THE DEAMINATION OF CYTOSINE: AN AB INITIO SCF MO STUDY

CENTRE FOR NEWFOUNDLAND STUDIES

TOTAL OF 10 PAGES ONLY
MAY BE XEROXED

(Without Author's Permission)

JIE XU

THE DEAMINATION OF CYTOSINE:

AN AB INITIO SCF MO STUDY

by

Jie Xu

A Thesis Submitted in Partial Fulfillment of the

Requirement for the Degree of

MASTER OF SCIENCE

Graduate Programme in Chemistry

Memorial University of Newfoundland

St. John's, Newfoundland

June 1990

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour metire des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-61850-7

Canada

ABSTRACT

In 1974, Lindahl and Nyberg studied the rate of deamination of cytosine residues in single and double stranded DNA. After studying the rate at several temperatures the activation energy for the conversion of cytosine to uracil was estimated to be 121 kJ/mol.

Possible mechanisms for deamination of cytosine are studied using ab initio SCF molecular orbital calculations. Ab initio calculations with STO-3G, 3-21G and 6-31G* basis sets have been performed to investigate the optimized structures and energies.

First the deamination of amidine is studied as a model compound and the calculations are then extended to cytosine. The results for both calculations indicate that the rate-determining step (second step) with OH^- has a barrier of 209 kJ/mol and with H_2O has a barrier of 174 kJ/mol. The results show that the reaction is thermodynamically favourable but kinetically not very favourable.

ACKNOWLEDGEMENTS

I am deeply and greatly indebted to my supervisor Dr. Raymond A. Poirier for all his patience and a lot of help during my studies. I would like to thank the Newfoundland and Laborador Computer Services (NLCS) and Memorial University of Newfoundland Computing Services for their generous allocation of computer time.

To my lovely son

TABLE OF CONTENTS

ABSTRACT.....ii

ACKNOWLEDGEMENTS.....iii

DEDICATION.....iv

TABLE OF CONTENTS.....v

LIST OF TABLES.....viii

LIST OF FIGURES.....xi

CHAPTER I INTRODUCTION.....1

1.1 General Introduction.....1

1.2 Experimental Results.....10

CHAPTER II MO THEORY AND COMPUTATION DETAILS.....24

2.1 MO Theory.....24

2.2 Basis Set Expansions.....26

2.3 Hartree-Fock Theory and SCF Theory.....28

2.4 Computational Methods.....30

CHAPTER III RESULTS AND DISCUSSION.....32

3.1 Method.....32

3.2 The Reaction of Amidine with OH^-.....33

3.3 The Reaction of Amidine with H_2O.....54

3.4 Effect of Basis Sets on the Deamination

of Amidine with OH^- and H_2O.....74

3.5 Comparison of the Deamination of Amidine

with OH^- and H_2O.....78

3.6 Optimized Structure of Cytosine.....79

3.7 The Reaction of Cytosine with OH^-.....79

3.8 The Reaction of Cytosine with H_2O.....107

3.9 Effect of Basis Sets on Cytosine with

OH^- and H_2O.....139

3.10 Comparison of the Deamination of Cytosine

with OH^- and H_2O.....143

3.11 Comparison of the Deamination of

Cytosine and Amidine with OH^-.....146

3.12 Comparison fo the Deamination of

Cytosine and Amidine with H_2O.....147

3.13 Comparison of the Experimental and

Calculated Results.....148

CONCLUSION.....151

APPENDIX.....153

REFERENCES.....157

LIST OF TABLES

Table

1. STO-3G, 3-21G and 6-31G* optimized geometries (under C_s symmetry) for the amidine-hydroxide complex (I_a)[a]

2. STO-3G, 3-21G and 6-31G* optimized geometries (under C_s symmetry) for the amidine-hydroxide complex (I_b)

3. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-hydroxide intermediate (I_c)

4. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-hydroxide transition state (I_d)

5. STO-3G, 3-21G and 6-31G* optimized geometries for the formamide ion-ammonia complex (I_e)

6. Total energies for the amidine-hydroxide structures[a] I_a -- I_e

7. Relative energies for the amidine-hydroxide structures[a] I_a -- I_e

8. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water complex (II_a)

9. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water complex (II_b)

10. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water intermediate (II_c)

11. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water intermediate (II_d)

12. STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water transition state (II_e)

13. STO-3G, 3-21G and 6-31G* optimized geometries for the formamide-ammonia complex (II_f)

14. Total energies for the amidine-water structures[a] II_a -- II_f

15. Relative energies for the amidine-water structures[a] II_a -- II_f

16. STO-3G and 3-21G optimized geometries (under C_s symmetry) for cytosine

17. STO-3G and 3-21G optimized geometries (under C_s symmetry) for the cytosine-hydroxide complex (III_a)

18. STO-3G and 3-21G optimized geometries (under C_s symmetry) for the cytosine-hydroxide complex (III_b)

19. STO-3G and 3-21G optimized geometries for the cytosine-hydroxide intermediate (III_c)

20. STO-3G and 3-21G optimized geometries for

the cytosine-hydroxide transition state (III_d)

21. STO-3G and 3-21G optimized geometries for the uracil ion-ammonia complex (III_e)

22. Total energies for the cytosine-hydroxide structures[a] III_a -- III_e

23. Relative energies for the cytosine-hydroxide structures[a] III_a -- III_e

24. STO-3G and 3-21G optimized geometries for the cytosine-water complex (IV_a)

25. STO-3G and 3-21G optimized geometries for the cytosine-water complex (IV_b)

26. STO-3G and 3-21G optimized geometries for the cytosine-water intermediate (IV_c)

27. STO-3G and 3-21G optimized geometries for the cytosine-water intermediate (IV_d)

28. STO-3G and 3-21G optimized geometries for the cytosine-water transition state (IV_e)

29. STO-3G and 3-21G optimized geometries for the uracil-ammonia complex (IV_f)

30. Total energies for the cytosine-water structures[a] IV_a -- IV_f

31. Relative energies for the cytosine-water structures[a] IV_a -- IV_f

LIST OF FIGURES

Figure

1.1 The structure of DNA. The pair of "ribbons" represents the sugar-phosphate backbone chains.

1.2 The two-dimensional structure of DNA.

1.3 Mechanism for the deamination of cytosine [10], B^-=buffer anion.

1.4 Mechanism for the deamination of cytosine [10], B^-=buffer anion.

1.5 Mechanism for the deamination of cytosine [14], HB^-= HSO_2^- ; $H_2PO_4^-$.

1.6 The tautomerism of cytosine.

3.1 The deamination of cytosine with OH^-.

3.2 The deamination of cytosine with H_2O.

3.3 The deamination of amidine with OH^-.

3.4 The deamination of amidine with H_2O.

3.5 The structures for the deamination of amidine with hydroxide. I_a and I_b are amidine-hydroxide complex (reactant), I_c is an intermediate, I_d is the transition state for the hydrogen rearrangement and I_e is the formamide ion-ammonia complex (product).

3.6 Schematic energy profile for the deamination of amidine with OH^-.

3.7 The structures for the deamination of amidine with water.

II_a and II_b are amidine-water complex (reactant),

II_c and II_d are intermediates, II_e is the transition

state for the hydrogen rearrangement and II_f is

the formamide-ammonia complex (product).

3.8 Schematic energy profile for the deamination of amidine with H_2O.

3.9 Numbering for the cytosine structure.

3.10 The structures for the cytosine-hydroxide complex (reactant).

3.11 The structure for the cytosine-hydroxide intermediate.

3.12 The structure of the transition state for the hydrogen

rearrangement (cytosine-hydroxide).

3.13 The structure for the uracil ion-ammonia complex (product).

3.14 Schematic energy profile for the deamination of cytosine with OH^-.

3.15 The structures for the cytosine-water complex (reactant).

3.16 The structures for the cytosine-water intermediate.

3.17 The structure of the transition state for the hydrogen

rearrangement (cytosine-water).

3.18 The structure for the uracil-ammonia complex (product).

3.19 Schematic energy profile for the deamination of cytosine with H_2O.

3.20 Reaction schemes for the formation of the cytosine-water and cytosine-hydroxide intermediates.

CHAPTER I

Introduction

1.1 General Introduction

The fundamental unit of any higher organism is the cell. Each daughter cell produced at cell division contains an identical copy of the genetic material. The instructions which enable these many intra-cellular reactions to take place, the genetic information, are encoded in the chromosomes.

The genetic information lies in the chemical substance known as deoxyribonucleic acid, or DNA, for this is the genetic material and the most important constituent of the chromosomes. DNA is a very long polymer known as a polynucleotide, made up of four types of building block units or nucleotides; adenine (A), thymine (T), guanine (G) and cytosine (C).

The first important feature of the Watson and Crick model for the structure of DNA is that (usually) the molecule is not just one polynucleotide chain but two such chains twisted around each other in the form of a double helix-rather like two pieces of string that have been twisted together (Fig.1.1).

Figure 1.1

The structure of DNA. The pair of "ribbons" represents the sugar-phosphate backbone chains. The points represent bases (i.e., adenine, thymine, guanine and cytosine).

The second important feature is that the pairing of the bases is not random. If the molecule is to be regular and undistorted then its dimensions are such that there is no room to accomodate a pair of the large purine bases (A and G) while a pair of pyrimidines (C and T) is too small. As a consequence one chain is always complementary to the other chain. The two-dimensional structure of DNA is shown in Fig. 1.2.

In a wide variety of organisms, both procaryotes and eucaryotes, chemicals which alter the structure of the genetic material DNA (or sometimes RNA) can cause heritable changes or mutations. So important is the exact sequence of the structural units in a gene, which may be up to 1500 structural units long, that the replacement of one structural unit by a different structural unit (mutation) may completely destroy the normal activity of the gene. It seems clear that gene mutation occurs by altering the sequence of nucleotides along the DNA molecule so that the functioning of the gene and, ultimately, the phenotype of the organism is affected. We now recognise that gene mutation can change the informational content of a gene in at least two ways, first, by a base substitution changing one base pair to a different base pair (or in a single-stranded nucleic acid by changing one base for another) and, second, by adding or deleting one or more base pairs.

Figure 1.2

The two-dimensional structure of DNA, where A=Adenine, T=Thymine, G=Guanine, C=Cytosine and P=Phosphate.

P
A
T
G
C
C
G
T
A
A
T

Agents which chemically alter the bases, which are adenine (A), thymine (T), guanine (G) and cytosine (C) in DNA, have been important tools for research in mutagenesis and carcinogenesis. Recently chemically modified DNA has been used to study the effect of altered base pairs on DNA duplex thermostability [1-2]. Of the many possible chemical alterations of DNA, including depurination, depyrimidination, alkylation and deamination, only deamination causes alterations which resemble the mismatched base pairs which might arise from a transition mutation.

Nucleotides consist of three components, i.e., ribose, base (purine and pyrimidine) and phosphate. The principal pyrimidines found in RNA are uracil and cytosine; in DNA they are thymine and cytosine. Cytosine and methylcytosine can undergo deamination to yield uracil and thymine, respectively. It has been confirmed by UV spectrophotometry that cytosine and cytidine deaminate to uracil and uridine, respectively, at all pH values. These products slowly degrade further in strongly alkaline solutions.

Spontaneous mutagenesis may involve a variety of distinct mechanisms [3-5]. One example of a spontaneous chemical change that has mutational consequences is the deamination of cytosine to uracil (see Scheme I).

Scheme I

The deamination of cytosine to uracil.

NH₂
H
N
O
N
H
H
+
H₂O
O
HN
O
N
H
H
H
+ NH₃

1.2 Experimental Results

The role which chemical mutagens play in the deamination of cytosine to uracil derivatives has been the subject of several investigations directed toward elucidation of *in vivo* mechanisms of mutagenic activity [6-10].

In 1966, R. Shapiro and R. S. Klein [10] reported that cytidine and cytosine are deaminated at 95° C by a variety of aqueous buffers of pH < 6.0. Based on the experimental data, they proposed two kinetically equivalent mechanisms for the deamination of cytosine which are outlined in Fig.1.3 and Fig.1.4. Both mechanisms combine specific hydrogen-ion catalysis with general base catalysis by buffer anion, B^- . The simple mechanism described in Fig.1.3 is analogous to the hydrolysis of an amide. However, little direct evidence has been reported to support its occurrence in known reactions involving nucleophilic displacement of the amino group of cytosine. The second mechanism is described in Fig.1.4 where following protonation, the buffer anion B^- adds to the 5,6 double bond with the resulting dihydrocytosine intermediate, undergoing deamination to give a uracil derivative. This route is similar to that established [11] for the reaction of the potent mutagen, hydroxylamine, with cytosine derivatives. Isolation of the components in reaction mixtures revealed only cytosine and uracil. No intermediates corresponding to buffer anion addition were detected, therefore, making it impossible to distinguish between these two mechanisms.

Figure 1.3

Mechanism for the deamination of cytosine [10], B^-=buffer anion.

NH2
H⊕
H2O,B⊖
NH2 OH
+BH
-NH3

Figure 1.4

Mechanism for the deamination of cytosine [10], B^-=buffer anion.

NH2
H⊕
B⊖
+H2O, -NH3
-BH

Schuster and Schramm [12] found that deamination of the amino groups of the cytosine, adenine, and guanine residues of the nucleic acids could be effected by nitrous acid. Jordan [13] showed that cytosine and guanine could be deaminated under strong acidic conditions. Quantitative data on the buffer-affected deamination of cytosine arabinoside was given by Notari [14] and Notari et al. [15-16], who reported that cytosine nucleosides underwent deamination in various buffers between pH 3.4 and 7.7. The work focused on the deamination of cytosine derivatives to the corresponding uracil derivatives after the formation of a reversibly hydrolyzing saturated 5,6-double bond adduct with bisulfite ion, (the bisulfite ion acting as a catalyst). Such a reaction may be responsible for mutagenic action by transforming the cytosine moieties in RNA and DNA [17-21].

Notari [14] proposed a mechanism for the hydrolytic deamination of cytosine in aqueous buffer, which is described in Fig.1.5. The kinetics of hydrolytic deamination of cytosine have been studied as a function of pH, temperature, and buffer composition. The reaction components were isolated, identified, and quantified. The prime criteria for catalysis by a nucleophile such as HB^- , are shown to be (a) sufficient nucleophilicity to attack the 6-position following the initial protonation at the 3-position and (b) the ability to subsequently donate a proton resulting in the complete saturation of the 5,6-double bond.

After studying the mechanism for the hydrolytic deamination of cytosine in

Figure 1.5

Mechanism for the deamination of cytosine [14], $HB^- = HSO_3^-$; $H_2PO_4^-$.

NH2
H⊕
HB⊖
BH
-H+
B⊖
B-
-NH3
+H2O
-HB⊖

aqueous buffer, Notari et al [16] investigated the intermolecular and intramolecular catalysis in deamination of cytosine nucleosides. The rates of deamination of cytosine arabinoside, cytidine and cytosine are compared in the presence and in the absence of catalytic buffers. The pseudo-first order rate constants for deamination of 1.8×10^{-3} M cytosine at 70^0C, pH 3.75 are 2.48--3.50×10^{-4} for several buffer components [16] and at 70^0C, pH 7.90 are 0.723--0.875 $\times10^{-4}$. The results indicate that pseudo-first order deamination rate constants, k_1, for cytidine and cytosine showed a linear dependency on buffer concentration. Both cytidine and cytosine were susceptible to general-acid and general-base catalysis by the buffer components. It is hypothesized that cytosine arabinoside is subject to general-base type of intramolecular catalysis by the 2'-hydroxyl group and thus only intermolecular general-acid catalysis is observed.

The solvolytic deamination of cytosine and cytidine are reported by Garrett and Tsau [22] in 1972. They confirmed using UV spectrophotometry and TLC that cytosine and cytidine deaminate to uracil and uridine, respectively, at all pH values [22]. The results show that the use of acidified solutions for the estimation of rate constants were preferred since the difference in λ_{max} values between cytosine and its product, uracil, was the greatest in this solution. The spectral changes of the acidified solutions of reacting cytosine followed the same pattern for all solutions more acidic than 0.1 N NaOH: a hypsochromic shift in λ_{max} from

cytosine (274nm) to uracil (259nm) [22]. The calculated first-order rate constants from the linear plots for the solvolytic deamination of cytosine and cytidine are given [22]. The first-order rate constants for the deamination of cytosine and cytidine become pH independent between pH 5 and 9 and indicate either water attack on the neutral species or its kinetic equivalent, hydroxyl-ion attack on the protonated species. The deamination rates characterized by the determined rate constant k_1 for both cytosine and cytidine above pH 9 are catalyzed by hydroxyl ions. A linear increase of log k_1 with increasing pH at a slope close to unity indicated a specific hydroxyl-ion catalyzed reaction. For cytosine, the energies of activation at the several pH values were (kJ/mol, pH): 53, 1.0; 51, 7.0; and 97, 9.8. For cytidine, they were: 49, 2.0; 54, 7.0; and 74, 11.0.

In 1974, Lindahl and Nyberg [23] studied the rate of deamination of cytosine residues in single- and double-stranded DNA. The activation energy for the conversion of cytosine to uracil was estimated to be 121 kJ/mol, based on the deamination of cytosine in different forms (poly(dG) ·poly(dC), poly(dC) and dCMP(dCMP=deoxycytidine 5'-monophosphate)), measured at several temperatures. They have also reported [24] an activation energy of 131 ± 8 kJ/mol for the depurination of DNA in neutral solution.

The heat-induced deamination conversion of 5-methylcytosine residues to thymine residues and of cytosine to uracil residues in single-stranded DNA was

studied by Ehrlich and Norris [25]. They investigated the rates of deamination of 5-methylcytosine and cytosine residues in single-stranded DNA at physiological pH and for temperatures of 70^0, 80^0, 95^0 and 110^0C to allow prediction of the rates at 37^0C. The results indicate that the calculated rates for the deamination of 5-methylcytosine and cytosine at 37^0C and pH 7.4 were 9.5×10^{-10} and 2.1×10^{-10} s^{-1}, respectively and that the Arrhenius plots of the temperature dependence of the deamination rates gave approximate activation energies of 100 and 113 kJ/mol for deamination of 5-methylcytosine and cytosine residues, respectively [25].

Another one step mechanism for mutation is the 1,3-hydrogen shift reaction. Tautomeric structures can be written for a variety of molecular formulae, such as amidine, formamide and acetaldehyde for example. The theoretical studies for amidine, formamide and other systems estimate that the barrier for 1,3-hydrogen shift range from 166 kJ/mol to 450 kJ/mol for the formic acid-formic acid and propene-propene reactions, respectively [26]. In aprotic solution, however, 1,3-hydrogen shifts are facilitated by direct involvement of one water molecule. The mechanism for the C-hydroxylmine-formamide tautomerism has been studied [27]. While in a gas phase the barrier for the hydrogen atom transfer is 251 kJ/mol [27], when reactant and product are solvated by one or two water molecules this barrier goes down to 46 kJ/mol and 8 kJ/mol, respectively [27]. The hydrogen atom transfer takes place through a chain of water molecules. As shown in Fig.1.6, 1,3-hydrogen shift, the migration of proton from 1-nitrogen,

leads to the formation of the enol tautomer of cytosine.

Figure 1.6

The tautomerism of cytosine.

NH2
N
O
N
H
H
H
H
NH2
N
O
H
N
H
H

CHAPTER II

MO Theory and Computational Details

2.1 MO Theory

The key to theoretical chemistry is molecular quantum mechanics. This is the science relating molecular properties to the motion and interactions of electrons and nuclei.

Ab initio calculations have long promised to become a major tool for the study of molecular problems of structure, stability, and reaction mechanism. The fundamental techniques, such as molecular orbital theory, date back to the earliest days of quantum mechanics more than half a century ago. However, widespread quantitative applications have only become practically possible in recent times, primarily because of explosive developments in computer hardware and associated achievements in the design of efficient mathematical algorithms. At the present time, the subject has become sufficiently mature and the technology sufficiently stable for practicing experimental chemists to use quantum mechanical methods directly for their own particular applications, rather than to rely on collaboration with a theoretician.

According to quantum mechanics [28], the energy and many properties of the stationary state of a molecule can be obtained by solution of the time-independent Schrodinger partial differential equation,

$$\hat{H}\Psi = E\Psi \tag{1}$$

Here $\hat{H}$ is the Hamiltonian, a differential operator representing the total energy, E is the numerical value of the energy of the st.. :, that is, the energy relative to a state in which the constituent particles (nuclei and electrons) are infinitely separated and at rest and Ψ is the wavefunction. The wavefunction depends on the position coordinates of all particles (which may take any value from $-\infty$ to $+\infty$) and also on the spin coordinates (alpha or beta spin).

Molecular orbital theory is an approach to molecular quantum mechanics which uses one-electron functions or orbitals to approximate the full wavefunction. A molecular orbital $\psi(x,y,z)$, is a function of the position coordinates x, y, z of a single electron.

The complete wavefunction for a single electron is the product of a molecular orbital and a spin function, $\psi(x,y,z)\alpha(\xi)$ or $\psi(x,y,z)\beta(\xi)$. It is termed a spin orbital, $\chi(x,y,z,\xi)$. To ensure antisymmetry, the spin orbitals may be arranged in a determinantal wavefunction.

With these features, we can write down a full many-electron molecular orbital wavefunction for closed-shell ground state of a molecule with n(even) electrons, doubly occupying n/2 orbitals:

$$\Psi = \frac{1}{\sqrt{n\,!}} \begin{vmatrix} \psi_1(1)\alpha(1) & \psi_1(1)\beta(1) & \psi_2(1)\alpha(1) & \cdots & \psi_{\frac{n}{2}}(1)\beta(1) \\ \psi_1(2)\alpha(2) & \psi_1(2)\beta(2) & \psi_2(2)\alpha(2) & \cdots & \psi_{\frac{n}{2}}(2)\beta(2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \psi_1(n)\alpha(n) & \psi_1(n)\beta(n) & \psi_2(n)\alpha(n) & \cdots & \psi_{\frac{n}{2}}(n)\beta(n) \end{vmatrix} \tag{2}$$

2.2 Basis Set Expansions

Above, we described how a many-electron wavefunction is constructed from molecular orbitals in the form of a single determinant. In practical applications of the theory, the individual molecular orbitals are expressed as a linear combination of a finite set of N prescribed one-electron functions known as basis functions. If the basis functions are ϕ_1 , ϕ_2 , ... , ϕ_N ,then an individual molecular orbital ψ_i can be written

$$\psi_i = \sum_{\mu=1}^{N} C_{\mu i}\phi_\mu \tag{3}$$

where $C_{\mu i}$ are the molecular orbital expansion coefficients. These coefficients provide the orbitals description with some flexibility, but clearly do not allow for complete freedom unless the ϕ_μ define a complete set. However, the problem of finding the orbitals is reduced from finding complete descriptions of the three-dimensional function ψ_i to finding only a finite set of linear coefficients for each

orbital.

In simple general versions of molecular orbital theory, atomic orbitals of constituent atoms are used as basis functions. Such treatments are often described as LCAO (linear combination of atomic orbital) theories.

Gaussian-type functions (GTF) were introduced into molecular orbital computations by Boys [29]. They are less satisfactory than STO's (Slater Type Orbitals) as representations of atomic orbitals, particularly because they do not have a cusp at the origin. Nevertheless, they have the important advantage that all integrals in the computations can be evaluated explicity without recourse to numerical integration. The basic advantage of GTF's is due to the fact that the product of any two Gaussians is also a Gaussian. Consequently, all integrals have explicit analytical expressions and may be evaluated rapidly. A similar theorem does not exist for STO's.

Very often STO's are expanded in terms of GTF's (i.e. a linear combination of GTF's is tailored to fit the shape of STO's (STO-NG basis set)). This method of using fixed linear combinations of GTF's and storing only AO integrals (to reduce the storage space needed) is called contraction. Without contraction many calculations on the electronic structure of large molecules would be quite impossible. In our calculation standard sets of GTF's are used.

2.3 Hartree-Fock Theory and SCF Theory

Up to this point, we have described how determinantal wavefunctions may be constructed from molecular orbitals, and how the orbitals may, in turn, be expanded in terms of a set of basis functions. The method for determining the expansion coefficients is in the realm of Hartree-Fock (HF) theory.

HF theory is based on the variational method in quantum mechanics [30].

We deal with these equations for closed-shell systems. The Hartree-Fock equations are

$$\sum_{\nu=1}^{N}(F_{\mu\nu} - \epsilon_i S_{\mu\nu})C_{\nu i} = 0 \qquad \mu=1,2,\ldots,N \tag{4}$$

with the normalization conditions

$$\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} C^*_{\mu i} S_{\mu\nu} C_{\nu i} = 1. \tag{5}$$

Here, ϵ_i is the one electron energy of molecular orbital ψ_i , $S_{\mu\nu}$ are the elements of an $N \times N$ matrix termed the overlap matrix,

$$S_{\mu\nu} = \int \phi_\mu^*(1)\phi_\nu(1)dx_1 dy_1 dz_1, \tag{6}$$

and $F_{\mu\nu}$ are the elements of another $N \times N$ matrix, the Fock matrix,

$$F_{\mu\nu} = H^{core}_{\mu\nu} + \sum_{\lambda=1}^{N}\sum_{\sigma=1}^{N} P_{\lambda\sigma}[(\mu\nu \mid \lambda\sigma) - \frac{1}{2}(\mu\lambda \mid \nu\sigma)]. \tag{7}$$

In this expression, $H^{core}_{\mu\nu}$ is a matrix representing the energy of a single electron

in a field which is fixed. Its elements are

$$H_{\mu\nu}^{core} = \int \phi_\mu^*(1)\hat{h}_{core}(1)\phi_\nu(1)dx_1dy_1dz_1 \ ,$$

$$\hat{h}^{core}(1) = -\frac{1}{2}(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}) - \sum_{A=1}^{M} \frac{Z_A}{r_{1A}}. \tag{8}$$

Here Z_A is the atomic number of atom A, and the summation is carried out over all atoms. The quantities $(\mu\nu \mid \lambda\sigma)$ appearing in (7) are two-electron repulsion integrals:

$$(\mu\nu \mid \lambda\sigma) = \int\int \phi_\mu^*(1)\phi_\nu(1)(\frac{1}{r_{12}})\phi_\lambda^*(2)\phi_\sigma(2)dx_1dy_1dz_1dx_2dy_2dz_2. \tag{9}$$

Here $(\mu\nu \mid \lambda\sigma)$ are given in chemist's notation. They are multiplied by the elements of the one-electron density matrix, $P_{\lambda\sigma}$,

$$P_{\lambda\sigma} = 2\sum_{i=1}^{occ} C_{\lambda i}^* C_{\sigma i} \tag{10}$$

The summation is over occupied molecular orbitals only. The factor of two indicates that two electrons occupy each molecular orbital, and the asterisk denotes complex conjugation (required if the molecular orbitals are not real functions).

The electronic energy, E^{ee} , is given by (11),

$$E^{ee} = \frac{1}{2}\sum_{\mu=1}^{N}\sum_{\nu=1}^{N} P_{\mu\nu}(F_{\mu\nu} + H_{\mu\nu}^{core}) \tag{11}$$

which, when added to (12), accounting for the internuclear repulsion,

$$E^{nr} = \sum_{A<B}^{M}\sum \frac{Z_A Z_B}{R_{AB}} \tag{12}$$

(where Z_A and Z_B are the atomic numbers of atoms A and B, and R_{AB} is their separation) yields an expression for the total energy.

The Hartree-Fock equations (4) are not linear since the Fock matrix $F_{\mu\nu}$ itself depends on the molecular orbital coefficients, $C_{\mu i}$, through the density matrix expression (10). Solution necessarily involves an iterative process. Since the resulting molecular orbitals are derived from their own effective potential, the technique is frequently called self-consistent-field (SCF) theory.

2.4 Computational Methods

During the past decade, the analytical calculation of the first derivatives of the energy has become practical and efficient for ab initio molecular orbital methods [31]. For SCF calculations, analytical gradients can provide an order of magnitude more information about a molecule than the energy computation alone. Analytical derivative calculations recently have been extended to methods that include correlation energy [32-35]. Energy gradients were first used to calculate vibrational force constants [36,37]. However, geometry optimization has now become the most widespread application.

Optimization of molecular geometry is one of the most important procedures in contemporary theoretical chemistry. The traditional application of geometry optimization has been the determination of equilibrium geometries from ab initio

calculations, and extensive literature [38] has demonstrated that ab initio structures are generally in good agreement with experimental data. This fact recommends ab initio geometry optimization as a powerful tool for investigating molecular structures which are not experimentally known.

Molecular orbital calculations generally involve three basic steps: (a) calculation of integrals, (b) solution of the self-consistent equations and (c) evaluation of the energy gradient.

CHAPTER III

Results and Discussion

3.1 Method

Closed shell SCF ab initio calculations were performed using the MUN-GAUSS program [39]. A minimal STO-3G [40,41] and split-valence 3-21G basis set [42-45] was used throughout. In addition, calculations were performed with the 6-31G* basis set [46] for the deamination of amidine. The geometries were optimized using gradient optimization techniques [47,48]. Except for transition state structures, all the geometry optimizations were performed using the optimally conditioned method of Davidon and Nazareth [49,50]. The criterion for convergence was the norm of the gradient length being less than 5.0×10^{-4} mdyn, a condition which normally gives bond lengths to within 0.0001 $\dot{A}$ and angles to within 0.01^0 of their true optimum values [51]. Transition state structures were determined using the VA05 sum of squares optimization technique [52]. The force constant matrix was calculated using VA05 to determine the order of some of the stationary points.

We have studied the deamination of cytosine using ab initio SCF MO calculations. The two mechanisms investigated in this work for the deamination of cytosine are outlined in Fig. 3.1 and Fig. 3.2. Fig. 3.1 shows that the nucleophilic attack by the hydroxide at position 4, followed by 1,3-hydrogen rearrangement and elimination of NH_3 to give the uracil anion. A similar mechanism with H_2O is shown in Fig.3.2 where the first step involves protonation of the nitrogen at position 3 and nucleophilic attack by hydroxide at position 4. The second step is followed by a 1,3-hydrogen rearrangement and subsequent elimination of NH_3, to give uracil.

Because of the size of molecules involved, as a first step, we investigated one (without intermediate in Fig.3.3) and two step mechanisms for the deamination of amidine as a model for cytosine (Fig. 3.3 and 3.4).

Unless otherwise specified, all calculated values reported in this chapter refer to the results from calculations using the 3-21G basis set.

3.2 The Reaction of Amidine with OH^-

A one step mechanism for the deamination of amidine with hydroxide has been studied. The structures involved in the amidine-hydroxide complex (reactant) and formamidine ion-ammonia complex (product) will be discussed in the two step mechanisms. The transition state structure (c.f. Table 1 in appendix) for the

Figure 3.1

The deamination of cytosine with OH^-.

NH2
+ OH⁻
HO⁻
NH2
IIIa/IIIb
NH2 OH
IIIc
IIId
NH3
IIIe
+ NH3

Figure 3.2

The deamination of cytosine with H_2O.

NH2
+ H2O
VIa / VIb
NH2 OH
VIc / VId
VIe
NH3
VIf
+ NH3

Figure 3.3

The deamination of amidine with OH^-.

NH2
C
N
H
H
+ OH⁻
HO⁻
NH2
Ia/Ib
NH2
OH
Ic
H
O
NH2
Id
O
NH3
Ie
+ NH3

Figure 3.4

The deamination of amidine with H_2O.

IIa/IIb IIc/IId IIe

IIf

one step mechanism shows that the O8-C1 bond is forming and C-N2 bond is breaking forming NH_3 as a leaving group. This structure is a first order saddle point on the potential energy surface, that is, the force constant matrix has only one negative eigenvalue. The results of these calculations indicate that the concerted reaction has a barrier of 424 kJ/mol. The one step mechanism was not considered any further because of the high barrier. The transition state structure and related information are given in the appendix.

The structures (two step mechanisms) involved in the reaction of amidine with hydroxide are shown in Fig.3.5. The corresponding geometries for all structures, amidine-hydroxide complex (I_a and I_b), intermediate (I_c), transition state (I_d) and formamide ion-ammonia complex (I_e) are given in Tables 1-5 and the corresponding energies are given in Table 6. The two conformations of the amidine-hydroxide complex (I_a and I_b) were constrained to have C_s symmetry. These two structures differ only in the orientation of the OH^- . The calculated O9-C1 bond lengts in I_a is found to be 3.54 $\dot{A}$. The calculated O9-C1 bond length in I_b is found to be 3.64 $\dot{A}$. Structure I_a is slightly more stable than I_b by 2 kJ/mol (c.f. Table 7).

The intermediate (I_c) resulting from addition of OH^- to amidine, has calculated O9-C1 and C1-N2 bond lengths of 1.52 $\dot{A}$. Comparison of the O9-C1 and C1-N2 bond lengths in I_a and I_c show that the O9-C1 bond length in I_c is getting shorter than that in I_a by 2.03 $\dot{A}$ and C1-N2 bond length in I_c is getting longer than that in I_a by 0.12 $\dot{A}$.

Figure 3.5

The structures for the deamination of amidine with hydroxide. I_a and I_b are amidine-hydroxide complex (reactant), I_c is an intermediate, I_d is the transition state for the hydrogen rearrangement and I_e is the formamide ion-ammonia complex (product).

O9
H2
N2
H6
C1
H3
N4
H5

Ia

O9
H1
H2
N2
C1
H3
N4
H5

Ib

H7
H6
N2
C1
O9
H8
H4
N3
H1

Ic

H6
H7
77
N2
H1
O9
C1
H8
H4
N3

Id

H7
H8
H9
N6
O9
C1
N4
H3
H5

Ie

Table 1 STO-3G, 3-21G and 6-31G* optimized geometries (under C_s symmetry) for the amidine-hydroxide complex (I_a)[a]

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4360	1.4043	1.4010
H3C1	1.1052	1.0898	1.0916
N4C1	1.2915	1.2658	1.2591
H5N4	1.0487	1.0170	1.0070
H1N2	1.0987	1.0366	1.0157
O9C1	3.2068	3.5401	3.7059
H6O9	1.0173	0.9971	0.9563
Bond angles			
H3C1N2	113.59	113.01	113.15
N4C1N2	127.58	127.10	126.93
H5N4C1	107.25	112.50	109.47
H1N2C1	109.22	115.37	112.16
O9C1N4	88.19	92.69	92.32
H6O9C1	85.80	80.51	81.66
Torsion angles			
H1N2C1H3	+140.58	+131.36	+129.88
H2N2C1H3	-140.58	-131.36	-129.88

a. Bond lengths are given in angstroms and bond angles in degrees for all tables.

Table 2 STO-3G, 3-21G and 6-31G* optimized geometries (under C_s symmetry) for the amidine-hydroxide complex (I_b)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4377	1.4034	1.4011
H3C1	1.1050	1.0904	1.0918
N4C1	1.2908	1.2652	1.2585
H5N4	1.0489	1.0171	1.0070
H1N2	1.0957	1.0336	1.0145
O9C1	3.2055	3.6367	3.7949
H6O9	1.0217	0.9997	0.9569
Bond angles			
H3C1N2	113.51	112.91	113.05
N4C1N2	127.60	127.34	127.07
H5N4C1	107.26	112.42	109.40
H1N2C1	109.18	116.65	112.96
O9C1N4	87.81	97.96	97.14
H6O9C1	127.46	111.08	110.64
Torsion angles			
H1N2C1H3	+140.27	+129.91	+128.94
H2N2C1H3	-140.27	-129.91	-128.94

Table 3 STO-3G, 3-21G and 6-31G* optimized geometries for amidine-hydroxide intermediate (I_c)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.5440	1.5211	1.4927
N3C1	1.4344	1.3869	1.3796
H4N3	1.0646	1.0379	1.0213
H6N2	1.0389	1.0102	1.0045
H7N2	1.0392	1.0100	1.0062
H8C1	1.1099	1.0854	1.0928
O9C1	1.4697	1.5165	1.4631
H1O9	0.9913	0.9684	0.9471
Bond angles			
N2C1N3	118.78	118.79	118.80
H4N3C1	99.82	106.04	105.41
H6N2C1	103.43	106.50	104.31
H7N2C1	104.30	107.31	105.72
H8C1N2	103.23	106.92	105.63
O9C1N2	101.94	100.04	101.89
H1O9C1	100.91	100.81	99.87

Table 3 continued

Torsion angles			
H4N3C1N2	57.26	50.17	52.58
H6N2C1N3	48.63	23.04	34.43
H7N2C1N3	303.01	265.90	283.61
H8C1N2N3	237.95	234.78	236.88
O9C1N2N3	127.65	124.95	126.32
H1O9C1N3	-9.49	-17.63	-18.39

Table 4 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-hydroxide transition state (I_d)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	2.0992	2.0449	2.0362
N3C1	1.3095	1.2910	1.2897
H4N3	1.0482	1.0199	1.0096
H6N2	1.0451	1.0270	1.0129
H7N2	1.0438	1.0248	1.0123
H8C1	1.1038	1.0756	1.0798
O9C1	1.4053	1.4193	1.3849
H1N2	1.5040	1.7675	1.9477
Bond angles			
N2C1N3	118.78	118.47	119.20
H4N3C1	104.53	109.76	107.35
H6N2C1	118.80	110.02	105.73
H7N2C1	118.74	104.79	101.60
H8C1N2	91.19	90.79	89.62
O9C1N2	80.29	86.04	90.17
H1N2C1	55.80	56.90	54.50

Table 4 continued

Torsion angles			
H4N3C1N2	81.12	81.68	84.02
H6N2C1N3	112. 23	99.89	81.09
H7N2C1N3	348.71	348.01	334.17
H8C1N2N3	240.28	238.58	239.28
O9C1N2N3	124.51	124.84	125.97
H1N2C1N3	232.99	231.88	225.94

Table 5 STO-3G, 3-21G and 6-31G* optimized geometries for the formamide ion-ammonia complex (I_e)

Bond lengths	STO-3G	3-21G	6-31G*
O9C1	1.2828	1.2700	1.2440
H3C1	1.1197	1.1118	1.1163
N4C1	1.3293	1.2975	1.2971
H5N4	1.0483	1.0188	1.0082
H9O9	1.4333	1.8265	2.0087
N6H9	1.0950	1.0270	1.0153
H7N6	1.0366	1.0101	1 0054
H8N6	1.0367	1.0102	1.0054
Bond angles			
H3C1O9	116.23	115.17	115.03
N4C1O9	125.96	127.75	128.29
H5N4C1	104.80	110.40	107.54
H9O9C1	113.30	125.25	126.63
N6H9O9	176.01	169.64	164.49
H7N6H9	103.70	109.12	104.69
H8N6H9	103.84	109.13	104.66

Table 5 continued

Torsion angles			
N4C1O9H3	180.03	180.05	180.03
H5N4C1H3	0.10	0.12	0.10
H9O9C1H3	0.85	3.43	4.51
N6H9O9C1	150.93	155.23	156.95
H7N6H9O9	52.69	59.14	55 17
H8N6H9O9	-52.69	-59.14	-55.17

Table 6 Total energies for the amidine-hydroxide structures[a] I_a -- I_e

Compound (OH^-)	Energies (hartrees)		
	STO-3G	3-21G	6-31G*
I_a	-221.255013	-223.135347	-224.418509
I_b	-221.254904	-223.134560	-224.417703
I_c	-221.385775	-223.176469	-224.443556
I_d	-221.329240	-223.153612	-224.417320
I_e	-221.407681	-223.236483	-224.511773

a. See Fig. 3.5 for the structures.

The transition state (I_d) for the 1,3-hydrogen rearrangement (c.f. Fig.3.5) indicates that the O9-C1 bond is shortening and C1-N2 bond is breaking forming NH_3 as a leaving group. The order of the stationary point, which was determined using VAO5, indicated that this structure is a first order saddle point on the potential energy surface, that is, the force constant matrix has one negative eigenvalue. The calculated O9-C1 bond length in the transition state (I_d) is found to be 1.42 $\dot{A}$,which is between the single O9-C1 bond (1.52 $\dot{A}$) in I_c and the O9-C1 double bond (1.27 $\dot{A}$) in I_e .

In the final product (I_e), the formamide ion-ammonia complex, the NH_3 is hydrogen bonded to the oxygen with a typically linear hydrogen bond.

The relative energies of the amidine-hydroxide complex, intermediate, transition state and product are summarized in Table 7. A schematic energy profile for the deamination of amidine with OH^- is shown in Fig.3.6. The results show that the second step has a barrier of 60 kJ/mol and that the product is more stable than the I_a by 268 kJ/mol. Therefore, these results indicate that this reaction is both thermodynamically and kinetically favourable.

3.3 The Reaction of Amidine with H_2O

The structures involved in the reaction of amidine with water are shown in Fig.3.7. The corresponding geometries for all structures, amidine-water complex

Table 7 Relative energies for the amidine-hydroxide structures[a] I_a -- I_e

Compound (OH^-)	Relative energies (kJ/mol)		
	STO-3G	3-21G	6-31G*
I_a	400.83	265.53	244.86
I_b	401.12	267.60	246.98
I_c	57.51	157.57	179.10
I_d	206.95	217.58	247.97
I_e	0.00	0.00	0.00

a. See Fig. 3.6

Figure 3.6

Schematic energy profile for the deamination of amidine with OH^-. $Am+OH^-$ and For^-+NH_3 energies are just the sum of the energies $E(Am+OH^-)=E(Am)+E(OH^-)$ and $E(For^-+NH_3)=E(For^-)+E(NH_3)$. Where $R_{AM...OH^-}$ corresponds to I_a, $I_{Am...OH^-}$ to I_c, $T_{Am...OH^-}$ to I_d and $P_{For^-...NH_3}$ to I_e.

KJ/MOL
300
250
200
150
100
50
0
Am + OH⁻
74
R Am··· OH⁻
110
T Am··· OH⁻
60
T Am··· OH⁻
268
218
For⁻ + NH3
67
P For⁻... NH3

Figure 3.7

The structures for the deamination of amidine with water. II_a and II_b are amidine-water complex (reactant), II_c and II_d are intermediates, II_e is the transition state for the hydrogen rearrangement and II_f is the formamide-ammonia complex (product).

IIa

IIb

IIc

IId

IIe

IIf

(II_a and II_b), intermediate (II_c and II_d), transition state (II_e) and formamide-ammonia complex (II_f) are given in Tables 8-13. The total computed energies are reported in Table 14. Two structures for the amidine-water complex corresponding to two different orientations of H_2O (II_a and II_b) are obtained. The O9-C1 bond length in II_a is calculated to be 3.15 $\dot{A}$. The calculated O9-C1 bond length in II_b is obtained to be 3.48 $\dot{A}$. Structure II_a is predicted to be more stable than the II_b by 93 kJ/mol (c.f. Table 15).

In the two conformers of the intermediate (II_c and II_d) resulting from addition of hydroxide at C1 and protonation at N3, the O9-C1 and O9-H1 bond lengths are calculated to be 1.43 and 0.97 $\dot{A}$ respectively. The HOCN torsion angles in II_c and II_d are -39^0 and 169^0, respectively. Structure II_d is slightly more stable than the II_c by 14 kJ/mol.

The calculated O9-C1 and N2-C1 bond lengths for the 1,3-hydrogen rearrangement transition state (II_e) are found to be 1.37 and 1.66 $\dot{A}$, respectively. Again the transition state structure (II_e) indicates that the O9-C1 bond is shortening and the N2-C1 bond is breaking forming NH_3 as a leaving group. The force constant matrix indicates that this is a first order saddle point. From Table 11-12 it is clear that the O9-C1 bond length decreases in going from intermediate to transition state by 0.06 $\dot{A}$ and the N2-C1 bond length increase in going from intermediate to transition state by 0.22 $\dot{A}$.

Table 8 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water complex (II_a)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4198	1.3491	1.3560
H3C1	1.0956	1.0798	1.0833
N4C1	1.2823	1.2719	1.2627
H5N4	1.0413	1.0070	1.0010
H6N2	1.0385	1.0086	1.0008
H7N2	1.0261	0.9934	0.9938
O9C1	3.1355	3.1509	3.3758
H1O9	0.9864	0.9665	0.9474
H2O9	0.9884	0.9863	0.9578
Bond angles			
H3C1N2	114.28	114.23	113.50
N4C1H3	124.92	123.51	123.33
H5N4C1	108.99	115.31	111.49
H6N2C1	109.61	117.02	116.68
H7N2C1	111.62	121.71	118.25
O9C1N4	63.43	60.49	58.93
H1O9C1	113.13	114.70	115.50
H2O9H1	100.99	108.10	105.24

Table 8 continued

Torsion angles			
N4C1H3N2	176.66	179.81	178.21
H5N4C1H3	-1.00	0.29	-0.97
H6N2C1N4	18.02	2.00	12.32
H7N2C1N4	142.52	177.77	159.31
O9C1N4N2	13.68	0.59	9.00
H1O9C1N2	85.96	96.24	94.17
H2O9H1C1	-54.23	-45.47	-42.56

Table 9 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water complex (II_b)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4672	1.4304	1.4225
H3C1	1.0976	1.0825	1.0866
N4C1	1.2723	1.2508	1.2464
H5N4	1.0466	1.0126	1.0048
H6N2	1.0340	1.0049	1.0037
H7N2	1.0340	1.0048	1.0037
O9C1	3.5562	3.4791	3.6300
H1O9	0.9873	0.9669	0.9482
H2O9	0.9876	0.9707	0.9510
Bond angles			
H3C1N2	116.81	116.26	116.85
N4C1H3	124.23	122.89	122.69
H5N4C1	108.98	114.54	110.95
H6N2C1	107.71	114.80	110.40
H7N2C1	107.74	114.83	110.42
O9C1N4	63.48	60.44	61.47
H2O9C1	65.31	64.91	64.88
H2O9H1	99.39	103.84	102.39

Table 9 continued

Torsion angles			
N4C1H3N2	180.07	180.07	180.03
H5N4C1H3	-0.04	-0.07	-0.02
H6N2C1H3	56.24	66.30	58.98
H7N2C1H3	-56.24	-66.30	-58.98
O9C1N4N2	-1.81	-1.29	-1.21
H1O9C1N2	181.86	181.54	181.28
H2O9H1C1	-1.90	-1.39	-1.25

Table 10 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water intermediate (II_c)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4845	1.4448	1.4376
N3C1	1.4865	1.4575	1.4491
H4N3	1.0314	1.0019	1.0018
H5N3	1.0328	1.0043	1.0036
H6N2	1.0328	1.0002	1.0010
H7N2	1.0336	1.0014	1.0024
H8C1	1.1020	1.0812	1.0859
O9C1	1.4381	1.4255	1.3868
H1O9	0.9902	0.9656	0.9479
Bond angles			
N2C1N3	111.47	117.15	118.08
H4N3C1	108.89	115.86	111.64
H5N3C1	108.89	114.42	110.53
H6N2C1	106.58	116.04	111.33
H7N2C1	107.49	113.53	110.31
H8C1N2	107.39	109.33	107.54
O9C1N2	107.78	104.33	105.45
H1O9C1	104.96	108.15	107.61

Table 10 continued

Torsion angles			
H4N3C1N2	-173.84	-87.39	-69.27
H5N3C1H4	244.84	134.25	120.48
H6N2C1N3	52.08	49.14	53.12
H7N2C1N3	299.78	272.35	292.14
H8C1N2N3	243.41	239.17	239.77
O9C1N2N3	126.13	119.71	120.71
H1O9C1N3	1.23	-38.52	-43.12

Table 11 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water intermediate (II_d)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.4852	1.4494	1.4380
N3C1	1.4848	1.4429	1.4342
H4N3	1.0326	1.0051	1.0030
H5N3	1.0338	1.0067	1.0042
H6N2	1.0321	1.0036	1.0019
H7N2	1.0334	1.0062	1.0038
H8C1	1.1029	1.0798	1.0849
O9C1	1.4302	1.4341	1.3982
H1O9	0.9891	0.9669	0.9490
Bond angles			
N2C1N3	112.40	113.52	113.65
H4N3C1	107.11	113.04	109.81
H5N3C1	106.74	112.23	109.38
H6N2C1	107.02	113.11	110.12
H7N2C1	107.25	112.72	109.91
H8C1N2	107.31	107.87	107.67
O9C1N2	109.23	107.85	108.30
H1O9C1	103.64	109.06	108.21

Table 11 continued

Torsion angles			
H4N3C1N2	-173.25	-171.07	-176.19
H5N3C1H4	248.55	234.23	243.33
H6N2C1N3	61.51	66.02	59.44
H7N2C1N3	309.50	298.20	301.09
H8C1N2N3	241.95	240.73	241.28
O9C1N2N3	120.98	121.09	122.06
H1O9C1N3	169.55	168.59	171.77

Table 12 STO-3G, 3-21G and 6-31G* optimized geometries for the amidine-water transition state (II_e)

Bond lengths	STO-3G	3-21G	6-31G*
N2C1	1.8121	1.6569	1.6253
N3C1	1.4075	1.4062	1.3904
H4N3	1.0135	0.9977	0.9920
H5N3	1.0139	0.9999	0.9930
H6N2	1.0328	1.0090	1.0046
H7N2	1.0325	1.0081	1.0038
H8C1	1.1112	1.0817	1.0883
O9C1	1.3236	1.3691	1.3253
H1N2	1.1425	1.2124	1.2085
Bond angles			
N2C1N3	111.19	109.59	112.53
H4N3C1	122.24	119.62	122.73
H5N3C1	119.55	113.73	117.67
H6N2C1	121.80	117.66	117.03
H7N2C1	121.25	112.67	115.81
H8N2C1	98.88	104.34	104.35
O9C1N2	87.52	92.15	93.48
H1N2C1	67.76	71.77	70.35

Table 12 continued

Torsion angles			
H4N3C1N2	-95.57	-113.25	-85.45
H5N3C1N2	88.20	102.49	89.01
H6N2C1N3	126.87	130.47	131.71
H7N2C1H3	345.68	357.60	358.35
H8C1N2N3	244.94	239.42	293.68
O9C1N2N3	121.88	119.02	119.45
H1N2C1N3	236.58	242.97	242.90

Table 13 STO-3G, 3-21G and 6-31G* optimized geometries for the formamide-ammonia complex (II_f)

Bond lengths	STO-3G	3-21G	6-31G*
O9C1	1.2184	1.2207	1.1983
H3C1	1.1041	1.0784	1.0871
N4C1	1.4296	1.3493	1.3460
H5N4	1.0255	0.9953	0.9930
H6N4	1.0263	0.9981	0.9955
H1O9	1.9371	2.1100	2.2973
N7H1	1.0336	1.0061	1.0043
H8N7	1.0319	1.0039	1.0030
H9N7	1.0320	1.0039	1.0029
Bond angles			
H3C1O9	123.87	121.18	121.98
N4C1O9	123.76	124.52	124.32
H5N4C1	112.75	121.74	121.66
H6N4C1	112.44	119.61	119.44
H1O9C1	113.17	94.79	96.11
N7H1O9	174.65	144.15	146.40
H8N7H1	104.40	112.53	107.32
H9N7H1	104.39	112.34	107.31

Table 13 continued

Torsion angles			
N4C1O9H3	176.09	179.95	180.06
H5N4C1H3	-36.24	0.02	0.23
H6N4C1H3	197.94	179.79	179.85
H1O9C1H3	3.68	0.27	0.77
N7H1O9C1	4.98	-3.33	-1.20
H8N7H1O9	133.45	121.03	122.53
H9N7H1H8	108.64	126.74	114.62

Table 14 Total energies for the amidine-water structures[a] II_a -- II_f

Compound (H_2O)	Energies (hartrees)		
	STO-3G	3-21G	6-31G*
II_a	-222.133981	-223.853695	-225.102779
II_b	-222.115357	-223.818174	-225.078524
II_c	-222.183429	-223.849440	-225.099238
II_d	-222.193799	-223.854666	-225.105232
II_e	-222.057560	-223.785155	-225.015579
II_f	-222.152181	-223.870066	-225.122167

a. See Fig.3.7 for the structures

In the final product (II_f), the formamide-ammonia complex, the NH_3 is hydrogen bonded to the oxygen with a typically linear hydrogen bond. Comparison of the structures of II_d, II_e and II_f show the O9-C1 bond getting shorter in going from the single O9-C1 bond in II_d (1.43 $\dot{A}$), to the O9-C1 double bond (1.22 $\dot{A}$) in II_f .

The relative energies of amidine-hydroxide complex, intermediates, transition state and formamide-ammonia complex are summarized in Table 15. The results indicate that the second step has a barrier of 182 kJ/mol and that the product is more stable than the complex (II_a) by 42 kJ/mol. Fig.3.8 illustrates that the reaction of amidine with water is thermodynamically favourable but kinetically not favourable.

3.4 Effect of Basis Sets on the Deamination of Amidine with OH^- and H_2O

Comparing some of the important geometrical parameters of the STO-3G, 3-21G and 6-31G* optimized geometries a number of trends appear. The 3-21G and 6-31G* basis set consistently give shorter N-C, H-C, H-N and H-O bond lengths than those from STO-3G basis set and longer O-C bond lengths of the order of 0.015-0.430 $\dot{A}$. The 3-21G and 6-31G* H3C1N2, N4C1N2, H5N4C1, H1N2C1, O2C1N4 and H6O2C1 bond angles differ by less than 3^0 , except in the transition state structures, where the angles vary by as much as 4^0. In addition,

Table 15 Relative energies for the amidine-water structures[a] II_a – II_f

Compound (H_2O)	Relative energies (kJ/mol)		
	STO-3G	3-21G	6-31G*
II_a	47.48	42.98	50.90
II_b	96.68	136.24	144.58
II_c	-82.04	54.14	60.20
II_d	-109.27	40.43	44.46
II_e	248.43	222.93	279.85
II_f	0.00	0.00	0.00

a. See Fig. 3.8

Figure 3.8

Schematic energy profile for the deamination of amdine with H_2O. Am+H_2O and For+NH_3 energies are just the sum of the energies E(Am+H_2O)=E(Am)+E(H_2O) and E(For+NH_3)=E(For)+E(NH_3). Where $R_{Am...H_2O}$ corresponds to II_a, $I^1_{Am...H_2O}$ to II_c, $I^2_{Am...H_2O}$ to II_d, $T_{Am...H_2O}$ to II_e and $P_{For...NH_3}$ to II_f.

KJ/MOL
240
200
160
120
80
40
0
Am + H2O
79
RAm---H2O
42
I'Am... H2O
11
I2Am... H2O
14
TAm--- H2O
182
223
PFor...NH3
34
For. +NH3

the most noticeable change in the geometry of the transition state involves a considerable narrowing of the OCN2 bond angle which increases in going from STO-3G, 3-21G to 6-31G*, where the OCN2 bond angle increases by 4-10^0.

3.5 Comparison of the Deamination of Amidine with OH^- and H_2O

Table 7 and Table 15 summarized the relative energies for the deamination of amidine with OH^- and H_2O, respectively. The schematic energy profile for the deamination of amidine with OH^- and H_2O are shown in Fig.3.6 and Fig.3.8, respectively. Comparison of the results of the reaction for the deamination of amidine with hydroxide and water indicate that amidine-hydroxide intermediate (I_c) is more stable than amidine-hydroxide complex by 110 kJ/mol and that the amidine-water intermediate (II_d) is only sightly more stable than the amidine-water complex by 3 kJ/mol. The second step with hydroxide has a barrier of 60 kJ/mol and with water has a barrier of 182 kJ/mol. The barrier for the deamination of amidine with hydroxide is significantly lower than the deamination of amidine with water. Both relative energies show that the reaction of amidine with OH^- is both thermodynamically and kinetically favourable and the reaction of amidine with H_2O is thermodynamically but kinetically not very favourable.

3.6 Optimized Structure of Cytosine

The optimized structure (c.f. Fig.3.9 for numbering) of cytosine using STO-3G and 3-21G basis set are reported in Table 16 and compared with the experimental crystal structure [53]. The computed bond distances in the ring are longer than the experimental distances, except for N4C1, C3C2, N3C3 and O1C4 which are shorter by 0.04, 0.01, 0.003 and 0.02 Å , respectively. The computed bond angles within the ring are in fairly good agreement with the experimental values, the largest discrepancy being a difference of 5.4^0 in the H1N2C1 angle.

3.7 The Reaction of Cytosine with OH^-

The analogous structures for the deamination of amidine with hydroxide were investigated for the deamination of cytosine with hydroxide. The structures involved in the reaction of cytosine with hydroxide are shown in Figs.3.10-3.13. The corresponding geometries for all structures, cytosine-hydroxide complex (III_a and III_b), intermediate (III_c), transition state (III_d) and uracil ion-ammonia complex (III_e) are given in Tables 17-21 and the total computed energies are reported in Table 22.

The structures for the cytosine-hydroxide complex (III_a and III_b) were constrained to have C_s symmetry. These two structures differ only in the orientation of the OH^-. Structure III_a is slightly more stable than III_b by 5.1 kJ/mol.

Figure 3.9

Numbering for the cytosine structure.

H1
H2
N2
C1
H3
N4
C2
C4
C3
O1
N3
H4
H5

Table 16 STO-3G and 3-21G optimized geometries (under C_s symmetry) for cytosine

Bond lengths	Experimental	STO-3G	3-21G
N2C1	1.330	1.3918	1.3439
C2C1	1.424	1.4715	1.4431
N4C1	1.337	1.3095	1.2973
H1N2	0.87	1.0141	0.9974
H2N2	0.86	1.0132	0.9943
H3C2	0.87	1.0764	1.0667
C3C2	1.342	1.3298	1.3376
H4C3	1.01	1.0876	1.0701
N3C3	1.357	1.3876	1.3544
H5N3	0.88	1.0198	0.9984
C4N4	1.364	1.4404	1.3686
O1C4	1.234	1.2209	1.2113
N3C4	1.374	1.4453	1.4147

Table 16 continued

Bond angles			
C2C1N2	--	116.98	119.15
N4C1N2	118.2	116.90	118.08
H1N2C1	123.0	119.16	117.64
H2N2C1	124.0	121.60	122.76
H3C2C1	123.0	120.37	121.78
C3C2C1	117.3	117.00	116.08
H4C3C2	122.0	123.76	122.43
N3C3C2	120.1	119.99	120.80
H5N3C3	117.0	119.88	121.33
C4N4C1	119.9	117.24	122.23
O1C4N3	119.8	119.51	118.90

Figure 3.10

The structures for the cytosine-hydroxide complex (reactant).

O2
H1
H6
H2
N2
C1
H3
N4
C2
C4
C3
O1
N3
H4
H5

III a

H6
O2
H1
N2
H2
C1
H3
N4
C2
C4
C3
O1
N3
H4
H5

III b

Figure 3.11

The structure for the cytosine-hydroxide intermediate.

IIIc

Figure 3.12

The structure of the transition state for the hydrogen rearrangement (cytosine-hydroxide).

III d

Figure 3.13

The structure for the uracil ion-ammonia complex (product).

III.

Table 17 STO-3G and 3-21G optimized geometries (under C_s symmetry) for the cytosine-hydroxide complex (III_s)

Bond lengths	STO-3G	3-21G
N2C1	1.4151	1.3848
C2C1	1.4945	1.4554
N4C1	1.3270	1.3012
H1N2	1.1266	1.0515
H3C2	1.0782	1.0649
C3C2	1.3224	1.3311
H4C3	1.0860	1.0712
N3C3	1.3985	1.3668
H5N3	1.0184	0.9968
C4N4	1.4310	1.3733
O1C4	1.2288	1.2194
O2C1	3.1378	3.5375
H6O2	1.0124	0.9943

Table 17 continued

Bond angles		
C2C1N2	114.44	114.89
N4C1N2	123.88	124.66
H1N2C1	109.33	117.16
H3C2C1	118.06	118.34
C3C2C1	120.11	118.22
H4C3C2	124.91	123.75
N3C3C2	119.18	119.81
H5N3C3	120.18	121.53
C4N4C1	119.19	122.90
O1C4N3	117.97	118.32
O2C1N4	83.90	94.81
H6O2C1	127.26	115.36
Torsion angles		
H1N2C1C2	+143.06	+133.06
H2N2C1C2	-143.06	-133.06

Table 18 STO-3G and 3-21G optimized geometries (under C_s symmetry) for the cytosine-hydroxide complex (III_b)

Bond lengths	STO -3G	3-21G
N2C1	1.4067	1.3828
C2C1	1.4953	1.4543
N4C1	1.3322	1.3039
H1N2	1.1404	1.0606
H3C2	1.0783	1.0649
C3C2	1.3223	1.3312
H4C3	1.0860	1.0713
N3C3	1.3985	1.3669
H5N3	1.0183	0.9967
C4N4	1.4267	1.3710
O1C4	1.2298	1.2201
O2C1	3.0990	3.4078
H6O2	1.0047	0.9896

Table 18 continued

Bond angles		
C2C1N2	114.80	115.27
N4C1N2	123.72	124.29
H1N2C1	108.74	114.92
H3C2C1	118.06	118.37
C3C2C1	120.16	118.17
H4C3C2	124.86	123.69
N3C3C2	119.21	119.87
H5N3C3	120.16	121.53
C4N4C1	119.26	122.94
O1C4N3	117.78	118.27
O2C1N4	82.90	88.32
H6O2C1	85.15	83 06
Torsion angles		
H1N2C1C2	+144.28	+135.81
H2N2C1C2	-144.28	-135.81

Table 19 STO-3G and 3-21G optimized geometry for cytosine-hydroxide intermediate (III_c)

Bond lengths	STO-3G	3-21G
N2C1	1.5190	1.4895
C2C1	1.5330	1.5019
N4C1	1.4722	1.4296
H1N2	1.0368	1.0105
H2N2	1.0368	1.0100
H3C2	1.0792	1.0686
C3C2	1.3130	1.3198
H4C3	1.0867	1.0745
N3C3	1.4268	1.3782
H5N3	1.0295	0.9937
C4N4	1.3592	1.3182
O1C4	1.2455	1.2451
O2C1	1.4537	1.4609
H6O2	0.9906	0.9689

Table 19 continued

Bond angles		
C2C1N2	107.64	107.94
N4C1N2	108.42	112.15
H1N2C1	103.93	107.16
H2N2C1	104.08	107.50
H3C2C1	119.47	117.95
C3C2C1	117.89	119.37
H4C3C2	123.61	122.43
N3C3C2	120.02	121.39
H5N3C3	111.54	122.12
C4N4C1	115.81	123.85
O1C4N3	113.74	114.91
O2C1N2	105.46	103.77
H6O2C1	101.30	102.47

Table 19 continued

Torsion angles		
N4C1 N2C2	233.36	231.42
H1N2C1N4	50.50	67.61
H2N2C1N4	304.24	311.21
H3C2C1N2	27.67	55.21
C3C2C1H3	-179.19	-179.48
H4C3C2H3	0.75	-0.10
N3C3C2H4	177.50	179.99
H5N3C3H4	-25.51	0.66
C4N4C1C2	30.80	-1.53
O1C4N3H5	23.82	0.08
O2C1N2N4	-117.86	-115.53
H6O2C1N4	-11.69	2.22

Table 20 STO-3G and 3-21G optimized geometry for cytosine-hydroxide transition state (III_d)

Bond lengths	STO-3G	3-21G
N2C1	2.2555	2.2292
C2C1	1.4999	1.4547
N4C1	1.3203	1.2998
H1N2	1.0618	1.0416
H2N2	1.0512	1.0282
H3C2	1.0784	1.0629
C3C2	1.3177	1.3227
H4C3	1.0847	1.0706
N3C3	1.4185	1.3836
H5N3	1.0229	0.9952
C4N4	1.4343	1.3664
O1C4	1.2273	1.2256
O2C1	1.4072	1.3838
H6N2	1.6895	1.9101

Table 20 continued

Bond angles		
C2C1N2	86.89	87.84
N4C1N2	114.92	112.59
H1N2C1	127.53	131.64
H2N2C1	92.83	96.67
H3C2C1	118.52	119.11
C3C2C1	118.88	117.38
H4C3C2	124.50	123.75
N3C3C2	119.97	120.02
H5N3C3	116.32	121.38
C4N4C1	117.41	122.39
O1C4N3	118.13	118.47
O2C1N2	82.78	87.31
H6N2C1	50.25	51.33

Table 20 continued

Torsion angles		
N4C1N2C2	234.58	236.65
H1N2C1N4	89.11	102.63
H2N2C1N4	348.22	348.82
H3C2C1N2	63.43	64.01
C3C2C1H3	-177.96	-176.61
H4C3C2H3	0.95	2.32
N3C3C2H4	176.57	181.18
H5N3C3H4	-27.02	-0.84
C4N4C1C2	-10.05	-5.86
O1C4N3H5	20.22	0.76
O2C1N2N4	-116.88	-119.37
H6N2C1N4	119.83	117.52

Table 21 STO-3G and 3-21G optimized geometries for the uracil ion-ammonia complex (III_e)

Bond length	STO-3G	3-21G
O1C1	1.2588	1.2495
C2C1	1.5172	1.4809
N1C1	1.3843	1.3445
C4N1	1.3890	1.3371
N2C4	1.4657	1.4229
C3N2	1.3959	1.3675
O2C4	1.2386	1.2333
H3N2	1.0181	0.9951
H2C3	1.0871	1.0731
H1C2	1.0787	1.0680
H5O1	1.5894	1.8843
N3H5	1.0589	1.0182
H6N3	1.0341	1.0074
H7N3	1.0342	1.0074

Table 21 continued

Bond angles		
C2C1O1	118.55	118.67
N1C1C2	120.73	118.34
C4N1C1	118.89	123.82
N2C4N1	118.80	116.81
C3N2C4	122.17	121.95
O2C4N1	125.57	127.20
H3N2C4	118.02	115.94
H2C3N2	116.07	116.41
H1C2C1	118.79	118.70
H5O1C1	122.00	115.90
N3H5O1	172.78	179.89
H6N3H5	103.73	110.71
H7N3H5	103.73	110.68

Table 21 continued

Torsion angles		
N1C1C2O1	180.12	180.05
C4N1C1C2	-0.28	- 0.11
N2C4N1C1	-0.27	0.16
C3N2C4N1	0.86	-0.13
O2C4N1C1	179.82	180.13
H3N2C4O2	-0.87	0.03
H2C3N2H3	0.88	-0.12
H1C2C1N1	180.50	180.03
H5O1C1C2	3.57	0.07
N3H5O1C1	155.67	156.93
H6N3H5O1	53.33	83.64
H7N3H5O1	-53.11	-58.38

Table 22 Total energies for the cytosine-hydroxide structures[a] III_a – III_e

Compound (OH^-)	Energies (hartrees)	
	STO-3G	3-21G
III_a	-461.659987	-465.318893
III_b	-461.661131	-465.320789
III_c	-461.837603	-465.403087
III_d	-461.709818	-465.323221
III_e	-461.833990	-465.449316

a. See Figures 3.10-3.13 for the structures

The O2-C1 bond length in III_a is calculated to be 3.54 $\dot{A}$. The calculated O2-C1 bond length in III_b is found to be 3.41 $\dot{A}$. Comparing both complex, III_a and III_b, the O2-C1 bond length in III_a is longer than that in III_b by 0.13 $\dot{A}$.

The intermediate of cytosine with hydroxide is analogous to that formed by amidine and hydroxide. This structure shows that nucleophilic attack by hydroxide at position 4 from below the plane of the cytosine ring. The calculated O2-C1, C1-N2 and H6-O2 bond lengths are found to be 1.46, 1.49 and 0.97 $\dot{A}$, respectively. Comparison of the O2-C1 and C1-N2 bond lengths show that the O2-C1 bond length in III_c is shorter than that in III_a by 2.08 $\dot{A}$ and that the calculated C1-N2 bond length in III_c is longer than that in III_a by 0.11 $\dot{A}$.

The transition state (III_d) for the 1,3-hydrogen rearrangement indicates that the O2-C1 bond is shortening and C1-N2 bond is breaking forming NH_3 as a leaving group. The calculated O2-C1 and C1-N2 bond lengths are found to be 1.38 and 2.23 $\dot{A}$, respectively. Again this structure is a first order saddle point on the potential energy surface. Comparison of the O2-C1 and C1-N2 bond lengths in the transition state (III_d) and the intermediate (III_c) show that the O2-C1 bond length is getting shorter in going from III_c to III_d by 0.08 $\dot{A}$ and C1-N2 bond length increases in going from III_c to III_d by 0.74 $\dot{A}$. Another noticeable change in geometry is the narrowing of the O2C1N2 bond angle, which decreases considerably in going from intermediate to transition state.

The structure of the product (III_e), uracil ion-ammonia complex, shows that the NH_3 is hydrogen bonded to the oxygen with a typically linear hydrogen

bond (the H3H5O1 angle is close to 180^0).

Comparing the structures of III_c , III_d and III_e shows that the O2-C1 bond is getting shorter in going from the single O2-C1 bond in III_c (1.46 $\dot{A}$) to the O-C double bond (1.25 $\dot{A}$) in III_e .

The relative energies of cytosine-hydroxide complex, intermediate, transition state and product are summarized in Table 23. The results indicate that the second step has a barrier of 209 kJ/mol and that the product is more stable than the cytosine-hydroxide complex (III_a) by 337 kJ/mol. From both Table 23 and Fig.3.14 it is indicated that the reaction of cytosine with hydroxide is thermodynamically favourable but kinetically not very favourable.

3.8 The Reaction of Cytosine with H_2O

The analogous structures for the deamination of amidine with water were investigated for the deamination of cytosine with water. The structures involved in the reaction of cytosine with water are shown in Figs.3.15-3.18. The corresponding geometries for all structures, cytosine-water complex (IV_a and IV_b), intermediates (IV_c and IV_d), transition state (IV_e) and uracil-ammonia

Table 23 Relative energies for the cytosine-hydroxide structures[a] III_a -- III_e

Compound (OH^-)	Relative energies (kJ/mol)	
	STO-3G	3-21G
III_a	458.86	342.57
III_b	453.66	337.46
III_c	-9.41	121.38
III_d	326.03	331.05
III_e	0.00	0.00

a. See Fig. 3.14

Figure 3.14

Schematic energy profile for the deamination of cytosine with OH^-. $Cy+OH^-$ and Ura^-+NH_3 energies are just the sum of the energies $E(Cy+OH^-)=E(Cy)+E(OH^-)$ and $E(Ura^-+NH_3)=E(Ura^-)+E(NH_3)$. Where $R_{Cy...OH}$ corresponds to III_b, $I_{Cy...OH}$ to III_c, $T_{Cy...OH}$ to III_d and $P_{Ura...NH.}$ to III_e.

KJ/MOL
Cy +OH⁻
400
94
Rcy...OH⁻
Tcy...OH⁻
300
216
209
200
Icy...OH⁻
100
337
Ura⁻+NH3
55
Pura⁻...NH3
0

Figure 3.15

The structures for the cytosine-water complex (reactant).

H6
O2
H1
H2
N2
H7
C1
H3
N4
C2
C4
C3
O1
H4
N3
H5

IVa

O2
H1
H7
H2
N2
H6
C1
H3
N4
C2
C4
C3
O1
H4
N3
H5

IVb

Figure 3.16

The structures for the cytosine-water intermediate.

H1
H2
N2
C1
O2
H7
N4
H3
C2
H6
C3
C4
H4
N3
O1
H5

IVc

H1
H2
N2
H6
C1
O2
H3
H7
N4
C2
C3
H4
C4
N3
O1
H5

IVd

Figure 3.17

The structure of the transition state for the hydrogen rearrangement (cytosine-water).

IVe

Figure 3.18

The structure for the uracil-ammonia complex (product).

IVf

complex (IV_f) are reported in Tables 24-29 and the total computed energies are given in Table 30.

Two structures for cytosine-water complex (IV_a and IV_b) corresponding to two different orientations of H_2O are obtained. The structure of IV_a shows that the water is out of the plane of the cytosine ring and two hydrogens, H1 and H2, are below the plane of the cytosine ring. The C1-O2 bond length in IV_a is calculated to be 3.19 $\dot{A}$ compared to 3.55 $\dot{A}$ in IV_b. Comparing complexes IV_a and IV_b, the C1-O2 bond length in IV_a is shorter than that in IV_b by 0.45 $\dot{A}$.

In the two comformers of the intermediate (IV_c and IV_d) resulting from addition of hydroxide at position 4 and protonation at position 3, the O2-C1 and H6-O2 bond lengths in IV_c and IV_d are calculated to be 1.43 and 0.97 $\dot{A}$, respectively. The H6O2C1 angles in both intermediates (IV_c and IV_d) differ only 1.2^0 but the difference in the O2C1N2 angles are as large as 4^0. The calculated H6O2C1N4 torsion angle are 63^0 and 162^0 in IV_c and in IV_d, respectively. The energy for cytosine-water complex (IV_a) is lower than both of those two intermediates (IV_c and IV_d), by 28 and 30 kJ/mol, respectively.

The structure of the transition state for the [1,3] hydrogen rearrangement again is found to be a first order saddle point on the potential energy surface. The structure indicates that the C1-O2 bond is shortening and C1-N2 bond is breaking forming NH_3 as a leaving group. The O2-C1, C1-N2 and H6-N2 bond lengths in transition state (IV_e) are calculated to be 1.36, 1.67 and 1.22 $\dot{A}$,

Table 24 STO-3G and 3-21G optimized geometries for the cytosine-water complex (IV_a)

Bond lengths	STO-3G	3-21G
N2C1	1.4073	1.3310
C2C1	1.4699	1.4434
N4C1	1.3143	1.3134
H1N2	1.0415	1.0143
H2N2	1.0232	0.9948
H3C2	1.0767	1.0667
C3C2	1.3301	1.3369
H4C3	1.0876	1.0701
N3C3	1.3886	1.3548
H5N3	1.0199	0.9984
C4N4	1.4395	1.3660
O1C4	1.2213	1.2126
O2C1	3.1356	3.1880
H6O2	0.9859	0.9660
H7O2	0.9891	0.9835

Table 24 continued

Bond angles		
C2C1N2	118.11	120.22
N4C1N2	116.72	118.15
H1N2C1	111.17	117.16
H2N2C1	113.52	121.77
H3C2C1	120.06	121.31
C3C2C1	117.41	116.57
H4C3C2	123.81	122.34
N3C3C2	120.08	120.99
H5N3C3	119.93	121.39
C4N4C1	118.04	122.64
O1C4N3	119.95	119.20
O2C1N4	61.91	59 77
H6O2C1	115.60	120.75
H7O2H6	101.31	108.56

Table 24 continued

Torsion angles		
N4C1N2C2	182.77	180.06
H1N2C1C2	195.67	180.71
H2N2C1C2	-33.33	-1.49
H3C2C1N2	3.35	0.10
C3C2C1H3	-179.53	-180.12
H4C3C2H3	0.32	0.05
N3C3C2H4	180.07	180.04
H5N3C3H4	0.06	0.03
C4N4C1C2	-1.15	-0.24
O1C4N3H5	0.00	-0.45
O2C1N4N2	14.30	0.89
H6O2C1N2	85.37	99.80
H7O2H6C1	-54.92	-50.17

Table 25 STO-3G and 3-21G optimized geometries for the cytosine-water complex (IV_b)

Bond lengths	STO-3G	3-21G
N2C1	1.4751	1.4272
C2C1	1.4710	1.4377
N4C1	1.2940	1.2795
H1N2	1.0343	1.0044
H3C2	1.0772	1.0661
C3C2	1.3302	1.3399
H4C3	1.0875	1.0696
N3C3	1.3903	1.3567
H5N3	1.0203	0.9997
C4N4	1.4598	1.3931
O1C4	1.2183	1.2064
O2C1	3.5922	3.5492
H6O2	0.9871	0.9668
H7O2	0.9873	0.9699

Table 25 continued

Bond angles		
C2C1N2	118.91	119.13
N4C1N2	115.60	118.01
H1N2C1	107.43	115.19
H3C2C1	120.21	121.12
C3C2C1	117.56	116.68
H4C3C2	124.12	123.22
N3C3C2	119.85	120.17
H5N3C3	119.98	121.00
C4N4C1	117.97	122.08
O1C4N3	120.92	120.59
O2C1N4	60.72	58.72
H6O2C1	65.63	72 59
H7O2H6	99.31	104.60

Table 25 continued

Torsion angles		
N4C1N2C2	179.95	179.96
H1N2C1C2	56.06	67.45
H3C2C1N2	0.00	-0.03
C3C2C1H3	-180.00	-180.00
H4C3C2H3	0.01	-0.03
N3C3C2H4	179.99	180.01
H5N3C3H4	0.04	-0.01
C4N4C1C2	-0.03	0.01
O1C4N3H5	-0.02	0.14
O2C1N4N2	-0.87	-0.58
H6O2C1N2	181.39	180.16
H7O2H6C1	-1.00	-0.36

Table 26 STO-3G and 3-21G optimized geometries for the cytosine-water intermediate (IV_c)

Bond lengths	STO-3G	3-21G
N2C1	1.4868	1.4387
C2C1	1.5293	1.4981
N4C1	1.4938	1.4663
H7N4	1.0263	0.9998
H1N2	1.0334	1.0016
H2N2	1.0326	1.0000
H3C2	1.0785	1.0670
C3C2	1.3136	1.3175
H4C3	1.0859	1.0698
N3C3	1.4191	1.3878
H5N3	1.0206	0.9963
C4N4	1.4334	1.3646
O1C4	1.2195	1.2181
O2C1	1.4355	1.4304
H6O2	0.9903	0.9658

Table 26 continued

Bond angles		
C2C1N2	109.74	110.79
N4C1N2	110.39	113.10
H7N4C1	112.01	115.91
H1N2C1	106.89	113.72
H2N2C1	106.81	114.97
H3C2C1	116.12	115.73
C3C2C1	122.21	122.08
H4C3C2	122.72	122.23
N3C3C2	122.97	122.68
H5N3C3	117.84	121.08
C4N4C1	122.04	129.61
O1C4N3	123.12	122.32
O2C1N2	104.09	102.83
H6O2C1	104.70	109.78

Table 26 continued

Torsion angles		
N4C1N2C2	239.51	237.57
H7N4C1N2	-72.41	-56.09
H1N2C1N4	64.35	86.87
H2N2C1N4	311.93	310.31
H3C2C1N2	45.08	53.50
C3C2C1H3	-180.69	-179.17
H4C3C2H3	0.32	0.57
N3C3C2H4	181.47	180.38
H5N3C3H4	23.95	0.08
C4N4C1C2	33.10	0.97
O1C4N3H5	-10.54	0.31
O2C1N2N4	-118.49	-117.38
H6O2C1N4	73.52	63.15

Table 27 STO-3G and 3-21G optimized geometries for the cytosin-water intermediate (IV_d)

Bond lengths	STO-3G	3-21G
N2C1	1.4908	1.4522
C2C1	1.5310	1.4986
N4C1	1.4868	1.4498
H7N4	1.0281	0.9995
H1N2	1.0335	1.0059
H2N2	1.0321	1.0033
H3C2	1.0787	1.0681
C3C2	1.3128	1.3169
H4C3	1.0861	1.0698
N3C3	1.4223	1.3855
H5N3	1.0223	0.9962
C4N4	1.4403	1.3639
O1C4	1.2183	1.2174
O2C1	1.4367	1.4330
H6O2	0.9894	0.9669

Table 27 continued

Bond angles		
C2C1N2	109.01	108.64
N4C1N2	111.08	113.89
H7N4C1	110.39	115.83
H1N2C1	107.18	113.19
H2N2C1	107.30	113.47
H3C2C1	116.64	116.75
C3C2C1	121.88	121.46
H4C3C2	122.80	122.27
N3C3C2	122.92	122.51
H5N3C3	116.63	121.13
C4N4C1	119.09	128.86
O1C4N3	123.09	122.23
O2C1N2	107.54	106.15
H6O2C1	103.57	108.60

Table 27 continued

Torsion angles		
N4C1N2C2	238.59	237.29
H7N4C1N2	-76.00	-63.76
H1N2C1N4	49.05	60 84
H2N2C1N4	296.71	291.13
H3C2C1N2	43.79	51.78
C3C2C1H3	-181.28	-180.26
H4C3C2H3	-0.23	0.25
N3C3C2H4	182.38	180.31
H5N3C3H4	28.52	1.85
C4N4C1C2	35.34	6.25
O1C4N3H5	-13.53	-0.13
O2C1N2N4	-116.25	-115.34
H6O2C1C2	72.14	72.30

Table 28 STO-3G and 3-21G optimized geometries for cytosin-water transition state (IV_e)

Bond lengths	STO-3G	3-21G
N2C1	1.8278	1.6664
C2C1	1.5172	1.4863
N4C1	1.4434	1.4251
H7N4	1.0207	0.9998
H1N2	1.0329	1.0090
H2N2	1.0334	1.0094
H3C2	1.0783	1.0675
C3C2	1.3187	1.3183
H4C3	1.0866	1.0697
N3C3	1.4094	1.3898
H5N3	1.0179	0.9968
C4N4	1.4224	1.3672
O1C4	1.2199	1.2154
O2C1	1.3253	1.3648
H6N2	1.1500	1.2224

Table 28 continued

Bond angles		
C2C1N2	105.59	108.10
N4C1N2	106.99	107.67
H7N4C1	116.22	115.56
H1N2C1	121.56	115.70
H2N2C1	121.80	114.67
H3C2C1	117.34	117.37
C3C2C1	121.03	119.95
H4C3C2	123.12	122.90
N3C3C2	122.42	121.92
H5N3C3	120.00	120.75
C4N4C1	126.76	127.63
O1C4N3	123.40	122.44
O2C1N2	86.59	91.51
H6N2C1	67.17	71.52

Table 28 continued

Torsion angles		
N4C1N2C2	241.92	239.20
H7N4C1N2	-72.44	-82.91
H1N2C1N4	227.51	229.84
H2N2C1N4	368.63	361.74
H3C2C1N2	78.71	75.61
C3C2C1H3	-180.44	-181.41
H4C3C2H3	0.07	-1.93
N3C3C2H4	179.01	178.91
H5N3C3H4	-8.23	-2.18
C4N4C1C2	-23.64	-15.12
O1C4N3H5	1.53	0.22
O2C1N2N4	-118.08	-117.49
H6N2C1C2	236.38	237.33

Table 29 STO-3G and 3-21G optimized geometries for the uracil-ammonia complex (IV_f)

Bond lengths	STO-3G	3-21G
O1C1	1.2250	1.2192
C2C1	1.4948	1.4549
N1C1	1.4314	1.3897
C4N1	1.4217	1.3752
N2C4	1.4238	1.3755
C3N2	1.4047	1.3783
O2C4	1.2188	1.2115
H3N2	1.0184	0.9978
H2C3	1.0870	1.0692
H1C2	1.0789	1.0686
H4N1	1.0205	1.0009
H5O1	1.9090	2.1314
N3H5	1.0339	1.0066
H6N3	1.0317	1.0044
H7N3	1.0318	1.0044

Table 29 continued

Bond angles		
C2C1O1	126.17	125.59
N1C1C2	113.70	114.30
C4N1C1	127.27	127.78
N2C4N1	112.56	113.29
C3N2C4	123.31	123.07
O2C4N1	123.95	123.54
H3N2C4	116.78	115.95
H2C3N2	114.49	115.29
H1C2C1	117.41	117.18
H4N1C1	117.08	116.49
H5O1C1	114.03	109.32
N3H5O1	169.75	155.27
H6N3H5	104.55	112.18
H7N3H5	104.56	112.12

Table 29 continued

Torsion angles		
N1C1C2O1	179.94	180.02
C4N1C1C2	0.12	0.00
N2C4N1C1	0.26	0.00
C3N2C4N1	-0.62	0.00
O2C4N1C1	180.21	179.99
H3N2C4O2	0.62	-0.02
H2C3N2H3	-0.73	0.00
H1C2C1N1	179.77	180.11
H4N1C1O1	0.15	-0.02
H5O1C1C2	-0.29	0.07
N3H5O1C1	1.76	-2.40
H6N3H5O1	124.82	119.85
H7N3H5H6	108.75	126.22

Table 30 Total energies for the cytosine-water structures[a] IV_a – IV_f

Compound (H_2O)	Energies (hartrees)	
	STO-3G	3-21G
IV_a	-462.531615	-466.034034
IV_b	-462.507550	-465.986737
IV_c	-462.568154	-466.023286
IV_d	-462.571961	-466.022668
IV_e	-462.450760	-465.956344
IV_f	-462.547534	-466.051247

a. See Figures 3.15-3.18 for the structure

respectively. Comparsion of the structures, IV_d, IV_e and IV_f, show the O2-C1 bond length is getting shorter in going from the single O2-C1 bond in IV_d (1.43 $\dot{A}$) to the O-C double bond (1.22 $\dot{A}$) in IV_f. From Table 27-28 it is shown that the O2C1N2 bond angle decreases in going from intermediate to transition state by about 15^0.

The relative energies of the cytosine-water complex, intermediates, transition state and product are summarized in Table 31. The results indicate that the second step has a barrier of 174 kJ/mol and that the product is more stable than VI_a by 45 kJ/mol. Fig. 3.19 illustrates that the reaction of cytosine with water is thermodynamically favourable but kinetically not very favourable.

3.9 Effect of Basis Sets on the Deamination of Cytosine with OH^- and H_2O

Comparing some of the more important geometrical parameters of the STO-3G and 3-21G optimized geometries a number of trends appear. The 3-21G basis set consistently gives shorter N-C, H-C, H-N, O1-C4 and C1-C2 bond lengths and longer C3-C2 and O2-C1 bond lengths than those from STO-3G basis set of the order of 0.010-0.070 $\dot{A}$. The STO-3G and 3-21G bond angles differ by less than 4^0, except in the transition state structures where the angle varies by as much as 7^0. In addition, the most important change in the geometry of the transition

Table 31 Relative energies for the cytosine-water structures[a] IV_a -- IV_f

Compound (H_2O)	Relative energies (kJ/mol)	
	STO-3G	3-21G
IV_a	[illegible].80	45.19
IV_b	104.98	169.37
IV_c	-54.14	73.41
IV_d	-64.13	75.03
IV_e	254.08	249.18
IV_f	0.00	0.00

a. See Fig. 3.19

Figure 3.19

Schematic energy profile for the deamination of cytosine with H_2O. $Cy+H_2O$ and $Ura+NH_3$ energies are just the sum of the energies $E(Cy+H_2O)=E(Cy)+E(H_2O)$ and $E(Ura+NH_3)=E(Ura)+E(NH_3)$. Where $R_{Cy...H_2O}$ corresponds to IV_a, $I_{Cy...H_2O}$ to IV_d, $T_{Cy...H_2O}$ to IV_e and $P_{Ura...NH_3}$ to IV_f.

KJ/MOL
200
125
100
75
50
25
0
Cy + H2O
84
Rcy...H2O
45
Icy...H2O
30
Tcy...H2O
174
Purа...NH3
Ura. + NH3
42

state involves a considerable narrowing of the O2C1N2 bond angle which increases in going from the STO-3G to 3-21G basis set, where the O2C1N2 bond angle increases by 5^0.

3.10 Comparison of the Deamination of Cytosine with OH^- and H_2O

The relative energies for the deamination of cytosine with OH^- and H_2O are summarized in Table 23 and Table 31. The schematic energy profile for the deamination of cytosine with hydroxide and water are shown in Fig.3.14 and Fig.3.19, respectively. Comparison of the results for the deamination of cytosine with hydroxide with the those of the deamination of cytosine with water indicate that the energy of cytosine-hydroxide intermediate is more stable than that of the cytosine-hydroxide complex by 216 kJ/mol whereas the cytosine-water intermediate is 30 kJ/mol higher in energy than the cytosine-water complex. This large difference can be explained according to the scheme given in Fig.3.20. As can be seen from this figure, the cytosine-hydroxide adduct possesses delocalization of the negative charge and may be alternatively represented as having two contributing canonical forms. By contrast, the cytosine-water intermediate is destabilized relative to the hydroxide case, as the intermediate can be seen to show no charge, and only one contributing canonical form. The second step with OH^- has a barrier of 209 kJ/mol and with H_2O has a barrier of 174 kJ/mol.

Figure 3.20

Reaction schemes for the formation of the cytosine-water and cytosine-hydroxide intermediates.

H2O
NH2
N
O
N
H
H
H
H
NH2
OH
HN
O
N
H
H
H

OH:
NH2
N
O
N
H
H
H
H
NH2
OH
N
O
N
H
H
H

From both reactions of cytosine with hydroxide and water, the results indicate that both reactions are thermodynamically favourable but kinetically not very favourable. Comparing both results show that the barrier for the deamination of cytosine with water is lower than the deamination of cytosine with hydroxide by 35 kJ/mol. (see Figs. 3.14 and 3.19).

3.11 The Comparison of the Deamination of Amidine and Cytosine with OH^-

The schematic energy profile for the deamination of amidine and cytosine with hydroxide are given in Figs 3.6 and 3.14, respectively. The relative energies for the deamination of amidine and cytosine with OH^- are summarized in Table 7 and Table 23, respectively. The structures for the cytosine with hydroxide are good agreement with the those of amidine with hydroxide for all cases, except for the intermediates, I_c and III_c, where the HOCN torsion angle in I_c is 2^0 and the H6O2C1N4 torsion angle in III_c is -17^0, which differ by 19^0. The products, formamide ion-ammonia complex (I_e) and uracil ion-ammonia complex (III_e), are more stable than amidine-hydroxide complex and cytosine-hydroxide complex by 268 and 337 kJ/mol, respectively. The results for the deamination of amidine and cytosine with hydroxide indicate that the energy for the amidine-hydroxide intermediate is more stable than the amidine-hydroxide complex by 110 kJ/mol but the energy for cytosine-hydroxide intermediate is lower than the cytosine-

hydroxide complex by as much as 216 kJ/mol. The second step with amidine has a barrier of 60 kJ/mol and with cytosine has a barrier of 209 kJ/mol. The barrier for the deamination of amidine with hydroxide is substantially lower than the deamination of cytosine with hydroxide. The different relative energies for the deamination of cytosine and amidine can be explained as: cytosine-hydroxide adduct possesses delocalization of the negative charge (see Fig.3.20) and may be alternatively represented as having two contributing canonical forms. Comparing the reaction for the deamination of amidine and cytosine show that the reaction of amidine is both thermodynamically and kinetically favourable and that the reaction of cytosine is thermodynamically favourable but kinetically not very favourable.

3.12 Comparison of the Deamination of Cytosine and Amidine with H_2O

The schematic energy profile for the deamination of amidine and cytosine with water are shown in Fig.3.8 and Fig.3.19, respectively. The relative energies of the deamination of amidine and cytosine with H_2O are summarized in Tables 15 and 31, respectively. The structures for the cytosine with water are in good agreement with the those of amidine with water for all cases, except for the intermediates, (II_d and IV_d), where the HOCN torsion angle in II_d is 169^0 and the H6O2C1N4 torsion angle in IV_d is 162^0 which differ by 7^0. Both Figs., Fig.3.8

and Fig.3.19, show that the second step with cytosine has a barrier of 174 kJ/mol and with amidine has a barrier of 182 kJ/mol which differ by only 8 kJ/mol in this case. Both products, II_f and IV_f, are more stable than the amidine-water complex and cytosine-water complex by 42 and 45 kJ/mol, respectively. Comparing both relative energies show that the reactions for the deamination of amidine and cytosine with water are thermodynamically favourable but kinetically not very favourable. The results indicate that the relative energies are very similar for both reactions.

3.13 The Comparison of the Experimental and Calculated Results

Shapiro and Klein [10] stated that there are two possible mechanistic explanations for the specific hydrogen-ion and general acid-base-catalyzed deamination of the cytosine derivatives (see Fig.1.3 and Fig.1.4). One mechanism is analogous to the classical amide hydrolysis, wherein water attack on the protonated cytosine moiety could be enhanced by general base attack or water to give hydroxyl-ion attack on the protonated species at the 4-position with subsequent loss of ammonia. The proposed alternative method of partial saturation at the 5-6-position of the ring of the protonated cytosine by the nucleophilic addition of the buffer anion with subsequent deamination by water attack (see Fig.1.4) and subsequent loss of the acid conjugate of the buffer anion to give a uracil derivative

was preferred. This mechanism was considered to be more consistent with the information in the literature [8,16,21,53,54] which show that hydrolytic deamination is facilitated when cytosine derivatives were saturated in the 5-6-position. Notari [14] also preferred this latter explanation, initially with the additional restriction that the base or anion that attacks the 6-position (Fig.1.5), following the initial protonation at the 3-position must have a donative proton which can saturate the 5-position and promote the nucleophilic attack of water on the positively charged 4-position leading to the loss of ammonium ion and a resultant 5-6-saturated uracil which can readily form a uracil derivative by the elimination of the original attacking anion. The present calculated results, for the deamination of cytosine with hydroxide, show in the first step nucleophilic attack by hydroxide at position 4 of the cytosine and the second step followed by [1,3] hydrogen rearrangement and elimination of ammonia to give uracil-ion (see Fig.3.1). A similar mechanism for the deamination of cytosine with water is shown in Fig.3.2. In the first step, nitrogen gets protonated at position 3 and nucleophilic attack by hydroxide at position 4. The transition state (IV_e) for the [1,3] hydrogen rearrangement in the second step is analogous to that formed by cytosine and hydroxide and subsequent elimination of ammonia to give uracil.

For the conversion of cytosine to uracil, the experimentally determined activation energies at pH 1.0, 7.0 and 9.8 were 53, 51 and 97 kJ/mol, as reported by Garrett et al [22]. Based on the deamination of cytosine in different forms, the energy of activation was estimated to be 121 kJ/mol [23]. These data indicate

that a significant amount of conversion of cytosine to uracil occurs during heat denaturation of DNA by standard procedures. Ehrlich studied the heat-induced deamination conversion of cytosine to uracil residues in single-stranded DNA. They gave an approximate activation energy of 113 kJ/mol for the deamination of cytosine residues using Arrhenius plots of the temperature dependence of the deamination rate [25]. The calculated results for the deamination of cytosine indicated that the second step of cytosine with water has a barrier of 174 kJ/mol and of cytosine with hydroxide has a barrier of 209 kJ/mol.

Notari et al [16] investigated the intermolecular and intramolecular catalysis in the deamination of cytosine nucleoside. They reported that both general-acid and general-base catalysis are required in the deamination of cytosine. To continue the present work, the calculations can be done to involve general-acid and general-base catalysis. Additionally, a theoretical study of the first step for the deamination of cytosine should provide useful insight. It is expected that the calculated activation energy should be more consistent with experimental results.

Conclusions

The theoretical calculations of the deamination of cytosine are based on SCF ab initio calculations. In summarizing our results the following conclusions can be drawn:

(1) The optimized geometrical bond lengths for the deamination of cytosine at 3-21G basis set are generally shorter than that from STO-3G bond lengths. The optimized bond angles differ by less than 5^0.

(2) Comparing the deamination of amidine with hydroxide and water, the results indicate that the second step with hydroxide has a barrier of 60 kJ/mol and with water has a barrier of 182 kJ/mol. The barrier for the deamination of amidine with hydroxide is significantly lower than the deamination of amidine with water by 122 kJ/mol. From both reactions it is clear that the reaction of amidine with hydroxide is both thermodynamically and kinetically favourable and the reaction of amidine with water is thermodynamically favourable and kinetically not very favourable.

(3) By comparing of the deamination of cytosine with hydroxide and water the results indicate that the second step with hydroxide has a barrier 209 kJ/mol and with water has a barrier of 174 kJ/mol. The barrier for the deamination of cytosine with water is lower than the deamination of cytosine with hydroxide by 35 kJ/mol.

(4) Comparing the deamination of cytosine and amidine with hydroxide the results indicate that the barrier for the reaction of cytosine with hydroxide is higher than the barrier for reaction of amidine with hydroxide by 149 kJ/mol. The different relative energies can be explained as: cytosine-hydroxide adduct possesses delocalization of the negative charge and may be alternatively represented as having two contributing canonical forms.

(5) The structures for the cytosine with water are in good agreement with the those of amidine with water. The relative energies for both reactions indicate that the barrier for amidine with water is 182 kJ/mol compared to 174 kJ/mol for cytosine with water. The results indicate that for both reactions, the deamination of amidine and cytosine with water, the relative energies are very similar.

Appendix

Figure 1.1

The transition state structure for the amidine with hydroxide (one step mechanism).

H6
H7
2
H9
O8
C1
N4
H3
H5

Table 1 STO-3G and 3-21G optimized geometry (under C_s symmetry) for the amidine-hydroxide transition state (one step mechanism)

Bond lengths	STO-3G	3-21G
N2C1	3.8632	3.6903
H3C1	1.1429	1.1494
N4C1	1.3039	1.3001
H5N4	1.0660	1.0385
H3N2	1.0342	1.0107
O8C1	2.2400	2.4198
H9O8	1.8576	2.2621
Bond angles		
H3C1N2	136.55	125.72
N4C1N2	112.40	120.58
H5N4C1	114.48	116.18
H3N2C1	110.59	106.61
O8C1N4	67.23	68.77
H9O8C1	87.81	69.47
Torsion angles		
H6N2C1H3	+123.46	+122.66
H7N2C1H3	-123.46	-122.66

Table 2 Total energies for the amidine-hydroxide transition state (one step mechanism)[a]

Transition state	STO-3G	3-21G
	-221.143114	-222.973982

a. See Fig.1 for the structure

REFERENCES

1. T. Kotaka and R. L. Baldwin, J. Mol. Biol., 9, 323 (1964).

2. C.D. Laird, B. L. McConanghy and B. J. McCarthy, Nature, 224, 149 (1969).

3. Von Borstel RC On the origin of spontaneous mutation Jpn J Genet 44: 102-105 (1969).

4. J.W. Drake, The molecular basis of mutation Holden Day, San Francisco, pp 177-185 (1970).

5. N.J. Sargentini, K.C. Smith, Carcinogenesis 2, 9, 863 (1981).

6. H. Schuster, J. Mol. Biol., 3, 447 (1961).

7. D. M. Brown and P. Schell, ibid., 3, 709 (1961).

8. J. S. Ullman and B. J. McCarthy, Biochemistry Biophysica Acta 294,396 (1973).

9.H. E. Johns, J. C. LeBlanc and K. B. Freeman, ibid., 13, 849 (1965).

10. R. Shapiro and R. S. Klein, Biochemistry, 5, 2358 (1966).

11. D. M. Brown, and J. H. Phillips, J. Mol. Biol. 11, 663, (1965).

12. H. Schuster and G. Schramm, Z. Naturforsch., 136, 697 (1958).

13. D. O. Jordan, " The Chemistry of Nucleic Acids", Butterworths, Washington, D. C., p.65, (1960).

14. R. E. Notari, J. Pharm. Sci., 56, 804 (1967).

15. R. E. Notari, M. L. Chin and A. Cardoni, Tetrahedron Lett.,40, 3499 (1969).

16. R. E. Notari, M. L. Chin and A. Cardoni, J. Pharm. Sci., 59, 28 (1970).

17. H. Hayatsu, J. Amer. Chem. Soc., 91, 5693 (1969).

18. H. Hayatsu and M. Yano, Tetrahedron Lett., 9, 755 (1969).

19. R. Shapiro, R. E. Servis and M. Welcher, J. Amer. Chem. Soc., 92, 422 (1970).

20. H. Hayatsu, Y. Wataya and K. Kai, ibid., 92, 724 (1970).

21. H. Hayatsu, Y. Wataya and K. Kai and S. Iida, Biochemistry, 9, 2858 (1970).

22. E. R. Garrett and J. Tsau, J. Pharm. Sci., 61, 1052 (1972).

23. T. Lindahl and B. Nyberg, Biochemistry, 13, 3405 (1974).

24. T. Lindahl and B. Nyberg, Biochemistry, 11, 3610 (1972).

25. M. Ehrlich, K. F. Norris, R. Y.-H. Wang, K. C. Kuo and C. W. Gehrke, Bioscience Reports, 6, 387 (1986).

26. R. A. Poirier, D. Majlessi and T. J. Zielinski, J. Comput. Chem., 7, 464 (1986).

27. A. Lledos and J. Bertran, J. Mol. Structure, 107, 233 (1984).

28. (a) E. Schrodinger, A. Physik, Representative general texts include, 79, 361, (1926); (b) E. C. Kemble, Fundamental principles of Quantum Mechanics, McGraw-Hill, New York, (1965); (c) I. N. Levine, Quantum Chemistry, 3rd ed., Allyn and Bacon, Boston, (1983); (d) F. L. Pilar, Elementary Quantum Chemis-

try, McGraw-Hill, New York, (1986).

29. S. F. Boys, Proc. Roy. Soc. (London), A200, 542 (1950); For a readable discussion of the properties and use of gaussian functions in quantum mechanics, see: I. Shavitt in Methods in Computational Physics, 2, Wiley, New york, (1962).

30. For a discussion, see reference 28c, pp. 172.

31. P. Pulay, in Applications of Electronic Structure Theory, H. F. Schaefer, III, Ed., Plenum, New York, (1977).

32. J. A. Pople, R. Krishnan, H. B. Schlegel and J. S. Binkley, Int. J. Quantum Chem., 13s, 225 (1979).

33. R. Krishnan. H. B. Schlegel and J. A. Pople, J. Chem. Phys., 72 4244 (1980).

34. B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi and H. F. Schaefer, III, J. Chem. Phys., 72 4652 (1980).

35. S. Kato and K. Morokuma, Chem. Phys. Lett., 65, 19 (1979).

36. P. Pulay, Mol. Phys., 17, 197 (1969); D. M. Bishop and M. Randic, J. Chem. Phys., 44, 2480 (1966).

37. S. Bratoz, Coll. Int. C. N. R. S., 82, 287 (1959).

38. P. W. Payne, J. Chem. Phys., 65, 5, 1920, (1976).

39. R. A. Poirier, A. Yadav and M. R. Peterson, MUNGAUSS, Dept. Chem. Memorial Univ. of Newfoundland, St. John's, Nfld., Canada and Dept. Chem. Univ. of Toronto, Ont., Canada.

40. W. J. Hehre, R. F. Stewart and J. A. Pople, J. Chem. Phys., 51, 2657, (1969).

41. W. J. Hehre, R. Ditchfield, R. F. Stewart and J. A. Pople, J. Chem. Phys., 52, 2769 (1970).

42. W. J. Pietro, B. A. Levi, W. J. Hehre and R. F. Stewart, Inorg. Chem., 19, 2225 (1980).

43. W. J. Pietro, E. S. Blurock, R. F. Hout, Jr., W. J. Hehre, D. J. DeFrees and R. F. Stewart, Inorg. Chem., 20, 3650 (1981).

44. J. S. Binkley, J. A. Pople and W. J. Hehre, J. Amer. Chem. Soc., 102 939 (1980).

45. M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro and W. J. Hehre, J. Amer. Chem. Soc., 104, 2797 (1982).

46. P. C. Hariharan and J. A. Pople, Theoret. Chim. Acta (Berl.) 28, 213 (1973).

47. P. Pulay, Mol. Phys., 17, 197 (1969).

48. H. B. Schlegel, Ph. D. Thesis, Queen's Univ. (1975).

49. W. C. Davidon and L. Nazareth, Technical Memos 303 and 306, (1977).

50. W. C. Davidon and L. Nazareth, Applied Mathematics Division Argonne National Laboratories, Argonne, IL., Math. Prog., 9, 1 (1975).

51. T. Z. Zielinski, R. A. Poirier, M. R. Peterson and I. G. Csizmadia, J. Comput. Chem., 3, 477 (1982).

52. M. J. D. Powell, Subroutine VAO5AD, AERE Subroutine Library, Harwell, Didcot, Berkshire, U. K.

53. D. L. Barker and R. E. Marsh, Acta Crystallogr., 17, 1581 (1964).

54. V. Lisy and J. Skoda, Chem. Commun, 31, 3020 (1966).

55. E. R. Garrett and J. Tsau, J. Pharm. Sci., 61, 1052 (1972).