INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

A Theoretical Study of Hydrogen Bonding Involving Bio-Molecules

by

Kathryn N. Rankin

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
June 2001

DALHOUSIE UNIVERSITY

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "A Theoretical Study of Hydrogen Bonding Involving Bio-Molecules" by Kathryn Rankin in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: July 26, 2001

External Examiner:

Research Supervisor:

Examining Committee:

DALHOUSIE UNIVERSITY

Date: August 15, 2001

Author: Kathryn Nadine Rankin
Title: A Theoretical Study of Hydrogen Bonding Involving Bio-Molecules

Department or School: Department of Chemistry

Degree: Ph. D. Convocation: October Year: 2001

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

[redacted]
Signature of Author

To my family.

Chapter Four.

Chapter Five.

Theoretical calculations play a fundamental role in advancing our understanding of the properties of hydrogen-bonded systems. The plethora of experimental data on systems containing hydrogen bonds, obtained from spectroscopic experiments, is complemented by electronic structure theory calculations. Contemporary density functional theory (DFT) has become a powerful and economic method in electronic structure theory as it enables the effects of electron correlation to be accounted for in the calculation and is extensively utilized in this thesis for the study of hydrogen bonding in a variety of chemical and biochemical systems.

The first study in this thesis, Chapter Three, involves the use of DFT to investigate the hydrogen halide dimers, $(HF)_2$, $(HCl)_2$ and $(HBr)_2$. Emphasis is focused upon an assessment of the performance of the hybrid DFT methods in predicting the geometry, binding energy and vibrational properties of these dimers. Overall, the hybrid density functional methods adequately predict the properties of the halide dimers and the BHandHLYP and B1LYP density functional methods offer a competitive alternative to the popular B3LYP method. In Chapter Four, the ability of hydrogen bonds to act as catalysts is examined. In the study of the aminolysis of 6-chloropyrimidine, derivatives of uracil stabilize the transition structures by the formation of multiple hydrogen bonds, thus catalyzing the aminolysis reaction. This study highlights the catalytic potential of hydrogen bonding and the importance of a well-chosen hydrogen bond acceptor. In the following chapter, Chapter Five, the direct aldol reaction as catalyzed by proline is investigated to assess the feasibility of the proposed reaction mechanism and the potential role of the solvent. This study is a simple example illustrating the potential of small molecules, such as proline, to act as catalysts in biochemical reactions. Chapter Six examines the structural and energetic effect the introduction of electron-withdrawing substitutents on C^{γ} has on the *cis-trans* isomerization in N-acetylproline methylamide. This study illustrates the important role the intramolecular N-H···N hydrogen bond plays in the *cis-trans* isomerization and the catalytic potential of hydrogen bonds.

Ψ	Wave function	$\hat{H}_{total}$	Total Hamiltonian
E	Energy	$\hat{T}$	Kinetic energy operator
$\hat{V}$	Potential energy operator	$\hat{H}_{elect}$	Electronic Hamiltonian
χ_i	i^{th} spin orbital	ψ_i	i^{th} molecular orbital
Φ	Trial wave function	$\hat{F}$	Fock operator
ε_i	Orbital energy of the i^{th} molecular orbital	$\hat{H}_{core}$	Core-electron Hamiltonian
$\hat{J}_j$	Coulomb operator	$\hat{K}_j$	Exchange operator
$c_{\mu i}$	Expansion coefficient	ϕ	Basis function
$F_{\mu\upsilon}$	Elements of Fock matrix	$P_{\mu\upsilon}$	Elements of density matrix
$S_{\mu\upsilon}$	Elements of overlap matrix	$(\mu\upsilon \mid \lambda\sigma)$	Two-electron integral
$\hat{H}_0$	Zero-order Hamiltonian	V	Pertubation
$\upsilon(r)$	External potential	ρ	Electron density
$T[\rho]$	Kinetic energy functional	$V_{ee}[\rho]$	Electron-electron repulsion functional
v_{eff}	Effective potential	E_{xc}	Exchange-correlation energy functional
U_{XC}^{λ}	Potential energy of exchange	λ	Coupling strength parameter
μ	Euler-Lagrange multiplier	φ_μ	μ^{th} atomic orbital
g_k	Primitive Gaussian functions	$d_{\mu k}$	Contraction coefficients
h	Planck's constant	m	Mass of particle
r	Distance between an electron and nuclei (or electron)	R	Nuclear-nuclear distance
Z	Nuclear charge	ε	Dielectric constant
D_0	Dissociation Energy	μ	Dipole moment

DFT	Density functional theory	HF	Hartree-Fock
MO	Molecular orbital	AO	Atomic orbital
SCF	Self-consistent field	RHF	Restricted Hartree-Fock
UHF	Unrestricted Hartree-Fock	ROHF	Restricted open-shell Hartree-Fock
LCAO	Linear combination of atomic orbitals	CI	Configuration interaction
CISD	Configuration interaction with singles and doubles	QCI	Quadratic configuration interaction
QCISD	Quadratic configuration interaction with singles and doubles	MP	Møller-Plesset
LDA	Local density approximation	LSDA	Local spin density approximation
GC	Gradient-corrected	GGA	Generalized gradient approximation
S	Slater-Dirac exchange functional	VWN	Correlation functional of Vosko, Wilk and Nusair
B	Becke's 1988 exchange functional	B3	Becke's three-parameter exchange functional
LYP	Lee, Yang and Parr's correlation Functional	P86	Perdew's 1986 correlation functional
PW91	Perdew and Wang's 1991 correlation functional	BHandH	Becke's half and half correlation functional
B1LYP	Becke's one-parameter hybrid exchange functional combined with the LYP correlation functional	STO	Slater-type orbital
GTO	Gaussian-type orbital	PES	Potential energy surface
TS	Transition structure	ZPVE	Zero-point vibrational energy
IRC	Intrinsic reaction coordinate	SCRF	Self-consistent reaction field
BSSE	Basis set superposition error	CP	Counterpoise correction
Å	angstrom	D	debye
PA	Proton affinity		

I would like to express my immense gratitude to my supervisor Dr. R. J. Boyd. Over the last four years he has acted as a mentor, providing the encouragement, support and guidance necessary to ensure my graduate experience has been a memorable one. As a person, he has helped me to mature and develop the skills necessary to become a proficient computational chemist.

I am gratefully for the dialogue and support from Fuqiang Ban and Nelaine Mora-Diez, who have been with me for the four years, and also Dr. Zhenming Hu. A special thank you to Dr. George Heard, who is always willing to answer my questions about the AIMPAC program, and Dr. Kent Worsnop who helped me to develop an appreciation of computers and DFT. Outside the lab, the members of the Burford group, past and present, have always provided comic relief.

A special note of thanks is necessary for Drs. Stacey Wetmore and James Gauld. Over the past four years, Stacey has provided me with uninterrupted friendship and support. Her conversations and advice have been a cherished part of my studies. In the two years that James has been at Dalhousie, he has taken on a unique role in my research. His attention to detail, guidance in research, and criticisms have had a notable influence on my personal and professional development. Although I do not always express it, I am very grateful to both of you for your help over the years. Thank You!

Finally, to my family. Thank you for believing in me and providing me with continuous support and encouragement. Even though we are miles apart, you are always in my heart and thoughts. And last but not least, Greg. You have been my dearest friend and companion through the good times and the bad over the last few years. Thank you for everything!

Chapter One.
Introduction

1.1 General Introduction

For decades, chemists have been interested in furthering their understanding of the fundamental nature of the chemical bond, which results as a consequence of the sharing or transfer of electrons between two or more atoms or molecules. An equal sharing of electrons between two atoms within the same molecule gives rise to short and directional intramolecular covalent bonds. However, if the interaction occurs between two separated molecules, van der Waals interactions arise which are weaker in strength and less directional in nature than covalent bonds. Although less than five years separated the time when G. N. Lewis[1] described covalent interactions and the time when Latimer and Rodenbush[2] proposed the idea of the hydrogen bond, a type of van der Waals interaction, the last decade has seen a surge of interest into the study of hydrogen

bonds due to their importance in a vast number of biological and chemical phenomena and the dramatic advances in computer technology.

Hydrogen bonding, a donor-acceptor interaction specifically involving hydrogen atoms,[3] is a relatively weak interaction with bond strengths that range up to 150 kJ mol^{-1}. As a molecule engages in hydrogen bonding, a variety of structural and energetic transformations result which may be quantified by spectroscopy.[4] Although the spectroscopic techniques have provided a multitude of information on hydrogen-bonded systems, additional information not easily accessible by experimental means may be obtained from theoretical calculations on model systems. Specifically, quantum chemistry enables chemists to study transient species, such as radicals, and to gain insight in the transition structures formed in chemical reactions.

Many of the quantum chemical methods traditionally used to study chemical systems are very costly in terms of the computation time and computer memory requirements even for systems containing as few as five to ten non-hydrogen atoms. This has severely limited the effectiveness by which biologically relevant chemical systems have been studied. However, the advent of density functional theory (DFT)[5,6] has alleviated the aforementioned restrictions and enabled investigations of notably larger systems. Not only are these calculations computationally feasible but also they maintain a level of accuracy comparable to traditional *ab initio* methods.

1.2 Thesis Overview

This thesis is composed of several projects linked by the general theme of hydrogen bonding. Prior to entering into a discussion of the projects, a brief description of the available quantum chemical methodologies and basis sets used for the determination of molecular structure and other electronic properties is given in Chapter Two. The application of *ab initio* and DFT methods to study the structures and spectroscopic properties of the prototypical hydrogen halide dimers $(HF)_2$, $(HCl)_2$ and $(HBr)_2$ is described in Chapter Three. Various hybrid DFT methods are compared and their potential for determining select geometrical and vibrational properties of the hydrogen-bonded systems is discussed.

Hydrogen bonds are an important part of the structure and function of biological molecules. Biomolecules contain a robust variety of functional groups that may generate multiple hydrogen bonding interactions, thus conferring additional stability to a chemical system of interest. In Chapter Four, molecules derived from the nucleobase uracil are employed as catalysts in the aminolysis of 6-chloropyrimidine and 2-amino-6-chloropyrimidine. The uracil derivatives are shown to stabilize the transition structures through the formation of multiple hydrogen bonds thus lowering the barrier to aminolysis. The computational study reveals a correlation between the barrier to aminolysis and the proton affinity of the carbonyl group of the base interacting with the incoming NH_3. This correlation is further explored using fluoro, imine and sulfur

derivatives of the uracil-derived bases, OCH_2, $OCH\text{-}NH_2$ and OCH-NH-HCO. This study illustrates the importance of a well-chosen hydrogen bond acceptor and the catalytic possibilities of hydrogen bonding.

Proline is one of the twenty naturally occurring amino acids that form the building blocks of biologically significant polymers and its conformational interconversion in polypeptides has been shown to be rate-limiting in protein folding.[7] In Chapter Five, proline is utilized as a catalyst in the direct aldol reaction between acetone and acetaldehyde. In this study, the proposed enamine mechanism is examined along with the potential role of the solvent. The results presented herein illustrate the potential of small molecules to act as catalysts in biochemical reactions. In Chapter Six, the *cis-trans* isomerization of prolyl peptides is examined using N-acetylproline methylamide as a model system. By examining the conformational and energetic consequences of the conformational interconversion and the effect of electron-withdrawing substituents on the proline ring, the importance of an intermolecular N···HN hydrogen bond is examined and the catalytic possibilities of hydrogen bonding explored.

Finally, in Chapter Seven, global conclusions drawn from the work presented within are summarized and several potential research projects relating to hydrogen bonding and catalysis are proposed.

1.3 References

(1) Lewis, G. N. *Valence and the Structure of Atoms and Molecules*, Chemical Catalog Co.: New York, 1923.

(2) Latimer, W. M.; Rodebush, W. H. *J. Am. Chem. Soc.* **1920**, *42*, 1419.

(3) Jeffery, G. A. *An Introduction to Hydrogen Bonding*, Oxford University Press: New York, 1997.

(4) Pimentel, G. C.; McClennan, A. L. *The Hydrogen Bond*, W. H. Freeman: San Francisco, 1960.

(5) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*, Oxford University Press: New York, 1989.

(6) Seminario, J. M.; Politzer, P. (editors) *Modern Density Functional Theory: A Tool for Chemistry*, Elsevier: New York, 1995.

(7) (a) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953. (b) Schmid, F. X.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 4764. (c) Schmid. F. X.; Baldwin, R. L. *J. Mol. Biol.* **1979**, *133*, 285.

Chapter Two.
Theoretical Background

2.1 Introduction

Chemistry is the branch of science concerned with the study of the properties and interactions of molecules. A molecule is in turn composed of electrons and nuclei, the motion of which is governed by the laws of quantum mechanics. From the laws of quantum mechanics, it is also known that many important and observable properties of a system may be obtained from the wave function, a function depending upon the coordinates of the particles and time. In science, the application of the laws of quantum mechanics to the study of chemical structure and reactivity is a branch of chemistry referred to as quantum chemistry.

Within computational quantum chemistry, there are two distinct categories of theoretical methods devoted to the evaluation of chemical phenomena: semi-empirical and *ab initio* electronic structure theory. The difference between these general classes lies in the nature of the approximations implemented in order to simplify the equations. Semi-empirical methods, which may be applied to the study of large systems, require few computational resources since they neglect or parameterize computationally demanding integrals using experimental data or the results of higher level calculations. The methods of electronic structure theory use the laws of quantum chemistry, fundamental constants and a few mathematical approximations in order to characterize a system of interest. Thus, although the *ab initio* methods of electronic structure theory are more computationally demanding, they may be utilized to characterize systems containing more transient species such as transition structures or radicals.

Throughout this thesis, the conventional *ab initio* and density functional theory (DFT) classes of electronic structure methods are utilized. The *ab initio* methods, which provide quantitative results for a variety of systems, differ in the nature of the approximations utilized to create the wave function. The density functional methods, which possess the accuracy of conventional *ab initio* methods and are computationally less expensive, model electron correlation using functionals based on the electron density of a system. This chapter is designed to provide the requisite theoretical background so that the reader can differentiate between the different quantum chemical methods and have a sense of the relative accuracy associated with such methods.

2.2 The Schrödinger Equation

According to the postulates of quantum mechanics,[1,2] the energy and other significant properties of a system are explicitly described by a wave function, Ψ, which satisfies the time-dependent Schrödinger equation (Equation **2.1**).

$$\hat{H}_{total}\Psi(\bar{r},t) = \frac{ih}{2\pi}\frac{\partial\Psi(\bar{r},t)}{\partial t} \tag{2.1}$$

For many applications of quantum mechanics to chemistry it is sufficient to utilize the time-independent Schrödinger equation,[3] obtained by writing the wave function as the product of a spatial and a time function. In the time-independent Schrödinger equation, expressed simply in Equation **2.2**, $\hat{H}_{total}$ is the Hamiltonian operator, Ψ is the wave function and E is the energy of the stationary state of the system.

$$\hat{H}_{total}\Psi = \mathrm{E}\Psi \tag{2.2}$$

The majority of the complexity associated with this equation is enclosed in the Hamiltonian that, for a general system, contains the kinetic and potential energy for all the particles. The Hamiltonian operator is expressed as

$$\hat{H}_{total} = -\sum_i \frac{h^2}{8\pi^2 m_i}\nabla_i^2 + \hat{V} \tag{2.3}$$

in which m, h and $\hat{V}$ represent the mass of the particle, Planck's constant, and the potential field in which the particle is moving, respectively.

Quantum chemistry involves the application of the time-independent Schrödinger equation to atoms and molecules in order to obtain knowledge about their properties. In order to apply the aforementioned equation to molecular systems, the Hamiltonian is more commonly expressed as in Equation **2.4**.

$$\hat{H}_{total} = \hat{T}_e + \hat{V}_{ee} + \hat{T}_n + \hat{V}_{ne} + \hat{V}_{nn} \tag{2.4}$$

In the above equation, $\hat{T}$ and $\hat{V}$ are the kinetic and potential energy operators, respectively, and the subscripts n and e refer to the nuclei and electrons. Unfortunately, the time-independent Schrödinger equation is insolvable except for the simplest of cases. Approximate solutions can be obtained for a wide range of chemical problems by the application of several rigorous mathematical approximations.

2.3 Fundamental Approximations

2.3.1 The Born-Oppenheimer Approximation

One of the central approximations to quantum chemistry, the Born-Oppenheimer or adiabatic approximation,[4,5] separates nuclear and electronic motion by assuming the nuclear motion is fixed and the electrons are moving in the field of fixed nuclei.

Assuming this to be the case allows the Schrödinger equation to be solved for the electrons in a given nuclear framework. In the electronic Schrödinger equation, (Equation **2.5**),

$$\hat{H}_{elect}\Psi(\bar{r}) = E_{elect}\Psi(\bar{r}) \tag{2.5}$$

the electronic Hamiltonian operator, $\hat{H}_{elect}$, is analogous to the Hamiltonian of Equation **2.1** except that there is no contribution from the nuclear kinetic energy and the nuclear potential energy is constant for a given nuclear configuration.

$$\hat{H}_{elect} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne} \tag{2.6}$$

The electronic Schrödinger equation is an eigenvalue equation, whose solutions (referred to as eigenvectors (Ψ_i)) are wave functions depending upon the spatial coordinates of the electrons. The eigenvalues (E_i) of $\hat{H}_{elect}$ describe the various electronic states of the system. The lowest eigenvalue corresponds to the ground state of the system.

2.3.2 The Orbital Approximation

In order to obtain accessible solutions to Equation **2.5**, the wave function (Ψ) must be properly expressed. This requires that the spin coordinates of the electrons be included in the wave function. As the motions of the electrons are assumed to be independent, each electron may be assigned to a unique spin orbital, χ_i. The spin orbitals

are expressed as a product of a spatial function, ψ_i, depending upon the coordinates of the i^{th} electron, and a spin function (α or β). This is referred to as the orbital approximation and the spatial orbitals referred to above are also known as molecular orbitals (MO).[4]

Using the one-electron spin orbitals described above, a wave function describing the full system of N non-interacting electrons may be derived. As originally proposed by Hartree,[6] the simplest form of this wave function is to express it as a product of spin and spatial functions (Equation **2.7**).

$$\Psi(1,2,...N) = \psi_1\alpha(1)\psi_2\beta(2)\psi_3\alpha(3)....\psi_N\beta(N) \qquad (2.7)$$

However, the resulting Hartree product is deficient since it does not account for the indistinguishability of electrons. It has been shown experimentally that, for particles with half-integral spins (fermions), the electronic wave function must be antisymmetric with respect to interchanging the spatial and spin coordinates of any two of the electrons. This is expressed mathematically in Equation **2.8**.

$$\Psi(1,2,...j,i,.....N) = -\Psi(1,2,...i,j,...N) \qquad (2.8)$$

Expressing the wave function for a $2N$-electron system as a Slater determinant,[7] as in Equation **2.9**, guarantees that the antisymmetry principle is satisfied since

interchanging any two row of the determinant (electrons) changes the sign of the determinant and if any two columns (orbitals) are identical, the determinant vanishes.

$$\Psi(1,2,3,....2N) = [(2N)!]^{-1/2} \begin{vmatrix} \psi_1\alpha(1) & \psi_1\beta(1) & \psi_2\alpha(1) & \cdots & \psi_N\beta(1) \\ \psi_1\alpha(2) & \psi_1\beta(2) & \psi_2\alpha(2) & \cdots & \psi_N\beta(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_1\alpha(2N) & \psi_1\beta(2N) & \psi_2\alpha(2N) & \cdots & \psi_N\beta(2N) \end{vmatrix} \quad \textbf{(2.9)}$$

The factor $(2N)!^{-1/2}$ is referred to as the normalization constant and is the consequence of imposing the condition outlined in Equation **2.10**.

$$\int \psi^* \psi d\tau = 1 \quad \textbf{(2.10)}$$

The antisymmetry requirement immediately leads to the Pauli exclusion principle,[8] which states that it is not possible for two electrons to occupy the same molecular orbital while having the same spin. In terms of the Slater determinant, two columns have to be identical in order for the Pauli principle to be violated.

2.3.3 The Variational Theorem

As described above, the simplest antisymmetric wave function often used to describe the ground state of an N electron system is a Slater determinant. Since the exact form of the wave function (Ψ) is not always known, it is approximated by trial functions (Φ). According to the variational theorem,[2-4] the energy obtained for any trial function will be greater than the true energy associated with the ground state.

$$E_{\Phi} = \frac{\int \Phi^* \hat{H} \Phi}{\int \Phi^* \Phi} d\tau \geq E_0 \qquad (2.11)$$

Varying the parameters associated with the trial function until the expectation value $\langle \Phi | \hat{H} | \Phi \rangle$ associated with the trial function reaches a minimum, provides an upper bound estimate of the ground state energy. An *ab initio* method is classified as a variational method if the energy it yields is always an upper bound to the exact energy.

2.4 The Independent Particle Model

2.4.1 The Hartree-Fock Method

The simplest *ab initio* technique available to obtain an approximate wave function is the Hartree-Fock (HF) method.[2,3] Often used as the basis for other computational methods, the HF method (also referred to as the independent particle method) uses a normalized single particle wave function or orbital to describe an electron.

According to the variational theorem, the best spin orbitals are those that minimize the electronic energy. By systematically varying the spin orbitals, the Hartree-Fock (HF) equations (Equation **2.12**) may be derived.

$$\hat{F} \psi_i = \varepsilon_i \psi_i \qquad i=1,2,...N \qquad (2.12)$$

In the HF equations, the eigenvalue (ε_i) is associated with the orbital energy of an electron in ψ_i, and $\hat{F}$ is the Hartree-Fock Hamiltonian operator, expressed mathematically in Equation **2.13**.

$$\hat{F} = \hat{H}_{core} + \sum_{j=1}^{N} \left(2\hat{J}_j - \hat{K}_j\right) \tag{2.13}$$

The core Hamiltonian operator, $\hat{H}_{core}$, the Coulomb operator, $\hat{J}_j$, and the exchange operator, $\hat{K}_j$, are defined in Equations **2.14**, **2.15**, and **2.16**, respectively.

$$\hat{H}_{core} = -\tfrac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{A=1}^{n}\sum_{i=1}^{N}\frac{Z_A}{r_{Ai}} \tag{2.14}$$

$$\hat{J}_j(1) = \int \psi_j(2)\frac{1}{r_{12}}\psi_j(2)dr_2 \tag{2.15}$$

$$\hat{K}_j\psi_i(1) = \left[\int \psi_j(2)\left(\frac{1}{r_{12}}\right)\psi_i(2)dr_2\right]\psi_j(1) \tag{2.16}$$

The HF equations must be solved iteratively until no appreciable change in the orbitals is achieved since the Fock operator ($\hat{F}$) depends on the exchange and Coulomb operators which are themselves functions of ψ_i. An initial guess is made for the orbitals and each iteration corresponds to a better guess of all the orbitals. The cycle continues until the orbitals do not change significantly between successive iterations. As each electron is described by a potential energy due to the nucleus, the resulting orbitals are

derived from their own potential. This is referred to as the self-consistent field (SCF) technique or Hartree-Fock (HF) procedure since this methodology was first described by Hartree[6] and Fock.[9]

2.4.2 The Roothaan-Hall Equations

Although the molecular orbitals of the HF equations (Equation **2.12**) have been determined numerically for atoms, this is not practical in the case of molecules. Roothaan[10] and Hall[11] provided a simplified method for solving the HF differential equations by proposing that the molecular orbitals, ψ_i, be constructed using a finite set of one-electron functions, φ_μ, referred to as basis functions. When the basis functions are taken to be atomic orbitals centered on the constituent atoms, this approximation is called the linear combination of atomic orbitals (LCAO) approximation. Thus, the resulting form of the molecular orbitals (MO) is outlined in Equation **2.17**, in which $C_{\mu i}$ are the MO coefficients.

$$\psi_i = \sum_{\mu=1}^{M} C_{\mu i}\varphi_\mu \tag{2.17}$$

Application of the variation principle leads to the Roothaan-Hall equations, Equation **2.18**, in which the energy is a minimum when the MO expansion coefficients

$C_{\mu i}$ satisfy Equation **2.18**. As with the HF equations, the Roothaan-Hall equations must be solved iteratively until self-consistency is attained.

$$\sum_{\mu=1}^{M} \left(F_{\mu\nu} - \varepsilon_i S_{\mu\nu}\right) C_{\mu i} = 0 \qquad \mu = 1, 2, \ldots M \tag{2.18}$$

In a system with M basis functions, ε_i is the one-electron orbital energy of the molecular orbital ψ_i, and $S_{\mu\nu}$ and $F_{\mu\nu}$ are elements of the overlap and Fock matrices, respectively. The elements of the Fock matrix are defined as

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda,\sigma} P_{\mu\nu}\left[(\mu\nu \mid \lambda\sigma) - \tfrac{1}{2}(\mu\lambda \mid \nu\sigma)\right] \tag{2.19}$$

where $H_{\mu\nu} = \langle \varphi_\mu \mid \hat{H}_{core} \mid \varphi_\nu \rangle$ is a matrix element of the core Hamiltonian operator representing the energy of a single electron moving in the field of the bare nuclei, $P_{\mu\nu} = 2\sum_{i=1}^{occ} C_{\mu i} C_{\nu i}$ is a matrix element of the density matrix and $(\mu\nu \mid \lambda\sigma)$ is a two-electron repulsion integral defined as in Equation **2.20**.

$$(\mu\nu \mid \lambda\sigma) = \iint \varphi_\mu(1)\varphi_\nu(1)\frac{1}{r_{12}}\varphi_\lambda(2)\varphi_\sigma(2)dr_1 dr_2 \tag{2.20}$$

However, the above equations are valid only when two electrons occupy the same spatial orbital. That is, all electron spins are paired in a closed shell state. This method is referred to as restricted Hartree-Fock (RHF). For an open-shell system, where the

number of α and β electrons are unequal due to the presence of one or more unpaired electrons, it is more appropriate to introduce different orbitals for different spins. In the unrestricted Hartree-Fock (UHF) method, the electrons with α and β spin are assigned to different spatial orbitals, generating two expressions for the molecular orbitals (Equation **2.21**).

$$\psi_i^{\alpha} = \sum_{i=1}^{M} C_{\mu i}^{\alpha} \varphi_{\mu} \quad \text{and} \quad \psi_i^{\beta} = \sum_{i=1}^{M} c_{\mu i}^{\beta} \varphi_{\mu} \tag{2.21}$$

The alternate method available for open-shell systems is the restricted open-shell Hartree-Fock (ROHF) method in which the doubly occupied orbitals are treated using the RHF formalism while the singly occupied orbitals are treated independently in a more complicated manner. The deficiency in the ROHF method lies in its inability to treat the interaction between paired and unpaired electrons, a problem corrected for in the UHF.

In general, the UHF method provides a superior description of open-shell systems and yields a lower energy than the ROHF method due to the extra flexibility inherent in the wave function. The major drawback of the UHF method over the ROHF method is that solutions to the UHF equations may not be pure spin states, but are often contaminated by higher states. A measure of the severity of these contributions can be obtained by determining the expectation value of the spin squared (S^2) operator. The difference between this value and the expectation value for a pure spin state [S(S+1)] is referred to as the spin contamination.

2.5 Electron Correlation

The main deficiency of Hartree-Fock theory is that it provides an inadequate description of the correlation between the electrons in a molecular system. In the HF method, the electrons are taken to be independent of each other or at least interacting only through some average, or effective potential. Thus, the probability of finding two electrons in the same region of space is assumed to be equal to the product of the individual probabilities. However, it is energetically preferable for electrons of paired spin to be well separated from each other and thus their motions are correlated. The energy associated with this phenomenon called the correlation energy[1,4] and is defined in Equation **2.22**.

$$E_{corr} = E_{exact} - E_{HF} \tag{2.22}$$

Although the correlation energy yields only a small contribution, about 1 %, to the total energy, it is very important for the calculation of molecular properties that involve bond breaking and formation. A number of types of calculations begin with a Hartree-Fock calculation and then correct for electron correlation. Methods which account for electron correlation, collectively referred to as post-SCF or post-HF methods, include configuration interaction (CI), many body perturbation theory and density functional theory (DFT).

2.5.1 Configuration Interaction

The wave function, expanded as a single determinant, in HF theory, is often inadequate for the calculation of many electronic properties. Configuration interaction (CI)[4,5] expands the wave function as a linear combination of Slater determinants (Φ) where c_i represents the expansion coefficient for the i^{th} determinant.

$$\Psi = \sum_i c_i \Phi_i \tag{2.23}$$

The first determinant, (Φ_o), is taken to be the HF determinant and the subsequent determinants, referred to as excited determinants, are generated by allowing the electrons in occupied orbitals of the HF determinant to be excited to the unoccupied (virtual) orbitals. The resulting CI wave function is represented in Equations **2.24** and **2.25**.

$$|\Psi> = c_o |\Psi_{HF}> + \sum_{a,r} c_a^r |\Psi_a^r> + \sum_{\substack{a<b,\\ r<s}} c_{a,b}^{r,s} |\Psi_{a,b}^{r,s}> + \sum_{\substack{a<b<c\\ r<s<t}} c_{a,b,c}^{r,s,t} |\Psi_{a,b,c}^{r,s,t}> + \cdots \tag{2.24}$$

$$\Psi = c_o \Phi_o + \sum_{ar} c_a^r \Phi_a^r + \sum_{\substack{a<b\\ r<s}} c_{ab}^{rs} \Phi_{ab}^{rs} + \sum_{\substack{a<b<c\\ r<s<t}} c_{abc}^{rst} \Phi_{abc}^{rst} + \ldots \tag{2.25}$$

In the above expressions, Φ_a^r represents an excitation generated by moving a single electron from the occupied orbital a to the virtual orbital r and double and triple excitations are represented by Φ_{ab}^{rs} and Φ_{abc}^{rst}, respectively.

In principle, an exact wave function can be obtained through an infinite expansion of the CI equation so that all possible excitations or configurations are included. Although the full CI expansion provides the most accurate upper bound to the ground state energy, it is time consuming and computationally expensive for even small systems. These obstacles may be overcome by truncating the CI expansion to include only those configurations that differ from the HF ground state determinant by no more than a set number of excitations. If the truncation occurs after two excitations, then single and double excitations are included and this method is called CISD (configuration interaction with singles and doubles). Although, most of the correlation energy is accounted for in the CISD method, the wave function that results from truncation of the CI equation is not size consistent. This means that the results obtained with a truncated CI wave function for a system of molecules infinitely separated from each other are not equal to the sum of the results calculated for each individual molecule. Size consistency is important for the comparison of results obtained for different systems with the same level of theory.

2.5.2 Quadratic Configuration Interaction

The problem of size consistency associated with the truncated CI methods can be overcome through the addition of supplementary terms to the truncated equations. Through the addition of quadratic terms of the operator, a method referred to as quadratic configuration interaction (QCI),[12] size consistency is restored. The most popular of the truncated CI expansions, with quadratic terms included, is the QCISD method.[12] This

method supplements the linear CISD expansion with a sufficient quantity of quadratic terms to ensure size consistency. As a result, the QCISD expansion contains contributions from all singly- and doubly-excited configurations in addition to a few quadruply-excited configurations. Although the QCI methods are size consistent, they are not variational.

2.5.3 Many Body Perturbation Theory

The methods discussed above are based on the variational principle. However, an alternative computational method exists for the inclusion of correlation effects referred to as perturbation theory.[2,4] In this technique, the total Hamiltonian is partitioned into an unperturbed and perturbed portion as outlined in Equation **2.26**.

$$\hat{H} = \hat{H}_o + \lambda V \tag{2.26}$$

In this equation, $\hat{H}_o$ is a unperturbed zero-order Hamiltonian which has known eigenfunctions and eigenvalues and λV is the perturbation. The eigenvalues and eigenfunctions of the full Hamiltonian are expanded in terms of a Taylor series in λ:

$$\mathrm{E}_{total} = \mathrm{E}_i^{(0)} + \lambda \mathrm{E}_i^{(1)} + \lambda^2 \mathrm{E}_i^{(2)} + \cdots \tag{2.27}$$

$$\left|\Psi_{total}\right\rangle = \left|\Psi_i^{(0)}\right\rangle + \lambda\left|\Psi_i^{(1)}\right\rangle + \lambda^2\left|\Psi_i^{(2)}\right\rangle + \ldots \tag{2.28}$$

where $E_i^{(n)}$ is the *n*th order energy for the *i*th MO. Equations **2.27** and **2.28** are subsequently substituted back into the electronic Schrödinger equation, the products expanded and the coefficients of equal powers of λ are equated to yield a series of equations representing progressively higher orders of perturbation. This type of perturbation theory is associated with the names Rayleigh and Schrödinger (RSPT) and is often referred to as many body perturbation theory (MBPT).

In computational chemistry, the most common method based on perturbation theory is Møller-Plesset perturbation theory (MP)[13] in which the HF Hamiltonian is used as the zero-order Hamiltonian. Thus, the energy derived from first-order equations is equivalent to the HF energy. The various orders of MP are obtained by setting $\lambda = 1$ and truncating Equations **2.27** and **2.28** to various orders. The calculation of the total energy to the second order is called MP2, to the third order MP3 and so on. The MP calculations are size consistent at any level of truncation and are notably faster than CI calculations, thus overcoming the major disadvantage of the variational methods. However, since this method is not variational, it may lead to an energy below the true energy.

The most familiar form of Møller-Plesset techniques used in the literature includes only the second-order energy contribution (MP2). In general, including third-order corrections leads to little enhancement in calculated results and often worse agreement with experiment is obtained at an increased computational cost. Thus, MP4 is

usually implemented to improve upon the second-order correction. Similarly, MP5 leads to little improvement over results obtained with MP4, which represents the oscillating behavior of the MP series. MP2 has been used to obtain a variety of electronic properties since it includes electron correlation at a reduced computational cost relative to other *ab initio* techniques.

2.6 Density Functional Theory

Another class of computational techniques, which includes the effect of electron correlation, is density functional theory (DFT).[14] Two theorems, put forth by Hohenberg and Kohn,[15] revealed that all the information about a system could be obtained by using the electron density, ρ(r), of a molecule. Specifically, the first theorem stated that the electron density of a molecule ρ(r) uniquely determines the external potential $\upsilon(r)$. This facilitates the description of the energy as a functional of the density as shown in Equation **2.29**.

$$E[\rho] = \int \rho(r)v(r)dr + T[\rho] + V_{ee}[\rho] \qquad (2.29)$$

In this equation, T[ρ] represents the kinetic energy functional and $V_{ee}[\rho]$ contains the classic electron-electron repulsion, as well as a non-classical contribution which is the major part of the exchange-correlation energy.

The second theorem, whose basis is the variational principle, implies that for a trial density $\tilde{\rho}(r)$ such that $\tilde{\rho}(r) \geq 0$ and $\int \tilde{\rho}(r) dr = N$, there exists an energy, $E[\tilde{\rho}]$, which is greater than the true energy. It can be shown that for the energy to be stationary with respect to changes in the density, Equation **2.30** must hold.

$$\delta E[\rho] - \mu \delta \left[\int \rho(r) dr - N \right] = 0 \tag{2.30}$$

In the above equation, μ is known as the Euler-Lagrange multiplier. This leads to the Euler-Lagrange equation:

$$\mu = \upsilon(r) + \frac{\delta T[\rho]}{\delta \rho(r)} + \frac{\delta V_{ee}[\rho]}{\delta \rho(r)} \tag{2.31}$$

In the above equation, $\frac{\delta T[\rho]}{\delta \rho(r)}$ and $\frac{\delta V_{ee}[\rho]}{\delta \rho(r)}$ are functional derivatives of the kinetic energy and electron-electron repulsion, respectively. This equation cannot be solved directly without knowledge of the aforementioned functionals. An indirect method for solving the Euler-Lagrange equations was introduced by Kohn and Sham[16] in which they approximated these functionals by introducing a determinantal function in which the kinetic energy is expressed as in Equation **2.32**,

$$T_s[\rho] = \left\langle \psi_i \left| \sum \left(-\frac{1}{2} \nabla_i^2 \right) \right| \psi_i \right\rangle \tag{2.32}$$

and which is subject to the constraint outline in Equation **2.33**.

$$\rho(r) = \sum_{i}^{N} |\psi_i(r)|^2 \tag{2.33}$$

This description of the kinetic energy and density holds true for a determinantal function that describes N non-interacting electrons described by the orbitals, ψ_i. This generates a potential, $\upsilon_{eff}(r)$, such that the Kohn-Sham N one-electron equations (Equation **2.34**) may be obtained.

$$\left[-\frac{1}{2}\nabla^2 + \upsilon_{eff}(r)\right]\psi_i = \varepsilon_i \psi_i \qquad i = 1,2,\ldots N \tag{2.34}$$

The energy of the interacting system can be expressed as

$$\mathrm{E}[\rho] = \int \rho(r)\upsilon(r)dr + \mathrm{T}_s[\rho] + J[\rho] + \mathrm{E}_{xc}[\rho] \tag{2.35}$$

where $\mathrm{E}_{xc}[\rho]$ is the exchange-correlation energy functional that is defined below.

$$\mathrm{E}_{xc}[\rho] = \mathrm{T}[\rho] - \mathrm{T}_s[\rho] + V_{ee}[\rho] - J[\rho] \tag{2.36}$$

The derivative of the exchange-correlation energy functional yields the exchange-correlation potential, $\upsilon_{xc} = \delta \mathrm{E}_{xc}[\rho] / \delta\rho(r)$. In Equation **2.35**, the only unknown parameter is the exchange-correlation energy functional. The difference between the

various DFT methods lies in the choice of the functional used to express the exchange-correlation energy.

DFT methods offer significant advantages over the methods previously discussed due to the fact that DFT, in addition to including electron correlation, is computationally more efficient and requires less computer time and memory relative to the post-HF methods. However, a DFT calculation can only be improved upon by obtaining a more accurate description of the exchange-correlation functional, a prominent area of research.

2.6.1 Local (Spin) Density Approximation (LSDA)

One of the basic approximations in DFT, the local density approximation (LDA), assumes that the density can be treated as that of a slowly varying uniform electron gas. Expressing the total density as a sum of the spin densities (α and β) yields the local spin density approximation (LSDA),[17] the simplest expression for the exchange-correlation energy functional. The exchange-correlation energy is expressed as

$$\mathrm{E}_{xc}^{LSDA}[\rho^{\alpha},\rho^{\beta}] = \int \varepsilon_{xc}^{LSDA}[\rho^{\alpha}(\bar{r}),\rho^{\beta}(\bar{r})]d(\bar{r}) \qquad (2.37)$$

where ε_{xc}^{LSDA} is the exchange-correlation energy distribution per unit volume, which depends on the density at the point of evaluation. The most commonly employed LSDA functional is a combination of the Slater-Dirac[18,19] exchange term (S) and the correlation functional of Vosko, Wilk, and Nusair (VWN).[20] This functional, SVWN, parameterizes

the exact uniform electron gas model. However, the LSDA functionals have been shown to overbind atoms in molecules and the energies obtained are not as good as those obtained with conventional *ab initio* methods. In addition, LSDA functionals yield errors that are in proportion to the size of the system under examination and generally cannot reproduce the effects associated with the creation or rupture of bonds. This is attributed to the fact that its homogeneous electron density associated with the uniform electron gas model is never achieved in molecular systems.

2.6.2 Generalized Gradient Approximation (GGA)

To account for the nonuniformity of the electron density, the density and the gradient of the density are incorporated into the exchange and correlation energy functionals. These functionals, referred to as gradient-corrected (GC) methods, depend upon the generalized gradient approximation (GGA) and have the general form of Equation **2.38**.

$$E_{xc}^{GGA}[\rho^{\alpha},\rho^{\beta}] = \int d(\vec{r})\varepsilon_{xc}[\rho^{\alpha}(\vec{r}),\rho^{\beta}(\vec{r}),\nabla\rho^{\alpha},\nabla\rho^{\beta}] \qquad (2.38)$$

This exchange-correlation functional can be separated into exchange and correlation functionals, as shown in Equation **2.39**, which are developed independently.

$$E_{xc}[\rho_{\alpha},\rho_{\beta}] = E_{x}[\rho_{\alpha},\rho_{\beta}] + E_{c}[\rho_{\alpha},\rho_{\beta}] \qquad (2.39)$$

The most popular correlation functionals include those of Perdew (P86),[21] Perdew and Wang (PW91),[22] and Lee, Yang and Parr (LYP).[23] The P86 functional introduced a cut-off parameter to go beyond a random phase for uniform, slowly varying electron gas and thus, improves upon its predecessor, the Langreth and Mehl functional.[24] PW91 also imposes a cut-off parameter within the random phase approximation and is a functional derived from first principles as it analytically represents the correlation energy for a uniform electron gas as a function of a density parameter r_s and the relative spin polarization ξ. Finally, the LYP functional uses the Colle-Salvetti[25] formula for E_c and replaces the local kinetic energy density with its second-order density gradient expansion.

In order to improve upon the "S" exchange functional, Becke proposed a gradient corrected exchange functional known as the Becke88 (B88) or simply B functional.[26] This functional is based on the fact that the exchange energy density for a many electron system must possess a certain asymptotic behaviour (-1/r) as r approaches infinity. The exact asymptotic behavior of the exchange-energy density is related to the total exchange energy by Equation **2.40**.

$$E_X = \frac{1}{2}\sum_{\sigma}\int \rho_\sigma U_X^\sigma d^3r \qquad \textbf{(2.40)}$$

The resulting exchange functional has the form outline in Equation **2.41**.

$$E_X = E_X^{LDA} - \beta \int \rho_\sigma^{4/3} \frac{\rho^{4/3} x^2}{(1 + 6\beta x \sinh^{-1} x)} d^3 r \qquad (2.41)$$

In the above functional, β is a parameter chosen to fit the known exchange energies of inert gas atoms and $x = \rho^{-4/3}|\nabla\rho|$. This functional is defined as a correction to the local LDA exchange functionals and successfully remedies many underlying deficiencies associated with the LDA functional.

An alternate approach to formulating the exchange-correlation energy uses the adiabatic connection method[27] to incorporate small amounts of exact exchange energy in the functional. Using the aforementioned method, the exchange-correlation energy is expressed as

$$E_{XC} = \int_0^1 U_{XC}^{\lambda} d\lambda \qquad (2.42)$$

in which λ, the coupling-strength parameter, controls the Coulomb interaction between electrons and U_{XC}^{λ} is the potential energy of the exchange-correlation at the coupling strength λ. This sums over a continuum of partially interacting systems, $0 \leq \lambda \leq 1$, from the noninteracting Kohn-Sham reference, $\lambda = 0$, to the fully interacting real system, $\lambda = 1$. For the Kohn-Sham reference, $\lambda = 0$, only exchange energy is present within E_{xc} and this energy is essentially equal to the Hartree-Fock exchange energy.[28]

The LSDA model substitutes a value from the local uniform electron gas theory for each value of U_{XC}^{λ}. This is often inappropriate near the λ=0 limit in bonds. As a result, the LSDA approximation in the noninteracting limit is the principal source of error. Thus, Becke created functionals which include a mixture of exact exchange and LSDA exchange to correct for the λ=0 case. These hybrid functionals are expressed as a linear combination of HF, LSDA and B exchange contributions, together with LSDA and non-local correlation contributions (usually P86, PW91 or LYP). The most popular of these hybrid functionals, the B3LYP functional,[29a,b] is given by

$$E_{xc}^{B3LYP} = E_{xc}^{LSDA} + a_0(E_x^{exact} - E_x^{LSDA}) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{LYP} \qquad \textbf{(2.43)}$$

where a_o=0.20, a_x=0.72 and a_c=0.81 are coefficients determined by semi-empirically fitting $E_{xc}[\rho]$ to experimental data (atomization energies, ionization potentials and proton affinities). The gradient corrections, ΔE_x^{B88} and ΔE_c^{LYP}, are corrections to the LSDA exchange and correlation energies, respectively. Hybrid functionals of this type are denoted as B3C, where B3 represents Becke's three-parameter functional[28,29a] and C represents the correlation functional.

Derived from the same methodology as the B3LYP functional, the B1LYP functional[28,29a] was created which performs essentially the same as the B3LYP functional but uses a single parameter to mix exact (HF) exchange and DF exchange in the

functional. By determining the ratio of HF/DF exchange prior to the calculation as 1/4, this functional is free of optimized parameters.

Another prominent hybrid functional, referred to as half-and-half theory (BHandH),[29a] approximates the exchange-correlation functional as

$$E_{xc} \cong \tfrac{1}{2} E_x + \tfrac{1}{2} E_{xc}^{LSDA} \tag{2.44}$$

which combines HF exact exchange with exchange derived from the LSDA approximation. This functional may be used independently or in conjunction with the correlation functional of LYP, yielding the BHandHLYP functional.

DFT has notable advantages over the other methods previously discussed. Specifically, it includes electron correlation at all levels and is computationally efficient. However, there is no systematic way to improve upon the calculation other than improving the exchange-correlation functional or by expanding the basis set. In addition, a lower energy by one DFT method does not guarantee that the functional used leads to more accurate molecular properties. Thus, all functional combinations must be tested to determine the best DFT method for a particular property. Various aspects of DFT are summarized in the numerous books[17,18,31] and review articles[27,32-37] that have appeared over the years.

2.7 Basis Sets

In the discussion of the Roothaan-Hall equations, it was stated that the molecular orbitals (ψ_i) are best described when they are expressed as a linear combination of atomic orbitals (LCAO). Thus, the description of the MOs is dependent upon an accurate description of the atomic orbitals (φ_i), which are mathematically modeled as a linear combination of a set of functions referred to as basis functions (ϕ^{BF}).[3,38]

$$\varphi^{AO} = \sum_i c_i \phi_i^{BF} \tag{2.45}$$

A finite group of basis functions, referred to as a basis set, should provide an adequate description of the AOs. Ideally, the basis functions should resemble the atomic orbitals and using Slater-type atomic orbitals (STOs)[39], which have an exponential decay outlined in Equation **2.46**,

$$\phi^{STO} \propto e^{-\xi|\vec{r} - \vec{R}|} \tag{2.46}$$

as basis functions reproduces most of the spatial properties of the AO, i.e. the cusp conditions at the nucleus. However, integration problems arise in the evaluation of the three- and four-center integrals and thus, Gaussian functions or Gaussian-type orbitals (GTOs)[40], with the exponential decay indicated in Equation **2.47**, are more commonly implemented.

$$\phi^{GTO} \propto e^{-\alpha|\vec{r}-\vec{R}|^2} \tag{2.47}$$

The GTOs lead to rapid integral evaluation due to the Gaussian product theorem which states that the product of two Gaussians on two different centers is a third Gaussian on a center between the first two. However, the GTOs are less satisfactory at describing the atomic orbitals since they are rounded in the region of the cusp and fall off too rapidly at large r. In order to achieve a compromise between the accuracy of STOs and the computational advantage of GTOs, STOs are commonly represented as a linear combination of GTOs. Thus, the atomic orbitals can now be expressed as

$$\phi_\mu = \sum_k d_{\mu k} g_k \tag{2.48}$$

where the coefficients $d_{\mu k}$ are fixed and the g_k, referred to as primitive Gaussian functions, are all of the same type (i.e., s, p, d, ...). A basis function of this type is a contracted Gaussian basis function. Since the number of variational parameters to be determined is reduced, the computational time is diminished.

In order to extend the variational flexibility in a minimal basis set, which uses the least number of basis functions to describe the atomic orbitals, the number of basis functions per atom is increased. This can be accomplished by using a split-valence basis set which uses twice the number of functions as a minimal basis set to describe the valence orbitals. This enables the valence AO to be modeled more accurately while

providing a minimal description of the core AO. Double-zeta split-valence basis sets, for example 6-31G, would contain basis functions composed of 6 primitives to describe the core orbitals (1*s*), a basis function of 3 primitives in addition to a basis function composed of a single primitive to describe the valence orbitals (2*s*, $2p_x$, $2p_y$, $2p_z$). By further dividing the valence region into three or more partitions, as in the 6-311G basis set, additional flexibility in a basis set is acquired. As this basis set only describes the s and p atomic orbitals, the basis set may become unbalanced.

To extend the accuracy of a double-zeta or triple-zeta split-valence basis set 6-31G and 6-311G, respectively, polarization functions (functions of higher angular momentum) can be added to account for distortion of the atomic orbitals in the molecular environment. For example, *d*-, *f*- or higher functions can be added to heavy atoms (atoms other than hydrogen or helium) and similarly, p-, d- or higher functions can be added to hydrogen and helium basis sets. This provides additional flexibility by allowing for electronic charge to be displaced away from the nucleus. Alternatively, diffuse functions (functions with small exponents) can be added to heavy atoms (6-31+G) or hydrogen (6-31++G). These functions account for diffuse electron clouds by allowing the orbitals to occupy larger regions in space. This is particularly useful to describe systems where electrons are loosely bound, such as anions.

The choice of basis set is an important consideration in a calculation. Not only does it have significant effects on the description of the atoms within the molecule of

interest, but it has implication on the amount of CPU time required to perform the calculation.

2.8 Potential Energy Surfaces

2.8.1 Geometry Optimizations

Using the aforementioned quantum methods, the primary goal of quantum chemistry is to obtain useful information about the electronic and geometrical properties of chemical systems. In order to investigate such properties of interest, an accurate description of the molecular geometry is required. The manner in which the geometry of a system changes with respect to its energy is represented by a potential energy surface (PES) diagram.

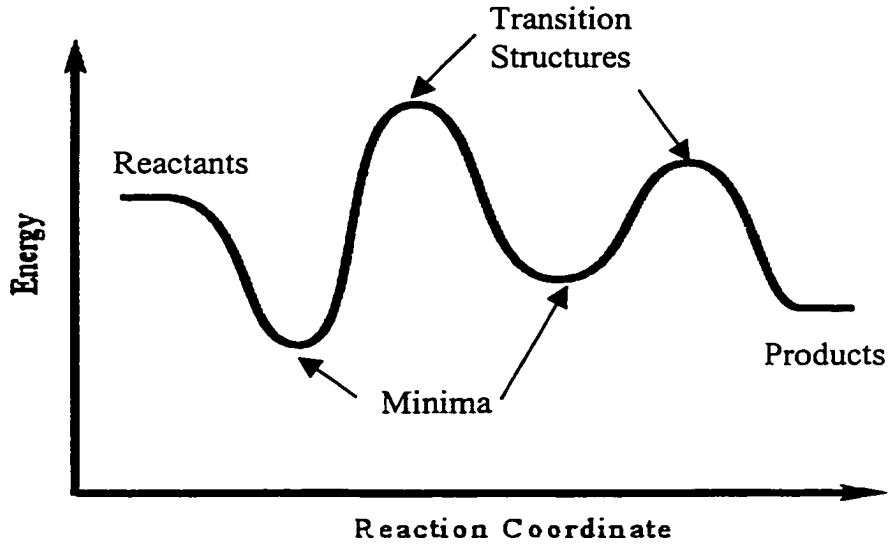


Figure 2.1: Schematic illustration of a potential energy surface (PES).

The calculation of a complete PES is impractical except for small systems. In order to develop an understanding of the chemistry of a system, one is interested in locating points of special significance on the surface referred to as stationary points. Stationary points are characterized by having the first derivatives of the energy, with respect to the 3N-6 internal coordinates, equal to zero. The two most vital types of stationary points are local minima (equilibrium structures) and first-order saddle points (transition structures). These stationary points are in turn characterized by the second derivatives of the energy (the force constants), which constitute the Hessian matrix or force-constant matrix.

$$\frac{\partial^2 E}{\partial r_i \partial r_j} \quad i, j = 1, 2, 3, \ldots 3N\text{-}6 \tag{2.49}$$

For a minimum, all the force constants or the eigenvalues of the Hessian matrix will be positive. A transition structure (TS), which connects two minima along a PES and occurs at a point of maximum energy, is identified by the one negative eigenvalue contained in the Hessian matrix.

2.8.2 Frequency Analyses

The stationary points are more readily identified on the basis of the number of imaginary frequencies the molecule possesses. An equilibrium structure is characterized by possessing all real frequencies while a transition structure contains one imaginary

frequency. Transition structures containing more than one imaginary frequency are possible, however, these species are generally not of chemical interest.

The harmonic vibrational frequencies of a structure are calculated by evaluation of the second derivative energy matrix (Hessian matrix or force constant matrix) by analytical or numerical techniques. Analytical energy gradients may be obtained on the HF, MP2, QCISD and DFT surfaces while the computationally more demanding analytic second derivatives are available for the HF, MP2 and DFT surfaces. Calculation of the Hessian matrix provides thermochemical analysis of the system, thus providing the zero-point energy of the system. The zero-point vibrational energy (ZPVE), which accounts for the residual molecular motion for a system at 0 kelvin, should be included when determining the relative energies of different species.

An additional advantage of calculating the Hessian matrix for the transition-structures lies in the fact that the vibrational mode associated with the negative eigenvalue describes the reaction coordinate connecting the reactants and products for a molecular process. The connectivity of the transition structure to the relevant stationary points can be followed by intrinsic reaction coordinate (IRC)[41,42] calculations.

In order to obtain improved relative energies for a system, the geometry of a molecule calculated at low levels of theory, is held fixed and electronic properties are calculated at a higher level of theory than that used to obtain the geometry. These

calculations are referred to as single-point calculations since a single geometry is used rather than optimizing all of the geometrical parameters. This enables the determination of more accurate properties of the system at a reduced computational cost.

2.8.3 Notation

As previously alluded to, an examination of a system or PES may involve multiple calculations and subsequent energy corrections. The usual notation used to describe a particular method or theory is outlined in Equation **2.50**.

method 1 / basis set 1 // method 2 / basis set 2 (2.50)

This denotes a single-point calculation using method 1 with basis set 1 on a molecule whose geometry and frequencies were determined using method 2 and basis set 2. For example,

B3LYP / 6-311+G(2df,p) // B3LYP / 6-31G(d,p) (2.51)

describes a B3LYP single point calculation using the 6-311+G(2df,p) basis set at the geometry optimized with the B3LYP method and the 6-31G(d,p) basis set. As the quantum chemistry literature contains references for a plethora of basis sets, it is often useful to refer to review articles[38] to obtain information about choosing basis sets for molecular calculations.

2.9 Solvent Effects

All the methods alluded to above are gas-phase calculations. While the chemical structure and reactivity attained from such calculations are often sufficient, they are inadequate in that they do not describe the characteristics of molecules in solution. In order to estimate the effect of a solvent on the electronic properties of a system of interest, a number of schemes have been developed which differ in the manner by which they model the solvent.

The theory of solvent effects stems from the quantum mechanics of composite systems. One family of models, referred to as the self-consistent reaction field (SCRF) methods, use a reaction field or a continuum of uniform dielectric constant (ε) to model the solvent and place the solute in a cavity within the reaction field. The different types of SCRF methods differ in the manner by which they define the reaction field and cavity.

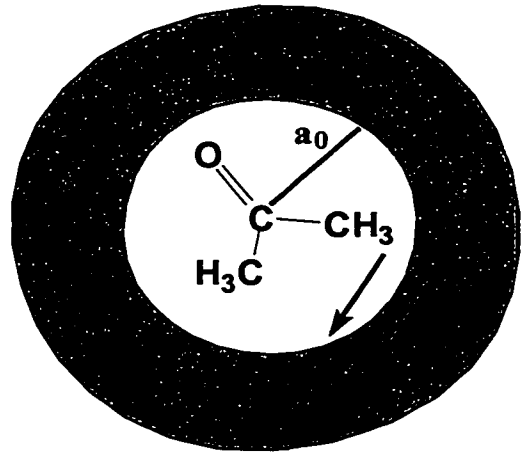


Figure 2.2: Schematic illustration of the Onsager model.

The simplest of the SCRF methods is the Onsager reaction field model.[43] In this method, schematically illustrated in Figure 2.2, the solute is assumed to exist as a fixed spherical cavity of radius a_o. As a dipole moment develops in the solute, an induced dipole moment is created within the solvent medium. A net stabilization results as the electric field applied due to the dipole moment in the solvent interacts with the dipole moment of the system of interest. Due to the physical robustness of the model, it has remained a widely used method to incorporate solvent effects. For a more detailed discussion on the various types of solvation models, one should examine the review articles[44-46] available in the literature.

2.10 References

(1) McQuarrie, D. A. *Quantum Chemistry*, University Science Books: California, 1983.

(2) Levine, I. N. *Quantum Chemistry*, Prentice Hall: New Jersey, 1991.

(3) Schrödinger, E. *Ann. Physik* **1926,** *79*, 361.

(4) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, John Wiley & Sons, Inc.: New York, 1986.

(5) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, MacMillan Publishing Co., Inc.: New York, 1982.

(6) (a) Hartree, D. R. *Proc. Cam. Phil. Soc.* **1928**, *24*, 89. (b) Hartree, D. R. *Proc. Cam. Phil. Soc.* **1928**, *24*, 111. (c) Hartree, D. R. *Proc. Cam. Phil. Soc.* **1928**, *24*, 246.

(7) (a) Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293. (b) Slater, J. C. *Phys. Rev.* **1930**, *35*, 509.

(8) Pauli, W. *Z. Physik.* **1925**, *31*, 57.

(9) Fock, V. *Z. Phys.* **1930**, *61*, 126.

(10) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(11) Hall, G. G. *Proc. Roy. Soc. (London)* **1951**, *A205*, 541.

(12) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

(13) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(14) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*, Oxford University Press: New York, 1989.

(15) Hohenberg, P.; Kohn, W. *Phys. Rev. B.* **1964**, *136*, 864.

(16) Kohn, W.; Sham, L. J. *Phys. Rev. A.* **1965**, *140*, 1133.

(17) Seminario, J. M.; Politzer, P. (editors) *Modern Density Functional Theory, A Tool for Chemistry*, Elsevier: New York, 1995.

(18) Dirac, P. A. M. *Proc. Cambridge Phil. Soc.* **1930**, *36*, 376.

(19) Slater, J. C. *Phys. Rev.* **1951**, *81*, 385.

(20) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(21) (a) Perdew, J. P. *Phys. Rev. B.* **1986**, *33*, 8822. (b) Perdew, J. P. *Phys. Rev. B.* **1986**, *34*, 7406.

(22) Perdew, J. P.; Wang, Y. *Phys. Rev. B.* **1992**, *45*, 13244.

(23) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785.

(24) (a) Langreth, D. C.; Mehl, M. *Phys. Rev. Lett.* **1981**, *47*, 446. (b) Langreth, D. C.; Mehl, M. *Phys. Rev. B.* **1983**, *28*, 1809. (c) Langreth, D. C.; Mehl, M. *Phys. Rev. B.* **1984**, *29*, 2310.

(25) Colle, R.; Salvetti, D. *Theor. Chim. Acta.* **1975**, *37*, 329.

(26) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098.

(27) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.

(28) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(29) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) A slightly modified form of the original three-parameter hybrid is implemented in the Gaussian programs: Stephens, P. J.; Devlin, F. J.; Chablowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(30) Adamo, C.; Barone, V. *Chem. Phys. Lett.* **1997**, *274*, 242.

(31) Sosa, C. P.; Carpenter, J. E.; Novoa, J. J. *Chemical Applications of Denisty-Functional Theory* (editors B. B. Laird, R. B. Ross and T. Ziegler), American Chemical Society, Washington D.C., 1996, pp.131

(32) Ziegler, T. *Chem. Rev.* **1991**, *91*, 651.

(33) Head-Gordon, M. *J. Phys. Chem.* **1996**, *100*, 13213.

(34) Baerends, E. J.; Gritsenko, O. V. *J. Phys. Chem.* **1997**, *101*, 5383.

(35) Hu, C. –H.; Chong, D. P. *Encycl. Comput. Chem.* **1998**, *1*, 664.

(36) St-Amant, A. *Rev. Comp. Chem.* **1996**, *7*, 187.

(37) Parr, R. G.; Yang, W. *Ann. Rev. Phys. Chem.* **1995**, *46*, 701.

(38) Davidson, E. R.; Feller, D. *Chem. Rev.* **1986**, *86*, 681.

(39) Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.

(40) Boyd, S. F. *Proc. Roy. Soc. (London)* **1950**, *A200*, 542.

(41) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154.

(42) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523.

(43) Onsager, L. *J. Am .Chem. Soc.* **1936**, *58*, 1486.

(44) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.

(45) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(46) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem. A.* **1998**, *102*, 7885.

Chapter Three.
A Density Functional Theory Study of the Hydrogen Halide Dimers

3.1 Introduction

Hydrogen bonding, a fundamental aspect of chemical structure and reactivity, has been the subject of intensive experimental and theoretical investigations. The formation of a X—H···Y hydrogen bond produces a low-frequency shift and an increase in the intensity of the fundamental X—H stretching vibration.[1-3] These two properties are characteristic of hydrogen-bonded systems and may be observed by infrared spectroscopy (IR), a major tool for the study of hydrogen bonds. The plethora of experimental data on hydrogen bonding, obtained by IR and other experimental techniques, is complemented

by information obtained from calculations performed using quantum chemical methods. In particular, *ab initio* methods have been shown not only to reproduce the principal properties of hydrogen-bonded systems but also to provide additional information that is not easily obtainable by other means. Moreover, the cadre of computational tools has been enhanced by the development of methods based on density functional theory (DFT). In addition to being computationally less expensive than the conventional electron correlation methods,[4] the DFT methods yield properties comparable in accuracy to those obtained using second-order Møller–Plesset perturbation theory (MP2) or quadratic configuration interaction (QCISD).[5]

One of the smallest and most extensively studied hydrogen-bonded system is $(HF)_2$. The structural and energetic parameters obtained from experimental studies[6-10] have stimulated several high level *ab initio* calculations.[11-17] Similarly, the heavier halide analogue $(HCl)_2$ has also been the subject of experimental studies[18-21] with the majority of theoretical investigations directed at an examination of the potential energy hypersurface (see overview in reference 22). Although $(HCl)_2$ is not regarded as a classically hydrogen-bonded system since the intermolecular interactions are dominated by dispersion forces,[23,24] it is geometrically similar to $(HF)_2$ and thus, the two dimers are often studied concurrently. DFT studies on $(HF)_2$[5,21,25-33,36] and $(HCl)_2$[25,27-29,34-36] have confirmed (see Table 3.1) that generalized gradient approximations (GGAs) yield more accurate results than the local density approximation (LDA) and Hartree-Fock (HF)

Table 3.1: Previous density functional calculations on the hydrogen halide dimers.

Reference	System, DF method and basis set
5	$(HF)_2$, FH/CO, FH/NH_3 SCF, MP2, SVWN, BLYP, BP86, B3LYP 6-31G(d,p), 6-31+G(d,p), 6-311++G(d,p), 6-311+G(2d,2p), 6-311+G(3df,3pd), DZ(d,p), cc-pVDZ, aug-cc-pVDZ, cc-pVTZ, aug-cc-pVTZ
25	$(HF)_2$, $FH/NCCH_3$, HCl/NCH, $HCl/NCCH_3$, HCl/H_2O, $(H_2O)_2$, FH/NCH B3LYP 6-31G(d,p), 6-31+G(d,p)
26	$(HF)_2$, $(H_2O)_2$, $(NH_3)_2$, C_2H_2/H_2O, $CH_4/H2O$ S-null, B-null, X_α, SVWN, BVWN, SLYP, BLYP, B3LYP cc-pVDZ, 6-31++G(2d,2p), aug-cc-pVDZ
27	$(HF)_2$, $(HCl)_2$, $(H_2S)_2$, FH/NH_3, $(HCN)_2$, $(H_2O)_2$ VWN, BP86, PW86P86 (7111/411/1) C, N, O, F; (7321/621/1) S, Cl; (41/1) H
28	$(HF)_2$, $(HCl)_2$, $(H_2O)_2$, FH/CO, FH/OC, FH/NH_3, ClH/NH_3, H_2O/NH_3, H_3O^+/H_2O HF, MP2, LDA, B3LYP, B97-1, PBE0, HCTH, BLYP, PBE, HCTH38 TZ2p, cc-pVTZ, cc-pVQZ, aug-cc-pVTZ, aug-cc-pVQZ
30	$(HF)_2$, $(HCl)_2$, Li^+/OH_2, F^-/OH_2, $(C_6H_6)_2$, $(C_2H_4)_2$, $(Ne)_2$, $(Ar)_2$, C_6H_6/Ne, C_6H_6/Ar, $(H_2O)_2$, BLYP, B3LYP, B3P86 6-31G(d), 6-31G(d,p), 6-31++G(d,p), 6-311G(d), D95(d,p), D95(2d,2p)
32	$(HF)_2$ + oligomers RHF, MP2, BLYP, B3LYP 6-311++G(d,p)
33	$(HF)_2$ HF, BLYP, B3LYP, CHA/BLYP, CHA/B3LYP 6-31G, 6-31G(d), 6-31G(d,p), 6-31++G(d,p), 6-311G(d,p), 6-311++G(3df,3pd)
34	$(HF)_2$ SVWN, SLYP, BVWN, BPL, BLYP, BP86, B3LYP, B3P86, BHandHLYP 6-311++G(d,p), 6-311++G(2d,2p), 6-311++G(3d,3p), 6-311++G(2df,2pd), 6-311++G(3df,3pd)
35	$(HCl)_2$, $(HCl)_3$,$(HCl)_4$ HF, MP2, BVWN 3-21+G(d,p), 6-311+G(d,p)
36	$(HF)_2$, $(HCl)_2$, $(H_2O)_2$, FH/HCN, FH/H_2O, CN^-/H_2O, OH^-/H_2O , HCC^-/H_2O, H_3O^+/H_2O, NH_3^+/H_2O MP2, B3LYP, BHandHLYP, PBE, VSXC, PBE1PBE, meta-GGA 6-311++G(d,p)

methods.[28] Despite this fact, no detailed study has been reported on the ability of the hybrid density functional methods to compute accurately the properties of these systems.

In this chapter, a study of the geometries, binding energies and vibrational properties of the series of HX (X=F, Cl and Br) dimers has been performed to assess the reliability and feasibility of the hybrid density functional methods.

3.2 Computational Details

Throughout this thesis, the Gaussian 98[37] suite of programs was used to perform the calculations. In this study, the seven density functional methods, described in detail in Chapter Two, considered included SVWN, B3LYP, B3P86, B3PW91, BHandH, BHandHLYP and B1LYP. Hartree-Fock (HF), *ab initio* second-order Møller-Plesset (MP2) and quadratic configuration interaction (QCISD) methods, with the core electrons frozen for the latter two methods, are employed for comparison.

Frequency analyses were performed to ensure that the stationary points are local minima and to correct the binding energies with zero-point vibrational energy (ZPVE). The binding energy was computed at the aforementioned levels of theory and was corrected for the basis set superposition error (BSSE) using the counterpoise (CP) correction as originally proposed by Boys and Bernardi.[38] As outline in Equation **3.1**, $D_{o(ZPVE + BSSE)}$, the binding energy of the dimer corrected for ZPVE and BSSE, is obtained

by subtracting the CP correction from the ZPVE corrected binding energy, $D_{o(ZPVE)}$. The CP correction is stated in Equation **3.2**, where the E_{m^i}'s represent the energy of the individual monomers, "*f*" frozen in their dimer geometry and the "*" represents monomers calculated with ghost atoms.

$$D_{o(BSSE+ZPVE)} = D_{o\,(ZPVE)} - CP \tag{3.1}$$

$$CP = \sum_{i=1}^{m} \left(E_{m_f^i} - E_{m_f^{i*}} \right) \tag{3.2}$$

On the basis of a preliminary study with a large variety of basis sets, we concluded that the 6-311+G(2df,p) basis set is sufficiently large for the present purposes. The basis set dependence of the optimized structures of $(HF)_2$ and $(HCl)_2$ has been investigated with a range of Pople and Dunning basis sets (see Tables A1.1 and A1.2 in Appendix A) and the inclusion of diffuse and two sets of d-functions on the heavy atoms is necessary for obtaining reliable geometrical parameters. Unless otherwise stated, bond distances are reported in angstroms (Å), angles are in degrees, energies are in kJ mol^{-1} and dipole moments (μ) are in debyes (D).

3.3 Results and Discussion

It is now firmly established that the lowest energy conformation for the hydrogen halide dimers is a bent structure of C_s symmetry, in which the bridging proton lies close

to the internuclear axis with the acceptor moiety almost perpendicular to this axis (Figure 3.1). In assessing the performance of various methods, three geometrical parameters were examined: the internuclear (X···X′) distance, the linearity of the hydrogen bond as reflected by θ_1 and the orientation of the proton acceptor molecule relative to the internuclear X···X′ axis as measured by θ_2.

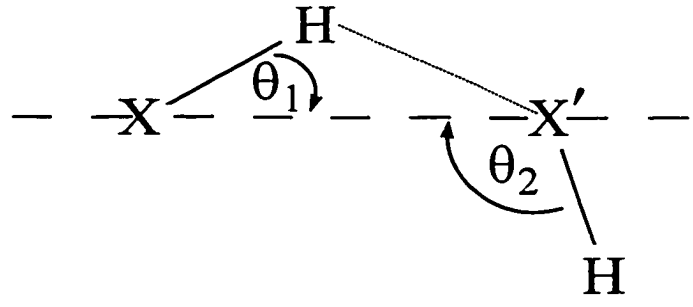


Figure 3.1: Schematic illustration of the conformation of the HX dimers.

A cyclic structure of almost C_{2h} symmetry has been obtained with the 6-31G(d,p) basis set and various DFT and MP2 methods.[2] This structure is in fact a transition structure corresponding to interconversion of the donor-acceptor units in the complex.

3.3.1 Geometries of the (HX) Dimers

The optimized geometrical parameters for $(HF)_2$, $(HCl)_2$ and $(HBr)_2$ are summarized in Tables 3.2, 3.3 and 3.4, respectively. The internuclear (F···F′) distance

for $(HF)_2$, experimentally determined to be 2.72 Å, is significantly underestimated by 0.17 Å with the SVWN method. Methods containing the B3 exchange functional (B3LYP, B3P86 and B3PW91) or the LYP correlation functional (BHandHLYP and B1LYP) predict (F···F′) distances within the range of experimental error (2.72 ± 0.03 Å). The BHandH method and the post-HF methods, MP2 and QCISD, all overestimate the (F···F′) distance by at least 0.03 Å.

Table 3.2: Optimized geometrical parameters for $(HF)_2$ obtained with the 6-311+G(2df,p) basis set and various theoretical methods.

Method	**R(F···F′)**	θ_1	θ_2	μ
SVWN	2.547	9.8	113.7	3.52
B3LYP	2.738	8.5	112.5	3.48
B3P86	2.704	8.6	111.1	3.45
B3PW91	2.751	8.6	112.3	3.44
BHandH	2.757	8.4	111.1	3.61
BHandHLYP	2.715	8.4	115.2	3.54
B1LYP	2.744	8.5	113.4	3.48
HF	2.828	8.6	112.2	3.64
MP2	2.759	7.3	116.3	3.66
QCISD	2.775	7.4	117.0	3.65
Experiment[a]	2.72 ± 0.03	7 ± 3	120 ± 2	3.2

[a] References 6 and 7.

With respect to the linearity of the hydrogen bond in $(HF)_2$, all hybrid density functional methods yield about 8.5° for θ_1. Although this agrees well with the experimental value (7 ± 3°), it is slightly larger than that obtained with MP2 or QCISD. This suggests that the linearity of the hydrogen bond may be overestimated by the hybrid density functional methods.

On the other hand, all DFT methods predict values for θ_2 that are well below the experimental value (120 ± 2°). Methods containing the B3 exchange functional (B3LYP, B3P86 and B3PW91) underestimate θ_2 by 7° to 9°. Similarly, the B1LYP and SVWN methods underestimate θ_2 by about 7°. The BHandHLYP, MP2 and QCISD methods underestimate θ_2 by 3 to 5°. Thus, the DFT methods predict the proton acceptor to be in an orientation more perpendicular to the internuclear axis than that observed experimentally.

The experimental results for $(HCl)_2$ indicate that the internuclear distance elongates to 3.75 – 3.84 Å, as θ_1 increases to 9° and θ_2 decreases to 90°. The (Cl···Cl′) distance, underestimated by at least 0.15 Å with the SVWN and BHandH methods, is elongated by about 0.06 Å with the B3LYP, B3PW91, BHandHLYP, B1LYP and MP2 methods. Reasonable agreement with experiment is attained with the B3P86 method. All DFT methods underestimate θ_1 by at least 2° but predict values of θ_2 with deviations of at

most 4° from 90°. Thus, the DFT methods overestimate the linearity of the hydrogen bond but adequately predict the perpendicular orientation of the proton acceptor relative to the internuclear axis.

Table 3.3: Optimized geometrical parameters for $(HCl)_2$ obtained with the 6-311+G(2df,p) basis set and various theoretical methods.

Method	R(Cl···Cl')	θ_1	θ_2	μ
SVWN	3.465	5.6	87.3	2.54
B3LYP	3.884	6.2	92.8	2.11
B3P86	3.773	5.5	92.5	2.18
B3PW91	3.889	6.1	93.1	2.12
BHandH	3.605	7.5	89.0	2.24
BHandHLYP	3.901	7.1	93.1	2.05
B1LYP	3.912	5.6	93.9	2.09
HF	3.898	4.8	93.7	1.85
MP2	3.883	8.3	87.9	2.03
QCISD	4.001	9.3	87.7	1.98
Experiment [a]	3.75 – 3.84	9	90	1.5

[a] References 3, 4, 18 and 21.

For $(HBr)_2$, the largest dimer considered, the internuclear distance is further elongated to 4.17 Å. This dimer contains a nearly linear hydrogen bond and the proton acceptor molecule is once again in a perpendicular orientation relative to the internuclear axis, as exemplified by the values for θ_1 and θ_2 of 3.4° and 92.7°, respectively.

Table 3.4: Optimized geometrical parameters for $(HBr)_2$ obtained with the 6-311+G(2df,p) basis set and various theoretical methods.

Method	R(Br···Br')	θ_1	θ_2	μ
SVWN	3.746	5.0	82.9	2.29
B3LYP	4.213	2.8	91.8	1.75
B3P86	4.071	3.5	89.4	1.84
B3PW91	4.184	2.9	91.1	1.78
BHandH	3.905	6.3	85.1	1.93
BHandHLYP	4.256	3.2	91.9	1.72
B1LYP	4.253	2.9	92.1	1.72
HF	4.660	11.4	87.5	1.43
MP2	4.164	3.9	90.1	1.75
QCISD	4.421	4.3	90.3	1.65
Experiment [a]	4.174	3.4	92.7	---

[a] References 23.

The SVWN and BHandH methods underestimate the (Br···Br') distance, while all remaining DFT methods predict a (Br···Br') distance within 0.1 Å of 4.17 Å. Unexpectedly, all methods, except SVWN and BHandH, predict reasonably well the linearity of the hydrogen bond yielding values of θ_1 within 0.5° of 3.4°. Similarly, all methods predict values of θ_2 that are at most 3° smaller than 92.7°.

Tables 3.2 to 3.4 also contain the dipole moment calculated for each dimer with all methods. Although the dipole moment is insensitive to the choice of the DFT method, all methods overestimate this parameter. In general, as the size of the halide atom increases from F to Cl and finally to Br, the internuclear distance elongates as expected. This elongation is accompanied by an enhancement in the linearity of the hydrogen bond and a more perpendicular arrangement of the proton acceptor. That is, the magnitude of θ_1 decreases from about 8.5° for $(HF)_2$ to about 3.4° for $(HBr)_2$ and θ_2 contracts from about 120° for $(HF)_2$ to about 90° for $(HBr)_2$.

3.3.2 Binding Energies of the (HX) Dimers

The calculated binding energies, corrected for ZPVE and BSSE, are summarized in Table 3.5. For $(HF)_2$, whose binding energy is moderate in strength at 12.7 kJ mol^{-1}, the SVWN and BHandH methods grossly overestimate this value. The DFT methods containing the LYP functional predict binding energies that deviate from the experimental value by 1.5 kJ mol^{-1} or less. The B3P86, B3PW91, MP2 and QCISD methods all underestimate the binding energy by 2 to 5 kJ mol^{-1}.

The binding energy, experimentally determined to be 5.9 - 8.4 kJ mol^{-1} for $(HCl)_2$, has not been reported for $(HBr)_2$. For $(HCl)_2$, the SVWN and BHandH methods predict binding energies larger than the other methods, although the BHandH value (9.1 kJ mol^{-1}) is in reasonable agreement with experiment. Methods containing the LYP

functional (B3LYP, BHandHLYP and B1LYP) predict binding energies of 1 – 2 kJ mol^{-1}, similar to that obtained with the QCISD method. The MP2 binding energy, although slightly larger than that predicted by those methods containing the LYP functional, is still underestimated.

Table 3.5: Binding energies (kJ mol^{-1}) for $(HF)_2$, $(HCl)_2$ and $(HBr)_2$ corrected for BSSE and ZPVE ($D_{o\ (BSSE\ +\ ZPVE)}$), with the 6-311+G(2df,p) basis set and various theoretical methods.

Method	$(HF)_2$		$(HCl)_2$		$(HBr)_2$	
	$D_{o\ (BSSE\ +\ ZPVE)}$	BSSE	$D_{o\ (BSSE\ +\ ZPVE)}$	BSSE	$D_{o\ (BSSE\ +\ ZPVE)}$	BSSE
SVWN	25.9	2.9	14.5	2.2	15.8	1.2
B3LYP	11.2	2.0	1.1	1.5	0.7	0.7
B3P86	10.4	2.1	1.3	1.8	1.5	0.8
B3PW91	8.2	2.0	3.2	1.7	-0.3	0.8
BHandH	22.9	2.6	9.1	1.9	8.6	1.0
BHandHLYP	12.7	2.0	1.6	1.5	0.9	0.7
B1LYP	11.1	2.0	0.9	1.5	0.3	0.7
HF	9.1	1.5	0.3	1.2	0.3	0.2
MP2	9.6	3.8	2.3	2.7	-1.6	1.7
QCISD	9.1	3.5	1.3	2.3	1.4	1.2
Experiment	12.70 ± 0.01[a]		5.9 – 8.4 [b]		N/A	

[a] Reference 9.
[b] Reference 21.

For $(HBr)_2$, the SVWN and B1LYP methods both predict binding energies larger and thus stronger than all other methods. The B3LYP, BHandHLYP and B1LYP methods predict binding energies similar in magnitude and smaller than 1 kJ mol^{-1}. Conversely, binding energies larger than 1 kJ mol^{-1} are obtained with the B3P86 and QCISD methods. Interestingly, the MP2 and B3PW91 methods predict that the dimers are unstable relative to the separated monomers.

The binding energies obtained for $(HCl)_2$ and $(HBr)_2$ using the hybrid DFT methods illustrate the inherent deficiency in these methods, namely the inability to accurately reproduce binding energies in complexes where dispersion interactions play an important role. The peculiar results obtained for the binding energies of the heavier halide dimers, $(HCl)_2$ and $(HBr)_2$, suggest that although the 6-311+G(2df,p) basis set yields adequate geometrical parameters for $(HF)_2$ and $(HCl)_2$ with the post-HF methods (see Appendix A), it is insufficient to describe the binding energy in these heavier halide dimers. A more comprehensive investigation to address the importance of relativistic effects and the possible correlation between the geometrical parameters, particularly the internuclear distance and θ_1, and binding energy in these larger dimers is warranted.

3.3.3 Vibrational Properties of the (HX) Dimers

The shift in the frequency of the X—H stretching mode, as a consequence of hydrogen bonding, is a quantity accessible by infrared spectroscopy. The average,

unscaled harmonic X—H frequency shift has been calculated for $(HF)_2$, $(HCl)_2$ and $(HBr)_2$ and is summarized in Table 3.6 along with the resulting elongation of the X—H bond. The experimental frequency shifts are 105.0, 53 and 60 cm^{-1}, for $(HF)_2$, $(HCl)_2$ and $(HBr)_2$, respectively.

Table 3.6: Frequency shift ($-\Delta\nu$) and change in the X—H donor bond length, upon dimerization for $(HF)_2$, $(HCl)_2$ and $(HBr)_2$, with the 6-311+G(2df,p) basis set and various theoretical methods.

	$(HF)_2$		$(HCl)_2$		$(HBr)_2$	
Functional	$\Delta\nu$	$\Delta r_{H—F}$	$\Delta\nu$	$\Delta r_{H—Cl}$	$\Delta\nu$	$\Delta r_{H—Br}$
SVWN	313.4	0.016	256.1	0.021	292.2	0.025
B3LYP	147.3	0.007	67.3	0.004	72.6	0.005
B3P86	174.9	0.008	101.7	0.009	111.1	0.009
B3PW91	153.9	0.007	73.7	0.005	82.9	0.006
BHandH	213.0	0.010	122.8	0.009	122.7	0.009
BHandHLYP	129.7	0.006	48.3	0.003	41.7	0.004
B1LYP	138.1	0.007	60.3	0.005	59.2	0.005
MP2	112.9	0.005	43.8	0.003	46.5	0.003
QCISD	83.1	0.005	17.8	0.002	14.5	0.001
Experiment	105.0 [a]	-----	53 [b]	-----	60 [b]	-----

[a] Reference 10.
[b] Reference 39.

All DFT methods overestimate the $(HF)_2$ frequency shift. The SVWN and BHandH methods predict frequency shifts at least twice as large as experiment. Methods containing the LYP functional (B1LYP, BHandHLYP and B3LYP) or the B3 functional (B3P86 and B3PW91) overestimate the frequency shift by 25 to 40 cm^{-1} and 50 to 70 cm^{-1}, respectively. Very reasonable agreement with experiment is attained with the MP2 method although the QCISD method underestimates the frequency shift by about 25 cm^{-1}.

For $(HCl)_2$ and $(HBr)_2$, the SVWN and BHandH methods overestimate the frequency shift. The B3P86 functional also overestimates the shift but yields a value similar to that observed with the BHandH functional. Methods containing the LYP functional (B3LYP, BHandHLYP and B1LYP) predict frequency shifts in reasonable agreement with experiment. However, the post-HF methods, MP2 and QCISD, underestimate the frequency shift for $(HCl)_2$ by 9 and 35 cm^{-1}, respectively and for $(HBr)_2$ by 15 and 45 cm^{-1}, respectively.

Upon closer examination of the change in the X—H bond length, the direct relationship between the elongation of the X—H bond upon complexation and the frequency shift is apparent. The linear correlation between the elongation of the X—H bond length and the frequency shift associated with the X—H stretching mode for the three dimers investigated in this study is illustrated in Figure 3.2.

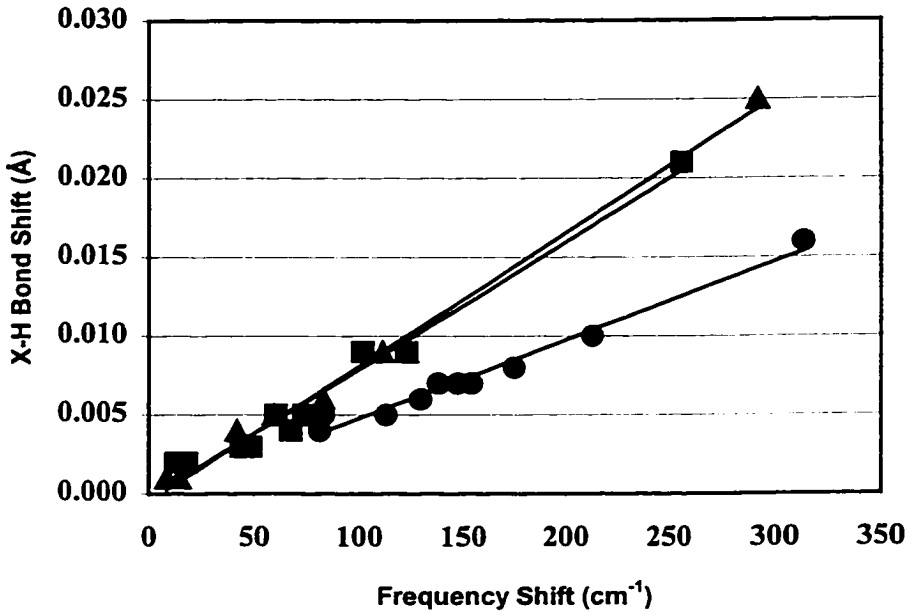


Figure 3.2: Correlation between the change in bond length and frequency shift for the X—H bond upon dimerization. {● $(HF)_2$; ■$(HCl)_2$; ▲$(HBr)_2$ }

3.3.4 Trends Amongst the (HX) Dimers

The hydrogen halides represent an interesting set of binary molecules. Through self-association, these molecules may create long chains in condensed phases or form smaller clusters in the gas phase. Thus, they are often employed as prototypes in the study of hydrogen bonding.

As one progresses down the periodic table, the size of the halide atom increases and the intermolecular (X···X′) distance present in the dimers lengthens from 2.72 Å for $(HF)_2$ to 3.80 Å for $(HCl)_2$ and finally to 4.17 Å for $(HBr)_2$. Upon dimer formation, the

X—H bond present in the proton donor elongates by up to 0.010 Å. The non-linearity of the hydrogen bond is measured in this study by θ_1 and is less than 10° for all dimers. In addition, the orientation of the acceptor molecule, as described by θ_2, approaches a more perpendicular orientation as the size of the halide ion in the dimer increases. That is, θ_2 contracts from 120° for $(HF)_2$ to around 90° for $(HCl)_2$ and $(HBr)_2$. Similar trends were observed in an analogous *ab initio* study[23] of $(HCl)_2$, $(HBr)_2$ and $(HI)_2$.

3.4 Conclusions

A study of the performance of a variety of hybrid density functional methods has been carried out on $(HF)_2$, $(HCl)_2$ and $(HBr)_2$. Assessment of the geometrical parameters for the dimers indicates that all hybrid functionals predict geometrical parameters in reasonable agreement with experiment. For $(HF)_2$, it was found that the methods containing the LYP correlation functional yield optimized geometrical parameters similar to each other and to the values obtained using the MP2 or QCISD methods. While a similar tendency was evident for $(HCl)_2$ and $(HBr)_2$, it was found that the BHandHLYP and B1LYP methods yield geometrical parameters in slightly better agreement with experiment.

For $(HF)_2$, the DFT binding energies, which are similar in magnitude to those obtained with the MP2 and QCISD methods, are too large by up to 4 kJ mol^{-1}. The DFT

binding energies for $(HCl)_2$ and $(HBr)_2$ are similar in magnitude to the MP2 and QCISD values but are in poor agreement with the experimental value for $(HCl)_2$.

Finally, the linear correlation between the shift in the X—H stretching frequency and the X—H bond length for the three dimers upon dimerization is illustrated. The B1LYP, B3LYP and BHandHLYP methods yield shifts in the stretching frequency which are in reasonable agreement with the experimental and MP2 values.

Through the study of the fundamental physical properties for a series of hydrogen-bonded dimers, it is apparent that the hybrid density functional methods perform with a reasonable degree of accuracy. Methods containing the LYP correlation functional tend to yield results similar not only to the MP2 and QCISD values but also to the experimental results. In general, the BHandHLYP and B1LYP methods offer a viable alternative to the popular B3LYP method.

3.5 References

(1) Liddel, U.; Wulf, O. R. *J. Am. Chem. Soc.* **1933**, *55*, 3574.

(2) Hilbert, G. E.; Wulf, O. R.; Hendricks, S. B.; Liddel, U. *J. Am. Chem. Soc.* **1936**, *58*, 1936.

(3) Hendricks, S. B.; Wulf, O. R.; Hilbert, G. E.; Liddel, U. *J. Am. Chem. Soc.* **1936**, *58*, 1991.

(4) Barrone, V.; Adamo, C.; Russo, N. *Chem. Phys. Lett.* **1993**, *212*, 5.

(5) Civalleri, B.; Garrone, E.; Ugliengo, P. *J. Molec. Struct.(Theochem)* **1997**, *419*, 227.

(6) Howard, B. J.; Dyke, T. R.; Klemperer, W. *J. Chem. Phys.* **1984**, *81*, 5417.

(7) Gutowsky, H. S.; Chuang, C.; Keen, J. D.; Klots, T. D.; Emilsson, T. *J. Chem. Phys.* **1985**, *83*, 2070.

(8) Pine, A. S.; Lafferty, W. J. *J. Chem. Phys.* **1983**, *78*, 2154.

(9) Miller, R. E. *Acc. Chem. Res.* **1990**, *23*, 10.

(10) Andrews, L.; Bondyberg, V. E.; English, J.A. *J. Chem. Phys.* **1984**, *81*, 3452.

(11) Klopper, W.; Lüthi, H. P. *Mol. Phys.* **1999**, *96*, 559.

(12) Tschumper, G. S.; Kelty, M. D.; Schaefer III, H. F. *Mol. Phys.* **1999**, *96*, 493.

(13) Peterson, K. A.; Dunning Jr., T. H. *J. Phys. Chem.* **1995**, *102*, 2032.

(14) Collins, C. L.; Morihashi, K; Yamaguchi, Y.; Schaefer III, H. F. *J. Chem. Phys.* **1995**, *103*, 6051.

(15) Klopper, W.; Quack, M.; Suhm, M. A. *J. Chem. Phys.* **1998**, *108*, 10096.

(16) Tschumper, G. S.; Yamaguchi, Y.; Schaefer III, H. F. *J. Chem. Phys.* **1997**, *106*, 9627.

(17) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgenson, P.; Taylor, P. R. *J. Chem. Phys.* **1999**, *111*, 9157.

(18) Ohashi, N.; Pine, A. S. *J. Chem. Phys.* **1984**, *81*, 73.

(19) Pine, A. S.; Howard, B. J. *J. Chem. Phys.* **1986**, *84*, 590.

(20) Powles, J. G.; Wojcik, M. *J. Chem. Phys.* **1983**, *78*, 5277.

(21) Blake, G. A.; Busarow, K. L.; Cohen, R. C.; Laughlin, K. B.; Lee, Y. T.; Saykally, R. J. *J. Chem. Phys.* **1988**, *89*, 6577.

(22) Hermida-Ramón, J. M.; Engkvist, O.; Karlström, G. *J. Comp. Chem.* **1998**, *19*, 1816.

(23) Latajka, Z.; Scheiner, S. *Chem. Phys.* **1997**, *216*, 37.

(24) Tao, F. -M.; Klemperer, W. *J. Chem. Phys.* **1995**, *103*, 950.

(25) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Phys. Chem.* **1995**, *99*, 10705.

(26) Sosa, C. P.; Novoa, J. J. *J. Phys. Chem.* **1995**, *99*, 15837.

(27) Jeanvoine, Y.; Bohr, F.; Ruiz-López, M. F. *Can. J. Chem.* **1995**, *73*, 710.

(28) Tuma, C.; Boese, A. D.; Handy, N. C. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3939.

(29) Maerker, C.; Schleyer, P. v. R.; Liedl, K. R.; Ha, T.-K.; Quack, M.; Suhm, M. A. *J. Comp. Chem.* **1997**, *18*, 1695.

(30) Hobza, S.; Sponer, J.; Reschel, T. *J. Comp. Chem.* **1995**, *16*, 1315.

(31) Quack, M.; Suhm, M. A. *Theor. Chim. Acta.* **1996**, *93*, 61.

(32) Hirata, S.; Iwata, S. *J. Phys. Chem. A.* **1998**, *102*, 8426.

(33) Salvador, P.; Fradera, X.; Duran, M. *J. Chem. Phys.* **2000**, *112*, 10106.

(34) Latajka, Z.; Bouteiller, Y. *J. Chem. Phys.* **1994**, *101*, 9793.

(35) Chandler, W. D.; Johnson, K. E.; Fahlman, B. D.; Campbell, J. L. E. *Inorg. Chem.* **1997**, *36*, 776.

(36) Rabuck, A. D.; Scuseria, G. E. *Theor. Chem. Acc.* **2000**, *104*, 439.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN 98, Gaussian , Inc.: Pittsburgh PA, 1998.

(38) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(39) Barnes, A. J.; Davies, J. B.; Hallam, H. E.; Howells, J. D. R. *J. Chem. Soc. Faraday Trans. 2* **1973**, 69, 246.

Chapter Four.

The Hydrogen-Bond Mediated Aminolysis of 6-Chloropyrimidine

4.1 Introduction

Hydrogen bonds are an essential feature of the structure and function of biological molecules. Although an individual hydrogen bond is relatively weak compared to a covalent bond, the cooperative nature of multiple hydrogen bonds confers added stability to a complex.[1,2] This, in addition to the complementary nature of hydrogen bonding, has made hydrogen bonding an important interaction in the self-assembly of molecules.[1,3,4]

Due to the specificity of the donor-acceptor units and the inherent weakness of the individual bonds within a multiply-hydrogen bonded complex, molecules capable of forming multiple hydrogen bonds have been employed as catalysts in organic

and biological reactions.[3,5-8] Recently, nucleobases such as uracil have been utilized[9,10] as catalytic agents due to the robust variety of functional groups associated with these molecules.[11] In one of the initial studies aimed at assessing the catalytic ability of the individual functional groups of the nucleobases, Horne and Melander[12] examined the aminolysis of pentafluorophenyl benzoate by propylamine, a reaction involving the formation of an amide bond. Designed to assess the ability of each nucleobase to stabilize an ionic transition state through hydrogen bonding, this study explicitly revealed the inherent bifunctional catalytic ability of the functional groups of the nucleobases and laid the foundation for further investigations into the catalytic possibilities of hydrogen bonding.

Recently, Tominaga et al.[10] accelerated the aminolysis of 6-chloropurine in benzene by the addition of derivatives of uracil. On the basis of ^{1}H NMR evidence, the catalytic behavior was proposed to occur as the result of multiple hydrogen bonding interactions between the uracils and 6-chloropurine derivatives. The latter were presumed to assist the formation of a reactive intermediate and subsequent stabilization of the transition state, an application of hydrogen bonds which is well established.[4-7,13-17] In order to develop a rationalization for the role of the hydrogen bonding interactions in the aforementioned aminolysis reaction and the potential role of solvent effects, the aminolysis of 6-chloropyrimidine (Figure 4.1a) was considered. In order to assess the catalytic potential of the third hydrogen bond present in the aminolysis of 2-amino-6-chloropurine, the aminolysis of 2-amino-6-chloropyrimidine (Figure 4.1b) was examined.

In addition, recent interest in the proton affinity of the proton donor[18-20] involved in hydrogen bonding has prompted an investigation into the role of the hydrogen bond acceptor and the correlation between the proton affinity (PA) of the group that interacts with the incoming NH_3 and the barrier to aminolysis. To examine this correlation, fluoro, imine and thioketo derivatives of the uracil-derived bases OCH_2, $OHC\text{-}NH_2$ and OHC-NH-CHO were utilized in the aminolysis of 6-chloropyrimidine (Figure 4.1a).

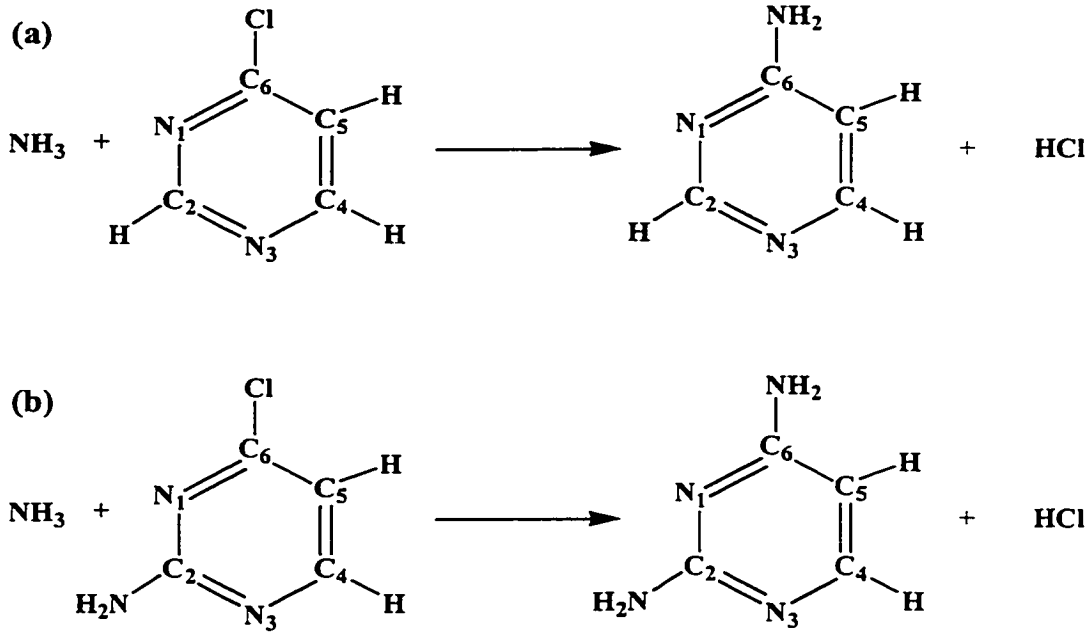


Figure 4.1: Schematic illustration of the aminolysis of **(a)** 6-chloropyrimidine and **(b)** 2-amino-6-chloropyrimidine.

4.2 Computational Details

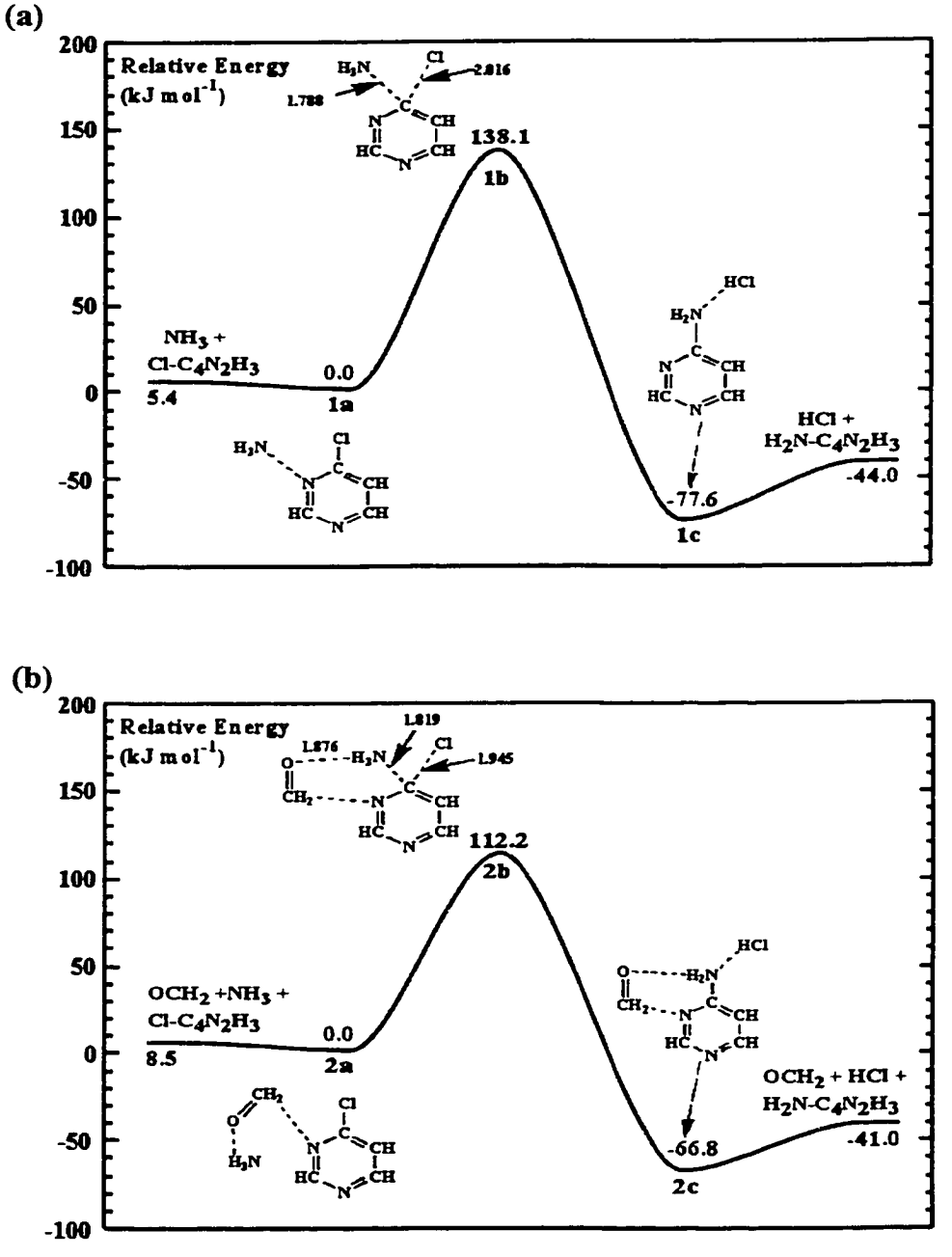
The potential energy surfaces of the aminolysis reactions were examined using density functional theory. Geometry optimizations, harmonic frequencies and zero-point vibrational energy (ZPVE) corrections were calculated with B3LYP/6-31G(d,p) (see Chapter Two). Improved relative energies, obtained by single point calculations at the B3LYP/6-311+G(2df,p) level using the above geometries, were corrected with the appropriate ZPVE, i.e., B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d,p) + ZPVE. All relative energies are in kJ mol^{-1} and bond lengths in angstroms (Å). Using the Onsager model (see Chapter Two), two solvents were considered: benzene with a dielectric constant of 2.28 and water with a dielectric constant of 78.39. The proton affinities (PA) of the uracil-derived bases were calculated as the difference between the protonated and non-protonated forms of the base. The total energies of all the molecules examined are summarized in Table B1.1 of Appendix B.

4.3 Results and Discussion

4.3.1 Initial Results

The isolated aminolysis of 6-chloropyrimidine (hereafter denoted as Cl-$C_4N_2H_3$) is illustrated schematically in Figure 4.2a. Initially, the reactants NH_3 + Cl-$C_4N_2H_3$ form the hydrogen bonded complex **1a** lying just 5.4 kJ mol^{-1} lower in energy. Aminolysis proceeds via transition structure (TS) **1b**, with a notable barrier of 138.1 kJ mol^{-1}, in

which the C···Cl bond is 2.016 Å and the C···NH_3 bond is 1.788 Å. The aminolysis products, HCl + H_2N-$C_4N_2H_3$, are produced as TS **1b** relaxes to generate the hydrogen bonded complex, **1c**, lying 77.6 kJ mol^{-1} lower in energy than the reactants.



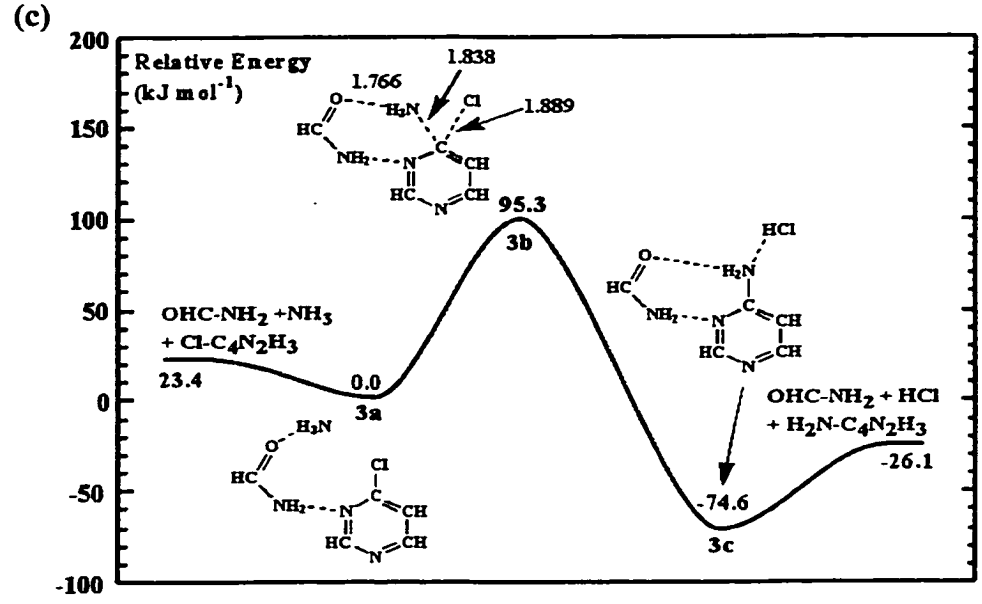


Figure 4.2: Schematic energy profile for the aminolysis of 6-chloropyrimidine **(a)** isolated, with **(b)** OCH_2 hydrogen bonded to the NH_3 moiety, and with **(c)** $OCH\text{-}NH_2$ hydrogen bonded to both the NH_3 moiety and the pyrimidine ring.

Based on the hydrogen bonding motif of uracil, the effect of OCH_2 and $OHC\text{-}NH_2$ on the aminolysis reaction of Figure 4.1a was examined. As illustrated in Figure 4.2b, the addition of the electron-donating OCH_2 species to NH_3 + $Cl\text{-}C_4N_2H_3$ yields complex **2a** lying 8.5 kJ mol^{-1} lower in energy. As OCH_2 remains bound to the incoming NH_3 via an O···HNH_2 hydrogen bond of 1.876 Å, aminolysis proceeds via TS **2b** with a barrier of 112.2 kJ mol^{-1}, 25.9 kJ mol^{-1} lower than that observed in the isolated system.

The OHC-NH_2 moiety, which contains both an electron-donating –OHC group and an electron-accepting –NH_2 group, interacts with NH_3 + Cl-$C_4N_2H_3$ as shown in Figure 4.2c. Initially, OHC-NH_2 forms the hydrogen-bonded complex **3a** lying 23.4 kJ mol^{-1} lower in energy than the reactants. As OHC-NH_2 remains hydrogen-bonded to the incoming NH_3, aminolysis proceeds via TS **3b** with a substantially reduced barrier of 95.3 kJ mol^{-1}. In TS **3b**, H_2NCHO is hydrogen bonded to both the incoming NH_3 by the O···H—NH_2 bond (1.766 Å) and to N1 of the pyrimidine ring by a N···HN bond of 1.906 Å.

In the transition structure of the isolated aminolysis (TS **1b**), the C···Cl and C···NH_3 distances are 2.016 Å and 1.788 Å, respectively. The addition of the OCH_2 moiety enables the formation of a strong H_2CO···HNH_2 bond in TS **2b** of 1.876 Å. This results in a shortening of the C···Cl distance to 1.945 Å and a lengthening of the C···NH_3 distance to 1.819 Å. Furthermore, the charges on the leaving Cl and incoming N, as illustrated in Figure 4.3, are smaller in magnitude than those observed in the TS for the isolated aminolysis. This suggests that the electron-donating ability of N in NH_3 is enhanced by hydrogen bonding to OCH_2, which enables the transition structure for aminolysis to occur earlier.

Enlarging the hydrogen bonding moiety from OCH_2 to OCH-NH_2 results in a further decrease in the C···Cl distance to 1.889 Å while the C···NH_3 distance is

elongated to 1.838 Å, indicating that the TS occurs even earlier when OCH-NH_2 is used as the base. In addition, the charges of the departing Cl and incoming N (Figure 4.3) are even smaller in magnitude than that observed in the TS involving OCH_2. Thus, the electron-donating –CHO group enhances the electron-donating ability of N in NH_3 by the formation of the O···H—N hydrogen bond. Furthermore, the electron-accepting $–NH_2$ group stabilizes the negative charge present on N1 of the pyrimidine ring, a consequence of the enhancement of the sp^3 hybridization of the carbon undergoing substitution.

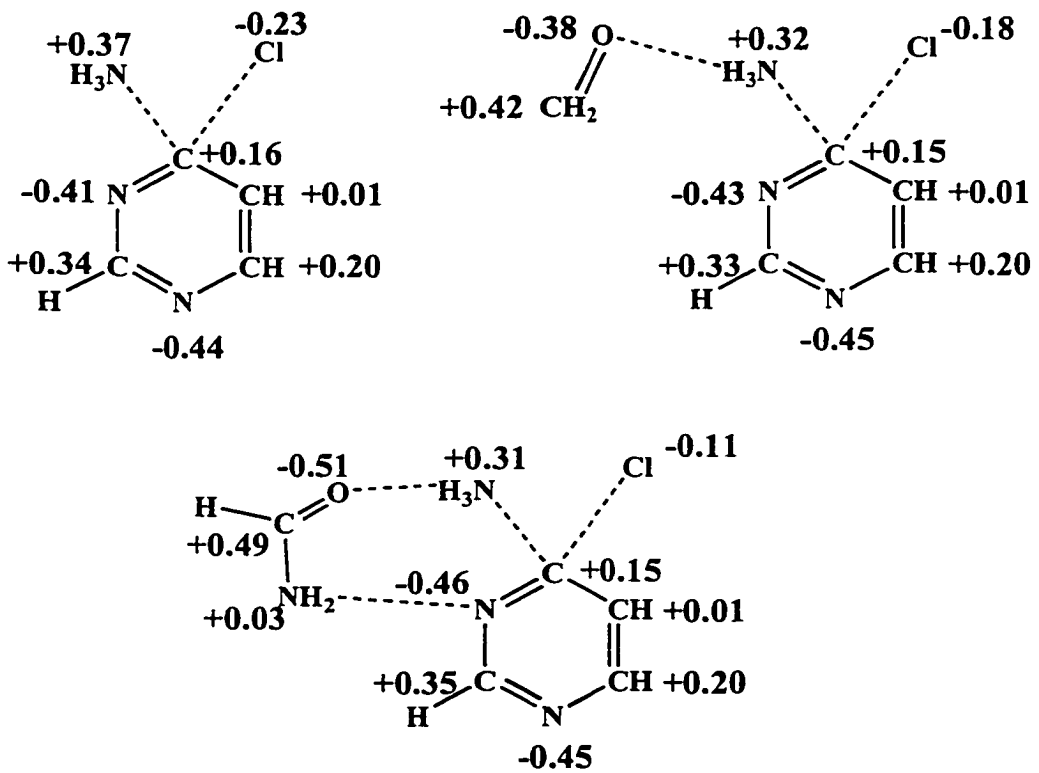


Figure 4.3: Schematic illustration of the transition structures **1b**, **2b** and **3b** showing the charge distribution on the heavy atoms as determined by Mulliken population analysis.

4.3.2 The Effect of Solvent

Thus, through the addition of small representative bases such as OCH_2 and OHC-NH_2, which form hydrogen bonds to the incoming NH_3 moiety and the pyrimidine ring, the barrier to aminolysis may be notably reduced. However, in the original study performed by Tominaga and coworkers,[10] the catalytic behavior of the uracil derivatives on the aminolysis of 2-amino-6-chloropurine was observed in benzene, a relatively nonpolar solvent. In order to assess the potential role solvent may play in stabilizing the partial ionic character present in the transition structures, the above aminolysis reactions were re-examined in benzene and water by using the Onsager model.

Table 4.1: Summary of the barrier to aminolysis (kJ mol^{-1}) in various media.

Reaction of Interest	*Gas Phase*	*Benzene*	*Water*
NH_3 + Cl-$C_4N_2H_3$	138.1	137.5	136.4
OCH_2 + NH_3 + Cl-$C_4N_2H_3$	112.2	111.9	110.8
OHC-NH_2 + NH_3 + Cl-$C_4N_2H_3$	95.3	95.0	94.6

In Table 4.1, the barriers to aminolysis for the various reactions are summarized in the three media of interest: gas phase, benzene and water. Benzene, a relatively non-ionizing solvent has a negligible effect on the barriers to aminolysis. The barrier to aminolysis are decreased by less than 1 kJ mol^{-1}. Water, a much stronger ionizing

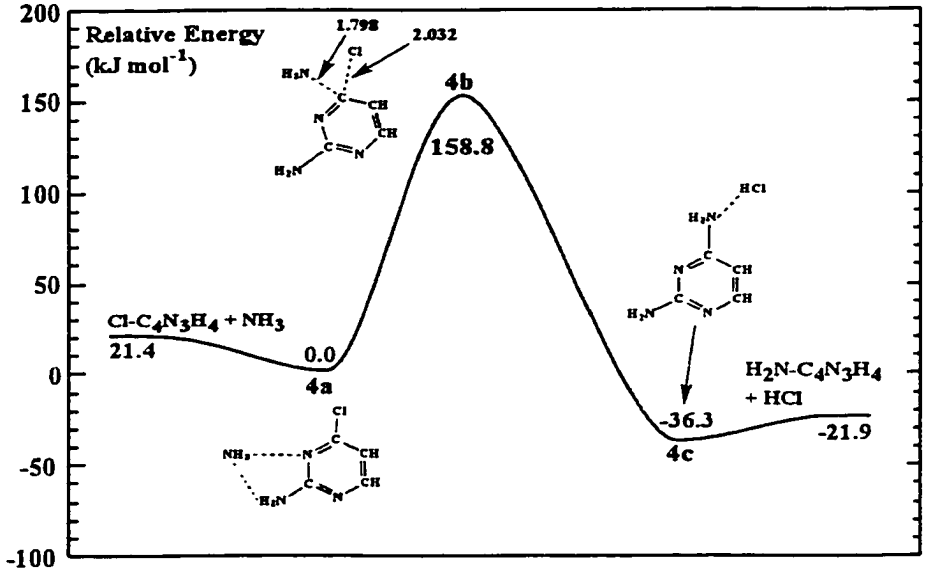
solvent has a slightly more pronounced effect on the aminolysis reaction but, once again, the decrease in the barriers is of a mere 1-2 kJ mol^{-1} in magnitude. Thus, the effect of the solvent on the barrier to aminolysis is much less than that due to specific hydrogen-bonding interactions present between the small bases (OCH_2 and $OHC\text{-}NH_2$), NH_3 and $Cl\text{-}C_4N_2H_3$.

4.3.3 The Role of the Third Hydrogen Bond

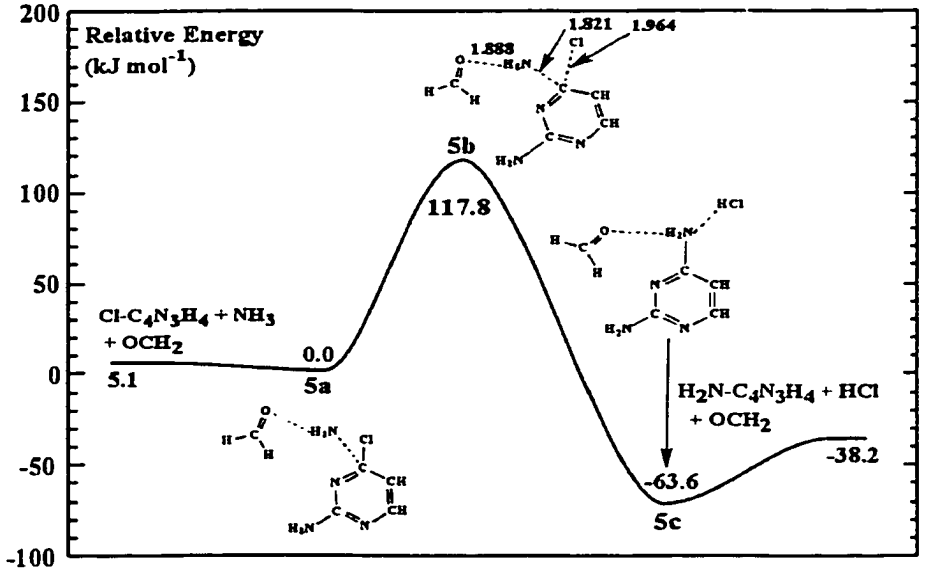
The above results illustrate the important role hydrogen bonding plays in catalyzing the aminolysis of 6-chloropyrimidine. However, Tominaga and co-workers[10] also examined the aminolysis of 2-amino-6-chloropurine. In order to assess the role of the electron-donating $-NH_2$ group present in the aforementioned aminolysis, the aminolysis of 2-amino-6-chloropyrimidine was examined (Figure 4.1b) using OCH_2, $OHC\text{-}NH_2$, OHC-NH-HCO and 1-methyluracil as bases.

In the isolated aminolysis of 2-amino-6-chloropyrimidine (NH_3 + $Cl\text{-}C_4N_3H_4$), the reactants generate the initial complex **4a** (Figure 4.4a) lying 21.4 kJ mol^{-1} lower in energy. Aminolysis proceeds via TS **4b** with a sizeable barrier of 158.8 kJ mol^{-1}. The addition of OCH_2 to NH_3 + $Cl\text{-}C_4N_3H_4$ (Figure 4.4b) forms complex **5a** lying 5.1 kJ mol^{-1} lower in energy. As OCH_2 remains bound to the incoming NH_3 moiety by a $O\cdots HNH_2$ hydrogen bond of 1.888 Å in TS **5b**, the aminolyisis proceeds with a barrier of 117.8 kJ mol^{-1}.

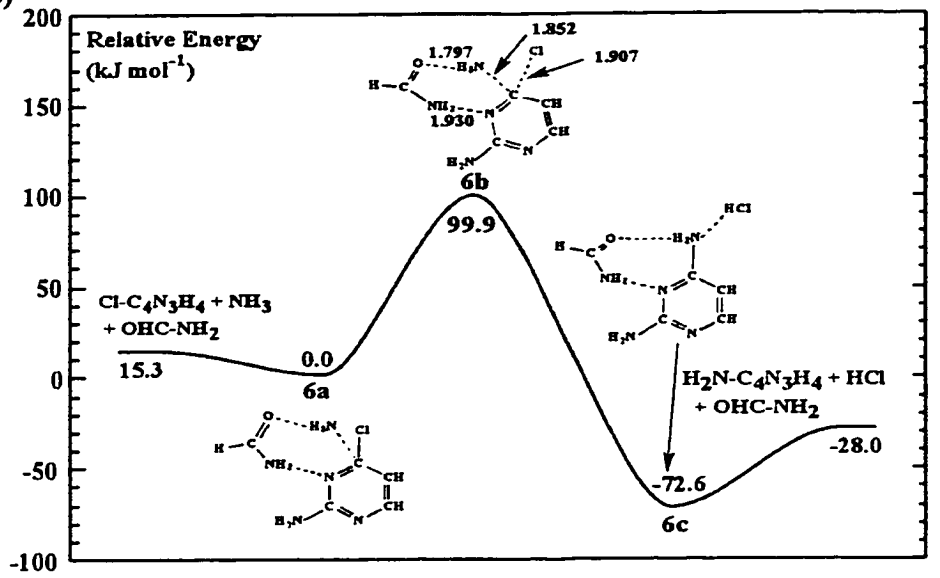
(a)


Relative Energy
(kJ mol^{-1})
200
150
100
50
0
-50
-100
1.798
2.032
4b
158.8
Cl-C$_4$N$_3$H$_4$ + NH$_3$
21.4
0.0
4a
-36.3
4c
H$_2$N-C$_4$N$_3$H$_4$
+ HCl
-21.9

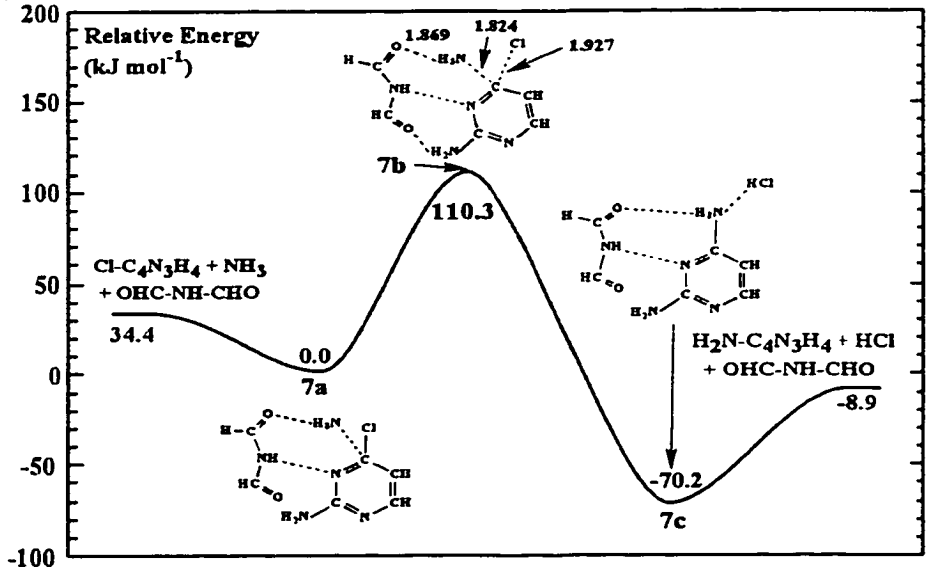
(b)


Relative Energy
(kJ mol^{-1})
200
150
100
50
0
-50
-100
1.821
1.964
1.588
5b
117.8
Cl-C$_4$N$_3$H$_4$ + NH$_3$
+ OCH$_2$
5.1
0.0
5a
H$_2$N-C$_4$N$_3$H$_4$ + HCl
+ OCH$_2$
-63.6
5c
-38.2

(c)


Relative Energy
(kJ mol^{-1})
200
150
100
50
0
-50
-100
1.797
1.852
1.907
1.930
6b
99.9
Cl-C$_4$N$_3$H$_4$ + NH$_3$
+ OHC-NH$_2$
15.3
0.0
6a
H$_2$N-C$_4$N$_3$H$_4$ + HCl
+ OHC-NH$_2$
-28.0
-72.6
6c

(d)


Relative Energy
(kJ mol^{-1})
200
150
100
50
0
-50
-100
1.869
1.824
1.927
7b
110.3
Cl-C$_4$N$_3$H$_4$ + NH$_3$
+ OHC-NH-CHO
34.4
0.0
7a
H$_2$N-C$_4$N$_3$H$_4$ + HCl
+ OHC-NH-CHO
-8.9
-70.2
7c

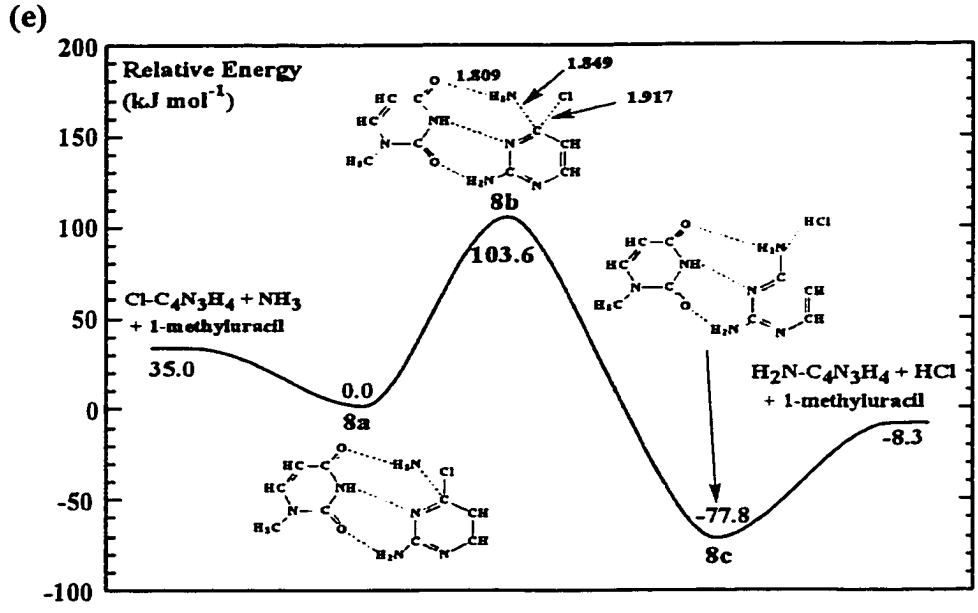


Figure 4.4: Schematic energy profile for the aminolysis of 2-amino-6-chloropyrimidine, **(a)** isolated, **(b)** with OCH_2 hydrogen bonded to the NH_3, **(c)** with OHC-NH_2 hydrogen bonded to both the incoming NH_3 and the pyrimidine ring, **(d)** with OHC-NH-HCO hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring, and **(e)** with 1-methyluracil hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring.

Complex **6a**, lying 15.3 kJ mol^{-1} lower in energy, is generated upon the addition of OHC-NH_2 to NH_3 + Cl-$C_4N_3H_4$ (Figure 4.4c). As the electron-donating –OHC group forms a shorter O···HNH_2 hydrogen bond of 1.797 Å to the incoming NH_3, aminolysis proceeds via TS **6b** with a barrier of 99.9 kJ mol^{-1}, in which the electron-accepting –NH_2 group forms a N···HN hydrogen bond of 1.930 Å to N1 of the pyrimidine ring.

With OHC-NH-CHO as the base, the initial complex **7a** lying 34.4 kJ mol^{-1} lower in energy is generated in the aminolyis of NH_3 + Cl-$C_4N_3H_4$ (Figure 4.4d). As OHC-NH-CHO remains bound to the incoming NH_3 via an O···HNH_2 bond (1.869 Å) and bound to Cl-$C_4N_3H_4$ via a N···HN bond (1.725 Å), aminolysis proceeds via TS **7b** with a noticeable increase in the barrier to 110.3 kJ mol^{-1}.

Finally, the addition of 1-methyluracil to Cl-$C_4N_3H_4$ (Figure 4.4e) generates the initial complex **8a** lying 24.7 kJ mol^{-1} lower in energy. Aminolysis proceeds via TS **8b** with a barrier of 103.6 kJ mol^{-1} in which 1-methyluracil forms a O···HNH_2 bond (1.809 Å) to NH_3 and a N···HN bond (1.792 Å) to N1 of the pyrimidine ring. In the latter two aminolysis reactions, the $-NH_2$ group of N1 of Cl-$C_4N_3H_4$ is engaged in a O···H—N bond of 2.079 Å and 2.025 Å for OHC-NH-CHO and 1-methyluracil, respectively.

In order to facilitate comparison between the two aminolysis reactions, OHC-NH-CHO and 1-methyluracil were utilized as bases in the aminolysis of 6-chloropyrimidine. The aminolysis of NH_3 + Cl-$C_4N_2H_3$ (6-chloropyrimidine) with OHC-NH-CHO as the base, proceeds via TS **9b** (Figure 4.5a) with a barrier of 100.8 kJ mol^{-1}. With 1-methyluracil as the base, the aminolysis of 6-chloropyrimidine proceeds via TS **10b** with a barrier of 95.2 kJ mol^{-1} (Figure 4.5b).

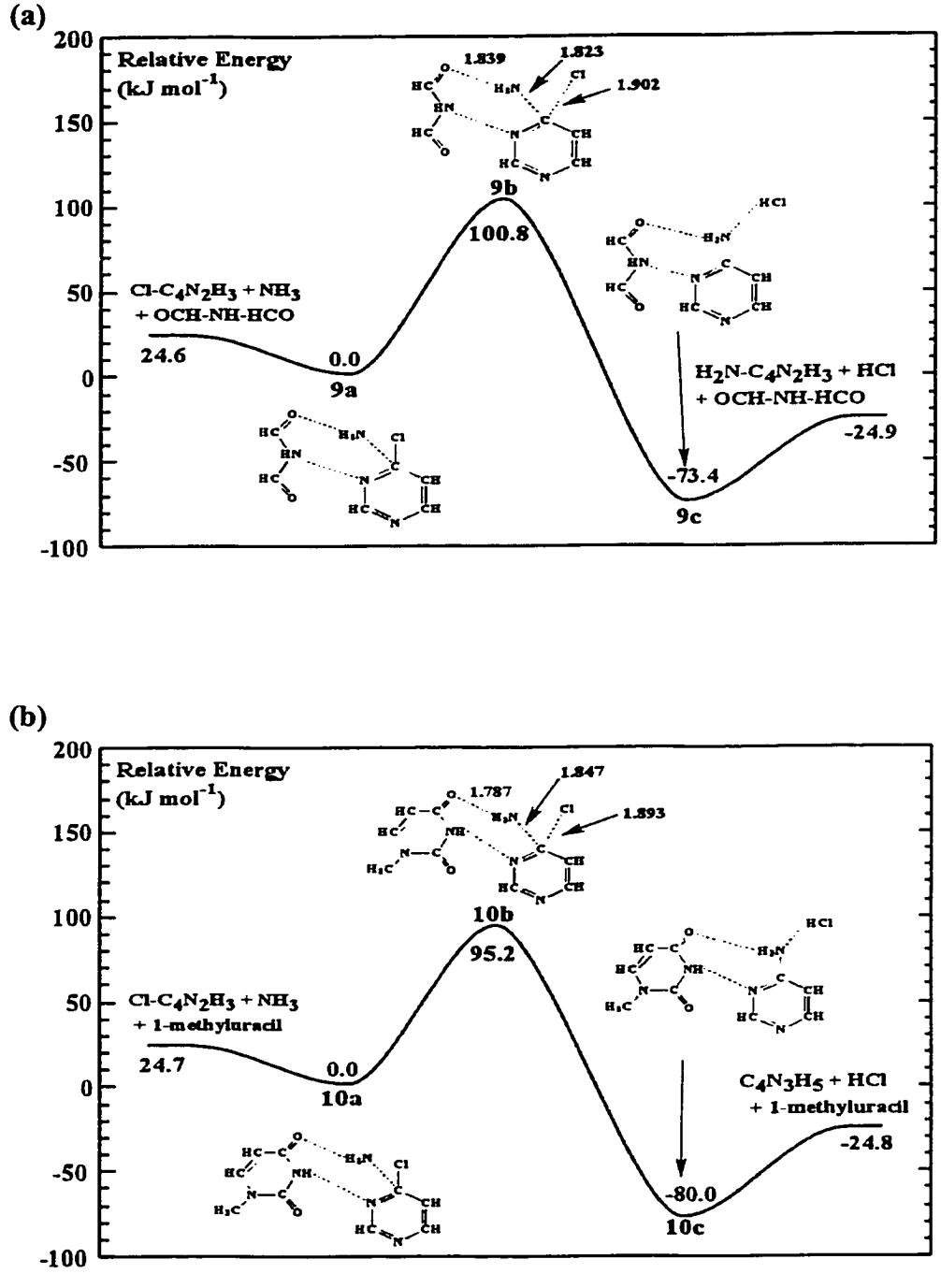


Figure 4.5: Schematic energy profile for the aminolysis of 6-chloropyrimidine with **(a)** OHC-NH-HCO hydrogen bonded to both the incoming NH_3 and the pyrimidine ring, and **(b)** 1-methyluracil hydrogen bonded to both the incoming NH_3 and the pyrimidine ring.

The barriers to aminolysis for 6-chloropyrimidine and 2-amino-6-chloropyrimidine are summarized in Table 4.2. As the barriers for the aminolysis of 2-amino-6-chloropyrimidine are larger than those observed for the analogous aminolysis of 6-chloropyrimidine, the presence of the $-NH_2$ group and its subsequent formation of an HN—H···O hydrogen bond does not play a catalytic role in the aminolysis reaction. Relative to the aminolysis of 6-chloropyrimidine, the presence of the amino group in 2-amino-6-chloropyrimidine increases the barriers by 5.6 kJ mol^{-1} and 4.6 kJ mol^{-1}, for OCH_2 and $OHC\text{-}NH_2$, respectively.

Table 4.2: Summary of the barriers to aminolysis (kJ mol^{-1}) for the two reactions of interest.

Model Base	$NH_3 + Cl\text{-}C_4N_2H_3$	$NH_3 + Cl\text{-}C_4N_3H_4$
Isolated	138.1	158.8
OCH_2	112.2	117.8
$OHC\text{-}NH_2$	95.3	99.9
OHC-NH-CHO	100.8	110.3
1-methyluracil	95.2	103.6

Due to the presence of the electron-donating $-NH_2$ group in the aminolysis of 2-amino-6-chloropyrimidine, the carbon undergoing substitution, C6, becomes less susceptible to nucleophilic attack by the incoming NH_3. As a consequence, the C···Cl distances in the TSs of the aminolysis of 2-amino-6-chloropyrimidine (Figure 4.6) are

elongated by ~0.02Å relative to that observed in the TSs for the aminolysis of 6-chloropyrimidine.

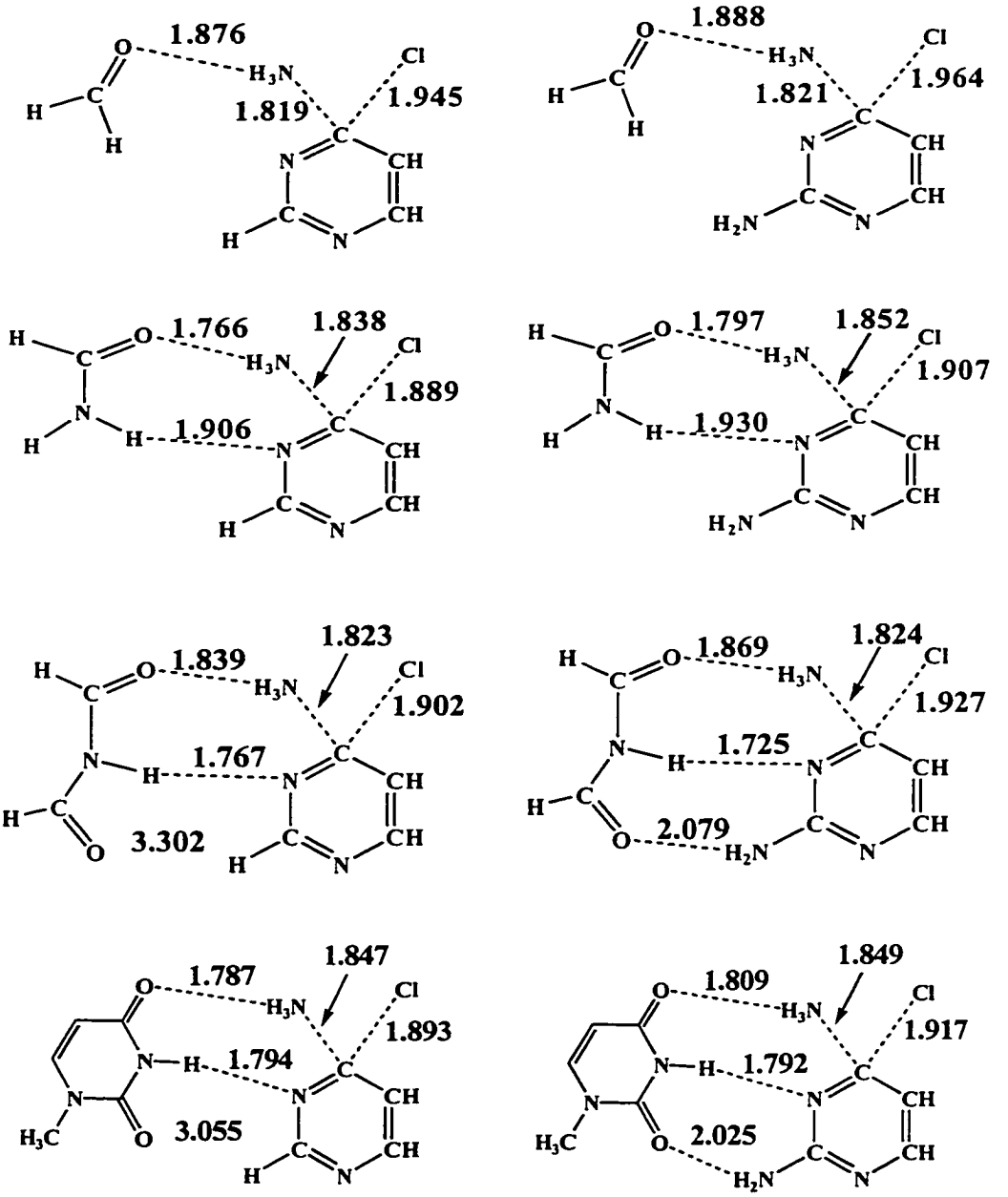


Figure 4.6: Pertinent geometrical parameters in the transition structures for the aminolysis reactions of 6-chloropyrimidine (left) and 2-amino-6-chloropyrimidine (right).

Enlarging the model base to OHC-NH-CHO, which forms a third hydrogen bond in the aminolysis of 2-amino-6-chloropyrimidine, leads to a barrier 9.5 kJ mol^{-1} larger than that observed in the aminolysis of 6-chloropyrimidine.

Due to the formation of the O···H—N hydrogen bond of 2.079 Å between the carbonyl group of the base and the $-NH_2$ group in the TSs for the aminolysis of 2-amino-6-chloropyrimidine (Figure 4.6), the electron donating ability of $-NH_2$ is enhanced. This results in an additional decrease in the electrophilicity of the C undergoing substitution. Similarly, the use of 1-methyluracil as the base in the aminolysis of 2-amino-6-chloropyrimidine leads to a barrier 8.4 kJ mol^{-1} larger than that observed in the corresponding aminolysis of 6-chloropyrimidine. In both these TSs, the C ···Cl distances are elongated to compensate for the decrease in the electrophilicity of C6. The Mulliken charges on the atoms in the TSs (not already presented in Figure 4.3) are summarized in Figure B1.1 of Appendix B.

Thus, the absence of a decrease in the barrier to aminolysis for 2-amino-6-chloropyrimidine and the minor changes in the transition structure geometry indicate that the third hydrogen bond does not play a catalytic role in the aminolysis reaction. However, the partial double bond character of the C–NH_2 bond provides a more rigid framework upon which the aminolysis reactions may proceed.

4.3.4 The Correlation with Proton Affinities

For the aminolysis of 6-chloropyrimidine, the barrier to aminolysis decreases from 112.2 kJ mol^{-1} for OCH_2 to 100.8 kJ mol^{-1} and 95.3 kJ mol^{-1} for the addition of OHC-NH-CHO and OHC-NH_2, respectively. For these three bases, the decrease in the barrier to aminolysis is correlated with a decrease in the O···HNH_2 distance from 1.876 Å to 1.839 Å and 1.766 Å. As summarized in Table 4.3, this is correlated with an increase in the proton affinity (PA) of the terminal carbonyl group of the base interacting with NH_3. As expected, the barrier to aminolysis is decreased further to 95.2 kJ mol^{-1} by the use of 1-methyluracil, whose terminal oxygen has the largest PA of the bases examined. However, there is a slight increase in the O···HNH_2 distance (1.787Å), a consequence of the electron-donating influence of the methyl group.

Thus, the barriers to aminolysis of 6-chloropyrimidine correlate with the PAs of the carbonyl oxygen of the base hydrogen bonding to the incoming NH_3 moiety. Although this correlation is prevalent in the aminolysis of 2-amino-6-chloropyrimidine, the model reaction of 6-chloropyrimidine was employed to further examine the correlation between the PA of the portion of the base that hydrogen bonds to the incoming NH_3 and the calculated barrier to aminolysis. To provide a range of PAs, derivatives of the carbonyl bases were examined in which the PA of the carbonyl-derived group was modified by fluoro, imine and sulfur substitution. The calculated PAs of the carbonyl oxygen in the uracil-derived bases are summarized in Table 4.3.

Table 4.3: Calculated proton affinities (kJ mol^{-1}) of the carbonyl oxygens in the uracil-derived bases.

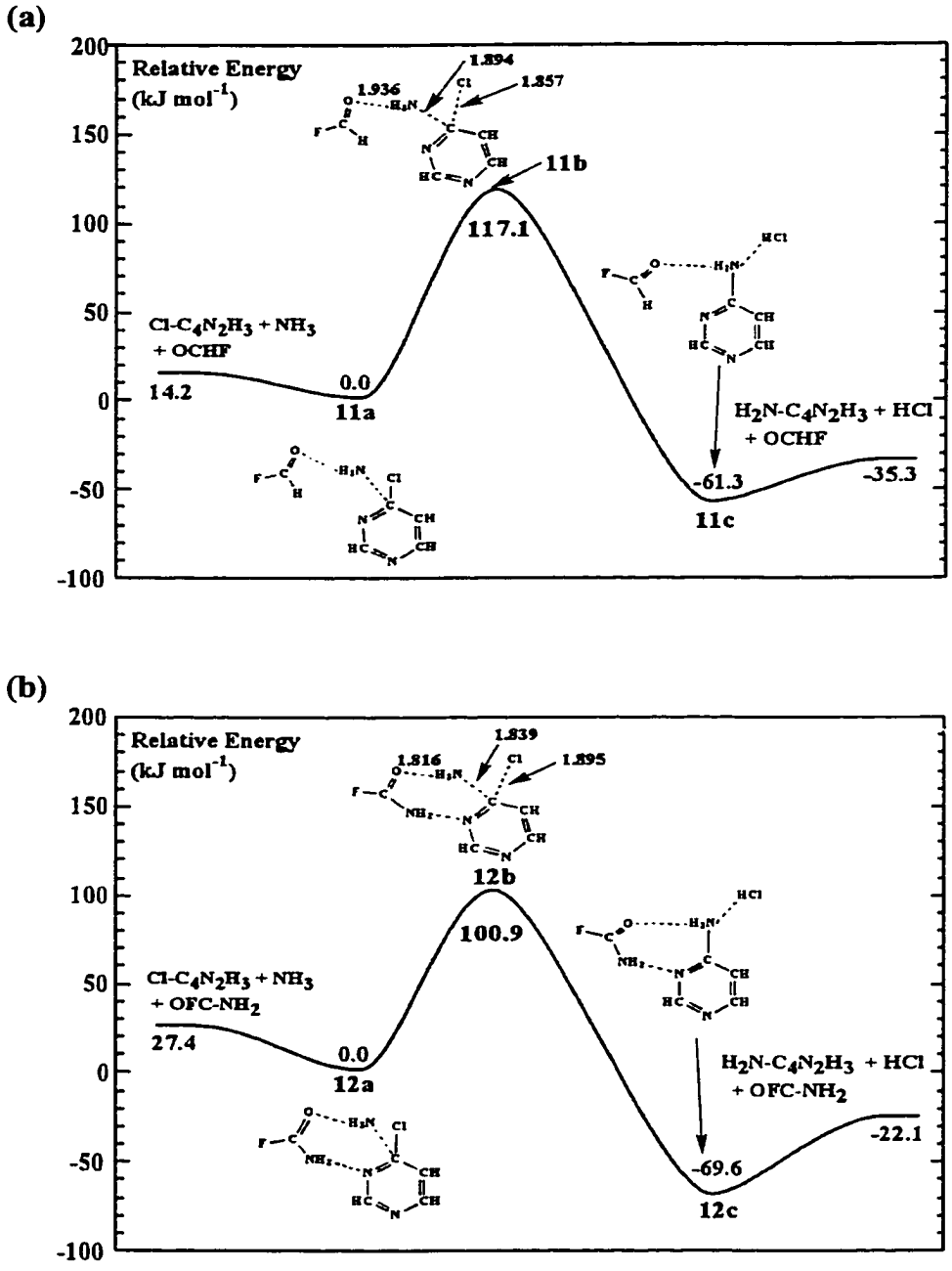
Base	Proton Affinities		
	X=O	X=S	X=NH
XCH_2	700.9	763.6	861.9
XHC-NH-CHO	789.9	801.8	896.2
$XHC-NH_2$	808.5	850.6	945.8
1-methyluracil [a]	861.5		
1-methyluracil [b]	831.2		
XCHF	646.8		
XFC-NH-CHO	701.1		
$XFC-NH_2$	743.7		

[a] Carbonyl oxygen that interacts with the incoming NH_3.
[b] Carbonyl oxygen that interacts with $-NH_2$ on 2-amino-6-chloropyrimidine.

4.3.4.1 Fluorine Substitution

By replacing a hydrogen adjacent to the carbonyl group by a fluorine, the fluorine-derived bases OCHF, $OFC-NH_2$ and OFC-NH-CHO are generated. The aminolysis of NH_3 + $Cl-C_4N_2H_3$, with OCHF acting as the base (Figure 4.7a), proceeds with a barrier of 117.1 kJ mol^{-1} and an O···HNH_2 bond of 1.936 Å in TS **11b**. With $OFC-NH_2$ as the base, the barrier to aminolysis of NH_3 + $Cl-C_4N_2H_3$ (Figure 4.7b) is reduced to 100.9 kJ mol^{-1} and a significantly shorter O···HNH_2 bond of 1.816 Å is present in TS **12b**. Finally, the aminolysis of NH_3 + $Cl-C_4N_2H_3$ with OFC-NH-CHO as the base (Figure 4.7c), proceeds with a barrier of 105.3 kJ mol^{-1} in which an O···HNH_2 bond of 1.888 Å is formed in the TS **13b**. Thus, as the barrier decreases from 117.1 kJ

mol^{-1} for OCHF, to 105.3 kJ mol^{-1} and 100.9 kJ mol^{-1} for OFC-NH-CHO and OFC-NH_2 respectively, the O···HNH_2 bond distance decreases from 1.936 Å to 1.888 Å and 1.816 Å, respectively.



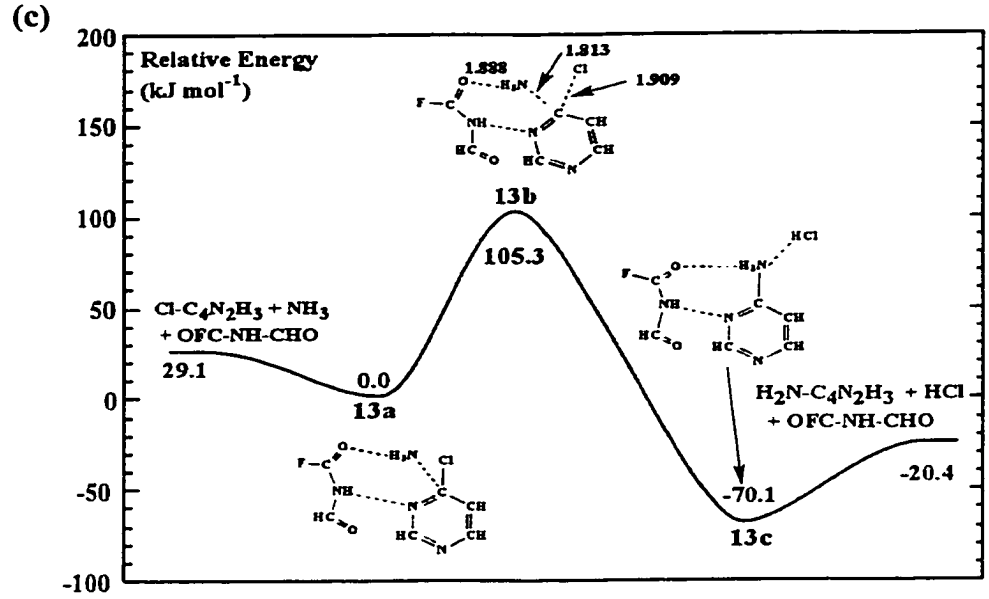


Figure 4.7: Schematic energy profile for the aminolysis of 6-chloropyrimidine, with **(a)** OCHF hydrogen bonded to the incoming NH_3 moiety, **(b)** OFC-NH_2 hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring, and **(c)** OFC-NH-CHO hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring.

This decrease in barrier height and hydrogen bond distance is correlated with an increase in the PA of the carbonyl group from 646.8 kJ mol^{-1} for OCHF to 701.1 kJ mol^{-1} and 743.7 kJ mol^{-1} for OFC-NH-CHO and OFC-NH_2, respectively (see Table 4.3). As the PA of the fluoro-substituted uracil bases is lower than that of unsubstituted uracil-derived bases, it follows that the barrier to aminolysis is larger by ~ 5 kJ mol^{-1} and the O···HNH_2 distances are longer by ~ 0.05 Å.

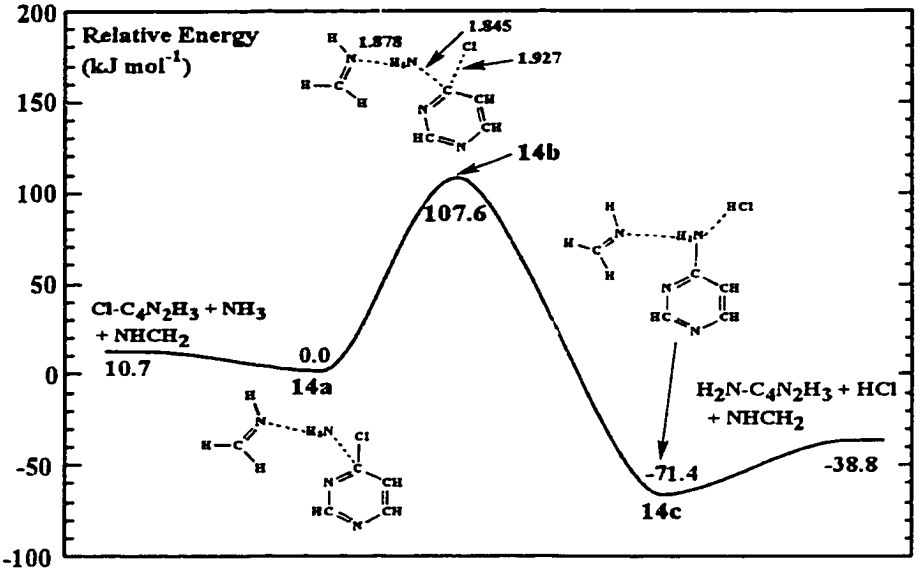
4.3.4.2 Imine Substitution

Due to the abundance of nitrogen containing species in biological systems and the evidence that the nitrogen in the imine group is a better proton acceptor than the carbonyl oxygen, and hence has a larger PA, the aminolysis of 6-chloropyrimidine was re-examined with the carbonyl oxygen replaced by an imine group to produce the bases, $HNCH_2$, HNCH-NH_2 and HNCH-NH-HCO. As $HNCH_2$ interacts with NH_3 + Cl-$C_4N_2H_3$ (Figure 4.8a), a N···HNH_2 hydrogen bond of 1.878 Å is present in TS **14b** and the aminolysis proceeds with a barrier of 107.6 kJ mol^{-1}. As the base is enlarged to HNCH-CH_2, the aminolysis of NH_3 + Cl-$C_4N_2H_3$ (Figure 4.8b), proceeds with a notably decreased barrier of 90.6 kJ mol^{-1} and a shorter N···HNH_2 bond of 1.800 Å in TS **15b**. The aminolysis of 6-chloropyrimidine, utilizing NHCH-NH-HCO as the base (Figure 4.8c), proceeds via TS **16b** with a barrier of 98.3 kJ mol^{-1} and an N···HNH_2 distance of 1.875 Å in the TS.

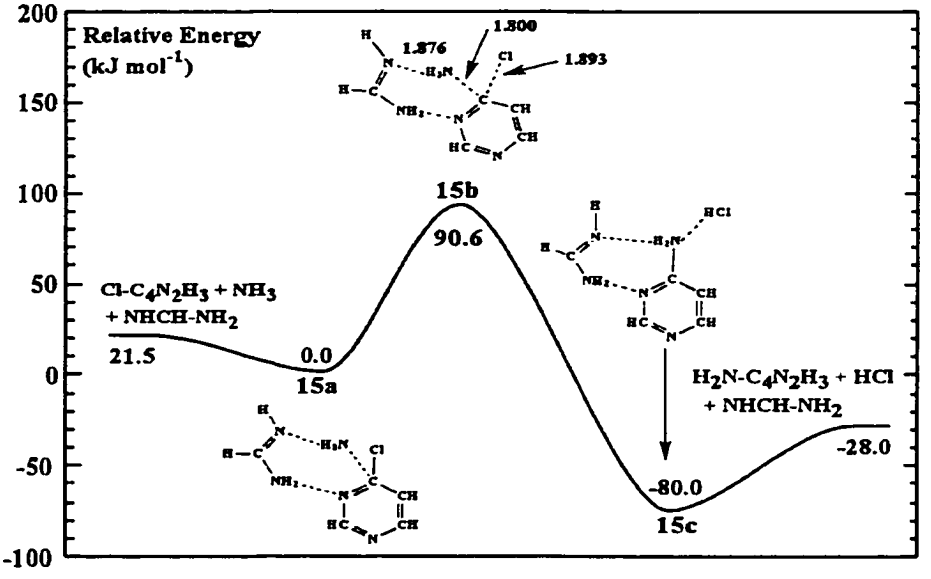
Thus, the barrier to aminolysis is reduced from 107.6 kJ mol^{-1} when $HNCH_2$ is used as the base to 98.3 kJ mol^{-1} and 90.6 kJ mol^{-1} with NHCH-NH-HCO and HNCH-NH_2, respectively. As the barrier decreases, the N···HNH_2 distance shortens from 1.878 Å with $HNCH_2$ as the base to 1.875 Å and 1.800 Å with NHCH-NH-HCO and HNCH-NH_2, respectively. The shortening of the N···HNH_2 distances is consistent with a sequential increase in the PA associated with the imine interacting with NH_3 from 861.9

kJ mol^{-1} for $HNCH_2$ to 896.2 kJ mol^{-1} and 945.8 kJ mol^{-1} for NHCH-NH-CHO and $NHCH-NH_2$, respectively (see Table 4.3).

(a)



(b)



(c)

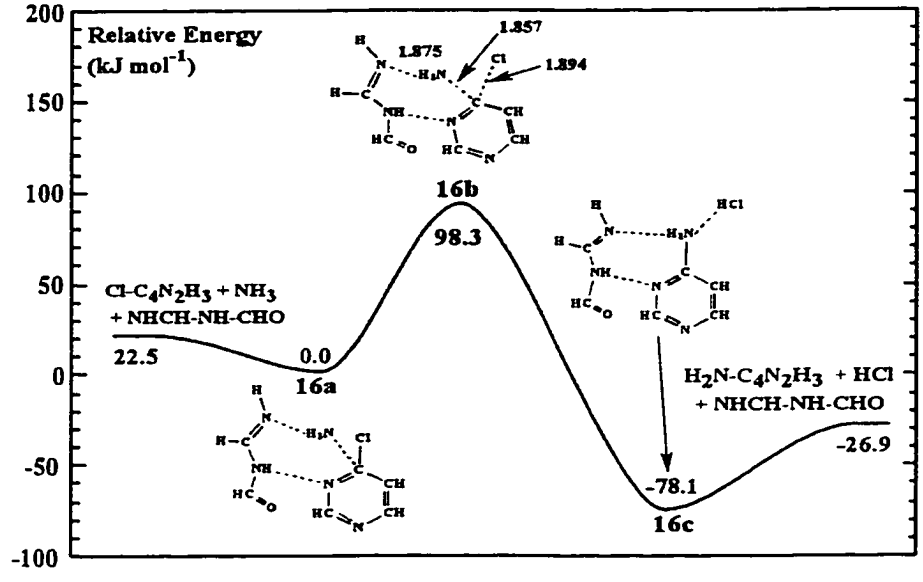


Figure 4.8: Schematic energy profile for the aminolysis of 6-chloropyrimidine, with **(a)** $NHCH_2$ hydrogen bonded to the incoming NH_3 moiety, **(b)** NHCH-NH_2 hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring, and **(c)** NHCH-NH-CHO hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring.

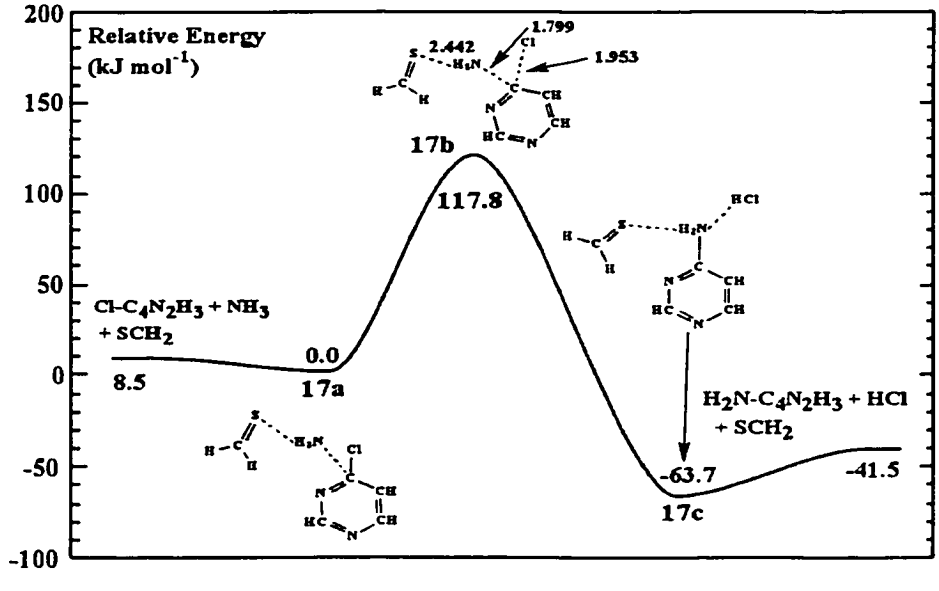
4.3.4.3 Sulfur Substitution

Sulfur may act as a hydrogen bond acceptor and is known to replace oxygen in this function. However, due to its larger size, sulfur is expected to act as a weaker hydrogen bond acceptor than oxygen. In order to determine if the correlation between the PA of the base that interacts with NH_3 and the barrier to aminolysis is maintained as the carbonyl-oxygen of the base is replaced by sulfur, the aminolysis of 6-

chloropyrimidine (Figure 4.1a) was examined with, SCH_2, SCH-NH_2 and SCH-NH-HCO as bases.

The aminolysis of 6-chloropyrimidine with SCH_2 as the base, Figure 4.9a, proceeds via TS **17b** with a barrier of 117.8 kJ mol^{-1} and an S···HNH_2 bond of 2.442 Å. The PA of the thioketo group is increased upon enlarging the base to SCH-NH_2 and the aminolysis proceeds via TS **18b** with a notably reduced barrier of 102.6 kJ mol^{-1} and a shortened S···HNH_2 bond (2.335 Å), Figure 4.9b. As the base is enlarged further to SCH-NH-HCO, which has a lower PA than that observed for SCH-NH_2, the aminolysis of 6-chloropyrimdine (Figure 4.9c) proceeds via TS **19b** with a slightly larger barrier of 106.9 kJ mol^{-1} and an elongation of the S···HNH_2 bond by 0.053 Å (2.388 Å).

(a)



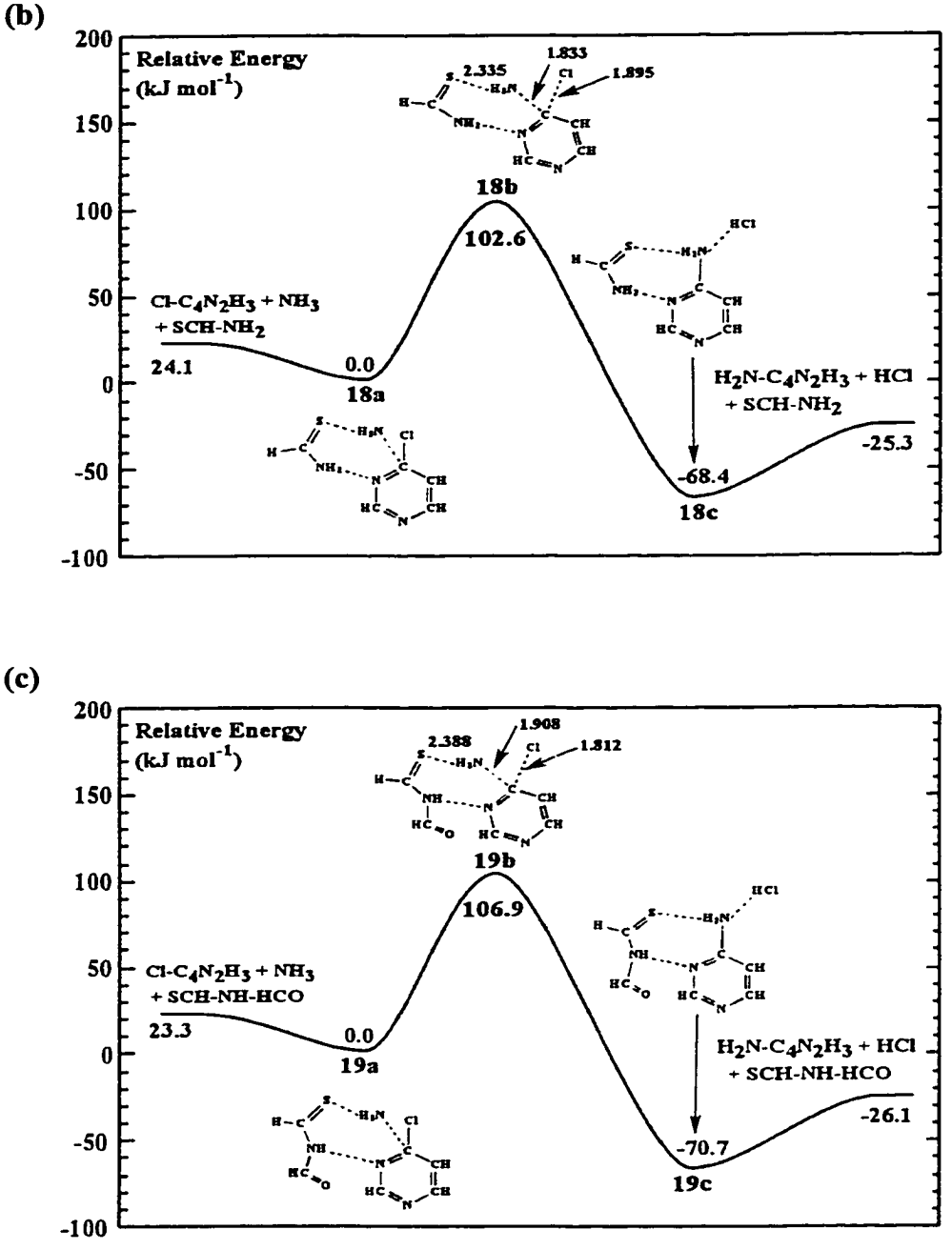


Figure 4.9: Schematic energy profile for the aminolysis of 6-chloropyrimidine, with **(a)** SCH_2 hydrogen bonded to the incoming NH_3 moiety, **(b)** SCH-NH_2 hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring, and **(c)** SCH-NH-CHO hydrogen bonded to both the incoming NH_3 moiety and the pyrimidine ring.

While the PAs of the thioketo bases (Table 4.3) are intermediate between those of the carbonyl- and the imine-derived bases, the barriers are slightly larger than those reported for the aminolysis of 6-chloropyrimidine employing OCH_2, $OCH\text{-}NH_2$ and OCH-NH-HCO as bases. The larger size of the sulfur atom results in an elongation of the $S\cdots HNH_2$ bond to the incoming NH_3, and as such, hydrogen bonding does not enhance the electron-donating ability of NH_3 to the same degree as was observed for the uracil-derived bases. Nonetheless, the correlation between the PA of the terminal thioketo group that hydrogen bonds to NH_3 and the barrier heights is evident.

4.4 Conclusions

In this chapter, the hydrogen bond mediated aminolysis of 6-chloropyrimidine and 2-amino-6-chloropyrimidine was investigated using DFT. While the isolated aminolysis of 6-chloropyrimidine (NH_3 + $Cl\text{-}C_4N_2H_3$) required 138.1 kJ mol^{-1}, the barrier was decreased to 112.2 kJ mol^{-1} by the use of OCH_2 which formed a $O\cdots HNH_2$ hydrogen bond to the incoming NH_3. The barrier was further reduced to 95.2 kJ mol^{-1} by enlarging the base of $OHC\text{-}NH_2$ which formed a hydrogen bond to the incoming NH_3 and a hydrogen bond to the N adjacent to the carbon undergoing subsitution. Although the original study was performed in DMSO, the effect of solvent on the aminolysis reaction is much less than that due to specific hydrogen bonding interactions.

Comparison of the barriers for aminolysis of 6-chloropyrimidine to those calculated for the aminolysis of 2-amino-6-chloropyrimidine reveals that the presence of the $-NH_2$ group enhances the electron density in the pyrimidine ring which in turn diminishes the electrophilicity associated with the carbon at which substitution occurs. While the formation of the third hydrogen bond does not act as a catalyst in the reaction, it does provide a more rigid skeleton upon which the aminolysis reaction may proceed.

Upon closer examination of the aformentioned reactions, it is evident that a correlation exists between the barrier to aminolysis and the $O\cdots HNH_2$ hydrogen bond distance, a consequence of the proton affinity (PA) of the carbonyl group of the base interacting with the incoming NH_3. In order to explore this correlation, the chemical nature of the proton acceptor, i.e. the base, was altered by fluoro-, imine- and sulfur substitution of the uracil bases OCH_2, $OHC\text{-}NH_2$ and OHC-NH-CHO.

Replacement of the H adjacent to the carbonyl group by a fluorine atom, decreases the calculated PA of the proton acceptor relative to that observed for the non-substituted bases and, hence, increases the barrier to aminolysis of 6-chloropyrimidine. Similarly, the imine derivatives of the carbonyl bases have larger PAs than the carbonyl-derived bases and therefore decrease the barrier to aminolysis. While the results for the fluorine and imine-derived bases provide evidence for the important role of the proton acceptor, thioketo substitution was also examined to extend the scope of the study. Although the sulfur-derived bases possess a proton affinity intermediate between the

carbonyl- and fluoro-derived bases, they generate elongated bonds to NH_3 which is reflected in the aminolysis barrier of 6-chloropyrimidine being slightly larger than that observed for the uracil-derived species.

Thus, the aminolysis of 6-chloropyrimidine and 2-amino-6-chloropyrimidine illustrates the ability of the functional groups in uracil to catalyze the reaction by the formation of multiple hydrogen bonds, which stabilize the transition structures. Thus the aminolysis reaction provides a clear example of the catalytic possibilities associated with the formation of multiple hydrogen bonds and illustrates the importance and flexibility associated with a well-chosen hydrogen bond acceptor.

4.5 References

(1) Kolotuchin, S. V.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9092.

(2) Beijer, F. H.; Sijbesma, R. P.; Vekemans, J. A. J. M.; Meijer, E. W.; Kooijman, H.; Spek, A. L. *J. Org. Chem.* **1996**, *61*, 6371.

(3) Kang, J.; Hilmersson, G.; Santamaría, J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 3650.

(4) Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. *Science* **1997**, *278*, 1601.

(5) Jubian, V.; Veronese, A.; Dixon, R. P.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1237.

(6) Jubian, V.; Dixon, R. P.; Hamilton, A. D. *J. Am. Chem. Soc.* **1992**, *114*, 1120.

(7) Kang, J.; Rebek, J., Jr. *Nature* **1997**, *385*, 50.

(8) Wang, B.; Sutherland, I. O. *Chem. Comm.* **1997**, 1495.

(9) Tominaga, M.; Konishi, K.; Aida, T. *J. Am. Chem. Soc.* **1999**, *121*, 7704.

(10) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.

(11) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010.

(12) Melander, C.; Horne, D. A. *J. Org. Chem.*, **1996**, *61*, 8344.

(13) Kelly, T. R.; Bridger, G. J.; Zhao, C. *J. Am. Chem. Soc.* **1990**, *112*, 8024.

(14) Huc, I.; Pieters, R. J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1994**, *116*, 10296.

(15) Wintner, E. A.; Conn, M. M.; Rebek, J., Jr. *Acc. Chem. Res.* **1994**, *27*, 198.

(16) Kirby, A. J. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 707.

(17) Philp, D.; Stoddart, J. F. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1154.

(18) Smith, D. M.; Golding, B. T.; Radom, L. *J. Am. Chem. Soc.* **1999**, *121*, 1383.

(19) Smith, D. M.; Golding, B. T.; Radom, L. *J. Am. Chem. Soc.* **1999**, *121*, 5700.

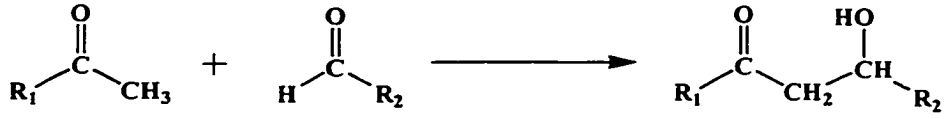
(20) Smith, D. M.; Golding, B. T.; Radom, L. *J. Am. Chem. Soc.* **1999**, *121*, 9388.

Chapter Five.
The Proline-Catalyzed Aldol Reaction

5.1 Introduction

The discovery and development of molecular species that may catalyze chemically important reactions has long been a subject of great research interest. In the previous chapter, derivatives of the nucleobase uracil catalyzed the aminolysis reaction by stabilizing the transition structures through the formation of multiple hydrogen bonds. Nature has created its own form of catalysts, enzymes, which enable a robust number of biologically relevant chemical transformations to occur with high catalytic efficiency and stereochemical control. In order to further the understanding of the fundamentals of molecular recognition and catalysis, considerable effort has been devoted to developing catalytic antibodies[1,2] or biocatalysts that rival natural enzymes in efficiency but catalyze an array of chemical reactions. This has resulted in the development of

novel biocatalysts for organic transformations such as the aldol reaction,[1] an example of which is shown below.



The aldol reaction is one of the fundamental chemical mechanisms for the formation of C—C bonds. As such, one of the current directives in organic chemistry is the development of catalytic variants of the aldol reaction. Previous approaches have generally utilized transition metal-based catalysts in combination with modified substrates.[3,4] Clearly, a more attractive synthetic alternative would involve the development of a catalyst which would enable the reaction between unmodified carbonyls; the *direct* aldol reaction.

In biological systems, the aldol reaction is catalyzed by a powerful group of enzymes referred to collectively as the aldolase enzymes.[5-8] These enzymes are classified depending upon their mode of catalysis. The Class II aldolase enzymes catalyze the aldol reaction using a zinc-cofactor. The Class I aldolase enzymes, however, utilize an enamine mechanism, the chemistry of which is dependent upon a chemically reactive lysine residue in the active site of the enzyme. Using the catalytic-antibody (reactive immunization) technology, antibodies of the Class I aldolase enzymes have been generated[9-15] which are programmed to function by a mechanism analogous to that used by the naturally occurring aldolase enzymes. Although the development of these

biocatalysts has demonstrated the importance of the amine group in catalyzing the aldol reaction, it was only recently that the first amine-based asymmetric Class I aldolase antibody was reported.

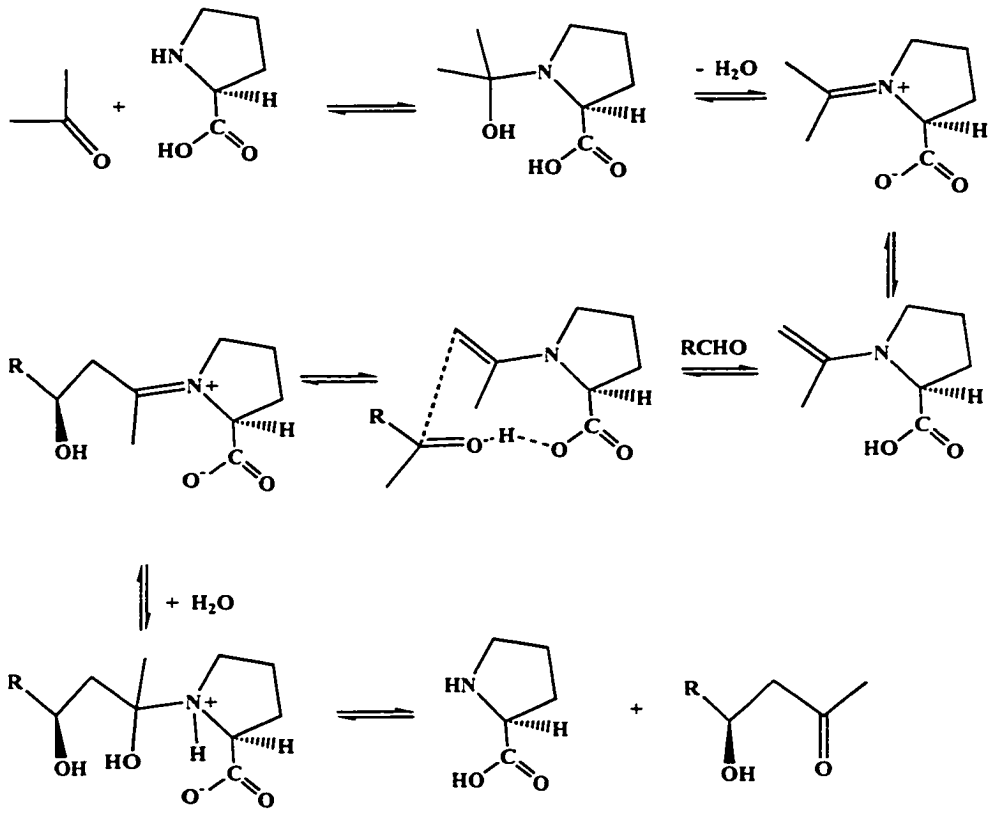


Figure 5.1: Proposed enamine mechanism of the proline-catalyzed aldol reaction.

List et al.[16] reported the catalysis of the aldol reaction between acetone and 4-nitrobenzaldehyde by the simple amino acid L-proline. They proposed a mechanism (Figure 5.1) in which the initial interaction between proline and acetone generates an

enamine intermediate, which then may react further with an aldehyde to yield the aldol product. The direct asymmetric aldol reaction between acetone and a variety of aldehydes was found to proceed with good yields and enantioselectivities, a consequence of the chiral nature of the proline catalyst.

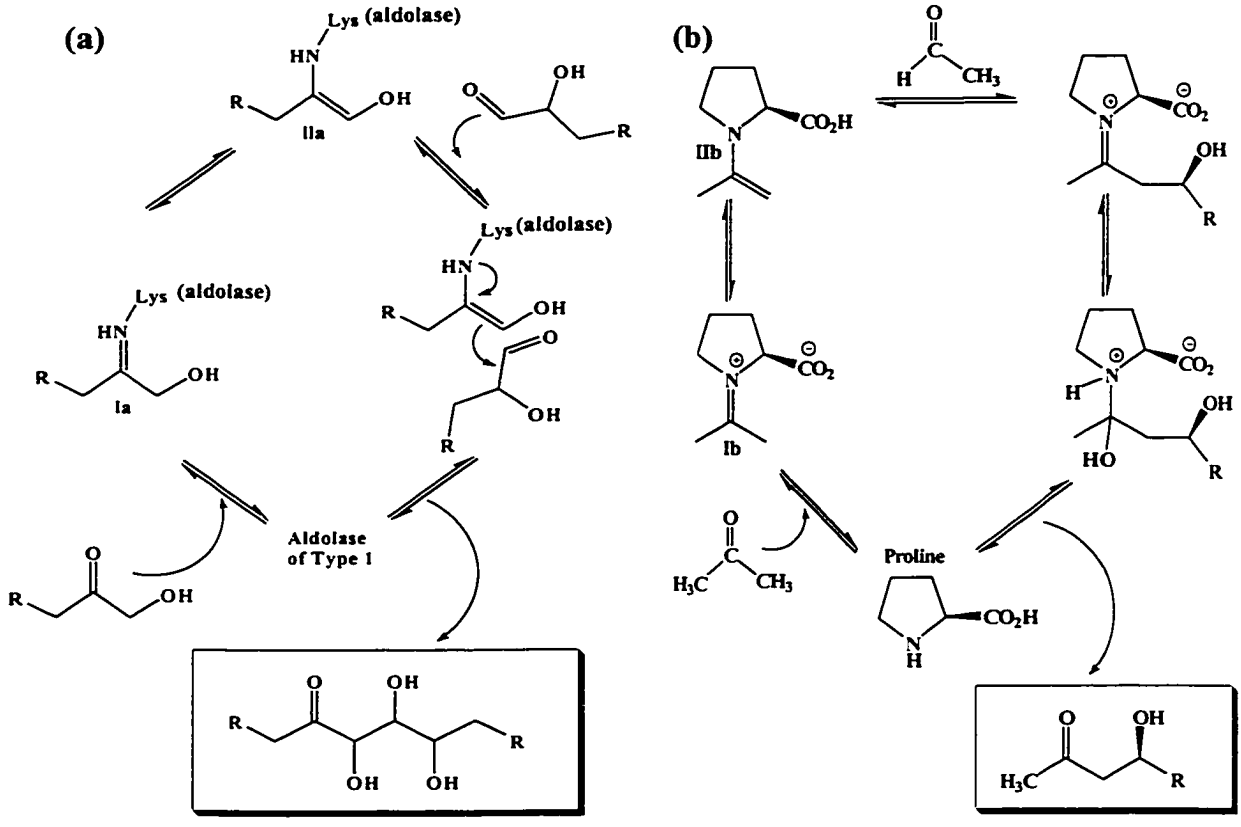


Figure 5.2: Catalytic cycle of the direct aldol reaction with **(a)** aldolases of Class I and **(b)** proline.

The close relation between the enamine mechanism of the Class I aldolases and the postulated mechanism of the proline-catalyzed direct aldol reaction is illustrated via the graphical comparison[17] in Figure 5.2. In both reaction mechanisms, the reaction of the ketone and the amino functionality provided by either the enzyme, in the case of Class I aldolases, or proline, generates the imine complex (**Ia** or **Ib**) that undergoes tautomerism yielding the enamine (**IIa** or **IIb**), a key intermediate in the reaction. The enamine then interacts with the aldehyde leading to the formation of the aldol product and its subsequent release from the original catalyst.

In this chapter, density functional theory calculations are employed to study the reaction between acetone and acetaldehyde, in the presence of proline, (Figure 5.3) in order to examine the feasibility of the proposed mechanism by which proline catalyses the aforementioned direct aldol reaction. In addition, the effect of an ionizing solvent on the mechanism has also been investigated.

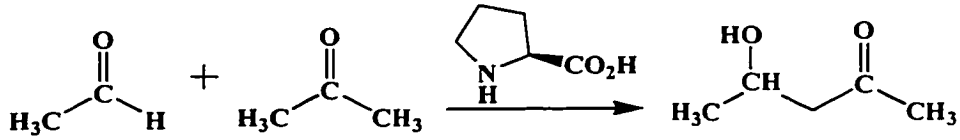


Figure 5.3: Schematic illustration of the aldol reaction catalyzed by proline.

5.2 Computational Details

The minima and transition structures present on the potential energy surfaces of the proline-catalyzed direct aldol reaction between acetone and acetaldehyde were examined using DFT. Geometry optimizations, harmonic vibrational frequencies and zero-point vibrational energy (ZPVE) corrections were obtained with B3LYP/6-31G(d,p). Relative energies, obtained by subsequent single point calculations performed at the B3LYP/6-311+G(2df,p) level, were corrected with the appropriate ZPVE, ie. B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d,p) + ZPVE. As the original study was performed in a DMSO/acetone mixture, the Onsager model was used to incorporate solvent effects using $\varepsilon = 46.7$, the dielectric constant for DMSO. For computational details, see Chapter Two. All bond lengths are in angstroms (Å) and energies in kJ mol^{-1}. The total energies of all molecules on the non-solvated and solvated surfaces (Tables C1.1 and C1.2, respectively) are summarized in Appendix C.

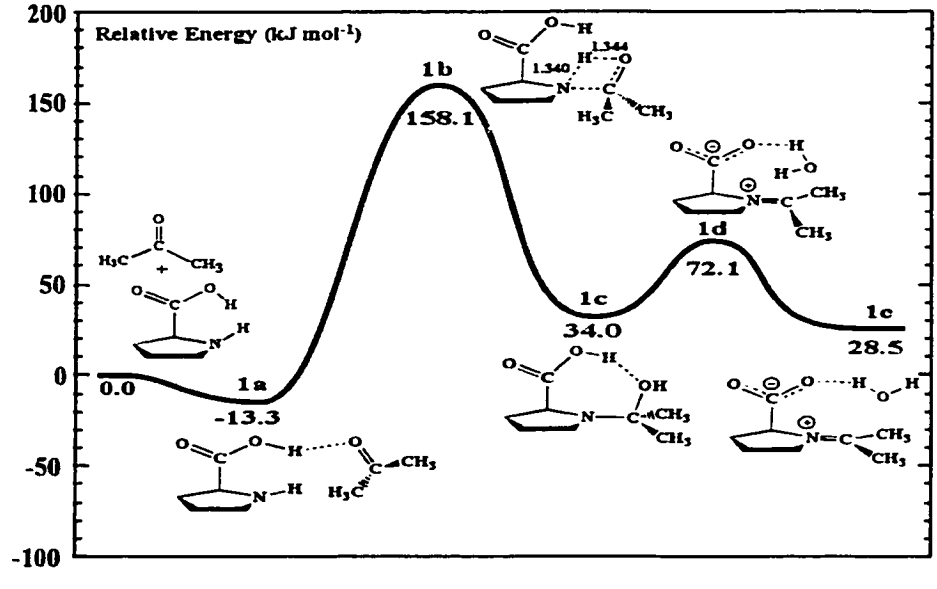
5.3 Results and Discussion

5.3.1 Non-Solvated Surface

Initially, acetone interacts with proline (Figure 5.4a) via an O···HO bond of 1.763 Å, forming complex **1a**, lying lower in energy by 13.3 kJ mol^{-1}. A formal C—N bond between acetone and proline forms as the proton on the N of proline is transferred

simultaneously to the carbonyl oxygen of acetone in transition structure (TS) **1b,** with a barrier of 158.1 kJ mol^{-1}. Substantial energy (124.1 kJ mol^{-1}) is released as the proton is completely transferred, producing complex **1c**. The formation of a water molecule, which interacts with the anionic carboxyl group via a short and strong $O^{-}\cdots HO$ hydrogen bond of 1.314 Å, results in a zwitterionic imine TS **1d** being generated which lies 72.1 kJ mol^{-1} higher in energy than the reactants. As the hydrogen bond between the water molecule and the carboxyl group elongates to 1.702 Å, complex **1e** is formed lying 43.6 kJ mol^{-1} lower in energy than TS **1d**.

(a)



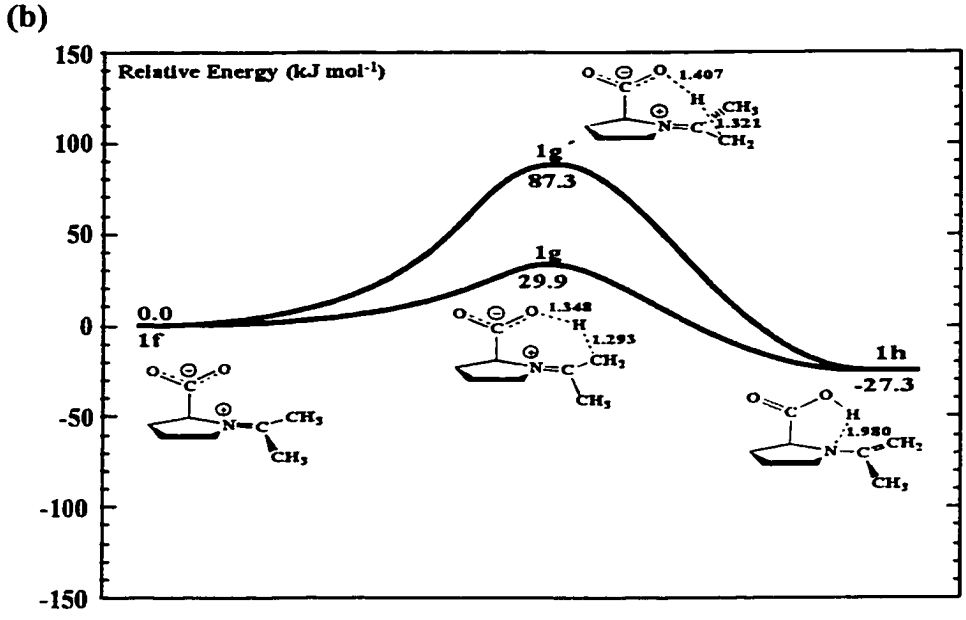


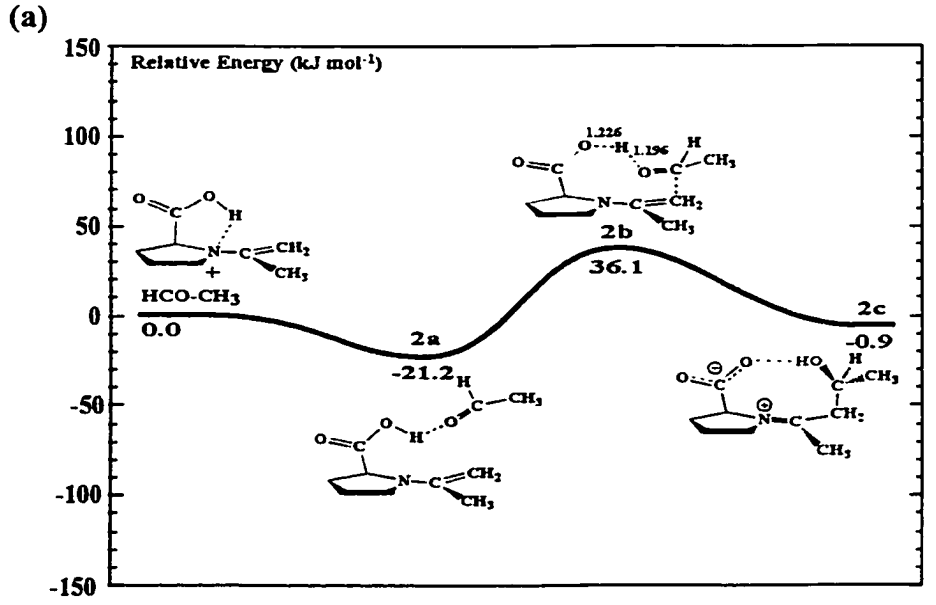
Figure 5.4: Schematic energy profiles of the **(a)** reaction of acetone with proline yielding the imine complex, and **(b)** two possible pathways of imine-enamine tautomerism.

Complete removal of the water molecule yields the isolated imine complex **1f**. As illustrated in Figure 5.4b, the imine complex converts to the enamine analogue by a proton transfer from the methyl group trans to the α-carbon of proline to the carboxyl group. The imine-enamine conversion passes through TS **1g**, with a barrier of 29.9 kJ mol^{-1}, before complete proton transfer yields the enamine complex **1h**, which is 27.3 kJ mol^{-1} more stable than its imine analogue.

Although this is the structurally and energetically preferred pathway for the imine-enamine tautomerism, an alternate pathway is available. The imine complex **1f** may tautomerize to its enamine structure **1h** via TS **1g'** with a barrier of 87.3 kJ mol^{-1}.

This higher-energy pathway involves proton transfer from the methyl group cis to the α-carbon of proline to the carboxyl group.

The addition of acetaldehyde to the enamine is summarized in Figure 5.5a. Initially, acetaldehyde interacts with the enamine moiety via an O···HO hydrogen bond (2.012 Å) to form the hydrogen-bonded complex **2a** lying 21.2 kJ mol^{-1} lower in energy. As the carbonyl carbon of acetaldehyde forms a C—C bond to the terminal carbon of the alkene, the proton of the carboxylic acid of proline is transferred to the carbonyl oxygen of acetaldehyde via TS **2b** with a barrier of 36.1 kJ mol^{-1}. Complete proton transfer to the attached acetaldehyde yields the zwitterionic imine complex **2c**, lying 0.9 kJ mol^{-1} lower in energy than the non-reacted enamine and acetaldehyde.



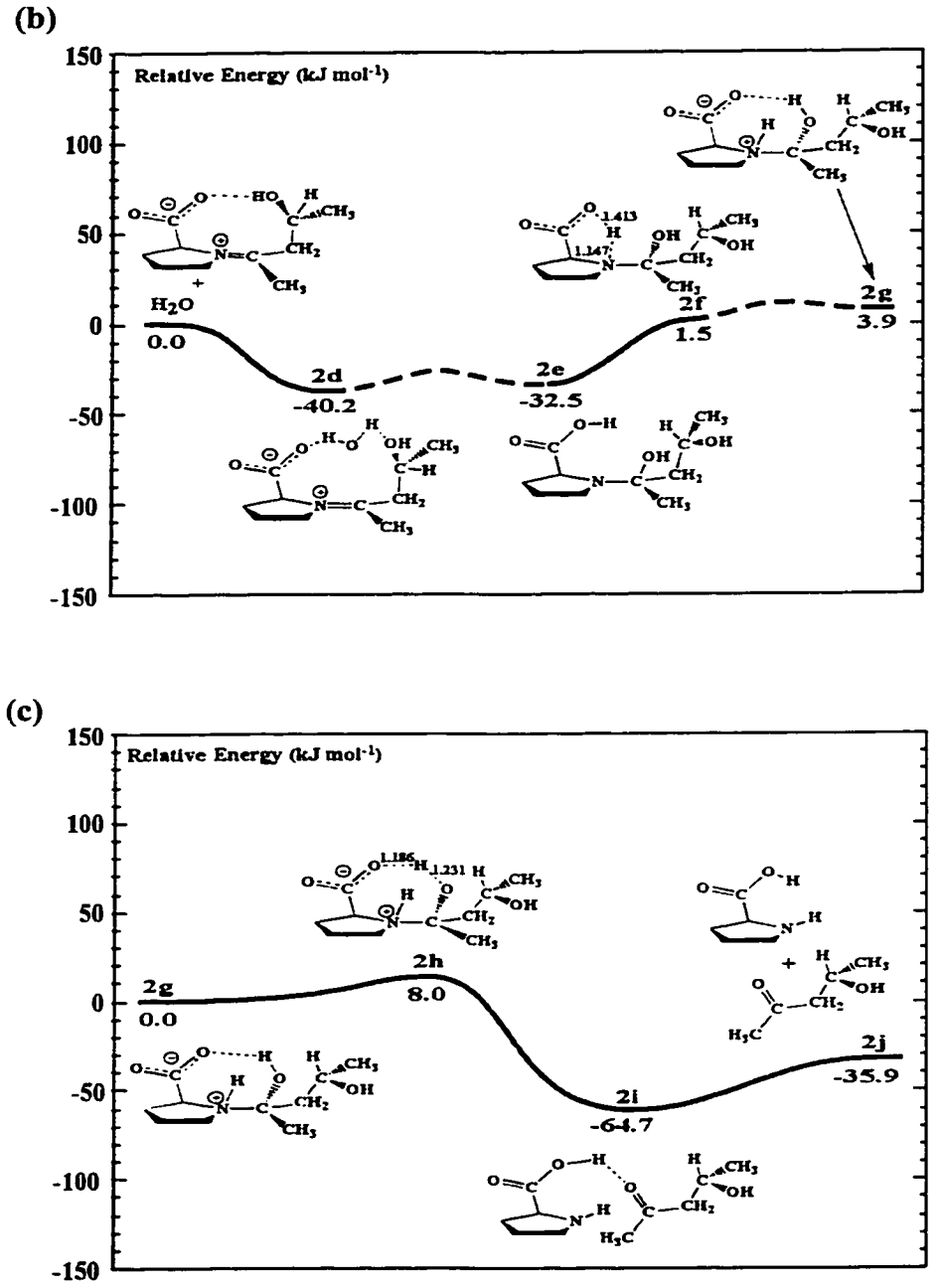


Figure 5.5: Schematic energy profiles of **(a)** the addition of acetaldehyde to the enamine complex, **(b)** the addition of water across the C=N bond of the enamine complex, and **(c)** the formation and release of the aldol product.

The introduction of a water molecule and its addition across the C=N bond of the substituted imine complex is illustrated in Figure 5.5b. The addition of H_2O to the substituted imine proline, complex **2c**, yields the doubly hydrogen-bonded complex **2d** lying 40.2 kJ mol^{-1} lower in energy. As the water molecule breaks apart and adds to the carbon of the C=N bond in the imine, complex **2e** is generated lying 7.7 kJ mol^{-1} higher in energy (32.5 kJ mol^{-1}). A transition structure connecting complexes **2d** and **2e** was not found. However, it is anticipated that should such a structure exist, it would possess a very small barrier. Proton transfer from the carboxylic acid moiety in complex **2e** to the N in proline proceeds via TS **2f**, with a barrier of 31.0 kJ mol^{-1}. Complete proton transfer yields the zwitterionic complex **2g**, lying 2.4 kJ mol^{-1} higher in energy than TS **2f**.

The final stage of the reaction, summarized in Figure 5.5c, involves the formation of the aldol product and its subsequent release from proline. As the hydroxyl proton in complex **2g** is transferred to the carboxylic anion of proline, the formal C—N bond between the aldol product and proline elongates, forming the hydrogen-bonded complex **2i** lying 64.7 kJ mol^{-1} lower in energy. The conversion from complex **2g** to **2i** proceeds via TS **2h** with a barrier of 8.0 kJ mol^{-1}. As the O···HO hydrogen bond of 1.796 Å between the aldol product and proline in complex **2i** breaks, the products are formed lying 28.8 kJ mol^{-1} higher in energy than the hydrogen bonded complex **2i**. The easy separation of the aldol product from proline is a feature desirable in an organic catalyst.

5.3.2 DMSO-Solvated Surface

Previously,[11] it has been suggested that enamine and C—C bond formation are rate-limiting with catalysts that utilize an analogous mechanism to that proposed for Class I aldolases. However, the theoretical calculations on the proposed mechanism of Figure 5.4 illustrate that the formation of the enamine (**1h**) from the imine (**1f**) requires a mere 29.9 kJ mol^{-1} of energy. Similarly, the formation of the C—C bond in Figure 5.5a, which involves the addition of acetaldehyde to the enamine (complex **2a** to **2c**), requires only 57.2 kJ mol^{-1} of energy. Thus, neither of these two steps are rate-limiting in the catalysis. However, as seen in Figure 5.4a, a significant barrier of 158.1 kJ mol^{-1} exists for the addition of acetone to proline, a key step in the proposed mechanism.

As the original study was performed in DMSO, a moderately ionizing solvent, it is possible that solvent may play a key role stabilizing zwitterionic structures and lowering key barriers in the reaction. The initial interaction between acetone and proline, in DMSO, is summarized in Figure 5.6a. Initially, acetone interacts with proline via an O···HO hydrogen bond (1.713 Å) forming complex **3a** lying 14.1 kJ mol^{-1} lower in energy than the reactants. In the presence of DMSO, the proton of the carboxylic acid moiety of proline, rather than the proton from the N of proline as observed in Figure 5.3a, is transferred to the carbonyl oxygen of acetone. This alternate mechanism, which also involves the formation of a formal C—N bond between the two moieties in TS **3b**, lying

26.6 kJ mol^{-1} higher in energy, has a barrier 130.7 kJ mol^{-1} smaller than that observed in the non-solvated surface.

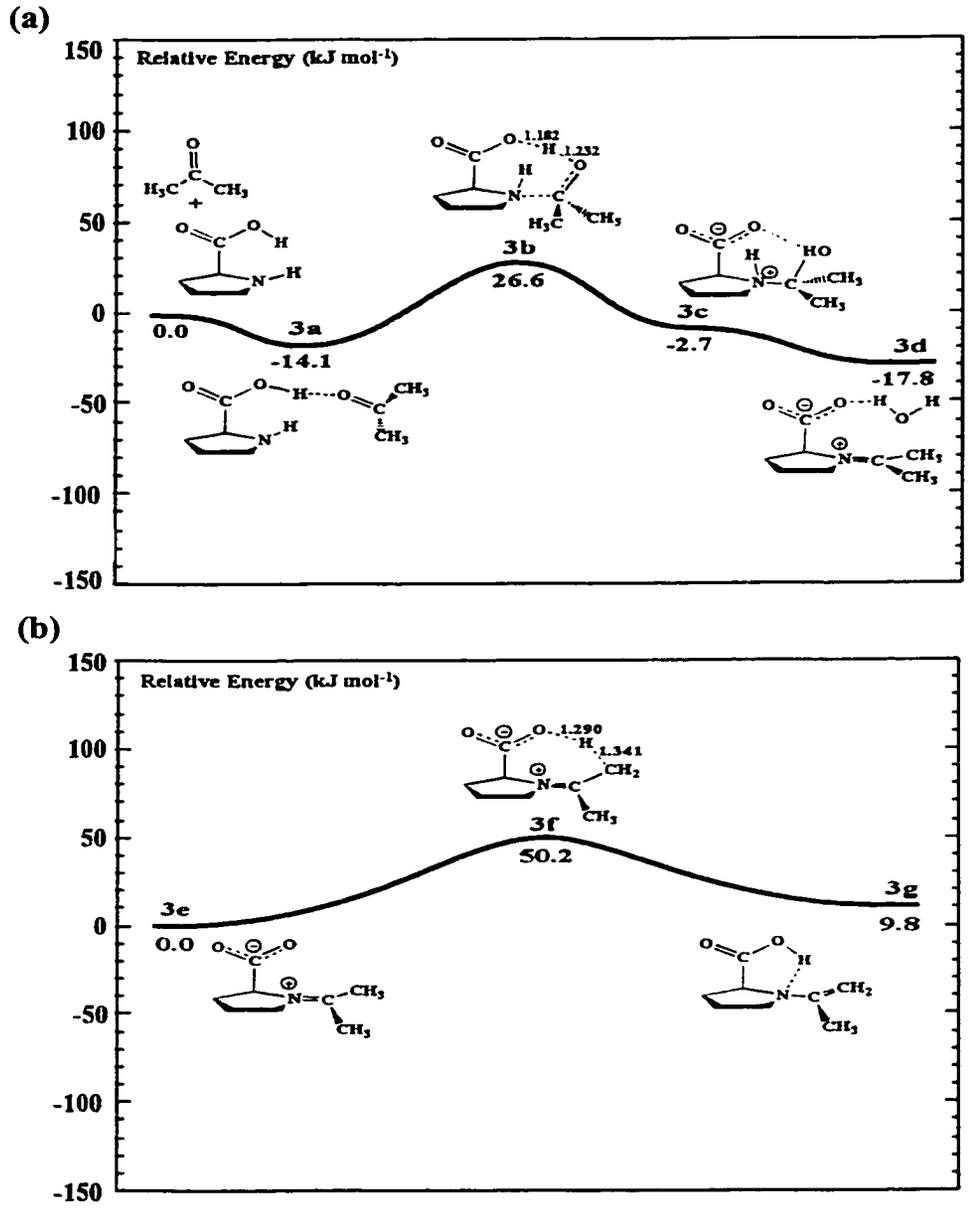
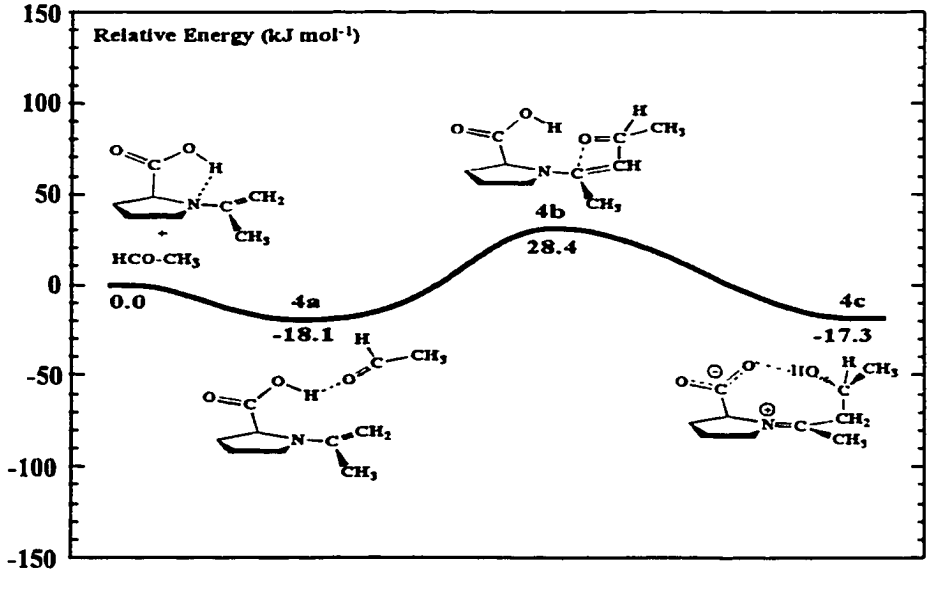


Figure 5.6: Schematic energy profile of the effect of DMSO on the **(a)** reaction of acetone with proline yielding the imine complex, and **(b)** imine-enamine tautomerism.

Complete proton transfer yields the zwitterionic structure **3c**, lying 2.7 kJ mol^{-1} lower in energy than the reactants. Removal of a water molecule yields the imine complex **3d** lying 17.8 kJ mol^{-1} lower in energy than the isolated proline and acetone. Conversion of the isolated imine molecule (**3e**) to the enamine molecule (**3g**), Figure 5.6b, proceeds via TS **3f** with a barrier of 50.2 kJ mol^{-1}, 20.3 kJ mol^{-1} higher in energy than that observed in the non-solvated surface.

The addition of acetaldehyde to the enamine, Figure 5.7a, generates complex **4a** lying 18.1 kJ mol^{-1} lower in energy than the isolated enamine molecule. The formation of a formal C—C bond between the enamine and acetaldehyde in the zwitterionic complex **4c** proceeds via TS **4b**, with a barrier of 46.5 kJ mol^{-1}. Relative to the non-solvated addition, this is a decrease of 10.8 kJ mol^{-1} in the barrier.

(a)



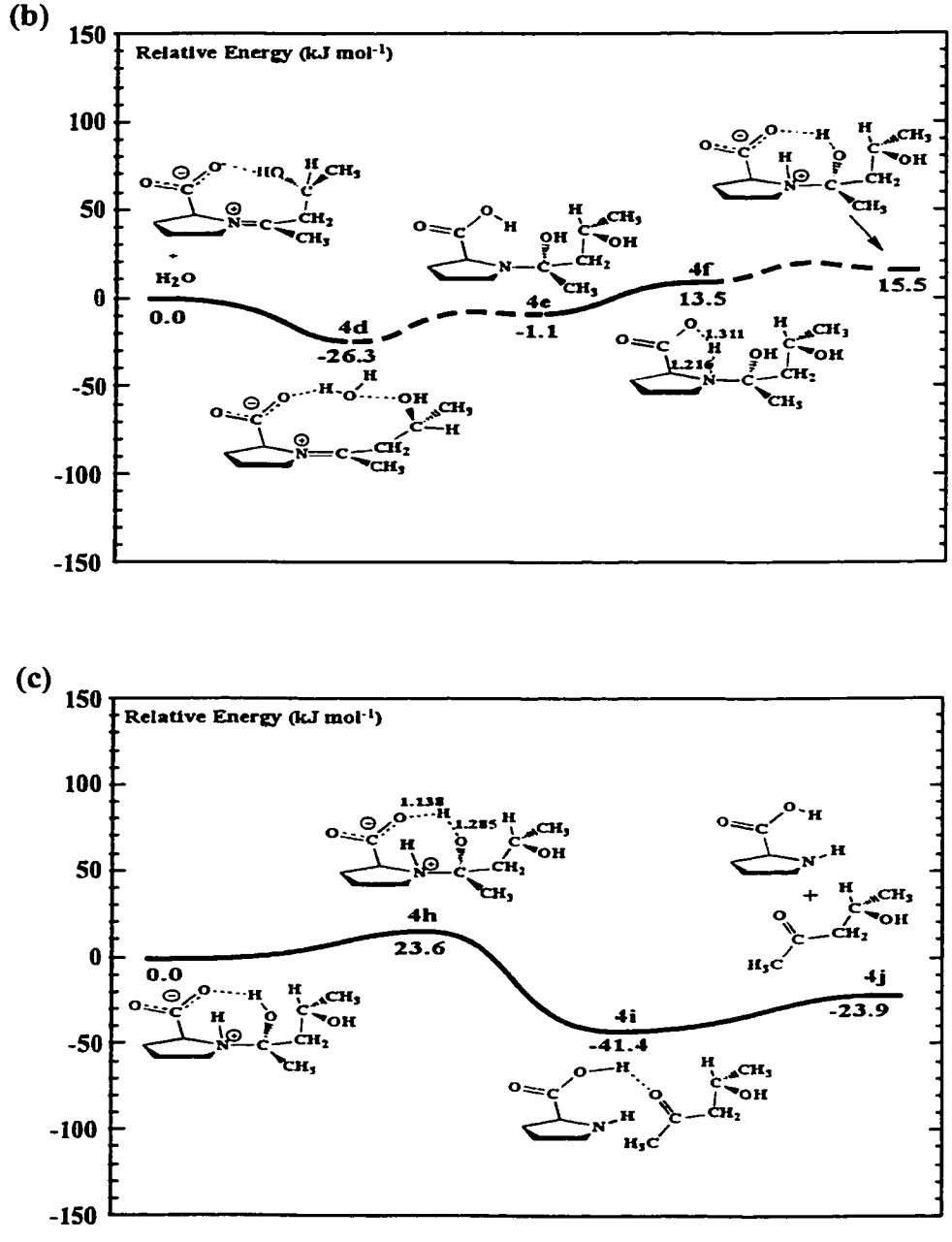


Figure 5.7: Schematic energy profiles of the effect of DMSO on **(a)** the addition of acetaldehyde to the enamine, **(b)** the addition of water across the C=N bond of the enamine complex, and **(c)** the formation of the aldol product and release of proline.

The final two segments of the mechanism, involving the addition of H_2O across the C=N bond of the imine and subsequent formation of the aldol product, are summarized in Figure 5.7b and 5.7c, respectively. The addition of H_2O to the imine complex **4c** generates the doubly hydrogen-bonded complex **4d** lying 46.3 kJ mol^{-1} lower in energy than the isolated imine. As the water molecule breaks apart, the hydroxyl anion adds to the carbon of the imine as the proton simultaneously adds to the carbonyl anion group yielding complex **4e**, lying 25.2 kJ mol^{-1} higher in energy. As with the analogous segment of the mechanism in the absence of solvent, a TS connecting these two minima was not located. As the proton of the carbonyl group is transferred to the proline nitrogen, complex **4g** is generated lying 15.5 kJ mol^{-1} higher in energy. This proton rearrangement proceeds via TS **4f** with a barrier of 12.4 kJ mol^{-1}. The carbonyl anion abstracts the hydroxyl proton from the carbon attached to the prolyl nitrogen in TS **4h**, lying 23.6 kJ mol^{-1} higher in energy. Complete proton transfer and the cleavage of the C—N bond yields the hydrogen bonded complex **4i** lying 41.4 kJ mol^{-1} lower in energy. As the O···HO hydrogen bond between the carboxylic group of proline and the carbonyl group of the aldol product elongates, the aldol product is released from proline.

These portions of the mechanism are thermodynamically endergonic and not greatly affected by inclusion of solvent. On the non-solvated surface, the addition of H_2O across the C=N bond of the imine and subsequent proton rearrangement, complex **2d** to complex **2g** (Figure 5.5b), requires 44.1 kJ mol^{-1} of energy. The analogous conversion (complex **4d** to complex **4g**) on the DMSO surface (Figure 5.7b) requires 41.8 kJ mol^{-1} of

energy. Both the solvated and non-solvated surfaces conclude with the release of the aldol product and the regeneration of the unsubstituted proline.

5.4 Conclusions

The direct aldol reaction between acetone and acetaldehyde, in which the amino acid proline functions as the catalyst during the enamine mechanism, was investigated using density functional theory. Previous studies of antibodies utilizing the analogous mechanism indicate that enamine formation and or C—C bond breaking/forming is rate limiting. The aforementioned processes require 29.9 and 57.2 kJ mol^{-1} of energy, respectively; not enough to inhibit the reaction. However, the calculations indicate that the initial complexation between proline and acetone requires substantial energy (171.4 kJ mol^{-1}) and would inhibit further progression of the reaction.

As the enamine mechanism involves the formation of charged species, the effect of solvent was examined using the Onsager model. In the presence of DMSO, the barrier for the initial complexation between proline and acetone is reduced to a 40.7 kJ mol^{-1} while the enamine and C—C bond formation steps require 50.2 and 46.5 kJ mol^{-1}, respectively. Thus, solvent plays a critical role in the direct aldol reaction by stabilizing ionic charges and providing an alternate, lower energy, pathway by which the reaction may proceed.

Proline efficiently catalyzes the direct aldol reaction using the enamine mechanism characteristic of nature's Class I aldolase enzymes. The aldol reaction between acetone and acetaldehyde, in which proline acts as an enzyme mimic, is a simple example illustrating the inherent potential of small organic molecules to act as chiral catalysts in asymmetric synthesis.

5.5 References

(1) Reviews: (a) Hasserodt, J. *Synlett* **1999**, *12*, 2007. (b) Stevenson, J. D.; Thomas, N. R. *Nat. Prod. Rep.* **2000**, *17*, 535. (c) Thomas, N. R. *Nat. Prod. Rep.* **1996**, *13*, 479. (d) Koeller, K. M.; Wong, C. –H. *Nature* **2001**, *409*, 232.

(2) Walsh, C. *Nature* **2001**, *409*, 226.

(3) Reviews: (a) Nelson, S. G. *Tetrahedron: Asymmetry* **1998**, *9*, 357. (b) Gröger, H.; Vogl, E. M.; Shibasaki, M. *Chem. Eur. J.* **1998**, *4*, 1137. (c) Bach, T. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 417. (d) Machajewski, T. D.; Wong, C. –H. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 1352.

(4) Yoshikawa, N.; Yamada, Y. M. A.; Das, J.; Sasai, H; Shibasaki, M. *J. Am. Chem. Soc.* **1999**, *121*, 4168 and references therein.

(5) Procter, G. *Asymmetric Synthesis*, Oxford University Press Inc.: New York, 1996.

(6) Wong, C. -H.; Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*, Elsevier Science Inc.: Oxford, 1994.

(7) Seoane, G. *Curr. Org. Chem.* **2000**, *4*, 283.

(8) Takayama, S.; McGarvey, G. J.; Wong, C. -H. *Chem. Soc. Rev.* **1997**, *26*, 407.

(9) Wagner, J.; Lerner, R. A.; Barbas III, C. F. *Science* **1995**, *270*, 1797.

(10) Barbas III, C. F.; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Bjornestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A. *Science* **1997**, *278*, 2085.

(11) Hoffmann, T.; Zhong, G.; List, B.; Shabat, D.; Anderson, J.; Gramatikova, S.; Lerner, R. A.; Barbas III, C. F. *J. Am. Chem. Soc.* **1998**, *120*, 2768.

(12) List, B.; Shabat, D.; Barbas III, C. F.; Lerner, R. A. *Chem. Eur. J.* **1998**, *4*, 881.

(13) Zhong, G.; Shabat, D.; List, B.; Anderson, J.; Sinha, S. C.; Lerner, R. A.; Barbas III, C. F. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 2481.

(14) Zhong, G.; Lerner, R. A.; Barbas III, C. F. *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 3738.

(15) Sinha, S. C.; Sun, J.; Miller, G.; Barbas III, C. F.; Lerner, R. A. *Org. Lett.* **1999**, *1*, 1623.

(16) List, B.; Lerner, R. A.; Barbas III, C. F. *J. Am. Chem. Soc.* **2000**, *122*, 2395.

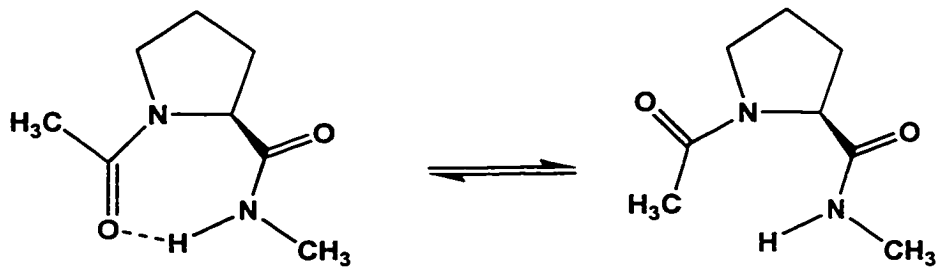
(17) Gröger, H.; Wilken, J. *Angew. Chem. Int. Ed.* **2001**, *40*, 529.

Chapter Six.
Intramolecular Catalysis of the Isomerization of Prolyl Amides

6.1 Introduction

As the basic building block in a variety of biologically important polymers, the amide group is one of the most significant functional groups in biology and chemistry. Of the twenty amino acids in naturally occurring polypeptides that are subject to nondegenerate *cis* and *trans* rotamers, proline is unique in that it has the potential to establish an imidic peptide bond through its prolyl nitrogen. Although the *cis* and *trans* rotamers of amides N-terminal to proline are energetically similar[1] (Figure 6.1), the *cis-trans* rotamer producing rotation about the C—N bond involving proline residues is the slow step in the folding of peptides[2-4] and has significant implications for the secondary

structure of proteins. Thus, the catalysis of the amide bond isomerization is an area of scientific interest with important biological and chemical connotations.



***trans* rotamer** ***cis* rotamer**

Figure 6.1: Amide isomerization in N-acetylproline methylamide.

The *cis-trans* isomerization of the amide bonds in peptides and proteins of biological systems is catalyzed by a class of enzymes referred to as peptidylprolyl isomerases[5] (PPIases). This group of enzymes represents the only type of biocatalyst known to date whose sole function is conformational interconversion. Although the exact mechanism by which the PPIases catalyze the *cis-trans* interconversion is yet to be elucidated, it has been proposed[6-8] that intramolecular catalysis of the amide isomerization by the formation of a weak hydrogen bond between the prolyl nitrogen, acting as a hydrogen bond acceptor, and the NH unit of the proline residue is an important part of the mechanism. In fact, intramolecular hydrogen bonding between the

prolyl nitrogen and nearby hydrogen donors is a common feature in structural protein chemistry.[8]

Experimental[8-12] and theoretical[6,13] studies often employ N-acetylproline methylamide as a model for the study of the *cis-trans* isomerization of proline-containing residues. These studies have shown that the rate of isomerization may be altered through the addition of substituents on various positions of the proline ring. Specifically, 5-tert-butyl groups on C^{δ} and methyl groups on C^{β} impose steric repulsions that impede the isomerization. Conversely, fluoro and hydroxy groups on C^{γ} have been shown to accelerate the rate of isomerization. However, the presence of the intramolecular hydrogen bond and its implications for the *cis-trans* isomerization of substituted prolines was not investigated in these studies.

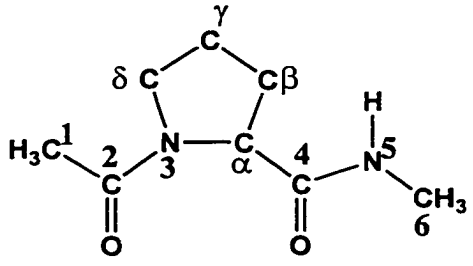


Figure 6.2: The numbering of the atoms of N-acetylproline methylamide.

In this chapter, the conformational preferences in transition structures during isomerization and the role of intramolecular hydrogen bonding in the model system of N-acetylproline methylamide (Figure 6.2) is investigated by DFT. In addition, the effect of electron-withdrawing fluoro and hydroxy subsitutents at C^{γ} in N-acetylproline methylamide on the barrier to isomerization and the formation of an intramolecular hydrogen bond are also investigated. Solvent effects are not considered in this study, as previous theoretical studies[13,14] have shown that amide bond rotation is retarded as the polarity of the solvent is increased. This has been attributed to an increase in the hydrogen-bond donating ability of the solvent, which stabilizes the ground state to a greater degree relative to the transition state.

6.2 Computational Details

The *cis-trans* isomerization of N-acetylproline methylamide was investigated using DFT. Geometry optimizations, harmonic vibrational frequencies and zero-point vibrational energy (ZPVE) corrections were obtained with B3LYP/6-31G(d,p) (Chapter Two). Relative energies, obtained by subsequent single point calculations performed at the B3LYP/6-311+G(2df,p) level, were corrected with the appropriate ZPVE, i.e. B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d,p) + ZPVE. All bond lengths are in angstroms (Å) and energies in kJ mol^{-1}. The total energies of all species in this study and the

charges on the atoms in the various structures (from Mulliken population analyses) are summarized in Table D1.1 and Figure D1.1 of Appendix D.

6.3 Results and Discussion

The *cis* rotamer of N-acetylproline methylamide is expected to contain a long intramolecular $N_3\cdots HN_5$ hydrogen bond between the NH group of the methylamide side chain and the prolyl nitrogen. In the *trans* rotamer, however, the carbonyl group in the acetyl unit is in closer proximity to the amide group and thus, the formation of the $N_5H\cdots O$ bond may be structurally favored over the $N_3\cdots HN_5$ bond.

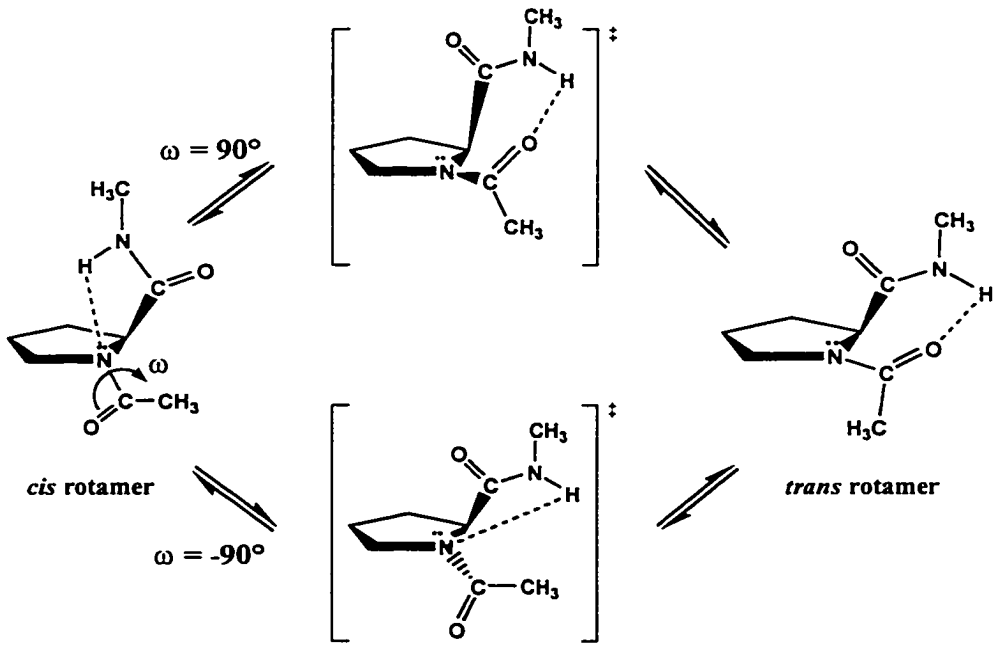


Figure 6.3: Possible reaction pathways for the *cis-trans* isomerization of proline-containing molecules.

In order for N-acetylproline methylamide to undergo conversion between the *cis* and *trans* structures, there exist two mutually exclusive, structurally accessible, pathways as illustrated in Figure 6.3. These two pathways differ in the direction by which the acetyl ($COCH_3$) group rotates, as described by the angle ω.

6.3.1 The Isomerization of N-acetylproline Methylamide

The *cis* rotamer of N-acetylproline methylamide contains an $N_3\cdots HN_5$ intramolecular hydrogen bond (2.277 Å) and is energetically less stable than the *trans* rotamer by 13.0 kJ mol^{-1}. In order to undergo isomerization from the *cis* to the *trans* rotamer, the acetyl ($COCH_3$) group may rotate in a clockwise or counterclockwise direction. Counterclockwise rotation (ω = -90°) results in the formation of TS **1a** lying 70.4 kJ mol^{-1} higher in energy than the *cis* rotamer. Clockwise (ω = 90°) rotation of the acetyl group generates transition structure (TS) **1b,** lying 70.5 kJ mol^{-1} higher in energy than the *cis* structure. In both the aforementioned TSs, rotation of the acetyl group enables the intramolecular $N_3\cdots HN_5$ hydrogen bond to shorten to 2.177 Å and 2.178 Å in TSs **1a** and **1b**, respectively (Figure 6.4).

Although the two reaction pathways are energetically similar, the transition structures are significantly different. In the *cis* and *trans* structures of N-acetylproline methylamide, the proline ring is essentially planar. Rotation of the acetyl group results in a deformation of the planarity of the proline ring by the puckering of C^{γ} either away from

(*exo*) or towards (*endo*) the C^{α} of proline in the transition structures. In addition, rotation of the acetyl moiety results in the geometry of the prolyl nitrogen (N3) going from a nearly planar sp^2 to essentially tetrahedral sp^3 hybridization. This deformation of the N3 geometry is typically associated with C2, the carbon of the acetyl moiety, moving away from (*exo*) or closer to (*endo*) C4, the carbon in the methylamide unit and is described by a virtual dihedral angle η. This parameter, which describes the imide nitrogen pyramidalization, is defined by the atoms (C^{α}-N3-C^{δ}-C2) and takes on values of $\eta = \pm 180°$ for a planar nitrogen and $\eta = \pm 120°$ for a tetrahedral nitrogen.

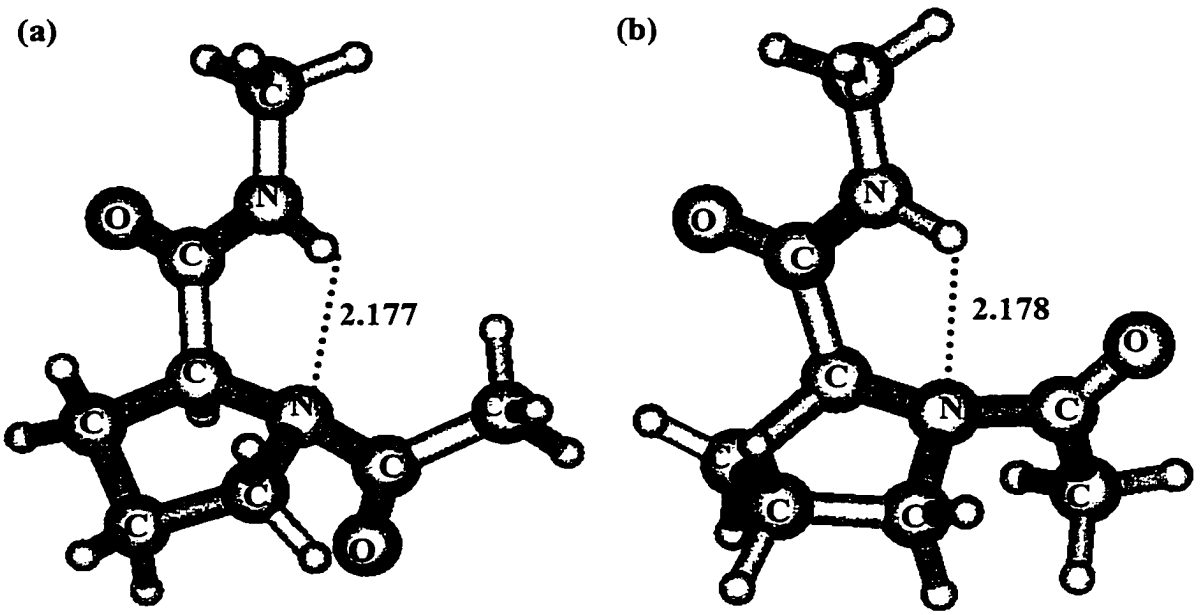


Figure 6.4: Transition structures obtained in the *cis-trans* isomerization of N-acetylproline methylamide for **(a)** counterclockwise (ω = - 90°) rotation of the acetyl group (TS **1a**) and, **(b)** clockwise (ω = 90°) rotation of the acetyl group (TS **1b**).

For the unsubstituted isomerization, counterclockwise rotation results in the C^{γ} distorting to adopt an *exo* conformation in TS **1a** while with clockwise rotation, C^{γ} distorts to adopt an *endo* conformation in TS **1b**. Relative to that observed in the *cis* rotamer (η = 161.7 °), rotation of the acetyl group results in N3 perturbing to adopt a more pyramidalized structure in TS **1a** (η = 123.1°) and in TS **1b** (η = 138.9°). In both TSs, the C—N bond to the acetyl group is elongated from 1.379 Å in the *cis* rotamer to 1.457 Å and 1.441 Å in TSs **1a** and **1b**, respectively. As C—N(sp^2) bond lengths (1.36-1.39Å) are considerably shorter than C—N(sp^3) bond lengths (1.45-1.47 Å), this, in addition to the increased negative charge of the prolyl nitrogen (see Figure D1.1 of Appendix D), is consistent with the change in the geometry of the prolyl nitrogen from sp^2 to an sp^3 hybridization.

Replacement of the H at the C^{γ} by various electron-withdrawing groups can be situated either *cis* or *trans* to the methylamide moiety on C^{α}. The orientation of these groups will have markedly different effects not only on the barrier to rotation but also on the puckering of the proline ring and the $N_3 \cdots HN_5$ intramolecular hydrogen bond. Thus, the effect of replacing the hydrogen on C^{γ} with fluoro and hydroxy groups on the barrier to *cis-trans* isomerization and the formation of the intramolecular N···HN hydrogen bond will be examined.

6.3.2 The Effect of Substituents on C^{γ} cis to the Methylamide Moiety

The barriers to isomerization for the N-acetylproline methylamide with electron-withdrawing fluoro and hydroxy substituents on the C^{γ} of the proline ring *cis* to the methylamide unit are summarized in Table 6.1. In both cases, clockwise rotation of the acetyl group (ω = 90°) produces transition states that are of higher energy than those obtained by counterclockwise rotation (ω = - 90°) of the acetyl group. In the TSs obtained by counterclockwise rotation, an $N_3\cdots HN_5$ hydrogen bond is always generated. As the barrier to rotation for the unsubstituted N-acetylproline methylamide is larger than that obtained upon replacement of the hydrogen at C^{γ} by electron-withdrawing fluoro and hydroxy substituents, this implies an accelerated rate of prolyl peptide isomerization accompanies the formation of the intramolecular $N_3\cdots HN_5$ hydrogen bond.

Table 6.1: Summary of the barriers to *cis-trans* isomerization (kJ mol^{-1}) for N-acetylproline methylamide with substituents on C^{γ} of the proline ring *cis* to the methylamide moiety.

Substituent	*ω = 90°*	*ω = - 90°*
unsubstituted	70.5	70.4
fluorine	83.1	66.7
hydroxy	67.3	64.9

As the acetyl moiety in the *cis* rotamer of N-acetyl-4-fluoroproline methylamide rotates by $\omega = -90°$, the C^{γ} of the proline ring puckers to adopt an *endo* conformation in TS **2a** (Figure 6.5), lying 66.7 kJ mol^{-1} higher in energy. This is accompanied by a shortening of the $N_3\cdots HN_5$ hydrogen bond from 2.311 Å in the *cis* structure to 2.166 Å in the TS. In TS **2a**, C2 is in an *exo* conformation and the prolyl nitrogen has distorted to enhance its sp^3 hybridization ($\eta = 125.4°$) relative to that observed in the *cis* rotamer ($\eta = -163.2°$). Conversely, clockwise rotation ($\omega = 90°$) of the acetyl group results in the formation of an $N_5H\cdots O$ hydrogen bond of 1.940 Å in TS **2b** (Figure 6.5) lying 83.1 kJ mol^{-1} higher in energy.

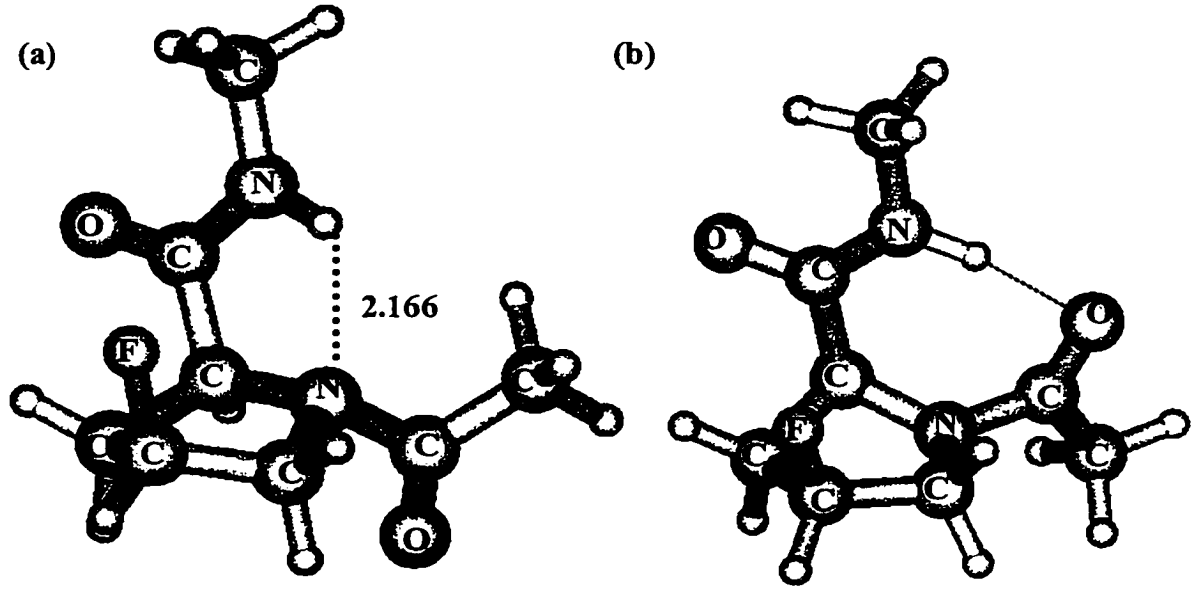


Figure 6.5: Transition structures obtained in the *cis-trans* isomerization of N-acetyl-4(*cis*)-fluoroproline methylamide by **(a)** counterclockwise rotation ($\omega = -90°$) of the acetyl group (TS **2a**) and **(b)** clockwise rotation ($\omega = 90°$) of the acetyl group (TS **2b**).

In order to facilitate the formation of the aforementioned hydrogen bond, C2 adopts an *endo* conformation thus distorting the planarity of the proline ring. This is accompanied by N3 adopting a more sp^3 hybridized conformation (η = -151.7°) by puckering below the plane of the ring. The manner in which N3 perturbs is different in the two TSs, as is evident by the different signs of the virtual angle η. The distortion of N3 in TS **2b** below the proline ring is a consequence of the shorter and stronger O···HN_5 hydrogen bond present in TS **2b**.

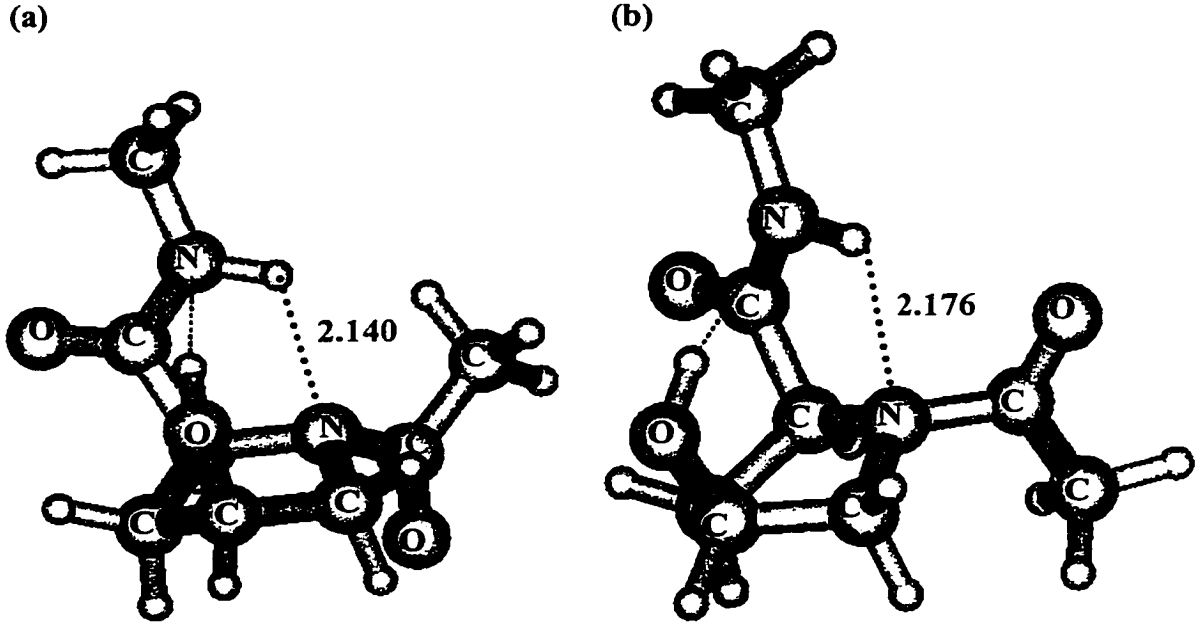


Figure 6.6: Transition structures obtained in the *cis-trans* isomerization of N-acetyl-4(*cis*)-hydroxyproline methylamide by **(a)** counterclockwise rotation (ω = -90°) of the acetyl group (TS **3a**) and **(b)** clockwise rotation (ω = 90°) of the acetyl group (TS **3b**).

Interestingly, while there is little change in the geometrical parameters associated with the amide bond of the methylamide unit, there is noticeable elongation of the C—N(3) bonds by at least 0.02 Å in both TSs. This elongation of the bond is consistent with the change in the pyramidalization of the prolyl nitrogen.

For N-acetyl-4-hydroxyproline methylamide, rotation of the acetyl moiety by ω = - 90° generates TS **3a**, lying 64.9 kJ mol^{-1} higher in energy, in which the $N_3\cdots HN_5$ bond is shortened by 0.142 Å to 2.140 Å relative to that present in the *cis* structure. Clockwise rotation of the acetyl group in N-acetyl-4-hydroxyproline methylamide results in the formation of TS **3b**, lying 67.3 kJ mol^{-1} higher in energy than the *cis* rotamer, in which the $N_3\cdots HN_5$ bond has shortened by 0.106 Å to 2.176 Å. Unlike that observed for 4-fluoro substitution, the 4-hydroxy substituent results in the formation of a long and weak $OH\cdots N_5$ hydrogen bond between the OH group and the N of the methylamide unit in the *cis* (2.312 Å) structure and both TSs, **3a** (2.353 Å) and **3b** (2.589 Å) (Figure 6.6).

This secondary hydrogen bond is important as it encourages the formation of the intramolecular $N_3\cdots HN_5$ bond. As observed for N-acetyl-4-fluoroproline methylamide, the C^{γ} of proline is in an *endo* conformation and N3 has deformed to attain a more tetrahedral geometry in both TS **3a** (η = 125.2 °) and TS **3b** (η = 111.4 °) relative to that observed in the *cis* rotamer (η = 161.3 °). In addition, the N—C bond between the prolyl

nitrogen and the acetyl group is elongated by 0.078 Å (1.457 Å) and 0.060 Å (1.441 Å) in TSs **3a** and **3b**, respectively, relative to that observed in the corresponding *cis* rotamer.

6.3.3 The Effect of Substituents on C^{γ} trans to the Methylamide Moiety

In Table 6.2, the barriers to isomerization are summarized for N-acetylproline methylamide with the electron-withdrawing substituents on C^{γ} of proline *trans* to the methylamide unit. As was observed for the substituents on C^{γ} *cis* to the amide unit, the counterclockwise rotation (ω = - 90°) of the acetyl group yields TSs of lower energy than that obtained by clockwise rotation (ω = 90°) of the acetyl group. However, unlike that observed when the electron-withdrawing subsitutents were *cis* to the methylamide unit, placement of subsitutents in a *trans* conformation results in lower barriers to isomerization, relative to that observed for the unsubstituted isomerization, only in the case of hydroxy substitution. Fluoro substitution results in a slight increase in the barrier to isomerization.

Table 6.2: Summary of the barriers to *cis-trans* isomerization (kJ mol^{-1}) for N-acetylproline methylamide with substituents on C^{γ} of the proline ring *trans* to the methylamide moiety.

Substituent	$\omega = 90°$	$\omega = -90°$
unsubstituted	70.5	70.4
fluorine	72.5	71.8
hydroxy	71.5	65.8

As the acetyl group in the *cis* rotamer of N-acetyl-4-fluoroproline methylamide rotates counterclockwise, TS **4a** (Figure 6.7) is generated lying 71.8 kJ mol^{-1} higher in energy. The $N_3\cdots HN_5$ hydrogen bond shortens from 2.281 Å in the *cis* structure to 2.166 Å in TS **4a**. The rotation of the acetyl moiety results in C^{γ} adopting an *exo* conformation in TS **4a** and the prolyl nitrogen adopts a more pyramidalized geometry (η = 121.7°) relative to that observed in the *cis* rotamer (η = 161.9°). Clockwise rotation of the acetyl group enables the formation of an $N_5H\cdots O$ hydrogen bond (1.892 Å) in TS **4b** (Figure 6.7), lying slightly higher in energy at 72.5 kJ mol^{-1}.

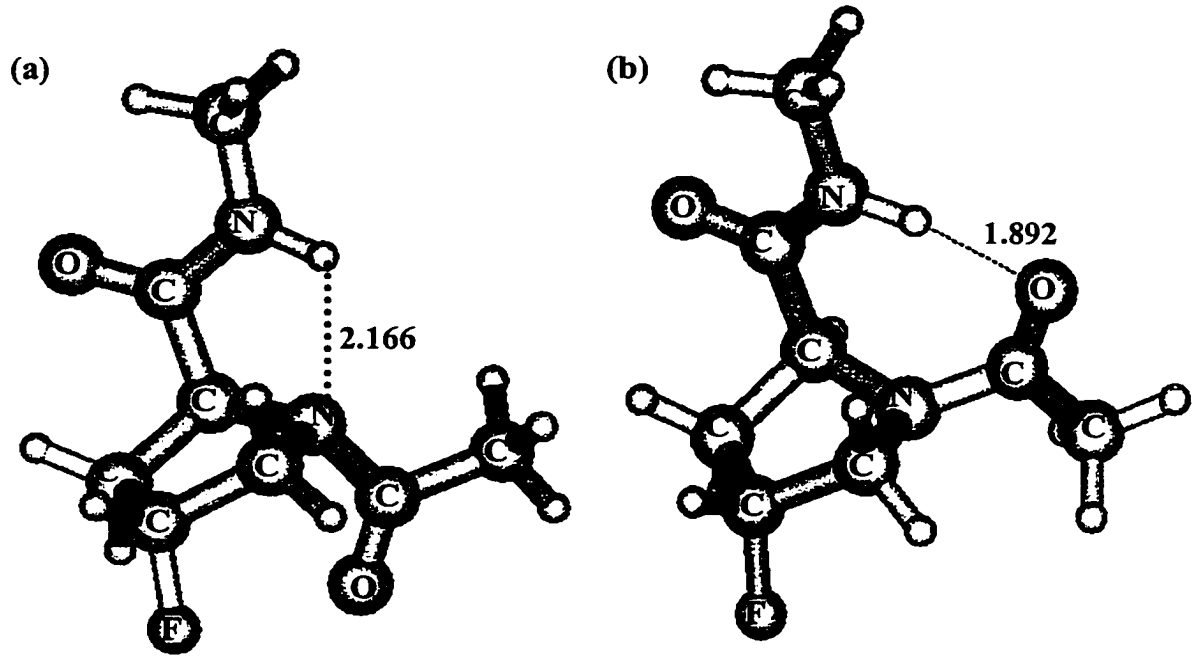


Figure 6.7: Transition structures obtained in the *cis-trans* isomerization of N-acetyl-4(*trans*)-fluoroproline methylamide, in which the 4-fluoro substituent is *trans* to the methylamide unit, by **(a)** counterclockwise rotation (ω = - 90°) of the acetyl group (TS **4a**) and, **(b)** clockwise rotation (ω = 90°) of the acetyl group (TS **4b**).

As was observed in TS **4a**, C^{γ} has adopted an *exo* conformation in TS **4b** and N3 has become more sp^3 like in hybridization (η = - 146.0°). However, as observed for the analogous rotation in which the 4-fluoro substituent was *cis* to the methylamide unit, the formation of the $N_5H\cdots O$ bond (1.892 Å) in TS **4b** is facilitated by the distortion in the planarity of the proline ring by C2 adopting an *endo* conformation, resulting in N3 puckering below the proline ring.

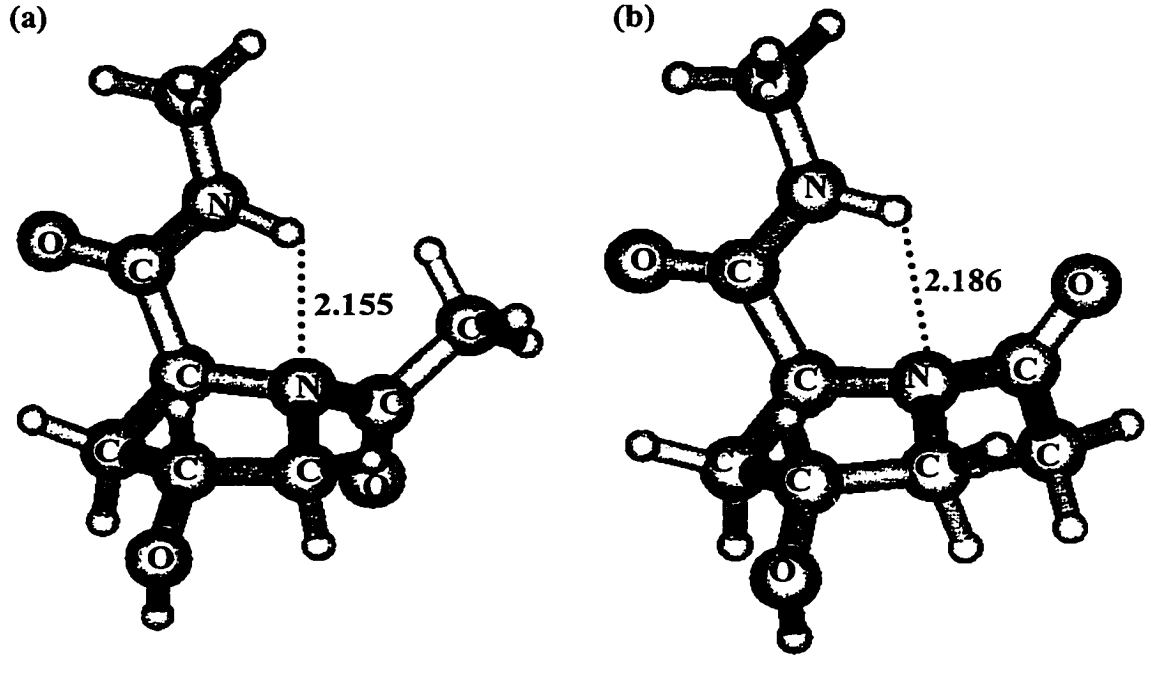


Figure 6.8: Transition structures obtained in the *cis-trans* isomerization of N-acetyl-4(*trans*)-hydroxyproline methylamide, in which the 4-hydroxy substituent is *trans* to the methylamide unit, by **(a)** counterclockwise (ω = - 90°) rotation of the acetyl group (TS **5a**) and, **(b)** clockwise rotation (ω = 90°) of the acetyl group (TS **5b**).

For the *cis-trans* isomerization of N-acetyl-4-hydroxyproline methylamide, in which the hydroxy substituent is *trans* to the methylamide unit, counterclockwise rotation (ω = - 90°) of the acetyl group enables the $N_3\cdots HN_5$ hydrogen bond present in the *cis* rotamer to shorten by 0.192 Å (2.155 Å) as TS **5a** (Figure 6.8) is generated, lying 65.8 kJ mol^{-1} higher in energy than the *cis* structure. The counterclockwise rotation of the acetyl group results in a distortion of the planarity of the proline ring as C^{γ} adopts an *endo* conformation and N3 perturbs to become more pyramidalized (η = 125.2°) relative to that observed in the *cis* rotamer (η = -163.0°).

Clockwise rotation of the acetyl group (ω = 90°) enables the formation of TS **5b** (Figure 6.8) lying slightly higher in energy that TS **5a** at 71.5 kJ mol^{-1}. As observed for TS **5a**, the $N_3\cdots HN_5$ hydrogen bond present in the *cis* rotamer (2.280 Å) has shortened to 2.186 Å in TS **5b** as C^{γ} adopts a slightly *endo* conformation and the prolyl nitrogen becomes more pyramidalized (η = 139.6°).

6.4 Conclusions

The *cis-trans* isomerization of N-acetylproline methylamide and its derivatives containing electron-withdrawing groups on the C^{γ} of proline, have been investigated using density functional theory. Relative to the unsubstituted isomerization of N-acetylproline methylamide, the presence of electron-withdrawing fluorine and hydroxyl

groups on the C^{γ} of proline results in a preferential rotation of the acetyl group in a counterclockwise direction to generate TSs with lower barriers to isomerization. This preferred direction of rotation enables the formation of an N···HN intramolecular hydrogen bond in the TSs and enhances the pyramidalization of the prolyl nitrogen, as is evident from the trend in the virtual dihedral angle (η) towards 120° and the elongation of the C—N bond length.

The effect of electron-withdrawing substituents on the overall barriers to isomerization is dependent upon the orientation of the substituents relative to the prolylamide group. The addition of fluoro and hydroxy subsitutents *cis* to the prolylamide group decreases the barrier to isomerization. However, the addition of the aforementioned substituents *trans* to the methylamide unit increases the barrier to isomerization for fluoro substitution while hydroxy substitution decreases the barrier to isomerization.

Thus, the presence of electron-withdrawing groups on the proline ring has significant geometrical and energetic consequences for the *cis-trans* isomerization. This study illustrates the important role the intramolecular N···HN hydrogen bond plays in the *cis-trans* isomerization of proline-containing molecules and the catalytic possibilities of hydrogen bonding.

6.5 References

(1) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397.

(2) Cox, C.; Lectka, T. *Acc. Chem. Res.* **2000**, *33*, 849.

(3) Karle, I. L. *Acc. Chem. Res.* **1999**, *32*, 693.

(4) Fischer, G. *Chem. Soc. Rev.* **2000**, *29*, 119.

(5) Rabenstein, D. L.; Shi, T.; Spain, S. *J. Am. Chem. Soc.* **2000**, *122*, 2401 (and references within).

(6) Pliyev, B. K.; Gurvits, B. Ya. *Biochemistry (Moscow)* **1999**, *64*, 738.

(7) Fischer, S.; Dunbrack R. L., Jr.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931.

(8) Fischer, S.; Michnick, S.; Karplus, M. *Biochemistry* **1993**, *32*, 13830.

(9) Cox, C.; Lectka, T. *J. Am. Chem. Soc.* **1998**, *120*, 10660.

(10) Eberhardt, E. S.; Panasik N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261.

(11) Cox, C.; Young, V. G., Jr.; Lectka, T. *J. Am. Chem. Soc.* **1997**, *119*, 2307.

(12) Beausoleil, E.; Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572 (and references within).

(13) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902.

(14) Jhon, J. S.; Kang, Y. K. *J. Phys. Chem. A* **1999**, *103*, 5436.

(15) Rablen, P. R. *J. Org. Chem.* **2000**, *65*, 7930.

Chapter Seven.
Conclusions and Future Work

7.1 Overview

In this thesis, molecules capable of forming hydrogen bonds have been examined in various systems of interest. The work may be divided into two broad sections. The first section, Chapter Three, examined the ability of the hybrid DFT methods to predict the structural and spectroscopic properties of the prototypical hydrogen halide dimers $(HX)_2$, X = F, Cl and Br. The second section, Chapters Four to Six, involved the application of the hybrid DFT method, B3LYP, to study the role of hydrogen bonding in systems containing biologically important molecules. General conclusions from each of these sections will be derived in addition to a description of the potential avenues of future work that results from these sections.

7.2 Conclusions and Future Work

7.2.1 Hydrogen Halide Dimers

The hydrogen halide dimers $(HX)_2$ X = F, Cl and Br, have been investigated with the 6-311+G(2df,p) basis set and a variety of density functional theory methods. The lowest energy conformation of these dimers is a bent structure of C_s symmetry in which the bridging proton is within 10° of the internuclear X···X′ axis (described by θ_1) and the proton of the acceptor molecule is nearly perpendicular (described by θ_2) to the aforementioned axis. All hybrid methods predict geometrical parameters {X···X′, θ_1, θ_2} in reasonable agreement with experiment. For the heavier halide dimers, $(HCl)_2$ and $(HBr)_2$, marginally better agreement was attained with the BHandHLYP or B1LYP methods. Although the binding energies, corrected for ZPVE and BSSE, were adequately determined for $(HF)_2$ using the DFT methods containing the LYP functional, the binding energies for $(HCl)_2$ and $(HBr)_2$ were poorly reproduced by all methods. The frequency shift associated with the X—H bond in all three dimers was also examined in this study. For all three dimers, reasonable agreement with the experimental results was obtained by methods containing the LYP functional. Overall, the parameters of interest were adequately predicted by the DFT methods and the BHandHLYP or B1LYP methods offer a viable alternative to the popular B3LYP.

However, this study has raised some important questions and areas that need to be addressed. The 6-311+G(2df,p) basis set was utilized in the study based on the results of a basis set study with the MP2 and QCISD methods. For $(HF)_2$ and $(HCl)_2$, this basis set adequately predicted the internuclear distance and angles of interest. However, the erroneous binding energies obtained for $(HCl)_2$ and $(HBr)_2$ suggest that although it may yield adequate geometrical parameters, it is insufficient to describe the binding in these systems. Thus, a more detailed study of the relationship between the basis set and the binding energy is warranted. In addition, it would be interesting to determine the nature of the relationship between the geometrical parameters and the binding energy. Is it more important to accurately determine the internuclear distance or θ_1 when the property of interest is the binding energy? In addition, as the subject of this study included some heavy halides, Cl and Br, it is possible relativistic effects may be very important and this contribution to the properties of interest should be examined in more detail.

7.2.2 Catalysis Involving Hydrogen Bonding

In Chapters Four to Six of this thesis, biological molecules and their hydrogen-bonding interactions in reactions of biochemical significance were investigated using the B3LYP method. In Chapter Four, the study of the aminolysis of 6-chloropyrimidine illustrated how the formation of multiple hydrogen bonding interactions between 6-chloropyrimidine and the derivatives of the nucleobase uracil stabilize the transition structures, thus catalyzing the aminolysis reaction. The crucial hydrogen bonding

interaction in this reaction is between the carbonyl group of the base and the incoming NH_3. By adjusting the proton affinity of the carbonyl group of the base through fluoro, imine and sulfur substitution, the barrier to aminolysis was distorted in a manner predictable on the basis of the proton affinity of the carbonyl group of the base. This study illustrates the importance of a well-chosen hydrogen bond acceptor and the catalytic possibilities of hydrogen bonding.

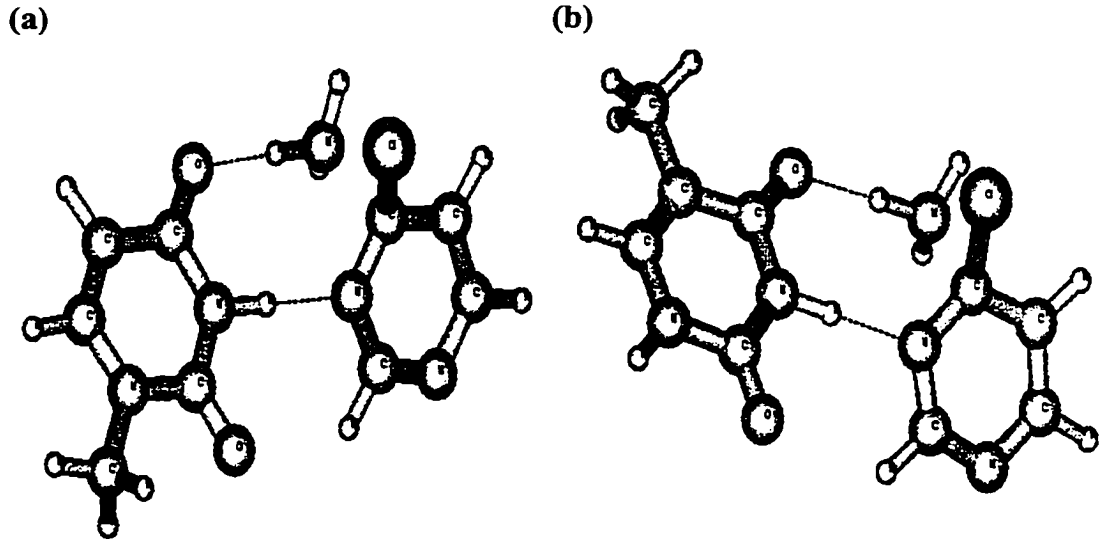


Figure 7.1: Schematic illustration of the transition structure in the aminolysis of 6-chloropyrimidine with, **(a)** 1-methyluracil in a *syn* conformation, and **(b)** 1-methyluracil in an *anti* conformation.

In the original study by Tominaga et al.,[1] the formation of the hydrogen bonds in the aminolysis of 6-chloropurine was observed by ^1H NMR. Thus, it would be of interest to calculate the corresponding proton chemical shifts in our model reaction to see if the

downfield shifts for the signals due to the formation of hydrogen-bonding interactions could be observed and reproduced. In addition, our study was incomplete in that in the aminolysis of 6-chloropyrimidine involving 1-methyluracil as the base, we only examined the *syn* conformation of the TSs. Tominaga and coworkers[1] also postulated an *anti* structure. Although the hydrogen bonding patterns in the two TSs are equivalent (Figure 7.1), in the *anti* configuration the electron-donating $-CH_3$ group would be in closer proximity to the carbonyl group which interacts with the incoming NH_3. In this case, the electron-donating effects may be stronger and have a more marked effect on the barrier to aminolysis than that observed in the *anti* conformation.

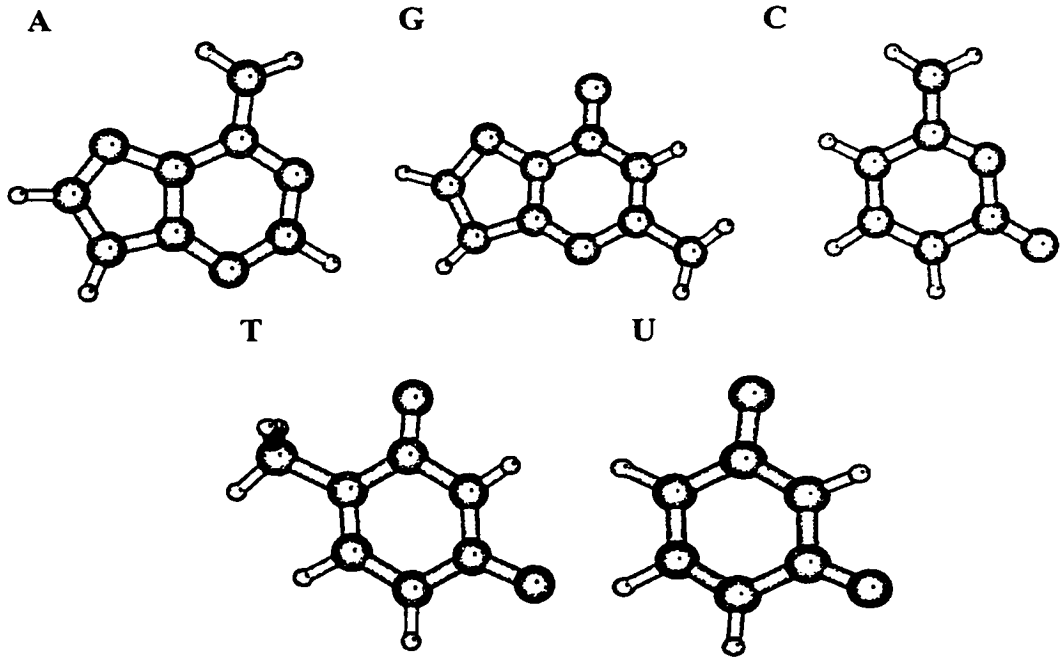


Figure 7.2: Illustration of the purine (adenine (A) and guanine (G)) and pyrimidine (cytosine (C), thymine (T) and uracil (U)) nucleobases.

This study explicitly focused on the use of uracil as the catalyzing agent. However, uracil is only one of the nucleobases present in biological systems and thus, it would be of interest to examine how the other purine and pyrimidine bases, Figure 7.2, would perform in the aforementioned aminolysis reaction. The purine and pyrimidine bases would alter the nature of the hydrogen bonding present in the TSs, a factor which would be reflected in the calculated barriers to aminolysis.

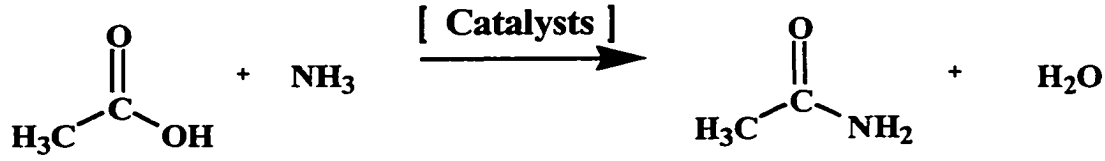


Figure 7.3: The aminolysis reaction of acetic acid by ammonia in the presence of various catalysts.

This would nicely complement a theoretical study of the ester aminolysis as catalyzed by nucleobases. The original experimental study by Melander and Horne[2] examined the aminolysis reaction of pentafluorophenylbenzoate by propylamine and was the first study to explicitly illustrate the catalytic potential of the functional groups of the nucleobases. Melander and Horne[2] postulated that the breakdown of the tetrahedral intermediate yielding the amide product is the rate-limiting step and that this step could be accelerated through the addition of various nucleobases which would stabilize the ionic TSs. Using the model reaction in Figure 7.3, in which the catalysts would be the

nucleobases schematically illustrated in Figure 7.2, correlations between the barriers to aminolysis, the hydrogen bond distances present in the TSs, and the experimentally observed rates of aminolysis could be drawn.

The aforementioned studies examine the ability of molecules capable of forming intermolecular hydrogen bonding to catalyze chemical reactions by stabilizing the transition structures. In Chapter Six, the *cis-trans* isomerization of prolyl amides, using N-acetylproline methylamide as a model system, was catalyzed by the formation of an intramolecular $N_3 \cdots HN_5$ hydrogen bond between the prolyl nitrogen and the methylamide unit. The addition of electron-withdrawing substituents on C^{γ} in N-acetylproline methylamide results in the preferential rotation of the acetyl group in a counterclockwise direction. This mode of rotation assists in the formation of the intramolecular $N_3 \cdots HN_5$ hydrogen bond in the TSs and an enhancement in the pyramidalization of the prolyl nitrogen, leading to a reduced barrier to *cis-trans* isomerization. This study illustrates the importance of the intramolecular hydrogen bond and the catalytic possibilities of hydrogen bonding.

Experimental studies[3,4] of the *cis-trans* isomerization of prolyl amides have also examined the effect of alkyl groups on the rate of isomerization. The study of alkylprolines has significant biological implications in that such systems may be used to develop and expand the knowledge about the relationship between conformation and

activity in addition to providing information on conformational effects on prolyl isomerization.[4-6] Thus it would be advantageous to extend the current study to examine how alkyl substitution on the proline ring affects the formation of the intramolecular $N_3\cdots HN_5$ hydrogen bond and the subsequent barrier to isomerization.

7.2.3 Catalysis Involving Amino Acids

In Chapter Five, the amino acid proline is once again the subject of study. Rather than using it as a model in the study of peptide isomerization of proline-containing molecules, in this study it is employed as a catalyst in the direct aldol reaction.

The enamine mechanism proposed by List and coworkers[7] for the direct aldol reaction between acetone and a variety of aldehydes was examined by DFT using the model reaction of acetone with acetaldehyde in the presence of proline. Proline acts as an enzyme mimic by providing the nucleophilic amino group, which interacts with acetone to yield the intermediate enamine complex. A very large barrier that may inhibit further progressions of the reaction accompanies this initial interaction. However, by inclusion of a moderately ionizing solvent (DMSO), an alternate, lower-energy pathway is available for the initial interaction between proline and acetone. Thus, the direct aldol reaction illustrates the catalytic potential of simple organic molecules in asymmetric synthesis.

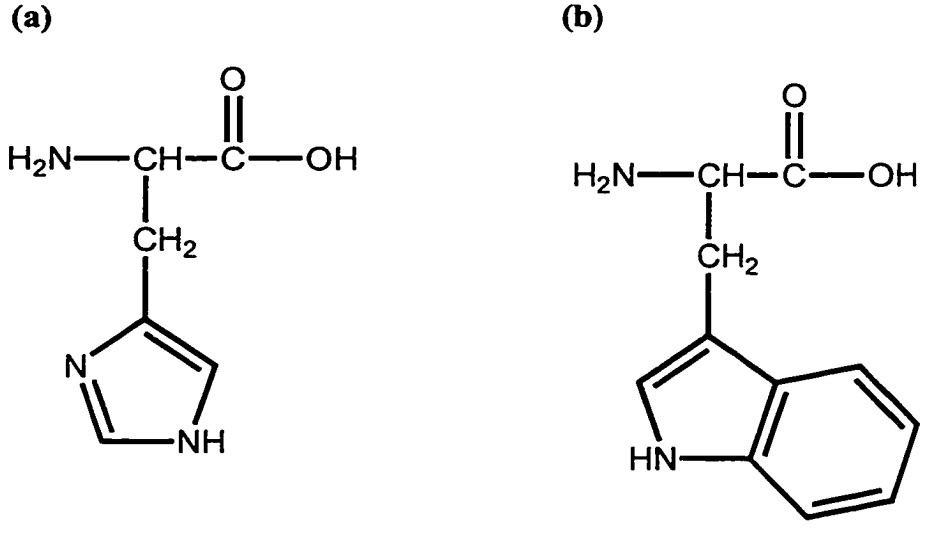


Figure 7.4: Schematic illustration of the amino acids **(a)** histidine and, **(b)** tryptophan.

As the amino and carboxylate groups present in proline both play important roles in the proposed mechanism, this raises the question as to whether another naturally occurring amino acid would be able to act as a catalyst in the aldol reaction. The amino acids tryptophan and histidine, Figure 7.4, both contain a carboxylic acid moiety and an amino group within a five-membered ring. However, in these two amino acids, the amino group and carboxylate group are separated to a greater degree than that encountered in proline, posing another question. What makes proline an effective catalyst? Is it the proximal location of the two groups or just their mere presence in the molecule? A theoretical examination into the ability of other amino acids as well as small molecules that contain the NH and COOH functional groups to catalyze the aldol

reaction could provide potential catalysts for use in organic synthesis as well as expanding the chemical basis for an efficient catalyst.

The DFT examination of the direct aldol reaction illustrates the stabilizing ability of ionizing solvent. As solvent plays a crucial role in this study by providing an alternate, lower energy reaction pathway for the aldol reaction, a more detailed examination into the effect of solvent would be insightful. By systematically varying the strength of the solvent, an alternate reaction medium may be revealed.

7.3 Summation

Throughout the work discussed within this thesis and that proposed for future research, the diversity and importance of hydrogen bonding has been explored. The work summarized provides an overview of the catalytic possibilities of intra- and inter-molecular hydrogen bonding involving biologically relevant molecules. It has laid the foundation for future theoretical and experimental studies of such processes, as outlined in the discussion of the future work. As all the projects presented within this thesis have been examined by experimental techniques, this thesis illustrates the complementary nature of theoretical quantum chemistry. Through the combination of experimental and theoretical chemistry, a more complete understanding of important chemical phenomena may be attained.

7.4 References

(1) Tominaga, M.; Konishi, K.; Aida, T. *J. Am. Chem. Soc.* **1999**, *121*, 7704.

(2) Melander, C.; Horne, D. A. *J. Org. Chem.* **1996**, *61*, 8344.

(3) Beausoleil, E.; Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572.

(4) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902.

(5) Baures, P.W.; Ojala, W. H.; Gleason, W. R.; Johnson, R. L. *J. Peptide Res.* **1997**, *50*, 1.

(6) Sharma, R.; Lubell, W. D. *J. Org. Chem.* **1996**, *61*, 202.

(7) List, B.; Lerner, R. A.; Barbas III, C. F. *J. Am. Chem. Soc.* **2000**, *122*, 2395.

Appendix A

Table A1.1: Optimized geometrical parameters[a] for $(HF)_2$ obtained at the MP2 and QCISD levels of theory with a variety of basis sets.

Basis Set	r(F···F')		θ_1		θ_2		r(H···F')	
	MP2	QCISD	MP2	QCISD	MP2	QCISD	MP2	QCISD
6-311G	2.707	2.730	15.2	14.5	109.3	111.4	1.816	1.836
6-311G(d)	2.681	2.705	17.7	17.0	100.1	102.2	1.824	1.844
6-311G(d,p)	2.709	2.725	15.5	15.5	98.3	99.4	1.840	1.859
6-311+G(d,p)	2.788	2.799	6.1	6.2	121.3	121.8	1.874	1.888
6-311++G(d,p)	2.787	2.799	6.0	6.0	121.6	122.1	1.873	1.887
6-311+G(2d,p)	2.767	2.781	7.7	7.8	114.9	115.6	1.851	1.885
6-311+G(2d,2p)	2.761	2.774	5.7	5.7	111.7	112.6	1.845	1.870
6-311+G(2df,p)	2.759	2.775	7.3	7.4	116.4	117.0	1.845	1.861
6-311+G(2df,2p)	2.758	2.772	5.6	5.7	112.2	113.0	1.841	1.864
6-311+G(3df,3pd)	2.738	2.749	3.8	3.9	115.6	116.6	1.818	1.833
cc-pVDZ	2.681	2.700	12.1	11.7	99.5	101.6	1.787	1.805
cc-pVTZ	2.717	2.734	8.8	8.7	105.3	106.9	1.810	1.830
aug-cc-pVDZ	2.754	2.767	6.6	6.6	110.7	111.0	1.832	1.848
aug-cc-pVTZ	2.745	2.756	6.3	6.7	111.0	111.9	1.826	1.842
Experiment[b]	2.72 ± 0.03		10 ± 6		117 ± 6		----	

[a] Bond lengths in Å, bond angles in degrees.
[b] Reference 6 and 7.

Table A1.2: Optimized geometrical parameters[a] for $(HCl)_2$ obtained at the MP2 and QCISD levels of theory with a variety of basis sets.

Basis Set	r(Cl···Cl′)		θ_1		θ_2		r(H···Cl′)	
	MP2	QCISD	MP2	QCISD	MP2	QCISD	MP2	QCISD
6-311G	4.172	4.216	7.6	8.2	110.6	108.9	2.861	2.898
6-311G(d)	4.035	4.104	6.1	6.2	104.0	103.2	2.760	2.825
6-311G(d,p)	4.026	4.098	5.5	5.8	102.2	101.6	2.760	2.830
6-311+G(d,p)	3.946	4.006	4.8	5.2	101.9	101.1	2.677	2.736
6-311++G(d,p)	3.936	3.996	4.0	4.0	102.4	102.1	2.666	2.722
6-311+G(2d,p)	3.911	4.014	8.2	9.0	92.9	92.9	2.653	2.760
6-311+G(2d,2p)	3.773	3.882	4.5	5.1	92.2	93.0	2.503	2.615
6-311+G(2df,p)	3.883	4.003	8.3	8.3	92.1	92.6	2.621	2.740
6-311+G(2df,2p)	3.749	3.866	4.3	4.8	92.0	93.2	2.475	2.594
6-311+G(3df,3pd)	3.775	3.882	5.5	6.9	90.3	90.0	2.508	2.619
cc-pVDZ	3.909	3.988	9.9	8.8	89.3	92.1	2.649	2.719
cc-pVTZ	3.786	3.909	6.8	7.4	90.2	91.5	2.521	2.646
aug-cc-pVDZ	3.817	3.898	6.0	5.9	90.7	91.9	2.535	2.615
aug-cc-pVTZ	3.744	3.853	5.7	6.0	89.9	91.6	2.473	2.583
Experiment	3.75 – 3.84[b]		9.0[c]		90.0[c]		----	

[a] Bond lengths in Å, bond angles in degrees.
[b] References 3 and 4.
[c] References 18 and 21.

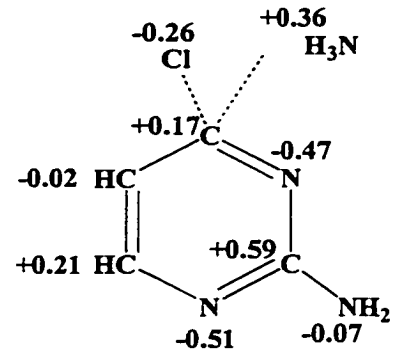
Appendix B

Table B1.1: B3LYP/6-311+G(2df,p) energies, with and without B3LYP/6-31G(d,p) ZVPE corrections (E_e and E_0, respectively), for all species referred to in the study.

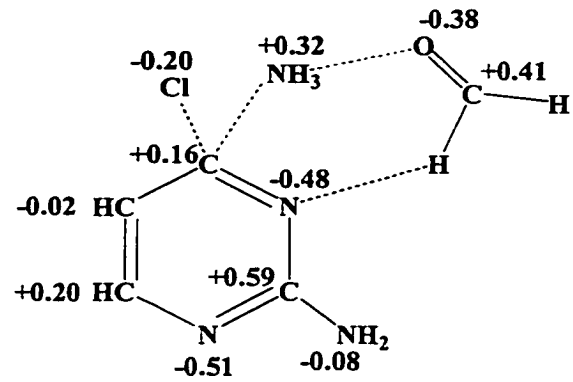
Species	Total Energy (hartrees)	
	E_0	E_e
Cl-$C_4N_2H_3$	-724.03848	-723.97098
NH_3	-56.58411	-56.54969
HCl	-460.83645	-460.82973
$C_4N_3H_5$	-319.80382	-319.70977
Cl-$C_4N_3H_4$	-779.43751	-779.35332
$C_4N_4H_6$	-375.20084	-375.09877
OCH_2	-114.54701	-114.52042
OHC-NH_2	-169.96249	-169.91709
OHC-NH-CHO	-283.31880	-283.27551
1-methyluracil	-454.17383	-454.28876
OCHF	-213.85047	-213.82943
OFC-NH_2	-269.25627	-269.21797
OFC-NH-CHO	-382.58913	-382.57248
SCH_2	-437.50905	-437.48423
SHC-NH_2	-492.91932	-492.87545
SHC-NH-CHO	-606.28824	-606.23426
$HNCH_2$	-94.66674	-94.62677
HNCH-NH_2	-150.06651	-150.00904
HNCH-HN-CHO	-263.43744	-263.37278
$HOCH_2^+$	-114.78724	-114.82790
HOHC-NH_2^+	-170.22501	-170.28484
HOHC-NH-CHO^+	-283.56836	-283.63711
1-methyluracil$^+$	-454.50196	-454.62946
1-methyluracil$^+$	-454.49043	-454.61757
$HOCFH^+$	-214.07576	-214.10971
HOFC-NH_2^+	-269.50122	-269.55301
HOFC-NH-CHO^+	-382.84563	-382.90630
$HSCH_2^+$	-437.77507	-437.81035
HSHC-NH_2^+	-493.19830	-493.25273
HSHC-NH-CHO^+	-606.53966	-606.60294
$HNHCH_2^+$	-94.95504	-95.00938
HNHHC-NH_2^+	-150.36926	-150.44189
HNHHC-NH-CHO^+	-263.71033	-263.79220

Figure B1.1: Schematic illustrations of the transition structures showing charge distribution on the heavy atoms, determined by Mulliken population analysis (hydrogens summed into heavy atoms).

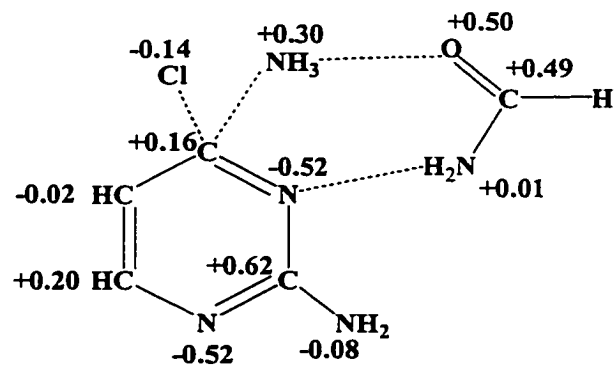
4b



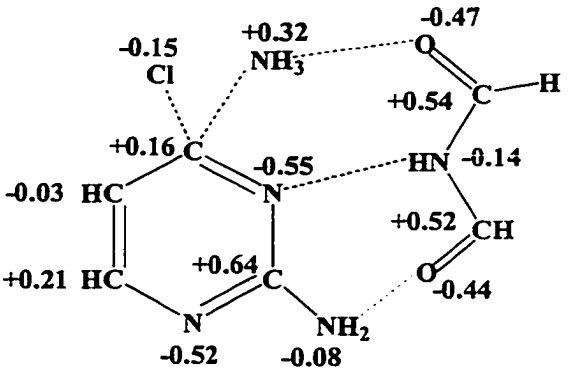
5b



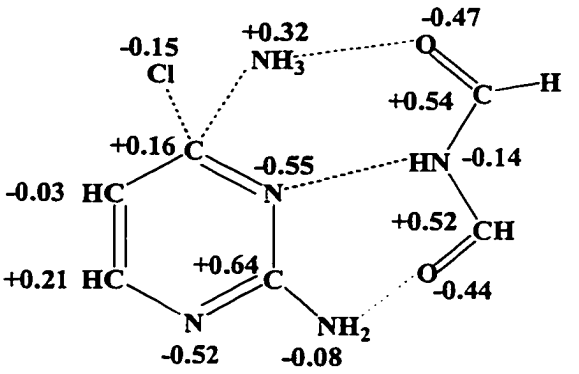
6b



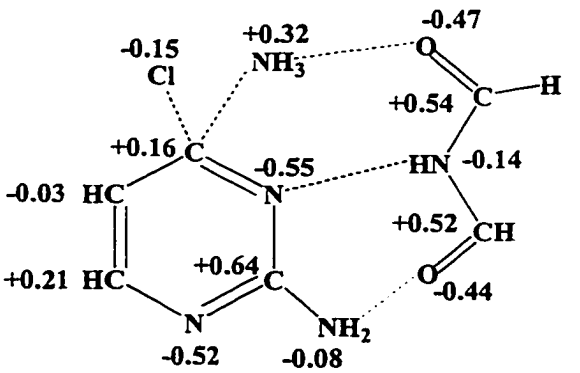
7b


-0.15
Cl
+0.32
NH3
-0.47
O
C
H
+0.54
+0.16
C
-0.55
HN
-0.14
-0.03
HC
N
+0.52
CH
+0.64
+0.21
HC
C
O
-0.44
N
NH2
-0.52
-0.08

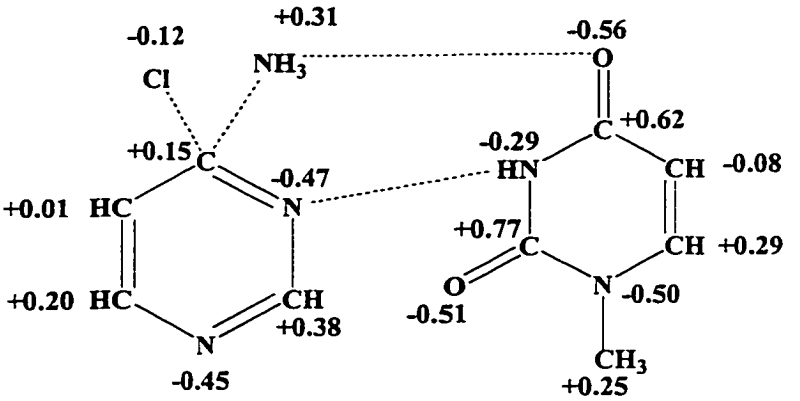
8b


-0.15
Cl
+0.32
NH3
-0.47
O
C
H
+0.54
+0.16
C
-0.55
HN
-0.14
-0.03
HC
N
+0.52
CH
+0.64
+0.21
HC
C
O
-0.44
N
NH2
-0.52
-0.08

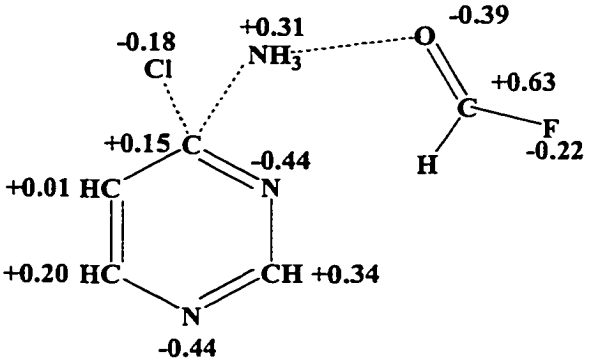
9b


-0.15
Cl
+0.32
NH3
-0.47
O
C
H
+0.54
+0.16
C
-0.55
HN
-0.14
-0.03
HC
N
+0.52
CH
+0.64
+0.21
HC
C
O
-0.44
N
NH2
-0.52
-0.08

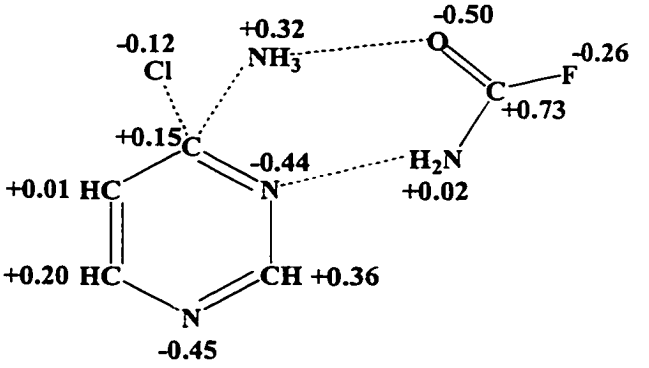
10b


-0.12
+0.31
-0.56
Cl
NH3
O
C
+0.62
-0.29
+0.15
C
HN
CH
-0.08
-0.47
+0.01
HC
N
+0.77
C
CH
+0.29
O
N
-0.50
+0.20
HC
CH
-0.51
+0.38
N
CH3
-0.45
+0.25

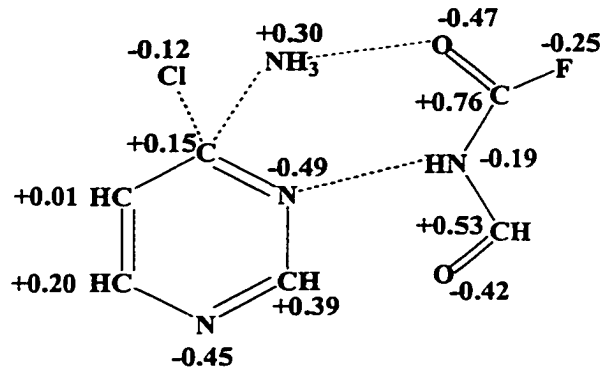
11b


-0.18
+0.31
-0.39
Cl
NH3
O
+0.63
C
F
+0.15
C
-0.44
-0.22
+0.01
HC
N
H
+0.20
HC
CH
+0.34
N
-0.44

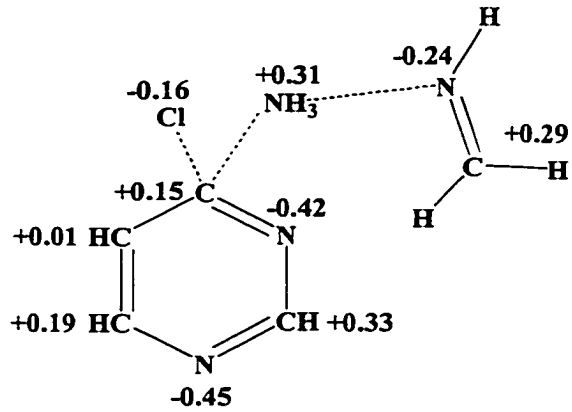
12b


-0.12
+0.32
-0.50
Cl
NH3
O
-0.26
F
C
+0.73
+0.15
C
-0.44
H2N
+0.01
HC
N
+0.02
+0.20
HC
CH
+0.36
N
-0.45

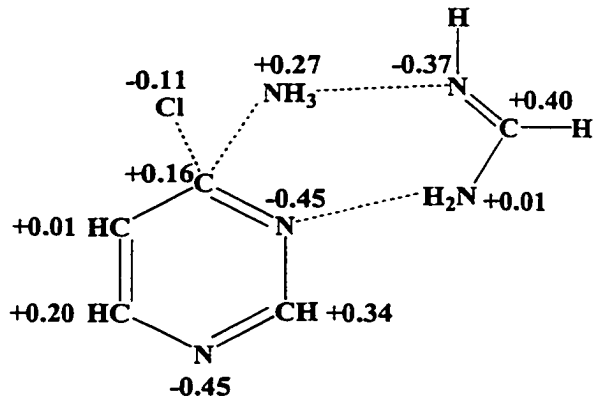
13b


-0.47
+0.30
-0.25
-0.12
NH3
O
F
Cl
+0.76
C
+0.15
C
-0.49
HN
-0.19
+0.01 HC
N
+0.53 CH
O
+0.20 HC
CH
-0.42
+0.39
N
-0.45

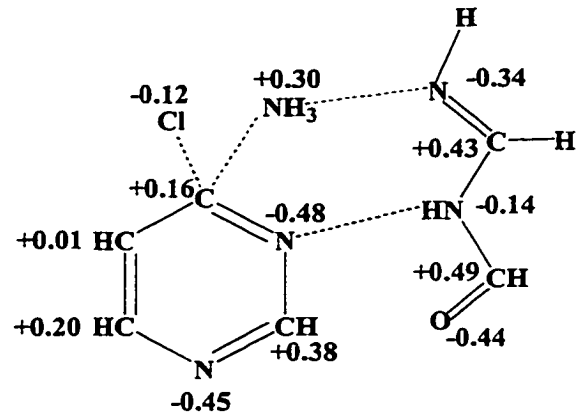
14b


H
-0.24
+0.31
-0.16
N
NH3
Cl
+0.29
C
H
+0.15
C
-0.42
H
+0.01 HC
N
+0.19 HC
CH +0.33
N
-0.45

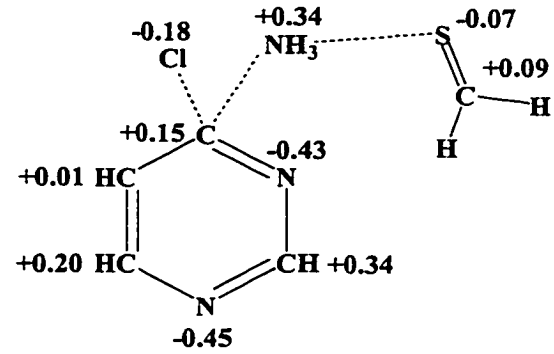
15b


H
+0.27
-0.37
-0.11
NH3
N
Cl
+0.40
C
H
+0.16
C
-0.45
H2N +0.01
+0.01 HC
N
+0.20 HC
CH +0.34
N
-0.45

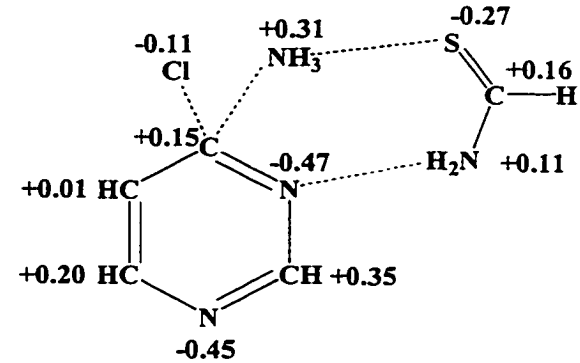
16b



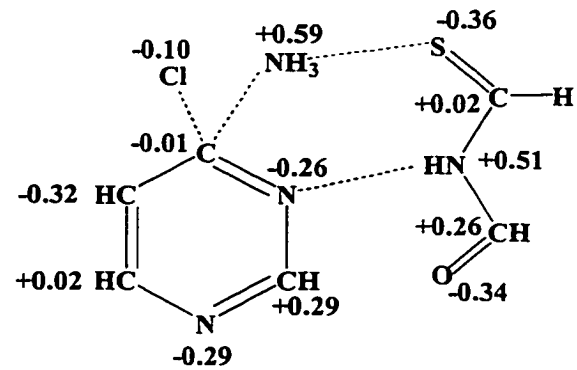
17b



18b



19b



Appendix C

Table C1.1: B3LYP/6-311+G(2df,p) energies, with and without B3LYP/6-31G(d,p) ZVPE corrections (E_e and E_0, respectively), for all species on the non-solvated surface.

Species	Total Energy (hartrees)	
	E_0	E_e
Proline	-401.30203	-401.15653
CH_3OCH_3	-193.22684	-193.14310
1a	-594.53553	-594.30468
1b	-594.46854	-594.23941
1c	-594.52099	-594.28669
1d	-594.50149	-594.27218
1e	-594.51886	-594.28877
1f	-518.03712	-517.83242
1g	-518.02195	-517.82102
1g'	-518.00019	-517.79915
1h	-518.04910	-517.84286
2a	-671.94605	-671.68285
2b	-671.92514	-671.66104
2c	-671.94291	-671.67512
2d	-748.42280	-748.12929
2e	-748.42213	-748.12634
2f	-748.40831	-748.11455
2g	-748.40881	-748.11247
2h	-748.40233	-748.10943
2i	-748.43056	-748.13714
Aldol product	-347.12871	-346.98230
H_2O	-76.46022	-76.43885
$HCO\text{-}CH_3$	-153.88728	-153.83197

Table C1.2: B3LYP/6-311+G(2df,p) energies, with and without B3LYP/6-31G(d,p) ZVPE corrections (E_e and E_0, respectively), for all species on the DMSO solvated surface.

Species	Total Energy (hartrees)	
	E_0	E_e
Proline	-401.29283	-401.14832
CH_3OCH_3	-193.22919	-193.14543
3a	-594.52998	-594.29881
3b	-594.51312	-594.28329
3c	-594.52944	-594.29447
3d	-594.53011	-594.30022
3e	-518.05579	-517.85063
3f	-518.03242	-517.83152
3g	-518.05315	-517.84688
4a	-671.95154	-671.68864
4b	-671.93908	-671.67095
4c	-671.95627	-671.68834
4d	-748.43398	-748.14036
4e	-748.42669	-748.13074
4f	-748.41818	-748.12517
4g	-748.42125	-748.12441
4h	-748.40869	-748.11542
4i	-748.43350	-748.14018
Aldol product	-347.13172	-346.98521
H_2O	-76.46334	-76.44198
$HCO\text{-}CH_3$	-153.89019	-153.83487

Appendix D

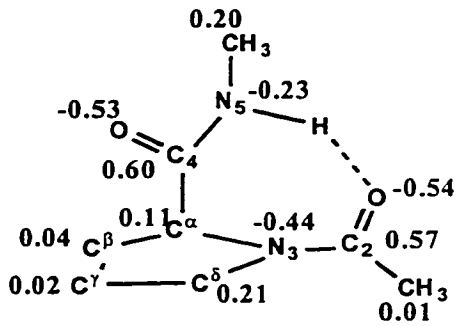
Table D1.1: B3LYP/6-311+G(2df,p) energies, with and without B3LYP/6-31G(d,p) ZVPE corrections (E_e and E_0, respectively), for all species examined.

Species	Total Energy (hartrees)	
	E_0	E_e
N-acetylproline methylamide (*cis*)	-573.46175	-573.23870
N-acetylproline methylamide (*trans*)	-573.46718	-573.24366
N-acetylproline methylamide (TS1a)	-573.43361	-573.21185
N-acetylproline methylamide (TS1b)	-573.43396	-573.21188
N-4-fluoro(*cis*)acetylproline methylamide (*cis*)	-672.73582	-672.52074
N-4-fluoro(*cis*)acetylproline methylamide (*trans*)	-672.73708	-672.52188
N-4-fluoro(*cis*)acetylproline methylamide (TS2a)	-672.70934	-672.49534
N-4-fluoro(*cis*)acetylproline methylamide (TS2b)	-672.70293	-672.48909
N-4-fluoro(*trans*)acetylproline methylamide (*cis*)	-672.73258	-672.51756
N-4-fluoro(*trans*)acetylproline methylamide (*trans*)	-672.43942	-672.52384
N-4-fluoro(*trans*)acetylproline methylamide (TS3a)	-672.70409	-672.49020
N-4-fluoro(*trans*)acetylproline methylamide (TS3b)	-672.70397	-672.48967
N-4-hydroxy(*cis*)acetylproline methylamide (*cis*)	-648.70543	-648.47780
N-4-hydroxy(*cis*)acetylproline methylamide (*trans*)	-648.70614	-648.47860
N-4-hydroxy(*cis*)acetylproline methylamide (TS4a)	-648.67979	-648.45309
N-4-hydroxy(*cis*)acetylproline methylamide (TS4b)	-648.67836	-648.45215
N-4-hydroxy(*trans*)acetylproline methylamide (*cis*)	-648.70451	-648.47736
N-4-hydroxy(*trans*)acetylproline methylamide (*trans*)	-648.71047	-648.48275
N-4-hydroxy(*trans*)acetylproline methylamide (TS5a)	-648.67865	-648.45229
N-4-hydroxy(*trans*)acetylproline methylamide (TS5b)	-648.67600	-648.45011

Figure D1.1: Schematic illustrations of the transition structures showing charge distribution on the heavy atoms, determined by Mulliken population analysis (hydrogens summed into heavy atoms).

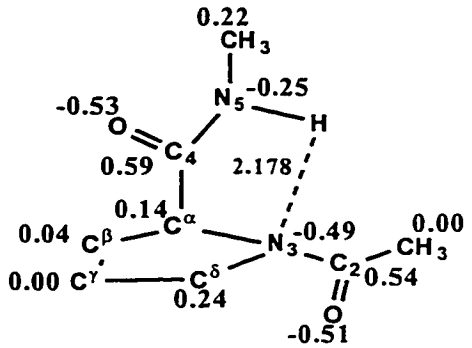
N-acetylproline methylamide

***cis* rotamer**



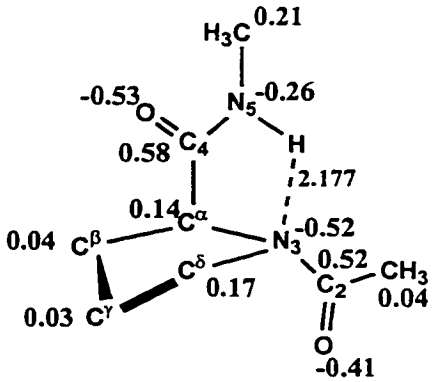
N-acetylproline methylamide

***trans* rotamer**



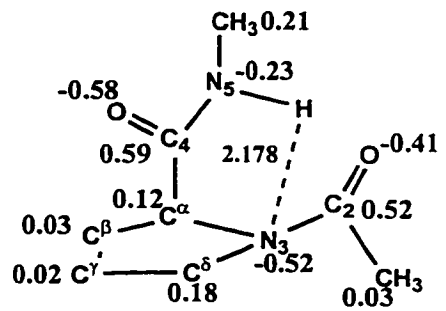
N-acetylproline methylamide

TS 1a



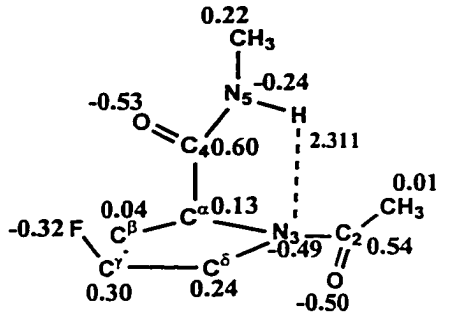
N-acetylproline methylamide

TS 1b



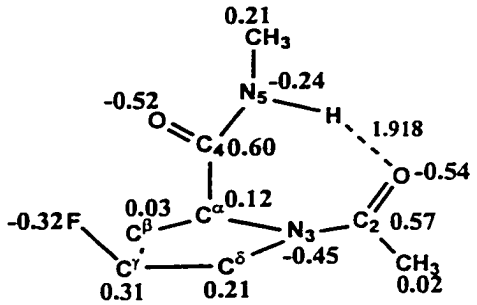
N-acetyl-4-fluoro(*cis*)proline methylamide

***cis* rotamer**



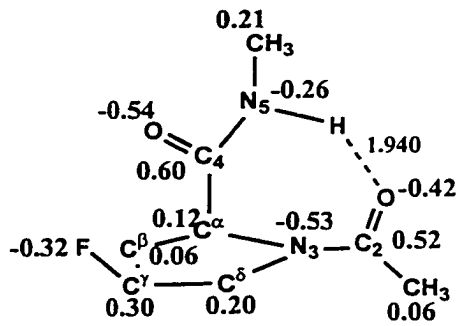
N-acetyl-4-fluoro(*cis*)proline methylamide

***trans* rotamer**



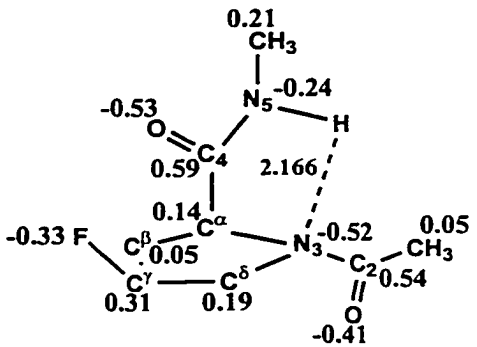
N-acetyl-4-fluoro(*cis*)proline methylamide

TS 2a



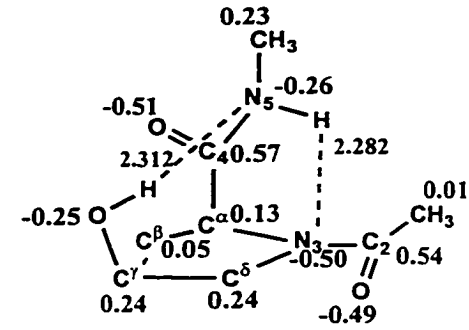
N-acetyl-4-fluoro(*cis*)proline methylamide

TS 2b



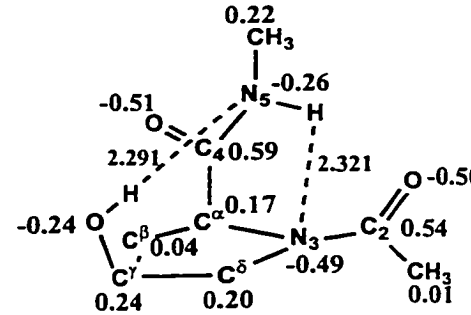
N-acetyl-4-hydroxy(*cis*)proline methylamide

***cis* rotamer**



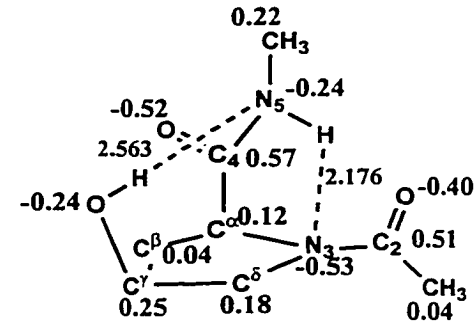
N-acetyl-4-hydroxy(*cis*)proline methylamide

***trans* rotamer**



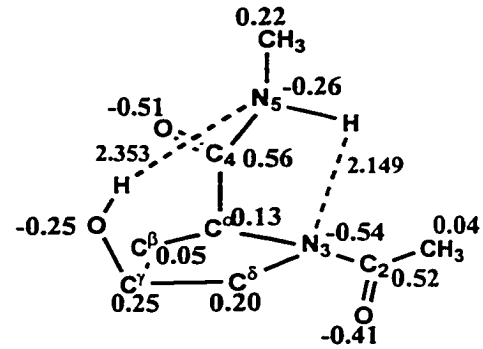
N-acetyl-4-hydroxy(*cis*)proline methylamide

TS 3a



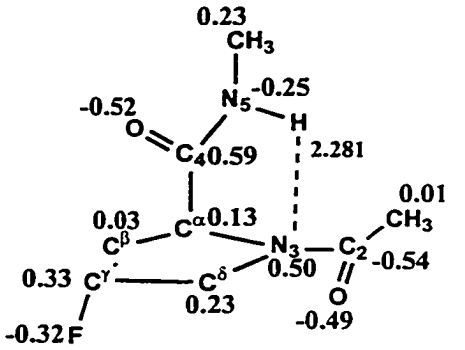
N-acetyl-4-hydroxy(*cis*)proline methylamide

TS 3b



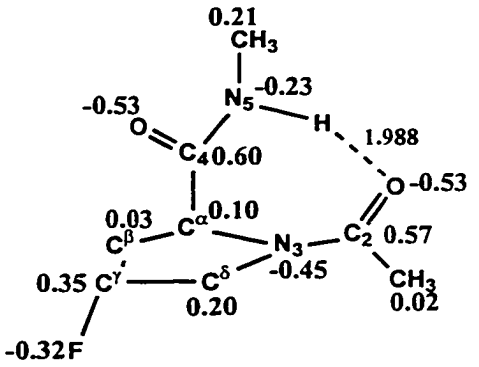
N-acetyl-4-fluoro(*trans*)proline methylamide

***cis* rotamer**



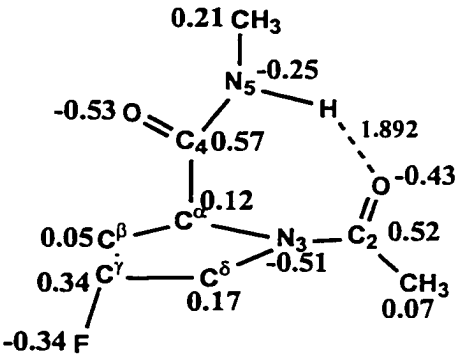
N-acetyl-4-fluoro(*trans*)proline methylamide

***trans* rotamer**



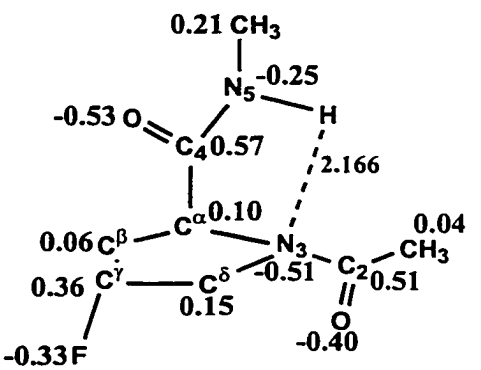
N-acetyl-4-fluoro(*trans*)proline methylamide

TS 4a



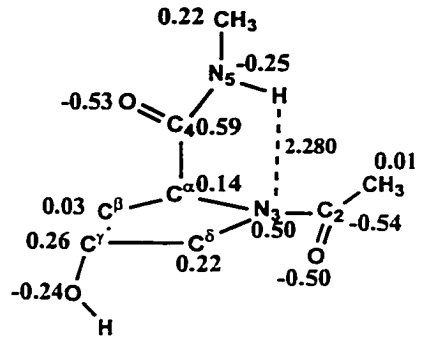
N-acetyl-4-fluoro(*trans*)proline methylamide

TS 4b



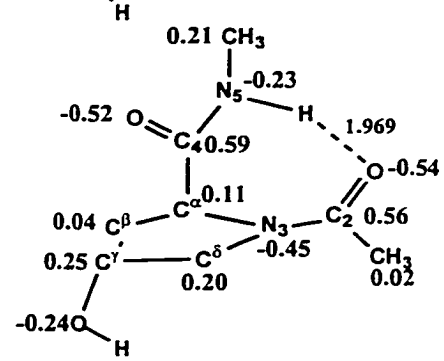
N-acetyl-4-hydroxy(*trans*)proline methylamide

***cis* rotamer**



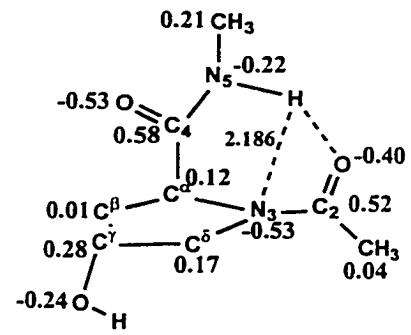
N-acetyl-4-hydroxy(*trans*)proline methylamide

***trans* rotamer**



N-acetyl-4-hydroxy(*trans*)proline methylamide

TS 5a



N-acetyl-4-hydroxy(*trans*)proline methylamide

TS 5b