, J.; Beltran, A.; Carda, M.; Krechl, J.; Monterde, J.; Silla, E. *J. Mol. Struct.: THEOCHEM* **1992**, *254*, 465.

(23) Tortajada, J.; Leon, E.; Luna, A.; Mó, O.; YáZez M. *J. Phys. Chem.* **1994**, *98*, 12919.

(24) (a) Poirier, R. A. *MUNgauss*, (*Fortran 90* version; Chemistry Department, Memorial University of Newfoundland, St. John's, Newfoundland, A1B 3X7 Canada. With contributions from 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.; and Jansen, G. *Comput. Phys. Commun.* **1994**, *81*, 293−317.

(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, revision B.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(26) Frisch, M. J.; Trucks, G. W.; Head-Gorden, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Rahavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92*, revision F.4; Gaussian, Inc.: Pittsburgh, PA, 1992.

(27) Shen, J. C.; Rideout, W. M.; Jones, P. A. *Nucleic Acids Res*. **1994**, *22*, 972.

(28) Coulondre, C.; Miller, J. H.; Farabaugh, P. J.; Gilbert, W. *Nature*. **1978**, *274*, 775.