A Computational Study of Nitrogen Heterocyclic Compounds

by

Christopher Ian Williams

Department of Chemistry, McGill University,

Montreal, Quebec, Canada

December 1996

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

This thesis has been prepared in the format of published or to be published papers, as described in the McGill University Faculty of Graduate Studies and Research *Guidelines Concerning Thesis Preparation:* The following text has been reproduced from those guidelines.

> Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly-duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.
>
> If this option is chosen, **connecting texts that provide logical bridges between the different papers are mandatory.** The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.
> The thesis must still conform to all other requirements of the "Guidelines for Thesis preparation." **The thesis must include:** A Table of Contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.
>
> Additional material must be provided where appropriate (e.g. in appendices) and in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported in the thesis.
>
> In case of manuscripts co-authored by the candidate and others, **the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent.** Supervisors must attest to the accuracy of such statements at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all the authors of the co-authored papers.

****

The applicant (C.I. Williams) is the primary author of all the chapters. Dr. Whitehead initiated the research, and helped with suggestions, insightful discussions, manuscript editing and support throughout the thesis. Dr. B.J. Jean-Claude, Dr. Li Pang and Slaven Suba contributed suggestions and helped edit the manuscripts of which they were co-authors. The execution, interpretation and reporting of the research was primarily performed by the applicant. Thus, this thesis represents the applicant's research.

****

Abstract:

The structures, energies and reactivities of nitrogen containing heterocyclic compounds are studied using computational methods. A method to calculate azole, azine and benzoazole heats of formation using *ab initio* total energies and isolobal reaction schemes is developed, and gives heats of formation to within 0-4 kcal/mol of experiment. Systematic errors in semi-empirical MNDO, AM1 and PM3 heats of formation of azoles, azines and benzoazoles are discussed, and correction terms suggested. Pentazole, pentazine, hexazine, pyrazepine, indazine, purine, azolotriazine and diazoazole heats of formation are then predicted. Isocyanate equilibrium geometries calculated with semi-empirical and *ab initio* methods are compared to experiment.

The cycloaddition reaction of isocyanates with alkenes is modelled using MNDO, AM1 and PM3. Localized bonding in diazoazole ring systems, and flexibility in the diazo group C=N=N bond angle are demonstrated. Thermolytic and photolytic 4-diazo-1,2,3-triazole reaction products are predicted by comparing the computed relative energies of the azolylidene intermediate singlet and triplet states.

Possible mechanisms for diazoazole cycloaddition reactions with ethyne, ynamine and cyanoethyne are discussed and modelled with semi-empirical methods. The regio-specificity of diazoazole cycloaddition reactions with ynamine, and the lack of diazoazole cycloaddition with cyanoethyne, are predicted and correlated with experiment.

Molecular Mechanics (MM+), semi-empirical and *ab initio* methods are used to generate benzotetrazepinone equilibrium geometries. The racemization of chiral benzotetrazepinone is found to occur through a concerted mechanism involving a chiral transition state. Mechanisms for the acid-induced ring opening of benzotetrazepinone (BTZ) are proposed and investigated using PM3. The computational results are compared with UV/VIS spectral evidence for the ring-

opening. The tendency for benzotetrazepinone to ring opening is correlated with the Hammett constants of the aryl-ring substituents using PM3 relative energies. The pK_a of protonated benzotetrazepinone is estimated.

A simple atom-atom potential method is used to model the stability and dynamics of small endohedral gas molecules inside fullerene cages. The results are shown to compare qualitatively with available *ab initio* results.

The problem of electron correlation is briefly addressed, and relationships between Configuration Interaction (CI) and Natural Spin Orbitals (NSOs) are discussed.

Résumé

Les structures, énergies et réactivités d'hétérocycles azotés ont été étudiées par des méthodes informatique. Nous avons développé une méthode pour calculer les chaleurs de formation d'azoles, d'azines et de benzoazoles via les énergies *ab initio* ainsi que pour les schémas de réaction *isolobale;* ce qui donne des chaleurs de formation à l'intérieur de 0-4 kcal/mol des valeurs expérimentales. Les erreurs systématiques des chaleurs de formation obtenues par les méthodes semi-empiriques MNDO, AM1 et PM3 pour les azoles, azines et benzoazoles sont discutées; des termes correctifs sont aussi suggerés. Les chaleurs de formation du pentazole, du pentazine, de l'hexazine, de la pyrazepine, de l'indazine, de la purine, de l'azolotriazine ainsi que du diazoazole ont été prédites.

Les géométries d'isocyanates en position d'équilibre ont été calculées par methods semi-empirique et *ab initio* et furent comparées aux valeurs expérimentales. La réaction de cycloaddition d'isocyanates avec des alcènes a été modélisée en utilisant MNDO, AM1 et PM3. Nous avons demontré la nature localisé des liaisons à l'interieur du système diazoazole ainsi que la flexibilité du groupe diazo (C=N=N). Les produits de réaction de thermolyse ainsi que de photolyse du 4-diazo-1,2,3-triazole ont été predits à l'aide des valeurs comparées des énergies relatives calculées pour les états triplets et singulets de l'intermédiaire azolylidene.

Les mécanismes possibles pour les réactions de cycloaddition des diazoazoles avec l'éthyne, l'ynamine et la cyanoéthyne sont discutés puis modélisés à l'aide de méthodes semi-empiriques. La régiospécificité des réactions de cycloaddition des diazoazoles avec l'ynamine ainsi que l'absence de réaction dans le cas de la cyanoéthyne sont prédites et corrélées avec les résultats expérimentaux.

Nous avons utilisé les méthodes de mécanique moléculaire (MM+), les méthodes semi-empiriques ainsi qu'*ab initio* pour générer les géométries d'équilibre de benzotétrazépinone (BTZ). Nous avons trouvé que la racémisation de

benzotétrazépinone chirales se produit via un mécanisme concerté avec un état de transition chiral. Nous avons étudié les mécanismes d'ouverture des cycles benzotétrazépinone induits par l'acide en utilisant la méthode PM3. Les résultats des calculs sont comparés avec les évidences de spectroscopie UV/VIS en ce qui concerne l'ouverture des cycles. La tendance des tétrazépinones à subir une ouverture de cycle est corrélée avec les constantes de Hammett pour les substituents aryles et les énergies relatives obtenues avec PM3. Nous avons estimé le pK_a de BTZ protonnées.

Une approche basée sur un potentiel atome-atome simple a été utilisée pour modéliser la stabilité et la dynamique de petites molécules endohédriques gazeuses à l'intérieur de cages fullerèniques. Les résultats obtenus se comparent qualitativement avec les résultats *ab initio* disponibles.

Les problèmes de corrélation électronique, d'interactions de configuration (CI) ainsi que de "spin orbitalaire naturel" (NSO) sont discutés.

Acknowledgments

I would like to thank Prof. M. A. Whitehead for his encouragement, guidance and support as a research supervisor.

I would also like to thank Dr. Bertrand J. Jean-Claude for collaboration, and help with the manuscripts.
Dr. Li Pang, for collaboration on the AAP potential chapter (Chapter 2).

I also extend my thanks to all my fellow co-workers in Dr. Whitehead's lab;
Slaven Suba and Dr. Nuria Gallego-Planas for discussions and friendship;
Dr. Fortunato Villamagna, for his suggestions in the initial stages of the research;
Denis Valois, for computing help, and;
Dr. Gérald Villeneuve for translating the abstract

Dr. Mike Schmidt is thanked for copy of the GAMESS software

I would also like to thank all the undergraduate students, graduate students, faculty members and staff members who have made my stay here at McGill enjoyable.

The Natural Science and Engineering Research Council of Canada (NSECR) and Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR) are thanked for funding.

McGill is thanked for various Teaching Assistantships

To my family, my friends

and my wife, Zovig

"Knowledge...prolongs life itself, and enlarges the sphere of existence"

— John Quincy Adams

Table of Contents

Chapter 1: Introduction

Chapter 2

Interaction and Dynamics of Endohedral Gas Molecules in C60 Isomers and C70

Chapter 3

Aromatic Nitrogen Heterocyclic Heats of Formation: A Comparison of Semi-Empirical and *Ab initio* Treatments

Chapter 4

Semi-Empirical Study of Isocyanate Geometries, and β-Lactam Formation Through Alkene-Isocyanate Cycloaddition Reactions

Chapter 5

A Comparison of Semi-Empirical and *Ab Initio* Calculations on Diazoazoles and Azolylidenes

Chapter 6

A Semi-Empirical and *Ab initio* Study of Diazoazole Cycloaddition Reactions with Ethyne, Ynamine and Cyanoethyne

Chapter 7

Equilibrium Geometries and Enantiomeric Interconversions in Tetrazepinone Ring Systems: A Molecular Mechanics, Semi-Empirical and *Ab initio* Treatment

Chapter 9

Electron Correlation and the Natural Orbital Representation: Relationships between the Fractional Orbital Occupancy Numbers, Configuration Interaction Expansion Coefficients and One-Particle Energies:

Chapter 10

Conclusions, Claims to Original Research and Future Work

List of Tables:

List of Figures:

Chapter 1: Introduction:

1.1 Overview of Thesis:

Computational methods used in this thesis describe the structures, reactivities and energies of various pharmacologically interesting chemical compounds. Each chapter is a self-contained research paper, as allowed by the McGill Faculty of Graduate Studies and Research. The requirements for a paper thesis are outlined on page ii; short connecting texts are included between the chapters, and the chapters themselves are modified from their published manuscript form, with short passages which cross-reference between chapters, and emphasize common themes throughout the thesis.

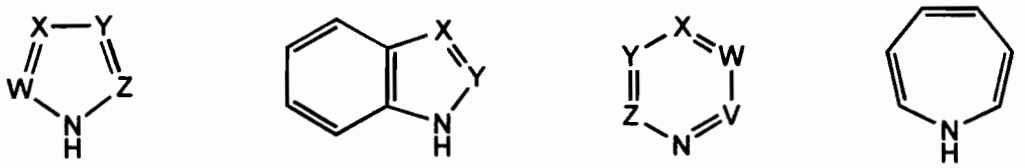
The applicant (C.I. Williams) is the primary author of all the chapters. Dr. Whitehead initiated the research, and helped with suggestions, insightful discussions, manuscript editing and support throughout the thesis. Dr. B.J. Jean-Claude, Dr. Li Pang and Slaven Suba contributed suggestions and helped edit the manuscripts of which they were co-authors. The execution, interpretation and reporting of the research was primarily performed by the applicant. Thus, this thesis represents the applicant's research.

Chapters two, three, five and seven have been published or accepted for publication; chapter eight has been submitted for publication; chapters four, six and nine will be submitted for publication. The chemical and computational details of each study will be discussed where appropriate. A general discussion of the theories and computational methods used in the thesis constitutes the remainder of the introduction.

Chapter two: a simple atom-atom potential method is developed to describe the energies and dynamics of endohedral guest molecules in fullerene cages.

Chapters three to eight: the structures, energies and reactivities of nitrogen heterocyclic compounds are studied. Chapter three presents a method to calculate

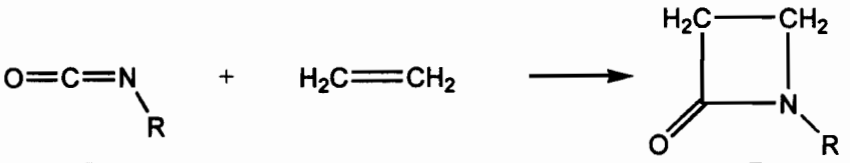
nitrogen heterocyclic heats of formation. Azoles (**1**), benzoazoles (**2**), azines (**3**), and azepine (**4**) are some of the ring systems considered.



1 **2** **3** **4**

V,X,Y,W,Z = C-H or N

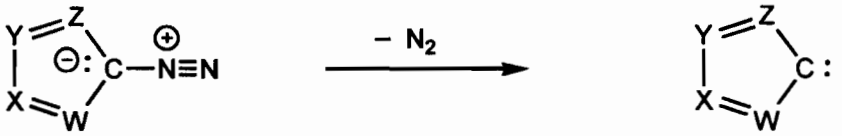
Chapter four is a study of isocyanate (**5**) equilibrium geometries, and isocyanate cycloaddition reactions with ethene (**6**) to form β-lactams (**7**). The structures of the transition states in these cycloaddition reactions are determined computationally, and the computed activation energy barriers are correlated with experimentally observed reactivity.



5 **6** **7**

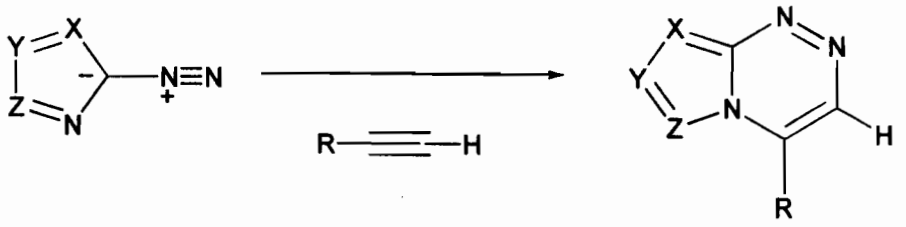
R = -H, $-CH_3$, -Cl, -F, $-C_6H_5$

Chapter five includes the heats of formation, structures and stabilities of diazoazoles (**8**) and azolylidenes (**9**). The results explain experimentally observed azolylidene addition reaction products, and explain the low temperature ESR measurements on azolylidenes.



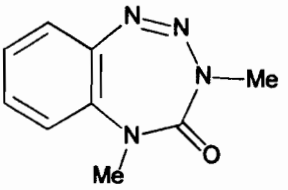
8 X,Y,W,Z = C-H or N **9**

Chapter six: possible mechanisms of diazoazole (**8**) cycloadditions with ethyne, ynamine and cyanoethyne to form azolotriazine (**10**) are computationally modelled. The results explain reactivity and regioselectivity in these reactions.



8 **10**

Chapters seven and eight are studies of the relatively unknown tetrazepinone compounds (**11**). In contrast to the well known molecules in chapters three to six, the stabilities of tetrazepinones, and their decomposition mechanisms are not well understood; almost no physical chemical measurements or mechanistic studies have been made on them. These two chapters advance the knowledge of these compounds, and the information has been useful to the organic chemists and oncologists, who are developing tetrazepinones for use as anti-tumor agents. These chapters exemplify the unique power of computational chemistry at investigating molecular structure, and properties of new compounds.



11

Chapter nine is a purely theoretical quantum chemistry study. Despite the success of computational chemistry, there remains to be developed a unique theoretical and computational approach, which can fully explain and predict all desired chemical properties. As a contribution to this ideal, the well known problem of Configuration Interaction (CI) and the Natural Spin Orbital Representation (NSO) of the total electronic wavefunction are discussed, in

conjunction with the problem of electron correlation in *ab initio* calculations. New original theoretical relationships between the CI expansion coefficients and the Natural Spin Orbital occupation numbers are developed.

1.2 Overview of Computational Methods:

The prediction of chemical properties by computation should facilitate research, by allowing an intimate study of molecular processes which are otherwise difficult, or impossible, to study experimentally. Dirac stated in 1929 that in principle, it is possible to study all of chemistry computationally, because

"the underlying physical laws necessary for the mathematical theory of the whole of chemistry are completely known." [1]

However, Dirac also recognized that "application of these physical laws leads to equations much too complicated to be soluble." [1] Modern computing has greatly changed this situation; however, the main limitation to all computational chemistry is still the number of atoms, the number of electrons, and the number and complexity of the equations to be solved. Consequently, many different approaches to modelling and calculating chemical systems have been developed, all of which are of limited application. Nevertheless, computational methods are increasingly used by chemists to solve structure and reactivity problems, and to predict properties of new molecules. The field of Computational Chemistry has consequently developed, and a number of journals (such as the "Journal of Computational Chemistry", the "Journal of Molecular Structure" the "Journal of Computer-Aided Molecular Design" and the "International Journal of Quantum Chemistry") are now devoted to the development and application of computational methods. Just as molecular modelling kits help chemists to visualize molecular structure, computational methods, such as molecular modelling, help chemists to enhanced their research activity with a variety of

quantitative results for molecular structure, from simple visualizations of complex macromolecules, to advanced electronic structures of small systems.

The computational method chosen to solve a particular chemical problem depends on the properties required, and the complexity of the chemical system. Faced with a particular 'real' chemical problem, the computational chemist judicially chooses which computational approach is appropriate for the study, and assesses the significance of the assumptions and the inherent errors of the method. A combination of computational approaches is often needed to give a real chemical insight into the problem.

The approaches to modelling chemical systems divide into molecular mechanics and quantum mechanical methods (see figure on page 5a). Molecular mechanics calculates molecular energies using empirically fitted model potential energy functions, based on the classical "ball-and-spring" picture of the molecule. They are not suited for calculation of electronic properties, and their accuracy depends on extensive parameterization. Their advantage over quantum-mechanical methods is their speed of computing. Molecular mechanics methods are used to explore conformational and non-bonded potential energy surfaces of systems for which they are well parameterized. Molecular mechanics methods can be applied to biologically interesting molecules, ensembles, and other systems too large for a complete quantum mechanically based study.

Quantum mechanics uses the Schrödinger wave equation to describe atomic and molecular electronic structure, and to calculate total energies, heats of formation, spectral properties, vibrational quantities, various electronic properties, and equilibrium and transition state molecular geometries. Quantum mechanical methods are too time consuming and expensive for large systems, but they can give detailed information about structure, reactivity, and reaction coordinates.

The molecular mechanical and quantum mechanical methods used in this thesis are in the HyperChem[2] and GAMESS[3] programs. An understanding of the theories employed in these packages is necessary, therefore the characteristics

Choice of Computational Method

Methods of Computation

Quantum Mechanics

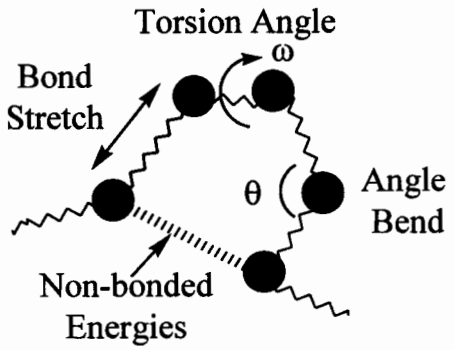
$H\Psi = E\Psi$

-computation of electronic structure

- *ab initio* and semi-empirical (MNDO, AM1, and PM3)

- computationally intensive (SCF procedure with many integrals)

Molecular Mechanics



a) MM+ force field used to generate input structures for QM calculations;

b) AAP potential used to calculate non-bonded fullerene/guest complexing energies

QM methods are the most suitable methods for calculating structure and reactivity of nitrogen heterocycles;

which define and differentiate the various molecular-mechanics and quantum mechanics methods are described.

1.3: Molecular Mechanics:

Molecular Mechanics (or Molecular modelling) methods use experimentally and theoretically parameterized classical harmonic oscillator type equations to define a force field, which describes the potential energy of a molecule as a function of its conformation. Molecular mechanical energies quantitatively measure the effect of bond stretch, bond angle distortion, and torsional angle rotation on the total energy of the molecule. Empirically and theoretically parameterized non-bonded atom-atom potentials are also included in these methods; consequently, they are suited to model collections of molecules, as in solute/solvent systems and interfacial monolayers.[4,5]

The recent foundations of molecular mechanics were laid by Allinger and co-workers,[6-10] with their MM family of molecular mechanics programs. Other models, such as AMBER[11a] and OPLS[11b,c] have been developed to deal with proteins and other biopolymers. The simplest form of the total molecular mechanics energy, common to all methods, divides the total molecular strain energy (E_{tot}) into bond stretch (E_{bonds}), bond angle bend (E_{angles}), torsion angle rotation ($E_{torsions}$) and non-bonded ($E_{non\text{-}bonded}$) energies.

(1) $$E_{tot} = E_{bonds} + E_{angles} + E_{torsions} + E_{non\text{-}bonded}$$

In (1), the bond stretch terms are commonly approximated by simple 1-D harmonic oscillator equations

(2) $$E_{bonds} = \sum_{bonds} k_b (l - l_o)^2$$

where k_b is the spring force constant for the particular bond, l is the actual length of bond, and l_o is the equilibrium or unstretched bond length. The k_b and l_o parameters can be optimized from experimental and theoretical results. When all the bond lengths are at their equilibrium lengths l_o, then $E_{bonds} = 0$, regardless of

the functional form of E_{bonds}. Cubic bond stretch terms are added to generate more refined force fields, as in MM2, and the MM+ force field in the HyperChem program.[2]

The bond angle distortion energies are described by harmonic-oscillator type equations similar in form to the bond stretch terms

$$E_{angles} = \sum_{angles} K_\theta(\theta - \theta_o)^2. \quad (3)$$

Here, K_θ is the bond angle bend force constant, θ is the bond angle and θ_o is the equilibrium bond angle. Other bond angle energy equations include cubic, quartic and other terms to describe the bond angle distortion energy. Bond-stretch/angle-bend cross terms are also employed in the MM2 family of molecular modelling programs.

Torsion angle (or dihedral angle) energies reflect the rotational symmetry and the energy of rotation about individual bonds. The total torsion angle energy is given as the sum over all the individual 1-4 torsion angle interactions, ω.

$$E_{tors} = \sum_{\omega} E_\omega \quad (4)$$

Each E_ω term may be expanded as a Fourier series in the torsion angle, ω,

$$E_\omega = \sum_{n} V_n(1 + s\cos(n\omega)), \quad (5)$$

where V_n is the energy of the n-fold barrier, and n represents the periodicity of the rotation (i.e. n=2 for ethene, n=3 for ethane). In ethene, the eclipsed form is the minimum, and s = -1; the staggered form is the minimum in ethane, and s = +1. Rotations of increasing complexity and periodicity are adequately described by inclusion of more terms in the Fourier expansion. The bond stretch, bond angle bend and torsion angle rotation terms constitute the molecular mechanics bonded interaction terms. Bonded interaction evaluation depends upon the connectivity[2] of the molecule, which is the arrangement of formal bonds connecting the atoms in the molecule.

In contrast, non-bonded energy terms depend upon interatomic distances, and are calculated between atoms whose connectivity is greater than 1-4. The

non-bonded terms in the molecular mechanics energy include van der Waals terms (E_{VDW}), electrostatic interaction terms (E_{CHG}), dipole interaction terms (E_{DIP}) and the H-bonded terms (E_{HH}). Each term has its own functional form. Van der Waals[12a] interactions are often described by a Lennard-Jones[12b] function,

$$E_{VDW} = \sum_{B>A}\left[\frac{a_{AB}}{R_{AB}^{12}} - \frac{b_{AB}}{R_{AB}^{6}}\right]. \quad (6)$$

Here, R_{AB} is the non-bonded distance between atoms A and B, and a_{AB} and b_{AB} are van der Waals parameters for the interaction of atoms A and B. Other functional forms, such as

$$E_{AB} = -a_{AB}R_{AB}^{-6} + b_{AB}\exp(-c_{AB}R_{AB}) \quad (7)$$

used in chapter two to describe van der Waals interactions between host and guest molecules in endohedral fullerene complexes, can also simulate van der Waals interactions. The electrostatic energy is usually simulated with a basic electrostatic interaction

$$E_{CHG} = \sum \frac{q_A q_B}{\varepsilon_o R_{AB}}, \quad (8)$$

where q_A and q_B are point charge approximations for ions A and B, R_{AB} is the A-B internuclear distance, and ε_o is the effective dielectric constant of the medium. The interaction between two dipoles, μ_i and μ_j, is given by

$$E_{DIP} = \frac{\mu_i \mu_j}{\varepsilon_o r_{ij}^3}((\cos(x) - 3\cos(\alpha_i)\cos(\alpha_j)); \quad (8a)$$

Here, r_{ij} is the distance between the two dipoles, x is the angle between the two dipoles, α_i and α_j are the angles between each dipole and the vector connecting the two dipoles, and ε_o is the effective dielectric constant.

Since most of the thesis (Chapters 3-8) focuses on the prediction of electronic structure dependent properties, molecular mechanics methods are unsuitable, except for generating input structures. Throughout this thesis, the MM+ force field[2] is used for molecular mechanics calculations. MM+ is an MM

type force field, and contains stretch-bend, torsion-bend and other energy cross-terms more complex than the general ones stated here. For details of the MM+ method, see the HyperChem documentation.[2] MM+ has an advantage over the United Atom force fields such as AMBER and OPLS, because it treats every atom explicitly, with a complex potential energy function. Consequently, it is ideally suited to generate geometries of the novel ground state structures studied in the forthcoming chapters. A detailed study in chapter 7 justifies this approach, by showing that quantum mechanical optimizations of MM+ input structures produce structures identical to those obtained by quantum mechanical optimization of X-ray input structures. Because the molecules considered in the thesis are all relatively small (<40 atoms), using such a complicated force field still has a minimal computational cost. Since the MM+ method is used primarily to generate input for quantum calculations, possible improvements to the MM+ method were not considered.

1.4 Electronic Structure and Quantum Mechanical Methods

All quantum mechanical methods solve the Schrödinger wave equation, to give the total molecular energy and wavefunction of the system. The total wavefunction, in principle, allows calculation of all properties related to molecular structure, such as spectra, dipole moments, polarizability and thermodynamic properties. A comprehensive review of the quantum mechanics relevant to chemistry is beyond the scope of this thesis; the reader is referred to a few excellent textbooks for thorough treatments.[13-22]

There are many quantum mechanical approaches to molecular structure, from simple Hückel MO based theories to extensive *ab initio* and density-functional methods.[23-26] These methods differ in their level of approximation and computational time and cost, but they all employ the Born-Oppenheimer approximation, which 'freezes' the nuclear positions during the electronic structure calculation. The nuclei of a molecule are a set of external point charges,

which interact with the electrons. This separates the total electronic and nuclear wavefunction of the molecule, Ω, into the wavefunction of the nuclei, and the wavefunction of electronic part of the system only, Ψ. In the Born-Oppenheimer approximation, the time-independent electronic Schrödinger wave equation is

$$H\Psi = E\Psi, \tag{9}$$

where E is the total electronic energy and Ψ is the total electronic wavefunction. H is the Hamiltonian which describes the various mathematical operators of the system. For an N electron system, in the field of stationary nuclei (A,B), the non-relativistic Hamiltonian which describes the energy of the system is (in atomic units, a.u. ; $h/2\pi = 1$, a_0 (Bohr radius) = 1, charge on electron, e =1).

$$H = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 \quad - \sum_{i=1}^{N}\sum_{A}\frac{Z_A}{R_{iA}} \quad + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \quad + \sum_{A}\sum_{B>A}\frac{Z_A Z_B}{R_{AB}}. \tag{10}$$

The first two terms in the Hamiltonian (equation 10) describe the electron kinetic and electron-nuclear attraction energies respectively. The third term is the electron-electron interaction operator. The summations *i* and *j* are over all the electrons in the system. The summations over A and B in the last term are over all nuclei in the system, and represent the nuclear-nuclear point charge repulsion energies. The nuclear-nuclear repulsion terms are non-quantum mechanical in nature, and are usually added to the electronic energy after solving the Schrödinger wave equation for the electrons.

1.4.1. Hartree-Fock Wavefunctions[13-16]

The Schrödinger equation cannot be solved exactly for systems with more than one electron; therefore, the total electronic wavefunction Ψ must be approximated. A common approximation for an N electron system is the Hartree-Fock (HF) wavefunction, which consists of an N x N Slater determinant of

orthonormal one-electron spin wavefunctions χ_i ;[16] ($\int \chi_i \chi_j \, dx = \delta_{ij}$; $\delta_{ij} = 1$ if i = j, $\delta_{ij} = 0$ if $i \neq j$).

$$(11) \qquad \Psi(x_1, x_2 \ldots x_N) = \sqrt{\frac{1}{N!}} \begin{vmatrix} \chi_1(x_1) & \cdots \chi_i(x_1) & \cdots & \chi_N(x_1) \\ \vdots & & & \\ \vdots & & \ddots & \\ \chi_1(x_N) & \cdots \chi_i(x_N) & \cdots & \chi_N(x_N) \end{vmatrix}.$$

Here, x_1 is the space and spin coordinate of electron 1, and $(1/N!)^{1/2}$ is a normalization constant, which ensures that

$$(12) \qquad \int \Psi^*(x_1 \ldots x_N) \Psi(x_1 \ldots x_N) dx_1 \ldots dx_N = 1$$

The dx_i in (12) refers to the space (dr_i) and spin ($d\omega_i$) coordinates of electron 'i'. The determinantal nature of the wavefunction is necessary because of the fermionic nature of electrons. The general HF wavefunction can be simplified to[16]

$$(13) \qquad \Psi^{HF}(x_1 \ldots x_N) = |\chi_1(x_1)\chi_2(x_2) \ldots \chi_i(x_i) \ldots \chi_{N-1}(x_{N-1})\chi_N(x_N)>,$$

which includes the normalization constant, and only shows the diagonal elements of the Slater determinant in (11). An N electron closed-shell spin-restricted HF wavefunction (RHF),[27] used for systems with no net spin (multiplicity = 1), consists of N/2 doubly occupied spatial orbitals ψ_i. Electrons with different spin are distinguished with $\alpha(\omega)$ and $\beta(\omega)$ spin functions. The spin orbital $\chi_i(x_i)$ in (13) is replaced by the product of a spatial orbital and a spin function; $\chi_i(x_i) = \psi_i(r_i)\alpha(\omega_i)$. Thus, the spatial orbitals, ψ_i, in $\psi_i(r_i)\alpha(\omega_i)$ and $\psi_i(r_i)\beta(\omega_i)$ are spatially and energetically equivalent. The RHF wavefunction, Ψ^{RHF} is

$$(14) \quad \Psi^{RHF} = |\psi_1(r_1)\alpha(\omega_1)\psi_1(r_2)\beta(\omega_2) \cdots \psi_{N/2}(r_{N-1})\alpha(\omega_{N-1})\psi_{N/2}(r_N)\beta(\omega_N)>$$

The Pauli exclusion principle is satisfied by the following spin integrals;

$$(15) \qquad \int \alpha(\omega)\alpha(\omega)d\omega = \int \beta(\omega)\beta(\omega)d\omega = 1; \int \alpha(\omega)\beta(\omega)d\omega = \int \beta(\omega)\alpha(\omega)d\omega = 0.$$

Systems with net spin (multiplicities greater than S=1) can be treated with a spin-unrestricted Hartree-Fock wavefunction (UHF),[28] which allows the spatial ψ_1^{α} and ψ_1^{β} orbitals to differ energetically and spatially.[16]

$$(16) \quad \Psi^{UHF} = |\psi_1^{\alpha}(r_1)\alpha(\omega_1)\psi_1^{\beta}(r_2)\beta(\omega_2) \cdots \psi_{N\alpha}^{\alpha}(r_{N\alpha})\alpha(\omega_{N\alpha})\psi_{N\beta}^{\beta}(r_{N\beta})\beta(\omega_{N\beta})>$$

In the UHF wavefunction above, N^{α} is the number of α spin electrons, and N^{β} is the number of β spin electrons. The number of α and β spin electrons need not be equal in a UHF calculation. UHF calculations can be used for systems with an odd number of electrons.

Insertion of either type of HF wavefunction into the Schrödinger wave equation (9) results in a set of one-electron Schrödinger wave equations, or "(Hartree-) Fock" equations. In the closed-shell RHF case, the one-electron Fock equation of electron 1 in the i^{th} orbital, with alpha spin, is

$$F(1)\,\psi_i(r_1)\alpha(\omega_1) = \varepsilon_i \psi_i(r_1)\alpha(\omega_1)\,. \tag{17}$$

Here, ε_i is the eigenvalue of the one-electron wavefunction ψ_i, and F(1) is the one-electron 'Fock' operator. Upon integration over the α and β spin coordinates, the closed-shell RHF Fock operator has the form

$$F(1) = H_1^{CORE} + \sum_{j}^{N/2}[2J_j(1) - K_j(1)]. \tag{18}$$

The first term in the Fock operator, H^{CORE}, consists of the one-electron kinetic and nuclear attraction operators, and is referred to as the "one-electron-core" operator.

$$H_1^{CORE} = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} \tag{18a}$$

The $J_j(1)$ term in the Fock operator in (18) is the two-electron coulomb repulsion operator, given by[13,16]

$$J_j(1)\psi_i(r_1) = \int \frac{\psi_j^*(r_2)\psi_j(r_2)}{r_{12}} dr_2\, \psi_i(r_1) \quad , r_{12} = |r_1 - r_2| \tag{18b}$$

The coulomb operator $J_j(1)$ represents the coulombic repulsion between electron 1 in the i^{th} orbital and the smeared out charge density ($|\psi_j(r_2)|^2$) of electron 2 in the j^{th} orbital. The factor of two in front of $J_j(1)$ in (18) arises because there are two electrons in each closed-shell RHF spatial orbital. The $K_j(1)$ term in (18) is the exchange operator, which is defined as[13,16]

$$K_j(1)\psi_i(r_1) = \int \frac{\psi_j^*(r_2)\psi_i(r_2)}{r_{12}} dr_2\, \psi_j(r_1)\,. \tag{19}$$

Unlike the coulomb operator, the exchange operator has no simple classical interpretation. Exchange is a net attractive interaction which arises from the

indistinguishabilty of electrons. Exchange occurs between electrons of like spin only; therefore, there is no factor of two in front of the K_j (1) term in (18), because there is no exchange between electrons of unlike spin. Exchange between electrons of unlike spin cancels out upon integration over spin. Spatial coordinates (dr_2) are used in (18b) and (19), because spin has been integrated out. Using the Fock operator, the RHF closed-shell one-electron energies ε_i in (17) are given by

$$(20)\quad \varepsilon_i = \int \psi_i^*(r_1) F(1) \psi_i(r_1) dr_1 = \int \psi_i^*(r_1) H_1^{CORE} \psi_i(r_1) dr_1 + \sum_j^{N/2} (2J_{ij} - K_{ij}),$$

where the coulomb integral, J_{ij}, is[13]

$$(21)\quad J_{ij} = \int \psi_i^*(r_1) J_j(1) \psi_i(r_1) dr_1 = \iint \frac{\psi_i^*(r_1)\psi_i(r_1)\psi_j^*(r_2)\psi_j(r_2)}{r_{12}} dr_1 dr_2$$

and the exchange integral, K_{ij}, is

$$(22)\quad K_{ij} = \int \psi_i^*(r_1) K_j(1) \psi_i(r_1) dr_1 = \iint \frac{\psi_i^*(r_1)\psi_j(r_1)\psi_j^*(r_2)\psi_i(r_2)}{r_{12}} dr_1 dr_2 .$$

The total electronic energy E_{El} in the closed-shell RHF case is[13,16]

$$(23)\quad E_{El} = 2\sum_i^{N/2} \varepsilon_i - \sum_i^{N/2}\sum_j^{N/2} (2J_{ij} - K_{ij}) .$$

Consequently, the total electronic energy is NOT the sum of the individual electron energies ε_i, because this would count all electron-electron interactions twice.

1.4.2. The MO-LCAO Approximation.[27]

The molecular orbitals (MOs) ψ_i,[27] are approximated as a linear combination of 'M' atomic orbitals (LCAO) or other suitable basis functions ϕ_μ, referred to as the basis set (henceforth denoted with ν,μ,σ,λ subscripts).

$$(24)\quad \psi_i(r_1) = \sum_\mu^M c_\mu^i \phi_\mu(r_1) \qquad \mu = 1...M$$

In (24), c_μ^i is the coefficient of the μ^{th} basis orbital in the i^{th} molecular orbital. Insertion of the LCAO expansion (24) into the Hartree-Fock equation (17) gives the following equation.[13,16]

$$(25) \qquad \sum_{\mu}^{M} c_\mu^i F(1)\phi_\mu(r_1)\alpha(\omega_1) \;=\; \varepsilon_i \sum_{\mu}^{M} c_\mu^i \phi_\mu(r_1)\alpha(\omega_1)$$

Multiplying by $\phi_\nu^*(r_1)$ on the left hand side of (25) and integrating over spin, gives (in the closed-shell RHF case), [16] the following integrals over dr_1

$$(26) \quad \sum_{\mu} c_\mu^i \int \phi_\nu^*(r_1)\left(H_1^{CORE} + \sum_{j}^{N/2}[2J_j(1) - K_j(1)]\right)\phi_\mu(r_1)dr_1 = \varepsilon_i \sum_{\mu} c_\mu^i \int \phi_\nu^*(r_1)\phi_\mu(r_1)dr_1 \,.$$

In equation (26), the expression in equation (18) has been used for the Fock operator, F(1). The core-Hamiltonian terms in equation (26) lead to "one-electron core" integrals of the following type.

$$(27) \qquad H_{\nu\mu}^{CORE} = \int \phi_\nu^*(r_1)(H_1^{CORE})\phi_\mu(r_1)dr_1 \;=\; (\nu|H_1^{CORE}|\mu)$$

The $(\nu|H_1^{CORE}|\mu)$ integral is the one-electron kinetic and nuclear-attraction energy.

Substitution of LCAO expansions $\psi_j(r_2) = \sum_{\lambda}^{M} c_\lambda^j \phi_\lambda(r_2)$ and $\psi_j^*(r_2) = \sum_{\sigma}^{M} (c_\sigma^j)^* \phi_\sigma^*(r_2)$

into the coulomb operator (equation 18b), and subsequent insertion into (26), gives coulomb integrals with the following form;

$$(28) \qquad \begin{aligned} \int \phi_\nu^*(r_1)J_j(1)\phi_\mu(r_1)dr_1 &= \sum_{\lambda}\sum_{\sigma} c_\lambda^j (c_\sigma^j)^* \iint \frac{\phi_\nu^*(r_1)\phi_\mu(r_1)\phi_\sigma^*(r_2)\phi_\lambda(r_2)}{r_{12}} dr_1 dr_2 \\ &= \sum_{\lambda}\sum_{\sigma} c_\lambda^j (c_\sigma^j)^* (\nu\mu|\sigma\lambda) \end{aligned} \,.$$

An analogous expansion of the exchange operator, (19), and subsequent insertion into (26), gives exchange integrals, which have the form[16]

$$(29) \qquad \begin{aligned} \int \phi_\nu^*(r_1)K_j(1)\phi_\mu(r_1)dr_1 &= \sum_{\lambda}\sum_{\sigma} c_\lambda^j (c_\sigma^j)^* \iint \frac{\phi_\nu^*(r_1)\phi_\lambda(r_1)\phi_\sigma^*(r_2)\phi_\mu(r_2)}{r_{12}} dr_1 dr_2 \\ &= \sum_{\lambda}\sum_{\sigma} c_\lambda^j (c_\sigma^j)^* (\nu\lambda|\sigma\mu) \end{aligned} \,.$$

The $(\nu\mu|\sigma\lambda)$ notation is used to simplify the two-electron coulomb and exchange integrals in (28) and (29). This follows the nomenclature used by Szabo and Ostlund,[16] where curved bracket (|) notation is used for spatial orbital (ϕ)

integrals. This is different from Dirac notation $\langle \nu\mu|\sigma\lambda \rangle$, which is reserved for spin orbital integrals.[16] Using equations (27)-(29), the Fock matrix **F** can be constructed. In the RHF case, elements of the Fock matrix, $F_{\nu\mu}$, are

$$F_{\nu\mu} = H_{\nu\mu}^{CORE} + \sum_{\lambda}\sum_{\sigma}\sum_{j}^{N/2} c_{\lambda}^{j}(c_{\sigma}^{j})^{*}[\,2(\nu\mu|\sigma\lambda) - (\nu\lambda|\sigma\mu)], \tag{30}$$

Equation (30) can be simplified further by introducing the density matrix **P**, whose elements are[16]

$$P_{\lambda\sigma} = 2\sum_{j}^{N/2} c_{\lambda}^{j}(c_{\sigma}^{j})^{*}\,; \tag{31}$$

The summation in (31) is over all occupied orbitals. Using the density matrix elements, equation (30) then becomes

$$F_{\nu\mu} = H_{\nu\mu}^{CORE} + \sum_{\lambda}\sum_{\sigma} P_{\lambda\sigma}[(\nu\mu|\sigma\lambda) - 1/2\,(\nu\lambda|\sigma\mu)]. \tag{31a}$$

By defining the overlap matrix **S**, whose elements are

$$S_{\mu\nu} = \int \phi_{\mu}^{*}(r_1)\phi_{\nu}(r_1)dr_1, \tag{32}$$

and expressing the coefficients (c_{μ}^{i}) of the i^{th} molecular orbital as a column vector $\mathbf{c}_i$, the one-electron wave equation in (26) can be converted to a matrix equation,[16]

$$\mathbf{F}\mathbf{c}_i = \varepsilon_i \mathbf{S}\mathbf{c}_i, \quad i = 1,2,3,.....M \tag{33}$$

where M is the number of basis functions. The one-electron matrix wave equations in (33) combine to give a single matrix equation for the entire electronic system, when the coefficient vectors $\mathbf{c}_i$ are grouped into a coefficient matrix **C**, and the eigenvalues ε_i into an eigenvalue matrix ε. The matrix form of the Hartree-Fock equation is often referred to as the *Hartree-Fock-Roothan* equation.[27] The Fock matrix **F** depends upon the density matrix **P**, which in turn depends upon the coefficient matrix **C**, so the Fock matrix in equation (34) is written as **F(C)** to emphasize its dependence on the coefficient matrix **C**.

$$\mathbf{F(C)C} = \mathbf{SC}\varepsilon \tag{34}$$

Insertion of the UHF wavefunction into the Hartree-Fock wave equation (17) results in the unrestricted Hartree-Fock-Roothaan, or Pople-Nesbet[28] matrix equations.

(35) $$\mathbf{F}^{\alpha}(\mathbf{C}^{\alpha}, \mathbf{C}^{\beta})\,\mathbf{C}^{\alpha} = \mathbf{S}\mathbf{C}^{\alpha}\,\varepsilon^{\alpha}\,; \qquad \mathbf{F}^{\beta}(\mathbf{C}^{\alpha}, \mathbf{C}^{\beta})\,\mathbf{C}^{\beta} = \mathbf{S}\mathbf{C}^{\beta}\,\varepsilon^{\beta}\,.$$

The two matrix equations in (35) couple, because both the $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ Fock matrices are dependent upon the $\mathbf{C}^{\alpha}$ and $\mathbf{C}^{\beta}$ coefficient matrices. The different $\mathbf{C}^{\alpha}$and $\mathbf{C}^{\beta}$ coefficients lead to different alpha and beta spin density matrices $\mathbf{P}^{\alpha}$ and $\mathbf{P}^{\beta}$, with elements[16]

(36) $$P^{\alpha}_{\lambda\sigma} = \sum_{j}^{N^{\alpha}} c^{j\alpha}_{\lambda}\,(c^{\alpha}_{\sigma j})^{*}\,; \qquad P^{\beta}_{\lambda\sigma} = \sum_{j}^{N^{\beta}} c^{j\beta}_{\lambda}\,(c^{j\beta}_{\sigma})^{*}\,; \qquad \mathbf{P}^{T} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}.$$

Here $\mathbf{P}^{\alpha}$ and $\mathbf{P}^{\beta}$ are the α and β spin-density matrices, $\mathbf{P}^{T}$ is the total density matrix, and N^{α} and N^{β} are the number of α and β spin electrons respectively. The elements of the $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ Fock matrices are[16]

(37) $$F^{\alpha}_{\nu\mu} = H^{CORE}_{\nu\mu} + \sum_{\lambda\sigma}[\,P^{T}_{\lambda\sigma}\,(\nu\mu|\sigma\lambda) - P^{\alpha}_{\lambda\sigma}(\nu\lambda|\sigma\mu)]$$

$$F^{\beta}_{\nu\mu} = H^{CORE}_{\nu\mu} + \sum_{\lambda\sigma}[\,P^{T}_{\lambda\sigma}\,(\nu\mu|\sigma\lambda) - P^{\beta}_{\lambda\sigma}(\nu\lambda|\sigma\mu)]$$

As with the spin-restricted HF case, the spin-unrestricted matrix equations in (35) must be solved iteratively, since both $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ depend upon the coefficient matrices $\mathbf{C}^{\alpha}$ and $\mathbf{C}^{\beta}$.

1.4.3. SCF Procedure[13-16]

Both equations (34) and (35) must be solved by iteration to self-consistency because **F(C)** depends upon **C**. A typical procedure to accomplish this is to find a transform matrix **X**, such that $\mathbf{X}^{\dagger}\mathbf{S}\mathbf{X} = \mathbf{1}$, and $\mathbf{X}\mathbf{X}^{-1} = \mathbf{1}$, and insert it into equation (34) to give

(38) $$\mathbf{X}^{\dagger}\,\mathbf{F(C)}\mathbf{X}\mathbf{X}^{-1}\,\mathbf{C} = \mathbf{X}^{\dagger}\mathbf{S}\mathbf{X}\mathbf{X}^{-1}\mathbf{C}\varepsilon$$

(39) $$(\,\mathbf{X}^{\dagger}\,\mathbf{F(C)}\mathbf{X}\,)(\mathbf{X}^{-1}\,\mathbf{C}) = \mathbf{X}^{-1}\,\mathbf{C}\varepsilon$$

Defining $\mathbf{F'(C)} = \mathbf{X}^{\dagger}\,\mathbf{F(C)}\mathbf{X}$ and $\mathbf{X}^{-1}\,\mathbf{C} = \mathbf{C'}$ and inserting into (39) gives

(40) $$\mathbf{F'(C)}\mathbf{C'} = \mathbf{C'}\varepsilon\,.$$

Equation (40) can be solved for $\mathbf{C'}$ by diagonalizing $\mathbf{F'(C)}$. The $\mathbf{C'}$ matrix is then used to compute a new **C** matrix (by the relation $\mathbf{X}\mathbf{C'} = \mathbf{C}$), which is then used to form a new Fock matrix **F(C)**. The iterative procedure is repeated until the difference between the old and new Fock matrices (**F(C)** and **F'(C)**) is less than a

preset convergence criteria. Different approaches to iterative solution of the Hartree-Fock-Roothaan equations exist,[16] but essentially they all follow similar principles, and are referred to as 'self-consistent-field'[16] (SCF) calculations.

Solution of the restricted or unrestricted Hartree-Fock equations by SCF procedures can be computationally expensive. The number of integrals required to compute an N electron Fock matrix varies as N^4 , so the number of integrals can reach millions for even small molecules. The computational cost of Hartree-Fock calculations has given rise to different approaches to their solution, such as the *ab initio* and semi-empirical methods.

1.5: Semi-Empirical Methods:

Semi-empirical methods depend on using experimental and theoretical data to parameterize and approximate certain integrals in the Fock matrix (34). This approach reduces the computational time and cost of a Hartree-Fock calculation, while maintaining chemical accuracy. The two-electron integrals are the most numerous and costly to calculate in any Hartree-Fock type electronic structure calculation. All semi-empirical methods attempt to accelerate HF calculations, by neglect or modification of these integrals. Semi-empirical methods are defined by the different approximations and parameterizations used to simplify the Fock matrix. An overview of semi-empirical approaches is given in various books and reviews.[29-30]

1.6 The MNDO, AM1 and PM3 Methods

In this thesis, the theoretically similar MNDO, AM1 and PM3 semi-empirical methods are used, because they have previously given accurate results for many chemical properties. The MNDO method (Modified Neglect of Differential Overlap) is the oldest method, and was introduced by Dewar et. al. in 1977.[31-34] MNDO was developed as an extension to Dewar's previous MINDO/3 method.[35-36] The AM1 (Austin Model 1) method, a modification of the MNDO

method by Dewar and co-workers,[37-38] corrected MNDO deficiencies. The PM3 (Parameterization Model 3) method, introduced by Stewart[30,39] in 1989, re-parameterized the MNDO method, using modified nuclear repulsion terms similar to those in AM1.

The MNDO, AM1 and PM3 methods are all NDDO (Neglect of Differential Diatomic Overlap)[40a] type theories, because they simplify the Fock matrix by systematic neglect of diatomic differential overlap integrals, $(\nu\lambda|\sigma\mu)$. For example, all the three- and four-centered two-electron integrals are not calculated.

The heat of formation of a molecule (ΔH_f) is approximated in MNDO, AM1 and PM3, by[30,32]

$$\Delta H_f = E_{El} + E_{Nuc} - \sum_A E_{El}^A + \sum_A \Delta H_f^A \quad . \tag{41}$$

E_{El} is the calculated molecular electronic energy, and E_{El}^A is the calculated electronic energy of the isolated atom A. E_{Nuc} is the calculated nuclear-nuclear repulsion energy. ΔH_f^A is the experimental heat of formation of atom A.

1.6.1. MNDO, AM1 and PM3 Electronic Energies

A Slater orbital basis set is used in MNDO, AM1 and PM3 calculations. Slater-type orbitals (STOs) differ from hydrogenic orbitals in radial form only; STO's have the general form[13]

$$\phi_{nml}^{STO} = R(r,n,\zeta)Y_{lm}(\theta,\Phi), \tag{42}$$

where $Y_{lm}(\theta,\Phi)$ are identical to the hydrogenic orbital spherical harmonics. The radial component of STOs, $R(r,n,\zeta)$, have the general form[13,30]

$$R(r,n,\zeta) = (2\zeta / a_o)^{n+\frac{1}{2}}[(2n)!]^{-\frac{1}{2}} r^{n-1} \exp(-\zeta r / a_o) . \tag{43}$$

where r is the distance of the electron from the nucleus, n is a parameter related to the principle quantum number, a_o is the Bohr radius, and ζ is the orbital exponent.

The most theoretically difficult and computationally expensive term in equation (41) is the electronic energy of the molecule. Complete solution of (41) by semi-empirical approaches depends upon using a set of parameters to construct elements of the Fock matrix. The Fock matrix elements are divided into one-center and two-center terms. The one-center terms are further divided into one-orbital and two-orbital terms. The diagonal elements of the Fock matrix are <u>one-center one-orbital</u> terms; following Dewar,[32] the one-center one-orbital diagonal element of the closed-shell Fock matrix, for a basis function μ on center A, is approximated as[32]

$$(44)\quad F_{\mu\mu} = H_{\mu\mu}^{CORE} + \sum_{\nu}{}^{A} P_{\nu\nu}[(\mu\mu|\nu\nu) - \frac{1}{2}(\mu\nu|\mu\nu)] + \sum_{B\neq A}\sum_{\lambda}{}^{B}\sum_{\sigma}{}^{B} P_{\lambda\sigma}(\mu\mu|\lambda\sigma).$$

Here $\sum_{\lambda}{}^{B}$ and $\sum_{\sigma}{}^{B}$ indicate summations over all basis functions (λ,σ) on center B, $\sum_{B\neq A}$ indicates a summation over all centers except A, and $\sum_{\nu}{}^{A}$ indicates a summation over all basis functions (ν) on center A. The notation used here is similar to that of Pilar,[30c] and modifies Dewar's notation[32] to be clearer. The <u>one-center two-orbital</u> terms, $F_{\nu\mu}$, are off-diagonal Fock matrix elements that occur between different ϕ_μ and ϕ_ν orbitals on the same nucleus, center A. In the RHF case they are approximated as[32]

$$(45)\quad F_{\mu\nu} = H_{\mu\nu}^{CORE} + \frac{1}{2}P_{\mu\nu}[3(\mu\nu|\mu\nu) - (\mu\mu|\nu\nu)] + \sum_{B\neq A}\sum_{\lambda}{}^{B}\sum_{\sigma}{}^{B} P_{\lambda\sigma}(\mu\nu|\lambda\sigma); \qquad \mu \text{ on A},\quad \nu \text{ on A}, \nu \neq \mu$$

The two-center terms are all off-diagonal terms; in the restricted Fock matrix, they are approximated as[32]

$$(46)\quad F_{\mu\lambda} = H_{\mu\lambda}^{CORE} - \frac{1}{2}\sum_{\nu}{}^{A}\sum_{\sigma}{}^{B} P_{\nu\sigma}(\mu\nu|\lambda\sigma); \qquad \mu,\nu \text{ on A};\ \lambda,\sigma \text{ on B};\quad A\neq B$$

Evaluation of each integral in equations (44)-(46) is now outlined.

<u>One-Electron Integrals:</u> The <u>one-center</u> one-orbital one-electron integral, $H_{\mu\mu}^{CORE}$ in (44), is given by[30,32]

$$(47a)\qquad H_{\mu\mu}^{CORE} = U_{\mu\mu}^{A} + \sum_{B\neq A} V_{AB}^{\mu\mu};$$

The $U^A_{\mu\mu}$ parameters are core-electron attraction plus kinetic energy integrals. These are computed for each element from experimental data, as outlined in reference 41. The $V^{\mu\mu}_{AB}$ term represents the attraction of the $(\mu^A\mu^A)$ electron density on center A to the core of center B. The expression for $V^{\mu\mu}_{AB}$ is a generalization of the expression used in the MINDO/3 method,[35,36] and has the form[32]

$$V^{\mu\mu}_{AB} = -Z_B\,(\mu^A\,\mu^A \mid s^B\,s^B). \tag{47b}$$

Here, Z_B is the core charge of nucleus B (nuclear point charge minus inner shell electron charges). Thus, the atomic core of center B is simulated as an "ss" orbital valence-shell charge distribution. The core charge distribution and the "ss" charge distribution both have no multipole moments higher than the monopole.[32] The $(\mu\mu|\ ss)$ term is identical to a two-electron repulsion integral, whose form will be discussed in the next section.

Orbitals on one center are orthogonal; therefore $U_{\nu\mu} = 0$ for ν,μ on A, $\nu \neq \mu$. The <u>one-center two-orbital</u> one-electron $H^{CORE}_{\nu\mu}$ term is given by[32]

$$H^{CORE}_{\nu\mu} = -\sum_{B\neq A} Z_B\,(\nu^A\mu^A\,|s^Bs^B); \qquad \nu,\mu \text{ on A},\ \nu \neq \mu \tag{47c}$$

The <u>two-center</u> one-electron integrals, $H^{CORE}_{\mu\lambda}$, are often referred to as the one-electron core resonance integrals. These integrals are proportional to the overlap of the atomic Slater orbitals, $S_{\mu\lambda}$, and are approximated as[32]

$$H^{CORE}_{\mu\lambda} = S_{\mu\lambda}\,\frac{1}{2}(\beta^A_\mu + \beta^B_\lambda) \qquad \lambda,\ \mu \text{ on different centers.} \tag{48}$$

Here, β^A_μ and β^B_λ are adjustable parameters, specific to each atomic orbital (μ, λ) and element (A, B). The $S_{\mu\nu}$ overlap integrals are evaluated analytically, using the STO orbital exponents.[32] The $H^{CORE}_{\mu\lambda}$ integrals are the main contributors to the bonding energy of the molecule.[32,42b]

<u>Two-Electron Integrals:</u> The two-electron integrals in the Fock matrix are divided into one-center terms and two-center terms. The one-center two-electron terms in (44) and (45) are the two-electron terms involving ν and μ only. The ν

and μ correspond to the s and p atomic orbitals on center A. There are five unique one-center two-electron integrals for each atom.[30]

(49) $(ss|ss) = G_{ss}$ $(ss|pp) = G_{sp}$ $(pp|pp) = G_{pp}$ $(pp|p'p') = G_{p2}$ $(sp|sp) = H_{sp}$

The notation used for these integrals is that used by Stewart in his first PM3 paper,[30] and subsequent publications. In (49), "s" represents an s orbital, while p and p' represent p orbitals that differ in angular orientation. These quantities are all part of the parameter set for each element, and are assigned values derived from atomic spectra.[30,42]

The two-center electron repulsions are the $(\nu\lambda \mid \sigma\mu)$ type terms in (44), (45) and (46), which occur between orbitals on different centers. Explicit evaluation of these integral using STOs would require too much computer time; therefore, the two-center electron-electron repulsion energy in MNDO, AM1 and PM3 are approximated as a sum of classical multipolar interactions.[32] Each pair of densities in the two-center integrals is approximated by a set of point charges that depends on their symmetry, as described below[2] and in reference 32.

Orbitals	**Multipole**	**Description**
|ss)	monopole	one -1 charge centered at the nucleus
|sp)	dipole	one -1/2 at ($-D_1$,) one +1/2 charge at ($+D_1$)
|pp)	monopole + linear quadrupole	+1/4 charge at the nucleus and +1/4 charge at the nucleus, and two -1/4 charges, located at ($+2D_2$) and ($-2D_2$)
|pp')	square quadrupole	two -1/4 charges and two +1/4 charges, forming a square of side $2D_2$ centered at the nucleus.

The (D_i) in the table above represent the distance of the point charge from the nucleus. The D_i are determined from STO orbital exponents, using the condition that the multipole moment of each multipole distribution is equal to that of the corresponding charge distribution. The simplest of the multipolar two-center terms is the ($s_A s_A | s_B s_B$) repulsion integral, which is approximated as[32]

$$(50) \qquad (s_A s_A | s_B s_B) = \left[(R_{AB})^2 + \frac{1}{4}\left(\frac{1}{G_{ssA}} + \frac{1}{G_{ssB}} \right)^2 \right]^{-1/2},$$

in MNDO, AM1 and PM3. All other two-centered electron repulsion integrals are more complicated functions of the ζ Slater orbital exponents, and the G_{ij} and H_{sp}

one-center two-electron repulsion integrals. The repulsion integrals are derived by ensuring the correct asymptotic behaviour of the repulsion integral in the limits of $R_{AB} \rightarrow \infty$, and $R_{AB} \rightarrow 0$. For example, if A and B are the same element (A = B) in (50), the resulting $(s_A s_A | s_A s_A)$ goes to G_{ssA} (the monocentric A (ss|s's') repulsion integral) as $R_{AA} \rightarrow 0$, and to 0 as $R_{AA} \rightarrow \infty$. For a complete discussion of MNDO two-centered electron-electron integral evaluation, see references 32 and 42.

1.6.2. MNDO, AM1 and PM3 Nuclear Repulsion Terms:

The inner shell electrons are not included explicitly in MNDO, AM1 and PM3 electronic structure calculations. The inner shell electrons are combined with the nuclear charge to form an atomic core. The inner-shell electrons effectively screen the bare nucleus, so the nuclear-nuclear repulsion terms are replaced with core-core repulsion terms. It is in the nuclear repulsion expressions where the functional forms of the MNDO, AM1 and PM3 methods differ. In MNDO, the nuclear-core repulsion terms between atoms A and B are approximated as[32]

$$(51) \qquad E_N(A,B) = Z_A Z_B (s_A s_A | s_B s_B)[1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})],$$

where $(s_A s_A | s_B s_B)$ is two-center (ss|ss) electron-electron repulsion integral. The α_A and α_B are atom-specific adjustable parameters. With the nuclear repulsion term above, MNDO overestimates the repulsion between atoms at their van der Waals distance.[37] To correct this trend, Dewar and co-workers[37-38] developed the AM1 (Austin Model 1) method. AM1 is identical to MNDO, except that the core-core repulsions are modified to better describe sterically crowded systems, strained ring systems, and hydrogen-bonding. To correctly account for hydrogen bonding, the core-core repulsion terms for N-H and O-H interactions in AM1 are modified from MNDO, and Gaussian terms are added.[37]

$$(52) \quad E_N(A,H) = Z_A Z_H (s_A s_A | s_H s_H)\left(1 + \frac{\exp(-\alpha_A R_{AH})}{R_{AH}} + \exp(-\alpha_H R_{AH})\right) + \frac{Z_A Z_H}{R_{AH}}\left(\sum_k a_{kA} \exp[-b_{kA}(R_{AH} - c_{kA})^2] + \sum_k a_{kH} \exp[-b_{kH}(R_{AH} - c_{kH})^2]\right)$$

Atom-atom core repulsion, other than those involving hydrogen bonding, are modified from MNDO only by the addition of Gaussian terms.[37]

$$E_N(A,B) = Z_A Z_B (s_A s_A | s_B s_B)\left(1 + \exp(-\alpha_A R_{AB}) + \exp(-\alpha_B R_{AB})\right)$$
$$+ \frac{Z_A Z_B}{R_{AB}}\left(\sum_k a_{kA} \exp[-b_{kA}(R_{AB} - c_{kA})^2] + \sum_k a_{kB} \exp[-b_{kB}(R_{AB} - c_{kB})^2]\right) \quad (53)$$

In the AM1 method, the summations in (53) are over k = 1→3. The PM3 parameterization of uses the same nuclear repulsion terms as AM1, except that fewer gaussians are used (k = 1→2).[30]

The parameters discussed in the previous section are summarized in the following table. Each element in the MNDO, AM1 and PM3 methods has its own set of parameters, a complete list of which may be found in the HyperChem manual.[2] The notation used here for these parameters is the same as used by Stewart,[30] and as employed in the HyperChem manual.[2]

Table 1.1: MNDO, AM1 and PM3 Parameter Types:

Parameter (unit)	comment	Parameter (unit)	comment
U_{ss} (eV)	"s" one-electron energy	ζ_s a.u.	"s" STO exponent
U_{pp} (eV)	"p" one-electron energy	ζ_p a.u.	"p" STO exponent
G_{ss} (eV)	(ss|ss)one-center two-electron repulsion	β_s (eV)	"s" one electron two center resonance
G_{pp} (eV)	(pp|pp)one-center two-electron repulsion	β_p (eV)	"p" one-electron two center resonance
G_{p2} (eV)	(pp|p'p')one-center two-electron repulsion	α $(Å)^{-1}$	core-core repulsion parameter
G_{sp} (eV)	(ss|pp)one-center two-electron repulsion	a_{1-x}	gaussian coefficients (AM1 x =3: PM3 x = 2) (none for MNDO)
H_{sp} (eV)	(sp|sp)one-center two-electron repulsion	b_{1-x} , c_{1-x}	gaussian exponents terms (AM1 x =3: PM3 x = 2) (none for MNDO)

1.7 *Ab initio* Methods:

1.7.1. Hartree Fock Ab initio Methods

Hartree-Fock *ab initio* methods solve the Schrödinger equation using only fundamental physical constants, such as the electron mass and Plank's constant. The variational principle states that any calculated energy will be greater than the true energy.

$$E^{true} \leq E^{calc} = \frac{\int \Psi^* H \Psi dx}{\int \Psi^* \Psi dx} \quad (54)$$

Therefore, any trial molecular Ψ will produce a calculated energy E^{calc} greater than the true energy E^{true}. Increasingly more refined and adjustable expansions for Ψ are better approximations for the true wavefunction Ψ^{true}, and will consequently yield calculated energies closer to the true energy. In practice, using increasingly more refined atomic basis functions ϕ for the Hartree-Fock wavefunction will lower calculated energy.

Ab initio calculations require the solution of many integrals for chemical accuracy, so they are often performed with Gaussian-type orbitals, which have the general form

$$g^{GTO}(r,k,l,m,\zeta) = Nx^k y^l z^m \exp(-\zeta r^2). \quad (55)$$

Here, the zeta (ζ) orbital exponents are optimized for each type of orbital in each element. The k, l, and m are integers which determine the angular behaviour of the orbital (e.g. for an 's' orbital, k,l and m = 0). N is a normalization factor.

Gaussian type orbitals are more easily evaluated than Slater type orbitals, because the product of two gaussian functions is another gaussian. This property greatly simplifies multi-center integral computation. However, gaussian orbitals are not a good approximation to atomic orbitals, and are a worse representation of atomic orbitals than Slater type orbitals. Therefore, a linear combination of primitive gaussian orbitals g^{GTO}'s are combined to form a contracted gaussian AO basis function ϕ^{GTO}, which more closely reproduces the STOs and the cusp at the nucleus in hydrogenic orbitals.[16]

(56) $$\phi^{GTO} = \Sigma d_i \, g_i^{GTO} .$$

The d_i in (56) are fixed throughout the SCF calculation, and are determined by fitting the contracted Gaussian to an optimized Slater type orbital.

Ab initio results are very sensitive to the form of the basis set.[43] Gaussian-type basis sets as classified by how many primitive g_i^{GTO} gaussians are used to describe each ϕ^{GTO}, and by how many contracted ϕ^{GTO} are used to form each AO. A minimal basis consists of one contracted ϕ^{GTO} for each AO, while double-zeta basis sets use two ϕ^{GTO} basis functions, each with the same k,l, and m, but different ζ 's, to describe each AO. Triple-zeta use three ϕ^{GTO} basis functions with the same k,l, and m, but with different ζ 's, to describe each AO. Split-valence basis sets use a minimal basis set for the core AO's, and a double-zeta or triple-zeta basis set for the valence electrons.

Gaussian orbital basis sets are further refined by the addition of p-type polarization functions to H and He, and d-type polarization functions to second row elements. Polarization functions have *l* quantum numbers one larger than the highest valence electron *l* quantum number of the element in its ground state.

The commonly used 3-21G, 4-31G and 6-31G basis sets,[44-49] with and without polarization functions, are employed in this thesis. These are all split-valence basis sets, which used one contracted GTO for each inner shell AO. The 3-21G basis set uses three primitive g_i^{GTO}s to form the contracted GTO inner shell AO. The 4-31G basis set uses four primitives g_i^{GTO}s to form the inner shell contracted GTO AO, and the 6-31G basis uses six primitive gaussians to form the contracted inner shell AO. The 3-21G, 4-31G and 6-31G basis sets all used one contracted GTO and one diffuse primitive GTO to approximate each valence orbital. The 3-21G basis uses two primitive g_i^{GTO} functions for the contracted valence GTO, while the 4-31G and 6-31 basis sets both use three primitive gaussians to form each contracted valence GTO.

The number of polarization functions in the basis set is indicated by asterixes after the basis set specification. Thus, a 6-31G basis set with d-polarization functions is denoted by 6-31G*; the same basis set with p- and d-polarization functions is denoted by 6-31G**. An alternate nomenclature, 6-31G(*d*) for 6-31G* and 6-31G(*d,p*) for 6-31G** can also be used.

It is instructive to note that the relative computational times for semi-empirical, 3-21G, 4-31G and 6-31G** Hartree-Fock calculations are approximately 1 for semi-empirical (MNDO, AM1 and PM3), 100 for STO-3G, 600 for 4-31G and 3000 for 6-31G**.[50]

1.7.2 Electron Correlation and Ab Initio Methods:[16]

The Hartree-Fock *ab initio* total energy is improved by increasing and optimizing the basis set. However, there is a limit to the accuracy obtained with a single determinant HF wavefunction; the Hartree-Fock limit. Further improvement of the *ab initio* total energy is achieved by going beyond the HF single-determinant approximation, and expanding the total electronic wavefunction Φ as a linear combination of Slater determinants Ψ

(57). $$\Phi = C_0\Psi_0 + \sum_s C_s\Psi_s + \sum_d C_d\Psi_d + \sum_t C_t\Psi_t + \cdots$$

Here, Ψ_0 is the ground state, or reference determinant. The other Ψ_i Slater determinants, are formed by promoting electrons from occupied states in Ψ_0, to virtual states. The singly excited determinant Ψ_s is formed by promoting one electron from an occupied state to a virtual state. The doubly excited Ψ_d determinant is formed by promoting two electrons from occupied to virtual states. The Ψ_i excited state determinants are different configurations of the ground state determinant in (57).[16] The Φ wavefunction is a post-Hartree-Fock wavefunction, and is often referred to as a Configuration Interaction (CI) expansion.[16] The C's in (57) are variationally determined coefficients. Ideally, the expansion includes an infinite number of configurations and basis functions. This unattainable upper limit to a CI calculation would, in principle, yields the exact total energy of the

system, E^{true}. In practice, the CI expansion must always be truncated; therefore calculated CI total energies are always an upper bound to the exact total energy

(58) $$E^{CI} \geq E^{true}.$$

However, CI calculations give energies closer to E^{true} than HF calculations. The difference between the HF energy E^{HF} and the true energy E^{true} is referred to as the 'correlation energy' E^{corr} ;

(59a) $$E^{corr} = E^{true} - E^{HF}$$

which can be approximated as

(59b) $$E^{corr} \approx E^{CI} - E^{HF}.$$

In practice, E^{corr} is about 0.5%-2.0% of the total energy; however, this can be 200-2000 kcal/mol of energy, depending upon the value of the total energy, and the magnitude of the correlation error. Consequently, the correlation energy can have a large effect on *ab initio* calculated reaction energies. Correlation effects can however cancel, as will be seen in chapter three.

Extensive CI calculations on even the smallest molecules are nearly impossible, even with the most powerful computers. However, a recent paper by Wulfov[51] suggests that it may be possible to extrapolate to the full CI limit using an efficient algorithm, on a personal computer. A fuller discussion of Configuration Interaction and correlation energy is in chapter nine.

All calculations in this thesis, except for those in chapter nine, are not performed with post-Hartree-Fock theories, because the size and number of compounds to be studied make such calculations too expensive. Relativistic effects are ignored, because the main properties studied arise from valence electrons in light elements, where relativistic effects are negligible.

1.8 Hessians and Gradients:[52-56]

Calculating the molecular energy for a molecule containing 'A' nuclei at a fixed set of coordinates, by either molecular mechanics, semi-empirical or *ab initio* methods, gives the total energy at a point on the 3A dimensional potential energy surface (PES) of the nuclei. Such calculations are denoted as "single-point" calculations. The nuclear coordinates are fixed in a single-point calculation. The single point energy E(**R**) of a given nuclear configuration is parametrically dependent upon the nuclear coordinates, represented by a row vector $\mathbf{R}^{\dagger} = (R_1,....R_{3A})$. A is the total number of nuclei. Following an incremental displacement of each nucleus by δR_{Ai}, the energy at $\mathbf{R}+\delta\mathbf{R}$ can be expanded as the first two terms in a Taylor series,[54]

$$E(\mathbf{R}+\delta\mathbf{R}) = E(\mathbf{R}) + \delta\mathbf{R}^{\dagger}\ \mathbf{g}(\mathbf{R}) + 1/2\ \delta\mathbf{R}^{\dagger}\ \mathbf{H}(\mathbf{R})\ \delta\mathbf{R}\,. \tag{60}$$

In equation (60), $\delta\mathbf{R}$ is a column vector of the δR_{Ai} increments, and $\mathbf{g}(\mathbf{R})$ is a column vector of gradients, whose elements, g_i , are[54]

$$g_i = \frac{\partial E(\mathbf{R})}{\partial R_i}\,. \tag{61}$$

The $\mathbf{H}(\mathbf{R})$ in equation (60) is called the Hessian matrix. The Hessian must be calculated to differentiate between different types of extrema. The Hessian is a 3A X 3A matrix of second-derivatives, whose elements are[54]

$$H_{ij} = \frac{\partial^2 E(\mathbf{R})}{\partial R_i\, \partial R_j}\,. \tag{62}$$

The gradient gives the forces on the atoms at that point on the Potential Energy Surface (PES). Near-zero-gradients ($\mathbf{g}(\mathbf{R}) = \mathbf{0}$) correspond to local extrema on the PES, and show that there are almost no forces on the nuclei in that

geometry. Gradients can be obtained numerically or analytically at little computational cost, as discussed in various papers on this topic.[52-56]

Diagonalization of the Hessian matrix leads to eigenvalues which reflect the curvature of the PES. For linear molecules, five of these eigenvalues will correspond to rotations and translations, while for non-linear molecules, six of these eigenvalues will correspond to rotations and translations. From the harmonic oscillator approximation, the square-root of each remaining vibrational Hessian matrix eigenvalue (excluding the rotational and translational components) is proportional to the vibrational frequency for that internal coordinate.

If all the Hessian eigenvalues are positive, the vibrational frequencies are all real, and the structure is at a local minima on the PES. This corresponds chemically to an equilibrium or intermediate structure.

Any negative Hessian eigenvalues correspond to imaginary vibrational frequencies, which indicate that the structure is at an extrema other than a local minima. Structures with only one imaginary vibrational frequency represent saddle-points (or cols) on the PES, and are of interest because they correspond to transition states in conformational and chemical changes.[57,58] Energies at cols correspond to activation energy barriers. Structures with more than one imaginary vibrational frequency can be shown to minimize to structures either one or no imaginary vibrational frequency;[59] therefore, they are of little chemical interest. The force constants and vibrational frequencies from Hessian calculations can also be used to compute thermodynamic properties such as entropies and heat capacities.

Like gradients, Hessians can be computed numerically or analytically, but unlike gradients, calculation of Hessians is computationally costly, and often requires more computational time than gradient-only driven geometry optimizations.

1.9 Geometry Optimizations.[60,61]

Geometry optimization calculations are performed to determine equilibrium, excited state and transition state molecular geometries. A geometry optimization is initiated by a single point energy and gradient calculation on a given input structure. If the gradient is greater than a preset threshold, the nuclei are moved in a direction predicted to reduce the energy. A new energy and gradient are then computed. The procedure is repeated until the gradient is reduced to almost zero. In practice, the criteria for geometric convergence is often that the root-mean square (RMS) of the gradient, defined as[2]

$$(63) \quad \text{RMS Gradient} = \left(\sum_{A} \left(\frac{\partial E}{\partial R_{XA}} \right)^2 + \left(\frac{\partial E}{\partial R_{YA}} \right)^2 + \left(\frac{\partial E}{\partial R_{ZA}} \right)^2 \right)^{1/2},$$

be below a preset threshold. At this point the geometry is said to be "optimized."

Molecular structures are "optimized" by performing a geometry optimization routine. Geometry optimization input structures may be obtained from experimental data, computation, or from chemically intuitive approximation.

An optimized structure represents a stationary point, or local extrema on the PES. Chemically interesting structures are almost always extrema on the PES. Geometry optimizations can be used to calculate equilibrium molecular geometries, sometimes with an accuracy that rivals experiment.[54]

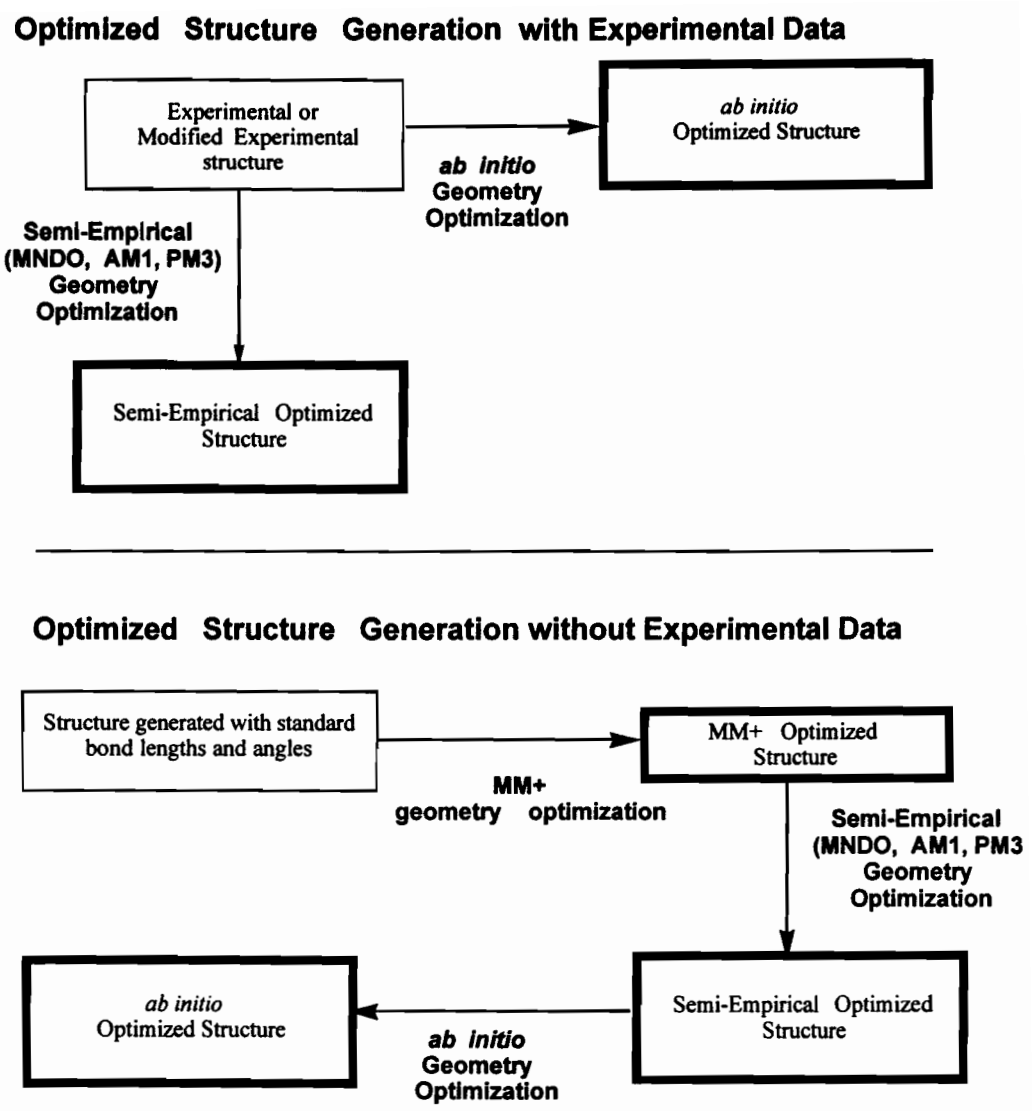
Reviews and outlines of different geometry optimization algorithms are in references 52-56. In this thesis, semi-empirical and *ab initio* geometry optimizations were performed with the Baker[60] and Schlegel[61] methods included

in the GAMESS[3] package. Molecular mechanics (MM+) geometry optimizations were performed with the Newton-Raphson block-diagonal method included in the HyperChem[2] package.

Geometry optimizations on molecules with many conformational degrees of freedom depend on the input structure, the geometry optimization algorithm, and the choice of coordinate system (internal or Cartesian). Geometry optimizations on highly flexible molecules may get caught in one of many local minima, which can be quite different from the ground state structure. A tree-branch method, developed by Villamagna,[5] overcomes the multiple minima problem in the prediction of hydrocarbon surfactant chain conformations. However, since all the molecules in the thesis are ring systems with few degrees of freedom, location of minima other than the ground state structure is improbable. Semi-empirical and *ab initio* geometry optimizations were found to invariably return identical ground state geometries, regardless of optimization algorithm or molecular coordinate system.

1.10 Computational Approach

The previously described theories, as implemented in the HyperChem and GAMESS computational packages,[2,3] were used to study the structures, energies and reactivities of the various compounds discussed at the beginning of the introduction. A typical procedure to model a chemical systems with these methods is outlined in **Figure 1.1**.



(plain boxes indicate input; bolded boxes indicate optimized structures,which can be used as input)

Figure 1.1

Outline of Modelling Procedure

In **Figure 1.1**, experimental structures (mostly X-ray) are input for geometry optimizations whenever possible. In the absence of experimental data, input for quantum calculations were generated using either MM+ minimized molecular structures, or modified quantum mechanically optimized structures. Inputs for transition state geometry searches were created by intuitive alteration of relevant

reactant and product structures. The energies and geometries resulting from the quantum calculations are analyzed to yield chemical information. The details of the computations used in each study are given in the respective chapters.

Throughout the thesis, energy differences are reported in kilocalories per mol (kcal/mol), as customary in quantum chemistry. The conversion of kilocalories to Joules is 1 kcal = 4 184 J. Bond distance are reported in angstroms (Å); 1 Å = 10^{-10} m. *Ab initio* total energies are reported in atomic units (a.u.); 1 a.u. = 2625 kJ/mol; to convert to kcal/mol, 1 a.u. = 627.5 kcal/mol.

Glossary:

Molecular Mechanics:

MM+	MM+ force field[2] based on Allinger's MM2 force field[6-9]
AMBER	Assisted Model Building and Energy Refinement [11a] -developed for proteins and nucleic acids...not used in the thesis
OPLS+	Optimized Potentials for Liquid Simulations[11b,c] - similar to AMBER; not used in the thesis

Semi-Empirical Quantum Mechanics

MNDO	Modified Neglect of Differential Overlap -a semi-empirical method developed by Dewar et. al.,[31-33] which simplifies the Fock matrix with experimental parameters, and by neglect of diatomic differential overlap terms.
AM1	Austin Model 1 -an imporved MNDO method, developed by Dewar et. al. in Austin, Texas.[37,38]
PM3	Paramaterization Model 3 -an MNDO/AM1 method, which was reparameterized by Stewart.[30,39]

References:

1. P. A. M. Dirac, *Proc. Roy. Soc. (London),* **123**, 714 (1929).

2. The HyperChem™ program and manual is available from Hypercube Inc., 419 Philip St., Waterloo, Ont. Canada N2L 3X2.

3. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen, S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** 52(1990).
   b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).

4. F. Villamagna and M. A. Whitehead, *J. Chem. Soc., Faraday Trans.* **90**, 47 (1994).

5. F. Villamagna, "*Modelling of Interfaces in Emulsion Explosives*" Ph.D Thesis, McGill University, 1992.

6. N.L.Allinger, E.L. Eliel, S.J. Angyal, G.A. Morrison, "*Conformational Analysis,*" (Wiley: New York) 1965.

7. N.L.Allinger, *Adv. Phys. Org. Chem.* **13**, 1, (1976).

8. N.L. Allinger and U. Burkett, "*Molecular Mechanics*" ACS Monograph 177, ACS (1982).

9. N.L.Allinger, Y.H. Yuh and H.H. Lii, *J. Am. Chem. Soc.*, **111**, 8551, 8566, 8576 (1989).

10. N.L.Allinger and H.H. Lii, *J. Am. Chem. Soc.*, **111**, 8566, 8576 (1989).

11. a)P.K. Weiner and P.A. Kollman et. al. *J. Am. Chem. Soc.*, **106**, 765 (1984).
    b) W.L Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).
    c) J. Pranata, S. Wierschke, and W.L. Jorgensen, *J. Am. Chem. Soc.*, **113**, 2810, (1991)

12. a) K.J. Laidler, and J.H. Meiser, "*Physical Chemistry,*" (Benjamin-Cummings: Don Mills, Ont.) p771, 1982.
    b)J.E. Lennard-Jones *Proc. R. Soc. Chim. Fr.*, 1113, (1963).

13. I.R. Levine, *"Quantum Chemistry"*, (Allyn and Bacon: Toronto) 1983.

14. P.W. Atkins, *"Molecular Quantum Mechanics"* (Oxford University Press: New York) 1986.

15. D.A. McQuarrie, *"Quantum Chemistry"*, (University Science Books: Mill Valley, California) 1983.

16. A. Szabo and N.S. Ostlund, "*Modern Quantum Chemistry*" (MacMillan: New York) 1982.

17. A. Hinchliffe, *"Computational Quantum Chemistry"* (Wiley: New York) 1988.

18. Guy H. Grant and W.G. Richards, *"Computational Chemistry"* (Oxford University Press: New York) 1995

19. G.C. Schatz and M.A. Ratner, *"Quantum Mechanics in Chemistry"* (Prentice Hall: Englewood Cliffs, New Jersey) 1993.

20. W.G. Richards and J.A. Horsley, *"Ab initio Molecular Orbital Calculations for Chemists"* (Clarendon Press: Oxford) 1970.

21. J.A. Pople and D.L Beveridge, *"Approximate Molecular Orbital Theory"* (McGraw-Hill : New York) 1970.

22. M.S. Dewar and R.C. Dougherty, *"The PMO Theory of Organic Chemistry"* (Plenum Press: New York) 1975.

23 S. Suba and M.A. Whitehead in , "*Recent Advances in Density Functional Methods, Part I*" D. P. Chong, ed. (World Scientific: Ney Jersey, U.S.A.) (Chapter Two) 1995.

24. Y. Guo and M.A. Whitehead, *Phys. Rev. A.,* **43**, 95 1991.

25. S. Manoli and M.A. Whitehead, *Phys. Rev. A,* **38**, 630, 1988.

26. a) W. Kohn and L.J. Sham, *Phys. Rev. A.,* **140a**, 1133 (1965).
b)E.S. Kryachko and E.V. Ludeña, *" Energy Density Functional Theory of Many-Electron Systems*" (Kluwer Academic : Dordrecht, Netherlands) 1990.
c) R.G. Parr and Y. Wang, "*Density Functional Theory of Atoms and Molecules"* (Oxford University Press: New York) 1989.

27. C.C.J. Roothan, *Rev. Mod. Phys.,* **23**, 69 (1951).

28. J.A. Pople and R.K. Nesbet, *J. Chem. Phys.*, **22**, 271 (1954).

29. a)G.A. Segal, ed., "*Semi-Empirical Methods of Electronic Structure Calculation*," (Plenum: New York) 1977.
b) T. Clark, "*Handbook of Computational Chemistry*," (Wiley: New York) 1985.
c) L. Sadlej, "*Semi-Empirical Methods of Quantum Chemistry*," (Wiley: New York) 1985.

30. (a) J.J.P. Stewart, *J. Comput. Chem.*, **10**, 209 (1989).
(b) R.L Johnston, P.J. Knowles and J.N. Murrell, *"Molecular Orbital Theory"* in *"Annual Reorts; Section C: Physical Chemistry. Vol 87"* (Royal Society of Chemistry: Cambridge) 1991.
(c) F. L. Pilar, *"Elementary Quantum Chemistry, 2nd Edition"* (McGraw-Hill: New York) 1990.

31. M.J.S. Dewar and M.L. McKee, *J. Am. Chem. Soc.*, **99**, 5231 (1977).

32. M.J.S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4899 (1977).

33. M.J.S. Dewar and H.S. Rzepa, *J. Am. Chem. Soc.*, **100** 58 (1978).

34. L.P. Davis, R.M. Guidry, J.R. Williams, M.J.S. Dewar and H.S. Rzepa, *J. Comput. Chem.*, **2**, 433 (1981).

35. R. C. Bingham, M.J.S. Dewar and D.H. Lo, *J. Am. Chem. Soc.*, **97**, 1285, 1294, 1302, 1307 (1975).

36. M.J.S. Dewar, *Science*, **187**, 1037 (1975).

37. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am Chem Soc.*, **107**, 3902 (1985).

38. M.J.S. Dewar and K.M. Dieter, *J. Am Chem. Soc.*, **108**, 8075 (1986).

39. a) J.J.P. Stewart, *J. Comput. Aided Mol. Design*, **4**, 1 (1990).
b) J.J.P Stewart, *J. Comput. Chem.*, **10**, 221 (1989).

40. a)J.A. Pople, D.P. Santry and G.A. Segal, *J. Chem. Phys.*, **43**,5129 (1965).
b) K. Ruedenberg, *Rev. Mod. Phys.*, **34**, 326 (1962).

41. L. Oleari, L. DiSipio and G. DeMichelis, *Mol. Phys.*, **10**, 97 (1966).

42. M.J.S. Dewar and W. Thiel, *Theoret. Chim. Acta (Berl.)*, **46**, 89 (1977).

43. E.R. Davidson and D. Feller, *Chem. Rev.* **86**, 681 (1986).

44. R. Ditchfield, W.J. Hehre and J.A. Pople, *J. Chem. Phys.*, **54**, 724 (1971).

45. R. Ditchfield, W.J. Hehre and J.A. Pople, *J. Chem. Phys.*, **56**, 2257 (1972).

46. J.S. Binkley, J.A. Pople and W.J. Hehre, *J. Am. Chem. Soc.*, **102**, 939 (1980).

47. M.M. Francl, W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. DeFrees, J.A. Pople *J. Chem. Phys.*, **77**, 3654 (1982).

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

49. K.D. Dobbs and W.J. Hehre, *J. Comput. Chem.*, **7**, 359 (1986).

50. W.J. Hehre, *J. Am. Chem. Soc.*, **97**, 5308 (1975).

51. A.L. Wulfov, *Chem. Phys. Lett.*, **255**, 300 (1996).

52. R. Fletcher, "*Practical Methods of Optimization,*" (John Wiley and Sons, New York) 1980.

53. P.E. Gill, W. Murray, and M.H. Wright, *"Practical Optimizations"* (Academic Press, Inc., New York) 1981.

54. M. Zerner, Appendix C in A. Szabo and N.S. Ostlund, "*Modern Quantum Chemistry*" (McGraw-Hill, New York) 1989.

55. S. Bell and J.S. Crighton, *J. Chem. Phys.*, **90**, 344, (1986).

56. H.B. Schlegel, "*Advances in Chemical Physics (Ab initio Methods in Quantum Chemistry),*" K.P. Lawley, ed. (Wiley: New York) 1987.

57. K.J. Laidler, "Chemical Kinetics" (Harper and Row : New York) 1987.

58. J.M. McIver, Jr., and A. Kormornicki, *J. Am. Chem. Soc.*, **94**, 2625, (1972).

59. J.N. Murrell and K. J. Laidler, *Trans. Faraday Soc.*, **64**, 371 (1968).

60. J. Baker, *J. Comput. Chem.*, **7**, 385 (1986).

61. H.B. Schlegel, *J. Comput. Chem*, **3**, 214 (1982).

The endohedral and exohedral complexing of small molecules with fullerenes has been of interest since the discovery of fullerene systems. In chapter 2, the accuracy of a simple atom-atom potential method for calculating these complexing energies is assessed. The atom-atom potentials give results quantitatively and qualitatively similar to *ab initio* calculations. This paper justifies the use of the atom-atom potential method for studying these non-bonded systems. This chapter also shows that high level *ab initio* calculations do not necessarily give results better than simpler and less computationally expensive methods.

Chapter 2:

Interaction and dynamics of Endohedral Gas Molecules in C_{60} Isomers And C_{70}

C. I. Williams, L. Pang and M. A. Whitehead ,

Published in *J. Phys Chem.* **97**, 11652 (1993).

2.1 Abstract

Atom-atom potential energies of endohedral gas fullerene complexes, $G@C_{60}$ and $G@C_{70}$ (G = H_2, N_2, O_2, F_2, HF, HCl, CO, H_2O, CH_4 and CO_2) have been calculated semi-empirically. The non-bonded energies give insight into the molecular interactions and dynamics of the gas guest molecules inside cavities of four C_{60} isomers and C_{70}. Only complexes where the gas molecules and the fullerene cavities have minimum van der Waals overlap are stable. The degree of molecular motion of the gas molecule inside the fullerene cavities is determined from the rotational energy barriers.

2.2 Introduction

The discovery of Buckminsterfullerene[1] opened up a new family of compounds. Since their discovery, C_{60} and other fullerene solids have been found to absorb metal elements [2] and give endohedral fullerene complexes; this led to the discovery of fullerene superconductors[3]. However, the large, closed spheroidal shapes of the fullerene molecules suggest that they could hold a variety of small guest atoms and/or molecules and form endohedral complexes defined by the formula $G@C_{60}$. Indeed, C_{60} fullerenes containing La, Ni, Na, K, Rb, Cs atoms[4-6] and helium, neon atoms[7-9] have been observed. Extensive calculations have been made on both the static and dynamic properties of guest atoms inside fullerene cages.[10-13]

The results of a semi-empirical atom-atom potential (AAP) treatment of fullerene-guest interactions is presented. The AAP method calculates mainly non-bonded interactions using a Buckingham potential akin to the non-bonded potentials used in molecular mechanics force fields. This method has described the molecular dynamics and complexing characteristics[14,15] of inclusion compounds. For fullerenes, the stability of each fullerene-guest system is determined by the change in potential energy as a function of the relative distance between the fullerene and the guest molecule. At infinite fullerene-guest separation, the

potential energy is assumed zero; energy changes that occur in the system as the fullerene and the guest are brought together are the *complexing energies*. There is particular interest in the complexing energies of four different C_{60} fullerene isomers with H_2O, CH_4, H_2, O_2, N_2, F_2, HF and HCl. The C_{60} isomer geometries were from MNDO calculations.[16] The C_{60} isomers studied were Buckminsterfullerene (abbreviated BF, I_h symmetry), Graphitene (abbreviated GR, D_{6h} symmetry) and two other isomers having C_{2v} symmetry and D_{2h} symmetry. (henceforth abbreviated SW1 and SW2 respectively). The geometry of C_{70} was from reference 17.

2.3 Methods

The calculations were performed on a 386 IBM compatible personal computer with an AAP programme developed in our laboratory. The programme was derived from Gavezotti's OPEC programme,[18] and written in the Pascal language for the IBM PC utility. The modification was made to allow the rotation and translation of the guest molecule along a predefined direction relative to the host system. Thus, translational motion has been defined along the z axis, and rotations around the x, y, and z axes. The atom-atom potential function is a two center function[19]

$$(1) \quad E_p = \sum_{i}^{HOST} \sum_{j}^{GUEST} E_{ij} .$$

The index i runs over the atoms in the fullerene host molecule, and j over the atoms in the guest molecule. The form of the potential E_{ij} is

$$(2) \quad E_{ij} = -A_{ij} r_{ij}^{-6} + B_{ij} \exp(-C_{ij} r_{ij})$$

where r_{ij} is the distance between the i^{th} atom in the guest molecule and the j^{th} atom in the host. The parameters A_{ij}, B_{ij}, and C_{ij} are empirical parameters given by Mirsky[20] and extended to the halogen elements by Gavezzotti.[18] The guest molecule geometries were constructed with known bond lengths and angles.[21] The

diatomic molecules lie on the x axis, making rotation around the x axis redundant. The origins of the diatomic molecule coordinate systems were defined to lie either at the midway point between the atoms (in N_2, O_2, F_2, CO) or at the coordinates of the heavy atom (in HF, HCl). The origin of the fullerene host molecule remained fixed at the gravimetric center of the fullerene cage throughout each calculation, **Fig. 2.1**. The direction of translational motion was in the z axis. The relative distance was the distance between the origins of the fullerene and the guest molecule coordinate systems. The rotational freedom of the guest molecules, inside the fullerene cavities, was investigated by calculating the atom-atom potential at different rotational angles between the fullerene cage and the guest molecule coordinate system.

The AAP results were analyzed in terms of the relative sizes of the guest molecules and the fullerene cavities. The size of a guest molecule or a cavity was derived from its molecular geometry and the van der Waals radii of its constituent atoms. The van der Waals radii are in **Table 2.1.**[22]

The size of a guest molecule is expressed as a van der Waals volume, or simply by the three maximum van der Waals lengths of the molecule along the x, y, and z axes, $l_{VDW(x)}$, $l_{VDW(y)}$, $l_{VDW(z)}$. These parameters can be calculated from,

$$l_{VDW(i)} = |i_{max} - i_{min}| + R_{VDW(imax)} + R_{VDW(imin)} \quad (i = x, y, z) \tag{3}$$

where i_{max} and i_{min} are the maximum and minimum i-coordinates of the molecule; $R_{VDW(imax)}$ and $R_{VDW(imin)}$ are the van der Waals radii of the atoms at i_{max} and i_{min}. The size of a fullerene cavity can be expressed by the van der Waals cavity diameters along the x, y and z axes, e.g. $d_{VDW}(x)$, $d_{VDW}(y)$ and $d_{VDW}(z)$, calculated from

$$d_{VDW}(i) = |i_{max} - i_{min}| - 2\,R_{VDW} \quad (i = x, y, z) \tag{4}$$

where i_{max} and i_{min} are the maximum and minimum i-coordinate of the fullerene; R_{VDW} is the van der Waals radius of the aromatic carbon. **Fig. 2.2** illustrates the $l_{VDW}(i)$ and $d_{VDW}(i)$ of the CO/BF-C_{60} complex on the xz plane. A complete list of van der Waals lengths is in **Table 2.2**. A complete list of fullerene van der Waals dimensions is in **Table 2.3**.

Table 2.1: Van der Waals Radii of Selected atoms (Å).

Atom	R_{VDW}
C (aromatic)	1.75
H	1.17
O	1.40
F	1.30
N	1.55
Cl	1.77

Table 2.2: Van der Waals Lengths of the Guest Molecules (Å).

Guest	Bond Length	$l_{VDW}(x)$	$l_{VDW}(y)$	$l_{VDW}(z)$
H_2	0.747	3.1	2.4	2.4
F_2	1.570	4.3	2.7	2.7
N_2	1.120	4.1	3.0	3.0
O_2	1.208	4.0	2.8	2.8
CO	1.103	4.4	3.7	3.7
HF	0.917	3.5	2.7	2.7
HCl	1.270	4.3	3.6	3.6
CO_2	1.159	5.1	3.7	3.7
H_2O	0.957	3.9	3.2	2.8
CH_4	1.090	3.7	4.2	4.2
COH_2 (C-H)	1.120	4.4	4.3	3.7
(C-O)	1.210			

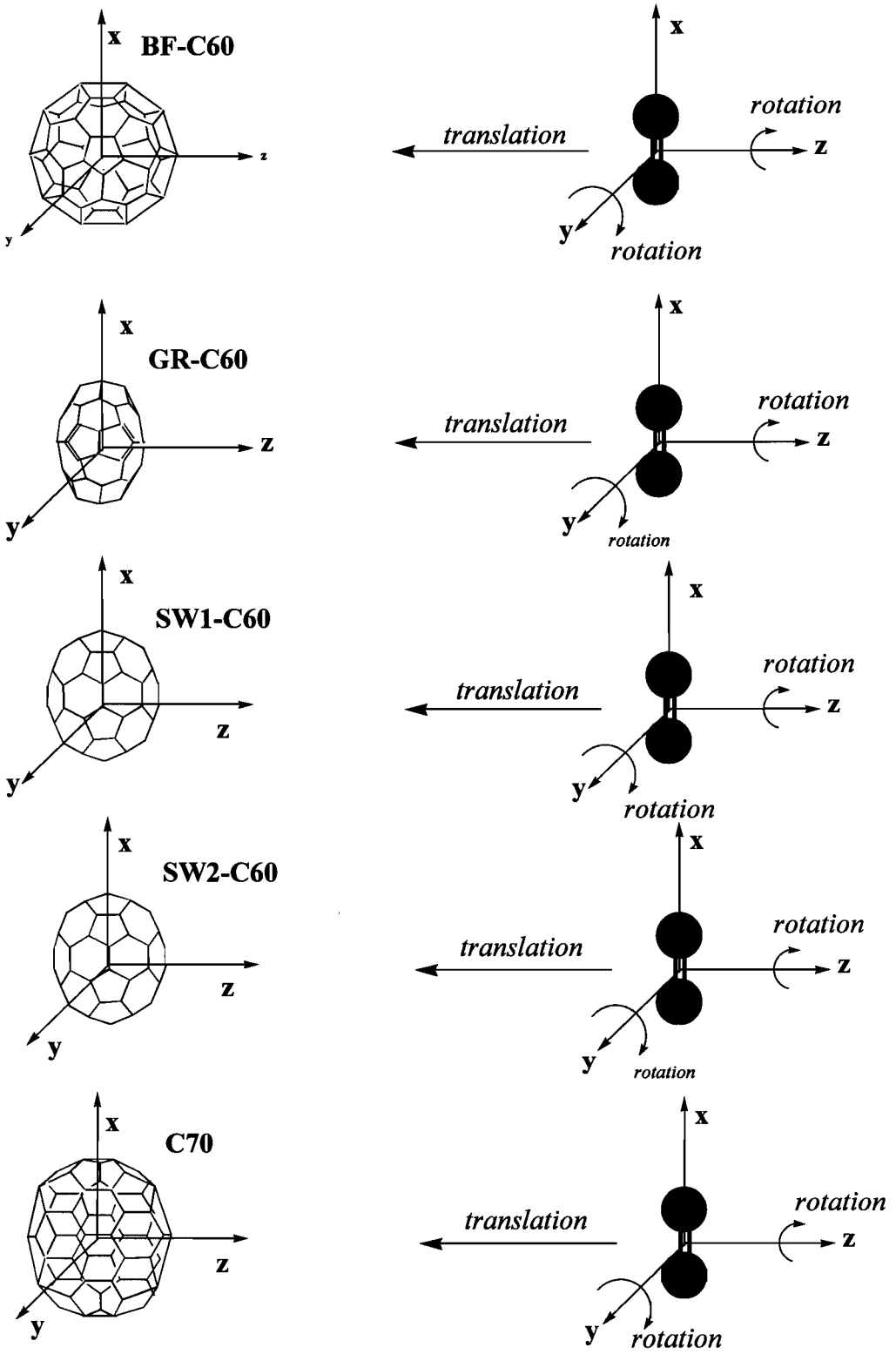


Figure 2.1: Orientations of fullerenes and the direction of Translational Motion used in translational plots

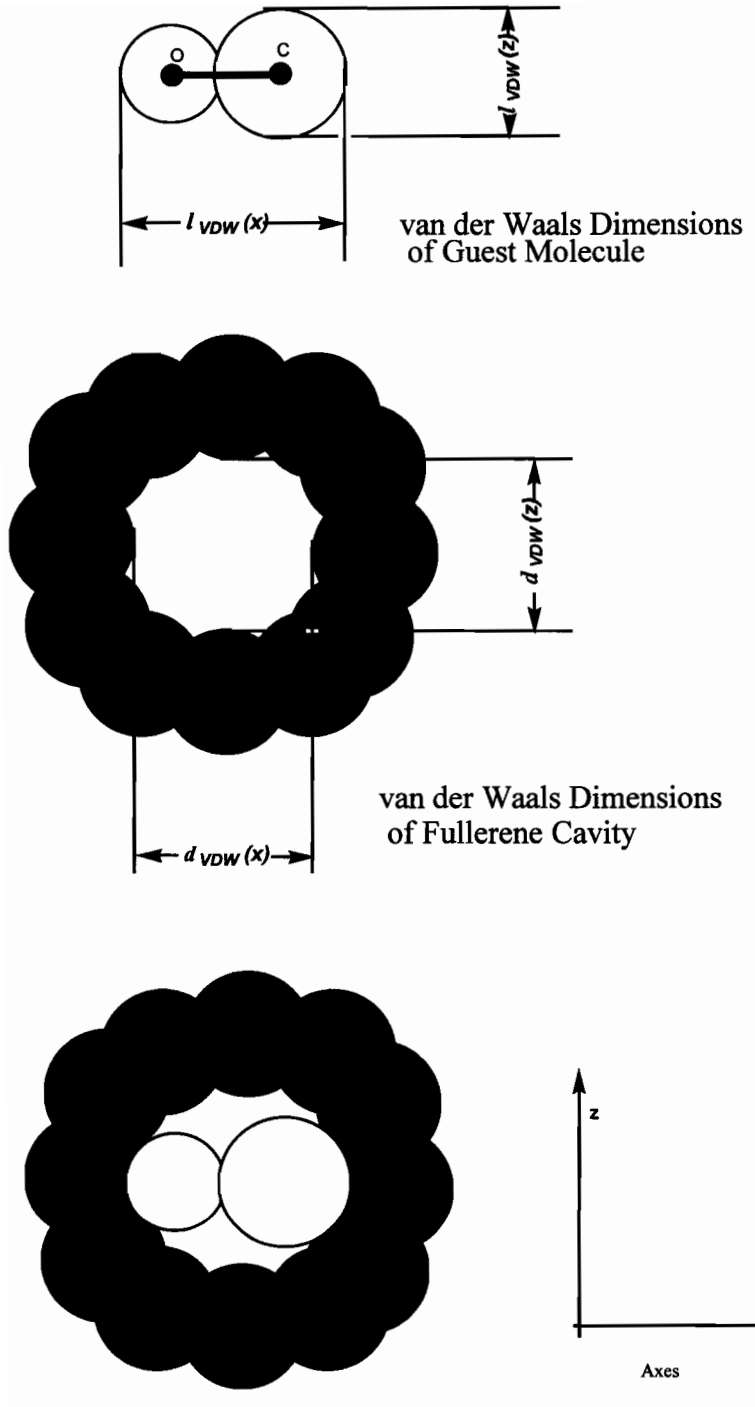


Figure 2.2: van der Waals dimensions of guest molecules and fullerene host

The difference between $l_{VDW}(i)$ and $d_{VDW}(i)$ is defined as the van der Waals overlap between a guest molecule and a fullerene cavity, along each of the three axes of the fullerene. It can be seen from equation (2) that small overlaps between the van der Waals radii of the guest molecule and the fullerene, result in a net lowering of the potential energy, and that large overlaps result in a net increase of the potential energy. The differences in van der Waals radii overlaps between the guest and the fullerene, as a function of rotational angle, were used to analyze the AAP results. Thus, changes in the van der Waals radii overlaps, at different guest molecule orientations, are correlated with the AAP potential as a function of rotational angle.

Table 2.3:Van der Waals Dimensions of Fullerenes (Å)

Fullerene	$d_{VDW}(x)$	$d_{VDW}(y)$	$d_{VDW}(z)$
BF-C_{60} (I_h)	3.3	3.3	3.3
GR-C_{60} (D_{6h})	3.9	1.2	4.6
SW1-C_{60} (C_{2v})	3.7	3.3	3.0
SW2-C_{60} (D_{2h})	3.9	3.3	2.7
C_{70} (D_{5h})	4.3	3.2	3.2

2.4: Results and Discussion

2.4.1. Translational Plots

Fig. 2.3 shows the plots of the atom-atom potential interactions of H_2, F_2, N_2 and O_2 in BF-C_{60}. At a relative distance of 10 Å there are almost no attractive or repulsive forces between the host and guest molecules, and the potential energy of the host/guest system is approximately zero. As the relative distance between the fullerene and the guest is decreased, attractive forces between them increase, resulting in a gradual net lowering of the potential energy. The potential energy continues to decrease until a minimum is reached, at about 6.6-6.8 Å, near the outside surface of the fullerene cage. At this point, the van der Waals radii of the

guest molecule and the fullerene are touching. This minimum represents the most stable exohedral interaction between the two molecules, with the guest molecule bound to the outside of the fullerene cage. As the relative distance between the guest and the fullerene is decreased by another 0.1-0.2 Å the van der Waals radii of the fullerene and the guest molecule begin to overlap slightly, resulting in a small potential energy increase. However, if the overlaps remain less than 0.2 Å attractive forces will still dominate, and the net potential energy will still be negative. When the relative distance between the guest molecule and the fullerene is decreased further, large overlaps between their van der Waals radii generate repulsive forces which raise the potential energy, resulting in a net positive potential energy. As the guest molecule passes through the surface of the fullerene (relative position in the range 1.5-5.0 Å), the distances between their atoms approach zero. At this point, the proximity of the atoms in the guest and in the fullerene suggest that bonded interactions will occur, and changes in the geometry of the fullerene and the guest would be expected. Since there are no terms in our AAP energy to describe these interactions, the AAP energies calculated when the guest passes through the fullerene surface are meaningless. Mowrey et. al.[23] have described the inclusion of He into C_{60}^{+} ions using a many body potential that contains terms which describe bond-breaking and bond formation. However, we have yet to include such terms in our potential, which at present is unable to describe the barrier to inclusion. When the guest molecule has passed through the outside surface of the fullerene cage (relative position < 1.5 Å), the bonded interactions disappear, and the AAP energies again become meaningful. At a relative position of 1.5-1.0 Å, large van der Waals overlaps between the guest and the fullerene give rise to repulsive forces that result in a net positive potential energy. As the guest molecule approaches the center of the fullerene cage (relative position = 0 Å), the van der Waals overlaps between the guest and the fullerene decrease to a minimum. If the fullerene cavity is sufficiently large, the van der Waals radii overlaps will decrease to a point where repulsive forces give way to

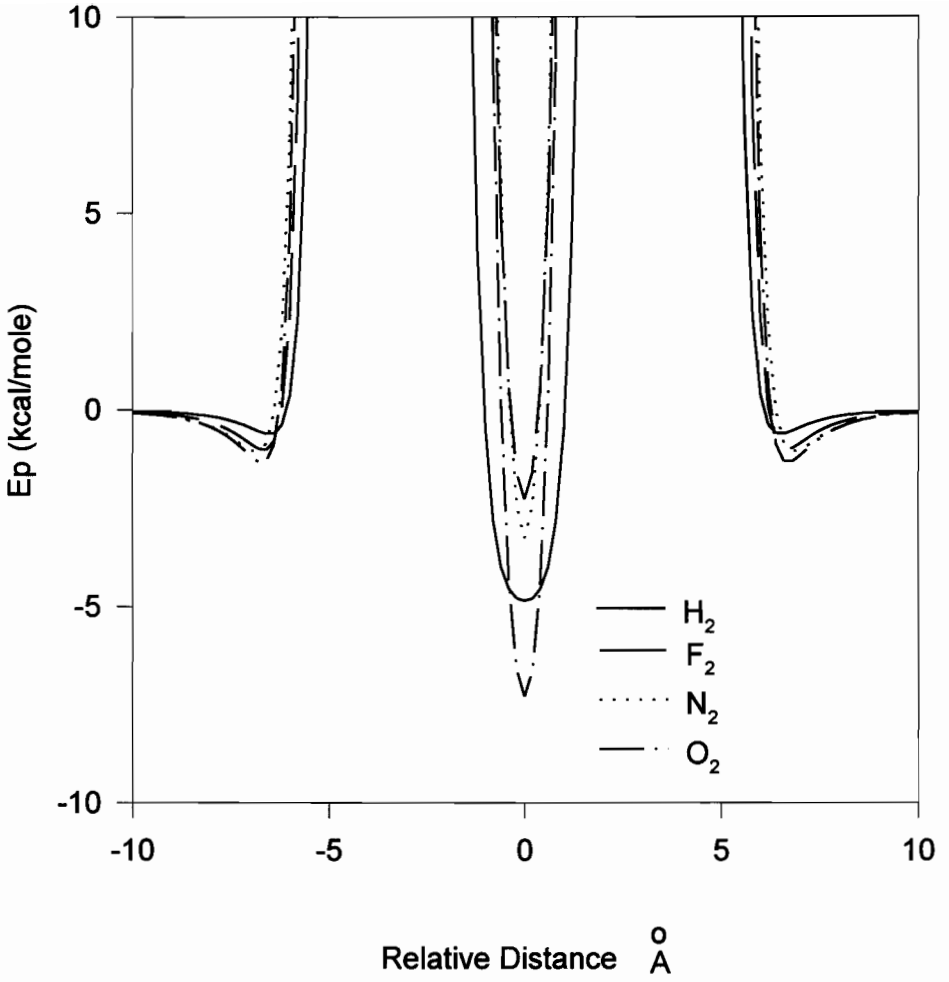


Figure 2.3: Translational Plots of homonuclear diatomics in BF-C60

attractive forces that lower the potential energy, resulting in a negative net potential energy at the endohedral minimum. This potential energy minimum represents the most stable endohedral guest-fullerene interaction. If the endohedral minimum of a particular guest-fullerene system is lower than the exohedral minimum the guest molecule is more energetically stable inside the fullerene cage than on the outside surface. Conversely, if the exohedral minimum is lower than the endohedral

minimum, the guest molecule is more energetically stable on the outside surface of the fullerene cage, than inside the cavity. Thus, the translational plots allow us to compare the energies of the exohedral and endohedral complexes. However, the AAP potential cannot be used to calculate the energy barrier for the insertion of a molecule into a fullerene cavity, because the AAP energy does not contain terms which allow for bond relaxation, or bonded guest-host interactions. The potential energies of the gas molecules under consideration in four C_{60} isomers and C_{70} at the endohedral and exohedral minima are in **Table 2.4**.

Table 2.4:
Potential Energies of Endohedral and Exohedral Gas-Fullerene Complexes. (kcal/mole)

Guest		BF-C_{60}	GR-C_{60}	SW1-C_{60}	SW2-C_{60}	C_{70}
H_2	endo	-4.85	-2.34	-4.81	-4.82	-4.68
	exo	-0.60	-0.53	-0.62	-0.63	-0.69
F_2	endo	-2.26	a	-4.02	-4.60	-7.96
	exo	-1.01	-0.85	-0.99	-1.01	-1.17
N_2	endo	-3.26	a	-4.02	-3.50	-8.39
	exo	-1.06	-0.91	-1.05	-1.08	-1.22
O_2	endo	-7.29	a	-8.22	-8.09	-11.10
	exo	-1.31	-1.13	-1.33	-1.35	-1.53
CO	endo	-3.04	a	-4.13	-3.67	-9.55
	exo	-1.23	-1.05	-1.21	-1.25	-1.41
HF	endo	-5.69	a	-6.10	-5.96	-6.65
	exo	-0.77	-0.67	-0.78	-0.80	-0.89
HCl	endo	-1.57	a	-2.14	-1.05	-8.68
	exo	-1.33	-1.15	-1.35	-1.37	-1.53
H_2O	endo	-8.32	a	-8.50	-8.51	-9.92
	exo	-1.21	-1.06	-1.22	-1.25	-1.39
CH_4	endo	-3.90	a	-4.62	-3.32	-11.88
	exo	-1.43	-1.21	-1.44	-1.46	-1.67
CO_2	endo	a	a	a	a	-11.75
	exo	-1.82	-1.54	-1.81	-1.83	-2.10
COH_2	endo	a	a	a	a	-4.24
	exo	-1.69	-1.41	-1.67	-1.68	-1.92

[a]Guest is unstable inside fullerene cavity

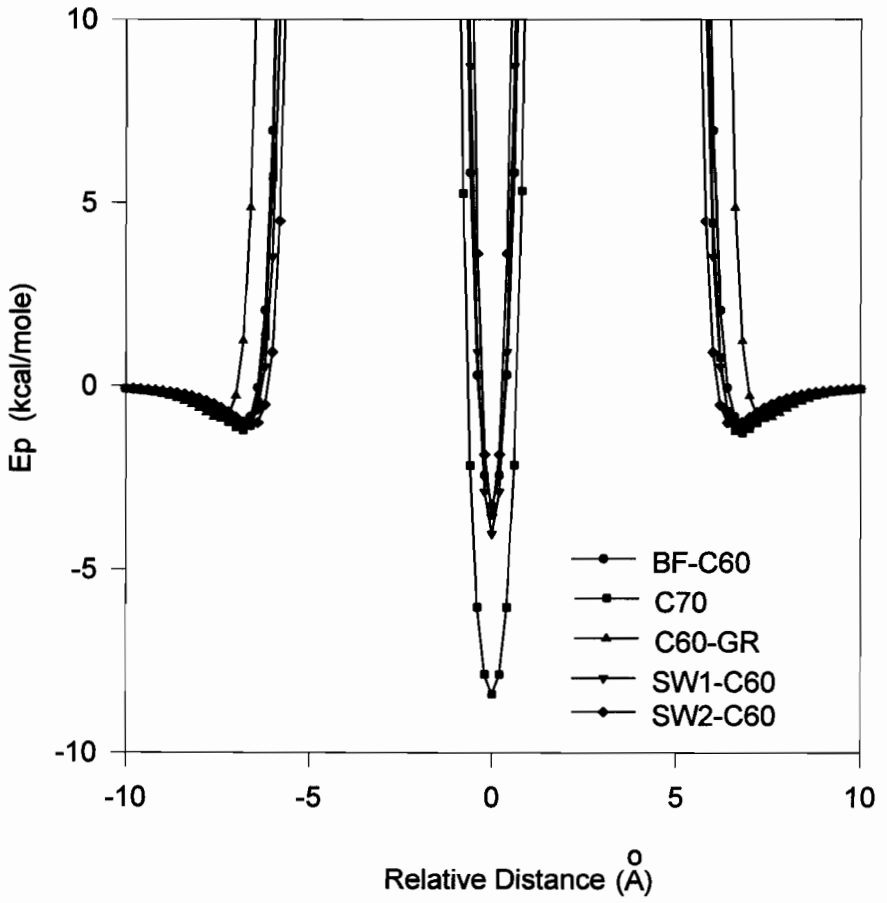


Figure 2.4: Translational Plots of N_2 in four C_{60} isomers and C_{70}

A comparison of the translational plots of the atom-atom potential energies of N_2 in I_h, D_{6h}, C_{2v} and D_{2h} isomers of C_{60} and in C_{70} are in **Fig. 2.4**. As the sizes of the fullerene cavities increase, the endohedral minima broaden and become more negative, reflecting the increased ability of the fullerne cage to accommodate the

guest molecule. Note that all the exohedral minima are at approximately the same energy. This is expected, because the van der Waals radii overlaps between the exohedrally bound guest molecule and the different fullerenes, are approximately equal, because all fullerenes that have roughly the same curvature.

To summarize the results of homonuclear diatomic-fullerene AAP interactions, the results predict that all the fullerenes except the D_{6h} isomer of C_{60} will form energetically favorable endohedral complexes with these diatomics. The cavity inside the D_{6h} isomer (GR) is too small to enclose even the small N_2, O_2, or F_2 gas molecules. However, the translational plots reveal that H_2 is sufficiently small to form a stable endohedral complex with the D_{6h} isomer. In all cases, except H_2, the C_{70} endohedral complex has a lower potential energy than any of the C_{60} complexes.

Translational plots of CO, HCl, HF, CH_4, H_2O, CO_2 and COH_2 were also generated and analyzed in **Figs. 2.5-2.6**. The AAP energies of all these guest-fullerene complexes at both the endohedral and exohedral minima are given in **Table 2.4**. All the heteronuclear diatomics were stable inside the BF-, SW1-, SW1-C_{60} and C_{70} cavities. Thus, both CO and HF are stable inside the SW2-C_{60} cavity, while HCl is 0.33 kcal/mole more stable on the outside surface of the SW2-C_{60} isomer than inside the cavity. All the heteronuclear diatomics are unstable inside the GR-C_{60} isomer cavity because of its small volume. All the heteronuclear diatomics have lower endohedral minima inside the C_{70} cavity than inside any of the C_{60} isomer cavity, because of the larger size of the C_{70} cavity. All the small molecules (three atoms or more) were unstable inside the GR-C_{60} isomer because of its small size. Only H_2O and CH_4 were stable inside the other C_{60} isomers, with H_2O having the greater stability in all cases. This was expected, since H_2O and CH_4 have smaller van der Waals lengths than CO_2 or COH_2. In C_{70}, the endohedral minima of all the small molecules were lower than the exohedral minima, indicating that the increased cavity size of C_{70} may accommodate these molecules.

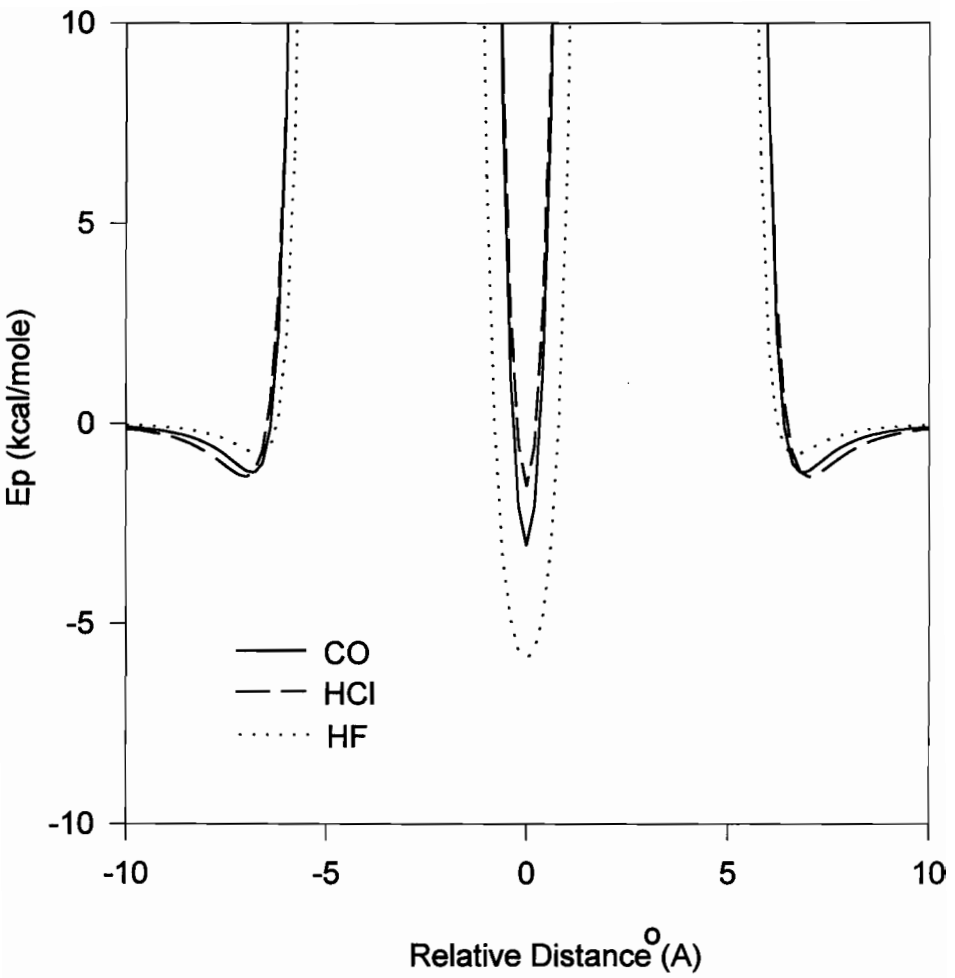


Figure 2.5: Translational Plots of CO, HF and HCl in BF-C_{60}

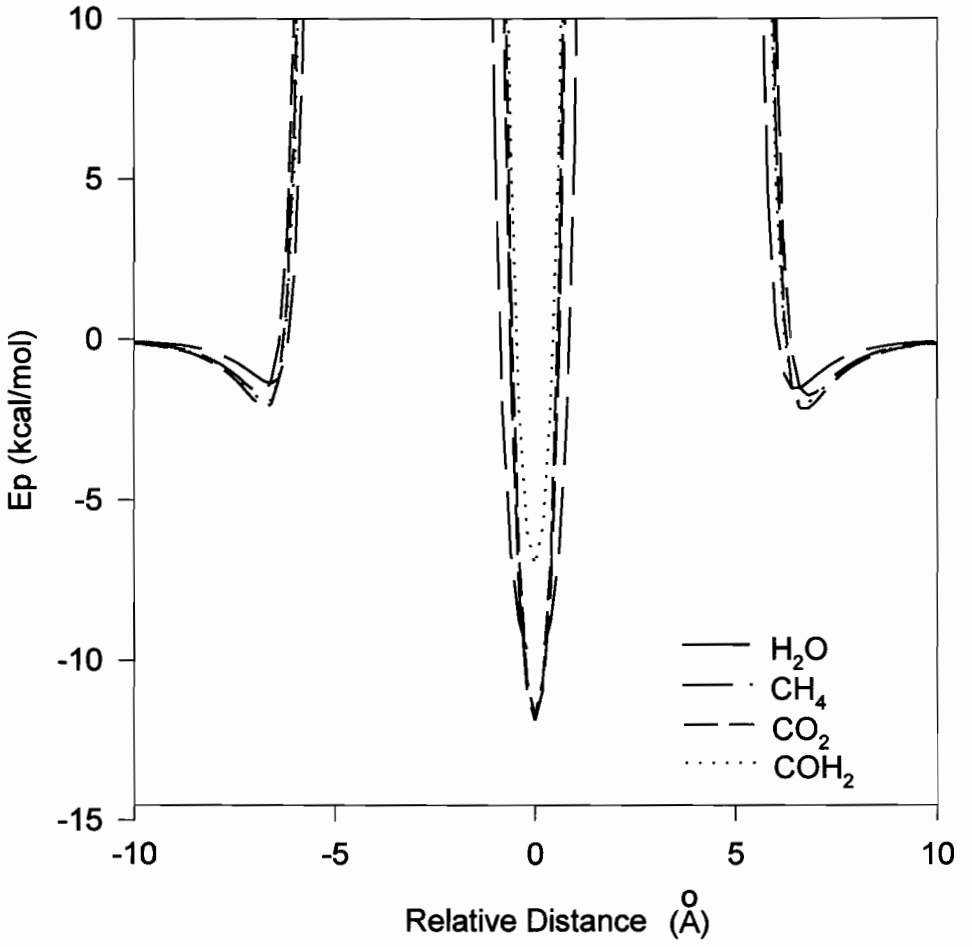


Figure 2.6: Translational Plots of Small Molecules in C_{70}

2.4.2. Rotational Plots

Plots of the endohedral minimum as a function of rotation angle are given for N_2 in the three C_{60} isomers and C_{70} in **Fig. 2.7**. These plots were obtained by calculating the AAP energy at every 2° increment while rotating the guest molecule around the y and z axes. In the SW1- and SW2-C_{60} isomers and in C_{70}, the endohedral AAP energy was found to change, as a function of rotation angle, resulting in an energy barrier to rotation. The size of the energy barrier is measured

at the maxima of the rotational plot. Since all the fullerenes (except GR-C_{60}) were defined to have their largest cavity diameters along the x axis (see **Fig. 2.1**), the guest molecules experience an increased van der Waals overlap with the fullerene cage when they are rotated to lie on either of the y or z axes. The absolute energy of the rotational barrier (kcal/mole) is found by subtracting the AAP energy at the endohedral minimum from the maximum AAP energy experienced during the rotation. It should be noted that rotation by 90° about the y axis places the guest on the z axis of the fullerene, and vice versa.

The absolute values of the rotational energy barriers experienced by the all the molecules studied are in **Table 2.5**. Note that there are no rotational barriers reported for the BF-C_{60}. It was found that all the molecules which form stable endohedral complexes with BF-C_{60} experience negligible energy changes (> 0.02 kcal/mole) when rotated around the axes of the I_h isomer. This suggests that while BF-C_{60} has icosohedral symmetry, for the purposes of the AAP energy, and to the degree in which the AAP energy may be resolved, the I_h isomer approximates spherical symmetry. The lack of rotational barrier in BF-C_{60} is exemplified by the rotational plots of homonuclear diatomics in **Fig. 2.8**. The plots show the AAP energy as approximately equal at all rotational angles. The lack of a rotational barrier was also observed in BF-C_{60} complexes with heteronuclear diatomics, water and methane.

Because of its small size, and resulting lack of van der Waals overlaps with the fullerene host, H_2 does not experience a rotational barrier in any of the fullerenes, except a small rotational barrier around the z axis in the GR-C_{60} isomer. This rotational barrier is not surprising, since the y axis cavity diameter in GR-C_{60} is only 1.2 Å, while the van der Waals length of H_2 is 3.1 Å. All the other homonuclear diatomics experience rotational barriers about both axes in SW1-, SW2-C_{60} and in C_{70}, because their van der Waals overlaps with the fullerene cavities change as a function of rotation angle.

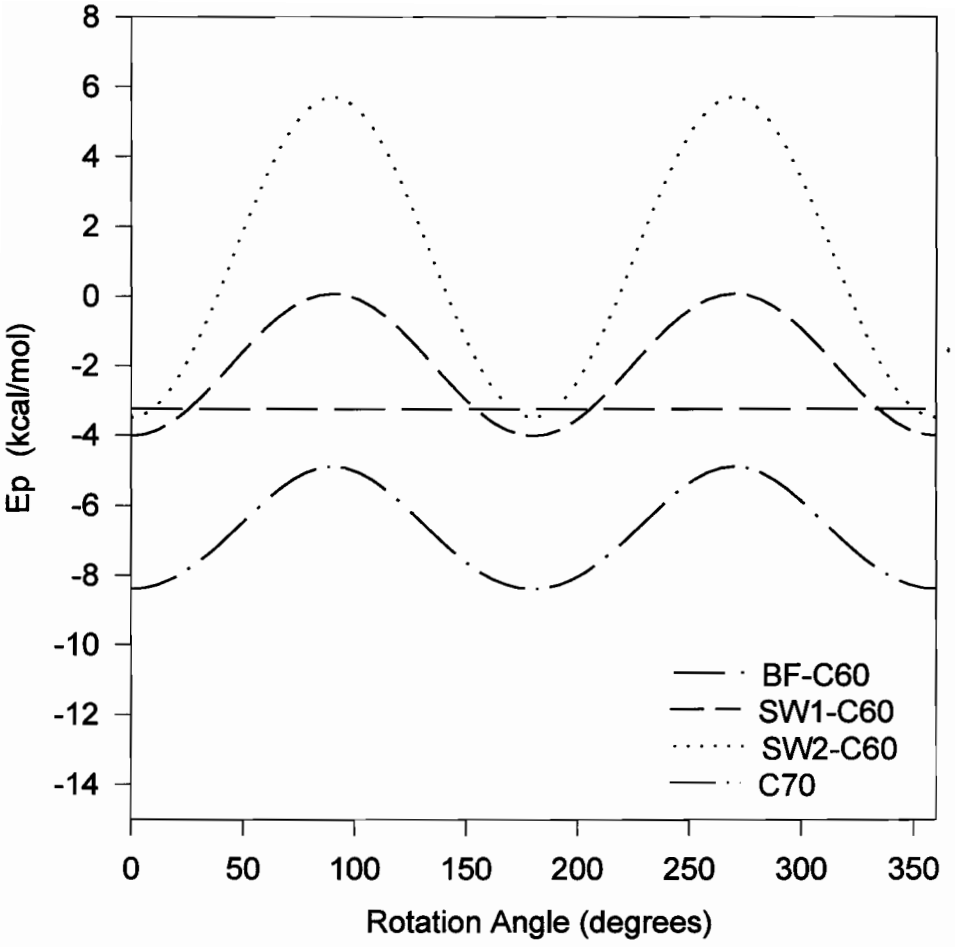


Figure 2.7: Rotational Plots of N_2 in three C_{60} isomers

Table 2.5:
Rotational Barriers in SW1-, SW2-C_{60} and C_{70} (kcal/mole) Around the y and z Axes.

Guest	SW1-C_{60}		SW2-C_{60}		C_{70}	
	y	z	y	z	y	z
H2	0.23	0.10	0.54	0.18	0.16	0.16
F2	6.07	2.68	13.91	4.66	4.97	4.96
N2	4.09	1.85	9.22	3.26	3.50	3.49
O2	4.08	1.82	9.33	3.19	3.27	3.26
CO	5.32	1.90	11.97	4.22	4.61	4.59
HF	1.17	0.38	2.72	0.88	0.91	0.89
HCl	4.44	1.59	a	a	3.86	3.64
H2O	1.12	0.22	4.73	0.77	1.69	0.75
CH4	5.95	4.34	13.26	5.15	4.72	4.71
CO2	a	a	a	a	28.51	28.49
COH2	a	a	a	a	27.44	27.46

a -Guest is unstable inside fullerene cage

All molecules experience higher rotational barriers when they are rotated to lie along the axis of the shortest fullerene cavity diameter. In the C_{70} fullerene, the y axis and the z axis rotational barriers are equal, because the y axis and z axis cavity diameters are equal (see **Fig. 2.1**). The absolute values of the rotational barriers decrease as the size of the fullerene cavity increases. This is exemplified by the reported energy barriers, and the rotational plot of homonuclear diatomics in C_{70}. Although the rotational barrier occurs at the same rotation angle in both SW2-C_{60} and C_{70}, the rotational barriers in C_{70} are much smaller. The absolute value of the rotational energy barrier also reflects the size of the guest molecule. For example, CO has the largest van der Waals length of the heteronuclear diatomics studied, and it experiences the largest rotational barriers. HF has the smallest van der Waals length of the heteronuclear diatomics, and it experiences the smallest

rotational barriers. In fact, HF represents a borderline case of hindered rotation; in the SW1- and SW2-C_{60} isomers, HF is almost free to rotate around the z axis, since the van der Waals length of HF is 3.5 Å while the y axis cavity diameters of the SW1 and SW2 isomers are both about 3.3 Å. In C_{70}, HF is again almost free to rotate, because the y and z axis cavity diameters are both 3.2 Å. Both CO and HCl experience hindered rotation in these fullerenes, because their van der Waals lengths are about 1.0 Å longer than either the y of z axes cavity diameters of these fullerenes.

Rotational plots of small molecule-fullerene endohedral complexes (exemplified by **Fig. 2.9**) revealed similar trends to those observed for the diatomics studied. Only CH_4 and H_2O were energetically stable inside the C_{60} cavities. As in the case of the diatomics, the greatest rotational barriers were experienced when the van der Waals overlaps between the guest at the host was at a maximum. H_2O had the smallest van der Waals dimensions of the small molecules studied, and it accordingly experienced the smallest rotational barriers. In the SW1-C_{60} and C_{70} fullerenes, H_2O experiences small rotational barriers, because its van der Waals dimensions are not much larger than the dimensions of the cavities. CH_4, like H_2O, is stable in SW1- SW2-C_{60} and C_{70}, but it experiences higher rotational barriers than H_2O because of its larger size. The other two small molecules studied (CO_2 and COH_2) form stable endohedral complexes with the C_{70} fullerene only. CO_2 experiences the large rotational barriers around both the y and z axes, because its x axis van der Waals length of 5.1 Å is much larger than any of the C_{70} van der Waals cavity diameters. In contrast, COH_2 experiences a large rotational barrier around the y axis only, because its x axis van der Waals length is 4.4 Å.

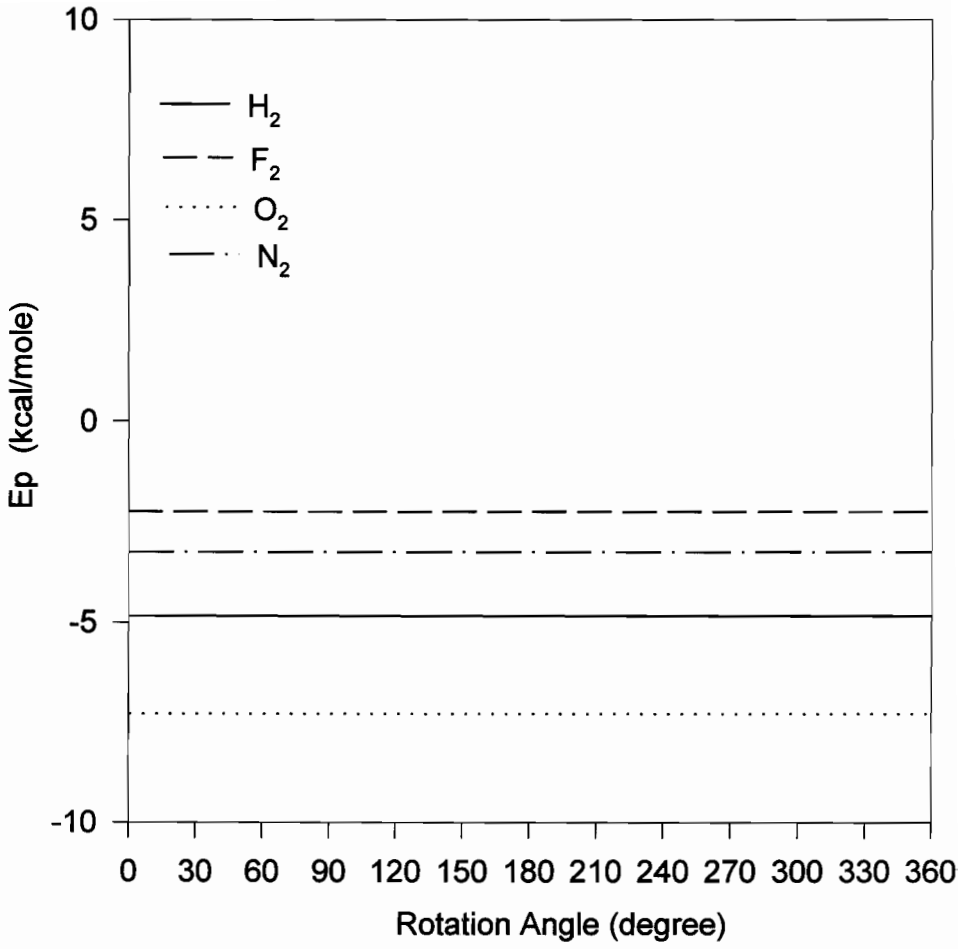


Figure 2.8: Rotational Plots of Homonuclear Diatomics in BF-C_{60}

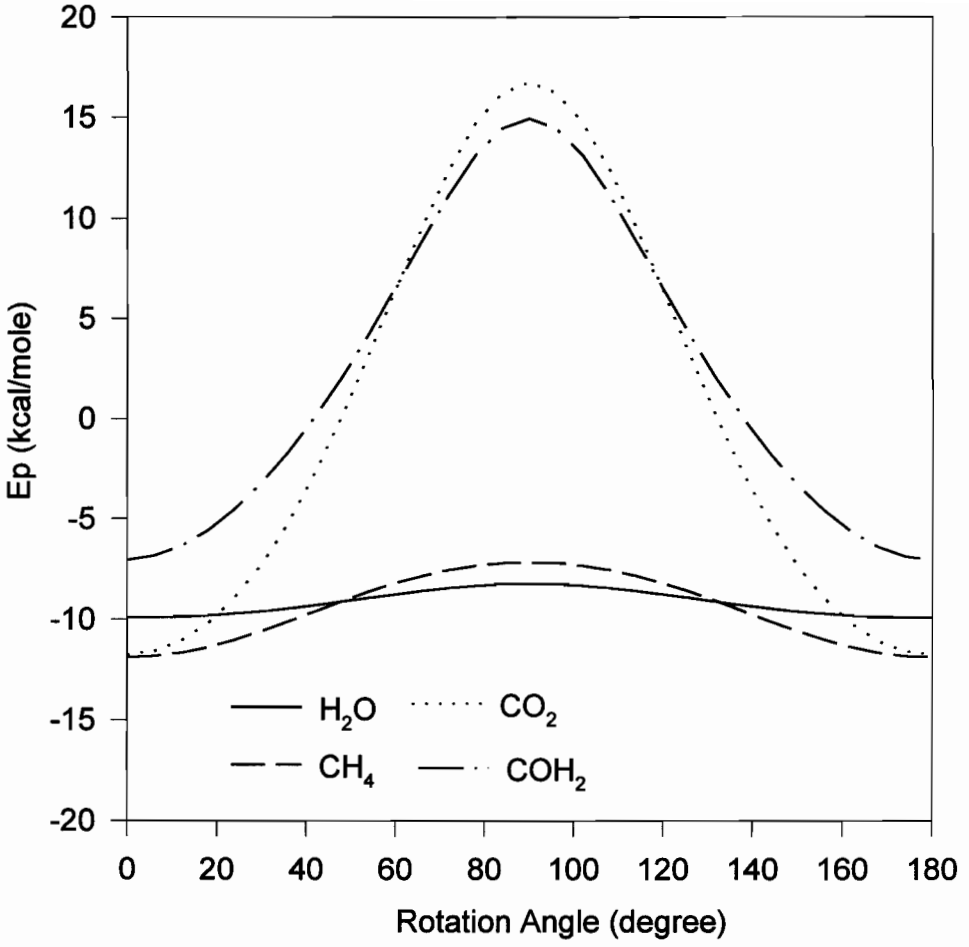


Figure 2.9: Rotational Plots of Small Molecules in C_{70}

2.4.3: Comparisons of AAP and ab initio Results:

Ab initio calculations on some endohedral gas have been performed by Cioslowsky et. al.[11] **Table 2.6** compares our results with the *ab initio* results. Although the same trend of stability is predicted by both methods, our results predict more negative complexing energies than the *ab initio* method. Possible

discrepancies between the AAP and *ab initio* methods could arise from the fact that the *ab initio* BF-C_{60} geometry yields a slightly smaller cage than that is produced by the MNDO geometry optimization. In addition, the bond lengths in the guest are not allowed to relax in our calculations, (as they are in the *ab initio* work), and the AAP potential does not contain dipole-dipole interactions between the polarizable fullerene cage and the polar guest molecules HF and CO. Nevertheless, both AAP and *ab initio* methods produce numbers with an order of magnitude which show that the formation of endohedral complexes is enegetically feasible. In some cases, the discrepancy between AAP and *ab initio* is only 4 kcal/mole. In the worst case, it is about 14 kcal/mole.

Discrepancies between the *ab initio* and AAP complexing energies may also arise from errors in the *ab initio* energies. As will be shown in the next chapter (Chapter 3), *ab initio* enthalpies of reaction can be in error by as much as 10 kcal/mol for even small systems; therefore, the *ab initio* results for fullerene/guest complexing energies may also be in error by similar quantities.

The advantage of using the AAP method over *ab initio* methods for these systems is the substantial reduction in computer CPU time. All calculations in this paper were performed in less than one hour of CPU time on a PC, as compared with the 6000 hours of workstation CPU time required to perform the *ab initio* calculations.[11]

Table 2.6:
Comparison of AAP and *ab initio* Endohedral Complexing Energies (kcal/mole) of gas/BF-C_{60} Fullerene Complexes.

Guest	AAP(this work)	*ab initio*[a]
HF	-5.69	-1.94
H2	-4.85	1.22
N2	-3.26	9.60
CO	-3.04	11.20

[a] The values are from Ref. 11.

2.5: Conclusion:

The AAP method was used to investigate the non-bonded potential energies of a series of fullerene-guest systems. The stability and the rotational dynamics of a guest molecule inside a fullerene cavity were found to reflect the van der Waals dimensions of both the guest molecule and the fullerene host. A small degree of overlap between the van der Waals radii of the fullerene host and the guest molecule introduces attractive, or stabilizing energies, while larger overlaps produce repulsive, or destabilizing energies. H_2, the smallest of the molecules studied, is stable inside all the fullerene cavities considered, and it is the only molecule that is stable inside the D_{6h} isomer of C_{60} (GR-C_{60}). All other diatomics have varying degrees of stability inside the other fullerene cavities, with the general trend suggesting that increased cavity size results in an increased stability of the endohedral complex. Of the small molecules studied (3 atoms or more), only CH_4 and H_2O appear to be stable inside any of the C_{60} isomers. Carbon dioxide and formaldehyde are too large to fit inside any of the C_{60} cavities, but they both fit inside the C_{70} cavity, when correctly oriented.

References:

1. H.R. Kroto, J.R Heath, S.C. O'Brien, R.F. Curl and Smally, R.E. *Nature*, **318**, 162 (1985).

2. J.E. Fischer, P.A. Heiney and A.B. Smith, Acc. Chem. Res., **25**, 112 (1992).

3. R.C. Haddon, *Acc. Chem. Res.*, **25**, 127 (1992).

4. J.R. Heath, S.C. O'Brien, Q. Zhang, Y. Lin, R.F. Curl, H.R. Kroto, F.K. Tittel, and R.E. Smalley, *J. Am. Chem. Soc.*, **107**, 7779 (1985).

5. F.D. Weiss, J.L. Elkind, S.C. O'Brien, R.F. Curl, and R.E. Smalley, *J. Am Chem. Soc.*, **110**, 4464 (1988).

6. D.M. Cox, D.J. Trevor, K.C. Reckmann and A. Kaldor, *J. Am. Chem. Soc.*, **108**, 2457 (1986).

7. T. Weiske, D.K. Bohme, J. Hrusak, W. Kratschmer H. and Schwarz, *Agnew. Chem., Int. Ed. Engl.*, **30**, 884 (1991).

8. M.M. Ross and J.H. Callahan, *J. Phys. Chem.*, **95**, 5720 (1991).

9. M. Saunders, H.A. Jimenez-Vazquez, R.J. Cross, and R.J. Poreda, *Science*, **259**, 1428 (1993).

10. J. Cioslowski and E.D. Fleischman, *J. Chem. Phys.*, **94**, 3730 (1991).

11. J. Cioslowski, *J. Am. Chem. Soc.*, **113**, 4139 (1991).

12. A.L.R. Bug, A. Wilson and G.A. Voth, *J. Phys. Chem.*, **96**, 7864 (1992).

13. P.P. Schmidt, B.I. Dunlap and C.T. White, J. Phys. Chem., **95**, 10537 (1991).

14. L. Pang and M.A. Whitehead, *Supramolecular Chemistry*, **1**, 81 (1992).

15. L. Pang and M.A. Whitehead, *J. Mol. Struct. (THEOCHEM)*, **257**, 143 (1992).

16. R.E. Stanton, *J. Phys. Chem.*, **96**, 111 (1992).

17. M. Froimowitz, *J. Comp. Chem.*, **12**, 1129 (1991).

18. A. Gavezzotti, *Nouveau Journal de Chimie*, **6**, 443 (1982).

19. A.J. Pertsin and A.I. Kitaigorodsky, "*The Atom-Atom Potential Method*" Springer-Verlag: New York 1987.

20. K. Mirsky, *"Proceedings of an International School on Crystallographic Computing*" Delft University Press: Twente 1978.

21. L.E. Sutton, (ed) *"Tables of Interatomic Distances and Configurations in Molecules and Ions*" The Chemical Society: London 1965.

22. R.C. Weast and S.M. Selby, "*The Handbook of Chemistry and Physics*" (Chemical Rubber Company, Cleavland, Ohio) 1992.

23. R.C. Mowrey M.M. Ross and J.H. Callahan *J. Phys. Chem.*, **96**, 4755 (1992).

The energies, structures and reactivities of nitrogen heterocyclic compounds are the research focus of this thesis. In chapter two, it was shown that simple models (i.e. the AAP model) can often give results which correlate well with high level *ab initio* results. However, the AAP method in chapter two does not treat chemical bonds, and cannot be used to study nitrogen heterocyclic energies and reactivities. Quantum mechanical methods are the only methods suited to compute these properties. The two most common quantum mechanical methods, semi-empirical and *ab initio*, are used in this chapter to generate nitrogen heterocyclic heats of formation. *Ab initio* methods, although in principle superior to semi-empirical, will be shown to suffer from basis set and correlation defects. Accurate nitrogen heterocyclic heats of formation can be obtained using *ab initio* total energies, but only by cancelling out basis set and correlation errors by averaging over different basis sets and a number of different *isolobal* reaction schemes. Semi-empirical methods are much less computationally intensive than *ab initio,* but they will be shown to yield qualitatively correct nitrogen heterocyclic heats of formation. Simple correction terms are developed to increase the accuracy of semi-empirical calculated heats of formation. Most importantly, the PM3 semi-empirical method will be shown to give the most accurate nitrogen heterocyclic heats of formation of all the semi-empirical methods considered. The accuracy of PM3 with nitrogen heterocyclics will justify its use in forthcoming chapters.

CHAPTER 3:

AROMATIC NITROGEN HETEROCYCLIC HEATS OF FORMATION: A COMPARISON OF SEMI-EMPIRICAL AND *AB INITIO* TREATMENTS

by C.I. Williams and M.A. Whitehead,

accepted *J. Mol. Struct. (THEOCHEM)*, October 1996.

3.1 Abstract:

Ab initio total energies and isolobal reaction schemes are used to calculate azole, azine and benzoazole heats of formation to within 0-4 kcal/mol of experiment. The *ab initio* heats of formation are compared with calculated semi-empirical (MNDO, AM1, PM3) heats of formation. The MNDO method systematically underestimates all nitrogen heterocyclic heats of formation. The AM1 method overestimates azole, and underestimates azine heats of formation. The PM3 method performs well on azoles and benzoazoles, but underestimates azine heats of formation. Overall, the PM3 method is the most accurate semi-empirical method for calculating these nitrogen heterocyclic heats of formation. Correction terms for the semi-empirical azine heats of formation are suggested, which bring semi-empirical values into agreement with experiment. Semi-empirical and *ab initio* methods are used to predict pentazole, pentazine, hexazine, pyrazepine, indazine, purine and azolotriazine heats of formation.

KEYWORDS: isolobal, ab initio, semi-empirical, azoles, azines

3.2 Introduction:

Heats of formation can be obtained experimentally, but when a compound is unstable or difficult to purify, experimental heats of formation become increasingly difficult to measure. For example, experimental heats of formation exist for azoles and azines (**Figs. 3.1** and **3.2**) with up to three nitrogen atoms, but little data exists for azoles with four or more nitrogen atoms. No experimental heats of formation exist for azines with four or more nitrogen atoms because of their instability.[1,2]

The numbering schemes in **Figs. 3.1-3.4** reflect that azoles, azines, benzoazoles and purines (**Figs. 3.1-3.4**) are similar in structure to cyclopentadiene, indene and benzene, except that N or N-H groups have been successively substituted for C-H and CH_2 groups. The first digit in the molecule number gives the total number of atoms in the ring, while the second digit gives the number of nitrogen heteroatoms. For example, benzene is a six-memberd ring with no N

heteroatoms, so it is numbered **60**. Pyridine, a six-membered ring with one N heteroatom, is numbered **61**. Letters after the numbers differentiate structural isomers, such as imidazole **52a** and pyrazole **52b,** which are both structural isomers of a five-membered ring with two nitrogen heteroatoms. Benzoazoles are indicated with a '**B**' before the number reflecting the azole ring. For example, indole is benzo-fused pyrrole (**51**), so indole is numbered **B51.** This numbering scheme was introduced to facilitate balancing the isolobal reaction schemes used to generate the *ab initio* heats of formation. The purines are numbered **P1** and **P2** for simplicity.

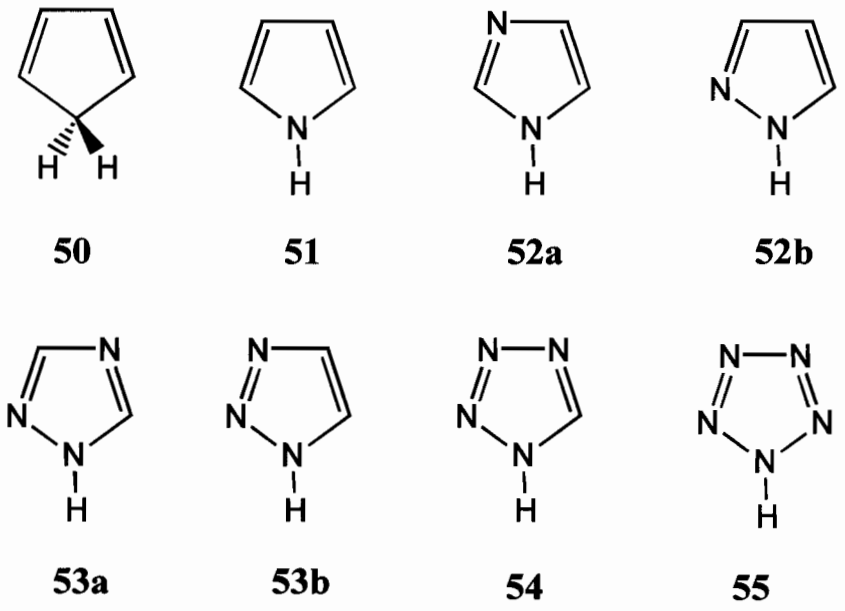


Figure 3.1: Cyclopentadiene and Related Azole Ring Systems

The numbering scheme is meant to reflect the structure of the azoles; the first digit indicates the total number of atoms in the ring, while the second digit indicates the number of nitrogen heteroatoms. Letters after the numbers differentiate structural isomers. For example, triazoles are 5-membered rings with 3 N heteroatoms, so they are numbered **53**; **53a** and **53b** refer to the 1,2,4- and the 1,2,3-triazole isomers respectively.

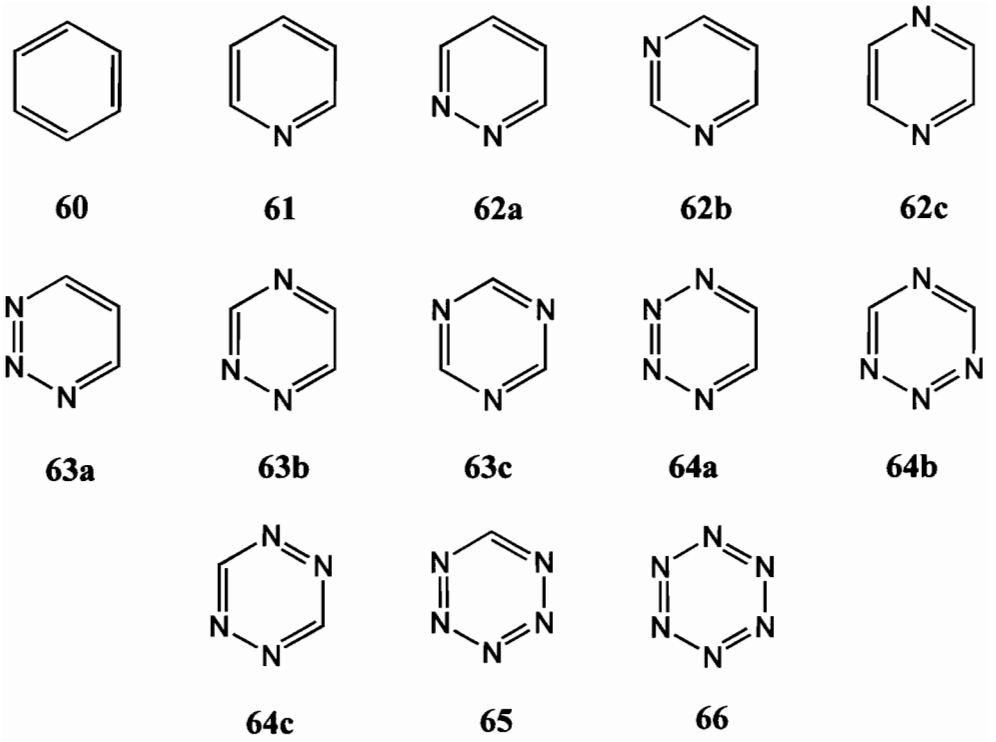


Figure 3.2: Benzene and Related Azines:

(see text and caption of **Figure 1** for explination of numbering scheme)

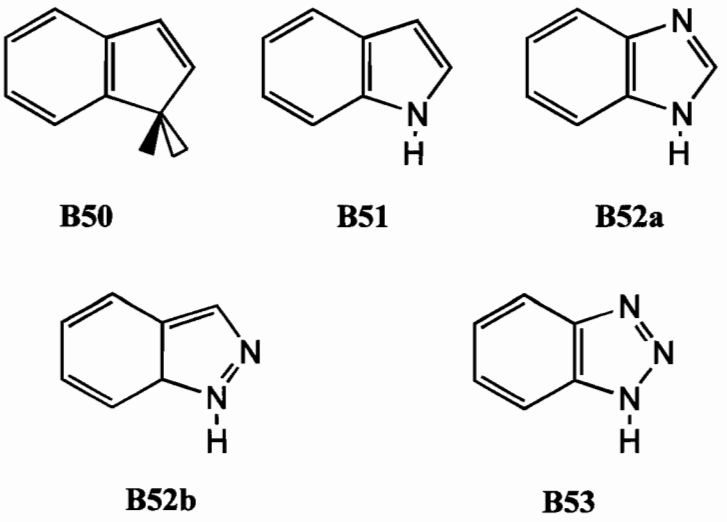


Figure 3.3 Indene and Related Benzoazoles:

(see text and caption of **Figure 1** for explination of numbering scheme)

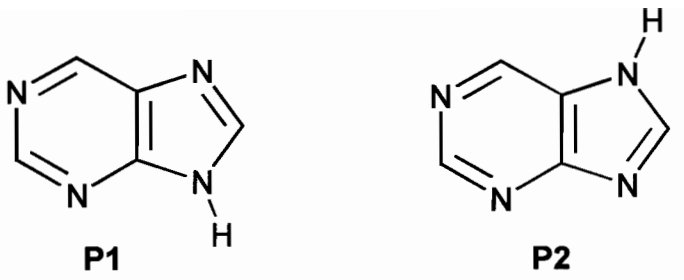


Figure 3.4: Purines

Because the experimental heats of formation are either difficult or impossible to obtain for some azoles and azines, their heats of formation will be calculated using both semi-empirical and *ab initio* methods. Well-established semi-empirical methods (MNDO, AM1 and PM3) automatically return heats of formations from their calculations (see chapter 1). *Ab initio* calculations return only total energies, so *ab initio* heats of formation must be calculated by some method which uses the *ab initio* total energies. One approach to calculating the *ab initio* heat of formation of a target molecule (X) is to formulate a hypothetical chemical reaction that includes the target molecule (X) and other molecules whose experimental heats of formation are known (A, B, C). The *ab initio* energy for the reaction scheme (E_{rxn}) is then calculated, and used with the known heats of formation to solve the target molecule heat of formation ($\Delta H_f X$) using Hess's law.

(1) $$A + B \rightarrow C + X$$

$$\Delta H_f X = E_{rxn} + \Delta H_f A + \Delta H_f B - \Delta H_f C$$

This approach takes advantage of experimental heats of formation. When the experimental heats of formation are accurate, the errors in the *ab initio* heats of formation calculated by this method arise solely from errors in the *ab initio* heats of

reaction. The *ab initio* reaction energy (E_{rxn}) for a given scheme and temperature (T) is

(2) $$E_{rxn} = (\Sigma E_p - \Sigma E_r) + (\Sigma ZPE_p - \Sigma ZPE_r) + (\Sigma TC_p - \Sigma TC_r) + 4(m - n)RT$$

where (E_p , E_r) are the total electronic energies of the products and reactants, (ZPE_p,ZPE_r) are the product and reactant zero-point energies, and (TC_p ,TC_r) are the product and reactant thermal corrections. The final $4(m - n)RT$ term in equation (2) is the translational contribution; here m and n are the product and reactant stoichiometric coefficients.

Calculation of E_{rxn} by eq. 2 is computationally costly, because zero-point energies and thermal correction must be calculated. However, equation (2) can be simplified; when the number of vibrational modes in the reactants and product is the same, and when the reactants and products are chemically similar, the zero-point energies and thermal corrections will almost cancel.[7] Also, when the number of molecules is the same on both sides of the reaction scheme ($m = n$), the translational contribution cancels. The simplified expression for E_{rxn} is

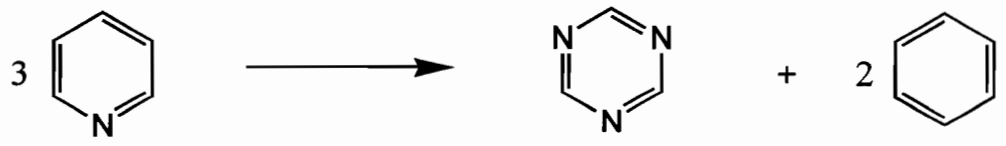
(3) $$E_{rxn} = (\Sigma E_p - \Sigma E_r).$$

Equation (3) is less accurate than equation (2), but it is more useful because it requires the calculation of *ab initio* total electronic energies only. An abundance of *ab initio* total energies already exist in the literature, so using equation (3) to calculate a desired heat of reaction will in many cases require only one or two *ab initio* calculations. However, the accuracy of heats of formation calculated with these heats of reaction will depend upon cancellation of the zero-point energies and thermal correction terms in equation (2), which will only occur if the product and reactant molecules are similar. The accuracy of *ab initio* reaction energies calculated by equation (3) will also depend upon cancellation of correlation energies between reactants and products. Thus, the main drawback to this approach is that the accuracy of the results depends upon the judicious choice of a reaction scheme. Therefore, isodesmic,[3] homodesmic,[4] isogyric[5,6] and isolobal[7] reaction

schemes, all discussed in the literature,[3-7] have been designed to cancel correlation effects, and in some cases, ZPE and TC terms.

Isodesmic reaction schemes (reactions where products and reactants have the same number, but different arrangements, of formal bond types) are the most commonly used type of reaction scheme for computing *ab initio* enthalpies of formation.[8,9] Isodesmic reaction schemes work well because the correlation energies of electrons in like bonds partially cancel, giving accurate heats of reaction and heats of formation even at the Hartree-Fock level.[8,9] However, isodesmic schemes do not give accurate results when aromatic and non-aromatic molecules are in the same reaction, because the formal bond count which defines an isodesmic reaction is ambiguous in aromatic systems with many resonance structures.

In this chapter, azole, azine, benzoazole and purine *ab initio* heats of formation will be calculated with *isolobal* reaction schemes. Isolobal reaction schemes have recently been used to successfully evaluate heats of formation of thienopyridine isomers,[7] so their use here is justified. *Isolobal* reaction schemes can be used to increase the number of nitrogen atoms in the a ring systems by successively replacing C-H groups with nitrogen atoms. For example, 1,3,5-triazine (**63c**) can be generated from pyridine (**Fig. 3.5**). The similarity of these molecules allows use of equation (3) to calculate the *ab initio* heat of this reaction.



61 **63c** **60**

Figure 3.5: *Isolobal* Reaction for Generation of 1,3,5,-Triazine from Pyridine

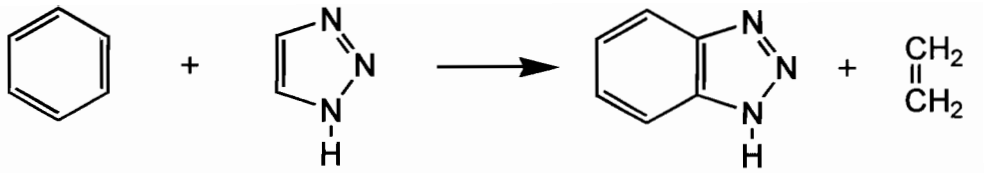
(Note that the numbering scheme can be used to check that the isolobal reaction is balanced; The reaction can be written as 3 (**61**) → **63** + 2 (**60**); It is balanced because 3(**61**) = **63** + 2(**60**) = 183.)

To increase the accuracy of the isolobal reaction energies, the following conditions will be imposed on the isolobal reaction schemes:

a) the number of five, six and seven membered ring systems must be the same on both sides of the reaction.
b) the number of localized and delocalized systems must be the same on both sides of the reaction, and
c) the number of N-C, N-N, C-C, C-H and N-H linkages must be the same on both sides of the reaction.

Condition (a) partially cancels correlation effects caused by ring systems, while condition (b) partially cancels correlation effects caused by localized and delocalized systems. Condition (c) cancels effects caused by specific atom-atom linkages. These conditions all help to cancel the TC and ZPE terms from equation (1), allowing use of equation (3) for the *ab initio* reaction energy. The equivalence of formal bond types is <u>not</u> a condition, because counting formal bonds is ambiguous in delocalized ring systems. A few reaction schemes which did not satisfy all the above conditions were also included to demonstrate the poor quality of the results obtained when these conditions are relaxed.

Benzofusion reaction schemes, which join two ring systems by elimination of an ethene molecule (**Fig. 3.6**), will be used in addition to *isolobal* schemes to calculate *ab initio* heats of formation of benzofused compounds. Benzofusion reaction schemes have been successfully used to calculate heats of formation of a number of fused ring systems.[4,7] Benzoazole and purine heats of formation will be generated by both isolobal and benzofusion reaction schemes.



60 **53b** **B53**

Figure 3.6: Benzofusion Reaction Scheme

In the absence of experimental data, the quality of the calculated heats of formation will be assesed by comparison with semi-empirical (MNDO, PM3 and AM1) heats of formation. Systematic errors in semi-empirical heats of formation will be discussed, and correction terms to semi-empirical heats of formation will be suggested. Semi-empirical and *ab initio* methods will be used to predict azolotriazine (**Fig. 3.8**) and pyazepine (**Fig. 3.9**) heats of formations.

3.3 Method:

All calculations were performed on a SUN Sparc10 workstation using the GAMESS[10] package. Full geometry optimizations were allowed, and RHF wavefunctions used. Experimental geometries (when available) were input; otherwise, geometries were created by substituting nitrogen atoms into existing geometries. Equilibrium geometries were verified as local minima by force constant analysis. *Ab initio* calculations used 4-31G and 6-31G** basis sets. *Ab initio* total energies are in **Table 3.A1 (Appendix)**.

Isolobal reaction schemes: Suitable isolobal reaction schemes were devised for each target molecule. Reaction schemes were written so that the target molecule would be a product. The heats of reaction and subsequent heats of formation were calculated for each reaction scheme using both the 4-31G and 6-31G** basis sets. The 4-31G and 6-31G** heats of reaction and the resulting target molecule heats of formation are all included for comparison (Appendix **Tables 3.A2-3.A5**). Errors in *ab initio* heats of formation which arise from the choice of reaction schemes were reduced by averaging the heat of formation over all reaction schemes. By Hess's law,[11] all the reaction schemes could be added together to give a 'master' reaction

scheme, so averaging the *ab initio* heats of formation is equivalent to calculating the heat of formation from the 'master' reaction scheme. The process gives better results, because it includes more molecules with known heats of formation than a single reaction scheme, and partially eliminates errors in the heats of formation caused by peculiarities in each reaction scheme. A few reaction schemes which do not satisfy all of the requirements (a-c) (asterixed '*') are included for comparison. Reaction schemes whose heats of formation are not included in the averaging process are tagged with a question mark " ? ".

3.4 Results and Discussion:

3.4.1. Azoles: Azoles have been extensively studied with semi-empirical and *ab initio* methods,[12-15] and, except for pentazole and the triazoles, experimental gas phase heats of formation are well known. Therefore, this series of molecules is perfect for testing the accuracy of the heats of formation generated by isolobal reaction schemes. The 4-31G and 6-31G** *ab initio* heats of formation, generated by isolobal reaction schemes, are compared with semi-empirical and experimental azole heats of formation in **Table 3.1**.

The *ab initio* reaction schemes and heats of reaction used to generate azole heats of formation are in **Table 3.A2** (Appendix). For any given isolobal reaction scheme, the 4-31G and 6-31G** reaction energies are within 8 kcal/mol of each other. This similarity in the 4-31G and 6-31G** reaction energies indicates that correlation and basis set effects partially cancel out by using these reaction schemes. In contrast, the 4-31G and 6-31G** reaction energies of the <u>*non-isolobal*</u> reaction cyclopentadiene + pyridine → pyrrole + benzene differ by 42 kcal/mol, and have opposite signs! Correlation energy errors are not effectively cancelled in this reaction scheme, because this scheme is neither isodesmic, nor does it satisfy conditions (a), (b) and (c). This reaction is included as an example of a poor choice of reaction scheme for calculating *ab initio* azole heats of formation. The isomerization reactions (**52a** → **52b** and **53a** → **53b**) are also poor choices, because

they do not satisfy all of the isolobal conditions; heats of formation derived from them are less accurate than those derived from reaction schemes which satisfy the isolobal conditions.

TABLE 3.1:
Semi-Empirical, *Ab initio* and Experimental Azole Heats of Formation:

Compound	ΔH_f (kcal/mol)						
	semi-empirical			This work			
	MNDO	AM1	PM3	4-31G	6-31G**	A	expt.
50	32.1	37.1	31.8	--	--	--	31.9[a]
51	32.5	39.9	27.1	29.8	26.6	28.2	26.0[a]
52a	33.3	50.8	31.3	32.8	37.3	35.1	35.0[a]
52b	45.4	65.6	48.8	43.2	41.9	42.6	44.0[a]
53a	43.7	76.9	51.8	49.8	42.2	45.7	46.0[b]
53b	49.8	86.4	67.9	64.4	63.4	63.9[c]	
54	53.8	109.7	86.2	84.1	77.7	80.9	80.1[a]
55	65.8	146.9	120.7	115.0	106.2	110.6	--

A = average of 4-31G and 6-31G** values;
a = ref. 14; b = estimated experimental (ref. 15);
c = value used as 'experimental' value in subsequent computations

Heats of formation generated from individual reaction schemes can be in error by as much as 10 kcal/mol, (**Table 3.A2**) emphasizing the fact that *ab initio* reaction energies can be in error by 10 kcal/mole or more. However, averaging over all the reaction schemes within each basis set (4-31G and 6-31G**) gives azole heats of formation which differ from experiment by 1-5 kcal/mol. This is better than the azole heats of formation calculated from one reaction scheme alone. Furthermore, averaging the 4-31G and 6-31G** heats of formation gives azole heats of formation remarkably close to experimental results for this modest level of theory (absolute errors ranging from 0.3 to 2.1 kcal/mol). The averaging process improves these results by reducing errors due to basis sets and electron correlation.

The PM3 method gives the best semi-empirical azole heats of formation. AM1 consistently overestimates azole heats of formation by 20-30 kcal/mol, while MNDO performs well on pyrrole and imidazole, but underestimates heats of formation of azoles with three or more nitrogen atoms. The errors in MNDO and AM1 nitrogen-heterocyclic heats of formation will be discussed again in Chapter Five. Corrections to MNDO and AM1 values have been suggested by other workers,[11-15] but the close fit between PM3 and experimental values suggests that PM3 is the most trustworthy semi-empirical method for azoles.

Because of the good fit obtained with the other azoles, the *ab initio* 1,2,3-triazole heat of formation is probably quite accurate. The average of the 4-31G and 6-31G** 1,2,3-triazole heats of formation (63.9 kcal/mol) will be used as an 'experimental' value in subsequent isolobal reaction schemes involving 1,2,3-triazole. The PM3 and isolobal *ab initio* heats of formation are in good agreement for pentazole, although the PM3 value is 10 kcal/mol higher than the *ab initio* value. However, since PM3 triazole and tetrazole heats are 6-7 kcal/mol higher than experimental values, a high value for the PM3 pentazole heat of formation is expected. Consequently, the *ab initio* value of 110.6 kcal/mol is the best estimate for the pentazole heat of formation from this work.

3.4.2. Azines: Azines have been extensively studied with semi-empirical and *ab initio* methods,[11-13] and experimental gas phase heats of formation have been reported for azines with up to three nitrogen heteroatoms. All semi-empirical methods perform well on pyridine, but underestimate other azine heats of formation by amounts which increase with both the number of nitrogen atoms in the azine ring, and the number of N-N linkages. Previous workers have suggested corrections to MNDO azine heats of formation, and this approach is now extended by suggesting corrections for AM1 and PM3 azine heats of formation. The correction terms (outlined in **Table 3.2**) bring the semi-empirical heats of formation in line with experimental values.

The isolobal reaction schemes and *ab initio* reaction energies used to calculate the *ab initio* azine heats of formation are in **Table 3.A3 (Appendix)**. The 6-31G** and 4-31G reaction energies are all within 10 kcal/mol of each other for any given reaction scheme, and the difference between the 4-31 and 6-31G** reaction energies is often only 2-4 kcal/mol. This suggests that the chosen reaction schemes successfully cancel basis set and correlation errors.

TABLE 3.2: Semi-Empirical, Corrected Semi-Empirical* and Experimental Azine Heats of Formation:

	ΔH_f(kcal/mol)*						
	MNDO		AM1		PM3		expt.
60	21.3	--	22.0		23.5		19.8[a]
61	28.1	(33.3)	32.0	(32.0)	30.4	(34.0)	34.6[a]
62a	43.6	(66.5)	55.3	(66.5)	56.0	(66.5)	66.5[a]
62b	35.0	(45.4)	43.9	(43.9)	38.0	(45.2)	47.0[a]
62c	37.8	(48.2)	44.2	(44.2)	39.3	(46.5)	46.9[a]
63a	55.7	(101.5)	81.6	(104.0)	81.7	(102.7)	--
63b	51.5	(79.6)	69.5	(80.7)	65.8	(79.9)	--
63c	40.0	(55.6)	58.0	(58.0)	46.2	(57.0)	53.9[b]
64a	70.2	(138.9)	110.5	(144.1)	110.1	(141.6)	--
64b	62.8	(113.8)	98.1	(120.5)	92.2	(116.8)	--
64c	67.7	(112.6)	97.5	(119.9)	94.6	(115.6)	--
65	84.3	(175.9)	142.0	(186.8)	139.2	(181.2)	
66	104.4	(241.8)	189.8	(257.0)	186.8	(249.8)	

a = ref. 14 b = ref. 16 *(corrected values in parentheses)

Corrections for semi-empirical azine heats of formation are calculated as follows:

-correction for C-N-C = (δ(**61**) + δ(**62b**) + δ(**62c**) + δ(**63c**)) / 8
-correction for N-N = δ(**62**); where δ(X) = ΔH_f X (expt.) - ΔH_f (semi-empirical)

		MNDO	AM1	PM3
Corrections: (kcal/mol)	per N-N	+22.9	+11.2	+10.5
	per C-N-C	+5.2	0.0	+3.6

Isolobal *ab initio* azine heats of formation agree well with experiment (**Table 3.3**). As with azoles, averaging the 4-31G and 6-31G** values improves the quality of the *ab initio* results, giving azine heats of formation which differ from experimental values by 1-4 kcal/mol. Where experimental values are not available (i.e. the tetrazines), corrected semi-empirical and *ab initio* azine heats of formation agree well, suggesting that the *ab initio* and corrected semi-empirical azine heats of formation are accurate and predictive. Corrected semi-empirical and *ab initio* azine heats of formation agree up to the tetrazines, but with pentazine and hexazine, the corrected semi-empirical heats of formation are higher than the *ab initio* results. This is probably because the semi-empirical correction term for the N-N link is too large; however, a better correction term could not be calculated because there is only one experimental heat of formation for an azine with an N-N linkage (1,2-diazine).

TABLE 3.3: Corrected Semi-empirical, *Ab initio* and Experimental Azine Heats of Formation:

	ΔH_f (kcal/mol)				
	corrected* semi-empirical			*ab initio* (This work)	expt.
	MNDO	AM1	PM3	A	
61	33.3	32.0	34.0	34.8	34.6[a]
62a	66.5	66.5	66.5	68.2	66.5[a]
62b	45.4	43.9	45.2	44.5	47.2[a]
62c	48.2	44.2	46.5	51.0	46.9[a]
63a	101.5	104.0	102.7	98.2	--
63b	79.6	80.7	79.9	82.0	--
63c	55.6	58.0	57.0	54.5	53.9[b]
64a	138.9	144.1	141.6	133.5	--
64b	113.8	120.5	116.8	111.8	--
64c	112.6	119.9	115.6	117.0	--
65	175.9	186.8	181.2	166.5	--
66	241.8	257.0	249.8	223.1	--

A = average of 4-31G and 6-31G** values (see **Table 3.A3**).
a = ref. 14 b = ref. 16 * see **Table 3.2** for semi-empirical correction terms

3.4.3 Benzo-Fused Compounds and Purines: *Ab initio* heats of formation of benzoazoles and purines were calculated with both isolobal and benzofusion reaction schemes. Reaction schemes, heats of reaction and subsequent heats of formation are in **Table 3.A4 (Appendix)**. The isolobal and benzofusion 4-31G and 6-31G** reaction energies are very close to each other, indicating that basis set and electron correlation errors almost completely cancel; therefore, these are excellent schemes for heats of formation calculations. Semi-empirical and *ab initio* benzoazole and purine heats of formation are compared with available experimental data in **Table 3.4**. The *ab initio* heats of formation are the best fit to experiment. PM3, which performs well on azoles, also performs well on benzoazoles. Errors in MNDO and AM1 benzoazole heats of formation are similar to MNDO and AM1 errors in azole heats of formation.

TABLE 3.4: Semi-Empirical and *Ab initio* Heats of Formation of Benzofused Azoles:

	ΔH_f (kcal/mol)				
	semi-empirical			This work *ab initio*	
Compound	MNDO	AM1	PM3	A	expt.
B50 Indene	38.3	45.8	40.7	37.5	39.0[a]
B51 Indole	44.2	55.2	42.6	40.8	37.4[a]
B52a Benzoimidazole	45.9	67.1	47.8	48.9	--
B52b Benzopyrazole	58.4	82.3	65.7	62.4	--
B53 Benzotriazole	64.0	104.2	85.6	81.1	83.0[a]
P1 Purine (1)	59.5	90.0	62.3	68.3	--
	69.9	--	*69.5*		
P2 Purine(2)	64.0	94.0	65.8	73.2	--
	74.4	--	*73.0*		

A = average of 4-31G and 6-31G** values (see **Table 3.A4**). a = ref.16
italics = corrected semi-empirical (see **Table 3.2** for correction terms)

Semi-empirical and *ab initio* purine heats of formation are not in good agreement because purines contain an azine moiety. However, by applying the previously developed semi-empirical azine heat of formation correction terms (see **Table 3.2**) to MNDO and PM3 purine heats of formation improves the MNDO and PM3 results. Corrected PM3 and MNDO purine heats of formation agree well with the *ab initio* purine heats of formation. Azine correction terms do not help the AM1 purine heats of formation, because purines contain azole rings which are poorly predicted by AM1. Interestingly, AM1 overestimates purine heats of formation by 20-30 kcal/mol, which is similar to the AM1 overestimation of azole heats of formation.

3.4.4. Assessment of Method: The scatter plot in **Fig. 3.7** compares the isolobal *ab initio* heats of formation with experimental values. The close fit between experiment and isolobal *ab initio* heats of formation over the range of compounds studied shows that the *ab initio* values are trustworthy, and can reasonably be applied to predict heats of formation of systems similar to these.

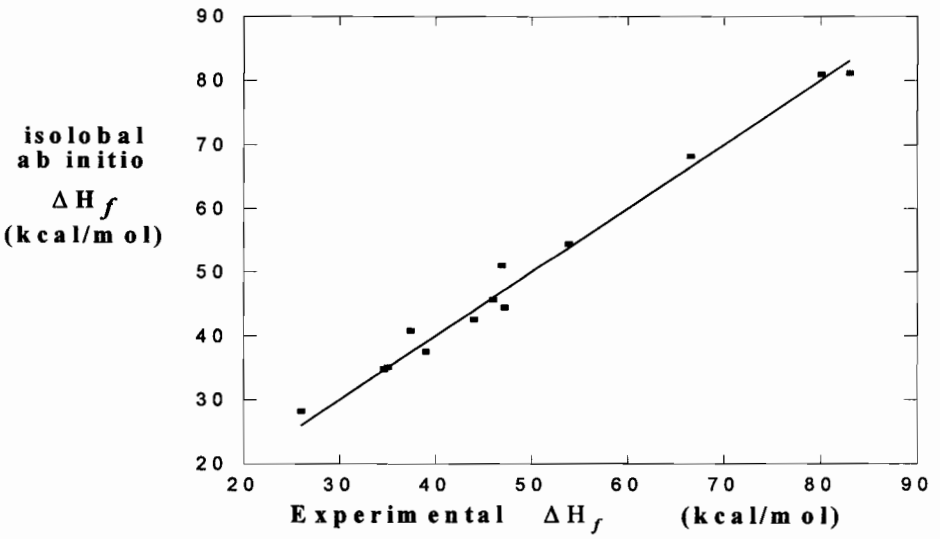
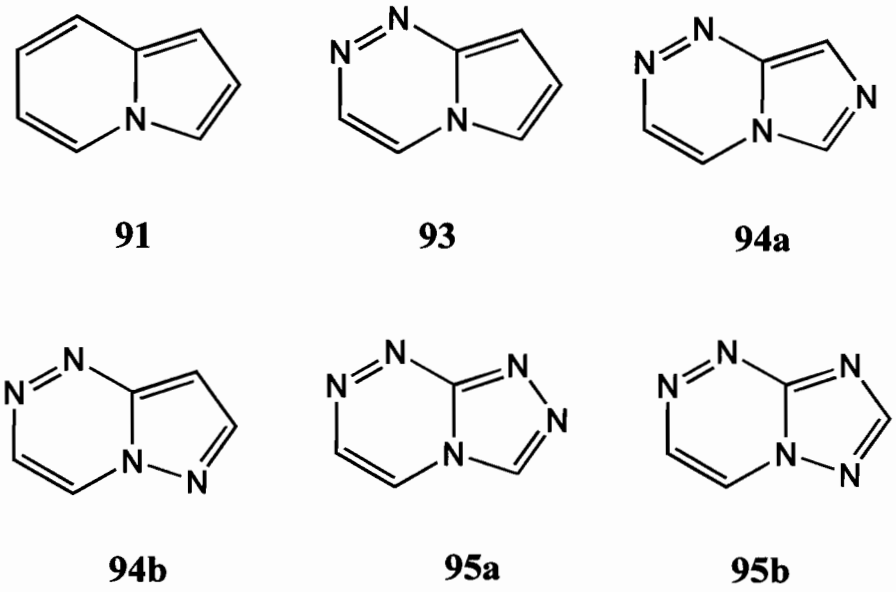


Figure 3.7:

Correlation between *ab initio* Isolobal and Experimental Azole, Azine and Benzoazole Heats of Formation (straight line = experimental)

3.4.5. Extension to Indazine and Azolotriazine Heats of Formation: Indazine (**91, Fig. 3.8**) and azolotriazines (**93-95b, Fig. 3.8**) are compounds similar to benzotriazoles and purines, except that a nitrogen atom is shared by both the five- and six-membered rings.



91 **93** **94a**

94b **95a** **95b**

Figure 3.8: Indazine and Some Related Azolotriazines

(see text and caption of **Figure 3.1** for numbering scheme explination)

Isolobal reaction schemes that generate indazine and azolotriazine from azoles and azines could not be found, because all the reaction schemes considered violate one or more of the conditions set forth for isolobal reactions. However, since all the azolotriazines can easily be generated from isolobal reactions involving indazine, the indazine heat of formation was calculated from the *non-isolobal* isomerization reaction

indole $\rightarrow$ indazine.

The calculated indazine heat of formation was used as the 'experimental' indazine heat of formation in subsequent isolobal reaction schemes. The *ab initio* indazine

heat of formation calculated by the isomerization reaction agrees with the corrected PM3 and MNDO indazine heats of formation (**Table 3.5**), suggesting that the *ab initio* value is reasonable, because MNDO and PM3 perform well on azoles and azines with one or two nitrogen heteroatoms. Reaction schemes used to generate azolotriazine heats of formation are in **Table 3.A5** (**Appendix**).

Table 3.5:
Semi-Empirical and *Ab initio* Indazine and Azolotriazine Heats of Formation

Compound	ΔH_f (kcal/mol)			
	semi-empirical			*ab initio*
	MNDO	AM1	PM3	A
Indazine **91**	55.7	72.3	50.5	54.5*
	60.9	--	*54.1*	
93	78.7	107.5	85.1	104.5
	106.8	--	*99.2*	
94a	81.4	120.7	91.8	115.5
	109.5	--	*105.9*	
94b	91.8	135.3	109.1	116.2
	119.9	--	*123.2*	
95a	93.0	148.5	117.4	138.5
	121.1	--	*131.5*	
95b	94.4	151.7	117.1	125.9
	122.5	--	*131.2*	

*taken as the 'experimental' indazine ΔH_f in subsequent isolobal reaction schemes

italics = corrected semi-empirical

A = average of 4-31G and 6-31G** results

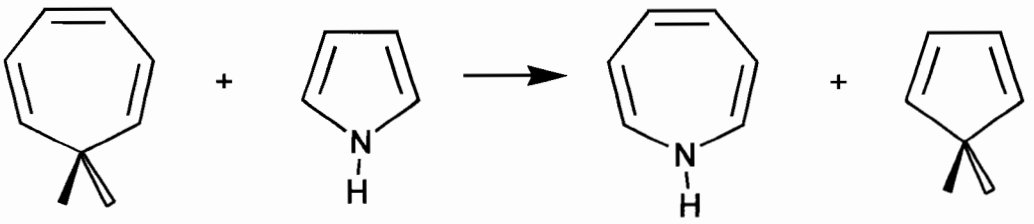
As with azines, the MNDO and PM3 azolotriazine heats of formation are consistently lower than the *ab initio* values, because MNDO and PM3 underestimate the energy of the azine portion of the azolotriazine. However, applying the MNDO and PM3 azine heat of formation correction terms to the azolotriazine heats of formation brings their values closer to the *ab initio* results. AM1 consistently gives azolotriazine heats of formation larger than the *ab initio* values, presumably because AM1 overestimates the energy of the azole portion of the azolotriazine. Thus, as with purines, the AM1 azine heat of formation correction terms were not applied to AM1 azolotriazine heats of formation, as they would not improve the AM1 results. Unfortunately, there are no experimental azolotriazine heats of formation with which to compare with the results.

3.4.6 Extension to Pyrazepine Heat of Formation

For indole, the good result obtained by the isolobal reaction scheme

indene + pyrrole → cyclopentadiene + indole

is of particular interest, because it provides a route to introduce the N-H functionality into a cyclo-olefin, as in the generation of pyrazepine from pyrrole and 1,3,5-cycloheptatriene (**Fig. 3.9**).



70 **51** **71** **50**

Figure 3.9:

Isolobal Reaction for Pyrazepine Heat of Formation

(see **Figure 3.1** for numbering scheme explination)

This reaction scheme was used to calculate the pyrazepine heat of formation. Pyrazepine and 1,3,5-cycloheptatriene can both either planar or non-planar, so calculations were performed on both planar and out-of-plane conformations of these molecules. The 6-31G** total energies show that 1,3,5-cycloheptatriene prefers to be out-of-plane by 4.3 kcal/mol (**Table 3.6**). Semi-empirical values parallel this, and predict that 1,3,5-cycloheptatriene prefers the non-planar conformation by 0.9 (MNDO) to 3.9 kcal/mol (PM3). The experimental 1,3,5-cycloheptatriene heat of formation is well predicted by the PM3 method.

Pyrazepine should prefer a non-planar conformation, because a planar pyrazepine is analogous to aromatic pyrrole, except that it would have 8 π electrons and be anti-aromatic. Surprisingly, the 6-31G** energy difference between the planar and the non-planar conformations of pyrazepine is smaller than the energy difference between the planar and non-planar conformations of 1,3,5-cycloheptatriene (2 kcal/mol). In contrast, semi-empirical methods predict that the planar to non-planar pyrazepine energy differences are slightly larger than the planar to non-planar 1,3,5-cycloheptatriene energy differences (1.9-4.3 kcal/mol).

Overall, the energy differences between planar and non-planar forms of pyazepine and 1,3,5-cycloheptatriene are small. The isolobal reaction scheme used the non-planar *ab initio* total energies of 1,3,5-cycloheptatriene and pyrazepine to calculate the pyrazepine heat of formation, since the experimental 1,3,5-cycloheptatriene heat of formation is undoubtedly for the non-planar conformation. The planar pyrazepine heat of formation can be estimated by adding the planar to non-planar energy difference to the non-planar pyrazepine heat of formation. The resulting *ab initio* and semi-empirical pyrazepine heats of formation are in **Table 3.6**.

The PM3 and *ab initio* pyrazepine heats of formation differ by only 5-7 kcal/mol, while the MNDO and AM1 pyrazepine heats of formation differ from the

ab initio values by 10-15 kcal/mol. There is no data evidence with which to compare these values, but because of the accurate azole, azine and benzoazole heats of formation predicted with PM3 and isolobal reactions, and the agreement between the *ab initio* and PM3 pyrazepine heats of formation, the *ab initio* and PM3 pyrazepine heats of formation are probably reliable estimates.

Table 3.6: Semi-Empirical and Ab initio Pyrazepine Heats of Formation :

Compound	ΔH_f (kcal/mol)						
	semi-empirical			*ab initio*			
	MNDO	AM1	PM3	A	B	C	expt.
1,3,5-cycloheptatriene							
(70)$_{planar}$	34.7	41.1	46.4	--	--	--	--
(70)$_{oop}$	33.8	38.3	42.5	--	--	--	43.2
pyrazepine							
(71)$_{planar}$	49.0	51.3	56.6	61.0	59.8	60.4	--
(71)$_{oop}$	47.3	49.8	52.3	60.2	57.8	59.0	--

A = 4-31G ΔH_f B = 6-31G** ΔH_f C = average of 4-31G and 6-31G** ΔH_f's

Table 3.7:
Isolobal Reaction Energies for Pyrazepine Heat of Formation Calculations

Target Compound Reaction Schemes	ΔH_{rxn} (kcal/mol)		ΔH_f	
	4-31G	6-31G**	4-31G	6-31G**
pyrazepine (**71**)				
70$_{oop}$ + **51** → **71**$_{oop}$ + **50**	22.9	20.5	60.2	57.8

3.5 Conclusions:

Isolobal reaction schemes were used to calculate *ab initio* azole, azine and benzoazole heats of formation. The best results were obtained when an average over a set of possible isolobal reaction schemes were taken, because many similar molecules with known heats of formation were included. This process eliminates effects of particular reaction schemes by cancelling electron correlation and basis set errors. Errors caused by basis sets were further minimized by averaging the results from different basis sets (4-31G and 6-31G**). All averaged isolobal heats of formation were within 0-4 kcal/mol of experiment. The calculated heats of formation are good enough to use in many thermodynamic applications, and the method is useful because it relies on *ab initio* total electronic energies only, allowing direct use of the multitude of *ab initio* results which already exist in the literature. This study considered only nitrogen as a heteroatom, but the method obviously applies to other heteroatoms (such as oxygen and sulphur), when there are some relevent experimental heats of formation. The accuracy of the heats of formation calculated with this method were improved by increasing the number of isolobal reactions; therefore, this method would achieve its full potential in the form of a computer program that would automatically generate suitable isolobal reaction schemes and resulting heats of formation from a data base of *ab intio* total energies and experimental heats of formation.

Uncorrected PM3 azoles and benzoazoles heats of formation are accurate, regardless of the number of nitrogen heteroatoms in the ring system. For azines and other six-membered nitrogen heterocyclic systems, simple correction terms can be added to the PM3 results to increase their accuracy and predictability. The PM3 correction terms are smaller than the MNDO correction terms, reflecting the higher accuracy of PM3. Reasonable agreement between the isolobal and corrected PM3 purine and azolotriazine heats of formtion indicate that these values are good estimates of these quantities. There is also agreement between the PM3 and *ab*

initio pyrazepine heats of formation. MNDO tends to seriously underestimate nitrogen heterocyclic heats of formation when two or more nitrogen heteroatoms are in the system. AM1 behaves erratically, underestimating azine heats of formation and overestimating azole heats of formation. Overall, these results suggest that PM3 is the most reliable semi-empirical method for nitrogen containing heterocyclics. Upcoming chapters will further compare the MNDO, AM1 and PM3 treatments of nitrogen heterocyclic compounds, so as to completely assess the results and conclusions of this chapter.

References:

1. F.R. Benson, "*The High Nitrogen Compounds*", (Wiley:, New York) 1984.

2. E. Fos, J. Vilarrasa and J. Fernandez, *J. Org. Chem.*, **50**, 4894 (1985).

3. W.J. Hehre, R. Ditchfield, L. Radom, and J.A. Pople, *J. Am. Chem. Soc.*, **92**, 4796 (1970).

4. P. George, M. Trachman, A.M. Brett and C. W. Bock, *J. Chem. Soc., Perkin II*, 1036 (1977).

5. M. Sana and G. Leroy, *J. Mol. Struct.*, **207.** 307 (1991)

6. J. A. Pople, B.T. Luke, M. Frisch and J.S. Brinkley, *J. Phys. Chem.*, **89**, 2198 (1985).

7. J.S. Webber and R.G. Woolley, *J. Mol Struct. (THEOCHEM)*, **341**, 171 (1995).

8. M. Sana, G. Leroy, D. Peeters, and C. Wilante, *J. Mol. Struct., (THEOCHEM)*, **164,** 249 (1988).

9. A.Y. Shibaev, and Y.V. Puzaznov, *Russ. J. Phys. Chem.*, **62**, 291 (1988).

10. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. HensenS. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, 10 (1990) 52. b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, J. *Comput. Chem.*, **14**, 1347 (1993).

11. K.J. Laidler, and J.H. Meiser, "*Physical Chemistry,*" (Benjamin-Cummings: Don Mills, Ont.) p59, 1982.

12. a) E. Fos, J. Vilarrasa, and J. Fernandez, *J. Org. Chem.*, **50**, 4894 (1985).
b) W.M.F. Fabian, *Z. Naturforsch*, **45a** 1328 (1990).

13. J.P. Ritchie, *J. Org. Chem.*, **54**, 3553 (1989).

14. S.G. Lias, J.F. Libman and R.D. Levine, *J. Phys.Chem.Ref. Data*, **13**, 695 (1984).

15. Y.I. Alexsandrov, T.R. Osipova and V.F. Yushkevich, *Thermodin. Org. Soedin.*, **42**, 1 (1982).

16. J.D. Cox and G. Pilcher, "*Thermochemistry of Organic and Organometallic Compounds*"; Academic Press: New York, 1970.

APPENDIX 3A:

TABLE 3.A1: *Ab initio* Total Energies Used in Study

		-E (a.u.)	
		4-31G	6-31G**
cyclopentadiene	(50)	192.51317	192.80242
benzene	(60)	230.37776	230.71386
1,3,5-cycloheptatriene			
(planar)	(70)planar	269.28798	269.68924
(out-of-plane)	(70)oop	269.29647	269.69607
ethene	(e)	77.92216	78.03884
pyrrole	(51)	208.50745	208.81970
imidazole	(52a)	224.47112	224.82430
pyrazole	(52b)	224.45073	224.80328
1,2,4-triazole	(53a)	240.41404	240.80693
1,2,3-triazole	(53b)	240.38208	240.77698
tetrazole	(54a)	256.31781	256.76012
pentazole	(55)	272.22237	272.71786
pyridine	(61)	246.33133	246.70461
1,2-diazine	(62a)	262.24489	262.65707
1,3-diazine	(62b)	262.28621	262.70034
1,4-diazine	(62c)	262.27852	262.68995
1,2,3-triazine	(63a)	278.16137	278.61661
1,2,4-triazine	(63b)	278.19295	278.64718
1,3,5-triazine	(63c)	278.24148	278.70076
1,2,3,4-tetrazine	(64a)	294.07028	294.57073
1,2,3,5-tetrazine	(64b)	294.10917	294.61119
1,2,4,5-tetrazine	(64c)	294.09823	294.59638
pentazine	(65)	309.97691	310.52541
hexazine	(66)	325.84356	326.44896

TABLE 3.A1 (cont): *Ab initio* Total Energies Used in Study

		-E (a.u.)	
		4-31G	6-31G**
Indene	(B50)	344.96902	345.47736
Indole	(B51)	360.95088	361.48292
Benzoimidazole	(B52a)	376.91618	377.48879
Benzopyrazole	(B52b)	376.88927	377.45995
Benzotriazole	(B53)	392.82132	393.43350
Purine1	(P1)	408.82924	409.48079
Purine2	(P2)	408.82057	409.47268
pyrazepine(planar)	(71)	285.25299	285.67753
(oop)		285.25418	285.68068
Indazine	(91)	360.92189	361.45763
azolotriazines			
	(93)	392.78344	393.39532
	(94a)	408.73885	409.39148
	(94b)	408.72543	409.37912
	(95a)	424.66556	425.36240
	(95b)	424.68385	425.38487

TABLE 3.A2: Azole Isolobal Reaction Schemes

Target Compound Reaction schemes	ΔHrxn (kcal/mol)		ΔHf (kcal/mol)	
	4-31G	6-31G**	4-31G	6-31G**
pyrrole (51)				
*?50 + 61 → 51 + 60	-25.5	16.6	-12.0	30.1
52a + 61 → 51 + 62c	10.3	6.8	33.0	29.5
52a + 60 → 51 + 61	6.3	3.4	26.5	23.6
			29.8	*26.6*
imidazole (52a)				
51 + 62c → 52a + 61	-10.3	-6.8	28.0	31.5
51 + 61 → 52a + 60	-6.3	-3.4	34.5	37.4
51 + 62b → 52a + 61	-5.5	-0.3	32.9	38.1
53a + 52b → 52a + 53c	7.3	14.0	35.5	42.2
			32.8	*37.3*
pyrazole (52b)				
51 + 62a → 52b + 61	-18.6	-19.5	39.3	38.4
62a + 53a + 51 → 2(52b) + 62c	-8.5	-8.4	41.5	41.6
52a + 62a → 52b + 62b	-13.1	-14.0	41.2	40.3
52a + 62a → 52b + 62c	-8.3	-7.5	46.3	47.1
*?52a → 52b	12.8	7.0	47.8	48.9
			43.2	*41.9*
1,2,4-triazole (53a)				
51 + 62a → 53a + 60	-24.8	-30.8	47.9	41.9
52a + 62a → 53a + 61	-18.4	-27.4	48.5	39.5
52a + 52b → 53a + 51	0.2	-7.9	53.2	45.1
			49.8	*42.2*
1,2,3-triazole (53b)				
62a + 52b → 53b + 61	-11.2	-13.2	64.6	62.6
53a + 52b → 53b + 52a	7.3	5.6	62.4	60.7
2(52b) → 53b + 51	7.5	6.2	69.5	68.2
62a + 53a → 53b + 62b	-5.9	-8.3	59.4	57.0
62a + 53a → 53b + 62c	-1.0	-1.8	64.4	63.8
*?53a → 53b	20.1	21.9	66.2	68.0
			64.4	*63.4*
tetrazole (54)				
62a + 53b → 54 + 61	-13.9	-19.3	79.8	74.4
2(52b) + 62a → 52a + 54 + 60	-12.8	-19.9	85.1	78.0
62a + 2(52b) → 54 + 51 + 61	-6.4	-13.1	87.5	80.8
			84.1	*77.7*
pentazole (55)				
54 + 2(52b) → 55 + 52a + 51	11.5	3.0	118.6	110.1
54 + 52b + 53a → 55 + 2(52a)	11.3	2.1	111.4	102.2
			115.0	*106.2*

TABLE 3.A3: Azine Isolobal Reaction Schemes :

Target Compound Reaction Schemes	ΔHrxn (kcal/mol)		ΔHf (kcal/mol)	
	4-31G	6-31G**	4-31G	6-31G**
pyridine (61)				
52a + 60 → 61 + 51	6.3	3.6	35.1	32.2
53a + 60 → 61 + 52b	6.1	8.1	34.9	36.9
			35.0	34.6
1,2-diazine (62a)				
52b + 61 → 62a + 51	18.6	19.5	71.2	72.1
52b + 62c → 62a + 52a	8.3	6.6	65.1	63.4
52b + 62b → 62a + 52a	13.1	14.0	69.3	70.2
53a + 60 → 62a + 51	24.8	30.8	64.6	70.6
			67.6	69.0
1,3-diazine (62b)				
*61a → 62b	-25.9	-27.1	40.5	39.3
61 + 61 → 62b + 60	-0.8	-3.1	48.6	46.3
52a + 61 → 62b + 51	5.5	0.3	47.9	43.9
			45.7	43.2
1,4-diazine (62c)				
61 + 61 → 62c + 60	-0.8	3.4	48.6	52.8
52a + 61 → 62c + 51	10.3	6.8	53.8	50.3
*62b → 62c	4.8	6.5	51.8	53.5
63c + 61 → 62c + 62b	5.0	9.4	46.6	51.0
			50.1	*51.9*
1,2,3-triazine (63a)				
2(62a) → 63a + 61	-1.8	-4.4	96.6	94.0
53a + 62a → 63a + 52a	16.6	14.2	94.1	91.7
52b + 62a → 63a + 51	16.8	17.4	101.3	101.9
*63c → 63a	50.3	52.8	101.7	104.2
			98.4	98.0
1,2,4-triazine (63b)				
62a + 62c → 63b + 61	-0.5	-3.0	78.3	75.8
62a + 61 → 63b + 60	3.5	0.4	84.8	81.7
52a + 62a → 63b + 51	9.8	3.8	85.5	79.3
62a + 62b → 63b + 61	4.3	3.5	83.2	82.4
53a + 61 → 63b + 51	28.2	31.2	82.8	85.8
* 63c → 63b	30.5	33.6	85.5	88.6
			82.9	81.0

TABLE 3.A3: (cont.)

Target Compound Reaction Schemes	ΔHrxn (kcal/mol)		ΔHf (kcal/mol)	
	4-31G	6-31G**	4-31G	6-31G**
1,3,5-triazine (63c)				
3(61) → 63c + 2(60)	-1.8	-9.2	62.4	55.0
2(62b) → 63c + 61	-0.2	-2.9	59.1	56.4
62b + 61 → 63c + 60	-1.1	-6.1	60.7	55.7
62b + 52a → 63c + 51	5.3	-2.7	60.3	52.6
62b + 53a → 63c + 52b	5.0	5.2	54.0	54.2
62c + 61 → 63c + 60	-5.9	-12.6	55.8	49.1
62c + 52a → 63c + 51	0.4	-9.2	57.3	46.7
62c + 53a → 63c + 52b	0.2	-4.4	49.1	44.5
			57.2	*51.8*
1,2,3,4 -tetrazine				
3(62a) → 64a + 2(61)	1.1	-5.5	131.4	124.8
2(52b) + 62a → 64a + 2(51)	38.4	33.5	141.0	136.1
2(53a) + 62a → 64a + 2(52a)	37.9	32.3	125.6	120.9
3(52b) + 61 → 64a + 3(51)	57.0	53.1	145.6	141.7
61 + 54 → 64a + 51	44.8	46.4	133.1	134.7
			135.3	*131.6*
1,2,3,5 -tetrazine				
2(53a) + 61 → 64b + 52a + 51	45.0	39.7	111.6	105.3
52b + 62a + 62b → 64b + 51 + 61	21.3	18.1	118.4	115.2
52b + 62a + 62c → 64b + 51 + 61	16.4	11.6	113.2	108.4
2(62a) + 62b → 64b + 2(61)	2.6	-3.7	113.5	107.2
2(62a) + 63c → 64b + 61 + 62c	7.7	5.7	113.1	111.1
			114.0	*109.5*
1,2,4,5 -tetrazine				
2(62a) → 64c + 60	8.7	2.4	121.9	115.6
2(53a) + 60 → 64c + 2(51)	58.2	52.8	118.0	112.6
			120.0	*114.0*

TABLE 3.A3:(cont.)

Target Compound Reaction Schemes	ΔHrxn (kcal/mol) 4-31G	ΔHrxn (kcal/mol) 6-31G**	ΔHf (kcal/mol) 4-31G	ΔHf (kcal/mol) 6-31G**
pentazine (65)				
54 + 53a + 60 → 65 + 2(51)	73.9	72.9	166.8	165.8
54 + 52b + 61 → 65 + 2(51)	67.8	64.8	174.5	171.5
2(62a) + 54 → 65 + 52b + 61	30.5	25.7	165.0	160.5
2(54) + 61 → 65 + 2(52b)	55.6	58.3	162.4	165.1
			167.2	*165.7*
hexazine (66)				
6(62a) → 66 + 4(61) + 62b	8.9	-15.9	222.5	197.7
6(62a) → 66 + 4(61) + 62c	13.8	-9.4	227.5	204.3
60 + 2(54a) → 66 + 2(51)	97.2	91.5	225.2	219.5
6(52b) + 60 → 66 + 6(51)	121.7	104.4	249.5	232.2
3(53b) + 60 → 66 + 3(51)	99.2	85.8	232.7	219.3
			231.5	*214.6*

TABLE 3.A4: Reaction Schemes for Benzoazoles and Purines:

Target Compound Reaction Schemes (e= ethene)	ΔHrxn (kcal/mol) 4-31G	6-31G**	ΔHf 4-31G	6-31G**
Indene (B50)				
60 + 50 → B50 + e	-0.2	0.05	39.0	39.3
B51 + 50 → B50 + 51	-7.8	-7.4	35.5	35.9
			37.3	*37.6*
Indole (B51)				
60 + 51 → B51 + e	7.6	7.4	40.9	40.7
B50 + 51 → B51 + 50	7.8	7.4	40.9	40.5
			40.9	*40.6*
Benzoimidazole (B52a)				
60 + 52a → B52a + e	6.6	6.6	48.9	48.9
B50 + 52a → B52a + 50	6.8	6.6	48.9	48.7
			48.9	*48.8*
Benzopyrazole (B52b)				
60 + 52b → B52b + e	10.7	11.5	62.0	62.8
B50 + 52b → B52b + 50	10.9	11.5	62.0	62.6
			62.0	*62.7*
Benzotriazole				
B2 + 53a + 52b → B6 + 52a + 51	9.9	16.0	76.3	82.4
B1 + 52a + 54a → B6 + 50 + 53a	5.9	5.7	82.1	81.9
60 + 53c → B6 + e	10.3	11.6	81.5	82.8
			79.9	*82.2*
purine1 (**p1**)				
62b + 52a → p1 + e	3.7	3.1	73.2	72.6
62c + 52a → p1 + e	-1.1	-3.4	68.3	66.0
indole + 52a + 62b → p1 + 51 + 60	-3.9	-4.2	69.7	69.4
indole + 63c → p1 + 60	-9.2	-6.9	62.3	64.6
			68.4	*68.2*
purine2 (**p2**)				
62b + 52a → p2 + e	9.2	8.2	78.7	77.7
62c + 52a → p2 + e	4.3	1.7	73.1	71.2
indole + 52a + 62b → purine2 + 51 + 60	1.5	0.8	75.1	74.4
indole +63c → p2 + 60	-3.7	-1.8	67.8	69.7
			73.7	*73.3*

ethene ΔH_f = 12.5 kcal/mol (ref. 16)

Table 3.A5: Indazine and Azolotriazine Isolobal Reaction Schemes

Target Compound Reaction Schemes (e= ethene)	ΔHrxn (kcal/mol) 4-31G	6-31G**	ΔHf 4-31G	6-31G**
*B50 → 91	18.2	15.9	53.3	55.6
93				
91 + 62a → 93 + 60	3.5	3.5	104.7	104.7
91 + 53a → 93 + 51	28.3	31.1	102.8	105.6
			103.8	*105.2*
94a				
91 + 62a + 52a → 94a + 60 + 51	8.7	8.7	118.9	118.9
91 + 53a + 52a → 94a + 2(51)	33.5	36.3	116.9	119.9
91 + 62a + 62b → 94a + 60 + 61	3.2	3.2	116.8	116.8
91 + 53a + 62c → 94a + 51 + 61	23.1	24.3	109.9	111.1
91 + 62a + 53c → 94a + 60 + 62b	3.4	6.1	111.5	114.2
			114.8	*116.2*
94b				
91 + 62a + 62a → 94b + 60 + 61	-14.3	-16.2	118.8	116.9
91 + 62a + 53a → 94b + 60 + 52a	4.1	2.7	116.3	114.9
91 + 53a + 53a → 94b + 51 + 52a	28.8	30.4	114.3	115.9
			116.5	*115.9*
95a				
91 + 62a + 53a → 95a + 60 + 51	18.9	16.1	140.1	137.3
91 + 2(53a) → 95a + 2(51)	43.6	43.7	138.1	138.2
			139.1	*137.8*
95b				
91 + 62a + 53a → 95b + 60 + 51	7.4	2.0	128.6	123.2
91 + 53a + 53a → 95b + 2(51)	32.1	29.6	126.6	124.1
91 + 62a + 62a → 95b + 2(60)	-17.4	-25.6	130.5	122.3
			128.6	*123.2*

The results in chapter three suggested that the PM3 semi-empirical method is the most trustworthy for nitrogen heterocyclic heats of fomation. In the next chapter, MNDO, AM1 and PM3 methods will again be compared with experiment and *ab initio* methods, to further assess their accuracy and predictive power. Isocyanate equilibrium geometries, predicted by all the semi-empirical methods, will be shown to agree well with experiment. The isocyanate equilibrium geometries, predicted by *ab initio*, are very sensitive to the inclusion of polarization functions in the basis set. The transition state geometries, and activation energy barriers, for the isocyanate cycloaddition reaction with alkenes to form β-lactams, will be calculated with the MNDO, AM1 and PM3 semi-empirical methods. The semi-empirical activation barriers and transition state geometries will be compared with those obtained from high level *ab initio* calculations. Again, the PM3 method will be shown to yield results superior to MNDO and AM1. PM3 results are comparable to post-Hartree-Fock MP2 and MP4 results for the same systems. This chapter further justifies the use of PM3 a for nitrogen-heterocyclic study, because PM3 yields results compaable to post-Hartree-Fock, at a small fraction of the post-Hartree-Fock computer cost.

CHAPTER 4:

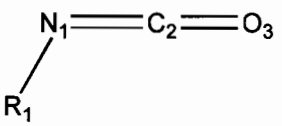
SEMI-EMPIRICAL STUDY OF ISOCYANATE GEOMETRIES, AND β-LACTAM FORMATION THROUGH ALKENE-ISOCYANATE CYCLO-ADDITION REACTIONS:

4.1 Abstract:

Semi-empirical and *ab initio* geometries and energies of ground state isocyanate molecules are compared with experimental results. Semi-empirical methods all produce isocyanate geometries that compare well with experiment. Without polarization functions, 3-21G, 6-31G, and 6-311G *ab initio* basis sets all perform poorly on isocyanate, and produce incorrect linear geometries. Inclusion of polarization functions improves the *ab initio* isocyanate geometries drastically, and brings the *ab initio* results in line with experiment. Semi-empirical methods (MNDO, AM1 and PM3) are used to study the β-lactam forming isocyanate cycloaddition reaction with various subsutituted alkenes. MNDO and AM1 transition state geometries for these reactions differ significantly from *ab initio* transition state geometries, while PM3 activation energy barriers and transition states are in better agreement with high level *ab initio* results than either MNDO or AM1. The PM3 method is the most accurate semi-empirical method for the hydrogen isocyanate heat of formation.

4.2 Introduction:

Isocyanates (**1**, **Fig. 4.1**) are useful in the syntheses of β-lactams[1] (**3,5,7, Fig. 4.2**) and imidazotetrazinones.[2] Isocayantes are of interest to us because they are also produced in the decomposition reactions of benzotetrazepinones,[3] which will be considered in chapters seven and eight. Isocyanates have been extensively studied with CNDO semi-empirical methods[4], as well as minimal and extensive *ab initio* calculations.[5-9]



1

R_1 =
a -H
b -Me
c -Cl
d -F
e $-C_6H_5$

Figure 4.1: Isocyanates

To date, the most extensive *ab initio* studies are the MP4SDTQ/6-31G** calculations of Mack et. al.,[9] and the MP2/TZ2P calculations of Handy et. al.[5]. These studies yielded isocyanate geometries in excellent agreement with experimental results.[10-14]

Isocyanate cycloaddition reactions with olefins to form β-lactams[1] (**Fig. 4.2**) have been studied at the *ab initio* RHF/6-31G* and MP2/6-31G* levels of theory by Cossio et. al.,[15] and Fu et. al.[16] Such state-of-the-art calculations can yield excellent results, but they become costly when the number and size of molecules to be studied becomes large. Semi-empirical methods (MNDO, AM1 and PM3) are theoretically less accurate than extensive *ab initio* calculations, but they are orders of magnitude faster, thus allowing a larger number of molecules to be studied. In chapter three, it was shown that semi-empirical methods (especially PM3) can yield useful heats of formation data for nitrogen heterocyclic compounds.

Semi-empirical methods are further tested in this chapter. MNDO, AM1, and PM3 equilibrium geometries for isocyanic acid (**1a**), methyl-isocyanate (**1b**) and chloro-isocyanate (**1c**) will be compared with the results from different levels of *ab initio* theory and experiment. Semi-empirical activation energy barriers and transition state geometries for isocyanate/olefin cycloaddition reactions will be compared with results from *ab initio* calculations, since no experimental transition state geometries of these systems exist.

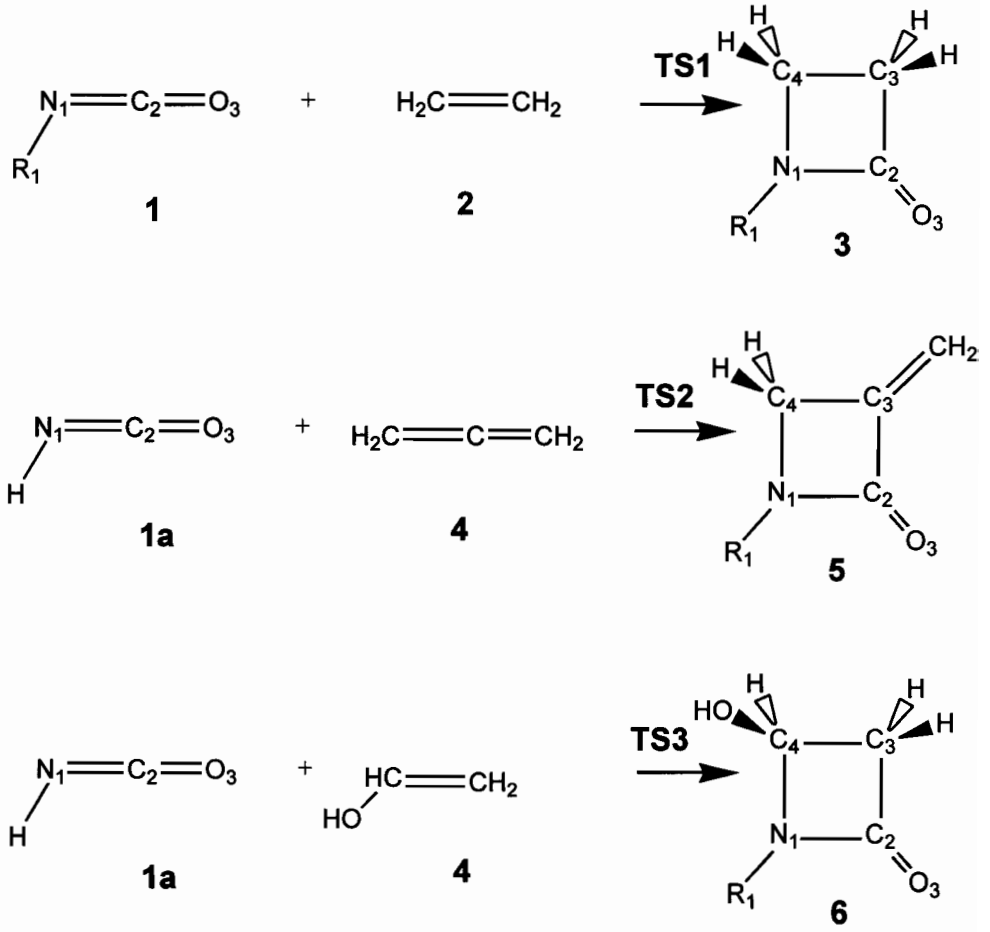


Figure 4.2:

Isocyanate Cycloaddition Reactions with Olefins to form β-Lactams

4.3 Calculations:

All calculations were performed on a SUN Sparc10 workstation using the GAMESS[17] package. Full geometry optimizations with the Baker[18] method were performed in all cases, and RHF wavefunctions used. Experimental values were used for input geometries. In all cases, the equilibrium geometries and transition states were verified as such by force constant analysis. *Ab initio* calculations used 6-31G and 6-311G and basis sets, with and without *p*- and *d*- polarization functions. Although isocyanate has been reported as planar, C_s symmetry was not enforced, because it would have biased any calculation that would otherwise report a non-planar geometry. Calculations which yielded planar optimized geometries when C_1 symmetry was specified were re-initiated with C_s enforced, using the C_1 optimized geometry as input.

4.4 Results and Discussion

4.4.1. Isocyanate Geometries

In **Table 4.1**, the MNDO, AM1 and PM3 generated equilibrium geometries are compared with *ab initio*[5,9] and experimental results. In all cases (experiment, semi-empirical and *ab initio*) the isocyanate molecules are planar with non-linear NCO bond angles. The oxygen is *trans* to the nitrogen substituent (R_1). The NCO bond angle in isocyanic acid and methyl isocyanate is assumed to be 180° by some experimental workers,[10-13] but more recent results[14] report a bent NCO bond, in agreement with our calculations. The NCO bond in chloroisocyanate is known to be bent.[10-12] The MNDO, AM1 and PM3 methods give a slightly more bent CNO bond angle than *ab initio* methods, with the PM3 method in closest agreement with experiment. The AM1 semi-empirical method was best at predicting the N=C and N-H bond distances in isocyanic acid (**1a**).

Table 4.1: Semi-Empirical and Experimental Isocyanate Geometries

	MNDO	AM1	PM3	RHF 6-31G*	MP2 6-31G*[a]	expt.
Hydrogen Isocyanate (**1a**)						
			Bond Lengths (Å)			
N1=C2	1.249	1.232	1.251	1.199	1.225	1.214[14a]
C2=O3	1.185	1.202	1.181	1.148	1.183	1.166[14a]
N1-H	0.998	0.985	0.985	0.993	1.004	0.995[14a]
			Bond Angles (degrees)			
C2=N1-H	120.2	127.2	123.7	125.1	124.8	123.9[14a]
N1=C2=O3	167.6	166.7	168.7	174.3	171.7	172.6[14a]
Methyl Isocyanate (**1b**)						
			Bond Lengths (Å)			
N1=C2	1.243	1.236	1.257	1.180	1.215	1.207[14b]
C2=O3	1.186	1.202	1.180	1.158	1.190	1.166[14b]
N1-C4	1.443	1.415	1.449	1.437	1.440	--
C4-H5	1.113	1.124	1.096	1.080	1.079	--
C4-H6	1.115	1.124	1.098	1.083	1.083	--
			Bond Angles (degrees)			
C2=N1-C4	130.1	132.2	129.3	142.9	137.6	140.1[14b]
N1=C2=O3	167.1	165.7	168.3	174.9	172.3	180.0*[14b]
N1-C4-H5	112.1	112.1	112.0	110.7	110.6	--
N1-C4-H6	107.3	106.1	105.9	108.6	108.7	--
Chloroisocyanate (**1c**)						
			Bond Lengths (Å)			
N1=C2	1.276	1.264	1.267	1.217	1.241	1.218[14c]
N1-Cl	1.708	1.658	1.695	1.694	1.709	1.703[14c]
C2=O3	1.178	1.194	1.174	1.141	1.181	1.165[14c]
			Bond Angles (degrees)			
C2=N1-Cl	116.5	121.5	124.4	120.2	121.9	119.5[14c]
N1=C2=O3	168.8	166.7	169.4	173.8	169.7	171.5[14c]

* angle of 180.0° was assumed by these authors.[14b]
a = MP2/6-31G** optimized geometry from ref. 9

PM3 gave the best match to experiment for the H-N=C bond angle. The semi-empirical methods predict an N-H bond distance closer to experiment than the high level *ab initio* results. The C=O bond distance was moderately well predicted by the semi-empirical methods, with errors ranging from 0.019 to 0.036 Å. Note that the MNDO and PM3 predictions for the C=O bond length are better than the RHF/6-31G** predictions, and are essentially identical to the MP2/6-31G** values.[9] Of all the theoretical methods, PM3 gave a C=O bond length which is closest to experiment.

The N=C bond distance was the least satisfactory semi-empirical prediction, with N=C bond lengths as much as 0.04 Å larger than experiment. However, the best semi-empirical result (AM1) over-predicts the C=N bond length by about the same amount that RHF/6-31G** under-predicts this bond length. The MP2/6-31G** gave the closest match to the experimental N=C bond length.

Ab initio and semi-empirical methods all predict that substitution of a methyl group (**1b**) in place of hydrogen in isocyanic acid had little effect on both N=C and C=O bond distances. This is verified by the experimental results. In addition, both *ab initio* and semi-empirical methods predict only a slight change (1^{o}) in the N=C=O bond angle upon methyl substitution. Unfortunately, the experimentally determined geometry of methyl-isocyanate assumes an N=C=O bond angle of 180^{o}, so there is no way to compare the theoretical predictions to experimental values.[14b] The semi-empirical and *ab initio* methods all show that the greatest effect of methyl substitution is the opening of the X-N=C bond angle, which *ab initio* predicts to open by 17^{o}, and semi-empirical methods predict to open by 5^{o} to 10^{o}. This parallels the experimental estimate of a 16.2^{o} opening of this bond angle upon methyl substitution.[14b]

For chloro-isocyanate (**1c**), all theoretical methods except HF/6-31** predict N=C and C=O bonds lengths drastically larger than the experimentally obtained values. The semi-empirical and MP2/6-31G** *ab initio* methods[9] predict increases of 0.015 to 0.028 Å in the N=C bond distance upon chloro-

substitution. The experimental results show a 0.004 Å increase in this bond length. The HF/6-31G** C=N bond length is much better than the MP/6-31G** result[9] for this bond length, even though MP2 is a higher level of *ab initio* theory than HF. This shows that going to a higher level of *ab initio* theory does not necessarily produce better results.

The MNDO and PM3 methods predict the N-Cl bond length quite well, (MNDO gives a better match to experiment than *ab initio* in this case) while the AM1 method seriously underestimates this bond length. The Cl-N=C bond angle predictions are quite different between the semi-empirical methods, with AM1 and PM3 overestimating and MNDO underestimating this bond angle. However, the semi-empirical Cl-N=C bond angles are comparable to the *ab initio* and experimental values.

Overall, the most glaring weakness of the semi-empirical methods in reproducing experimental isocyanate geometries is the drastic overestimation of the N=C bond distance. Another weakness of the semi-empirical methods is their underestimation of the N=C=O bond angle. However, the semi-empirical methods do match experiment and *ab initio* results by predicting a non-linear N=C=O bond angle. With regards to other geometric parameters, the semi-empirical results are comparable to the *ab initio* results. The post-HF MP2 and MP4 methods used by workers in references 5,9,15 and 16 require more than 6000 times the amount of computer CPU time required by the semi-empirical methods.

4.4.2. Effect of Polarization Functions on ab initio Optimized Isocyanate Geometries:

Mack et. al.[9] found that a 3-21G basis set, without polarization functions, predicts a linear H-N=C bond angle. Modelling isocyanates with *ab initio* methods reveals that inclusion of polarization functions in the *ab initio* gaussian orbital basis sets is necessary to correctly model isocyanates. This point can be exemplified by examining the effect of polarization functions on the *ab initio* geometries of isocyanate (**1a**), **Table 4.2.** Here, the results of HNCO geometry

optimization using a 6-311G basis set, with and without polarization functions (**Table 4.2**) are given. Without polarization functions, many *ab initio* basis sets incorrectly predict an H-N=C bond angle of almost 180°. This total disagrees with experiment. The error occurs with the 3-21G, 6-31G and TZV basis sets, as well as the 6-311G basis set results given here. In addition, the N=C=O bond is predicted to be almost 180° in the absence of polarization functions.

Addition of one *p*-type polarization function to the hydrogen changes the *ab initio* N=C=O bond angle to 174.8°, which is closer to experiment, and is qualitatively correct in its prediction of a non-linear N=C=O bond angle. The *p*-type polarization function changes the H-N=C bond angle to 146.9°, which is closer to experiment than the 180° angle predicted in the absence of polarization functions, but is still drastically wrong, and inferior to the semi-empirical results. The *p*-type function also brings all the bond distances into better agreement with experiment.

Addition of one *d*-type polarization function to each of the heavy atoms yields predictions of the H-N=C and N=C=O bond angles in very good agreement with experimental results. One *d*-type function brings the N=C and N-H bond distances into better agreement with experiment, but inclusion of the *d*-type function only results in a short C=O bond and a long H-N bond. This is in direct contrast with the non-polarized and *p*-only polarized results, which produce a long C=O bond and a short N-H bond relative to experiment.

Addition of one *p*-type and one *d*-type polarization function yields isocyanate geometries very similar to the geometries produced by one *d*-type polarization function alone; the difference between the two cases is a slight lengthening of the H-N and N=C bonds, a slight shortening of the C=O bond, and a slight decrease of the H-N=C and N=C=O bond angles. Addition of three *p*-type and three *d*-type polarization functions yielded results very similar to those obtained with only one of each type of polarization function.

Table 4.2:
Effect of Polarization Functions on *Ab Initio* Isocyanate Geometry Prediction

Polarization Functions	6-311G Basis set					6-31G
# p-type	0	1	0	1	3	3
# d-type	0	0	1	1	3	3
			Bond Lengths (Å)			
N1=C2	1.163	1.179	1.195	1.196	1.198	1.202
N1-H4	0.972	0.980	0.991	0.994	0.993	0.995
C2=O3	1.182	1.175	1.140	1.139	1.137	1.141
			Bond Angles (degrees)			
C2=N1-H4	179.9	146.9	125.7	123.9	122.2	121.2
N1=C2=O3	180.0	174.8	174.3	174.4	174.1	174.5

In total, these results indicate that although high level *ab initio* results for the isocyanates give slightly better results than the semi-empirical methods, one must be very cautious when using low level *ab initio* theories. Even without polarization functions, the RHF/6-311G *ab initio* calculation is about >1000 times more computationally costly than semi-empirical calculations; however, it yields results which are qualitatively incorrect, and far inferior to the semi-empirical results.

4.4.3. Isocyanate Heats of Formation

Little thermodynamic data exists for isocyanates,[9] but the heat of formation of hydrogen isocyanate (**1a**) is known.[19] The hydrogen isocyanate heat of formation has also been computed by isolobal reaction schemes using extensive *ab initio* total energies.[20] In addition, the methyl isocyanate heat of formation was calculated using the 6-31G** energy of the isodesmic reaction scheme

$$HNCO + CH_3NH_2 \rightarrow CH_3NCO + NH_3 ,$$

in conjunction with ammonia (-11.0 kcal/mol), methylamine (-6.7 kcal/mol) and HNCO (-27.2 kcal/mol) experimental heats of formation.[19]

Semi-empirical, experimental and *ab initio* isocyanate heats of formation are compared in **Table 4.3**. There are no experimental or *ab initio* heats of formation for **1c-1e** with which to compare the semi-empirical results. Comparison of experimental, *ab initio* and semi-empirical **1a** and **1b** heats of formation shows that semi-empirical methods correctly predict **1a** and **1b** to be exergonic with respect to the elements, but there are consistent errors of 10-17 kcal/mol in semi-empirical isocyanate heats of formation. PM3 is in closest agreement with *ab initio* and experiment, while MNDO is in worse agreement.

Table 4.3: Semi-Empirical Isocyanate Heats of Formation:

		ΔH_f (kcal/mol)				
		MNDO	AM1	PM3	*ab initio*[20]	expt.
Isocyanates						
1a	(-H)	-10.8	-15.2	-15.3	-22.4	-24.4
		-21.4	*-24.6*	*-22.5*		
1b	(-Me)	-16.2	-13.9	-18.5	-25.7	
		-26.8	*-23.4*	*-25.7*		
1c	(-Cl)	5.7	10.4	-14.4		
1d	(-F)	-11.3	-5.2	-6.6		
1e	(-C6H5)	14.1	18.2	11.8		

corrected semi-empirical heats of formation (in italic) calculated from
corrected ΔH_f = semi-empirical ΔH_f + correction term

correction terms (kcal/mol)

MNDO	AM1	PM3
- 10.6	-9.5	-7.2

Because there is only one experimental isocyanate heat of formation, the *ab initio* isocyanate heats of formation were used to calculate correction terms for the MNDO, AM1 and PM3 isocyanate heats of formation. The correction terms were calculated by simply averaging the differences between the semi-empirical and *ab initio* **1a** and **1b** heats of formation. The correction terms and resulting 'corrected' semi-empirical isocyanate heats of formation are also in **Table 4.3**. These correction terms will be used to compare semi-empirical and *ab initio* isocyanate + olefin → lactam heats of reaction.

4.4.4. Cycloaddition Reactions with Alkenes:

Cycloaddition reactions between isocyanates and alkenes to form lactams (**Fig. 2**) have been extensively studied with high level *ab initio* theory by Cossio et. al.,[15] and Fu et. al.[16] Both groups found the reaction to occur through a concerted [2+2] suprafacial mechanism, without involvement of intermediates. All transition states, except for the **1a** + **6** → **7** reaction transition state (**TS3**), had C_s symmetry. The MNDO, AM1 and PM3 semi-empirical methods parallel the *ab initio* results and predict concerted [2+2] suprafacial planar transition states, all of which (except for **TS3**) have C_s symmetry.

The *ab initio* and semi-empirical reaction energies are compared in **Table 4.4**. Semi-empirical reactant and product heats of formation used to calculate the heat of reaction are in **Table 4.A1** (**Appendix**). Corrected semi-empirical reaction energies take in to account the previously discussed corrections to the hydrogen isocyanate and the methyl isocyanate heats of formation. Semi-empirical olefin heats of formation agree well with experiment, therefore no correction terms are required for their heats of formation. No experimental β-lactam heats of formation exist with which to compare the semi-empirical β-lactam heats of formation.

Although there are discrepancies between the semi-empirical and *ab initio* reaction energies, all methods predict a net exothermic reaction. In some cases there is excellent agreement between *ab initio* and semi-empirical reaction energies, but this agreement is inconsistent and apparently random. PM3 and MNDO agree well with *ab initio* **1c** + **2** → **3c** reaction energies, while AM1 agrees well *ab initio* **1a** + **2** → **3a** reaction energies. The corrected PM3 **1a** + **2** → **3a** heat of reaction agrees well with the *ab initio* value. Overall, the semi-empirical MNDO method predicts more exothermic reaction energies than do the *ab initio* or the other semi-empirical methods, probably because of MNDO's overestimation of 4-membered ring stabilities.[21] The differences between semi-empirical and *ab initio* reaction energies range from 0.7 - 21 kcal/mol. It should be noted that within different levels of *ab initio* theory there are discrepancies of 3-7 kcal/mol between reaction energies.

Table 4.4:
Comparison of *ab initio* and Semi-Empirical β-Lactam Reaction Energies

	ΔH_{rxn} (kcal/mol)					
Isocyanates with ethene					6-31G*	
	MNDO	AM1	PM3	RHF	RMP2	RMP4
(-H)	-31.5	-10.2	-23.1	-10.9	-15.0	-14.1
	-20.9	*-0.7*	*-15.9*			
(-Me)	-29.7	-7.8	-22.4	--	--	--
(-Cl)	-29.0	-8.9	-22.7	-22.3	-29.7	
(-F)	-31.2	-8.0	-25.3	--	--	--
$(-C_6H_5)_{ip}$	-27.8	-10.4	-24.5	--	--	--
HNCO with						
(C_2H_3OH)	-33.5	-9.7	-32.7	-13.1	-16.4	--
(C_3H_4)	-36.7	-11.0	-28.4	-18.5	-22.4	--

4.4.5. Activation Barriers:

The semi-empirical activation barriers are in **Table 4.5,** along with the *ab initio* values from the literature for comparison. For R = H, the MNDO and

AM1 activation barriers lie between the RMP2/6-31G** and the RHF/6-31G* values,[15,16] while PM3 is 5.8 kcal/mol smaller than the RMP2 barrier, and compares quite well with the very high level and computationally expensive RMPSDTQ/6-31G** value.[15] PM3 reaction barriers for the **1a** + **4** → **5** and the **1a** + **6** → **7** reactions are a comparable 4-6 kcal/mol smaller than the RMP2 results, suggesting that the PM3 results would agree well with results from RMPSDTQ/6-31G** calculations.

All methods predict a decrease in the activation barrier upon substitution of -Cl or -F for -H, although the semi-empirical methods predict a smaller decrease than do the *ab initio* methods. Little change in the activation barrier upon substitution of $-CH_3$ for -H is shown by both semi-empirical and *ab initio* methods. Semi-empirical activation energy barriers also change little upon substitution of an aryl ring for -H in isocyanate. There are no *ab initio* values to compare with the semi-empirical aryl isocyanate reaction barriers.

Table 4.5: Activation energies of Isocyanate/ethene cycloadditions

	$\Delta E_{activation}$ (kcal/mol)					
	semi-empirical			*ab initio* (6-31G*)		
	MNDO	AM1	PM3	RHF	RMP2	RMP/SDTQ
TS1						
$R_1NCO + C_2H_4$						
a) R_1 = -H	48.9	48.1	37.8	57.1	43.6	40.7
b) R_1 = $-CH_3$	51.4	48.8	36.8	59.1	40.9	--
c) R_1 = -F	37.2	38.1	28.6	32.2	15.8	--
d) R_1 = -Cl	35.7	34.9	31.5	39.5	22.1	--
e) R_1 = $-C_6H_5$	50.5	47.5	36.4	--	--	--
TS2						
$HNCO + C_3H_4$	--	--	38.4	60.0	44.9	
TS3						
$HNCO + C_2H_3OH$	--	40.2	26.7	40.6	30.2	

4.4.6. Transition State Geometries:

Although both semi-empirical and *ab initio* methods predict concerted mechanisms, semi-empirical transition state geometries differ considerably from the *ab initio* geometries.[15,16] *Ab initio* transition states are slightly asynchronous, with C2-C3 bonds 0.2-0.3 Å shorter than N1-C4 bonds at the transition state. Semi-empirical methods all predict an opposite asynchronous transition state, with C2-C3 bonds longer than C4-N1 bonds (**Table 4.6**). The AM1 transition states are all markedly asynchronous, with C2-C3 bonds 0.6 - 0.8 Å longer than the N1-C4 bonds. The PM3 results are less exaggerated, with C4-N1 bonds only 0.3 - 0.4 Å shorter than the C2-C3 bonds. Interestingly, for R = -F, MNDO predicts a transition state with a geometry very similar to *ab initio*, but for all other substituents, MNDO transition states resemble AM1 transition states (i.e. pronounced asyncronicity).

Table 4.6 Semi-empirical Transition State C1-C3 and N2-C4 Bond Distances

			BOND LENGTHS (Å)			
	MNDO	AM1	PM3	MNDO	AM1	PM3
TS1a				**TS1b**		
C2-C3	2.411	2.433	2.212	C2-C3	2.247	2.363	2.171
N1-C4	1.705	1.649	1.864	N1-C4	1.843	1.716	1.920
TS1c				**TS1d**		
C2-C3	1.849	2.237	2.151	C2-C3	2.514	2.324	2.180
N1-C4	2.442	1.806	1.984	N1-C4	1.700	1.763	1.926
TS1e							
C2-C3	1.848	2.287	2.170				
N1-C4	2.303	1.778	1.932				
TS2 (C_3H_4)				**TS3** (-OH)			
C2-C3			2.253	C2-C3		1.634	2.050
N1-C4			1.856	N1-C4		2.316	2.037

Despite the differences between semi-empirical and *ab initio* transition state geometries, trends in semi-empirical transition state geometries reflect trends in the *ab initio* results; substitution of -F and -Cl on isocyanate shortens C1-C3 bonds and lengthens N2-C4 bonds in *ab initio* transition states, which parallels the trend observed with semi-empirical methods. Overall PM3 gave the closest match to *ab initio* TS geometries. Overlays of PM3 and *ab initio* **TS1a** stick diagrams (**Fig. 4.3**) highlights the similarities and differences between the two. The greater asynchronicity of semi-empirical as compared with *ab initio* transition states is not surprising, and has been observed in other cycloadditions involving heteroatoms.[22]

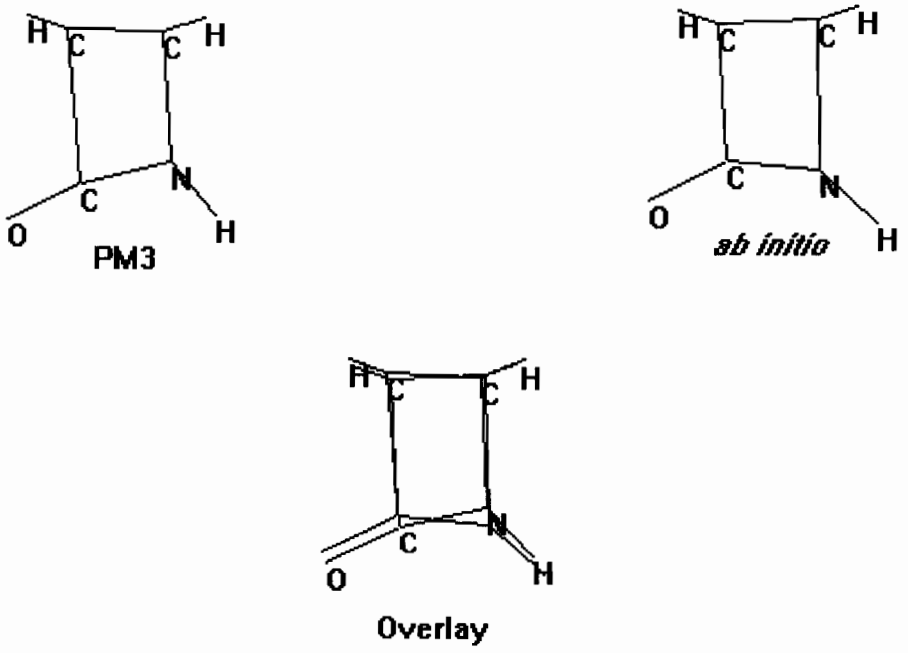


Figure 4.3:

Overlays of PM3 and *Ab Initio* TS1a Transition State Geometries

4.5 Conclusions:

Isocyanate, methyl-isocyanate and chloro-isocyanate geometries were modelled with semi-empirical and *ab initio* quantum mechanical methods. Semi-empirical methods verified the experimental findings of a bent N=C=O bond in all cases, although they were poor at predicting the C=O and N=C bond lengths. Without polarization functions, *ab initio* methods yielded drastically incorrect predictions of linear H-N=C bond angles for these molecules. However, addition of polarization functions to *ab initio* basis sets corrects this error, and reasonable geometries result.

The semi-empirical methods gave seemingly reasonable and consistent heats of formation for isocyanate and methyl-isocyanate, which can easily be corrected to match experiment and *ab initio*. When the semi-empirical correction terms for isocyanate heats of formation are applied to isocyanate + olefin → lactam reactions, the resulting corrected PM3 reaction energies agree well with high level *ab initio*.

Overall, the PM3 treatment of the isocyanate + olefin → lactam reaction was in best agreement with high level *ab initio*. With either hydrogen isocyanate or methyl isocyanate as a reactant, PM3 activation barriers agree well with post-Hartree-Fock values.

Recently, Jursic[23] has shown that performing *ab initio* density-functional stationary point calculations on semi-empirical optimized structures in the acetylene-oxazolium cycloaddition reaction produced a reaction coordinates essentially identical to that produced by full geometry optimization with *ab initio* methods. The time required to perform the complete optimizations with *ab initio* density-functional method is orders of magnitude greater than that required for the semi-empirical optimizations. In this work, the semi-empirical MNDO and AM1 transition state geometries were markedly more asynchronous than the corresponding *ab initio* transition states. The PM3 transition states showed great similarity with *ab initio* transition states, suggesting the PM3 transition state

geometries are useful input for either single point or geometry optimization *ab initio* transition state calculations. For these systems, PM3 yields transition state geometries and activation barriers comparable to high-level *ab initio* results at a small fraction of the computer cost. These considerations justify the continued use of semi-empirical methods (especially PM3) for the remainder of this thesis.

References:

1. a) E.F Moricioni and P.H. Mazzocchi, *J. Org. Chem.*, **31**, 1372 (1966).

   b) A.K. Mukerjee and R.C. Shrivastava, *Synthesis*, 32 (1973).

   c) for a review see G. Luckacs and M. Ohno, eds., "*Recent Progress in the Synthesis of Anitbiotics,*" p 565-612 (Springer-Verlag: Berlin) 1990.

2. G. Ege, and K. Gilbert, *Tetrahedron Lett.*, **44** 4253 (1979).

3. B.J. Jean-Claude, personal communication

4. B.M. Rode, W. Kosmus and E. Nachbaur, *Z. Naturforsch A*, **29**, 650 (1974)

5. N.C. Handy, W.H. Green, et. al., *Mol. Phys.*, **78**(2), 319 (1993).

6. R. Bonaccorsi, C. Petrongolo, et. al., *J. Chem. Phys.*, **48**(4), 1500 (1968).

7. K. Yokoyama, S. Takane, and T. Fueno, *Bull. Chem. Soc. Jpn.*, **64**, 2230 (1991).

8. D. Poppinger, L. Radom and J.A. Pople, *J. Am. Chem. Soc*, **99**, 7806 (1977).

9. H.G. Mack and H. Oberhammer, *Chem. Phys. Lett.*, **157**(5), 436 (1989).

10. W.H. Hocking and M.C.L. Gerry, *J. Mol. Spectr.*, **42**, 547 (1972).

11. W.H. Hocking, *Diss. Abstr. Int. B.*, **35**, 2681 (1974).

12. R.G. Shulman, L.H. Jones et. al., *Chem. Phys. Lett.*, **17**, 180 (1950).

13. D.W.W. Anderson, D.W.H. Rankin and A. Robertson, *J. Mol. Spectr.*, **41**, 385 (1972).

14. a) K. Yamada, *J. Mol. Spectr.*, **79**, 323 (1980).

    b) R.G. Lett and W.H. Flygare, *J. Chem. Phys.*, **47**, 4730 (1967).

    c) W.H. Hocking, M.L. Williams and M.C.L. Gerry, *J. Mol. Spectr.*, **58**, 250 (1975).

15. F. P. Cossio, B. Lecea, X. Lopez, G. Roa, A. Arrieta and J.M. Uglade, *J. Chem. Soc., Chem. Commun.,* 1450 (1993).

16. D. Fang, Z.F. Xu, X.Y. Fu, *J. Mol. Struct. (THEOCHEM)*, **333**, 159 (1995).

17. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** (1990) 52.

b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).

18. J. Baker, *J. Comput. Chem*, , **7**, 385 (1986).

19. JANAF Thermodynamic Tables, 3rd Edition.; (Dow Chemical: Midland, MI) 1985.

20. E.V. Borisov, A.M. Mebel, B.A. Knysav, V.B. Zabrodin and A.A. Korkin, *Izv. Akad. Nauk., Ser. Khim.*, **7**, 1585, (1992).

21. M.J.S. Dewar and W. Thiel, *J. Am. Chem. Soc.,* **99**, 4907 (1977).

22. B.S. Jursic and Z. Zdravkovski, *J. Mol. Struct., (THEOCHEM)*, **315**, 85, (1994).

23. B.S. Jursic, *J. Mol. Struct.(THEOCHEM),* **365**, 55, (1996).

Appendix 4.A

Table 4.A1: Semi-Empirical Product and Reactant Heats of Formation:

		ΔH_f (kcal/mol)		
		MNDO	AM1	PM3
2	(C_2H_4)	15.4	16.5	16.6
3a	(-H)	-26.9	-8.9	-21.8
3b	(-Me)	-30.5	-5.2	-24.3
3c	(-Cl)	-7.9	18.0	-20.5
3d	(-F)	-27.1	3.3	-15.3
3e	(C_6H_5)ip	1.7	24.3	3.9
3f	(C_6H_5)oop	2.0	29.1	7.9
4	(C_3H_4)	43.9	46.1	47.1
5a		-3.6	19.9	3.4
6	(C_2H_3OH)	-34.3	-30.3	-21.7
7a		-76.8	-55.2	-69.7

The results of chapters three and four suggested that the PM3 method would be the most suitable semi-empirical method for studying nitrogen heterocyclic compounds. The testing of the semi-empirical method continues in this chapter, with the study of diazoazoles and azolylidenes. Diazoazole structures and heats of formation will be calculated with *ab initio* atom equivalent methods, and the MNDO, AM1 and PM3 semi-empirical methods. Interestingly, the systematic MNDO and AM1 azole heat of formation errors discussed in chapter three resurface as similar trends in the MNDO and AM1 diazoazole heats of formation. The PM3 and *ab initio* diazoazole heats of formation are in good agreement, suggesting that these are the first trustworthy estimates of diazoazole heats of formation. With azolylidene stabilities, PM3 is again shown to be superior to MNDO and AM1, yielding results which agree with *ab initio* and experiment. The PM3 and *ab initio* calculations on 4-triazolylidene are used to explain the different products obtained from photolysis and thermolysis of 4-diazo-1,2,3-triazole. In total, this chapter is further justification for the continued use of PM3 to investigate nitrogen heterocyclic structure and reactivity.

CHAPTER 5:

A COMPARISON OF SEMI-EMPIRICAL AND *AB INITIO* CALCULATIONS ON DIAZOAZOLES AND AZOLYLIDENES

by C.I Williams and M.A. Whitehead

Theoretical Chemistry Laboratory, Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6

and

Bertrand J. Jean-Claude

Department of Oncology, McGill University, 3655 Drummond St., Ste. 701, Montreal, Quebec, Canada H3G 1Y6

Accepted into *J. Mol. Struct. (THEOCHEM)*, May, 1996.

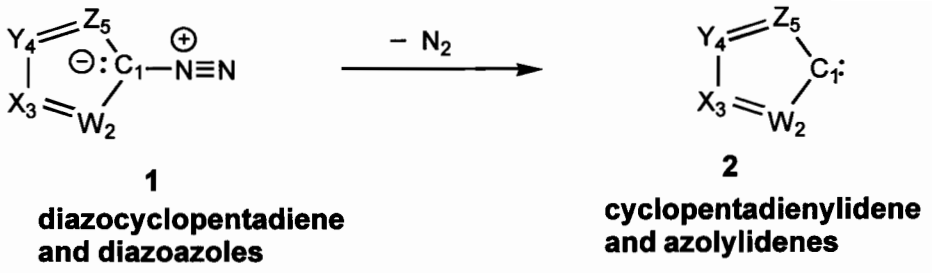
5.1 Abstract:

Semi-empirical (MNDO, AM1, PM3) and *ab initio* (R/UHF/6-31G**) methods are used to calculate the geometries, energies and heats of formation of a series of diazoazoles and azolylidenes. Unsubstituted azoles, cyclopentadiene, diazo-cyclopentadiene and aminoazoles are studied for comparison. *Ab initio* and semi-empirical conformational analyses demonstrate the floppy nature of the diazo C=N=N bond angle. Ring bond lengths in diazoazoles show that bond localization is greater in diazoazole rings than in azole or aminoazole rings. *Ab initio* atom equivalent diazoazole heats of formation compare well with the PM3 semi-empirical values. Calculations predict that the triplet spin state is the most stable for all cyclic azolylidenes. All singlet state azolylidenes except 4-triazolylidene are stable high energy cyclic species. Singlet state 4-triazolylidene is unstable and spontateously undergoes ring scission to diazoacetonitrile, while triplet state 4-trizaoylidene is stable as a cyclic structure. The results for 4-triazolylidene are used to explain why different products are obtained from thermolysis and photolysis of 4-diazo-1,2,3-triazole.

5.2 Introduction:

Diazoazoles (**1b-1f, Fig. 5.1**) were synthesized by Thiele in 1892,[1] and their structure determined in the 1950's by analogy with the suggested structure of diazocyclopentadiene. Interest in the diazoazoles has increased because of their synthetic usefulness: they can be used to prepare fused heterocyclics, via cycloaddition reactions with alkenes, alkynes and isocyanates,[2-7] while reactions between diazoazoles and 1,1-dimethoxyethene produce interesting spiro-bicyclic compounds via [3+2] cycloadditions, which are similar to cycloaddition reactions between diazocyclopentadiene (**1a, Fig. 5.1**) and alkynes.[8,10] Some substituted diazoazoles internally cyclize because of the flexibility of the C=N=N bond angle.[11] Diazoazoles can lose nitrogen upon thermolysis or photolysis to form azolylidenes, which are carbenic intermediates (**2b-2f, Fig. 5.1**) that can react with points of unsaturation in alkenes.[2] Azolylidenes exist in both singlet and triplet states,[12] but it is believed that the singlet state prevails.

Because of the relative instability, and sometimes explosive nature of these compounds, few X-ray structural data or experimental enthalpies of formation exist for unsubstituted diazoazoles.[2a] No experimental enthalpies of formation or X-ray structural data exist for azolylidenes. Therefore, semi-empirical and *ab initio* methods are used to study the geometries and the stabilities of diazoazoles and azolylidenes; azoles, aminoazoles, cyclopentadiene and diazo-cyclopentadiene will also be studied for comparison. In particular, the C=N=N bond angle flexibility will be studied, because the flexibility of this group is believed to play a role in diazoazole cycloaddition and internal cyclization reactions. The degree of aromaticity in azoles, diazoazoles, aminoazoles, diazocyclopentadiene, cyclopentadiene and the cyclopentadienyl anion will be assessed by comparing azole ring bond lengths. Heats of formation will be generated by semi-empirical and atom equivalent *ab initio* methods.[13,14] Geometries, relative stabilities and heats of formation of singlet and triplet state azolylidenes will also be discussed.



1 and 2

a	W,X,Y,Z = C-H	
b	W = N	X,Y,Z = C-H
b'	X = N	W,Y,Z = C-H
c	W,Z = N	X,Y = C-H
c'	W,Y = N	X,Z = C-H
d	W,X = N	Y,Z = C-H
d'	X,Y = N	Z,W = C-H
e	W,X,Y = N	Z = C- H
e'	W,X,Z = N	Y = C-H
f	W,X,Y,Z = N	

Figure 5.1:

Diazocyclopentadiene, Diazoazoles and Corresponding Azolylidenes

5.3 Calculations:

Calculations were performed on a SUN Sparc10 workstation using the GAMESS[15] package. Preliminary input geometries were generated by functionalizing azoles with a diazo group whose bond lengths and angles were the same as in diazomethane. Alternate inputs used aromatic diazonium bond lengths to generate the diazo group. Both preliminary geometries gave the same structure when optimized at the semi-empirical level. Fully optimized MNDO, AM1 and PM3 geometries were used as input for the 6-31G** geometry optimizations. The 6-31G** calculation would always return the same structure regardless of input geometry (MNDO, AM1 or PM3); RHF wavefunctions were used for singlet state molecules; UHF wavefunctions were used for triplet state azolylidenes, whose input geometries were generated by eliminating the diazonium nitrogen atoms from

the corresponding diazoazole geometries. Equilibrium geometries were verified as local minima by force constant analysis. *Ab initio* calculations used 6-31G* and 6-31G** basis sets. Although diazoazoles have been reported as planar, C_s symmetry was not enforced (because it would bias any calculation that would otherwise report a non-planar geometry). Calculations which yielded planar optimized geometries were re-initiated with enforced C_s symmetry. Some structures were initiated with C_{2v} symmetry where possible. All diazoazoles were found to be planar. Diazocyclopentadiene and some azoles (**1d, 1c, 1f**) were found to have C_{2v} symmetry. *Ab initio* total energies used in this study are in appendix **Table 5.A1**. *Ab initio* heats of formation were generated using both Castro's[13] and Yala's[14] methods.

5.4 Results and Discussion:

5.4.1. Diazoazole Geometries: Diazoazoles with C_{2v} symmetry have 'linear' (C=N=N bond angles (180°)), while the others have near 'linear' (C=N=N bond angles (>178°)) diazo groups. This compares well with the 179.1° C=N=N X-ray value for 3-diazoindazole.[16] All methods predict carbon-diazo C-N distances to be between 1.30 Å and 1.34 Å, which is closer to 1.32 Å aliphatic C-N diazo bond of diazomethane than the average 1.402 Å C-N bond of other aromatic diazonium compounds.[2,2a] Diazo N=N bond distances range between 1.13Å (semi-empirical) and 1.08 Å (6-31G**), which is close to the X-ray values for 3-diazoindazole[16] (X-ray: C-N = 1.338 Å; N=N 1.110 Å) and 4-acetyl-3-diazo-2,5-diphenylpyrrole[2a] (X-ray: C-N = 1.338 Å; N=N 1.110 Å). The semi-empirical and *ab initio* methods all predict a lengthened diazo C=N bond and a shortened N=N diazo bond with increasing number of nitrogen atoms in the azole ring.

5.4.2. Diazo Group Flexibility: 3-D plots in **Fig. 5.2** show semi-empirical and *ab initio* diazoazole energies as a function of the C=N=N bond angle and the W-C=N=N torsion angle. Semi-empirical methods predict a highly flexible C=N=N

bond angle, which requires only 20 kcal/mol to distort from the equilibrium value of 180° to 135°. *Ab initio* predicts slightly less flexibility in this bond angle, requiring 28 kcal/mol for the same distortion. In all cases, distortions of <30° require 6 kcal/mol or less.[10] Semi-empirical results predict almost free rotation about the W-C=N=N torsion angle at any given value of the C=N=N bond angle. *Ab initio* energies about the same torsion angle are slightly affected by the C=N=N bond angle, as shown by the maxima along the W-C=N=N torsion angle axis at C=N=N values of 160° or less. The largest *ab initio* barrier to free rotation around the W-C=N=N torsion angle (at a C=N=N bond angle of 135.0°) is 11.6 kcal/mol.

The differences between the semi-empirical and *ab initio* results are minor considering the magnitude of the C=N=N bond angle distortions. *Ab initio* results predict a more restricted W-C=N=N torsion angle than do semi-empirical calculations, but both methods suggest a 'floppy' C=N=N group, whose bond angle may easily bend by >30°. The floppiness of the diazo group is probably what allows substituted diazoazoles to undergo autocyclizations and cycloaddition reactions.

5.4.3. Bond Lengths and Delocalization in Diazoazoles: Variations in bond lengths around ring systems have been systematically used to assess their aromatic character in various chemical environments;[16b] bond delocalization results in zero to minor bond length variations around a ring system, while bond localization results in large bond length variations, corresponding to localized single and double bonds. Variations in 6-31G** ring bond lengths in diazocyclopentadiene, cyclopentadiene and the cyclopentadienyl anion are compared graphically in **Fig. 5.3**. Cyclopentadiene is a highly localized system, and therefore, the bond lengths vary in a drastic zig-zag fashion reflecting localized single and double bonds. In contrast, the cyclopentadienyl anion is completely delocalized, so the ring bond lengths are constant throughout the system.

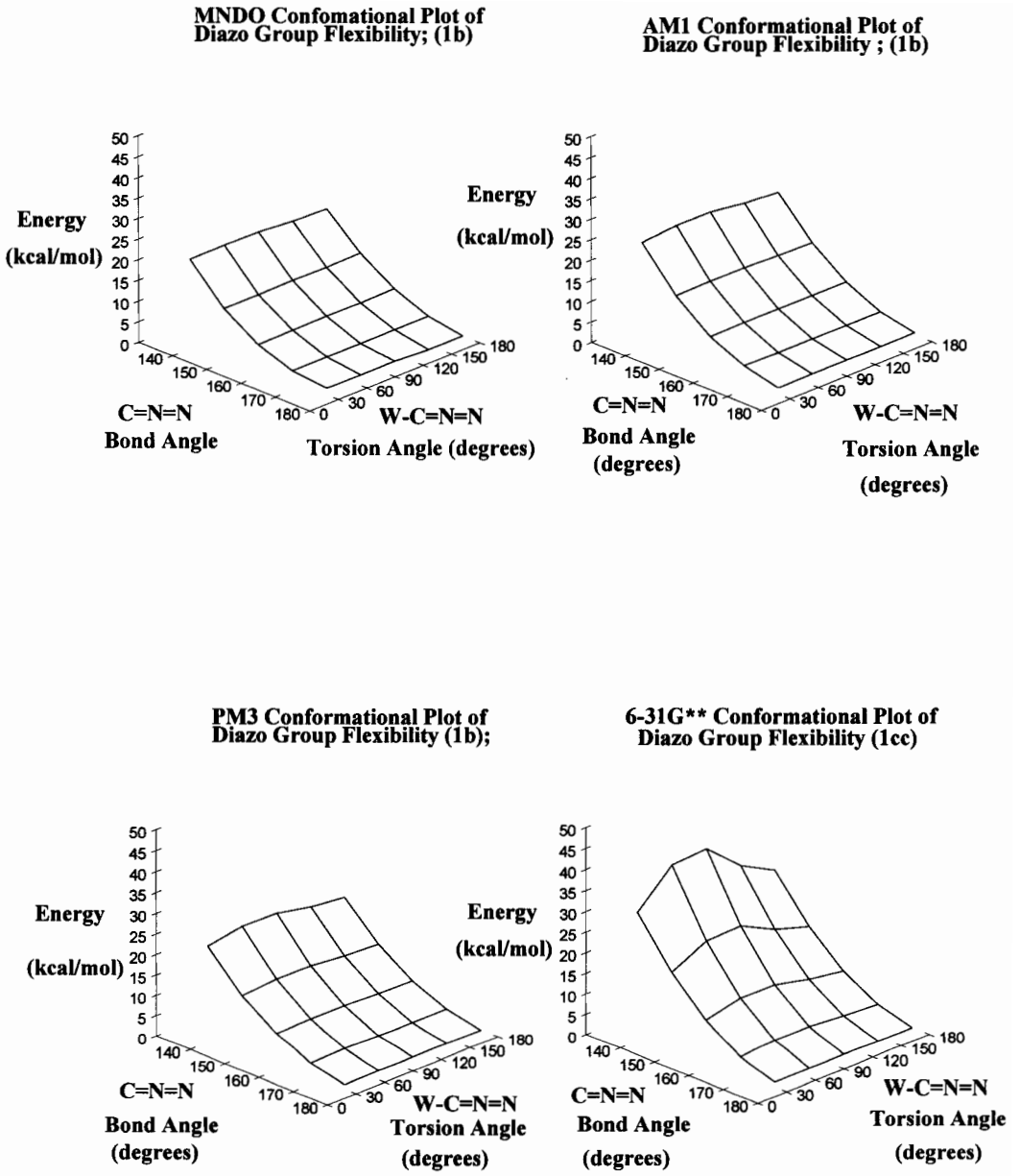
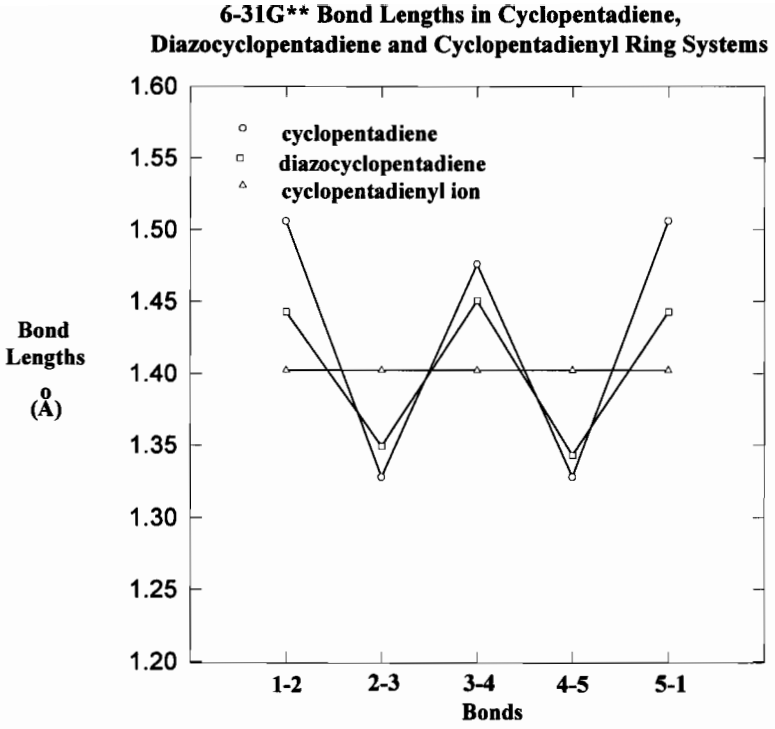
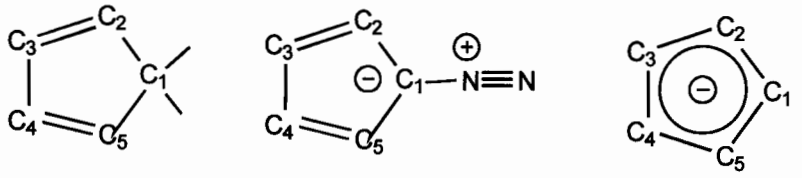


Figure 5.2:

Diazo Group Flexibility

Energy vs. Bond Angle and Torsion Angle





cyclopentadiene diazocyclopentadiene cyclopentadienyl anion

Figure 5.3:

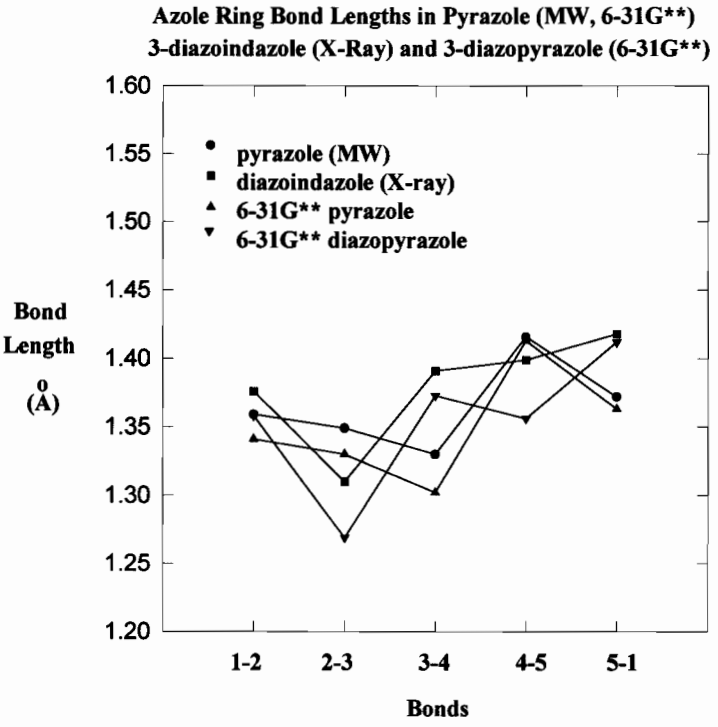
6-31G** Ring Bond Length Variations in Cyclopentadiene, Diazocyclopentadiene and the Cyclopentadienyl Anion

Diazocyclopentadiene ring bond length variations closely resemble those in cyclopentadiene, although the magnitude of the differences between the 'single' and 'double' bond lengths are smaller in diazocyclopentadiene. However, the comparison does reveal that bond localization in diazocyclopentadiene is similar to that in cyclopentadiene; thus, there is little aromatic character in diazocyclopentadiene, and the ring system is best described as an alternant hydrocarbon, rather than an aromatic system like the cyclopentadienyl anion.

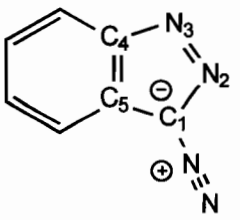
Experimental azole ring bond lengths for 3-diazoindazole (X-ray)[16] and pyrazole (microwave (MW))[16c] are compared with the calculated 6-31G** azole ring bond lengths of pyrazole and 3-diazopyrazole (**Fig. 5.4**). Although numerical discrepancies exist between the calculated and the experimental results, the plot in **Fig. 5.4** reveals that unsubstituted pyrazole C1-N2, N2-N3 and N3-C4 bond lengths vary only slightly (from 1.34-1.35 Å for C1-N2 to 1.34-1.31 Å for N3-C4). Lengths of the corresponding bonds in 3-diazopyrazole and 3-diazoindazole vary in a zig-zag fashion, reflecting the single-bond/double-bond alternation drawn in the diazopyrazole chemical structure. Similar variations in other X-ray azole ring bond lengths have been attributed to increased bond localization.[17] Thus, as with diazocyclopentadiene, diazoazoles appear to be more localized and less aromatic than their unsubstituted azole counterparts.

The effects of an amino and diazo substitution on azole ring bond lengths are compared in **Fig. 5.5a** and **Fig. 5.5b**. Computed (6-31G**) ring bond lengths in imidazole, pyrrole and their 2-amino derivatives show that amino substitution has little effect on azole ring bond lengths, indicating that bond delocalization is similar in both unsubstituted azoles and aminoazoles. However, diazo substitution alters the pattern of azole ring bond lengths substantially. Azole ring bond length variations in diazoazoles bear closer resemblance to the zigzag pattern of bond length variation found in diazocyclopentadiene and cyclopentadiene. This effect was observed for all the diazoazoles in this study, in both the *ab initio* and semi-empirical results. Thus, at least by bond length criteria, it appears that azole rings

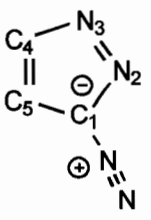
in diazocyclopentadiene and diazoazoles are more localized than their unsubstituted azole, cyclopentadienyl anion and aminoazole counterparts.



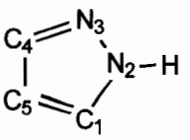
Atom numbering schemes



3-diazoindazole



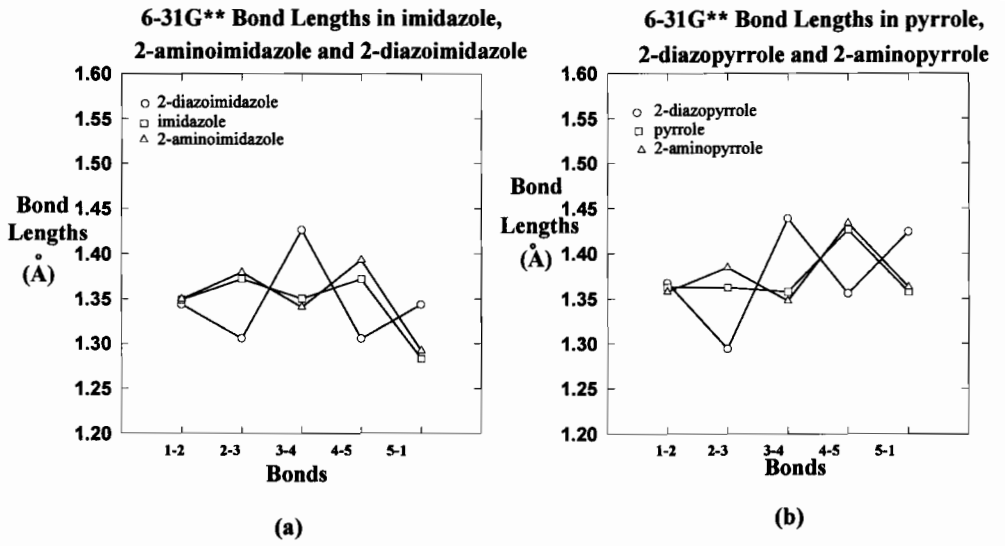
diazopyrazole

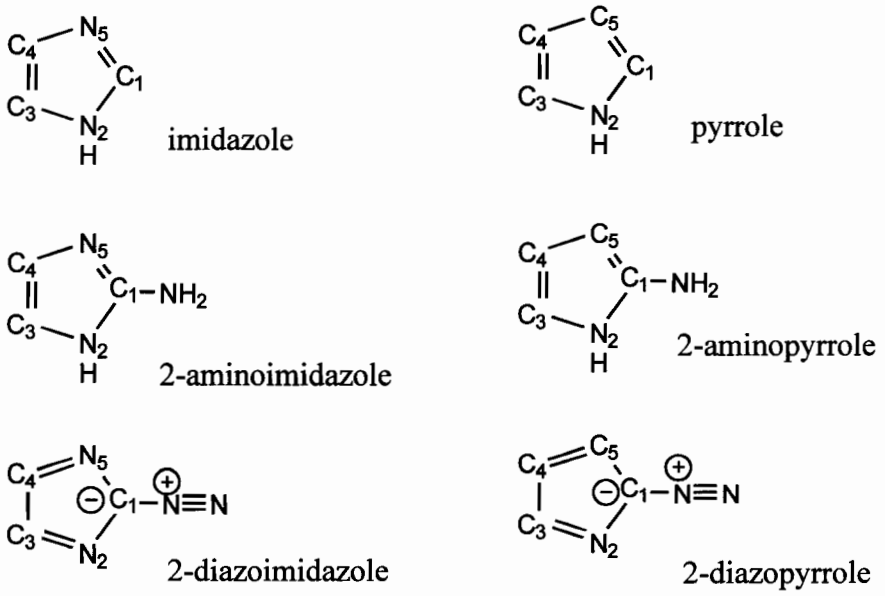


pyrazole

Figure 5.4:

Comparison of Experimental and 6-31G** Azole Ring Bond Lengths in Unsubstituted Azoles and Diazoazoles





Figures 5.5a and 5.5b:

Comparison of 6-31G** Azole Ring Bond Lengths in Unsubstituted Azoles, Aminoazoles and Diazoazoles

5.4.4. Azole Heats of Formation: Experimental and semi-empirical heats of formation of unsubstituted azoles were discussed at length in chapter three, and the results are reiterated here (**Table 5.1**). Semi-empirical PM3 azole heats of formation fit experimental data well,[18] MNDO performs well on azoles as long as N=N-N linkages are absent in the azole ring,[19] while AM1 consistently overestimates azole heats of formation by 20-30 kcal/mol.[20] Corrections to MNDO and AM1 azole heats of formation have been suggested,[19a] but in light of the close fit between experimental and PM3 data, and the discussion in Chapter Three, the PM3 results are probably a more trustworthy semi-empirical method for azoles, and possibly diazoazoles.

Table 5.1:
Semi-Empirical and Experimental Azole Heats of Formation:

Compound	ΔH_f (kcal/mol)			
	MNDO	AM1	PM3	expt.
cyclopentadiene	32.1	37.1	31.8	31.9[a]
pyrrole	32.5	39.9	27.1	26.0[b]
imidazole	33.3	50.8	31.3	35.0[b]
pyrazole	45.4	65.6	48.8	44.0[c]
1,2,3-triazole	49.8	86.4	67.9	63.9[e]
1,2,4-triazole	43.7	76.9	51.8	46.0[d]
tetrazole	53.8	109.7	86.2	80.1[a]

(a) ref. 21 (b) ref. 22 (c) ref. 23
(d) ref. 24 (estimated experimantal value)
(e) *ab initio* value from *isolobal* reaction schemes (Chapter Three)

5.4.5. Diazoazole Heats of Formation: Atom equivalent methods can be used to estimate diazoazole heats of formation from *ab initio* total energies. Castro[13] and

Yala[14] have both introduced simple methods by which heats of formation can be calculated from 6-31G* total energies and a small set of *ab initio* atom equivalents. *Ab initio* atom equivalent heats of formation are calculated from *ab initio* 6-31G* total energies (E_{el}) using the equation

$$\Delta H_f = E_{el} - \Sigma\, k_s B_{is},$$

where k_s is the number of atoms of a given type, and B_{is} is the atom equivalent parameter for the atom type. The summation is over all atoms in the molecule. Atom equivalent parameters from both Castro's work (A, ref. 13) and Yala's work (B, ref.14) were used, because these methods work well on cyclopentadiene, (heats of formation errors of less than 0.5 kcal/mol from experiment) and diazomethane (errors 3-4 kcal/mol from experiment). Large errors (>6 kcal/mol) arise when these atom equivalent methods are applied to highly delocalized systems, such as azoles and benzene. However, both the C-N diazo bond lengths and the diazoazole ring bond lengths suggest that the diazoazole ring systems have alternant hydrocarbon character; therefore, the atom equivalent methods should give reasonably accurate *ab initio* diazoazole heats of formation, with errors less than that observed for the atom equivalent *ab initio* benzene heat of formation.

Atom equivalent *ab initio* and semi-empirical diazoazole and diazocyclopentadiene heats of formation are in **Table 5.2**. Values of diazomethane heats of formation are also given for comparison. The semi-empirical and *ab initio* diazocyclopentadiene heats of formation agree well, although PM3 gives a result which is a bit low. However, the low PM3 result is to be expected, since PM3 underestimates the diazomethane heat of formation by 10 kcal/mol. MNDO, which performs well on cyclopentadiene and gives the best semi-empirical result for diazomethane, gives the closest match to the *ab initio* diazocyclopentadiene heat of formation.

As the number of nitrogens in the diazoazole azole ring increases, the PM3 heats of formation give the closest match to the *ab initio* heats of formation. The trends in semi-empirical diazoazole heats of formation reflect the trends observed in semi-empirical azole heats of formation. MNDO and PM3 values agree well for diazopyrrole and diazoimidazole, and moderately well for diazopyrazole, as they do for unsubstituted pyrrole, imidazole, and pyrazole. However, as with triazoles and tetrazole, MNDO diazotriazole and diazotetrazole heats of formation are consistently lower than the PM3 values by magnitudes similar to the differences between MNDO and PM3 triazole and tetrazole heats of formation. AM1 consistently gives diazoazole heats of formation which are 20-30 kcal/mol higher than the PM3 values, reflecting the AM1 errors in azole heats of formation. The differences in MNDO, AM1 and PM3 diazoazole heats of formation are probably due to the inability of MNDO and AM1 to accurately treat azole rings; the differences between the PM3, AM1 and MNDO diazoazole heats of formation are the same as the differences between PM3, AM1 and MNDO unsubstituted azole heats of formation.

Table 5.2: Semi-Empirical and *Ab initio* Diazoazole and Diazocyclopentadiene Heats of Formation:

Compound	ΔH_f(kcal/mol)					
	MNDO	AM1	PM3	6-31G*		expt.[21]
				A	B	
1a	98.1	105.3	94.3	101.8	99.8	--
1b	101.7	120.3	97.5	103.2	99.2	--
1b'	102.9	116.7	99.3	100.2	96.3	--
1c	105.2	135.9	102.1	100.0	93.8	--
1c'	104.4	132.2	101.8	99.2	93.0	--
1d	111.9	141.3	118.1	131.5	125.3	--
1d'	113.5	139.6	121.0	123.8	117.0	--
1e	119.4	165.6	138.0	137.4	128.9	--
1e'	112.9	156.8	121.6	119.1	111.1	--
1f	124.8	190.8	156.1	152.3	141.5	--
diazomethane	67.3	62.6	61.0	67.1	68.1	71

No experimental data exist for comparison with the calculated diazoazole heats of formation; however, the close fit between *ab initio* and PM3, the known accuracy of PM3 azole heats of formation, and the excellent *ab initio* atom equivalent treatment of diazomethane and cyclopentadiene all suggest that the calculated diazoazole heats of formation are good estimates. Although these are NOT experimental values, the close match between PM3 and *ab initio* gives credibility to the PM3 values, justifying the use of PM3 on these molecules.

5.4.6. Azolylidenes

The only previous theoretical work on the azolylidenes used Huckel molecular orbital (HMO) calculations,[26,27] which predict that azolylidenes without a nitrogen atom adjacent to the carbenic center will exist as ground state triplets.[27] All other azolylidenes exist as ground state singlets. In contrast, semi-empirical and UHF/6-31G** energy differences between the triplet and singlet states of azololylidenes and cyclopentadienylidene (**Table 5.3**) predict that the triplet state is the most stable spin state for cyclopentadienylidene and for all the azolylidenes. These values are not to be treated as exact, because spin contamination in UHF calculations is known to affect these energies; however, since the triplet states are more energetically stable in all cases, the qualitative conclusion is that the ground state of all these species is a triplet spin state. This is not surprising, since cyclopentadienylidene is known to exist for hours as a stable triplet state in hexachlorobutadiene solvent.[25]

Except for 4-azolylidene (discussed below), all singlet and triplet state azolylidenes are cyclic. The geometries of the singlet and triplet state azolylidenes show that the triplet state is less strained. *Ab initio* and semi-empirical singlet state structures (**Figs. 5.6** and **5.7**) all have large Z-C-W bond angles (> 130°), and small C-W-X bond angles (<100°), indicating that the singlet state azolylidene ring system is a highly strained distorted pentagon. Triplet state geometries of the same

species are much less strained, with ring bond angles closer to ideal pentagon bond angles of 108°. Thus, the high energies and highly distorted geometries of singlet state azolylidenes suggest that these are less stable than the triplet states.

TABLE 5.3:
Semi-Empirical and *Ab initio* Singlet/Triplet Cyclopentadienylidene and Azolylidene Energy Differences:

	ΔH(triplet-singlet)kcal/mol;			
Compound	MNDO	AM1	PM3	6-31G**
2a	-40.2	-39.4	-39.9	-58.3
2b	-36.2	-38.8	-39.6	-44.8
2b'	-38.4	-38.7	-40.0	-56.2
2c	-37.6	-41.2	-39.6	-34.7
2c'	-38.9	-41.0	-39.8	-49.8
2e	-33.6	-31.3	**	**
2e'	-37.0	-39.1	-36.3	-2.3
2f	-32.5	-30.6	-11.8	-3.6

** = cyclic structure of singlet state azolylidene is unstable with these methods; cyclic structures spontaneously open to diazoacetonitriles

	ΔH_f (kcal/mol);		
	PM3	*Ab initio*	
		Castro	Yala
diazoacetronitrile	95.2	97.6	97.6

The higher stability of the triplet state azolylidenes is also supported by ESR analysis of low temperature photolyzed diazoimidazole-4-carboxamide. The experimental results show an increasing concentration of the tripet state as a function of time, suggesting that the relaxed state of the azolylidene is the triplet state.[11] Although singlet states of azolylidenes are initally produced, relaxation to the triplet state is slow because it involves geometric as well as spin relaxation.

Electron spin resonance (ESR) measurements on reactions between imidazolylidenes and benzene indicate that the rate determining step is the addition of the azolylidene in a p^2 singlet state to the benzene;[25] Thus, in most cases, azolylidenes will probably react or degrade before relaxing to the triplet state.

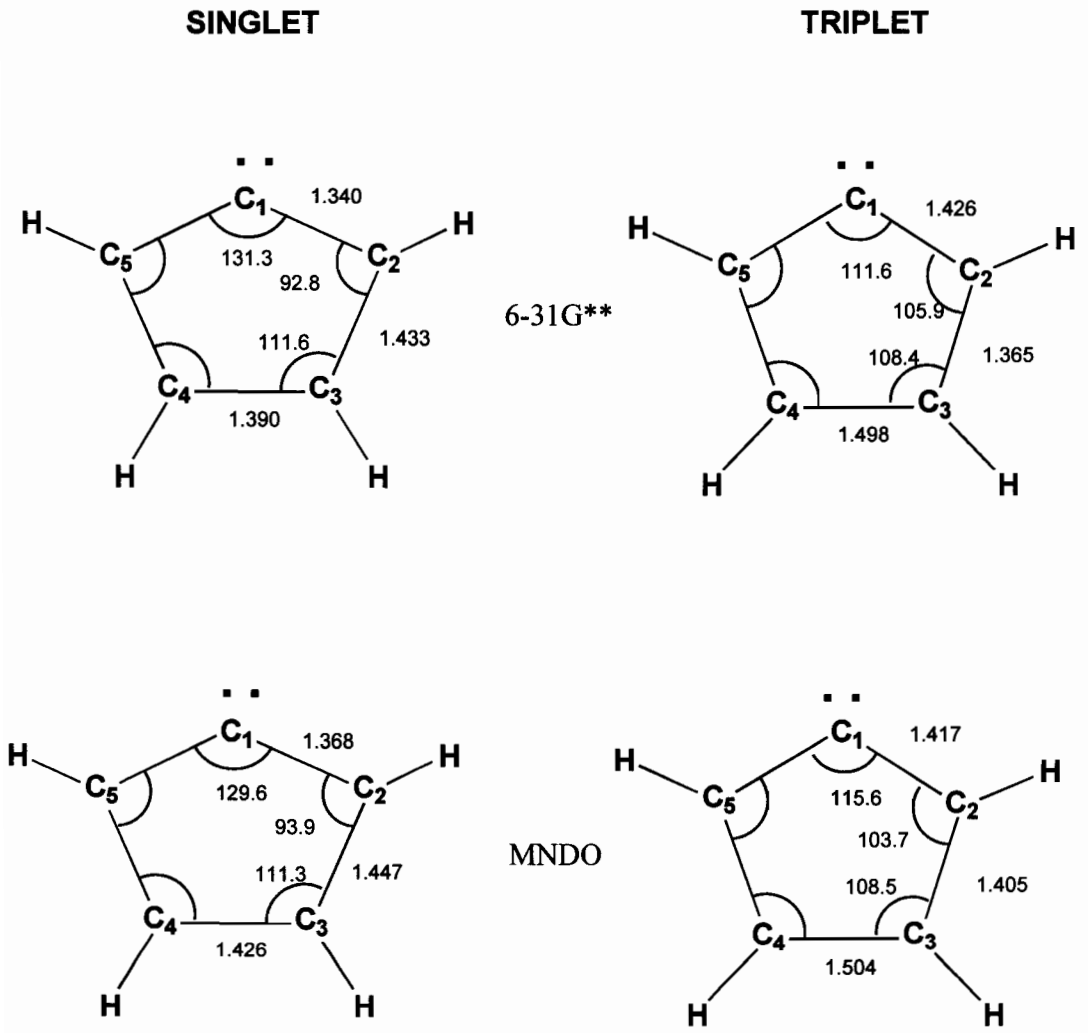


Figure 5.6:

Comparison of MNDO and 6-31G** Singlet and Triplet State Cyclopentadienyl Geometries:

(all bond lengths in Å; all bond angles in degrees)

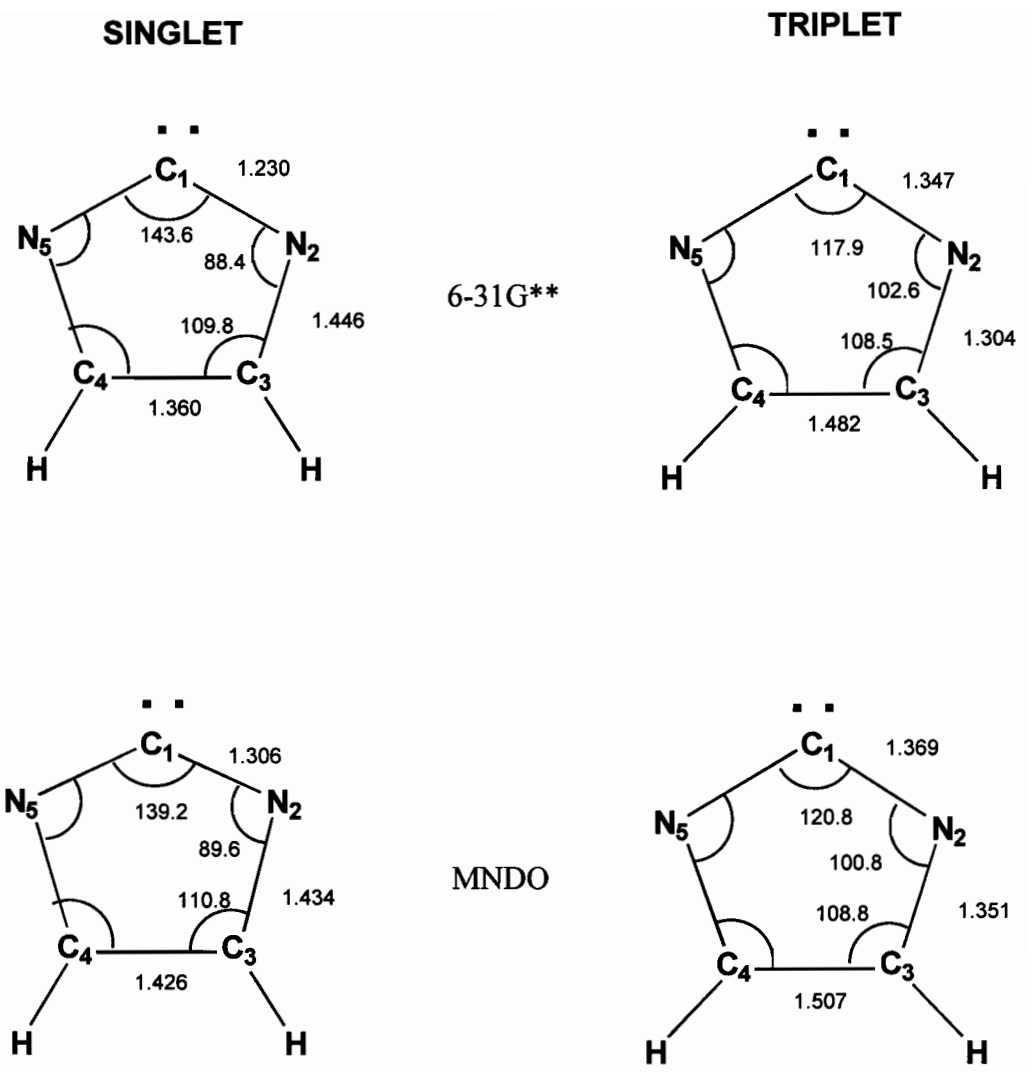


Figure 5.7:

Comparison of MNDO and 6-31G** Singlet and Triplet State 2-imidazolylidene Geometries.

(all bond lengths in Å; all bond angles in degrees)

5.4.7 Ring Scission in 4-Triazolylidene: Upon geometry optimization in the singlet state, all methods produce cyclic structures for every azolylidene except for 4-triazolylidene. PM3 and *ab initio* geometry optimizations of cyclic singlet state 4-azololylidene spontaneouly produce diazoacetonitrile (**Fig. 5.8**) instead of cyclic 4-triazolylidene. The similarity between the *ab initio* and PM3 diazoacetonitrile structures, and the good agreement between *ab initio* atom equivalent and PM3 diazoacetonitrile heats of formation (**Table 5.3**) lend confidence to the PM3 and *ab initio* results. MNDO and AM1 singlet state geometry optimizations of 4-triazolylidene give high energy cyclic structures. All methods predict stable cyclic structures for triplet state 4-triazolylidene.

These *ab initio* and PM3 results explain the experimental fact that 4-triazolylidenes generated by photolysis almost exclusively undergo ring scission to produce diazoacetonitriles, which lose nitrogen and react as cyanocarbenes. Due to spin rules, photolysis must initially produce the singlet state 4-triazolylidene exclusively. *Ab initio* and PM3 predict that the singlet state 4-triazolylidene is unstable, and thus it will spontaneously undergo ring scission before it can either react as a cyclic singlet state, or relax to a stable cyclic triplet state. Therefore, reactions under photolytic conditions should produce ring scission products only, as is observed experimentally.[28]

In contrast, 4-azolylidene generated by thermolysis retains its cyclic azolylidene structure, and reacts in a manner analogous to other azoloylidenes.[2,28] Thermolysis can directly produce a triplet state 4-triazolylidene, which both *ab initio* and PM3 predict to have a stable cyclic structure. The cyclic triplet state 4-triazolylidene can undergo reactions as a cyclic azoloylidene.

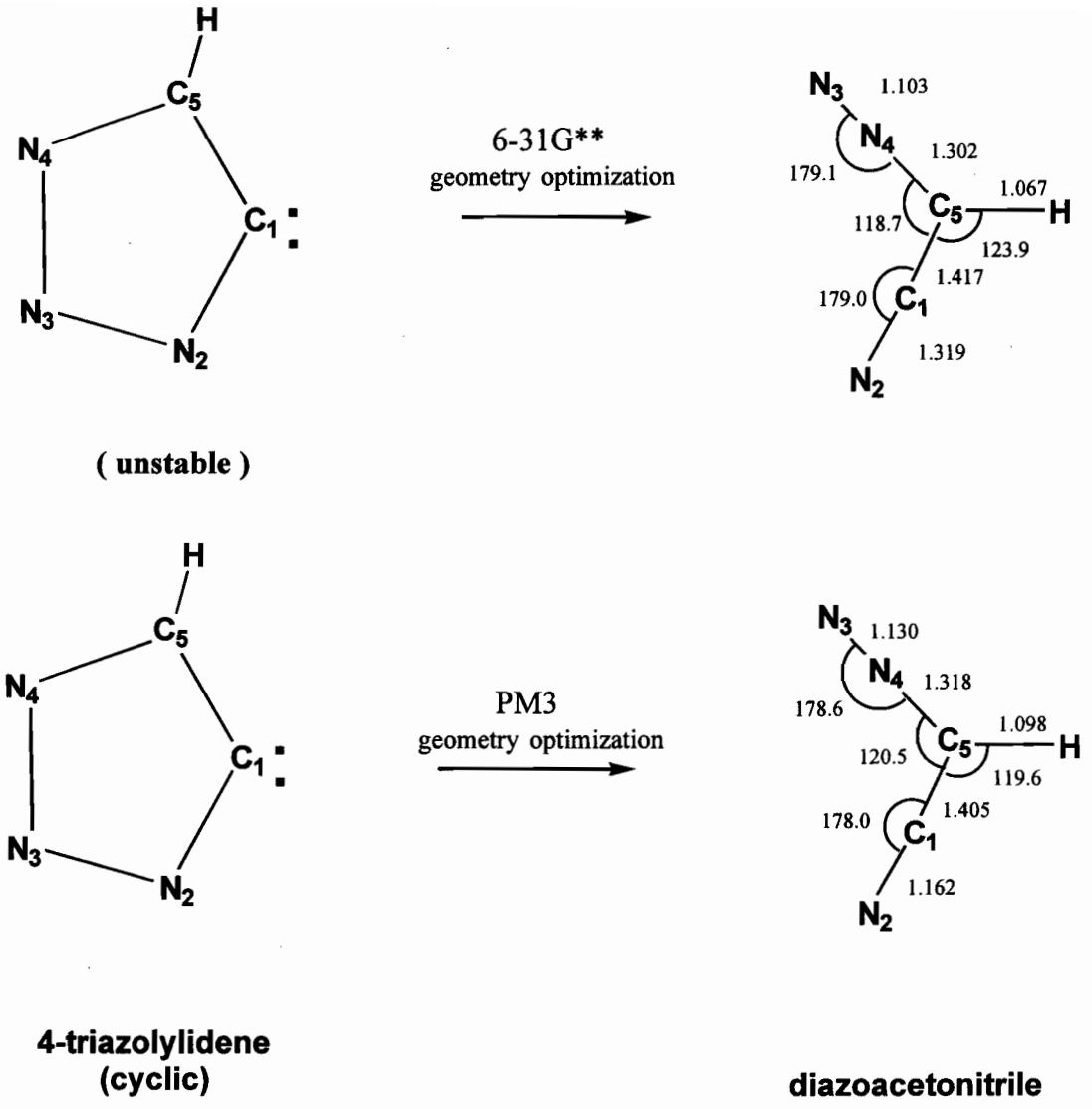


Figure 5.8:

Instability of Singlet State 4-Triazolylidene: PM3 and *ab initio* geometries of singlet state 4-triazolylidenes are unstable, and spontaneously undergo ring scission to give diazoacetonitriles

5.5 Conclusions:

Semi-empirical and *ab initio* methods predict high flexibility of the diazoazole C=N=N bond angle. Ring system bond lengths in diazoazoles and diazo-cyclopentadiene indicated a smaller degree of delocalization in these systems than in unsubstituted azoles or amino azoles. Bond localization in diazoazole rings allows the successful use of *ab initio* atom equivalent methods for calculating diazoazole heats of formation. The *ab initio* and PM3 heats of formation agree well throughout the series of diazoazoles, confirming the accuracy of PM3 in these systems. Errors in MNDO and AM1 diazoazole heats of formation follow trends similar to the AM1 and MNDO errors in unsubstituted azole heats of formation.

Both semi-empirical and *ab initio* methods predict that the triplet spin state is the ground state for all cyclic azolylidenes. All methods predict high energy cyclic structures for singlet state azolylidenes, except for 4-triazoazolylidene, where PM3 and *ab initio* geometry optimizations spontaneously open the ring to produce diazoacetonitrile. The spontaneous ring scission of singlet state 4-triazoazolylidene and the stability of a cyclic triplet state 4-triazoazolylidene predicted by *ab initio* and PM3 account for the fact that reactions under photolytic conditions produce ring scission products only, while thermally induced reactions give products resulting from additions of cyclic 4-triazoylidene.

References:

1. J. Thiele, *Justus Liebigs Ann. Chem.* **270**, 46 (1892).

2. a) G. Cirrincione, A.M. Almerico, E. Aiello and G. Dattolo, *Adv. Heterocycl. Chem.*, **48**, 65 (1990).

   b) S. Sorriso, in "*The Chemistry of Diazonium and Diazo Groups*" (S. Patai, ed.), p95. John Wiley & Sons, New York, 1978.

3. M. Tisler and B. Stanovnik, *Heterocycles*, **4**, 1115 (1976).

4. R.N. Butler, *Chem. Rev.* **75**, 242 (1975).

5. M. Tisler and B. Stanovnik, *Kim. Geterosikl. Soedin.*, 579 (1980).

6. M.H. Elnagdi, E.M. Zayed, and S. Abdou, *Heterocycles*, **19**, 559 (1982).

7. A. Albert, *Adv. Heterocyclic Chem.*, **40**, 129 (1986).

8. H. Durr, A.C. Ranade, and I. Halberstadt, *Synthesis*, 878 (1974).

9. G. Ege, and K. Gilbert, *Tetrahedron Lett.*, **44**, 4253 (1979).

10. M. Cabre, J. Farras, J. Fernandez Sanz and J. Vilarrasa, *J. Chem. Soc., Perk. Trans. 2*, 1943 (1990).

11. J.K. Horton, M.F.G. Stevens, *J. C. S. Perkin 1*, 1433 (1981).

12. H.B. Ambroz, B.T. Golding, T.J. Kemp and V.S. Shukla, *J. Chem. Soc., Chem. Commun.*, 414 (1982).

13. E.A. Castro, *J.Mol. Struct. (THEOCHEM)*, **304**,93 (1994).

14. Zhang Yala, *J. Mol. Struct. (THEOCHEM)*, **207**, 217 (1990).

15. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen, S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** (1990) 52
    b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).

16. a) I. Leban, M. Tisler and B. Stanovnik, *Acta Crystallogr. Sect. B*, **B34**, 293 (1978).

b) T.M Krygowski, A. Ciesielski, and M. Cyranaski, *J. Mol. Struct. (THEOCHEM)*, **374**, 227 (1996).

c) L. Nygaard, D. Christen, J.T. Nielsen, E.J. Pedersen, O. Snerling, E. Vestergaard and G. O. Sorensen, *J. Mol. Struct.*, **22**, 401, (1974).

17. a) F.R. Benson, "*The High Nitrogen Compounds*", John Wiley & Sons, New York, 1984.

b) M. Sana, G. Leroy, M.T. Nguyen and J. Elguero, *Nouv. J. Chim.*, **3**, 607 (1979)

18. V. A. Ostravoskii, G.B. Erusalimskii and M.B. Shcherbina, *Zh. Org. Khim.*, **29**, 1297 (1993).

19. MNDO errors in 5- and 6-membered aromatic nitrogen containing heterocycles have been observed previously by other workers [a) E. Fos, J. Vilarrasa, and J. Fernandez, *J. Org. Chem.*, **50**, 4894 (1985); b) W.M.F. Fabian, *Z. Naturforsch*, **45a** 1328 (1990); and c) J.P. Ritchie, *J. Org. Chem.*, **54** 3553 (1989)]. The authors of (a) attributed MNDO errors to the underestimation of electronic repulsions between nitrogens in N-N linkages; however, MNDO heats of formation for pyrazole and 1,2,4-triazole are quite good (within 6 kcal/mol of the experimental values), suggesting that the 18-27 kcal/mol errors in the MNDO 1,2,3-triazole and the tetrazole heats of formation are due to poor treatment of the N=N-N linkage, rather than the N-N linkage.

20. The AM1 errors are consistent with a previous study (M. Meyer, *J. Mol. Struct.*, **304**, 45 (1994)), which found AM1 to consistently overestimate imidazole and methyl-substituted imidazole heats of formation by 20-30 kcal/mol.

21 J.A. Dean, ed., "*Lange's Handbook of Chemistry*," 14th Edition, (McGraw-Hill, New York) 1992.

22. S.G. Lias, J.F. Libman and R.D. Levine, *J. Phys. Chem. Ref. Data*, **13**, 695 (1984).

23. Cox, J.D. and G. Pilcher, "*Thermochemistry of Organic and Organometallic Compounds*"; Academic Press: New York, 1970.

24. Y.I. Alexsandrov, T.R. Osipova and V.F. Yushkevich, *Thermodin. Org. Soedin.,* **42**, 1 (1982) (estimated experimental value).

25. A. Padwa ,T. Kumagai and A.D. Woolhouse, *J. Org. Chem.*, **48**, 2330 (1983).

26. T.J. Amick, *Diss. Abstr. Int. B,* **44**, 1113 (1983).

27. T. J. Amick and H. Shechter, *Tetrahedron Lett.*, **27**, 901 (1986).

28. H.K. Hui, *Diss. Abstr. Int. B*, **44**, 2557 (1983).

Appendix 5.A

TABLE 5. A1: *Ab initio* Total Energies (a.u.)

Compound	E(a.u.)	
diazoazoles	6-31G*	6-31G**
1a	-300.46593	-300.47365
1b	-316.46538	-316.47557
1b'	-316.46994	-316.47108
1c	-332.47247	-332.47617
1c'	-332.47376	-332.47762
1d	-332.42233	-332.43846
1d'	-332.43460	-332.44187
1e	-348.41496	-348.41688
1e'	-348.44340	-348.44534
1f	-364.39342	-364.39342
diazomethane	-147.84378	-147.84782
N2	-108.94395	-108.94395
azolylidenes	6-31G**	
	singlet/RHF	triplet/UHF
2a	-191.46639	-191.55932
2b	-207.49146	-207.56290
2b'	-207.46684	-207.55645
2c	-223.50603	-223.56139
2c'	-223.48333	-223.56277
2e	**	-239.45153
2e'	-239.46989	-239.47361
2f	-255.39551	-255.40136

** = cyclic structure unstable; opens to diazoacetonitrile : 6-31G** E(a.u.) = -239.57986

The results from chapter five showed that the PM3 semi-empirical method gives reasonable results when applied to diazoazoles. The MNDO and AM1 systematic errors in azole heats of formation are paralleled in the diazoazole heats of formation. Since the MNDO and AM1 systematic errors appear to be constant, MNDO, AM1 and PM3 energy differences in diazoazole reactions with alkynes should be similar to one another, as the systematic errors will partially cancel. To investigate this possibility and to further justify the use of these semi-empirical methods for studying nitrogen heterocyclic reactivity, the MNDO, AM1 and PM3 methods were all used to study diazoazole cycloaddition reactions with ethyne, ynamine and cyanoethyne. This chapter will test if MNDO, AM1 and PM3 can predict the regio-specificity and relative reactivity in these systems.

CHAPTER 6:

A SEMI-EMPIRICAL AND *AB INITIO* STUDY OF DIAZOAZOLE CYCLOADDITION REACTIONS WITH ETHYNE, YNAMINE, AND CYANOETHYNE

by C.I. Williams, M.A. Whitehead
Theoretical Chemistry Laboratory, Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6

and

Bertrand J. Jean-Claude
Department of Oncology, McGill University, 3655 Drummond St., Ste. 701, Montreal, Quebec, Canada H3G 1Y6

6.1 Abstract:

Semi-empirical AM1, MNDO and PM3 methods were used to elucidate mechanisms of diazoazole cycloaddition reactions with ethyne, ynamine, and to explain the lack of reaction with cyanoethyne. Ethyne was found to react by either a [7+2] cycloaddition, which directly forms azolotriazine products, or by a [3+2] cycloaddition which forms spirobicyclic intermediates. The lowest energy reaction path for ynamine was neither the [7+2] or the [3+2] concerted cycloadditions, but rather a two-step mechanism involving an azo-coupled type intermediate. The proposed mechanism for ynamine/diazoazole cycloaddition predicts both the experimentally observed regioselectivity and the increased reactivity of ynamine relative to ethyne. The relative reactivities of ynamine, enamine and 1,2-diaminoethene with diazoazoles were also compared. No transition states were found for cyanoethyne/diazoazole cycloadditions, which explains why these reactions have not been observed experimentally. *Ab initio* (RHF/6-31G**) methods were used in certain instances to compare with the semi-empirical results.

6.2 Introduction:

Cycloaddition reactions between diazoazoles (**1, Fig. 6.1**) and points of unsaturation in alkenes, alkynes and ynamines (recently reviewed by Cirrincione et. al.[1]) are useful in the synthesis of azolotriazines (**2, Fig. 6.1**). Experimentally, ynamines and alkynes both react with diazoazoles, but ynamines are more reactive.[1] Ynamine reactions with diazoazoles produce exclusively one aminoazolotriazine isomer (**Fig. 6.1**). No cyanoethyne cycloaddition reactions with diazoazoles have been reported. The reactions are net 1,7 cycloadditions, but the actual mechanisms are still being debated; experimental evidence suggests that diazoazole -RC≡CR reactions may occur either by a [3+2] cycloaddition, which forms spirobicyclic intermediates (**3, Fig. 6.1**) that quickly rearrange to azolotriazines, or by a direct [7+2] cycloaddition.[2-6] Mechanisms involving aziridine intermediates have also been proposed,[6] but MNDO, AM1 and 3-21G *ab initio* studies predict these intermediates to be of high energy;[7] therefore they will not be discussed here.

The reliability of semi-empirical methods (AM1, MNDO, PM3) in nitrogen heterocyclic chemistry has been extensively investigated, and they have been shown to yield reasonable intermediates, transition states and products in nitrogen-heterocyclic reactions.[8-15] In chapters three through five, the MNDO, AM1 and PM3 semi-empirical methods were used to determine nitrogen heterocyclic heats of formation, and to study the reactivities of azolylidenes and isocyanates. The results in these chapters suggest that the PM3 method is the most accurate for nitrogen heterocyclic compounds. Here, the AM1, MNDO and PM3 methods are applied to the diazoazole cycloaddition reactions. AM1, MNDO and RHF/3-21G calculations have already been performed on possible reactions intermediates in diazoazole/-H2C=CRR (R= -OH, -H) cycloaddition reactions,[7] but no transition states were reported, and some aspects of the regioselectivity were not discussed. Also, the PM3 method was not used by those workers. Consequently, a detailed

semi-empirical investigation of possible diazoazole/H-C≡C-R (R= -H, -NH2, -CN) reaction mechanisms was performed. In some cases, semi-empirical energy differences are compared with RHF/6-31G** predictions. The relative reactivity of ynamine, enamine and 1,1-diaminoethene cycloaddition with diazoazoles is also discussed.

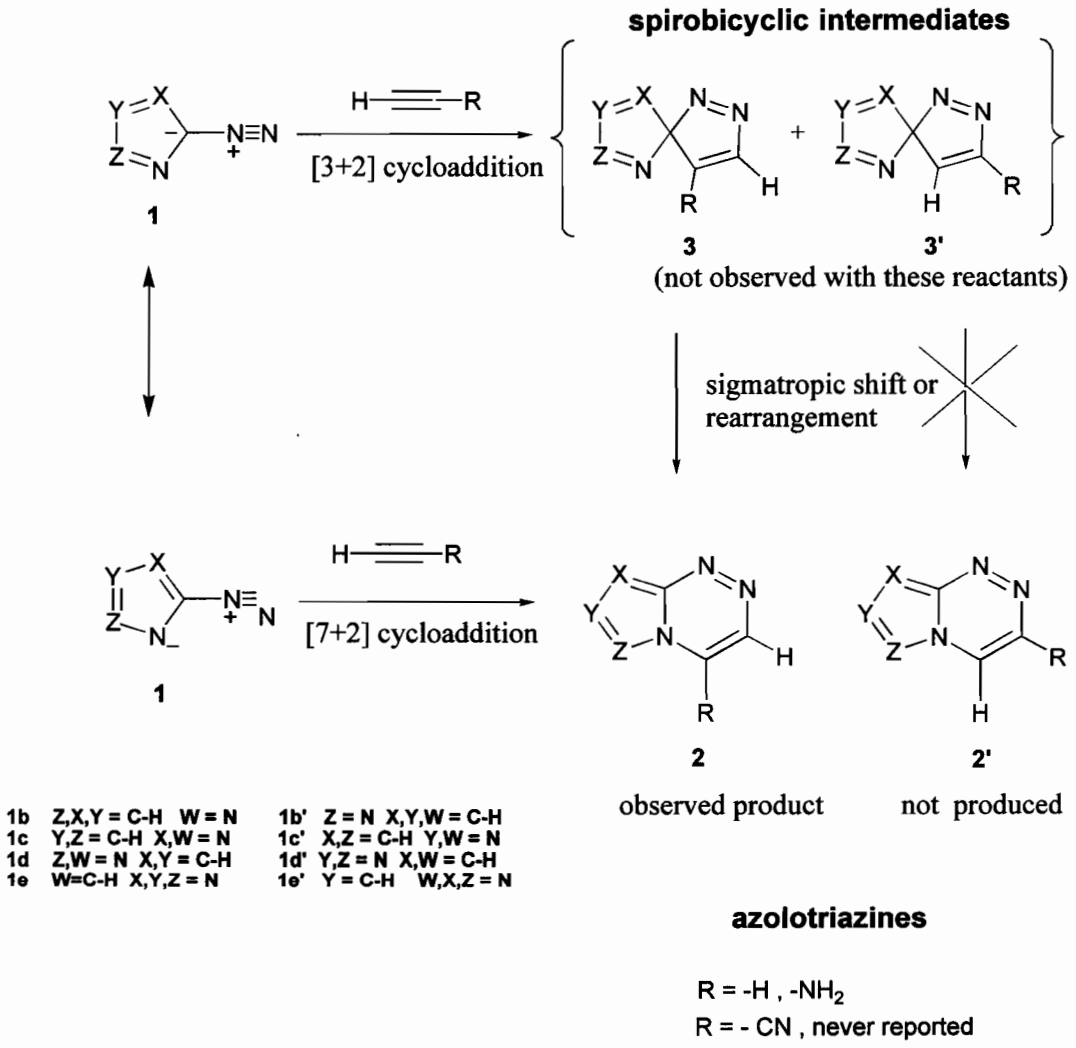


Figure 6.1:

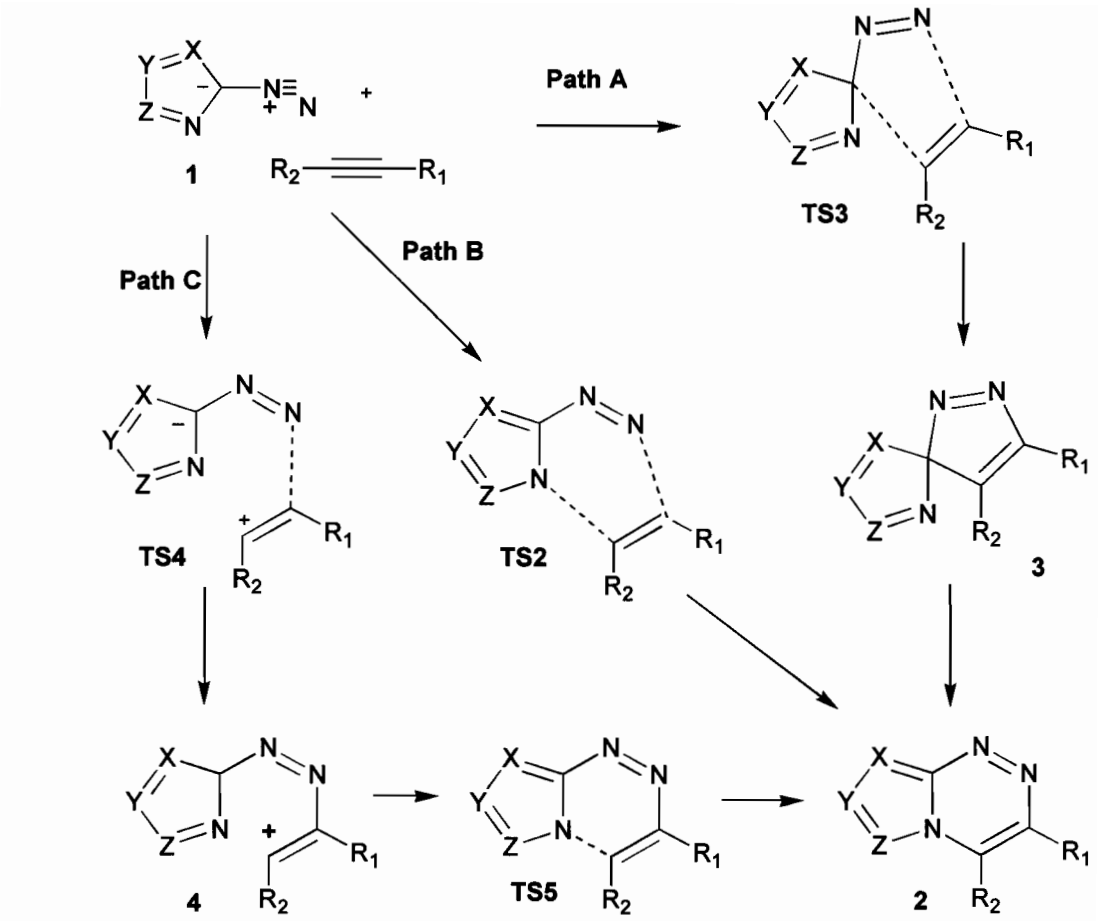
Diazoazole Cycloaddition Reactions with C≡C Triple Bonds

6.2.1. Cycloadditions Pathways: The three pathways of diazoazole/-C≡C- cycloaddition reactions studied here are summarized in **Fig. 6.2** (Paths **A-C**). Each path was studied in detail with ethyne, ynamine and cyanoethyne as sources of the C≡C bond to investigate the effects of electron withdrawing and electron donating substituents. Both possible orientations of the $-NH_2$ and -CN substituents are considered. All intermediates, transition states and products with ynamine and cyanoethyne having the substituent orientation R_2 = $-NH_2$ or -CN are unprimed (ie TS2h). Structures with R1 = $-NH_2$ or -CN are primed (ie **TS2h'**). In the case of ynamine, the unprimed structures are the experimentally observed products.

Path A starts with a concerted [3+2] cycloaddition, which forms a spirobicyclic intermediate (3) through spiro transition state (**TS3**). It is believed that spiro intermediates spontaneously undergo rearrangements or sigmatropic shifts to form azolotriazines. Spirobicyclics are not very stable, but they have been isolated and characterized from a cycloaddition reaction between 3-diazo-4-methyl-5-phenylpyrazole and 1,1-dimethoxy-ethene.[3,4] These reactions are analogous to [3+2] cycloaddition reactions between alkynes and diazocyclopentadiene, which also yield stable, isolable spirobicyclic products.[2]

Path B is a concerted [7+2] in-plane cycloaddition, which directly forms azolotriazine products (**2**) through planar transition states (**TS2**). **Path B** is impossible for diazoazoles with no nitrogen atoms α to the diazo group.

Path C: The first step of **Path C** is the formation of an open-chain azo-coupled intermediate (**4**) through a **TS4** type transition state. The azolotriazine is then formed by ring closure through a **TS5** type transition state. Intermediates of type **4** have never been isolated, but they may be stabilized by electron donating groups at the R_2 position. Type **4** intermediates may exist as trans (**E**) or cis (**Z**) isomers, all of which will be investigated.



ethyne (a - g)

a	X,Y,Z = C-H	R1, R2 = H
b	X = N Y,Z = C-H	R1, R2 = H
c	Y = N X,Z = C-H	R1, R2 = H
d	X,Y = C-H Z = N	R1, R2 = H
e	Y,Z = N X = C-H	R1, R2 = H
f	X,Y = N Z = C-H	R1, R2 = H
g	X,Z = N Y=C-H	R1, R2 = H

ynamines (h - n)

R2 = NH2 , R1= -H		R1 = NH2 R2 = H
h	X,Y,Z = C-H	h'
i	X = N Y,Z = C-H	i'
j	Y = N X,Z = C-H	j'
k	Z = N X,Y = C-H	k'
l	X,Z = N Y = C-H	l'
m	X,Y = N Z = C-H	m'
n	Y,Z = N X = C-H	n'

cyanoethyne (o - t)

	R1 = H R2 = CN	R1 = CN R2 = H
o	X,Y,Z = C-H	o' X,Y,Z = C-H
p	X = N Y,Z = C-H	p' X = N Y,Z = C-H
q	Y = N X,Z = C-H	q' Y = N X,Z = C-H
r	Z = N X,Y = C-H	r' Z = N X,Y = C-H
s	X,Z = N Y = C-H	s' X,Z = N Y = C-H
t	Y,Z = N X = C-H	t' Y,Z = N X = C-H

Figure 6.2:

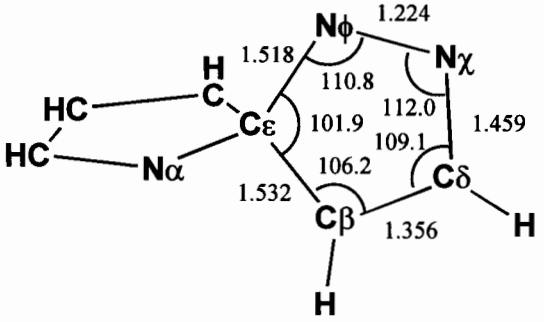
Possible Mechanisms, Intermediates and Transition States in Diazoazole Cycloaddition Reactions with Ethyne, Ynamine and Cyanoethyne:

6.3 Calculations:

Calculations were performed on a SUN Sparc10 workstation using the GAMESS[16] package. Full geometry optimizations were allowed, and RHF wavefunctions used. Equilibrium geometries of products, reactants and intermediates were verified as local minima by force constant analysis. *Ab initio* type calculations used the 6-31G** basis set, and RHF wavefunctions. Although diazoazoles have been reported as planar, C_s symmetry was not enforced, because it would bias any calculation that would otherwise report a non-planar geometry. Transition states were located by initiating saddle point geometry searches on trial structures whose Hessians had the desired curvature. Transition states were verified by the existence of only one imaginary vibrational eigenvalue.[17] The correct precursors and products for each transition state were checked by either running intrinsic reaction coordinate calculations initiated from the transition state in question, or by performing local minima geometry searches initiated with transition state structures that had been slightly distorted along the axes of the forming and/or breaking bonds. Structures and energies varied smoothly throughout the reaction coordinates in all cases. Semi-empirical heats of formation and 6-31G** total energies of diazoazoles and other reactants are listed in appendix **Table 6.A1**.

6.4 Results and Discussion:

6.4.1. Ethyne Path A: Spirobicyclic intermediates (**3a-g**) and transition states (**TS3a-g**) were easily located with all three semi-empirical methods. Heats of reaction and activation energies are in **Table 6.1**. Representative intermediate (**3a**) and transition state (**TS3**) structures from each semi-empirical method are in **Fig. 6.3** and **Fig. 6.4**. All methods predict that the spiro-bicyclic formations are exothermic. The 6-31G** and PM3 diazoazole → spirobicyclic intermediate reaction energies are in good agreement throughout the series. AM1 reaction energies differ from the *ab initio* values by 4-5 kcal/mol, while MNDO differs from the *ab initio* values by 20 kcal/mol.



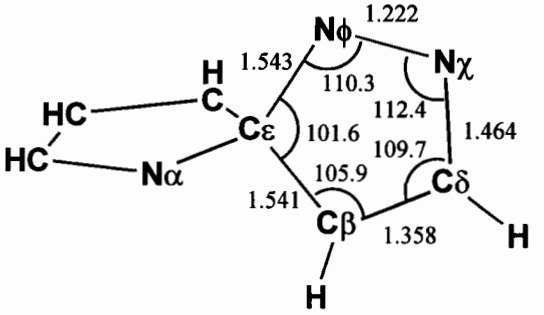
MNDO

Torsion Angles:

Nα Cε Cβ Cδ : 117.8

Nα Cε Nϕ Nχ : -121.5

Nϕ Nχ Cδ Cβ : 0.0



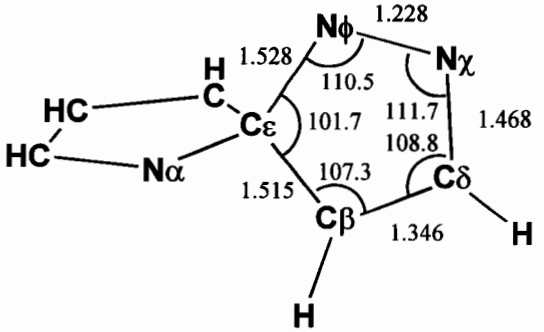
AM1

Torsion Angles:

Nα Cε Cβ Cδ : 122.5

Nα Cε Nϕ Nχ : -122.2

Nϕ Nχ Cδ Cβ : -0.5



PM3

Torsion Angles:

Nα Cε Cβ Cδ : 118.4

Nα Cε Nϕ Nχ : -120.3

Nϕ Nχ Cδ Cβ : -0.2

Figure 6.3: Representative Diazoazole/ethyne Spirobicyclic Intermediates (**3b**)

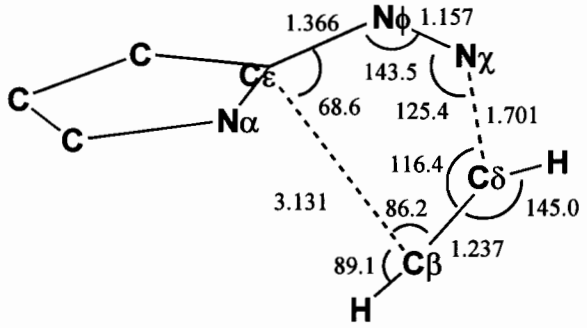
(all bond lengths in Å; all bond angle in degrees)

Table 6.1:
PATH A Diazoazoles + Ethyne Reaction Energies and Activation Barriers:

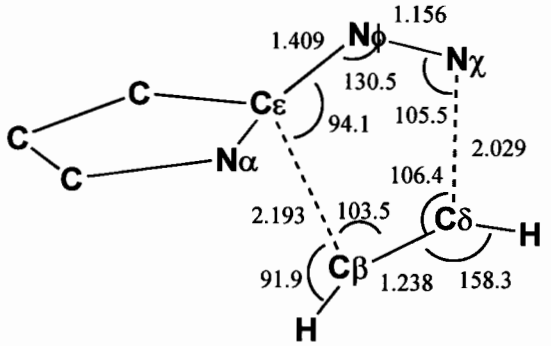
	ΔH_{rxn} (kcal/mol)			
	MNDO	AM1	PM3	6-31G**
Diazoazoles (1) + ethynes → spirobicyclics (3)				
1a → 3a	-59.1	-30.6	-37.7	-41.3
1b → 3b	-57.9	-29.8	-35.1	-38.9
1c → 3c	-57.1	-29.9	-33.9	-17.4
1c' → 3d	-56.5	-29.6	-33.4	-33.5
1d → 3e	-56.3	-29.0	-33.7	-32.6
1e → 3f	-54.6	-28.0	-31.5	-34.6
1e' → 3g	-56.2	-29.5	-32.4	-32.5
[3+2] Activation Energies E_a (kcal/mol)				
TS3a	34.9	35.4	37.5	
TS3b	33.7	33.5	38.9	
TS3c	31.5	36.0	38.7	
TS3d	35.0	37.5	38.9	
TS3e	34.4	38.4	40.2	
TS3f	32.5	34.9	39.2	
TS3g	33.6	34.2	39.9	

The PM3 and AM1 (**TS3a-g**) [3+2] cycloaddition reaction transition state geometries are quite similar throughout the series of diazoazoles; they are slightly asynchronous, with Nχ-Cδ bond distances shorter than Cε-Cβ bond distances. MNDO predicts highly asynchronous transition states, quite different from the AM1 or PM3 structures. The number of nitrogen atoms in the azole ring has little effect on the transition state geometries, or on the activation barriers. MNDO

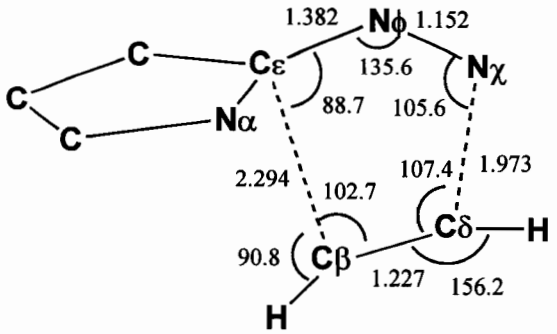
predicts the lowest average [3+2] activation barrier (33.7 kcal/mol), while PM3 predicts the highest (38.9 kcal/mol).



MNDO
Torsion Angles:
NαCεCβCδ : 116.3
NαCεNϕ Nχ : -89.9
Nϕ NχCδ Cβ : -1.4



AM1
Torsion Angles:
NαCεCβCδ : 123.2
NαCεNϕ Nχ : -109.4
Nϕ NχCδ Cβ : -0.86



PM3
Torsion Angles:
NαCεCβCδ : 120.1
NαCεNϕ Nχ : -108.6
Nϕ NχCδ Cβ : -0.17

Figure 6.4:

Representative Ethyne/Diazoazole [3+2] Transition States (**TS3b**)

(all bond lengths in Å; all bond angle in degrees)

No transition state for the spirobicyclic → azolotriazine sigmatropic shift could be found. Since the rearrangement might occur through a stepwise **Path C** type mechanism, this possibility was also explored. Presumably the stepwise spirobicyclic rearrangement would begin with an opening of the ring at the Cε-Cβ bond to form an open-chain structure similar to the (**4**) intermediate in **Path C**. However, no transition states for the spirobicyclic → open-chian intermediate conversion were found, and repeated attempts to open the spirobicyclic ring led to reactants, **TS3** transition states, or open-chain intermediates. Therefore, although the [3+2] cycloaddition appears straight-forward, the mechanism for the spirobicyclic → azolotriazine rearrangement is still not computationally treated.

Ethyne Path B: PM3, MNDO, and AM1 energies for diazoazole → azolotriazine reactions (**Table 6.2**) are not in as good agreement with the 6-31G** values as were the diazoazole → spirobicyclic intermediates reaction energies; however correlation between semi-empirical and *ab initio* energies was close enough for the semi-emprical values to be trusted. The azolotriazines are much more thermodynamically stable than either the reactants or the corresponding spirobicyclic intermediates.

Transition states for the [7+2] cycloaddition (**TS2a-g**) were easily located. Representative MNDO, AM1 and PM3 transition state geometries are in **Fig. 6.5**. Activation energies are in **Table 6.2**. As with the [3+2] transition states, the number of nitrogens in the azole ring has little effect on activation energy barriers.

MNDO predicts the smallest average [7+2] activation barrier (28.4 kcal/mol), while PM3 predicts the largest (40.1 kcal/mol). PM3 **TS2** transition states are slightly more asynchronous than PM3 **TS3** spiro-transition states. The AM1 and MNDO **TS2** transition states differ substantially from the PM3 geometries; they are highly asynchronous, with short Nχ-Cδ bond distances and very long Nα-Cβ bond distances.

Table 6.2:
PATH B Diazoazoles + ethyne Reaction Energies and Activation Barriers:

	ΔH_{rxn} (kcal/mol)			
	MNDO	AM1	PM3	6-31G**
Diazoazoles (1) + ethynes → products (2)				
1b → 2a	-80.9	-67.6	-63.1	-61.4
1c → 2b	-81.8	-70.0	-61.0	-36.9
1c' → 2c'	-80.9	-66.3	-60.7	-57.7
1d → 2d	-78.0	-60.8	-59.7	-72.4
1e → 2e	-78.8	-62.2	-57.9	--
1e' → 2f	-76.4	-59.9	-55.2	-73.9
→ 2g	-77.8	-63.1	-54.9	-59.8
[7+2] Activation Energies E_a (kcal/mol)				
TS2a	28.9	33.1	40.1	
TS2b	30.2	33.6	41.5	
TS2c	28.0	32.1	39.1	
TS2d	28.4	31.9	39.9	
TS2e	26.3	30.2	38.1	
TS2f	28.3	32.4	40.6	
TS2g	28.4	32.1	41.1	

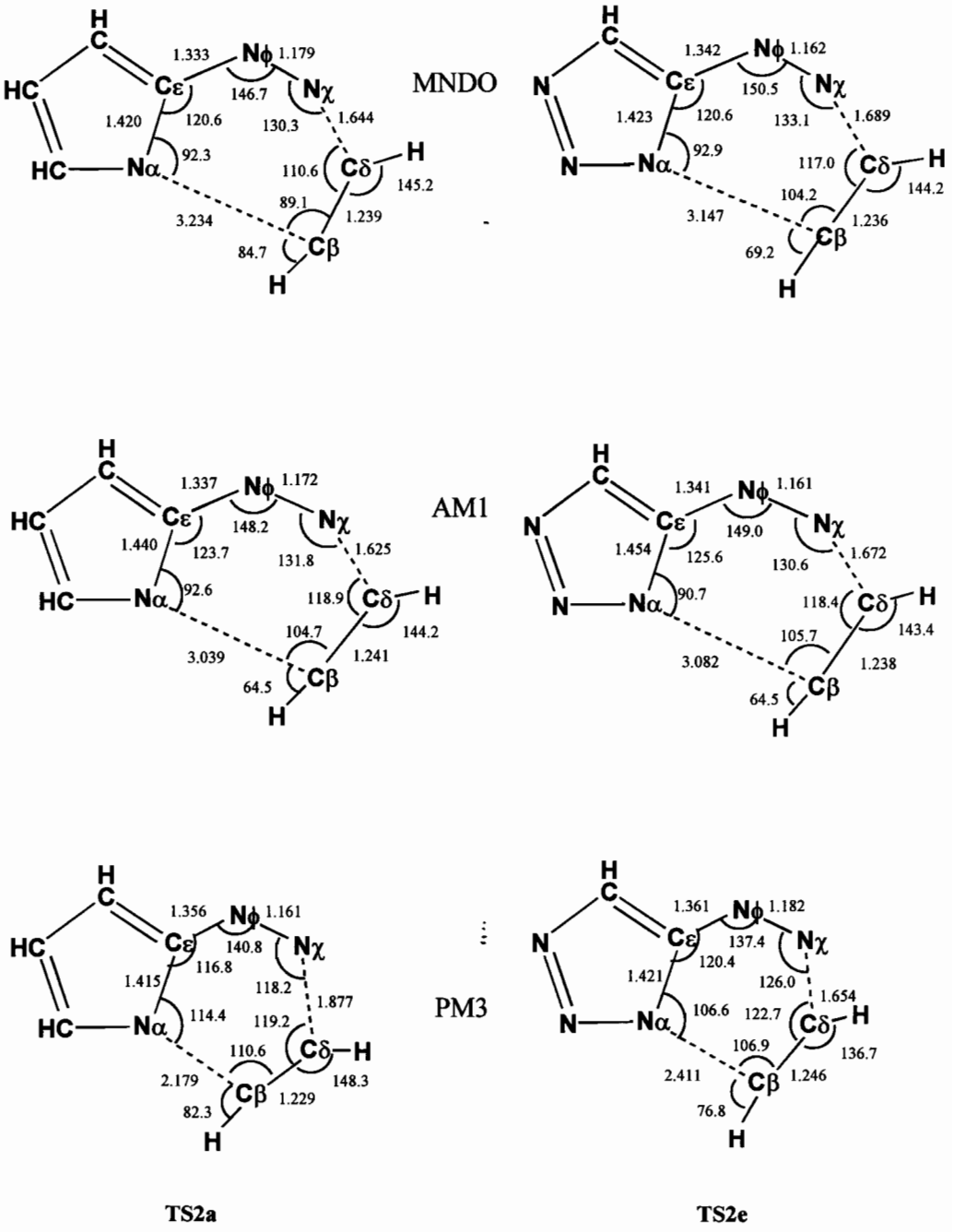


Figure 6.5:

Representative Diazoazole/Ethyne [7+2] Cycloaddition

Transition States (all bond lengths in Å; all bond angle in degrees)

Ethyne Path C: Intemediates with azo-coupled side chains coplanar with the azole ring (**Fig. 6.6**) were unstable, and these types of structures all revert to either

reactants, **TS2** type transition states, or products upon geometry optimization. However, azo-coupled intermediates (**4a-4g**) could be located when the azo side-chain was in an out-of-plane conformation (henceforth labelled $\mathbf{4a}_{oop}$-$\mathbf{4g}_{oop}$, **Fig. 6.7**). Reaction energies for diazoazole + ethyne $\rightarrow \mathbf{4}_{oop}$ were typically 28 kcal/mol (MNDO) to 52 (PM3) kcal/mol. The out-of-plane structures were investigated further, since they might be intermediates involved in a stepwise spirobicyclic $\rightarrow$ azolotriazine rearrangement. A transition state was found for closure of the ($\mathbf{4}_{oop}$) structure to azolotriazines, but no transition state could be found for closure of the ($\mathbf{4}_{oop}$) type structures to spiro-bicyclics.

6.4.2. Mechanisms of Diazoazole/ethyne cycloaddition: Semi-empirical reaction coordinates for diazoazole-ethyne cycloadditions are in (**Fig. 6.8**). The relatively high activation energy barriers of both [7+2] and [3+2] additions reflect the slow rate at which these reactions proceed.[1] The detailed energy profile of the spirobicyclic rearrangement could not be determined. However, since out-of-plane azo-coupled structures ($\mathbf{4}_{oop}$) structures easily close to azolotriazines, the ($\mathbf{4}_{oop}$) $\rightarrow$ azolotriazine step is included as an approximation of a possible stepwise spirobicyclic rearrangement. The activation barrier for ($\mathbf{4}_{oop}$) $\rightarrow$ azolotriazine is only 1 kcal/mol, so it is not included explicitly in the reaction coordinate. A transition state for the spirobicyclic open-chain ($\mathbf{4}_{oop}$) conversion could not be found, but for the sake of argument, the conversion was assumed to be barrierless; this step is therefore denoted by a dotted line on the reaction coordinate. Even with a zero activation barrier for the spiro $\rightarrow \mathbf{4}_{oop}$ conversion, the $\mathbf{4}_{oop}$ type intermediate is of such high energy that the overall activation energy barrier for the rearrangement is quite large (>80 kcal/mol).

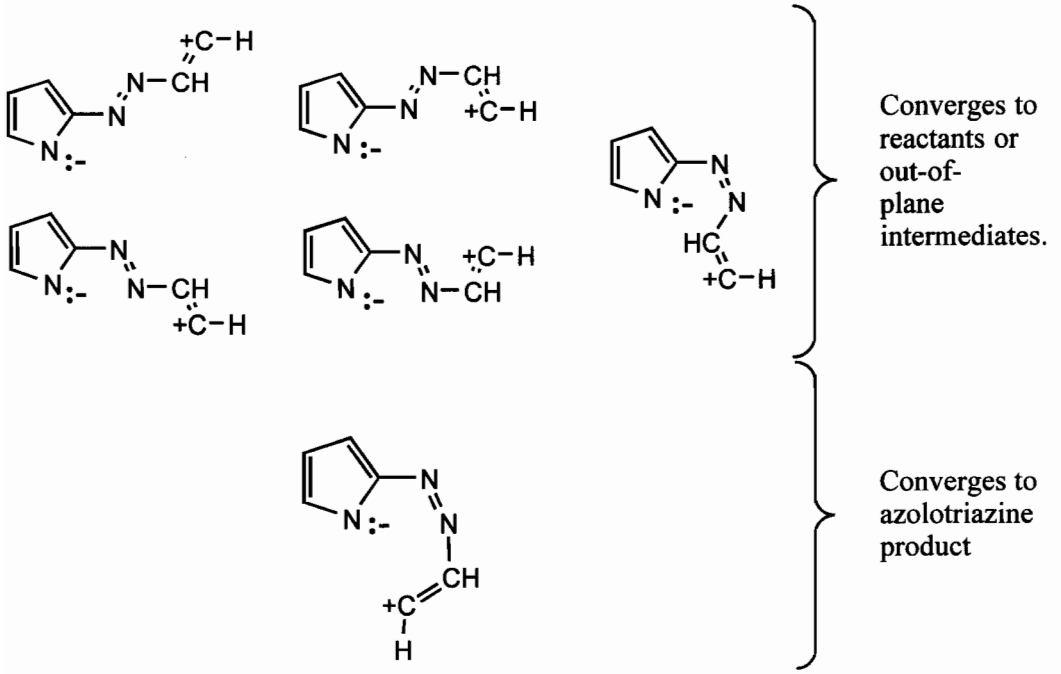


Figure 6.6:

Conformation of Planar Azo-Coupled Intermediates from Diazoazole/ethyne Reactions ($\mathbf{4}_{ip}$)

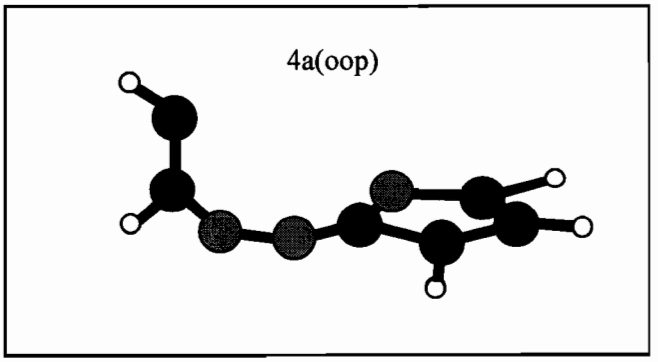
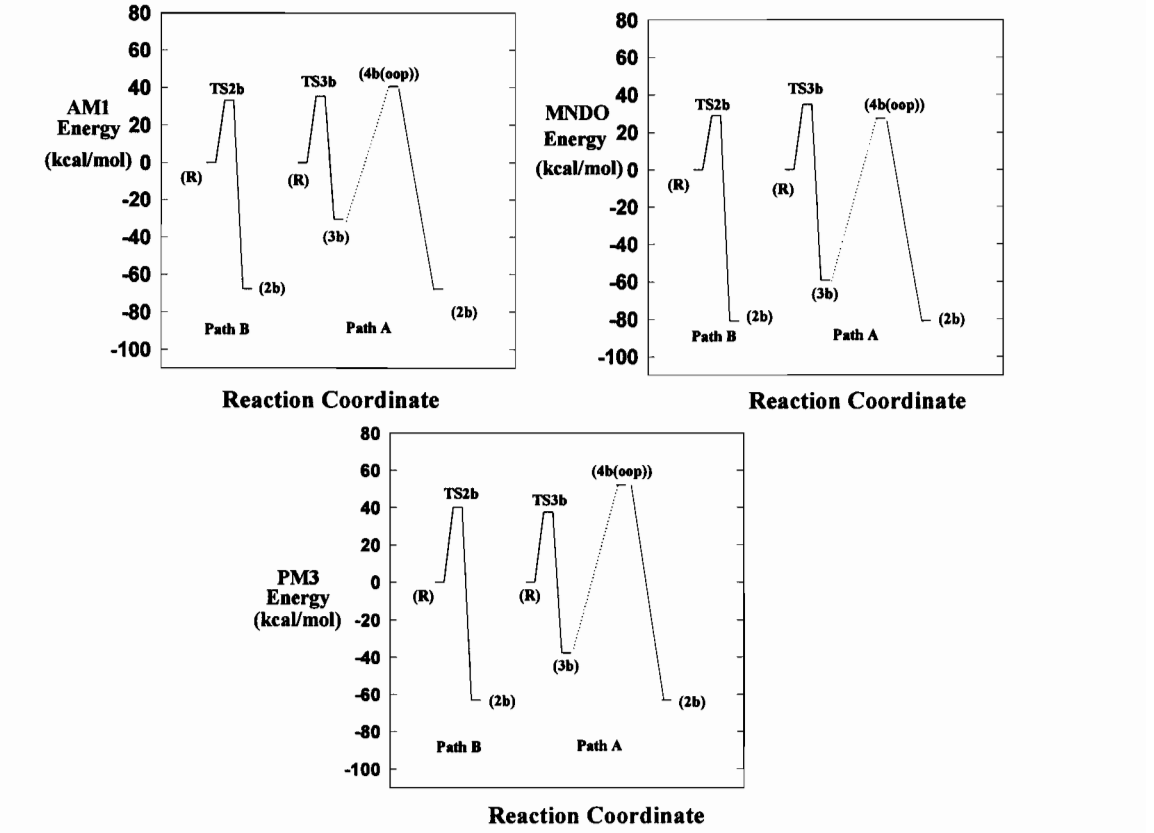


Figure 6.7:

Out-of-Plane Conformation of Diazoazole/ethyne Azo-Coupled Intermediate



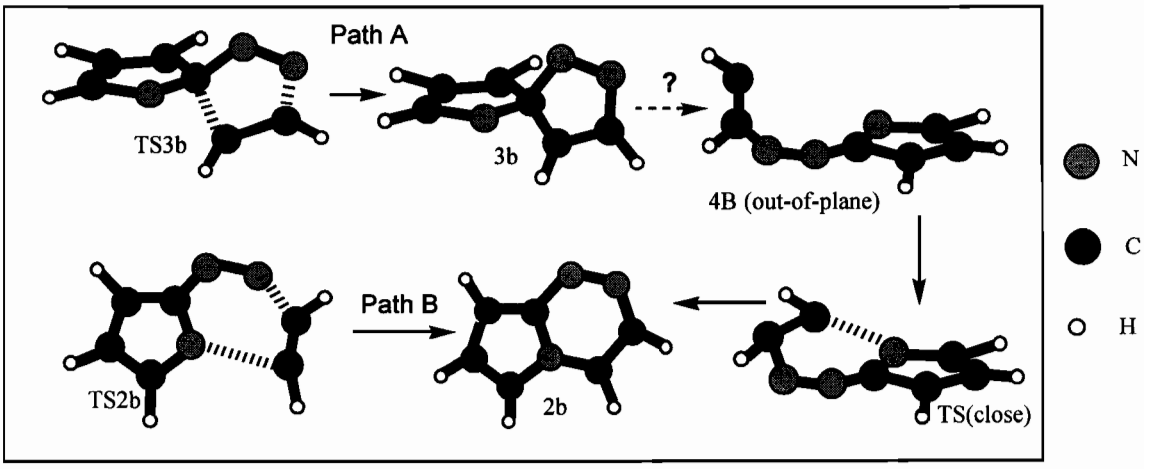


Figure 6.8:

Representative Reaction Coordinates in Diazoazole/Ethyne Cycloaddition Reactions

Within each method, activation energies for the [7+2] and [3+2] cycloadditions are similar; PM3 favours the [3+2] cycloaddition by an average of 1.5 kcal/mol, while AM1 and MNDO both favour the [7+2] cycloaddition by averages of 3.7 and 5.3 kcal/mol respectively. The differences between the [7+2] and [3+2] cycloaddition reaction barriers are small within each method, therefore both cycloadditions could occur. Although [3+2] cycloadditions can occur, the spirobicyclic → azolotriazine rearrangement is energetically unfavourable, and probably does not occur; therefore, formation of azolotriazines from ethyne and diazoazole probably occurs by the **Path B**, direct [7+2] cycloaddition. The **Path C** mechanism is unfeasible for ethyne, since all in-plane azo-coupled intemediates are unstable.

6.4.3 Investigation of the [7+2] Transition State Col:

The asynchronicity of PM3 **TS2** transition states was investigated by constructing a potential energy surface (PES) around a sample PM3 transition state col (TS2c). The data for the plots were generated by constraining the Nχ-Cδ and Nα-Cβ bond lengths but optimizing all other geometric parameters around the col. When viewed along the Nδ-Cχ and Cε-Cβ axes independently (**Fig. 6.9**), the plots clearly show shortening of the Nδ-Cχ bond before shortening of the Cε-Cβ at the point where the col is crossed.

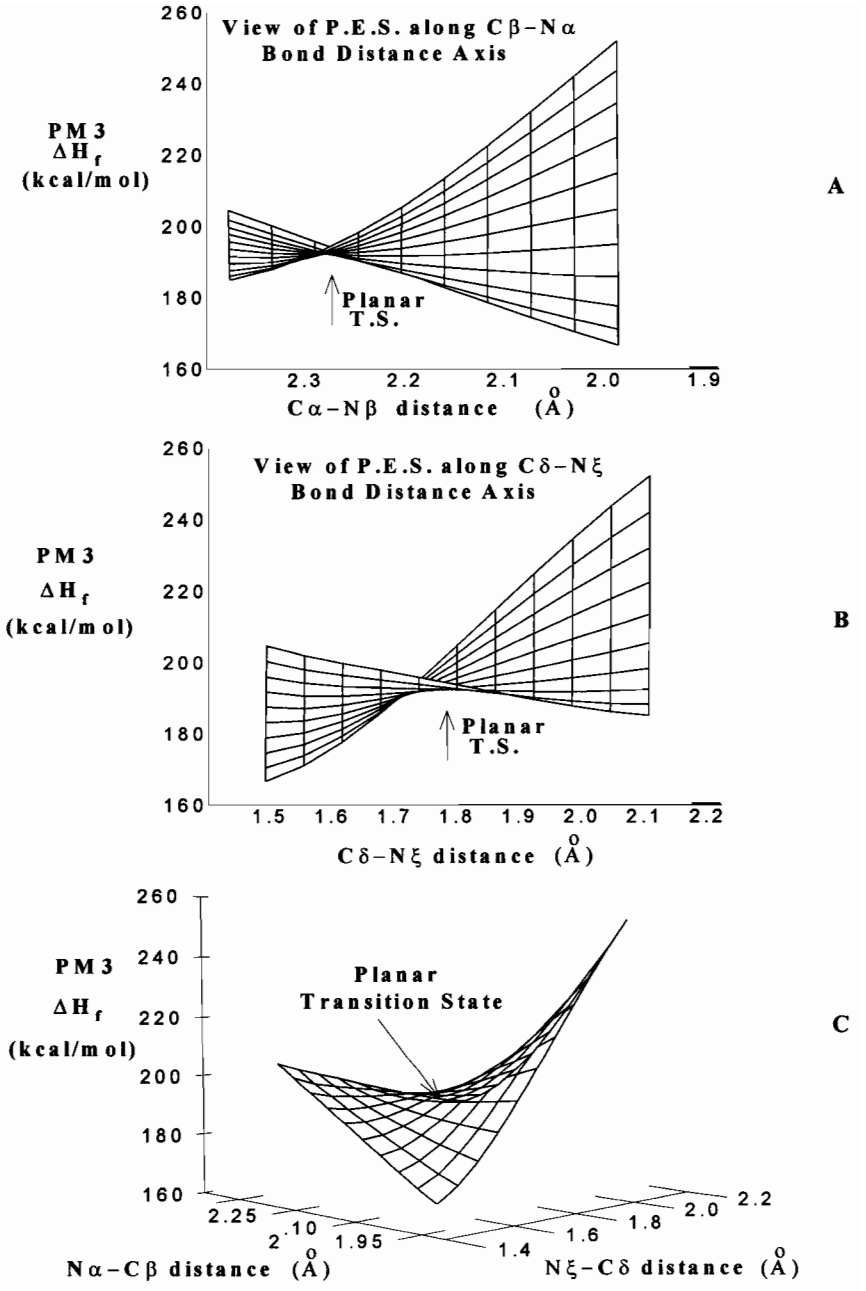


Figure 6.9: Potential Energy Surface near the **TS2c** [7+2] Transition State Col: Views **A** and **B** are side views of plot **C**, illustrating the asynchronicity of the addition. **Cσ-Nχ** bond distances are short (view **B**), and **Cα - Nβ** bond distances are long (view **A**) at the transition

6.4.4. Ynamines:

Ynamine Path A: Relative to the diazoazole/ethyne reaction energies, all the methods predict increased thermodynamic stability for spirobicyclic intermediates (**3h-3n'**) resulting from diazoazoles reactions with electron-rich ynamine (**Table 6.3**): structures were similar to the alkyne/diazoazole spirobicyclics. No transition states (**TS3**) for the formation of spiro- bicyclic intermediates were found for either R_1 = $-NH_2$ or R_2 = $-NH_2$. Attempts at locating these transition states resulted in either reactants or out-of-plane conformations of azo-coupled intermediates (discussed below).

Ynamine Path B: PM3 and MNDO predict comparable thermodynamic stability of both the R_1 = $-NH_2$ and R_2 = $-NH_2$ isomers of azolotriazine (**2h-n'**). AM1 slightly favours the experimentally observed R_2 = $-NH_2$ isomer, but not by enough energy to convincingly explain the experimentally observed selectivity (**Table 6.4**). However, because of their structural similarity, it would be expected that the azolotriazine isomers are of comparable thermodynamic stability. Products **2h-2n'** are about 4 kcal/mole more stable if $-NH_2$ has a pyramidal rather than planar geometry, suggesting there is little conjugation between the NH_2 lone pair and the rest of the azolotriazine aromatic system. There is a 4 kcal/mol maximum barrier to free rotation of the amino group.

Table 6.3:
Reaction Energies for Diazoazoles + Ynamines → Spiro-bicyclic Intermediates:

	ΔH_{rxn} (kcal/mol)		
	MNDO	AM1	PM3
Diazoazole (1) + ynamine → spiro-intermediates (3)			
1b → 3h	-65.7	-36.8	-46.5
→ 3h'	-65.7	-32.0	-46.2
1c → 3i	-64.6	-36.7	-44.7
→ 3i'	-65.7	-32.3	-45.5
1c' → 3j	-65.1	-35.7	-44.9
→ 3j'	-64.9	-32.1	-44.8
1d → 3k	-64.5	-35.9	-45.1
→ 3k'	-64.8	-31.6	-45.2
1e → 3l	-63.0	-35.3	-43.0
→ 3l'	-63.3	-31.2	-43.5
1e' → 3m	-63.6	-36.5	-43.2
→ 3m'	-65.1	-32.5	-44.4

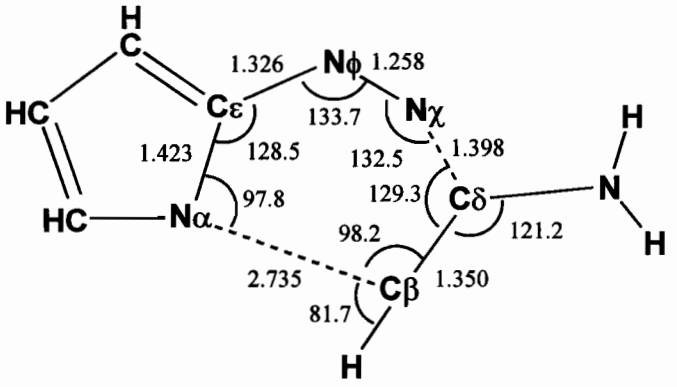
In-plane [7+2] TS2 type transition states with ynamines (**Fig. 6.10**) are only stable when R_1 = -NH_2. Activation energies (**Table 6.4**) are comparable to those of diazoazole/ethyne [7+2] cycloadditions. PM3, MNDO and AM1 geometries of ynamine **TS2** transition states mirror the diazoazole/ethyne **TS2** transition state structures, although asynchronicity is more pronounced in the ynamine **TS2** transition states. **TS2** transition states do not exist when R_2 = -NH_2, and attempts to locate **TS2** type transition states where R_2 = -NH_2 resulted in either products, reactants, or in-plane type **4** azo-coupled intermediates.

Ynamine Path C: The most striking result for **Path C** with ynamines is the stability of planar azo-coupled intermediates. Semi-empirical energies of planar azo-coupled intermediates and transition states involved in **Path C** (**4h-4n'**, **TS4h-n'**, **TS5h-n'**) depend heavily upon both the azo side-chain conformation, and the position of amino group. Representative structures and activation energies are in

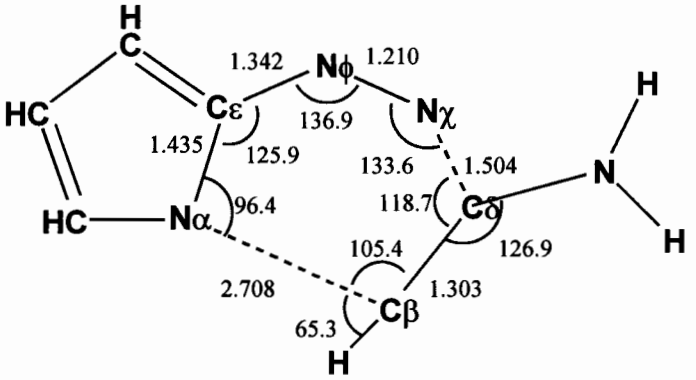
Figure 6.11 and **Table 6.5** respectively. Unlike planar azo-coupled intermediates in ethyne/diazoazole reactions, planar azo-coupled intermediates ($\mathbf{4}_{ip}$) resulting from diazoazole/ynamine reactions are stable and of low energy if R_2= -NH_2. These intermediates have linear Cδ-Cβ-Nγ bond angles, and hydrogens on the ynamine nitrogen are in a plane perpendicular to the plane of the rest of the intermediate (**Fig. 6.11**). However, minimization of in-plane azo-coupled intermediates ($\mathbf{4}_{ip}$) with R_1 = -NH_2 return either separated reactants, products, or **TS2** type transition states; therefore, the $\mathbf{4}_{ip}$ type intermediates with R_1 = -NH_2 are unstable, just like their diazoazole/ethyne counterparts.

Table 6.4: PATH B Diazoazoles + ynamines → aminoazolotriazines; Reaction Energies and Activation Energies:

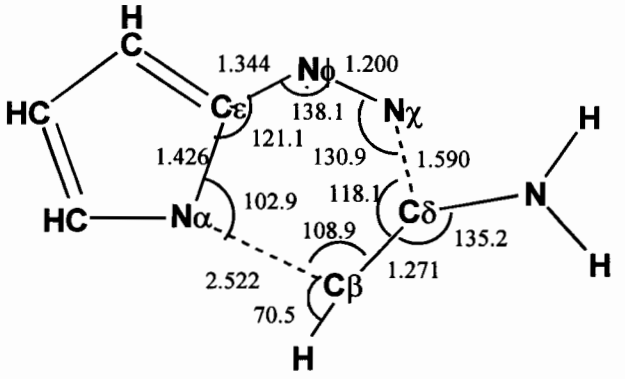
	ΔH_{rxn} (kcal/mol)		
	MNDO	AM1	PM3
Diazoazole (1) + ynamine → products (2)			
1b → 2h	-82.3	-66.3	-68.2
→ 2h'	-83.8	-64.7	-72.0
1c → 2i	-82.0	-66.1	-64.6
→ 2i'	-82.4	-63.8	-61.0
1c' → 2j	-82.7	-65.4	-66.2
→ 2j'	-84.4	-64.0	-70.3
1d → 2k	-82.7	-62.6	-67.1
→ 2k'	-81.0	-57.5	-68.8
1e → 2l	-84.0	-64.7	-66.0
→ 2l'	-82.3	-59.7	-67.6
1e' → 2m	-82.3	-62.7	-63.6
→ 2m'	-79.3	-56.7	-65.0
→ 2n	-80.9	-63.1	-61.2
→ 2n'	-81.5	-61.2	-65.4
[7+2] Activation Energies E_a (kcal/mol) (R1 = -NH2 only; R2 = -NH2 are unstable)			
TS2h'	24.3	39.2	40.3
TS2i'	22.7	38.2	39.7
TS2j'	25.6	35.6	38.1



MNDO



AM1



PM3

Figure 6.10:

Representative Diazoazole/ynamine [7+2] Cycloaddition Transition States **(TS2h)**

(all bond lengths in Å; all bond angle in degrees)

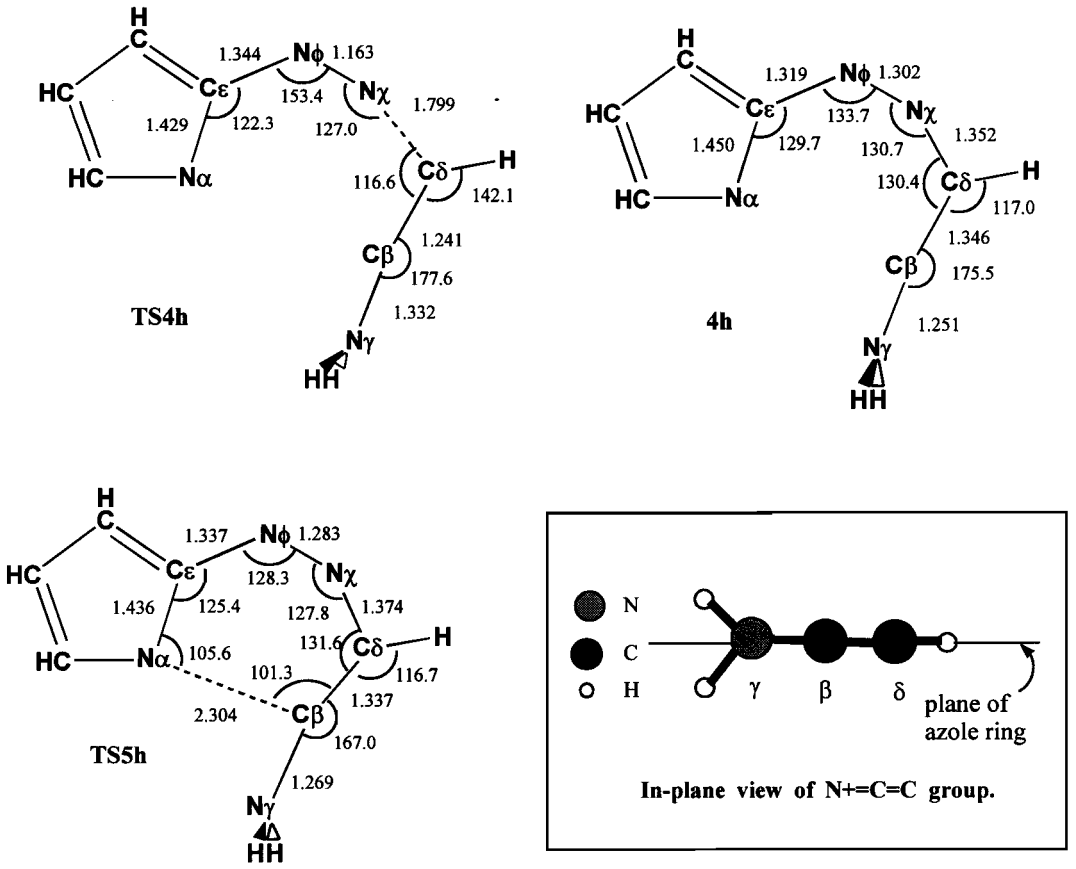


Figure 6.11:

Representative Diazoazole/ynamine PM3 Path C Structures **(TS4h, 4h TS5h)**

(all bond lengths in Å; all bond angle in degrees)

The relative stabilities of the R_1 = -NH_2 and R_2 = -NH_2 isomers of the $\mathbf{4}_{ip}$ intermediates can be explained by formal bonding. Electron donation from -NH_2 in the -R_2 position stabilizes the $\mathbf{4}_{ip}$ intermediate through resonance in the H_2N^+=C=C group. Instability of $\mathbf{4}_{ip}$ type intermediates with R_1 = -NH_2 is also expected from formal bonding, because stabilization of the Cβ atom by resonance with the amino nitrogen is not possible in this isomer. Type $\mathbf{4}_{ip}$ intermediates with R_1 = -NH_2 do not form, and **Path C** is inaccessible to ynamines with this substituent orientation. It should be noted that transient intermediates with structures remarkably similar to that of the ($\mathbf{4}_{ip}$) intermediate have been proposed by Leonard et. al. in 5-methyladenine translocative rearrangements.[18a]

Table 6.5: PATH C Diazoazoles + Ynamines; Reaction Energies and Activation Barriers:

	ΔH_{rxn} (kcal/mol)		
	MNDO	AM1	PM3
(R_2 = -NH_2) in-plane (4_{ip}) comformation			
1b → 4h	-17.8	-15.9	-5.0
1c → 4i	-15.9	-14.1	-3.3
1c' → 4j	-19.7	-17.9	-6.4
1d → 4k	-19.1	-17.6	-7.5
1e → 4l	-18.4	-16.6	-6.4
→ 4m	-18.3	-15.9	-7.9
1e' → 4n	-22.3	-20.4	-10.4
out-of-plane (4_{oop}) comformations			
1b → 4h	-21.8	-13.7	-8.7
(R_1 = -NH_2 ; in plane conformation = unstable)			
1b → 4h'	17.2	47.6	42.7
Activation Energies E_a (kcal/mol)			
TS4h	13.8	14.0	14.9
TS4i	13.6	14.0	14.6
TS4j	12.2	12.6	13.3
TS4k	12.0	11.8	12.6
TS4l	11.7	11.6	12.2
TS4m	11.9	12.4	13.0
TS4n	10.2	9.6	10.6
TS5h	1.1	6.0	1.9
TS5i	1.0	5.4	1.9
TS5j	0.9	6.3	1.7
TS5k	0.6	5.8	1.2
TS5l	0.4	5.5	1.1
TS5m	0.8	6.1	1.0
TS5n	0.3	5.7	0.7

Interesting similarities in the **TS4**, **TS5**, **4** and **TS2** geometries can be revealed by superimposing their skeletal structures (**Fig. 6.12**). Geometry changes in the azo-coupled side chain are small through the course of the (**TS4** → **4** → **TS5**) portion of **Path C**, and changes in the azole ring are minimal. Orientation of the Cβ=Cδ portion of the ynamine moiety in **TS2** is approximately an average of the orientations assumed by the same moiety in structures **TS4, TS5**, and **4**. However, **TS2** is a transition state, with one large imaginary vibrational eigenvalue, and **4** is an intermediate. Transition states of the type **TS2** cannot be located with R_2 = -NH_2; geometries in the vicinity of the **TS2** transition state revert to $\mathbf{4}_{ip}$ intermediates, since electron donation by the amino group stabilizes the 4ip structure, creating a true intermediate. Resonance is not possible when the amino group is in the R_1 position, so when ynamines with R_1 = -NH_2 assume geometries similar to $\mathbf{4}_{ip}$, they revert to (**TS2**) transition states.

Equilibrium structures of out-of-plane azo-coupled intermediates ($\mathbf{4}_{oop}$) could be located for both R_1 = -NH_2 and R_2 = -NH_2 isomers. Conformations of diazoazole + ynamine out-of-plane azo-coupled intermediates are similar to the **4a-4g**$_{oop}$ conformations found in the study of ethyne + diazoazole reactions. Azo-coupled out-of-plane ($\mathbf{4}_{oop}$) intermediates with R_1 = -NH_2 are 50-60 kcal/mol less stable than corresponding intermediates with R_2= -NH_2. $\mathbf{4}_{oop}$ structures with R_2 = -NH_2 are of comparable in energy to their in-plane $\mathbf{4}_{ip}$ structures (**Table 6.5**).

6.4.5. Mechanisms of Diazoazole/ynamine Cycloadditions: Examples of MNDO, AM1 and PM3 reaction coordinates for the **Path B** and **Path C** (**Fig. 6.13**) mechanisms show that all the semi-empirical methods favour **Path C**. The **Path C** mechanism, possible only when R_2 = -NH_2, yields the experimentally observed amino-azolotriazine isomer. Product isomers with R_1 = -NH_2 may only form through the unfavourable high energy concerted pathway (**Path B**) involving transition states **TS2h'-TS2n'**, and are therefore never observed experimentally. The low energy barriers of the **Path C** ynamine mechanism explain the greater reactivity of ynamines as compared with unsubstituted alkynes. Alkynes react through concerted [7+2] cycloadditions, with a high energy barriers of 28.4-41.1

kcal/mol, while ynamines are more reactive because they react through **Path C**, whose largest activation energy barrier is only 9.6-14.9 kcal/mol. The activation barrier heights reflect the temperatures at which these reactions are known to occur; reactions between alkynes and diazoazoles are carried out at temperatures around 30-40 °C,[1,19] while reactions between ynamines and diazoazoles give better yields, and can be carried out at lower temperatures (-10 to -30 °C).[20] The reaction barrier for the rate determining step in **Path C** decreases slightly with increasing number of nitrogen atoms in the azole ring, reflecting the experimentally observed order of diazoazole/ynamine reactivity (i.e. diazotriazoles > diazoimidazole, diazopyrazole > diazopyrrole).[1]

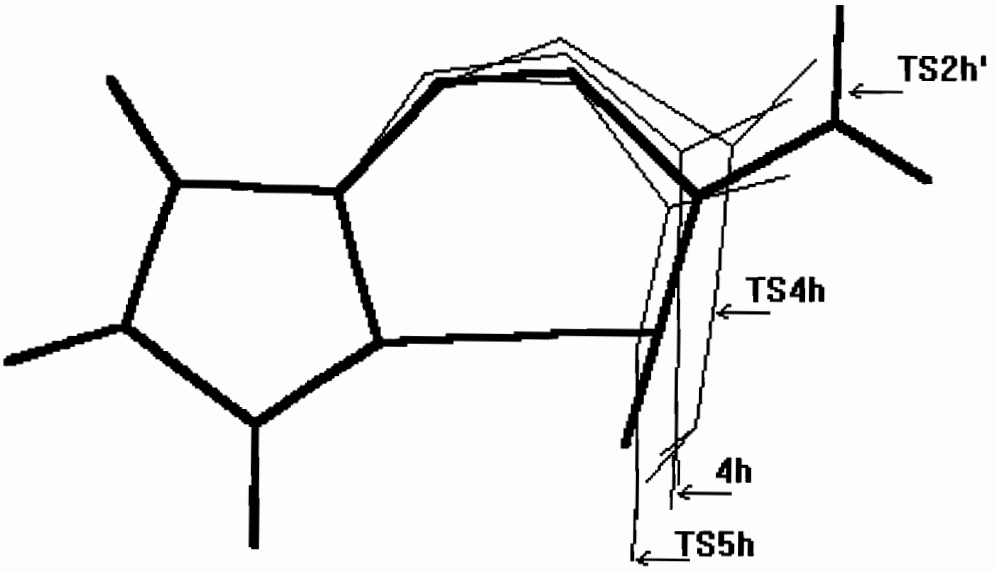
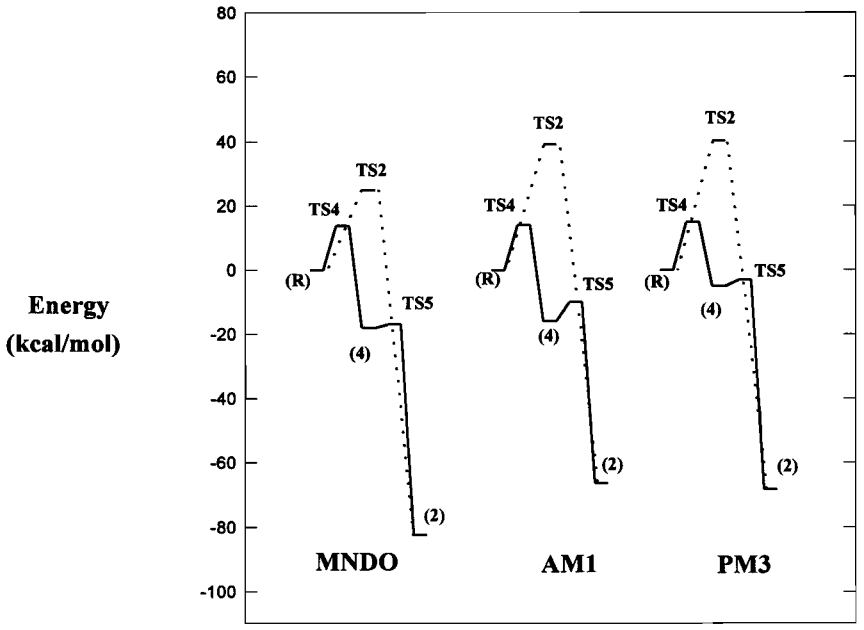


Figure 6.12:

Superimposed skeletal structures of **(TS4h, 4h, TS5h and TS2h')**

(TS2h' in bold)



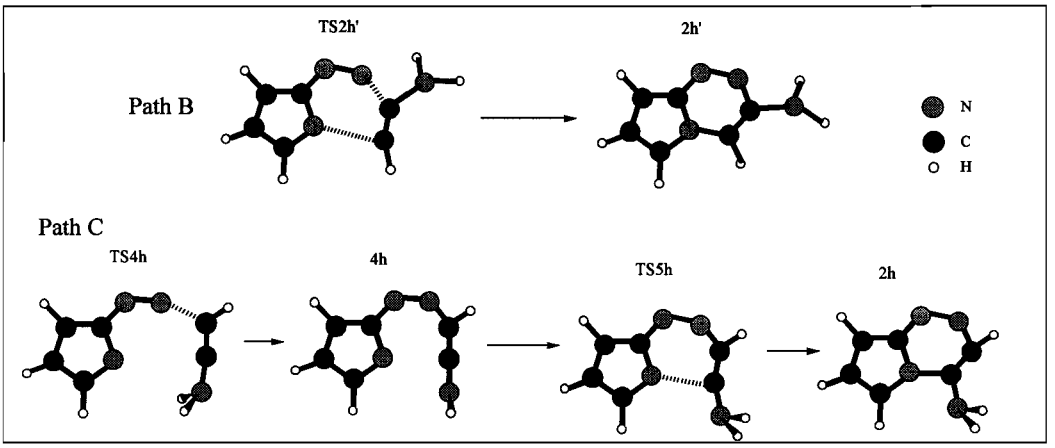
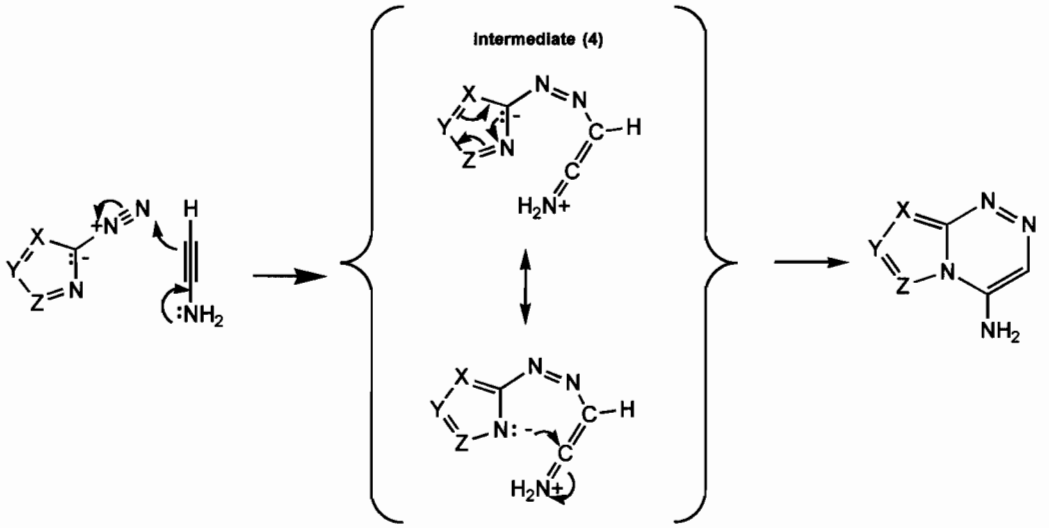


Figure 6.13:

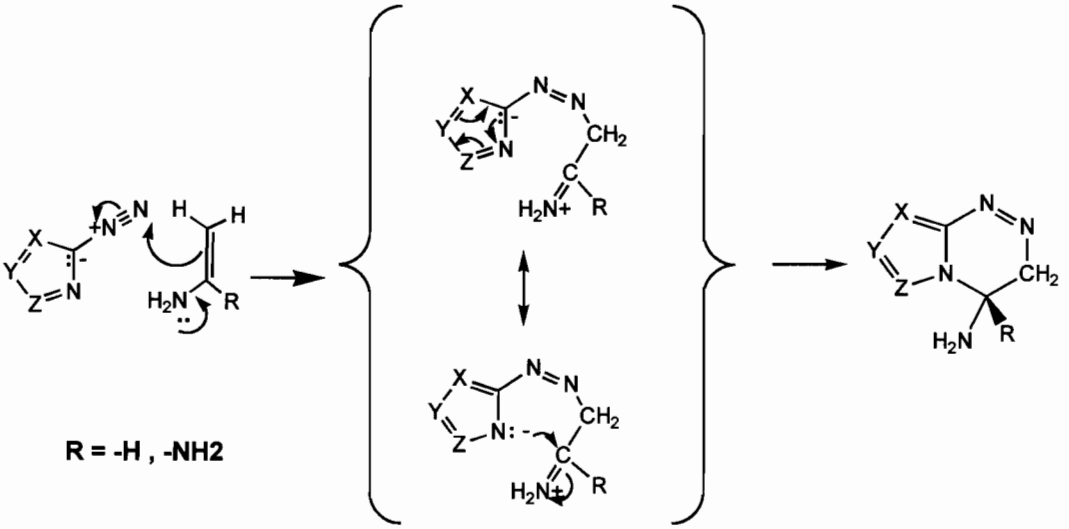
Representative Structures and Reaction Coordinates in Ynamine/Diazoazole Reactions

6.4.6. Reactivities of enamine and 1,1-diaminoethene: The first step of the **Path C** diazoazole + ynamine cycloaddition mechanism is the nucleophilic attack of the terminal diazo nitrogen by the ynamine, which forms the **4** type intermediate (**Fig. 6.14**). The **4** type intermediate then resonates with the azole ring, concentrating negative charge on the α N atom, which undergoes nucleophilic attack of the ynamine carbon ajacent to the amino group. Since the first step resembles the C alkylation reactions of enamines, the **Path C** mechanism was applied to diazoazole cycloaddition with enamine and 1,1-diaminoethene. Only the PM3 method was used. Both these reactions are found to occur through the **Path C** mechanism, and the order of reactivity (ynamine > 1,1-diaminoethene > enamine) reflects experimental observations.[1] The **Path C** mechanism also accounts for the observed stereospecificity in enamine and 1,1-diaminoethene diazoazole cycloaddition reactions.

6.4.7. Cyanoethyne: The hypothetical cyanoacetylene + diazoazole was studied to compare with the ynamine + diazoazole and ethyne + diazoazole reactions. Reaction energies in **Table 6.6** show that the cyano-azolotriazine are thermodynamically stable. The semi-empirical methods showed no marked thermodynamic preference for either the R_2 or R_1 positions of -CN in the azolotriazine. No spiro [3+2] (**TS3**) or planar [7+2] (**TS2**) transition states were found. High energy out-of-plane ($\mathbf{4}_{oop}$) type intermediates were found, with the R_2 = -CN isomers thermodynamically preferred by 10-15 kcal/mol (**Table 6.6**). Cyano substituted $\mathbf{4}_{oop}$ intermediates share similar conformational features with the ethyne and ynamine $\mathbf{4}_{oop}$ intermediates. However, calculations with cyanoethyne failed to find stable in-plane $\mathbf{4}_{ip}$ intermediates with cyano substituents, verifying the formal bonding prediction that in-plane **4** type intermediates with electron withdrawing substituents are unstable regardless of the -CN substituent position. Since no reasonable transition states were found for any of the proposed cyanoethyne/diazoazole cycloaddition mechanisms, these reactions probably do not occur. This was expected, since no cyanoethyne + diazoazole cycloaddition reactions have been reported.



Path C mechanism for Diazoazole/ynamine cycloaddition Reaction



Analagous Path C mechanism for enamine and 1,1-diaminoethene/diazoazole cycloaddition reactions

Figure 6.14:

Comparison of Proposed Mechanism (Path C) of Diazoazole/Ynamine Cycloaddition with Similar Possible Mechanism for Enamine and 1,1-Diaminoethene Reactions

Table 6.6: Energies for Diazoazoles + Cyanoethyne Reactions

			ΔH_{rxn} (kcal/mol)		
			MNDO	AM1	PM3
Diazoazole (1) + cyanoethyne → products (2)					
1b	→	2o	-72.3	-61.5	-61.7
	→	2o'	-78.9	-63.2	-63.4
1c	→	2p	-70.1	-59.7	-56.7
	→	2p'	-72.9	-61.3	-58.0
1c'	→	2q	-72.0	-59.5	-58.7
	→	2q'	-74.6	-61.2	-60.1
1d	→	2r	-72.3	-56.1	-58.7
	→	2r'	-71.7	-55.6	-59.0
1e	→	2s	-69.0	-54.2	-54.5
	→	2s'	-71.6	-56.1	-56.6
1e'	→	2t	-66.4	-51.9	-52.0
	→	2t'	-70.4	-57.0	-53.3
Diazoazole (1) + cyanoethyne → spirobicyclics (3)					
1b	→	3o	-52.0	-27.4	-34.8
	→	3o'	-52.6	-26.6	-35.3
1c	→	3p	-51.5	-27.3	-33.4
	→	3p'	-51.8	-26.2	-33.7
1c'	→	3q	-50.3	-26.3	-32.3
	→	3q'	-51.0	-25.7	-32.9
1d	→	3r	-50.1	-25.7	-32.7
	→	3r'	-50.7	-25.1	-33.3
1e	→	3s	-47.4	-23.7	-29.5
	→	3s'	-48.2	-23.1	-30.2
1e'	→	3t	-49.8	-26.2	-31.1
	→	3t'	-50.2	-24.9	-31.5
Diazoazole (1) + cyanoethyne → azo-coupled intermediated (out-of-plane) (4)					
1b	→	4o	36.1	38.7	40.1
	→	4o'	47.2	51.6	55.3

6.5 Conclusions:

In summary, the calculations predict the following for diazoazole reactions with carbon-carbon triple bonds:

1) Ethyne can react through either a [7+2] or a [3+2] concerted mechanism; however, the rearrangement from the spirobicyclic to the azolotriazine is not computationally described. The large activation barriers diazoazole/ethyne cycloadditions reflect the temperatures required for these reactions, and the low yields obtained.

2) Ynamine does not react by either a [7+2] or a [3+2] concerted mechanism, but rather through a two step mechanism, involving a metastable intermediate. The low activation barriers account for the low temperatures at which these reactions occur, and the high yields obtained. The semi-empirical determined reaction coordinates also account for the experimentally observed regioselectivity in ynamine/diazoazole cycloaddition reactions.[1]

3) The inability to locate any [3+2], [7+2] or azo-coupled transition states for the reaction of diazoazoles with cyanoethyne suggests that cyanoethyne does not undergo cycloaddition with diazoazoles. This computational result correlates with the fact that no cyanoethyne/diazoazole cycloaddition reaction products have ever been reported in the literature.

This chapter has shown that the MNDO, AM1 and PM3 semi-empirical methods can be used to ascertain relative reactivities and regiospecificities in chemical reactions involving these nitrogen-rich heterocyclic systems. In combination with the results in chapters three, four and five, the results in this chapter further justify the use of MNDO, AM1 and especially PM3 to study the structures, energies and reactivities of nitrogen-rich heterocyclics.

References:

1. G. Cirrincione, A.M. Almerico, E. Aiello and G. Dattolo, *Adv. Heterocycl. Chem.*, **48**, 65 (1990).

2. H. Durr, A.C. Ranade, and I. Halberstadt, *Synthesis*, 878 (1974).

3. A. Padwa and T. Kumagai, *Tetrahedron Lett.*, **22**, 1199 (1981).

4. A. Padwa ,T. Kumagai and A.D. Woolhouse, *J. Org. Chem.*, **48**, 2330 (1983).

5. H. Durr and H. Schmitz, *Chem. Ber.*, **111**, 2253 (1978).

6. A. Padwa and M. Tohidi, *J.C.S. Chem. Commun.*, 295 (1984).

7. M. Cabre, J. Ferras, J. Fernandez-Sanz, and J. Vilarrasa, *J. Chem. Soc., Perkin. Trans. 2,* 1943 (1990).

8. M.J.S. Dewar, *Pure Appl. Chem.*, **44**, 767 (1975).

9. A.R. Katritzky, K. Yannakopoulou, E. Anders, J. Stevens and M. Szafran, *J. Org. Chem.*, **55**, 5683 (1990).

10. M. Meyer, *J. Mol. Struct.*, **304**, 45 (1994).

11. E. Fos, J. Vilarrasa, and J. Fernandez, *J. Org. Chem.*, **50**, 4894 (1985).

12. W.M.F. Fabian, *Z. Naturforsch*, **45a** 1328 (1990).

13. J.P. Ritchie, *J. Org. Chem.*, **54** 3553 (1989).

14. B.S. Jursic and Z. Zdravkovski, *J. Mol. Struct., (THEOCHEM),* **312**, 11 (1994).

15. B.S. Jursic and Z. Zdravkovski, *J. Mol. Struct., (THEOCHEM),* **309**, 241 (1994).

16. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen, S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** (1990) 52.

    b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14,** 1347 (1993).

17. For discussions on transition state theory see
a) W.J. Hehre, L. Random, P.v.R Schleyer and J.A. Popel, "*Ab Initio Molecular Orbital Theory*", Wiley, New York, 1986.
b) W.J Albery, *Adv. Phys. Org. Chem.*, **28**, 139 (1993).

18. R. S. Hosmane, V. Bakthavachalam and N.J. Leonard, *J. Am. Chem. Soc.*, **104**, 235 (1982).

19. T. Yamazaki and H. Schechter, *Tetrahedron Lett.,* 1417, (1973).

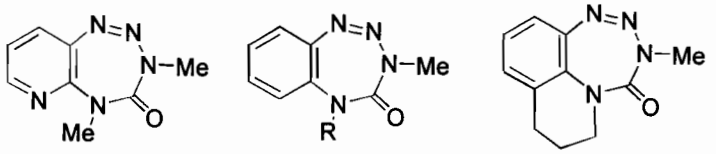
20. W.L. Magee, C.B. Rao, J. Glimka, H. Hui, T.J. Amick, D. Fiscus, S. Kakodar, M. Nair and H. Schechter, *J. Org. Chem.,* **52**, 5538 (1987).

Appendix 6.A

Table 6.A1:
Semi-empirical Heats of Formation and *ab initio* Total Energies of Diazoazoles and other Small Molecules in Study:

Compound	ΔH_f (kcal/mol)			total energy (6-31G**)
	MNDO	AM1	PM3	(a.u.)
Acetylene	57.9	54.8	50.7	-76.82184
Ynamine	63.1	54.1	58.1	--
cyanoacetylene	83.1	85.8	88.3	--
diazoazoles				
1b	101.7	120.3	97.5	-316.47557
1b'	102.9	116.7	99.3	-316.47108
1c	105.2	135.9	102.1	-332.47617
1c'	104.4	132.2	101.8	-332.47762
1d	111.9	141.3	118.1	-332.43846
1d'	113.5	139.6	121.0	-332.44187
1e'	119.4	165.6	138.0	-348.41688
1e	112.9	156.8	121.6	-348.44534

The results of chapters three through six strongly suggest the MNDO, AM1 and especially the PM3 methods are useful tools for prediction of the structure and reactivity of well known nitrogen heterocyclic compounds. Thus, these methods should also give useful structure and reactivity information about the relatively unknown tetrazepinone (TZP) class of compounds (**1-3**) synthesized at McGill by Jean-Claude and Just in 1991.[1,2]



1 **2** **3**

The research focus in chapter seven and eight shifts towards TZPs and related compounds, because of their promising anti-tumor properties,[1,2] and the absence of computational studies in the literature. Experimental work to date is contained in references 1-3, and all the completed computational work presented in chapters seven and eight[4,5] has been either submitted or accepted for publication. Chapters seven and eight will show how the PM3 method can accurately predict tetrazepinone structure and reactivity.

References:

1. B.J. Jean-Claude and G. Just, *J. Chem. Perkin Trans.* **1**, 2525 1991.

2. B.J. Jean-Claude, Ph.D Thesis, McGill University, 1992.

3. B.J. Jean-Claude and C.I. Williams, *Org. Mag. Reson.* accepted, October 1996.

4. C.I. Williams, M.A. Whitehead and B.J. Jean-Claude, *J. Mol. Struct. (THEOCHEM)*, accepted , (August 1996)

5. C.I. Williams, M.A. Whitehead and B.J. Jean-Claude, *J. Mol. Struct. (THEOCHEM)*, submitted (October 1996)

CHAPTER 7:

EQUILIBRIUM GEOMETRIES AND ENANTIOMERIC INTERCONVERSIONS IN TETRAZEPINONE RING SYSTEMS: A MOLECULAR MECHANICS, SEMI-EMPIRICAL AND *AB INITIO* TREATMENT

by C.I. Williams, M.A. Whitehead*
Theoretical Chemistry Laboratory, Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6
and

Bertrand J. Jean-Claude
Department of Oncology, McGill University, 3655 Drummond St., Ste. 701, Montreal, Quebec, Canada H3G 1Y6

7.1 Abstract:

Molecular Mechanics (MM+), semi-empirical (MNDO, AM1, PM3) and *ab initio* (RHF/3-21G, RHF/6-311G) methods are used to generate optimized theoretical benzotetrazepinone geometries. The optimized structures are compared to X-ray data. The PM3 and *ab initio* methods most successfully reproduce X-ray benzotetrazepinone geometries. Important structural features of the non-planar benzotetrazepinone conformation are discussed, and a procedure to generate optimized benzotetrazepinone geometries in the absence of X-ray data is presented. Semi-empirical and *ab initio* optimized geometries of the hypothetical non-benzofused (lone) tetrazepinone ring are compared with the optimized geometry of the benzofuzed counterpart. The results indicate that both the lone and benzofused tetrazepinone rings have similar non-planar conformations. The racemization of chiral benzotetrazepinone is studied with semi-empirical methods and found to occur through a concerted mechanism with a chiral transition state. The energy barrier to benzotetrazepinone racemization depends upon the size of the alkyl substituents on the N3 and N5 tetrazepinone ring atoms.

7.2 Introduction:

The benzo- and benzoalkyl-fused compounds **1a-d**, (**Fig. 7.1**) recently synthesized by Jean-Claude et. al.[1] represent a new class of molecules containing the novel seven-membered tetrazepinone ring system. There are no experimental examples of the other benzotetrazepinones shown (**1e-f**), or of the non-benzofused (henceforth refered to as lone) tetrazepinone ring (**2**, **Fig. 7.2**); however, these predicted compounds will also be studied. X-ray structural determinations of nitrobenzotetrazepinone (**1c**) and benzoalkyltetrazepinone (**1d**) indicate that the tetrazepinone part of the molecule has a non-planar conformation, with the N2, N3, C4 O14, C12 and C13 atoms significantly out of the benzene ring plane.[1] The X-ray structure also indicates that the N3 atom is almost perfectly tetrahedral (bond angles near 109°, and an improper torsion angle near 120°), and that the N5 atom is moderately tetrahedral (improper torsion angle near 157.3°). These tetrahedral nitrogens constitute two chiral centers in the tetrazepinone ring, and the X-ray unit cell contains both enantiomeric forms of the benzotetrazepinone.[1] However, whether or not the tetrazepinone ring is locked in its enantiomeric form is unknown, because no attempt has been made to resolve the enantiomers. Benzotetrazepinones could convert between enantiomeric forms by N3 and N5 nitrogen inversions and 'flipping' of the tetrazepinone ring.

The work of other researchers,[2-24] and the results in chapters three to six demonstrate that semi-empirical (MNDO, AM1, PM3) and *ab initio* methods can sucessfully provide useful structural and energetic information on a wide variety of nitrogen containing ring systems. Therefore, these methods will be used to study tetrazepinone structure and conformational flexiblity in this chapter. Benzotetrazepinone synthesis and reactivity will be treated in chapter eight.

This chapter will discuss important tetrazepinone ring structural features, such as the tetrahedral character of the N3 and N5 atoms, and co-planarity within different regions of the tetrazepinone ring. Molecular mechanics, semi-empirical and *ab initio o*ptimized benzotetrazepinone geometries will be compared with X-

ray data, and the ability of each theoretical method to reproduce important benzotetrazepinone structural features will be assesed.

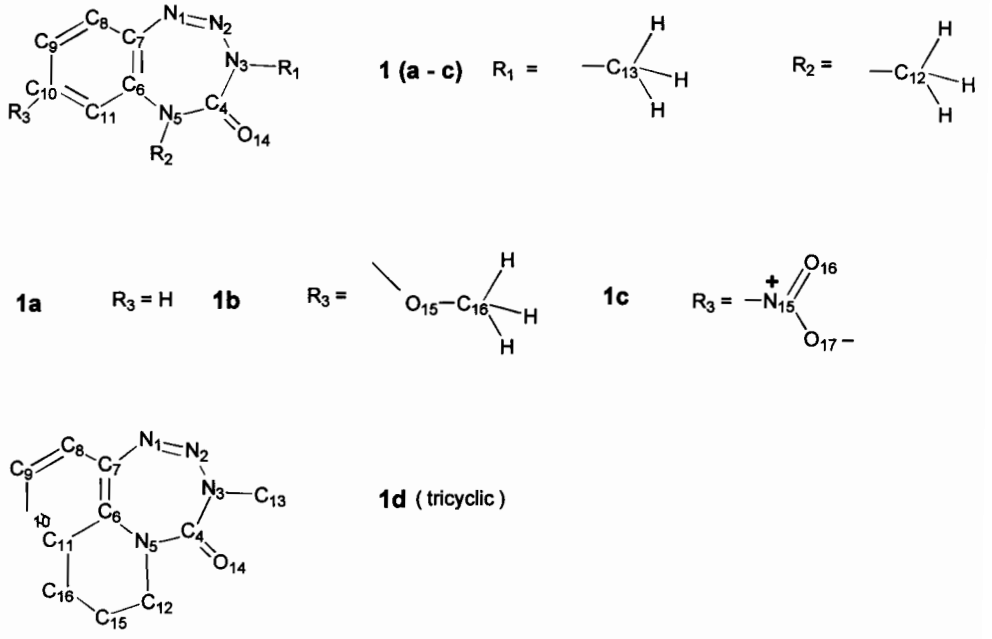
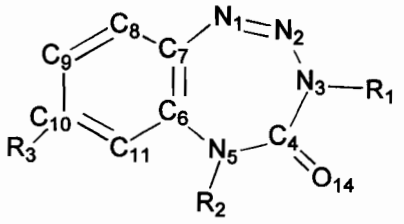


Figure 7.1a:

Recently Synthesized Benzotetrazepinones



1e R_3 , R_1 , R_2 = H

1f R_3 = H ; R_1 , R_2 = t-B

Figure 7.1b:

Hypothetical Tetrazepinones Predicted in this Study

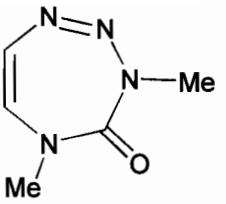


Figure 7.2:

Hypothetical Lone Tetrazepinone Ring Predicted in this Study

The ability of theoretical methods to predict *a priori* , in the absence of X-ray data, the out-of-plane equilibium conformation of the tetrazepinone ring will also be addressed. A short procedure for generating optimized benzotetrazepinone geometries in the absence of experimental data will be outlined, and optimized geometries generated with this procedure compared with optimized geometries obtained from X-ray input. The procedure will also be used to generate optimized non-benzofused (lone) tetrazepinone rings (**2**). Structural features of the optimized lone tetrazepinone will be compared with the corresponding structural features in the optimized benzotetrazepinones.

Different possible conformations of the tetrazepinone ring in benzotetrazepinone will be investigated. Benzotetrazepinone chirality, and a mechanism for benzotetrazepinone enantiomeric conversion (racemization) will be discussed, along with the effect of N3 and N5 substituents on the racemization energy barrier.

7.3 Calculations:

Molecular Mechanics calculations were performed on a PC using the MM+ force field supplied in the HyperChem[25] package. All semi-empirical and *ab initio* calculations were performed using the GAMESS package[26] on a Sun Sparc10 workstation. Full geometry optimization was allowed in each calculation, and the optimizations were carried out in Cartesian (as opposed to internal)

coordinates. Equilibrium geometries and transition states were verified by force constant analysis. All quantum mechanical calculations used RHF wavefunctions.

7.4 Results and Discussion:

7.4.1 Out-of-Plane Equilibrium Geometries:

Theoretical geometries of nitrobenzotetrazepinone (**1c**) and the tricyclic benzoalkyltetrazepinone (**1d**) were generated by optimizing the X-ray structures. The computationally determined bond lengths, bond angles and torsion angles of **1c** are compared with X-ray **Table 7.A1** (**Appendix**). For clarity, only the heavy atoms of the aromatic and tetrazepinone ring systems are given. In lieu of discussing each geometric parameter individually, the ability of each computational method to predict important benzotetrazepinone structural features, and the effect of each structural feature on the out-of-plane benzotetrazepinone conformation, will be discussed.

The tetrahedral character of the N3 and N5 atoms can be measured using the N3-C4-N2-C13 and N5-C4-C6-C12 improper torsion angles, henceforth refered to as ΦN3 and ΦN5. Angle ΦN3 represents the angle between the N3-C13 bond and the plane of the N3-C4-N2 atoms, while angle ΦN5 represents the angle between the N5-C12 bond and the plane of the N5-C4-C6 atoms. A value of 180° for either of these angles is assumed to be a perfectly trigonal planar nitrogen atom, while a value of 120° is assumed to be a perfectly tetrahedral nitrogen atom. The degree of N3 and N5 tetrahedral character is important because tetrahedral N3 and N5 atoms can be chiral centers, while trigonal planar N3 and N5 atoms cannot.

7.4.2. Tetrahedral Character of N3: The X-ray values of 128.5° and 129° for the ΦN3 improper torsion angles in **1c** and **1d** indicate that the N3 atom is almost

perfectly tetrahedral in both compounds. Compared with X-ray, all computational methods underestimate the tetrahedral character of N3. PM3 and *ab initio* methods give satisfactory results, which deviate from X-ray ΦN3 improper torsion angles by only 8°- 10°. MNDO, AM1 and MM+ predict the N3 atom to be much more planar than in the X-ray structures, with ΦN3 improper torsion angles 30° larger than X-ray values.

The tetrahedral character of N3 correlates with the bond lengths around N3. X-ray and computed C4-N3 and N3-C13 bond lengths are in good agreement, and all correspond to single bonds. The X-ray N3-N2 bond length of 1.446 Å also corresponds to a single bond, so X-ray bond lengths around N3 correlate well with the tetrahedral character of N3. PM3 and *ab initio* N2-N3 bond lengths match X-ray, and also indicate N-N single bonds. This correlates with the substantial tetrahedral character predicted for N3 by both the PM3 and *ab initio* methods. In contrast, AM1, MNDO and MM+ predict N3-N2 bond lengths which are closer to that of an N=N double bond. Double bonding between N2 would N3 would imply greater sp^2 hybridization and trigonal planar character of the N3 atom, so AM1, MNDO and MM+ failures to predict substantial tetrahedral character for N3 also assign too much double bond character to the N2-N3 bond.

7.4.3. Tetrahedral Character of N5: PM3, *ab initio* and X-ray results all predict the N5 atom to be less tetrahedral than the N3 atom. The MNDO method correctly reproduces the X-ray value for the ΦN5 improper torsion angle in nitrobenzotetrazepinone **1c**, but it erroneously predicts almost equal tetrahedral character for N3 and N5 in this compound. Both MNDO and AM1 erroneously predict greater tetrahedral character for N5 than for N3 in benzoalkyltetrazepinione **1d**.

The tetrahedral character of N5 correlates with X-ray and calculated bond lengths around N5. All computational methods successfully reproduce X-ray N5-C12 and N5-C6 bond lengths, both of which correspond to N-C single bonds. In

contrast, X-ray and *ab initio* values of the C4-N5 bond lengths (1.32 Å to 1.36 Å) are smaller than ideal C-N single bond lengths (1.47 Å), and closer to the ideal C=N double bond lengths (1.30 Å); this suggests a substantial amount of double-bonding between N5 and C4. The partial double bond between N5 and C4 relates to the lower tetrahedral character of N5 as compared with the N3. The N5-C4 bond length is well predicted by the *ab initio* methods only; all other methods predict N5-C4 bond lengths which are larger than X-ray, and close to N-C single bond lengths. Consequently, all semi-empirical methods predict a more tetrahedral ΦN5 improper torsion angle than is found in the *ab initio* or X-ray structures. PM3 still correctly predicts more tetrahedral character for N3 than for N5, but AM1 reverses this order. MNDO and MM+ both erroneously predict N3 and N5 to have similar tetrahedral character.

7.4.4. Plane A:

Coplanarity of the Triazine Moiety (N3-N2-N1-C7): The pseudo-coplanarity of the N3-N2-N1-C7 portion of the tetrazepinone ring can be measured with the N3-N2-N1-C7 torsion angle. X-ray results show that this group deviates only slightly from planarity, with an N3-N2-N1-C7 torsion angle of less than 1°. This tetrazepinone structural feature is well reproduced by all computational methods, with N3-N2-N1-C7 torsion angles ranging from 4.7° to 0.02°. The plane of the triazine group will be referred to as **Plane A**.

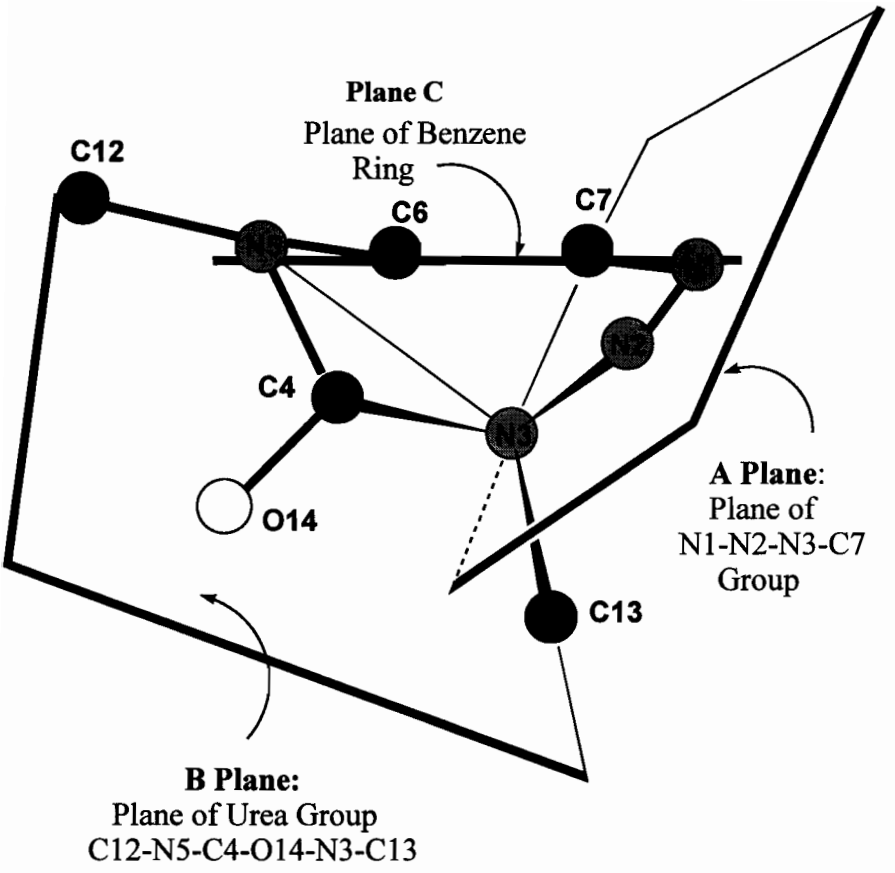
7.4.5. Plane B:

Coplanarity of the Urea Moiety (C13-N3-C4-O14-N5-C12): If the urea moiety, (C13-N3-C4-O14-N5-C12) were perfectly planar, the O14-C4-N3-C13 and O14-C4-N5-C12 torsion angles would be 0° and the C13-N3-C4-N5 and C12-N5-C4-N3 torsion angles would be 180°. X-ray values for the O14-C4-N3-C13 and O14-C4-N5-C12 torsion angles range from -2.5° to 1.73° , and X-ray C13-N3-C4-N5 and C12-N5-C4-N3 torsion angles range from 177° to 178° , which indicates a

significant degree of coplanarity in this group. This structural feature will be referred to as **Plane B.**

The pseudo-coplanarity of the urea moiety is well predicted by both *ab initio* and PM3. The largest *ab initio* or PM3 deviation from the X-ray torsion angles which define **Plane B** is only 7°. AM1 predicts a moderate deviation from in this moiety (20°), while MNDO and MM+ are both in drastic disagreement with the X-ray values, and predict large deviations from coplanarity in this moiety (> 40°). Thus, only PM3 and *ab initio* correctly predict the experimentally observed structural feature **Plane B**.

7.4.6. Influence of the C6-C7-N1-N2 and C4-N5-C6-C7 Torsion Angles on Tetrazepinone Ring Conformations: Coplanarity within both the C7-N1-N2-N3 and urea portions of the tetrazepinone allows use of three planes to describe the benzotetrazepinone molecule; **Plane A** , the plane of the C7-N1-N2-N3 group, **Plane B,** the plane of the urea moiety, and **Plane C**, the plane of the benzene ring (**Fig. 7.3**). A schematic which exaggerates the C6-C7 bond distance is included, to emphasize that the C6-C7-N1-N2 (α) and C4-N5-C6-C7 (β) torsion angles essentially govern the out-of-plane benzotetrazepinone conformation. *Ab initio* values of 52.6 ° and 50.4 ° for the C4-N5-C6-C7 (β) torsion angles are in best agreement with experimental value of 50.0 ° . MM+ and AM1 also give good values for this torsion angle, but PM3 and MNDO overestimate this torsion angle by 10°-14°. However, the PM3 value of the C6-C7-N1-N2 (α) torsion angle is in best agreement with X-ray values for both compound **1c** and **1d**.



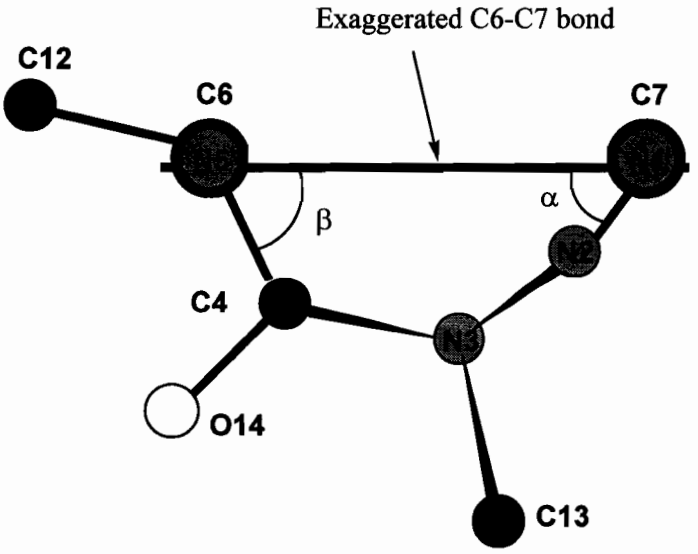


Figure 7.3:

Views of the Tetrazepinone Out-of-Plane Conformation;

Planes and Relevant Torsion Angles in Tetrazepinone

Despite numerical differences, all computational methods adequately predict an overall conformation of the tetrazepinone ring similar to the X-ray structure, with ground state stationary points having out-of-plane tetrazepinone ring conformations. The *ab initio* 3-21G method gave the best match to X-ray of all computational methods. PM3 gave the closest match to X-ray of the semi-empirical methods, while MM+ gave the worst match to X-ray of all methods. These results are exemplified in **Fig. 7.4**, where MM+, PM3 and *ab initio* optimized geometries of **1c** are overlayed with the X-ray geometry of **1c**. Three views of the overlay are given, in each case, to emphasize where optimized structures agree and disagree with the X-ray results. PM3 and *ab initio* optimized geometries coincide well with the X-ray structure. The MM+ optimized structure does not coincide with the X-ray structure as well the PM3 and *ab initio* optimized structures, but overall it still produces an out-of-plane benzotetrazepinone conformation which is qualitatively similar to the X-ray structure. The AM1 and MNDO overlays are not given, but the quality of coincidence between their optimized structures and the X-ray structure is intermediate of that between the PM3 and MM+ optimized structures. It should be noted that MNDO puts the nitro group prependicular to the plane of the benzene ring in compound **1c**, as does the MNDO treatment of nitrobenzene. This error will not be discussed further. The MM+ fails drastically with many structural features; however, the following section demonstrates that MM+ structures are useful as inputs to semi-empirical geometry optimizations.

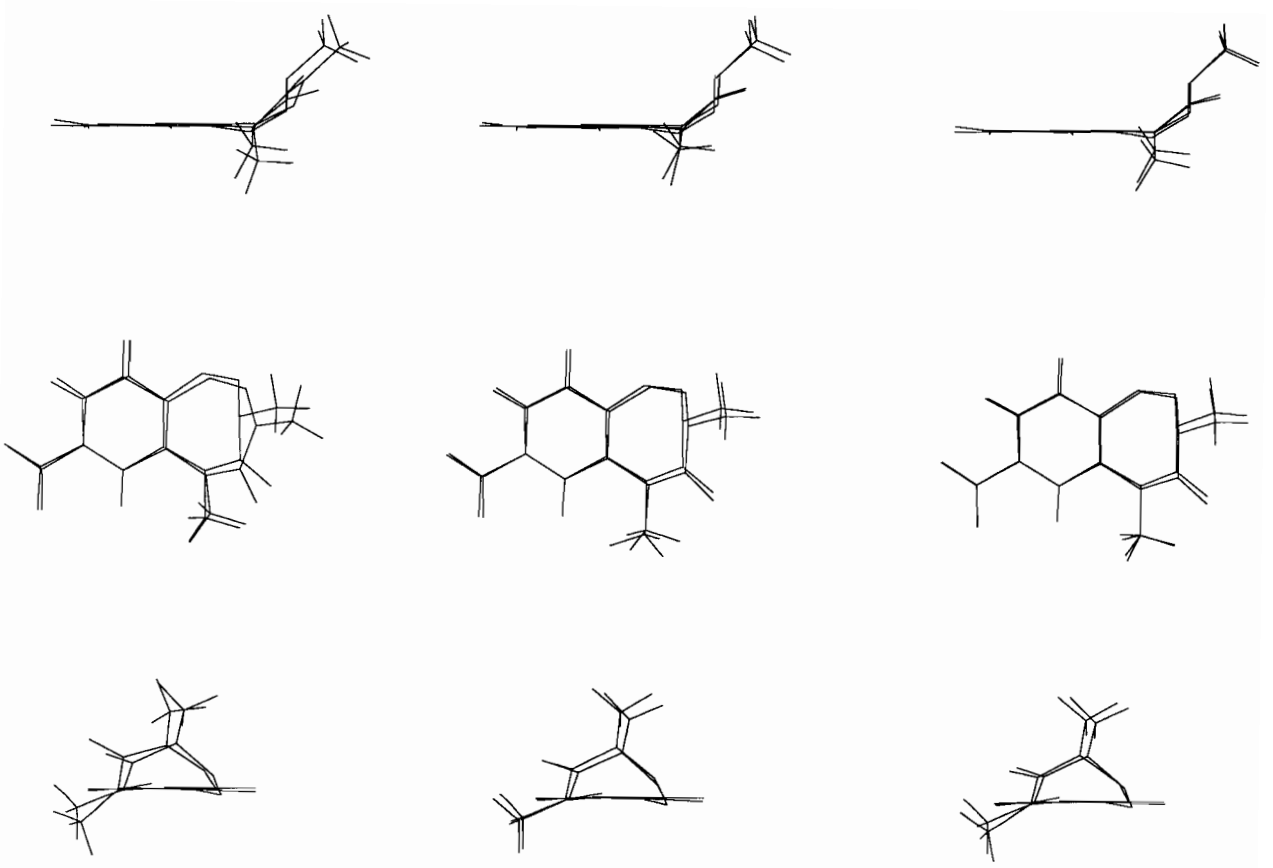


Figure 7.4:

Overlays of MM+, PM3 and 3-21G Optimized and X-Ray

Nitrobenzotetrazepinone Geometries

7.4.7. A Priori Generation of Tetrazepinone Ring Structures:

Since all the computational methods were successfully reproduced an out-of-plane benzotetrazepinone equilibrium geometries with X-ray geometries as input, the possiblity of reproducing optimized benzotetrazepinone geometries *a priori* (in the absence of X-ray data) was investigated.

MM+, semi-empirical and *ab initio* optimized structures were generated without X-ray input as outlined in **Fig. 7.5**. This procedure was applied to nitrobenzotetrazepinone (**1c**) and benzoalkyltetrazepinone (**1d**). The resulting optimized structures were compared with structures generated by optimization of X-ray geometries.

A priori MM+ Optimized Structures: Inputs for the MM+ calculations were created by simply drawing a planar benzotetrazepinone system in the HyperChem[26] visualizer. MM+ optimization of this input structure gave a planar tetrazepinone ring, which was initially interpreted as a failure for MM+ method; however, when the planar MM+ structure was then slightly distorted out-of-plane by manually adjusting a few torsion angles by 1-5°, subsequent MM+ minimization resulted in an out-of-plane structure 8.7 kcal/mol more stable than the planar structure. Small changes to MM+ structures must always be made, and the resulting structures re-minimized, otherwise errors and incorrect predictions will result.

The out-of-plane MM+ generated optimized structure was compared with a MM+ structure generated by MM+ minimization of the X-ray structure. Initially, the structures differed slightly, but re-minimization of each with a more stringent criteria for geometry convergence (maximum RMS gradient of 0.001 kcal/mol.Å), resulted in identical structures. Thus, the MM+ optimized tetrazepinone geometry can be determined without X-ray input. This is not surprising, because relatively few minima should exist for these ring systems, and local mimima with structures drastically different from the true minimum are unlikely.

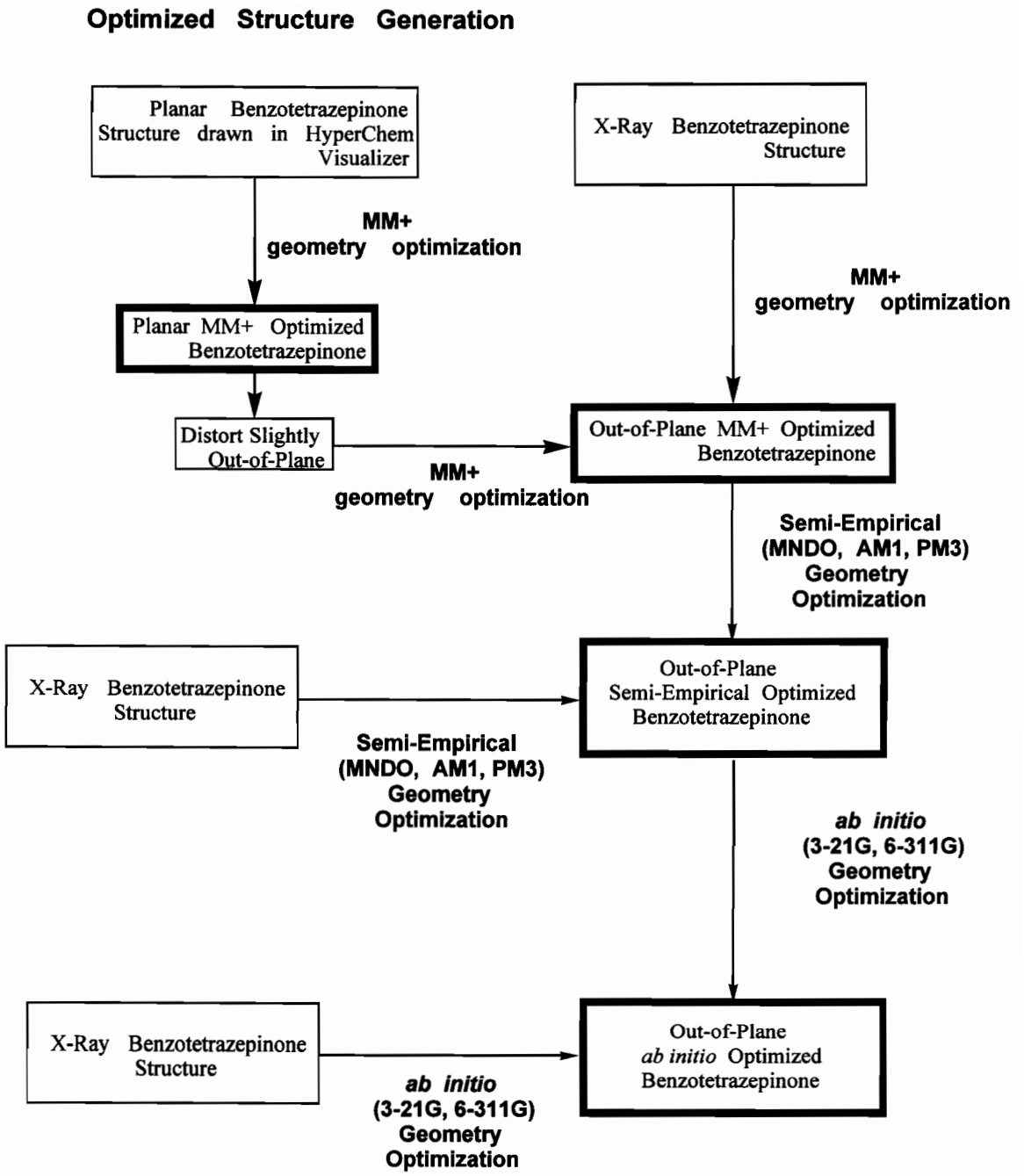


Figure 7.5:

Procedure for Obtaining Optimized Tetrazepinone Geometries with and without X-Ray Data. (plain boxes indicate input; bolded boxes indicate optimized structures, which can be used as input)

7.4.8. A priori Semi-Empirical and Ab initio Structures: Semi-empirical optimizations were initiated with both the X-ray structures, and the out-of-plane MM+ optimized structure above. The criteria for geometric convergence of semi-empirical and *ab initio* geometry optimizations was set at a maximum RMS gradient of 0.001 kcal/(mol.Å). Regardless of input (X-ray or optimized MM+), the resulting semi-empirical optimized structures were identical. 3-21G and 6-311G optimizations of **1c** and **1d** initiated with both X-ray and PM3 semi-empirical optimized structures also produced identical optimized benzotetrazepinone geometries. Therefore, by following the procedure outlined in **Fig. 7.5,** it is possible to generate the MM+, semi-empirical and *ab initio* optimized tetrazepinone geometries in the absence of X-ray input.

7.4.9. Hypothetical Non-fused (Lone) Tetrazepinone Rings:

The results of the structure generations above show that X-ray structures are not required to generate the MM+, semi-empirical or the *ab initio* optimized structures. Since the procedure in **Fig. 7.5** produced optimized benzotetrazepinone geometries identical to those obtained by optimization of X-ray structures, this procedure was used to generate optimized lone tetrazepinone ring structures (**2**), for which there are no experimental data. The effect of benzofusion on the lone tetrazepinone ring geometry was then assessed by comparing computed bond lengths, bond angles and torsion angles of the lone tetrazepinone ring (**2**) with computed values of the corresponding geometric parameters in the **1c** benzotetrazepinone (**Table 7.A2**). For convenience, the heavy atom numbering scheme is the same in both the lone and benzofused tetrazepinones.

All methods predict that only a few structural features are substantially affected by benzofusion. The only tetrazepinone bond length greatly affected by benzofusion is the C6-C7 bond length. The C6-C7 bond length is close to that of a true C=C double bond (1.335 Å) in the lone tetrazepinone, while in the benzofused compound (1.390-1.420 Å) it is closer to an ideal benzene C-C bond (1.41 Å). The C6-C7 bond is greatly affected by benzofusion because it is the only bond shared by the two ring systems in the benzofused compound. The *ab initio* and PM3 bond angles most affected by benzofusion are the C7-C6-N5 (126°-128° in the lone tetrazepinone, 116° -121° in benzotetrazepinone) and the N1-C7-C6 (125° in the lone tetrazepinone, 123° -124° in benzotetrazepinone) bond angles. These bond angles are larger in the lone tetrazepinone than in the benzofused tetrazepinone presumably because rehybridization opens up the angles to reflect the shorter C6-C7 bond length in the lone tetrazepinone. MNDO and AM1 predict larger differences in lone and benzofused tetrazepinone bond angles (as large as 10°) than do *ab initio* or PM3. All methods predict that the C4-N5-C6-C7 torsion angle β (which is important to the out-of-plane tetrazepinone conformation) is substantially affected by benzofusion. PM3 and *ab initio* predict small 17° and 11° increases, while MNDO and AM1 predict much larger 40° -45° increases in this torsion angle on going from the benzo-fused to the lone tetrazepinone ring. Only MNDO and AM1 predict a change of greater than 10° in the N2-N1-C7-C6 torsion angle (α), which also affects the out-of-plane tetrazepinone conformation. PM3 and *ab initio* predict small changes (< 10°) in this torsion angle upon benzofusion. The greatest differences between lone- and benzofused torsion angles are those predicted by MNDO, while the smallest differences between lone and benzofused geometries are those predicted by PM3 and *ab initio*.

PM3 and *ab initio* predict little effect of benzofusion on the tetrahedral chararacter of N3. However, both *ab initio* and PM3 predict that the N5 atom is

more trigonal planar in the lone tetrazepinone ring than in the benzofused tetrazepinone. The AM1 and MNDO improper torsion angles ΦN3 and ΦN5 are both greatly affected by benzofusion. Interestingly, both MNDO and AM1 predict a greater tetrahedral character for the N3 atom in the lone tetrazepinone ring than in the benzofused ring. MNDO, AM1 and *ab initio* all predict near perfect trigonal planar character for N5 in the lone tetrazepinone ring.

Overall, benzofusion has little effect on the tetrazepinone ring geometry. The largest differences between lone- and benzofused tetrazepinone are in the AM1 and MNDO geometries. PM3 differences between the lone and benzofused-tetrazepinone geometries are not as drastic, while the a*b initio* 3-21G predicts the smallest conformational differences between benzo-fused and non-benzo fused systems.. However, regardless of the differences between the methods, they all predict that the lone non-benzofused tetrazepinone is in a non-planar conformation similar to its benzofused counterpart.

7.4.10. N3 and N5 Nitrogen Inversions and Ring Flipping:

Possible equilibrium structures and various conformations of benzotetrazepinone **1a**, along with absolute configurations of the N3 and N5 atoms and their lone pair lobes are depicted in **Fig. 7.6**. Both the (**RS**)-**1A** and (**SR**)-**1A** enantiomers are represented in the X-ray unit cell, but there is no experimental evidence for any of the other structures in **Fig. 7.6**. For the (**RS**)-**1A** structure to racemize to its (**SR**)- **1A** enantiomer, both the N3 and N5 nitrogens must invert their configurations, and the tetrazepinone ring must flip. Simple inversion of the N3 and N5 atoms without flipping the tetrazepinone ring

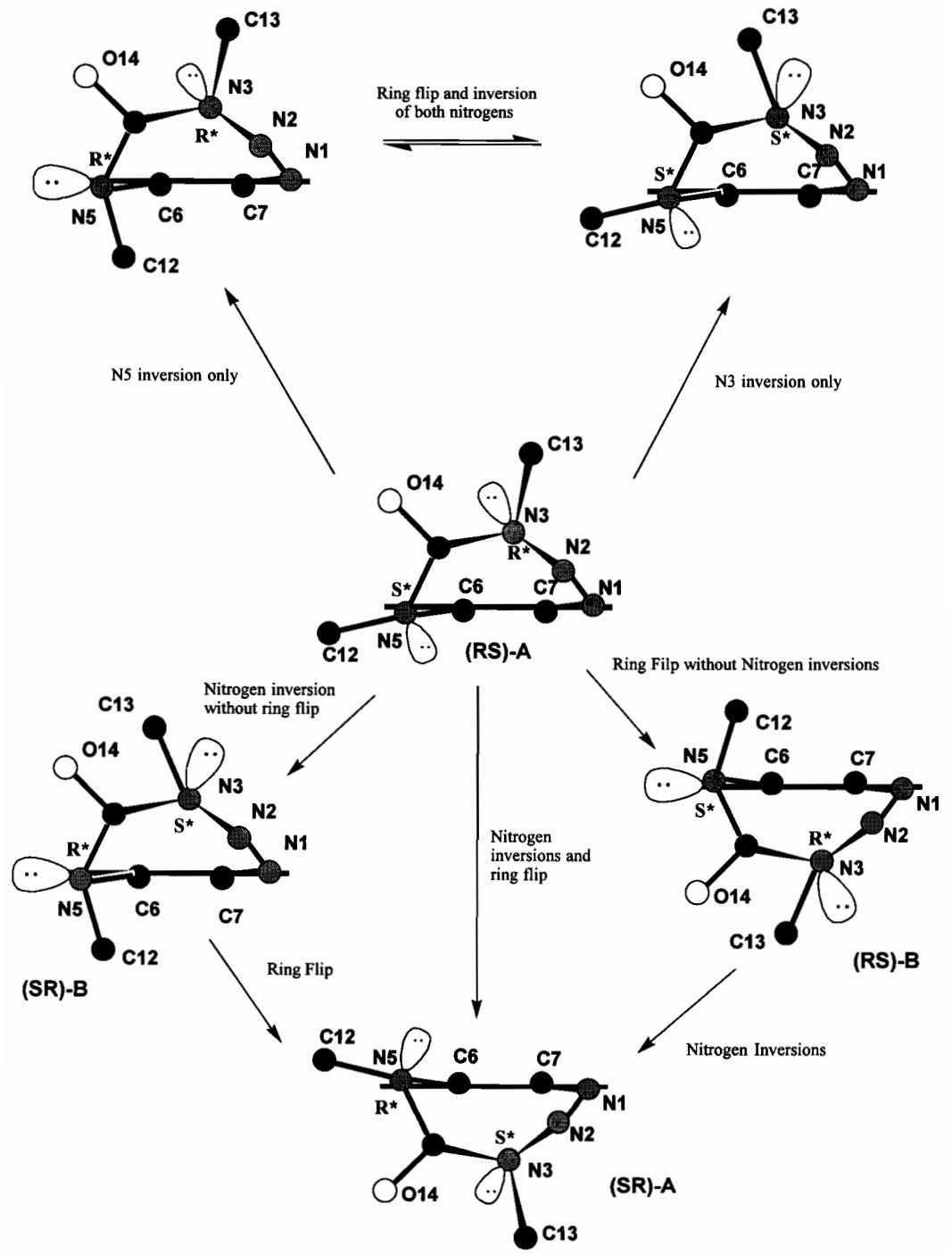


Figure 7.6:

Possible Tetrazepinone Stereoisomers and Routes for their Interconversion

results in hypothetical structure (**SR**)-**B**, which is not the enantiomer of (**RS**)-**1A**. Flipping of the tetrazepinone ring without inverting the nitrogens results in structure (**RS**)-**B**, which is also not the enantiomer of (**RS**)-**1A**. The nitrogen inversions change the absolute configurations of the nitrogens but do not correctly orient the tetrazepinone ring, while flipping of the tetrazepinone ring correctly orients the ring system, but does not change the absolute configurations of the N3 and N5 atoms. Therefore, N3 inversion, N5 inversion and tetrazepinone ring flipping must all occur for racemization to be complete.

Inversion of only one asymmetric center in (**RS**)-**1A** will necessarily result in a diastereomer; Inversion of only N3 results in structure (**SS**)-**C**, while inversion of only N5 results in structure (**RR**)-**D**; both are diastereomers of **1A**. (**RR**)-**D** and (**SS**)-**C** can be converted to each other through ring flipping and nitrogen inversion. Although structures (**RR**)-**D** and (**SS**)-**C** can both exist, they are diastereomers of (**RS**)-**1A**, and are therefore probably not invloved in the enantiomeric conversion.

Following these arguments, all possible structures discussed above were investigated by manually generating relevent input structures for PM3 geometry optimization. Equilibrium structures could be found only for (**RS**)-**1A** and (**SR**)-**1A**. All other structures either converted to (**RS**)-**1A** or (**SR**)-**1A** upon geometry opimization or the SCF would not converge. Therefore, it seems that none of the above structures are involved in the tetrazepinone enantiomeric interconversion, and racemization must follow a concerted mechanism.

7.4.11. Concerted Enantiomeric Interconversions: Concerted benzotetrazepinone racemization could occur through either an achiral planar transition state or a non-planar chiral transition state. A planar transition state seems reasonable because it would maintain planarity of both the urea and C7-N1-N2-N3 moieties throughout the ring flipping process. Such a transition state would imply that ring flipping

and nitrogen inversions occur simultaneously. A non-planar chiral transition state is also feasible, but such a transition state could mantain coplanarity of the urea group only when both nitrogens invert simultaneously; a non-planar chiral transition state would imply sequential nitrogen inversions and ring flipping, which could be resolved into distinct processes.

7.4.12. Planar Enantiomeric Conversion Transition State: The planar MM+ benzotetrazepinone geometry (**Fig. 7.5**) was used as input for semi-empirical and *ab initio* geometry optimizations of the planar tetrazepinone, which were confined to planarity by enforcing C_s symmetry. Stationary point structures were found in all cases, but force constant analysis revealed that all *ab initio* and semi-empirical planar benzotetrazepinone structures have more than one imaginary vibrational frequency, indicating that these are neither intermediates nor true transition states. However, for further comparison with benzotetrazepinone racemization mechanisms involving non-planar transition state(s), the energy difference between the 'planar' and equilibrium benzotetrazepinones is taken as the 'barrier' to ring flipping through a planar structure.

7.4.13. Non-planar Enantiomeric Conversion Transition State: A true non-planar transition state (with only one imaginary vibrational frequency) was found for the benzotetrazpinone racemization. As depicted in **Fig. 7.7**, the transition state is chiral. A complete geometry of a sample transition state is in **Table 7.A3.** Barriers to enantiomeric conversion through the 'planar' tetrazepinone structure and the non-planar transition states are compared in **Table 7.1.** Even if the planar transition states were valid, the non-planar interconversion would be favoured by 5-10 kcal/mol.

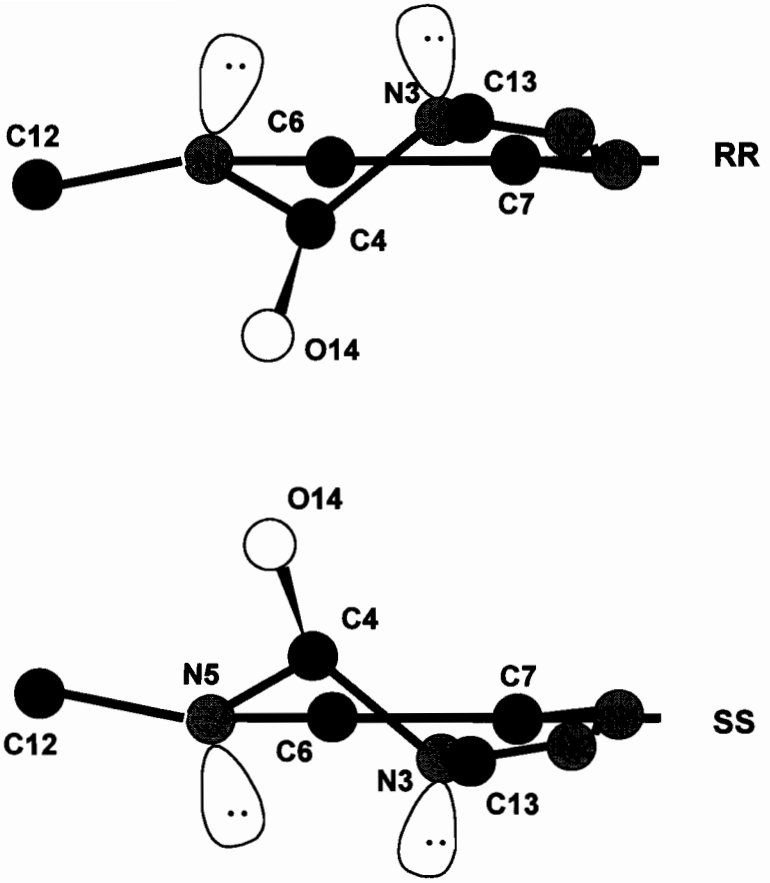


Figure 7.7:

Transition State for Benzotetrazepinone Enantiomeric Interconversion

Both (RR)- and (SS)- configurations are shown

Table 7.1: Barriers to Concerted Enantiomeric Conversion through Planar and Non-Planar Transition States :

Structure	Energy Relative to Equilibrium Structure (kcal/mol)				
	MNDO	AM1	PM3	MM+	3-21G
'Planar'*	18.2	5.3	10.5	8.7	9.2
Non-planar TS	8.7	4.2	4.0	-	-

* no valid structures; they contain more than one imaginary vibrational frequency.

The individual nitrogen inversion and the flipping of the tetrazepinone ring were resolved by investigating structures along the PM3 intrinsic reaction coordinate (IRC) of the interconversion path. Stick diagrams of the structures along the IRC path (**IRC1-15**) in **Fig. 7.8** show the tetrazepinone ring at different points in the enantiomeric conversion. The structure number refers to its position on the IRC path; **IRC1** and **IRC15** correspond to the **(SR)-1A** and **(RS)-1A** equilibrium structures respectively, while structure **IRC8** corresponds to the transition state for the enantiomeric conversion of **1A**. All other structures along the IRC path are between the transition state and the equilibrium structures.

The flipping of the tetrazepinone ring and the indiviual nitrogen inversions can be resolved by monitoring changes in the N3-C4-N2-C13 (ΦN3) and N5-C4-C12-C6 (ΦN5) improper torsion angles through the IRC path. The changes in ΦN3 and ΦN5 are defined as

$$\Delta N3 = | \Phi N3_{\text{at IRC point}} | - | \Phi N3_{\text{at equilibrium}} |$$

$$\Delta N5 = | \Phi N5_{\text{at IRC point}} | - | \Phi N5_{\text{at equilibrium}} |$$

Each nitrogen is considered to invert when its improper torsion angle is 180°, so the individual N3 and N5 inversions appear as peaks in plots of ΔN3 and ΔN5 versus IRC structure number (**Fig. 7.9**). IRC structures **1-6** correspond to the N3 inversion, because the improper torsion angle ΦN5 does not change much, while ΔN3 goes through a maxima. The structures **10-15** correspond the N5 inversion, because the ΦN3 improper torsion angle does not change much, while ΔN5 goes through a maxima. Comparing the plots of ΔN5 and ΔN3 with the energy at each IRC structure shows that the individual nitrogen inversions are shoulders (rather than saddle points) on the energy profile (**Fig. 7.9**). Since the N5 atom is less tetrahedral than the N3 atom, the shoulder for N5 inversion is about 1.5 kcal/mol less than the shoulder for N3 inversion.

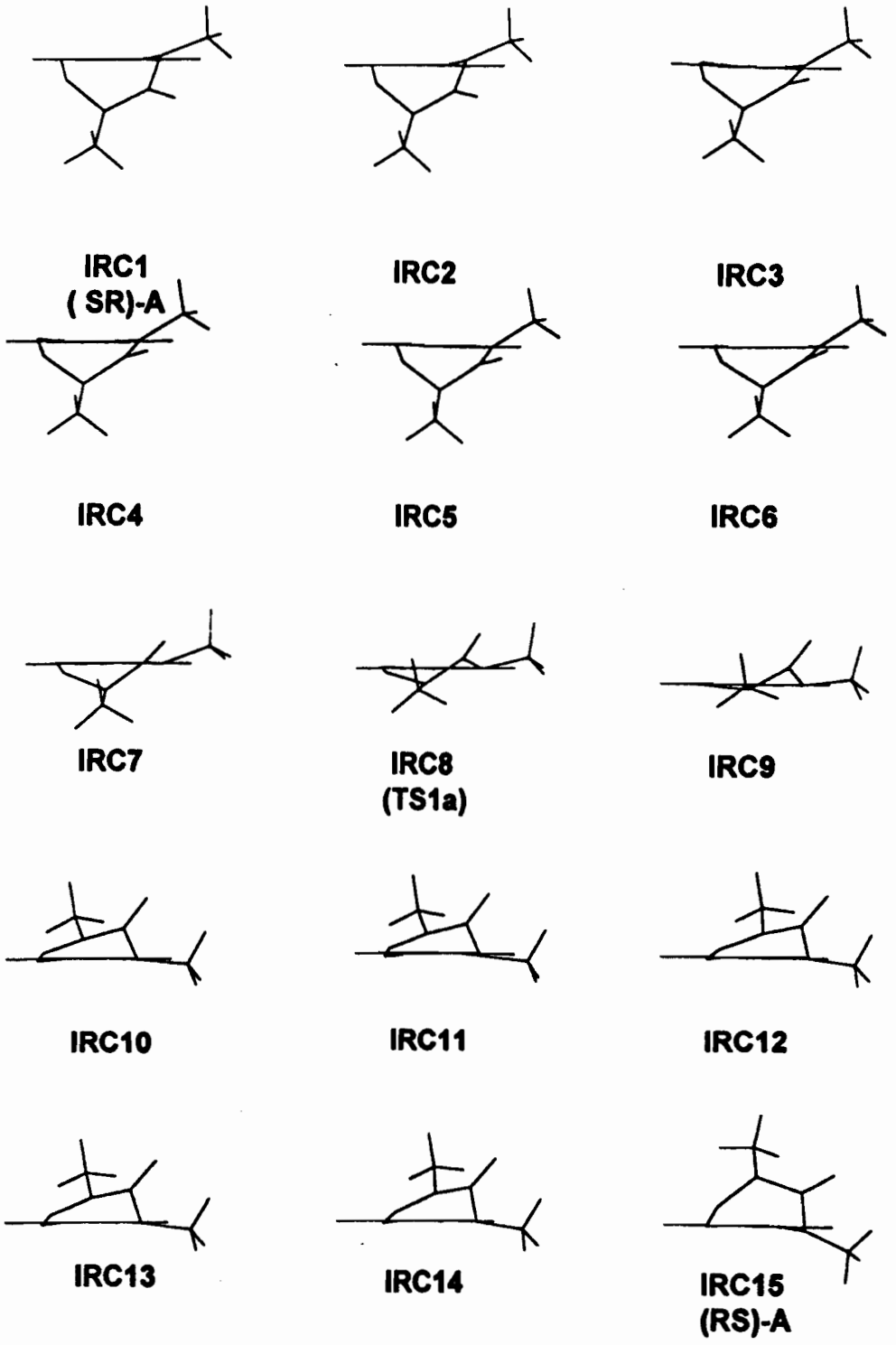


Figure 7.8:

Structures along the IRC Path for Tetrazepinone Enantiomeric Interconversion

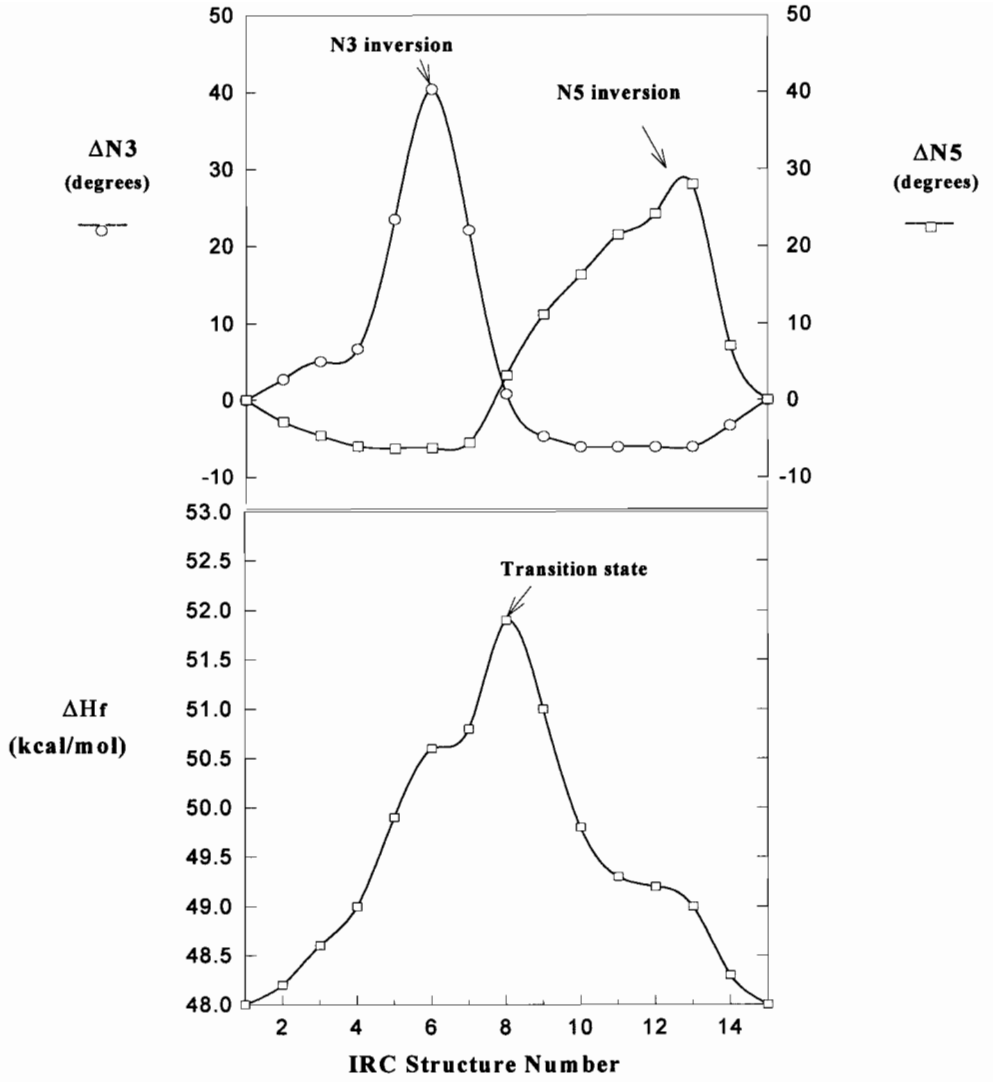


Figure 7.9:

Plots of ΔN3, ΔN5 and the PM3 Heat of Formation at Different Structures along the Racemization Intrinsic Reaction Coordinate (IRC) Path.

The maxima in the **Fig. 7.9** energy plot (IRC structures **7-9**) represents the ring flipping step. The N3 and N5 improper torsion angles do not change

drastically in these structures, but unlike the IRC structures **1-6** and **10-15**, the N3 and C4 atoms are on opposite sides of the benzene ring plane in IRC structures **7-9**. Interestingly, at structure **IRC8** (the racemization transition state) the absolute values of the improper torsion angles ΦN3 and ΦN5 are close to their equilibrium values. Thus, although racemization is concerted, the necessary steps for the interconversion (ie the nitrogen inversions and the tetrazepinone ring flip) appear to occur sequentially.

7.4.14. Effect of R1 and R2 Substituents on Tetrazepinone Racemization Energy Barriers:. Benzotetrazepinone racemization energy barriers with different N3 and N5 substituents are compared in **Table 7.2.** With methyl and -H groups as substituents on N3 and N5, the AM1, MNDO and PM3 energy barriers to benzotetrazepinone racemization are quite small (0.1-4 kcal/mol), as are the barriers to benzoalkyltetrazepinone (**1d**) racemization (4-8 kcal/mol). Barriers this small would make it impossible to resolve the enantiomeric forms of the tetrazepinones at room temperature. However, energy barriers to racemization with *tert*-butyl groups as N3 and N5 substituents (14-16 kcal/mol) may be large enough to allow enantiomeric resolution.

Table 7.2:
Effect of N3 and N5 Substituents on Racemization Energy Barriers

		Energy Barrier (kcal/mol)		
N3 substituents	N5	MNDO	AM1	PM3
H	H	0.1	0.1	2.9
CH3	CH3	8.7	4.2	4.0
CH3 (cycloalkyl)		8.3	4.2	5.0
t-But	t-But	16.3	15.0	14.4

7.5 CONCLUSIONS:

The out-of-plane X-ray geometry of the benzotetrazepinone ring system has been successfully predicted by *ab initio* calculations. PM3 is the best semi-empirical method for these molecules, giving optimized geometries almost as good as *ab initio*, and successfully reproducing most of the important tetrazepinone structural features. The MNDO and AM1 methods failed to accurately reproduce several important structural features, although they did yield out-of-plane benzotetrazepinone conformations similar to the X-ray structures. The MM+ method was the least successful at reproducing the X-ray structures, but optimized MM+ structures were adequate as input for more advanced calculations.

All semi-empirical and *ab initio* methods predict that the planar form of the tetrazepinone ring is not a valid structure (intermediate or transition state) because vibrational analyses of planar tetrazepinone rings return Hessians with more than one imaginary vibrational frequency. All semi-empirical methods predict that benzotetrazepinone racemization occurs through a concerted mechanism involving a non-planar transition state. However, the low energy barrier to racemization of tetrazepinones with methyl substituents on N3 and N5 would make it impossible to resolve the enantiomers with these substituents. Semi-empirical methods all predict that bulky susbstituents on N3 an N5 will hinder the enantiomeric interconversion.

REFERENECES:

1. a) Jean-Claude, B.J. and Just, G., *J. Chem. Perkin Trans.* **1**, 2525 (1991).

   b) B.J. Jean-Claude, et. al. submitted to *Heterocycles.*

2. A.R. Katritzky, K. Yannakopoulou, E. Anders, J. Stevens, and M. Szfran, *J. Org. Chem.*, **55**, 5683 (1990).

3. B,D Waldkowski, R.H. Smith, and C.J. Michejda, *J. Am. Chem. Soc.,* **113**, 7893 (1991).

4. K.Nielsen, I. Sotofte, and H. Johansen, *Acta Chemica Scandinavica*, **47**, 943, (1993).

5. D. Danovich, V. Zakrzewski, and E. Domnina, *J. Mol. Struct., (THEOCHEM)*, **187**, 297 (1989).

6. M. de la Concepcion Forces-Forces, F. Hernandez Cano, R.M. Claramunt, D. Sanz, J. Catalan, F. Fabero, A. Fruchier, and J. Elguero, *J. Chem. Soc. Perkin Trans. 2*, 237, (1990).

7. M. Sana, G. Leroy, M.T. Nguyen and J. Elguero, *Nouv. J. Chem.*, **3**, 607 (1979).

8. P.R. Lowe, C.E. Sansom, C.H. Schwalbe, M.F.G. Stevens and A.S. Clarke *J. Med. Chem.*, **35**, 3377 (1992).

9. A.S. Clarke, M.F.G. Stevens, C.E. Sansom and C.H. Schwalbe, *Anti-Cancer Drug Design*, **5**, 63 (1990).

10. M. Meyer, *J. Mol. Struct., (THEOCHEM),* **304**, 45, (1994).

11. W. M. F. Fabian, *Z. Naturforsch*, **45a**, 1328 (1990).

12 J. P. Ritchie, *J. Org. Chem.*, **54**, 3553, (1989).

13. A. K. Mishra and S. Dogra, *Indian J. Chem.*, **27a**, 851 (1988).

14. J. Mavri and D. Hadzi, *J. Mol. Struct.*, **224**, 285 (1990).

15. B. Vernon Cheney, *J. Org. Chem.*, **59**, 773 (1994).

16. A. L. Llamas-Saiz, C. Foces-Foces, O. Mo, M. Yanez, E. Elguero and J. Elguero, *J. Comput. Chem.*, **16**, 263, (1995).

17. B.S. Jursic, and Z. Zdravkovski, *J. Mol. Struct. (THEOCHEM),* **309**, 241 (1994).

18. E. Fos, J.Vilarrasa and J. Fernandez, *J. Org. Chem.*, **50**, 4894 (1985).

19. M. Cabre, J. Farras, J. Fernandez-Sanz and J. Vilarrasa, *J. Chem. Soc. Perkin Trans. 2*, 1943 (1990).

20. B. S. Jursic and Z. Zdravkovski, *J. Mol. Struct. (THEOCHEM)*, **331**, 229 (1995).

21. E. Pop, M.-J. Huang, M.E. Brewster and N. Bodor, *J. Mol. Struct (THEOCHEM)*, **315**, 1 (1994).

22 J. Barluenga, T.L. Sordo, J.A. Sordo, S. Fustero and J. Gonzalez, *J. Mol. Struct. (THEOCHEM)*, **315**, 63 (1994).

23. J. Frau, J. Donoso, B. Vilanova, F. Munoz and F. Garcia Blanco, *Theor. Chim. Acta.,* **86**, 229 (1993).

24. B. S. Jursic and Z. Zdravkovski, *J. Mol. Struct. (THEOCHEM),* **315**, 85 (1994).

25. HyperChem™ is available from Hypercube Inc., 419 Philip St., Waterloo, Ont. Canada N2L 3X2.

26. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** (1990) 52.

b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).

Appendix A.7

Table 7.A1:
Computed and X-ray Bond Lengths in Nitro-Substituted Benzotetrazepinone (1c):

	Bond Lengths (Å)						
	6-311G	3-21G	MNDO	AM1	PM3	MM+	X-RAY
N1-N2	1.226	1.223	1.217	1.214	1.217	1.232	1.255
N2-N3	1.419	1.469	1.360	1.357	1.446	1.359	1.446
N3-C4	1.395	1.405	1.430	1.430	1.460	1.421	1.419
N3-C13	1.466	1.475	1.489	1.464	1.488	1.458	1.489
C4-O14	1.218	1.210	1.221	1.245	1.216	1.233	1.228
C4-N5	1.372	1.374	1.436	1.437	1.457	1.417	1.361
N5-C12	1.468	1.476	1.473	1.449	1.482	1.401	1.466
N5-C6	1.412	1.404	1.425	1.415	1.451	1.420	1.407
C6-C7	1.389	1.388	1.425	1.427	1.409	1.408	1.399
C8-C11	1.386	1.387	1.415	1.410	1.398	1.409	1.383
C7-N1	1.418	1.418	1.429	1.427	1.445	1.410	1.409
C7-C8	1.388	1.386	1.413	1.408	1.397	1.406	1.384
C11-C10	1.378	1.371	1.409	1.399	1.398	1.409	1.389
C8-C9	1.378	1.373	1.406	1.388	1.389	1.404	1.375
C9-C10	1.383	1.378	1.409	1.402	1.399	1.407	1.374
N15-C10	1.451	1.451	1.505	1.490	1.503	1.403	1.437
O16-N15	1.225	1.242	1.209	1.201	1.214	1.277	1.227
O17-N15	1.227	1.244	1.209	1.201	1.215	1.277	1.205

Table 7.A1 (cont):
Computed and X-ray Improper Torsion Angles in Nitro-Substituted Benzotetrazepinone(1c):

	Improper Torsion Angles (degrees)						
	6-311G	3-21G	MNDO	AM1	PM3	MM+	X-RAY
N5C4C12C6	163.2	166.9	156.1	145.6	143.0	171.8	157.3
N3C4N2C13	-138.8	-131.5	-156.6	-158.4	-136.9	-166.7	-128.5

Table 7.A1 (cont):
Computed and X-ray Bond Angles in Nitro-Substituted Benzotetrazepinone (1c):

Bond Angles (degrees)							
N1-N2-N3	120.5	118.1	125.6	129.0	121.9	128.9	117.3
N2-N3-C4	120.1	116.3	125.1	125.6	117.3	122.2	114.5
N3-C4-N5	116.4	114.7	115.2	119.9	118.0	121.7	116.2
C4-N5-C6	121.7	122.6	116.3	119.2	116.3	119.2	121.5
C7-N1-N2	124.4	124.4	128.6	130.0	126.6	128.3	122.5
C8-C7-N1	115.0	115.1	117.0	116.6	116.8	117.3	115.3
N5-C6-C11	120.1	119.8	123.5	120.8	122.0	120.2	119.9
N2-N3-C13	109.7	108.1	111.9	116.6	111.4	116.6	107.9
N3-C4-O14	121.5	122.5	122.1	119.6	120.9	119.6	120.5
C6-N5-C12	119.3	119.5	120.4	117.1	116.7	119.7	119.0
C7-C8-C9	120.7	120.9	120.0	121.2	119.9	120.7	121.5
C6-C11-C10	119.4	120.0	118.3	120.1	120.0	121.6	117.7
C11-C10-N15	118.5	118.3	118.6	119.2	119.7	119.8	117.9
C10-N15-O17	118.1	117.2	119.0	118.8	119.3	118.4	117.1
C10-N15-O16	118.2	117.5	119.0	118.9	119.2	119.3	119.0

Table 7.A1 (cont):
Computed and X-ray Torsion Angles inNitro-Substituted Benzotetrazepinone (1c):

Torsion Angles (degrees)							
	6-311G	3-21G	MNDO	AM1	PM3	MM+	X-RAY
N1N2N3C4	66.8	73.2	45.2	33.8	63.9	40.5	73.6
N2N3C4N5	-49.0	-58.4	-5.9	-2.4	-34.7	-11.1	-59.7
N3C4N5C6	-25.8	-18.1	-63.5	-53.6	-42.0	-49.0	-19.7
C4N5C6C7	52.6	50.3	64.9	56.9	61.8	55.1	50.0
N5C6C7N1	0.0	0.2	0.0	5.5	1.1	0.1	4.8
C6C7N1N2	-39.9	-39.5	-36.0	-27.1	-41.9	-35.5	-46.2
C7N1N2N3	-1.2	-2.7	-4.7	-2.3	-4.2	-0.0	-0.9
N1N2N3C13	-154.3	-155.2	-158.1	-167.8	-159.1	-153.3	-157.9
C12N5C6C11	32.3	32.6	40.5	16.3	21.6	44.0	20.8
O14C4N3C13	-4.9	-5.5	-18.5	-14.3	-9.9	-0.3	-2.8
O14C4N5C12	9.0	5.7	40.1	13.9	4.6	38.1	1.8
N3C4N5C12	171.3	175.4	139.5	161.5	174.9	139.1	176.4
C8C7N1N2	143.7	144.2	146.8	154	140.8	117.4	141.4
C9C8C7N1	177.2	178.1	180.0	177.9	179.2	179.1	174
C10C11C6N5	-177.4	-176.9	-177.2	-173.3	-176.2	-178.1	-177.9
N1C7C6C11	-177.4	-176.0	-178.9	-178.7	-178.1	-179.0	-170.1
N15C10C9C8	-179.6	-179.5	-179.5	-179.4	-179.9	-179.7	-179.2
O16N15C10C9	-0.9	-0.6	-92.3	-1.2	-6.7	-2.7	-5.6
O17N15C10C9	179.1	179.4	88.1	178.8	173.2	178.0	174.4

Table 7.A2:
Comparison of Optimized Benzotetrazepinone and Lone Tetrazepinone Ring geometries:

	BOND LENGTHS (Å)							
	MNDO		AM1		PM3		3-21G	
	LONE	BENZO	LONE	BENZO	LONE	BENZO	LONE	BENZO
N1-N2	1.219	1.216	1.214	1.214	1.217	1.216	1.223	1.223
N1-C7	1.419	1.429	1.405	1.427	1.427	1.445	1.410	1.417
N2-N3	1.379	1.359	1.366	1.356	1.454	1.445	1.481	1.469
N3-C4	1.449	1.429	1.437	1.430	1.465	1.459	1.395	1.404
N3-C13	1.485	1.489	1.465	1.463	1.487	1.488	1.471	1.475
C4-N5	1.420	1.436	1.416	1.436	1.446	1.457	1.375	1.374
C4-O14	1.223	1.221	1.249	1.245	1.217	1.215	1.212	1.209
C7-C6	1.358	1.425	1.356	1.426	1.345	1.408	1.320	1.388
N5-C6	1.411	1.424	1.384	1.414	1.429	1.450	1.390	1.403
N5-C12	1.478	1.473	1.449	1.449	1.484	1.481	1.471	1.472
BOND ANGLES (DEGREES)								
N2-N1-C7	126.5	128.6	129.7	129.9	127.1	126.6	124.7	124.3
N1-N2-N3	120.5	125.6	128.9	129.0	120.7	121.8	117.3	118.0
N2-N3-C4	116.3	125.0	126.0	125.6	116.3	117.2	118.3	116.2
N2-N3-C13	111.7	111.9	114.9	116.6	110.4	111.3	108.3	108.0
C4-N3-C13	117.2	118.8	112.5	114.4	115.2	115.9	115.1	115.3
N3-C4-N5	115.0	115.1	123.0	119.8	118.9	118.0	114.9	114.7
N3-C4-O14	123.1	122.1	119.0	119.6	121.1	120.9	122.4	122.4
N5-C4-O14	121.8	122.7	117.7	120.3	119.9	121.0	122.5	122.8
N1-C7-C6	127.3	122.5	128.1	123.7	125.5	122.7	125.1	123.9
C4-N5-C6	123.4	116.3	126.7	119.2	120.7	116.4	123.3	122.6
C4-N5-C12	118.6	119.0	114.9	113.8	116.7	116.0	117.7	116.4
C7-C6-N5	126.1	116.7	128.2	120.3	124.3	118.7	126.4	121.7
TORSION ANGLES (DEGREES)								
C7N1N2N3	-4.2	-4.7	-4.6	-2.3	-5.5	-4.2	-6.4	-2.7
N2N1C7C6	-29.5	-36.0	-12.2	-27.1	-37.0	-41.8	-37.3	-39.9
N1N2N3C4	68.5	45.2	34.0	33.8	66.1	63.9	72.6	73.2
C13N3C4N5	150.7	161.1	174.9	161.2	176.6	169.7	171.3	173.4
C13N3C4O14	-27.5	-18.4	-1.9	-14.2	-2.1	-9.8	-5.5	-5.4
N3C4N5C12	162.4	139.4	168.7	161.5	163.5	174.8	179.8	175.3
O14C4N5C12	15.8	40.1	8.2	13.9	15.3	4.6	2.9	5.7
N1C7C6N5	0.9	8.1	0.6	5.5	0.3	1.1	5.7	0.2
C4N5C6C7	22.5	64.9	10.2	56.9	44.4	61.8	39.8	50.3
IMPROPER TORSION ANGLES								
N5C4C12C6	177.2	156.1	176.8	145.6	149.1	142.9	171.6	166.9
N3N2C4C13	138.5	156.6	149.6	158.4	134.4	136.9	131.6	131.5

Table 7.A3: Geometries of MNDO, AM1 and PM3 Transition States for Enantiomeric Interconversion (IRC8):

	MNDO	AM1	PM3
Bond lengths (Å)			
C13-N3	1.484	1.465	1.489
N1 -N2	1.215	1.208	1.215
N1 -C7	1.426	1.421	1.439
N2 -N3	1.371	1.357	1.419
N3 -C4	1.450	1.434	1.461
C7 -C8	1.425	1.414	1.407
C7 -C6	1.429	1.425	1.408
N5 -C6	1.420	1.406	1.442
N5 -C12	1.478	1.452	1.486
N5 -C4	1.423	1.412	1.450
C8 -C9	1.396	1.385	1.384
C6 -C11	1.429	1.422	1.405
C9 -C10	1.407	1.392	1.391
C4 -O14	1.221	1.250	1.217
BOND ANGLES (DEGREES)			
N2 -N1 -C7	129.6	132.4	132.6
N1 -N2 -N3	121.4	129.4	126.1
C13-N3 -N2	112.5	115.5	111.4
N2 -N3 -C4	114.8	125.0	118.8
N1 -C7 -C6	128.5	126.8	129.2
C8 -C7 -C6	120.2	119.8	119.9
C12-N5 -C4	116.2	114.5	113.6
C7 -C8 -C9	121.0	121.5	120.9
C7 -C6 -N5	123.3	124.8	123.9
C6 -C11-C10	121.7	121.9	121.5
C9 -C10-C11	120.6	120.6	120.2
N3 -C4 -N5	114.4	123.1	116.7
N3 -C4 -O14	123.2	117.8	121.3
TORSION ANGLES (DEGREES)			
C7 -N1 -N2 -N3	-0.8	-2.0	-3.6
N2 -N1 -C7 -C6	-21.2	-7.9	-5.6
N1 -N2 -N3 -C13	161.9	179.7	176.9
N1 -N2 -N3 -C4	60.2	30.3	39.3
N2 -N3 -C4 -N5	-84.1	-41.6	-70.5
C6 -C7 -C8 -C9	1.9	0.9	0.3
N1 -C7 -C6 -N5	3.9	2.2	3.6
C8 -C7 -C6 -C11	-2.1	-0.9	-0.6
C12-N5 -C6 -C7	-179.3	-178.9	-171.7
C12-N5 -C4 -O14	42.7	17.4	19.9
C7 -C8 -C9 -C10	-0.5	-0.3	-0.2

In chapter seven, the PM3 method was shown to be the most reliable semi-empirical method for benzotetrazepinone structure prediction. The results in chapters three to six show that the PM3 is also the most reliable semi-empirical method for nitrogen heterocyclic energies. Therefore, benzotetrazepinone acid-catalyzed decomposition is studied with PM3 exclusively in chapter eight

CHAPTER 8:

A SEMI-EMPIRICAL PM3 TREATMENT OF BENZOTETRAZEPINONE DECOMPOSITION IN ACID MEDIA

by
C.I. Williams, M.A. Whitehead*

Theoretical Chemistry Laboratory, Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6
and

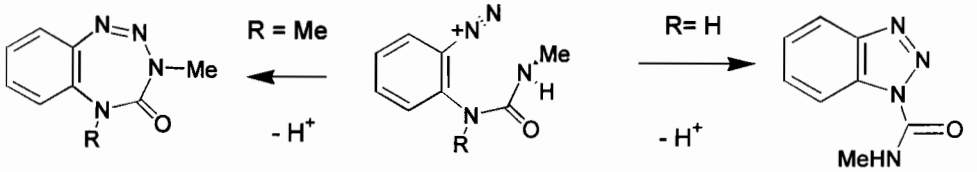
Bertrand J. Jean-Claude
Department of Oncology, McGill University, 3655 Drummond St., Ste. 701, Montreal, Quebec, Canada H3G 1Y6

8.1 Abstract:

A mechanism for the acid-induced ring opening of benzotetrazepinone (BTZ) is proposed, and investigated with the semi-empirical PM3 method. The PM3 results predict that the ring opening is pH dependant, and influenced by the aryl ring substituent. Yields of benzotetrazepinones produced from the internal cyclization of N-(2-diazoniumphenyl)-N-(methyl)-N'-methyl ureas (DPMU) are correlated with the computational results. The PM3 energies relate the tendency for DPMU to undergo ring closure, to the Hammett constants of the aryl-ring substituents. Linear free energy reaction constants of ρ=3.36 and ρ^+ = 2.25 are calculated for the ring closure reaction of DPMU. The pK_a of protonated benzotetrazepinone is estimated to be -1.0, which is similar to the experimental pKa's of structurally analogous benzamides and phenylureas. The cyclization of some DPMUs to N-carbamoyl benzotriazoles (NCBT) is also explained.

8.2 Introduction:

The benzotetrazepinones **1** (BTZ) were synthesized by Jean-Claude et. al.[1a,b] by internal cyclization of N-(2-diazoniumphenyl)-N-(methyl)-N'-methyl ureas (DPMU)[1] **2** with R = -Me (**Fig. 8.1**). Benzotetrazepinones represent a new class of compounds which contain the novel seven-membered tetrazepinone ring system. A number of aryl ring substituted BTZs and pyridotetrazepinones (PTZs) have been made (**Fig. 8.2**). Attempts to synthesize BTZ **1f** with R = -H fail, because DPMU **2** with R = -H cyclizes to N-carbamoyl benzotriazoles NCBT **3** (**Fig. 8.1**). No attempts have been made to synthesize the non-benzofused (lone) tetrazepinone rings **1i-j**.



1 **2** **3**

BTZ **DPMU** **NCBT**

Benzotetrazepinone N-(2-diazophenyl)-N-(methyl)-N'-methyl urea N-carbamoyl-benzotriazole

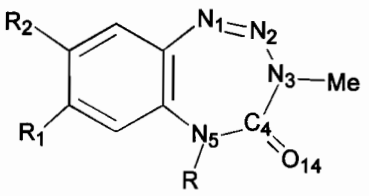
Figure 8.1: Tetrazepinone Synthesis
(only DPMUs with R = -Me cyclize to tetrazepinones)

BTZ percent yields and stabilities depend on the aryl ring substituents. The tricyclic tetrazepinone **1d**, *p*-nitro-BTZ **1c**, *m*-chloro-BTZ **1h** and PTZ **1e** are obtained in good yield (>50%), and are stable enough for X-ray structural determination. Unsubstituted BTZ **1a** and *p*-methoxy-BTZ **1b** give lower yields (≤ 50%), are not stable enough for X-ray structural determination, and readily decomposing to methyl isocyanate and benzotriazoles **4** (**Fig. 8.3**).

The substituents on N5 also affect BTZ stabilities. Attempts to synthesize BTZs from DPMU's with R sustituents bulkier than CH_3 fail (**1g**), and instead benzotriazoles and methyl isocyanate are obtained as products in these reactions. Nitro-BTZ with R = Me does not decompose to benzotriazoles and methyl

isocyanate, but *p*-nitro-BTZs with more bulky R substituents do. PTZs are not known to decompose by losing methyl isocyanate.

Recently Synthesized Tetrazepinones



1

1 BTZ: Benzotetrazepinones

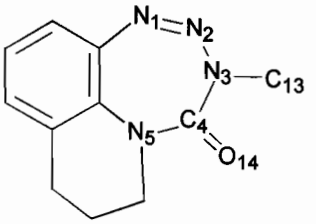
R3 = -Me

a R1 = H, R2 = H

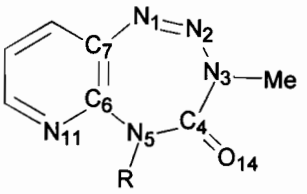
b R1 = OCH3, R2 = H

c R1 = NO2, R2 = H

h R1 = H, R2 = Cl



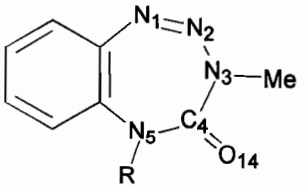
1d Tricyclic Benzotetrazepinones



1e PTZ: Pyridotetrazepinone

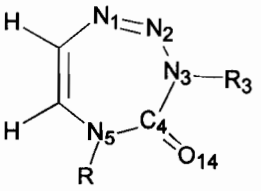
Hypothetical Predicted Tetrazepinones

(not yet experimentally observed)



1f R = H

1g R = cyclopentyl -(C_5H_9)

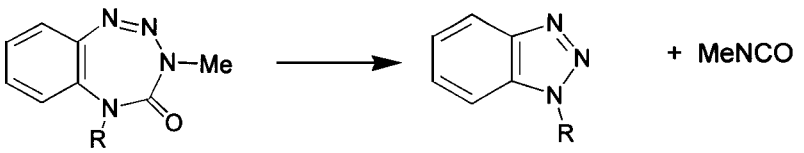


Non-benzofused (lone) tetrazepinone

1i R = -H **1j** R = -Me

Figure 8.2:

Representative Tetrazepinones



1 4

Figure 8.3:

Decomposition of BTZs into

Benzotriazoles (**4**) and Methyl Isocyanate

In acidic media, UV/VIS and ^{15}N NMR data show that tetrazepinones undergo a retro-synthetic ring opening to DPMUs.[1c] UV/VIS spectra show that compound **1c** (R_1 = -NO_2) remains in closed chain tetrazepinone form at relatively low pH, while tetrazepinones **1a** (R_1 = -H) and **1c** (R_1 = -OCH_3) readily ring open to DPMUs at moderate pH. Thus, electron-withdrawing groups stabilize closure of the tetrazepinone ring.

X-ray structural determinations of unprotonated tetrazepinones **1c, 1d** and **1e** show the tetrazepinone atoms in a non-planar comformation that significantly deviates from the aromatic ring plane. X-ray and computed tetrazepinone ring bond lengths show that the bonds are localized. The tetrazepinone ring is clearly neither aromatic nor anti-aromatic. X-ray and theoretical optimized geometries of BTZs were discussed completely in chapter 7.[1d]

The results in chapters three to six all suggest that the PM3 method is superior to the MNDO and AM1 methods for nitrogen-rich heterocyclic energetics. PM3 was shown to be the most reliably semi-empirical method for reproducing benzotetrazepinone experimental geometries in chapter seven. Therefore, the PM3 semi-empirical method was used exclusively to study (a); the effect of aryl ring substituents on the acid-induced BTZ ring opening to DPMU, (b); the formation of BTZs from DPMUs, (c); the effect of aryl ring substituents on BTZ synthetic yields, and (d); the formation of NCBT from DPMUs with R = -H.

The size and number of molecules to be studied precluded using *ab initio* methods for all the reactants, products, intermediates and transition states. *Ab initio* calculations (3-21G*, 6-31G**) were performed to compare with some of the PM3 results in some cases, but the PM3 method will save computer time and still yield useful qualitative and quantitative information. [2-24]

8.3 Calculations:

All semi-empirical calculations were performed using the GAMESS package[25] on a Sun Sparc10 workstation. Full geometry optimization was allowed in each calculation, and all optimizations were carried out in Cartesian as opposed to internal coordinates. All quantum mechanical calculations used RHF wavefunctions. Transition states and equilibrium geometries were verified as such by force constant analysis.

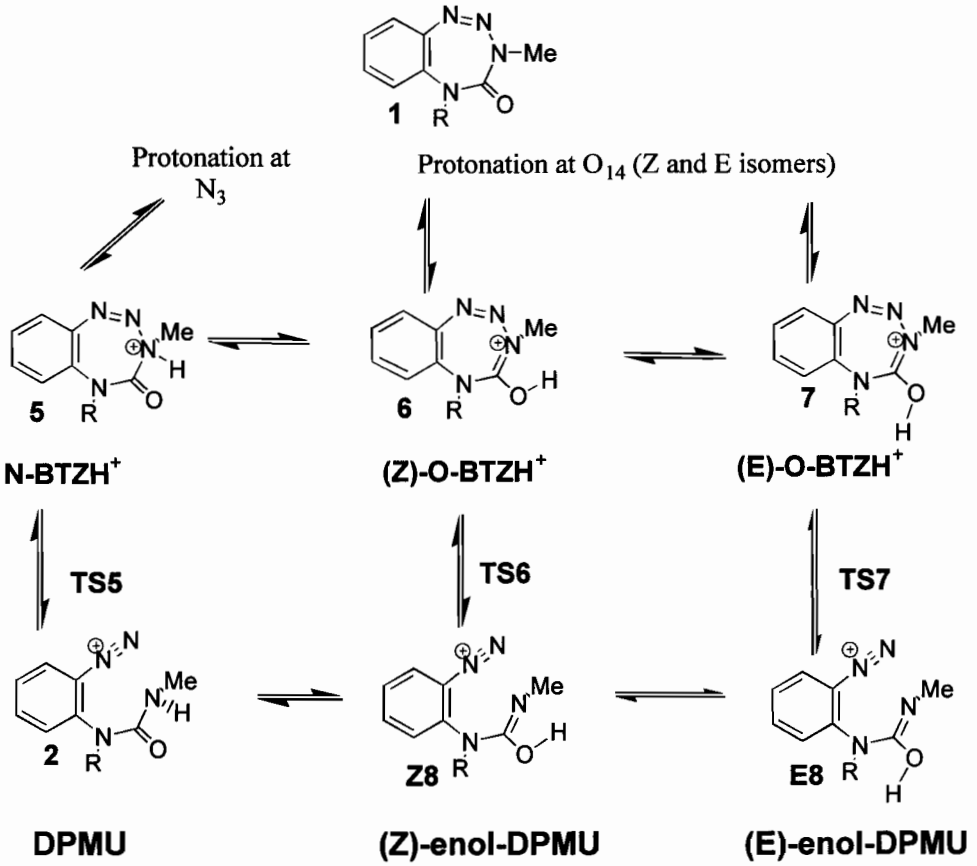
8.4 Results and Discussion:

8.4.1. Acid-Induced BTZ Ring Opening to DPMU:

Protonation of BTZs: Protonation of the tetrazepinone ring could occur at N1, N2, N3, O14 or N5; however, protonations at N3 or O14 were considered exclusively, because only these would result in retro-synthetic ring opening at N2-N3. Protonation of O14 is probably favoured, because amides and ureas are known to protonate at oxygen.[26a] However, protonation of N3 is reasonable, because X-ray and theoretical benzotetrazepinone structures indicate that the N3 atom is almost tetrahedral, and therefore, the lone pair is available for protonation. BTZs precipitate out of DPMU/BTZ reaction mixtures only upon addition of base, so BTZ protonation is reversible and pH dependent, regardless of protonation site.

Three paths for BTZ ring opening were considered (**Scheme 1, Fig. 8.4**); (a), protonation at N3 to form N-BTZH$^+$ **5** which ring-opens through transition state **TS5** to the keto form of DPMU **2**; (b), protonation at O14 with a hydrogen (Z) to N3, forming (Z)-O-BTZH$^+$ **6** which ring-opens through O14 transition state **TS6**

to the (Z)-enol form of DPMU **Z8,** and (c); protonation at O14 with a hydrogen (E) to N3, forming (E)-O-BTZH$^+$ **7** which ring-opens through transition state **TS7** to the (E)-enol form of DPMU **E8**. These pathways are not necessarily independent, and tautomeric equilibra could exist between species in different pathways, as indicated in **Fig. 8.4**.



Scheme 1

Figure 8.4:

Possible Pathways for BTZ Ring Opening

8.4.2. Optimized Geometries of Scheme 1 Structures: Representative PM3 optimized geometries of the N-BTZH$^+$ **5** and O-BTZH$^+$ **6** intermediates shows that N3 protonation results in a substantial lengthening of the N2-N3 bond, and large conformational changes in the tetrazepinone ring. O14 protonation at (Z) or (E) to N3 results in only small extensions of the N2-N3 bond length, and minor changes

in the tetrazepinone ring conformation (**Fig. 8.5**). Structures **7**, **TS7** and **E8** will not be considered in detail, because PM3 energies show that O14 protonation (Z) to N3 is energetically favoured over O14 protonation (E) to N3.

Overlays of PM3 optimized N-BTZH+ **5** and (Z)-O-BTZH+ **6** structures with the PM3 optimized structures of unprotonated BTZ **1** (**Fig. 8.6**) clearly show the O protonated species has a geometry similar to the unprotonated BTZ, while the N3 protonated species has a geometry significantly different from the neutral BTZ. Conformational relocation of the N3 methyl group and lengthening of the N3-N2 bond are the structural features most affected by N3 protonation. The large BTZ structural changes that occur upon N3 protonation, and the higher PM3 energy of N-BTZH$^+$ as compared with O-BTZH$^+$ all suggest that benzotetrazepinones initially protonate at O14; this correlates with experimental evidence that shows amides and ureas protonate at O and not at N.[26a]

Prelimilary 3-21G* and 6-31G** *ab initio* calculations show that the N3 protonated species is unstable, because *ab initio* geometry optimizations of N3 protonated BTZs result in extensions of the N3-N2 bond to > 3 Å. These results parallel the *ab initio* results of Micheja et al.,[26-30] who found that O protonation is favoured over N3 protonation in formyl triazines. *Ab initio* calculations show that N3-protonated (Z)-formyl triazines are unstable; geometry optimization results in the N2-N3 bond lengthening to > 3 Å.[29-30] Protonation of N3 in 1,2,3-triazines is energetically feasible, and is also accompanied by a lengthening of the N3-N2 bond. These result are analogous to our preliminary *ab initio* results for N3 protonated benzotetrazepinones, and are consistent with the elongated N2-N3 bond predicted by PM3.

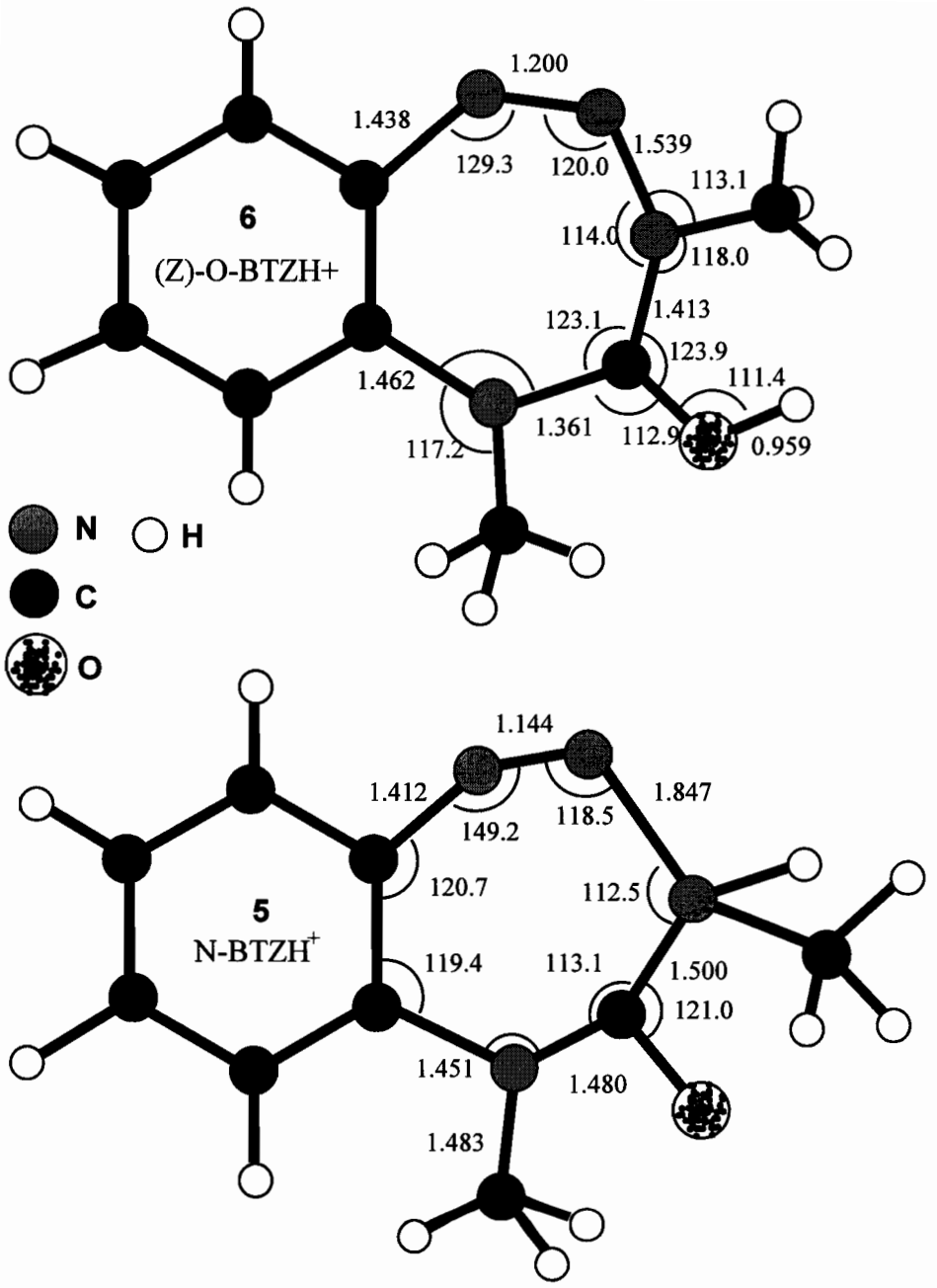
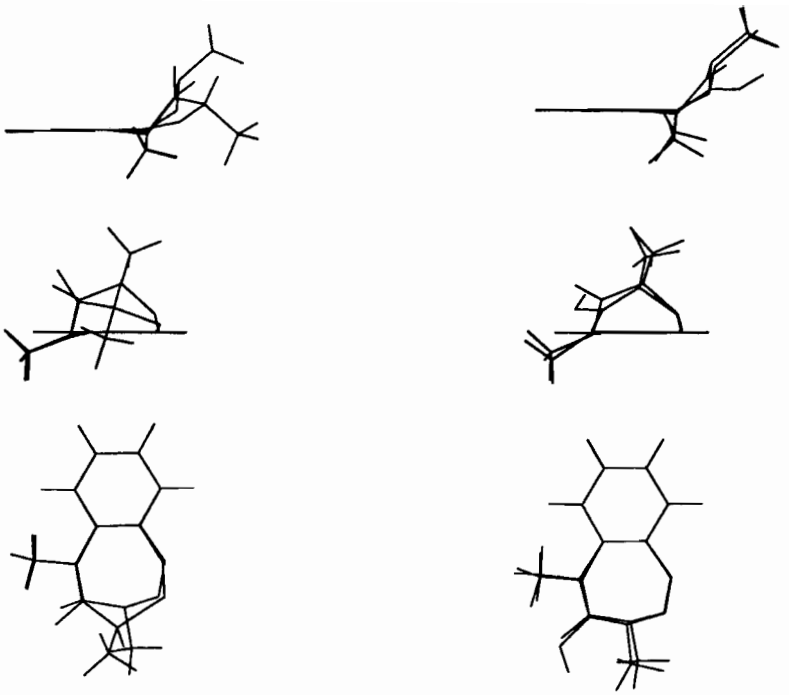


Figure 8.5:
Representative Geometries of N3 and O14 Protonated BTZs
(all bond in lengths Å ; all bond angles in degrees)



N-BTZH+ Z-O-BTZH+

Figure 8.6:

Overlays of PM3 optimized N3 and (Z)-O14 Protonated BTZs with PM3 optimized Unprotonated BTZ

Representative **TS5** and **TS6** transition states for the ring openings of N-BTZH$^+$ and Z-O-BTZH$^+$ are in **Fig. 8.7.** The N2-N3 distances in **TS5** and **TS6** are similar, but overlays of the N3 protonated transition state **TS5** and the neutral BTZ **1** (**Fig. 8.8**) show that the **TS5** transition state conformation differs substantially from that of unprotonated BTZ. The **TS5** more closely resembles the N-BTZH$^+$ **5** protonated intermediate than it does the neutral benzotetrazepinone.

Conformational similarities between the O14 protonated ring opening transition state **TS6** and the unprotonated BTZ are shown in the overlay. The **TS6** transition state also resembles the (Z)-O-BTZH+ protonated intermediate **6.** Thus, O14 protonation followed by ring opening through **TS6** involves smaller changes

in molecular geometries than N3 protonation followed by ring opening of through **TS5**.

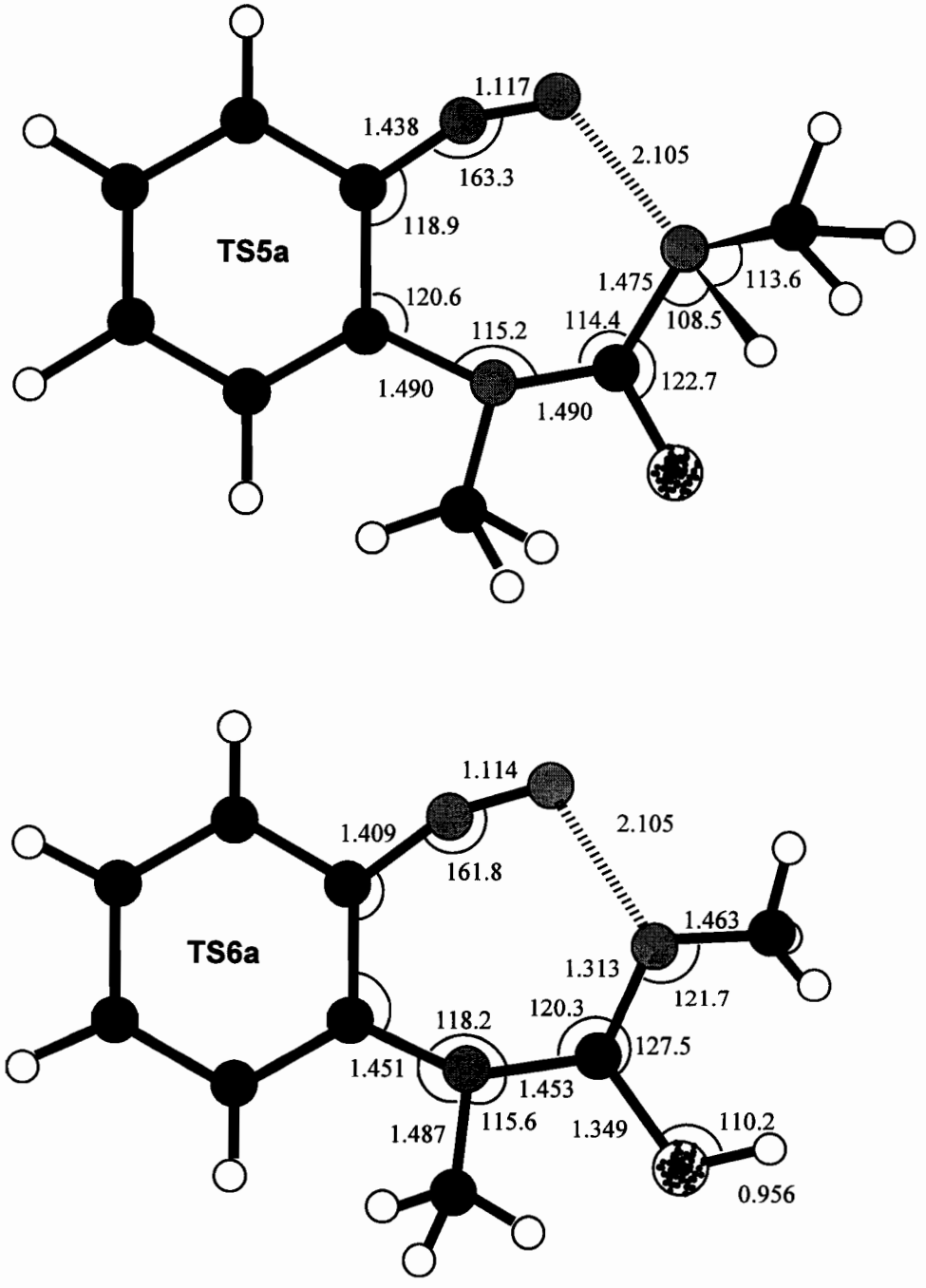
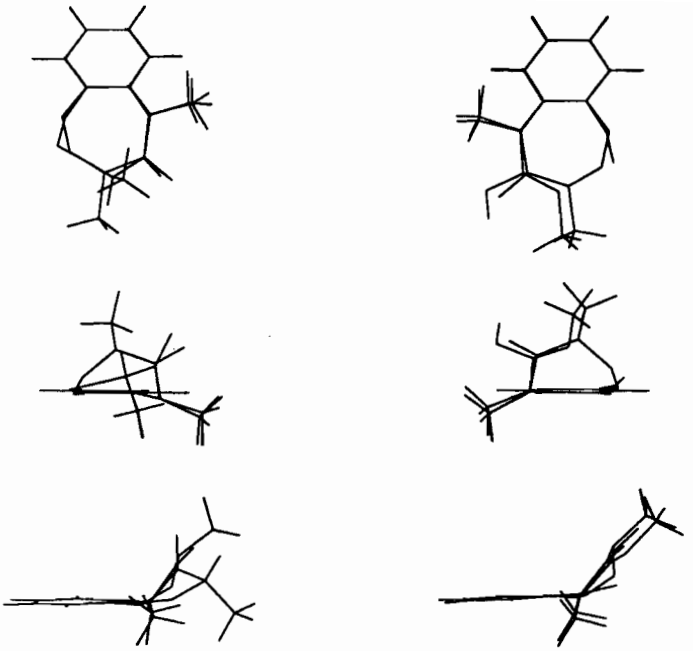


Figure 8.7:
PM3 Transition States of N3 and O14 Tetrazepinone Ring Opening
(all bond lengths in Å; all bond angles in degrees)



TS5 **TS6**

N3 Protonated O14 Protonated

Figure 8.8:

PM3 optimized N3 and (Z)-O14 Protonated BTZs

Chain-Opening Transition State Structures Overlayed with PM3 optimized

Unprotonated BTZ Structures

Representative geometries of keto and enol DPMU (**2** and **Z8** , **Fig. 8.9**) show similarity in the side chain conformations. Both species exhibit a great deal of planarity within the urea group. The plane of the urea group is nearly perpenduicular to the benzene ring (urea plane/benzene plane torsion angle = 82°-88°). As expected, the C4-N3 bond length in the enol form of DPMU corresponds to that of an N=C double bond, while the N3-C4 bond length corresponds to that of an N-C single bond in the keto form of DPMU. The keto-DPMU **2** geometry is similar to the X-ray structure of 2-amino-phenyl urea, the amino synthetic precursor to DPMU prior to diazotization.[33,34]

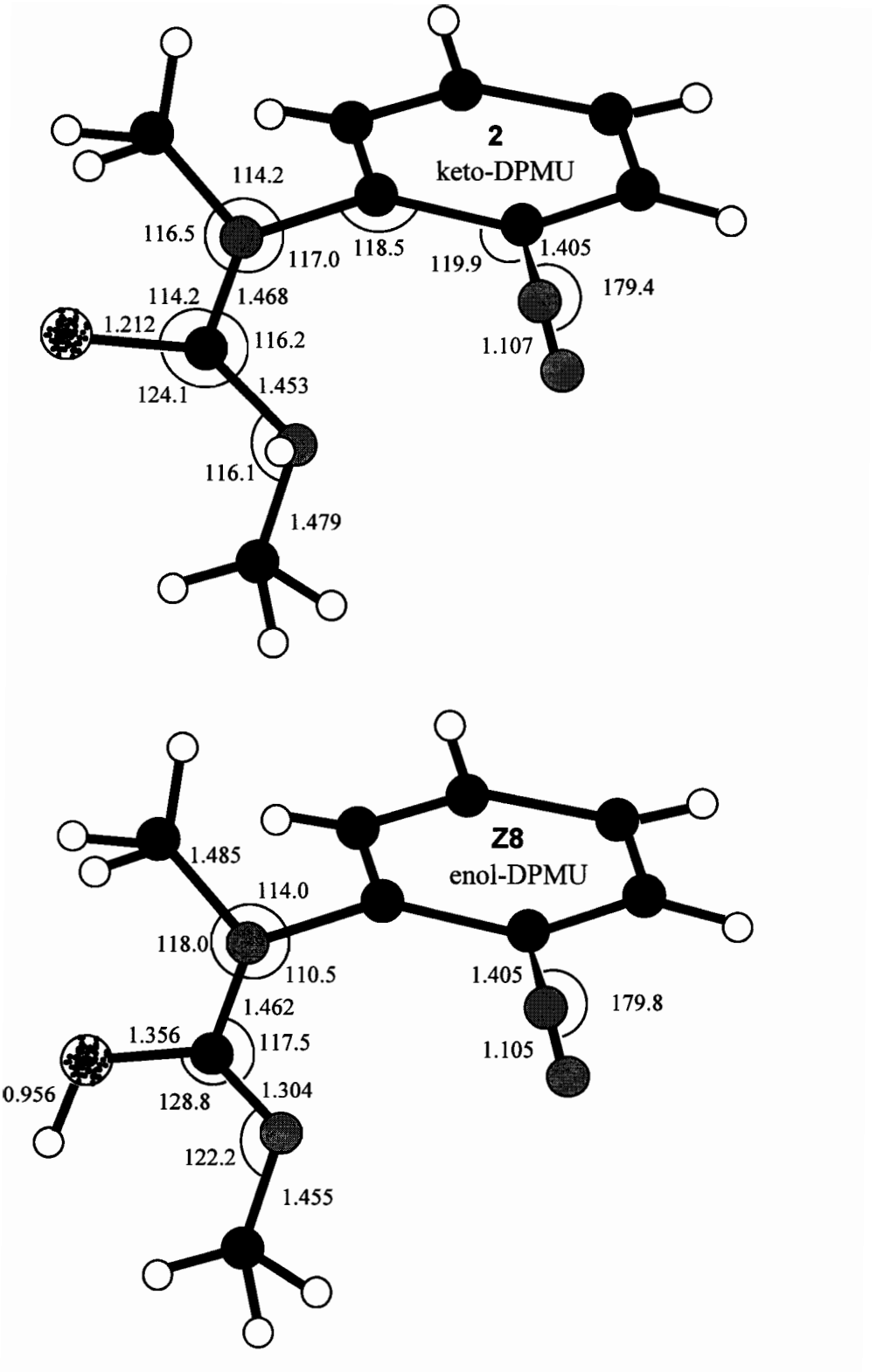


Figure 8.9: Open-Chain DPMUs (Keto and Enol Forms)

(all bond lenghts in Å; all bond angles in degrees)

8.4.3. Energetics of Benzotetrazepinone Ring Opening in Acid Media:

The energetics of the proposed ring opening mechanisms in **Fig. 8.4** were investigated by comparing the PM3 intermediate and transition state heats of formation. Some of the results are summarized in the reaction coordinate in **Fig. 8.10**. The relative energies (ΔH_{rel}) are computed as

(1) $$\Delta H_{rel} = \Delta Hx - \Delta H_f(H^+) - \Delta H_f\,(BTZ).$$

Here ΔHx is the PM3 heat of formation of the compound in question, $\Delta H_f(BTZ)$ is the PM3 heat of formation of the unprotonated parent tetrazepinone and $\Delta H_f(H^+)$ is the experimental H^+ heat of formation (367.2 kcal/mol).

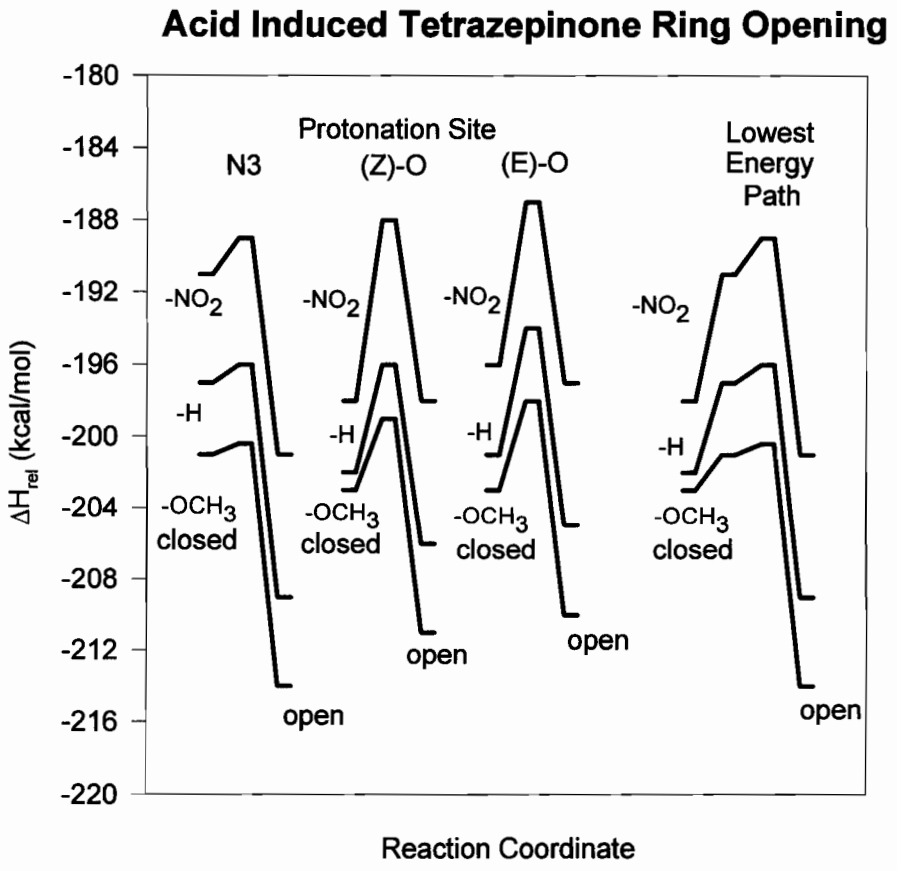


Figure 8.10:

PM3 Reaction Coordinate for Mechanism in Figure 4

The specifics of the protonation step, although important, were not considered explicitly; thus, the energy barriers to protonation are taken as zero in **Fig. 8.10.** The relative energies of the species in **Scheme 1, Fig. 8.4,** are similar; this is consistent with the possibilty of tautomeric equilibra between species as indicated in **Fig. 8.4**. PM3 favours protonation of the tetrazepinone at O14 over protonation at N3, which is expected by analogy with the amide protonations.[26a] The amide form of the open-chain DPMU **2** is favoured over the enol form **Z8**, which agrees with experiments that show the keto form is favoured over the enol form in unprotonated amides.[26a]

Activation barriers to N-BTZH$^+$ ring opening are small (0.5-2 kcal/mol) because severe tetrazepinone ring distortion in N-BTZH$^+$ weakens the N2-N3 bond, and facilitates bond scission. The low ring-opening activation barrier, the extended N2-N3 bond length of PM3 optimized N-BTZH$^+$ and the *ab initio* predictions of an unstable N-BTZH$^+$ all indicate that the existance of an N3 protonated species is questionable; even if an N3 protonated species could form, the previous considerations suggest that it would be unstable and quickly ring-open.

The N2-N3 bond in the O14 protonated (Z)-O-BTZH$^+$ **6** is not as lengthened as it is in N-BTZH$^+$ **5**, so activation barriers to ring opening of O14 protonated intermediates are much higher (4-12 kcal/mol) than those for N-protonated intermediates. PM3 **TS5** transition state heats of formation are 0.2-1 kcal/mol lower than the **TS6** and **TS7** transition state heats of formation.

The similarity in PM3 energies of the species in **Scheme 1**, **Fig. 8.4**, suggest that different energetically similar pathways are available for ring opening. The lowest energy path, outlined in **Fig. 8.11,** begins with (Z)-O-BTZH$^+$, which tautomerizes to the N-BTZH$^+$, and quickly ring opens to DPMU because of the low N-BTZH$^+$ ring opening activation barrier. This path leads directly to the thermodynamically more stable keto form of DPMU. Sample reaction coordinates for the lowest energy path are given in **Fig. 8.10**. Ring opening of the (Z)-O-BTZH$^+$ intermediate to enol-DPMU **Z8** through the **TS6** transition state, followed

by tautomerization to keto-DPMU **2** is a feasible alternate path that is only 0.2-1 kcal/mol higher in energy that the path involving N-BTZH+ **5** and **TS5**.

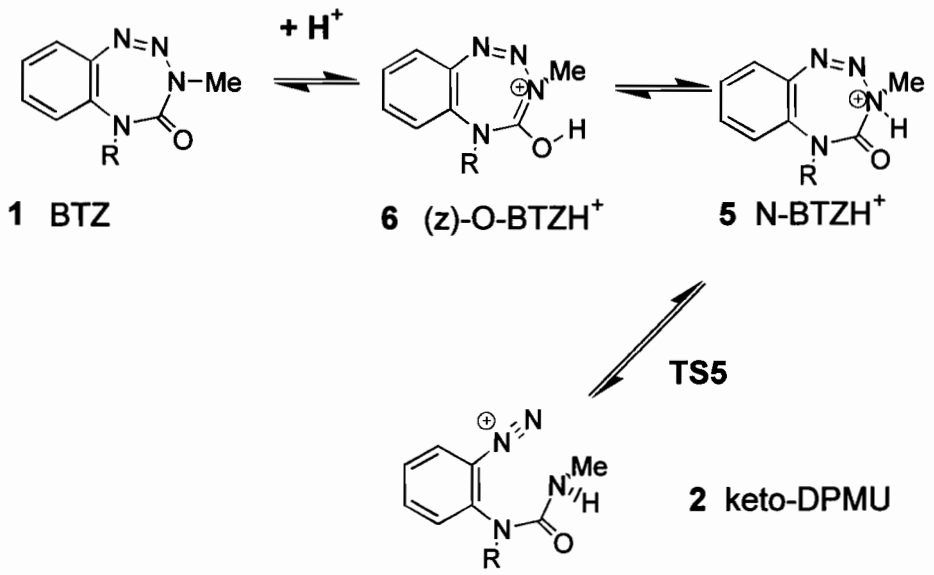


Figure 8.11.
Lowest Energy Path for Acid-Induced BTZ
Ring Opening as Suggested from PM3 Results

8.4.4. pH Dependence of Tetrazepinone Ring Opening:

The high energy of N-BTZH$^+$ **5** and its long N2-N3 bond length suggest that the N-BTZH$^+$ **5** intermediate (if it forms) quickly ring opens to keto-DPMU. Thus, the computational evidence suggests that the (Z)-O-BTZH$^+$ **6** protonated intermediate will be the predominant protonated BTZ species. The keto-form of DPMU **2** will be favoured over the enol-DPMU species. With these considerations, acidic media equilibria can be written in terms of (Z)-O-BTZH$^+$ **6**, the unprotonated neutral BTZ **1**, and the keto-DPMU **2**. The equilibria are given below, along with the corresponding expressions for their equilibrium constants K_a and K_{open} (**Fig. 8.12**).

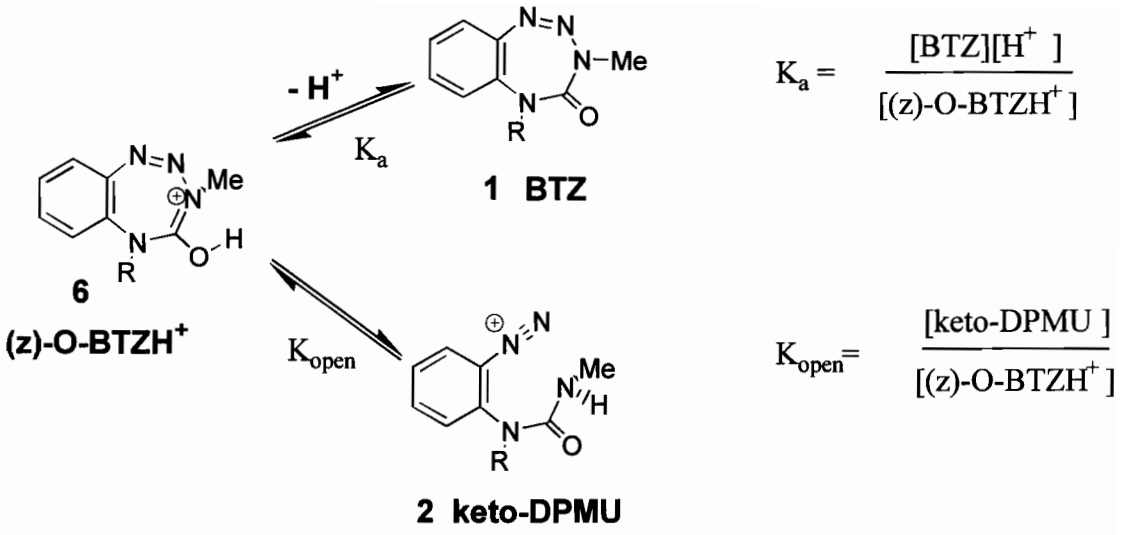


Figure 8.12: Equilibra in Acidic BTZ Solution

Using the equilibrium expressions in **Fig. 8.12**, K_a can be written in terms of K_{open}

$$K_a = \frac{[BTZ][H+]\ K_{open}}{[keto-DPMU]}. \quad (2)$$

Note that (Z)-O-BTZH$^+$ behaves like a weak acid in **Fig. 8.12**, so it will have a corresponding pK_a. Taking the negative $\log_{10}$ of equation (2) yields the following expression for pKa;

$$pK_a = pH - \log\frac{[BTZ]}{[keto-DPMU]} - \log K_{open}. \quad (3)$$

The pKa's of (Z)-O-BTZH$^+$s are unknown; however, if pH and pKa are grouped into $\Delta pH = (pH - pKa)$, equation (3) can be rewritten in a form which relates the [BTZ]/[keto-DPMU] ratio to ΔpH and the equilibrium constant K_{open}.

$$\Delta pH - \log K_{open} = \log[BTZ]/[keto-DPMU] . \quad (4)$$

Values for K_{open} may be approximated using PM3 (Z)-O-BTZH$^+$ **6** and keto-DMPU **2** enthalpy differences (henceforth refered to as $\Delta H_{(2\text{-}6)}$).

(5) $$\Delta H_{(2\text{-}6)} \approx -RTlnK_{open}$$

The % keto-DPMU can be plotted as a function of ΔpH, using PM3 calculated $\Delta H_{(2\text{-}6)}$ and equations (4) and (5). The ΔpH dependence was determined for R_1 = -H and R_1 = $-NO_2$ substituents (**Fig 8. 13**). The calculated pH dependence curves are compared with experimental points obtained from UV/VIS spectra.[1b] For simplicity, the unsubstituted and *p*-nitro substituted (Z)-O-BTZH+ compounds were assumed to have the same pKa. The experimantal points are plotted as pH = ΔpH. The pKa of (Z)-O-BTZH+ taken was -1.0 to maximize coincidence with the experimental points.

The experimental points in **Fig. 8.13** are of low quality because of the low intensities and broad spreading of the diazonium UV/VIS spectral peaks used to experimentally assay open-chain diazonium content. Despite this, the experimental results unequivocally demonstrate the same substituent dependence of the BTZ ring opening as the calculated results. Differences between the experimentally predicted and the computed pH dependence could arise because the computed pH dependance of tetrazepinone ring opening considers only one possible decomposition mechanism, and assmues the same pK_a for the *p*-nitro and unsubstituted benzotetrazepinone.

The correlation between the calculated an experimental results can be used to extract estimates of protonated benzotetrazepinone pKa's. The plots in **Fig. 8.13** used a pK_a = -1.0 to maximize correlation between the calculated and experimental pH dependence of ring opening. This suggests that the pKa of (Z)-O-BTZH$^+$ is near -1.0. This pKa value is not unreasonable for protonated BTZs, and may be expected, because protonated benzamides and phenyl ureas, (which are similar in

structure to DPMUs and BTZs), have experimental pKa's that vary between -1.0 to -3.3.[26b]

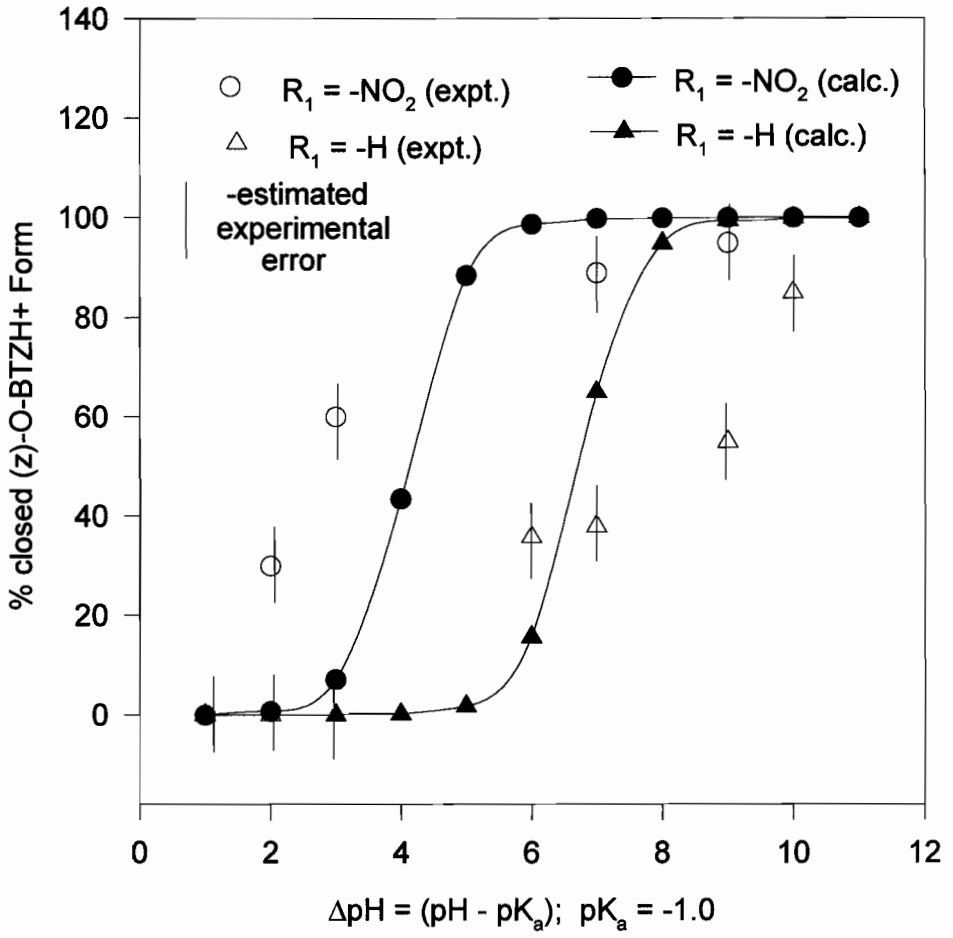


Figure 8.13:

Comparison of Experimental and Calculated pH Dependence of Benzotetrazepinone Ring Opening

8.4.5. Hammet Correlation of Substituent Effects in Benzotetrazepinone Ring Openings: Ring-closure of DPMU to $BTZH^+$ is analagous to many intramolecular

diazonium coupling reactions,[35-37] some of which have been shown to follow a Hammett-like substituent dependence in their reactivities. The aryl-ring substituent dependence of PM3 calculated DPMU ring-closure energies was investigated using both the Hammet[38]

(6) $$\rho\sigma = \log\frac{K'}{K} \quad ,$$

and the modified Hammett

(7). $$\rho^{+}\sigma^{+} = \log\frac{K'}{K} \quad ,$$

linear free energy relationships (LFER). Both of these LFERs are used in experimental physical chemistry to quantitatively measure aromatic system reactivity as a function of aryl ring substituent. The Hammett and modified Hammett constants (σ and σ+) describe the electron-donating/electron-withdrawing characteristics of the substituent. The magnitude and sign of the reaction constants (ρ and ρ^{+}) are used to measure the sensitivity of a reaction to aryl ring substitutent effects, and extract qualitative information about reactants, products and transition states.

The ring closure of DPMU to $BTZH^{+}$ is the reverse of the ring opening reaction in **Fig. 12**; therefore, $K_{close} = 1/K_{open}$. In equation (6) and (7), K is K_{close} for the unsubstituted benzotetrazepinone, and K' is K_{close} for the substituted benzotetrazepinone. Using equation (5), the open $\rightarrow$ close enthalpy differences $\Delta H_{(6\text{-}2)}$ can be recast in a form convenient for comparison with the Hammett equation.

(8) $$(u\Delta H_{(6\text{-}2)} - s\Delta H_{(6\text{-}2)}) / (2.303RT) = \log\frac{K'}{K}.$$

The s- and u- prefixes refer to 'substituted' and 'unsubstituted' BTZs enthalpy differences. Equation (8) was evaluated for a wide variety of substituents with the temperature arbitrarily set at T=300 K. The respective Hammett equations were plotted using σ and σ+ values from ref. 19, and logK/K' values computed with PM3. Plots of equations (6) and (7) were constructed (**Fig. 8.14a** and **8.14b**).

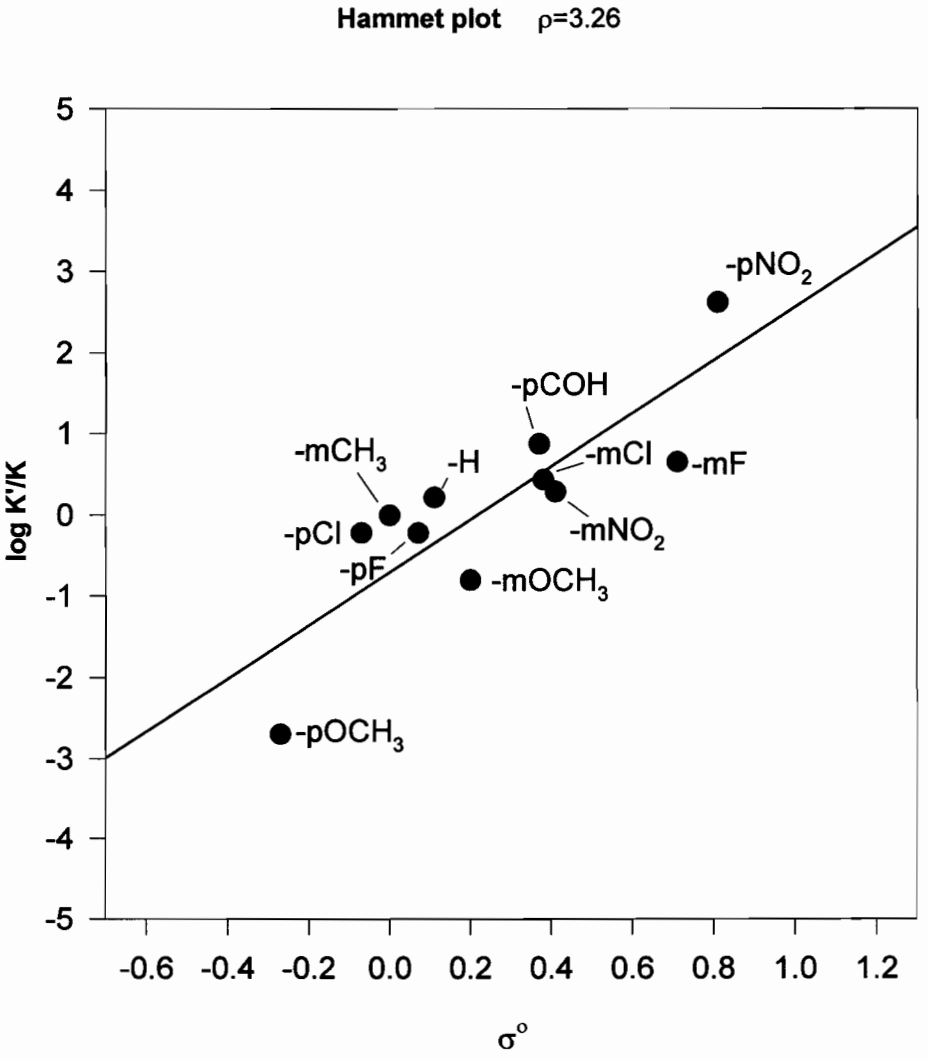


Figure 8.14a

Hammet Plot for DPMU Ring Closure

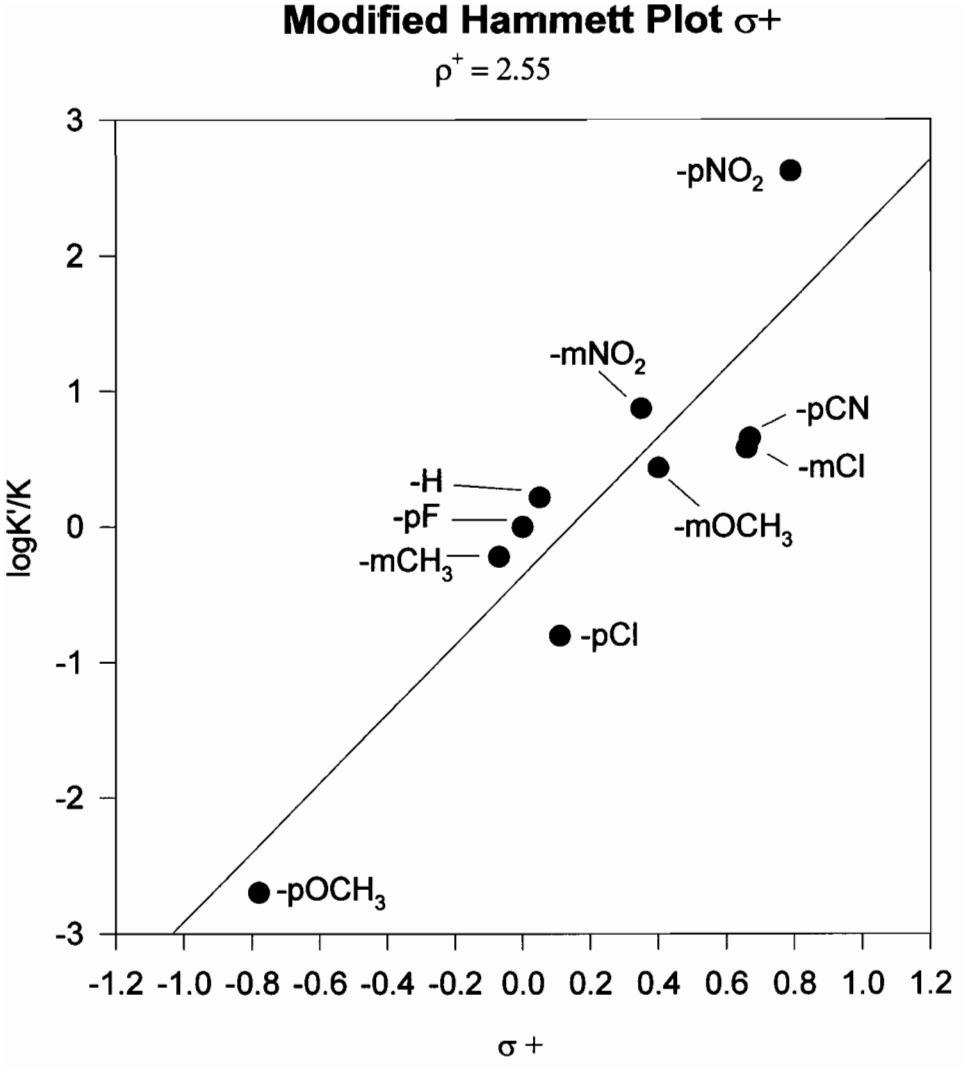


Figure 8.14b.

Modified Hammett Plot (σ+) of DPMU Ring Closure

Although there is some scatter in both plots (**Figs. 8.14a** and **8.14b**), the data clearly shows a Hammett like dependence of the PM3 calculated K_{close}. Hammett and modified Hammett reaction constants of $\rho = +3.26$ and $\rho+ = +2.55$ were calculated by linear regression analysis of the data. The Hammett plots slopes (ρ and $\rho+$) are both positive, consistent with a reaction rate increased by electron withdraw. The magnitudes of these reaction constants are relatively large, consistent with a highly substituent dependent reaction. Standard errors of 0.708 and 0.427 for ρ and $\rho+$ respectively indicate a better fit between the PM3 data and the $\sigma+$ substituent constants. This is expected, because the $\sigma+$ substituent constants are more appropriate for reactions where a positive charge is directly bonded to the phenyl ring.[38,39] In addition, the calculated $\rho+$ of +2.55 is similar to that of other diazonium coupling reactions, such as the intramolecular cyclization of 2-diazobenazamide.[39,40]

The open-chain DPMU is more resonance stabilized than either the out-of-plane benzotetrazepinone or the ring closure transition state, because of resonance between the diazonium group and the aryl ring in DPMU. The large positive $\rho+$ predicted by PM3 supports the traditional qualitative interpretation that a large positive $\rho+$ indicates the reactant is more resonance stabilized than either the transition state or the product.[38]

8.4.6. BTZ Experimental %Yields: A semi-quantitative relationship for tetrazepinone percent yields can be obtained by correlating experimental percent yields with PM3 enthalpy differences between the unprotonated BTZs and their corresponding N3 protonated, O14 protonated and open-chain DPMU forms. Enthalpy differences ($\Delta H_{rxn(2\text{-}1)}$, $\Delta H_{rxn(Z7\text{-}1)}$ and $\Delta H_{rxn(3\text{-}1)}$) were calculated for all tetrazepinones in **Fig. 8.2,** where experimental yields are available. The enthalpy differences all correlate linearly with the experimental yields, although some better than others (**Figs. 15a-c**). There is no extensive theoretical justification for this correlation, but it may be useful as a rough guide for practical synthesis.

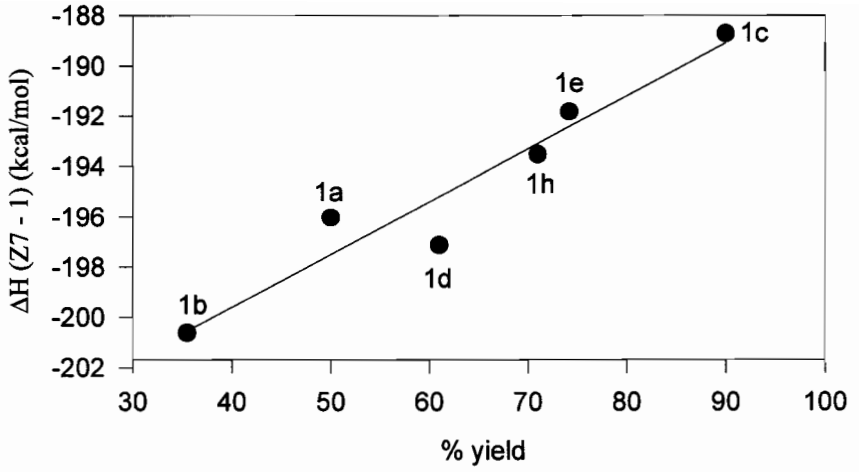


Figure 8.15a:

Z7 - 1 Relative enthalpy difference vs. percent yield

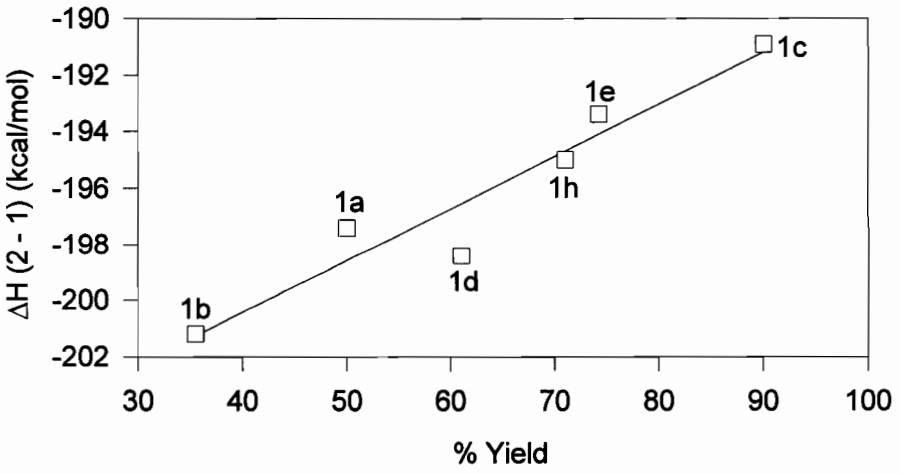


Figure 8.15b

2 - 1 Relative enthalpy difference vs. percent yield

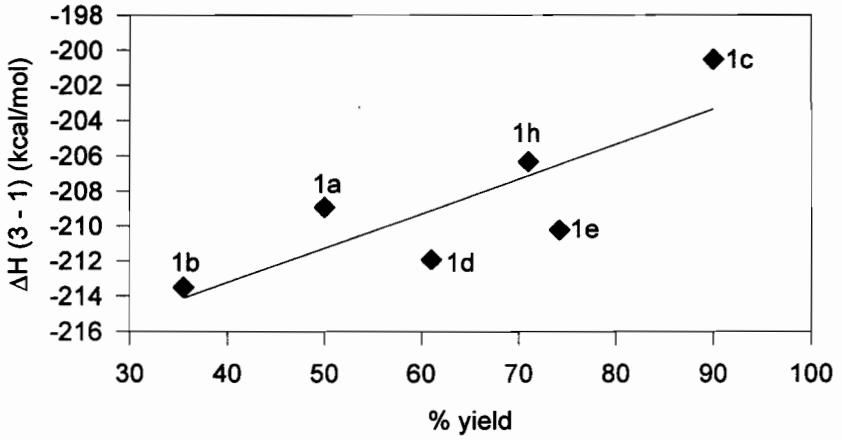


Figure 8.15c

3 - 1 Relative enthalpy difference vs. percent yield

8.4.7. Cyclization of DPMUs to N-Carbamoyl Benzotriazoles:

Attempts to synthesize BTZs with R = -H as an N5 substituent fail, because the DPMU cyclizes to N-carbamoyl benzotriazole (**Fig. 8.1**).[1a] Two different mechanisms were proposed to explain these results (**Fig. 8.16**). (a); the *keto* path, where cyclization at N5 of the keto DPMU (**3**) through transition state **TS10** forms the protonated species **10f** if R = -H, or the quaterary N+ species **10a** if R=-Me, and (b); the *enol* path, where the enol form of the DPMU **9** undergoes cyclization through **TS11** to the O protonated N-carbamoylbenzotriazole **11**. Ring closure to NCBT is never observed in internal cyclization of DPMUs with R = $-CH_3$, but structures in **Fig. 8.15** with R = $-CH_3$ were nevertheless investigated, for comparison with the R =-H structures.

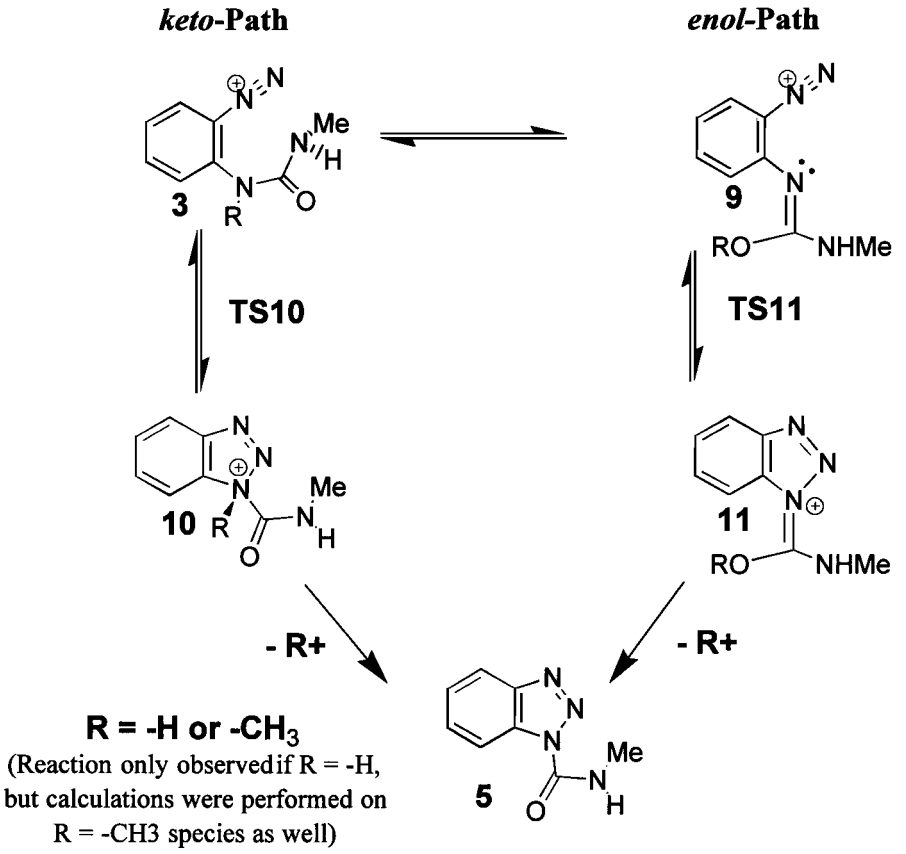


Figure 8.16:

Proposed Mechanisms for Ring Closure of DPMUs to NCBTs

PM3 heats of formation of the species in **Fig. 8.16** were calculated for R = -H and R = $-CH_3$. The reaction coordinate in **Fig. 8.17** summarizes the results for the *enol-* and *keto-* paths. The reaction coordinate in **Fig. 8.17** also includes the ring closure of enol-DPMU to (Z)-O-BTZH+ reaction coordinate for comparison.

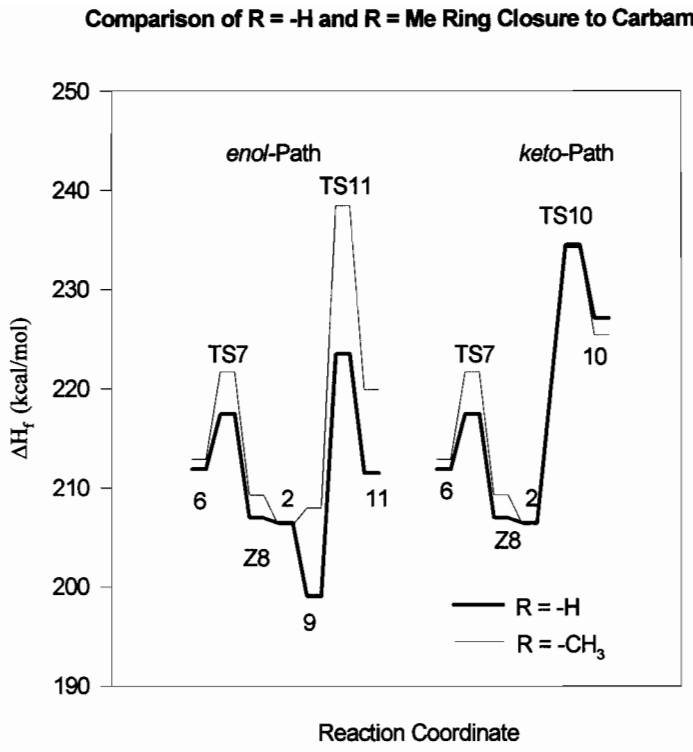


Figure 8.17:

Reaction Coordinates for *enol*- and *keto*-Paths for

DPMU → NCBT Ring Closure

The reaction coordinate in **Fig. 8.17** clearly shows that formation of the protonated (Z)-O BTZH+ is favoured over formation of either the quaterary N-carbamoyl benzotriazole **10a** or the protonated NCBT **10f**. The *keto*- path reaction coordinates are almost identical for R = $-CH_3$ and R = -H. Transition state barriers for ring closure to BTZ through the enolate **TS7** transition state are substantially lower than those for ring closure to quaternary or protonated N-carbomyl benzotriazoles, presumably because of strain effects. In addition, optimized geometries of **10** type species showed long N2-N5 bond lengths (> 1.8 Å) similar

to the elongated N2-N3 bonds in N-BTZH+ **5**, casting doubt on the existence of a stable **10** type species. By the *keto*-path, the R= -H and R = $-CH_3$ DPMUs should both cyclize to BTZs with similar yields.

The most striking difference between R = $-CH_3$ and R = -H in the **Fig. 8.17** reaction coordinate is in the *enol*-path. With R = -H, the enol-DPMU **9f** is more stable than the corresponding keto-DPMU **3f,** while with R = -Me the enol-DPMU **9a** is less stable than the corresponding *keto*-DPMU **3a**. Similar energy differences have been found with preliminary 3-21G** and 6-31G** *ab initio* calculations.

The large difference in the stabilities of **9** as a function of the N5 R = -H or R = $-CH_3$ substituent can be better understood from structural considerations. Representative bond lengths and angles of **9f** are in **Fig. 8.18,** and two overlay views of **9a** and **9f** are in **Fig. 8.19.**

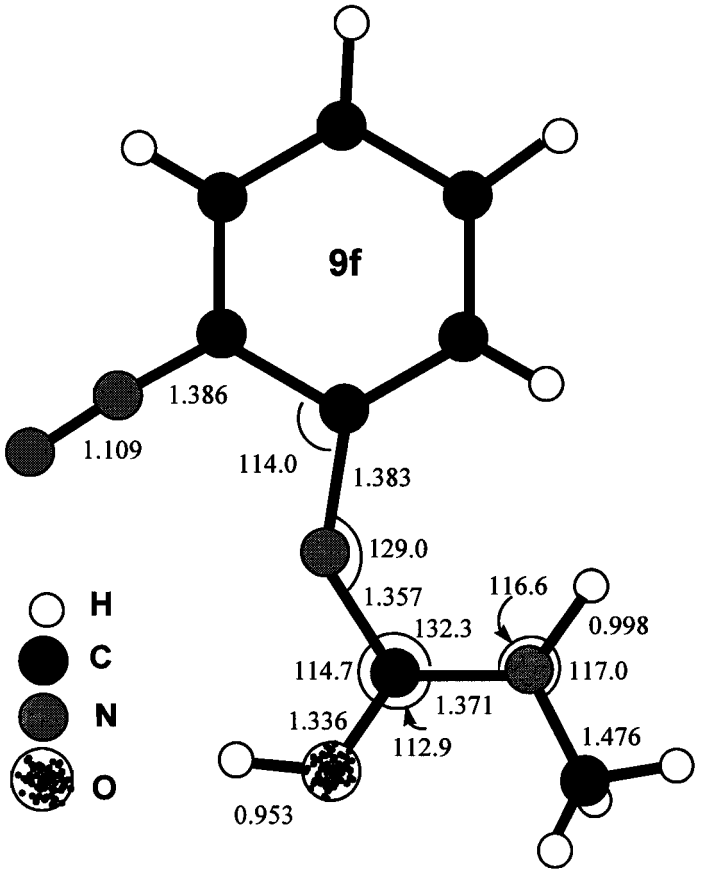


Figure 8.18:

Geometry of enol-DPMU **9f**

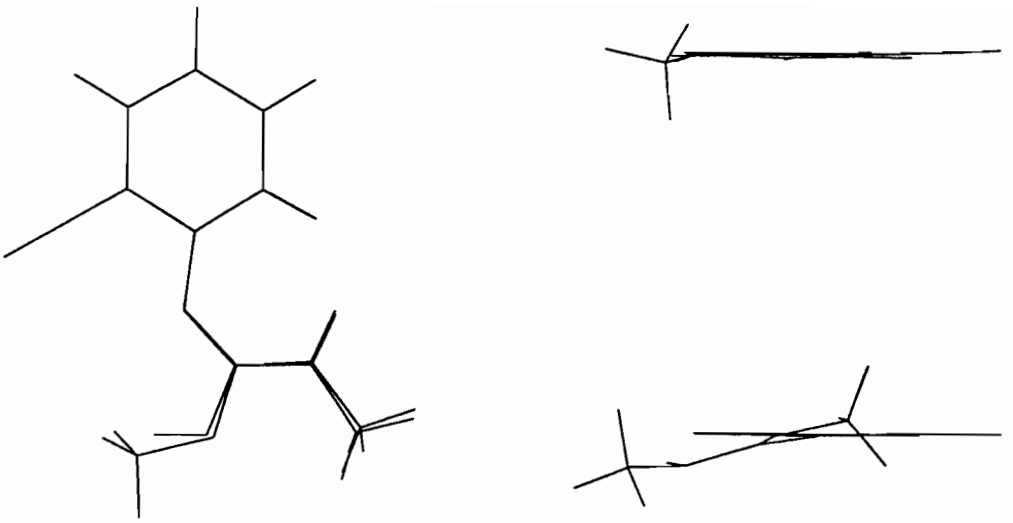


Figure 8.19:

Overlays of enol-DPMUs 9a and 9f

The overlay views in **Fig. 8.19** show that the bond lengths and angles are almost identical in both the **9f** (R = -H) and **9a** (R = $-CH_3$) enol-DPMUs. However, the side-chain in the **9f** species is in-plane with the benzene ring, so the enhanced stability of this species is probably due to conjugation with the benzene ring. The in-plane geometry of **9f** would also facilitate ring closure to the planar O-protonated NCBT **11f** species.

In contrast, the side-chain in the R = $-CH_3$ **9a** species is 11.2° out-of-plane with the benzene ring. This may be due to steric hinderence by the methyl group which prevents rotation of the enolate side-chain into the aromatic ring plane. The stability of **9a** is not enhanced by resonance, and the orientation of the side-chain in **9a** would hinder ring closure to O-methoxy NCBT **11a**. The O-methoxy NCBT **11a** is much less stable than the corresponding (Z)-O-BTZH+ **6a** for DPMU with

R= $-CH_3$, explaining why only BTZs are produced from cyclizations of DPMUs with R = $-CH_3$.

The reaction coordinate in **Fig. 8.16** shows that O-protonated NCBT **11f** has a stability equal to that of (Z)-O-BTZH$^+$ **6f** for DPMUs with R = -H. This suggests that an equal mixture of NCBT and BTZ should be obtained when DPMUs with R = -H cyclize. However, although DPMUs with R = -H can cyclize to BTZs, there will be a fast equilibrium between the DPMU **2f**, the protonated BTZ **6f**, and the protonated NCBT **11f**. Proton loss from (Z)-O-BTZH$^+$ is reversible because addition of base is necessary to precipitate the BTZ product out of the reaction mixture; however, proton loss from O-protonated NCBT **11f** is irreversible, because addition of base is unnecessary for precipitation of NCBT from the reaction mixture. Irreversible proton loss from **11f** drives the equilibrium towards the N-carbamoyl benzotriazole product, explaining why NCBT is the only compound produced from the internal cyclization of DPMU with R = -H. Thus, BTZ **1f** is difficult, if not impossible, to make by this synthetic approach.

8.5 Conclusions:

The PM3 semi-empirical method was used to successfully study the acid induced ring-opening of benzotetrazepinones (BTZs). The results indicate that BTZ most likely protonates at O as opposed to N3, and that ring-opening could occur from either an O or N3 protonated species. Protonated BTZ pKa's (pKa = -1.0) were estimated by comparison of the PM3 results with the experimentally determined pH dependence of BTZ ring opening. The PM3 energies of DPMU ring-closures to BTZs show a Hammett like aryl ring substituent dependence. The PM3 energies were used to determined LFER reaction constants (ρ and ρ+) which support the proposed mechanism, and are consistent with experimentally determined reaction constants for similar intramolecular azo-coupling reactions.[36] PM3 enthalpy differences were semi-quatitatively related to experimental percent yields. The mechanism proposed in **Fig. 8.4** and the subsequent PM3 calculations are consistent with all experimental observations. The PM3 method was also used

to explain why NCBT and not BTZ are produced from internal cyclizations of DPMUs with R = -H

References:

1. a) Jean-Claude, B.J. and Just, G., *J. Chem. Perkin Trans. **1**, 2525 (1991).*

   b) B.J. Jean-Claude, Ph.D Thesis, McGill University, (1992).

   c) Bertrand J. Jean-Claude and C.I. Williams, *Org. Mag. Reson.* accepted, October, 1996

   d) C.I. Williams, M.A. Whitehead and B.J. Jean-Claude, *J. Mol. Struct. (THEOCHEM)*, accepted , (August 1996)

2. Katritzky, A.R.; Yannakopoulou, K.; Anders, E.; Stevens, J.; and Szfran, M.; *J. Org. Chem.*, **22**, 5683 (1990).

3. Waldkowski, B.D., Smith, R.H. and Michejda, C.J., *J. Am.Chem. Soc.,* **113**, 7893 (1991).

4. Nielsen, K., Sotofte, I., and Johansen, H., *Acta Chemica Scandinavica,* **47**, 943, (1993).

5. D. Danovich, V. Zakrzewski, and E. Domnina, *J. Mol. Struct., (THEOCHEM),* **187**, 297 (1989).

6. M. de la Concepcion Forces-Forces, F. Hernandez Cano, R.M. Claramunt, D. Sanz, J. Catalan, F. Fabero, A. Fruchier, and J. Elguero, *J. Chem. Soc. Perkin Trans.* **2**, 237, (1990).

7. M. Sana, G. Leroy, M.T. Nguyen and J. Elguero, *Nouv. J. Chem., **3**, 607 (1979).*

8. P.R. Lowe, C.E. Sansom, C.H. Schwalbe, M.F.G. Stevens and A.S. Clarke, *J. Med. Chem.,* **32**, 3377 (1992).

9. A.S. Clarke, M.F.G. Stevens, C.E. Sansom and C.H. Schwalbe, *Anti-Cancer Drug Design*, **2**, 63 (1990).

10. M. Meyer, *J. Mol. Struct., (THEOCHEM),* **304**, 45, (1994).

11. W. M. F. Fabian, *Z. Naturforsch*, **42a,** 1328 (1990).

12. J. P. Ritchie, *J. Org. Chem.,* **24**, 3553, (1989).

13. A. K. Mishra and S. Dogra, *Indian J. Chem.*, **27a**, 851 (1988).

14. J. Mavri and D. Hadzi, *J. Mol. Struct.,* **224**, 285 (1990).

15. B. Vernon Cheney, *J. Org. Chem.,* **29**, 773 (1994).

16. A. L. Llamas-Saiz, C. Foces-Foces, O. Mo, M. Yanez, E. Elguero and J. Elguero, *J. Comput. Chem.,* **16**, 263, (1995).

17. B.S. Jursic, and Z. Zdravkovski, *J. Mol. Struct.(THEOCHEM), **309**, (1994).*

18. E. Fos, J. Vilarrasa and J. Fernandez, *J. Org. Chem.*, **20**, 4894 (1985).

19. M. Cabre, J. Farras, J. Fernandez-Sanz and *J. Vilarrasa, J. Chem. Soc. Perkin Trans. **2**, 1943* (1990).

20. B. S. Jursic and Z. Zdravkovski, *J. Mol. Struct. (THEOCHEM), **331**, 229 (1995).*

21. E. Pop, M.-J. Huang, M.E. Brewster and N. Bodor, *J. Mol. Struct (THEOCHEM), **312**, 1 (1994).*

22. J. Barluenga, T.L. Sordo, J.A. Sordo, S. Fustero and J. Gonzalez, *J. Mol. Struct. (THEOCHEM), **312**, 63* (1994).

23. J. Frau, J. Donoso, B. Vilanova, F. Munoz and F. Garcia Blanco, *Theor. Chim. Acta., **86**, 229 (1993).*

24. B. S. Jursic and Z. Zdravkovski, *J. Mol. Struct. (THEOCHEM), **312**, 85 (1994).*

25. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** 52 (1990).

b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.,* **14**, 1347 (1993).

26. A.M. Schmiedekamp, I.A. Topol, S.K. Burt, H. Razafinjanahary, H. Chermette,T. Pfaltzgraff and C.J Michejda, *J. Comput. Chem.*, **12**, 875 (1994).

26a J. Zabicky (Ed.), "The Chemistry of Amides" John-Wiley & Sons. Toronto (1970).

27. B. D. Wladkowski, R.H. Smith Jr., and C.J. Michejda, *J. Am. Chem. Soc.*, **113**, 7893 (1991).

29. A. Schmeidekamp, R.H. Smith Jr., and C.J. Michejda, *J. Org. Chem.*, **23** 3433 (1988).

30. R.H. Simth Jr., B.D. Waldkowski, J.A. Herling, T.D. Pfaltzgraff, B. Pruski, J. Klose and C.J. Michejda, *J. Org. Chem.*, **27**, 654 (1992).

31. K. Vaughan, *J. Chem. Soc., Perkin II*, 17 (1977).

32. R.H Smith Jr., C. L. Denlinger, R. Kupper, and C.J. Micheja, *J. Am. Chem. Soc.,* **106** 1056 (1984).

33. Bertrand J. Jean-Claude, J. F. Britten and G. Just, *Acta. Cryst. C.,* **49**, 1070 (1993).

34. Bertrand J. Jean-Claude and G. Just, *Magn. Reson. Chem.*, **30**, 571 (1992).

35. K.H. Saunders and R.L.M. Allen "*Aromatic Diazo Compounds*" (Edward Arnold: London) 1985.

36. H. Zollinger, "*Azo and Diazo Chemistry; Aliphatic and Aromatic Compounds*" (Interscience : New York) 1961.

37. H. Zollinger, "*Azo and Diazo Chemistry I: Aromatic and Heterocyclic*" (VCH: New York) 1995.

38. a)N.S. Issacs, *Physical Organic Chemistry*, (John Wiley and Sons: New York) 1987.

b) T.H. Lowry and K.S. Richardson "*Mechanism and Theory in Organic Chemistry*", (Harper-Row: New York) 1976.

c) A. Pross, "*Theoretical and Physical Principles of Organic Reactivity*" (Wiley: New York) 1995.

39. J.A. Dean, ed., "*Lange's Handbook of Chemistry*," 14th Edition, (McGraw-Hill, New York) 1992.

40. R.W. Taft, *Steric Effects in Organic Chemistry*, (M.Newman, ed.) (Wiley: New York) 1956.

In chapters two through eight, no attempt was made to develop new theoretical approaches to calculate molecular properties. The methods used in these chapters were sufficiently accurate to compute the structures and reactivities of the compounds. Indeed, useful chemical data can definitely be obtained from existing theories of molecular structure and computational algorithms. Increased accuracy over the methods used in chapters three to eight can only be obtained by moving to extensive and expensive post-Hartree-Fock electronic structure calculations. All post Hartree-Fock type calculations (see chapter 1) attempt to increase the accuracy of *ab initio* calculations by including the effects of electron correlation. Approaches to include electronic correlation effects are the Moller-Plesset (MP)[1] perturbation theory, density-functional exchange-correlation potentials,[2] and of course a Full CI.[3] However, all these methods are still computationally unfeasible for the size and number of molecules that chemists routinely study with modest computer resources. Consequently, despite the present success of Computational Chemistry, new approaches to molecular electronic structure are still required. Because electron correlation effects contribute to the energy in *ab initio* electronic energy calculations, it is a topic of great interest in modern Quantum Chemistry. Therefore, electron correlation will be addressed from a new standpoint in the following chapter.

1. a) C. Moller and M.S. Plesset, *Phys. Rev.* **46**, 618 (1934). b) J.A. Pople, J.S. Binkley and R. Seeger, *Int. J. Quantum Chem. Symp.* **10**, 1 (1976). c) R. Krishnan and J.A. Pople, *Int. J. Quantum Chem.* **14**, 91 (1976).

2. a) J. Keller, and J.L. Gazquez, *Phys. Rev. A*, **20**, 1289 (1979). b) O. Gunnarsson and B.I. Lundqvist, *Phys, Rev. B,* **13**, 4274 (1976). c) S.H. Vosko, L. Wilks and M. Nusair, *Can. J. Phys.*, **58,** 1200 (1980).

3. a)I. Shavitt in H.F. Schaefer, (ed.) "*Modern Theroretical Chemistry*", Vol. 3, p 189, (Plenum Press: New York, 1977). b) J.A. Pople, R. Seeger and R. Krishnan, *Int. J. Quantum Chem. Symp.*, **11**, 149 (1977). c) S.R. Langhoff and E.R Davidson, *Int. J. Quantum Chem.*, **8**, 61 (1974).

CHAPTER 9:

ELECTRON CORRELATION AND THE NATURAL ORBITAL REPRESENTATION: RELATIONSHIPS BETWEEN THE FRACTIONAL ORBITAL OCCUPANCY NUMBERS, CONFIGURATION INTERACTION EXPANSION COEFFICIENTS AND ONE-PARTICLE ENERGIES:

by C.I. Williams, S. Suba, and M.A. Whitehead
Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6.

9.1 Abstract:

The natural spin orbital (NSO) representation of the Configuration Interaction (CI) one-density is an infinite set of partially occupied one-particle states, whose occupation probabilities are related to the CI expansion coefficients. The correlation one-density, which is the difference between the Hartree-Fock (HF) and CI one-densities, is used with the virial theorem to calculate the electron correlation energy using one-particle energies only. Equations relating the NSO occupation probabilities to the NSO one-particle energies are derived, together with an approximate equation to relate the two-particle energies of the occupied orbitals to the two-particle energies of the virtual orbitals. Initial numerical results from SDCI calculations on C and Ne indicate the approximations to be accurate enough to systematically calculate correlation energies from Hartree-Fock eigenvalues and one-particle energies.

9.2 Introduction:

Ab initio quantum chemical calculations will be quantitatively predictive for large molecular systems when a computationally efficient and accurate treatment of the electron correlation energy is included. Although the correlation energy represents only a small fraction of the total energy, (0.5-1 %), in chemical terms, it often exceeds 200 kcal/mol; this reduces the predictive capabilities of calculations that do not include correlation. In chapter three, correlation and basis set errors cancelled out using *isolobal* reaction schemes. However, in many calculations, only one or two energy differences are needed, consequently the technique used in chapter three does not apply. Thus, *ab initio* quantum chemistry needs an accurate and computationally efficient treatment of electron correlation.

Theoretically, correlation energy can be accounted for by improving the Hartree-Fock wavefunction with Configuration Interaction expansions (CI),[1-3] Moller-Plesset perturbation theory (MP),[4-6] or Coupled Cluster Theory (CCA).[7-9] Alternately, the correlation energy can be included in Density-Functional Theory (DFT)[10-16a] by various exchange-correlation functionals which depend on the electron one-density. Self-Interaction Corrected Density-Functional Theories[16b] can treat electron correlation directly with the Vosko-Wilk-Nusair[15] correlation term, but they remain too computationally expensive for large systems. No one method has emerged universally applicable and computationally feasible.

DFT methods are computationally superior to CI type expansions, because they depend on the one-density, which relates to the physically observable electron density. However, the theoretically correct form of the exchange-correlation functional has not been found, and density functional methods all rely on approximate functionals. In contrast, CI expansions give a theoretically correct , although computational difficult, wavefunction.

Correlations in many-body quantum systems have been examined using the correlation one-density.[17,18] Here, the one-particle states are transformed into the natural spin orbital (NSO)[19] representation, and fractionally occupied to

minimize the total energy. The physical interpretation of correlation in this representation is simple; *particles that exhibit mutually repulsive interactions will relax into higher one-particle states to minimize their mutual repulsive energy.* The extent to which the electrons occupy the higher energy one-states is determined by the energy of the one-states. In principle, correlations could be effectively described by the single particle wavefunctions and their fractional occupation numbers. Therefore, this work will investigate the relationships between the correlation energy, HF and CI wavefunctions, and the correlation one-density, to develop a novel treatment of the correlation energy.

9.3 Hartree-Fock, CI, and Correlation One-Densities:[20]

In the Hartree-Fock (HF) approximation for an N electron system, the wavefunction (Ψ^{HF}) is formed from a single Slater determinant[20] of N orthonormal, one-electron orbitals, (ϕ_i).

1) $$\Psi^{HF}(x_1 \ldots\ldots x_N) = \sqrt{\frac{1}{N!}}\, |\phi_1(x_1)\ldots\phi_N(x_N)>$$

The one-electron density of the HF wavefunction, ρ^{HF} , is the sum over the squared ϕ_i orbitals.

2) $$\rho^{HF} = \sum_{i=1}^{N} \phi_i^* \phi_i$$

Correlation energy may be introduced with a CI type wavefunction,[20] which is an infinite linear combination of Slater determinants constructed by exciting electrons from the occupied to the virtual HF orbitals. The CI type wavefunction may be represented by grouping the types of excitations, singly, doubly or triply, etc,

3) $$\Psi^{CI} = c_o \Psi^{HF} + \sum_s c_s \Psi_s + \sum_d^N c_d \Psi_d + \sum_t^N c_t \Psi_t + \ldots \text{higher excitations}$$

where the $\{c_i\}$ coefficients are variationally determined. The one-particle orbitals of the CI wavefunction may be transformed into one-particle Natural Spin Orbitals[19] (NSO's) $\{\chi_i, \chi_a\}$. The the {i, j, k..} subscripts refer to the occupied orbitals, which are analogous to the occupied orbitals in the ground state HF wavefunction; the {a, b, c...} subscripts refer to virtual orbitals, which are analogous to the unoccupied in the ground state HF wavefunction. The one-density of the CI wavefunction, ρ^{CI}, may be expressed in a form similar to the HF one-density,

4) $$\rho^{CI} = \sum_i^N n_i \chi_i^* \chi_i + \sum_a^\infty n_a \chi_a^* \chi_a .$$

Since the NSO's are orthogonal, the one-density of the HF wavefunction can be transformed to the NSO representation,

5) $$\rho^{HF} = \sum_{i=1}^N \chi_i^* \chi_i .$$

The correlation one-density is the difference between the HF and CI densities,

6) $$\rho^{corr} = \sum_{i=1}^N (n_i - 1) \chi_i^* \chi_i + \sum_a^\infty n_a \chi_a^* \chi_a$$

The correlation one-density obeys the following sum rule:

7) $$\int \rho^{corr}(x_1) dx = 0$$

Because of the conservation of electrons, the following equation is also satisfied,

$$8) \qquad \sum_{i}^{N}(1-n_i) \;=\; \sum_{a}^{\infty} n_a$$

Physically, equation (8) means that the effect of electron correlation is to minimize the energy of the system by moving a fraction of the electron density from the occupied states ($1 - n_i$) into the virtual states (n_a). The NSO's themselves have the property that they minimize the depletion of electrons from the Fermi sea of occupied states, thereby optimizing the convergence of the CI expansion.[19] The NSO representation is attractive, because it allows this simple interpretation of electron correlation.

The occupation numbers of the NSO's may also be related to the coefficients of the CI expansion. If the ground state wavefunction is a single RHF Slater determinant, constructed from NSO's, different determinants in the CI expansion can be generated by exciting electrons from the N occupied orbitals to the virtual orbitals. If only the coefficients of the determinants, and not the orbitals themselves, are allowed to vary in this process, the optimized coefficients of the expansion are easily related to the occupation numbers of the NSO's. Determinants in the CI expansion which contain an NSO in question (occupied or virtual), will contribute $(c_i)^2$ to the occupation number of that NSO, where $\{c_i\}$ is the variationally determined coefficient of that determinant. If the determinant does <u>not</u> contain the NSO in question, it contributes zero to the occupation number of the NSO. Normalizing the CI wavefunction gives,

$$9) \qquad 1 \;=\; (c_o)^2 + \sum_{s}(c_s)^2 + \sum_{d}(c_d)^2 + \sum_{t}(c_t)^2 + \ldots \text{higher excitations}$$

and from the conservation of electrons,

10) $$N = \sum_{i=1}^{N} n_i + \sum_{a}^{\infty} n_a .$$

The occupation numbers of the lowest occupied orbitals can be related to the optimized CI coefficients by,

11) $$n_i = 1 - \sum_{a}^{\infty} (c_i^a)^2 + \sum_{j \neq i}^{N} \sum_{a}^{\infty} \sum_{b \neq a}^{\infty} (c_{ij}^{ab})^2 + \ldots \text{higher coefficients}$$

This equation will hold whether or not the CI one-particle wavefunction has been transformed to the NSO representation. For the virtual orbitals, the exact expression for the occupation numbers $\{n_a\}$ is

12) $$n_a = \sum_{i=1}^{\infty} (c_i^a)^2 + \sum_{i=1}^{N} \sum_{j \neq i}^{N} \sum_{b \neq a}^{\infty} (c_{ij}^{ab})^2 + \ldots \text{higher coefficients}$$

The absence of the ground state determinant coefficient (c_o) in the expression for $\{n_a\}$ reflects the absence of virtual orbitals in the ground state determinant.

9.4 One-Particle Correlation Energies:[20]

The total electronic energy of a multi-electronic system can be partitioned into the kinetic energy, $\langle T \rangle$, the electron-nuclear attraction energy, $\langle V_n \rangle$, and the electron-electron repulsion energy, $\langle V_{ee} \rangle$. The correlation energy (E^{corr}) is defined as the difference between the HF and CI total energies,

13) $$E^{corr} = E^{CI} - E^{HF}.$$

By analogy with the HF and CI total energies, the correlation energy can be split into kinetic, nuclear attraction, and electron-electron repulsion terms

14) $$E^{corr} = T^{corr} + Vn^{corr} + Vee^{corr}.$$

Partitioning the correlation energy this way allows use of the virial theorem; for atoms

15) $$E^{corr} = - T^{corr}$$

For molecules, (15) will be modified by a factor which depends upon forces on the nuclei.[20] For the moment, the situation is simplified by dealing with atoms only. However, it is true that for both atoms and molecules, the correlation energy may be expressed in terms of one particle energies only, when the total energy satisfies the virial theorem. For atoms, T^{corr} may be expressed in terms of the correlation one-density,

16a) $$T^{corr} = \int -1/2\nabla^2 \rho^{corr}(x_1;x_1')dx_1.$$

Upon integration, (16a) becomes

16b) $$T^{corr} = \sum_{i=1}^{N} (n_i - 1)T_i + \sum_{a=1}^{\infty} (n_a)T_a$$

where T_i and T_a are the kinetic energies of the occupied and virtual orbitals respectively. This is a straightforward and simple way to compute the correlation energy.

The formula in (16b) is deceptively simple, because both the NSO's and the occupation numbers have to be found, and these are normally only obtained after a CI calculation. Consequently, this approach is useful only if the NSO's and the occupation numbers can be obtained, or at least approximated, in a computationally efficient way that by-passes the Full CI calculation. For example, if one could fractionally populate the virtual orbitals of a standard HF type calculation in a manner which reliably yields a good estimate of the

correlation energy, this would be an improvement to existing methods. One way to go about this is to investigate the relations between the occupation numbers and the one-particle energies, in both the NSO and HF-MO representations.

9.5 Two-Particle Correlation Energies:

An alternate expression for the correlation energy, using Slater-Condon[20] rules, expresses the correlation energy as a sum over the cross terms between the ground and doubly excited states,

$$17) \qquad E_{corr} = \frac{1}{c_o}\sum_{d}^{\infty} c_d \, \langle \Psi^{*HF} | H | \Psi_d \rangle$$

where H is the electronic Hamiltonian, and the sum is over all the doubly excited determinants. Since no one-particle terms exist between determinants which differ by two or more spin orbitals, the total Hamiltonian in (17) can be replaced by the electron repulsion operator, $|r_1 - r_2|^{-1}$. Consequently, (17) can be used to express the correlation energy in terms of two-particle energies only. Combining (17) with (14) allows the one-electron and two electron terms to be grouped on different sides of the equation,

$$18) \qquad \frac{1}{c_o}\sum_{d}^{\infty} c_d \, \langle \Psi^{*HF} | |r_1 - r_2|^{-1} | \Psi_d \rangle - V_{ee}^{corr} = \; T^{corr} + V_n^{corr}$$

In equation (18) the two-electron terms on the LHS are dependent upon the CI coefficients explicitly, while the one-electron terms on the RHS may be expressed in terms of occupation numbers and one-particle energies. The difference between the CI and HF electron repulsion energies is V_{ee}^{corr} , which is given by the following integrals,

$$19) \qquad V_{ee}^{corr} = \int \Psi^{*CI} \, | r_1 - r_2 |^{-1} \Psi^{CI} dx_1 dx_2 - \int \Psi^{*HF} \, | r_1 - r_2 |^{-1} \Psi^{HF} dx_1 dx_2$$

The first integral in (19) is the electronic interaction energy of the CI wavefunction (V_{ee}^{CI}) and the second integral is the electronic interaction energy of the Hartree-Fock wavefunction (V_{ee}^{HF}). Insertion of the CI wavefunction from equation (3) into the first integral of (19) yields the following expression, which has been grouped into cross-determinantal terms, and energy terms involving one determinant only,

20)

$$V_{ee}^{CI} = (c_o)^2 <\Psi^{*HF}||r_1-r_2|^{-1}|\Psi^{HF}> + \sum_v (c_v)^2 <\Psi^{*v}||r_1-r_2|^{-1}|\Psi^{v}>$$

$$+ c_o \sum_v c_v <\Psi^{*HF}||r_1-r_2|^{-1}|\Psi^{v}> + c_o \sum_v c_v <\Psi^{*v}||r_1-r_2|^{-1}|\Psi^{HF}>$$

$$+ \sum_v \sum_u c_v c_u <\Psi^{*v}||r_1-r_2|^{-1}|\Psi^{u}>$$

The ground state Ψ_o is equated with the HF Ψ^{HF}. The summations over {u,v} denote a summation over all types of excitations. However, as a consequence of Brillouin's theorem[21] and the Slater-Condon[21] rules, the only non-zero cross-terms involving the HF ground state determinant are those between the ground state and the doubly excited determinants. The approximation that all other cross-terms (the last term in 20) are zero will be made. This will not be justified at the moment, but the resulting approximation to V_{ee}^{CI} will be re-stated; The first term in (20) is $(c_o)^2$ V_{ee}^{HF}, so by analogy, the integrals in the second term of (20) will be denoted by V_{ee}^{v}, to emphasize the fact that they are single determinant terms. In addition, the Hermitian nature[20] of $|r_1 - r_2|^{-1}$ allows grouping the third and fourth terms,

21) $$V_{ee}^{CI} \approx (c_o)^2 V_{ee}^{HF} + \sum_v (c_v)^2 V_{ee}^{v} + 2 c_o \sum_d c_d <\Psi^{*HF}||r_1-r_2|^{-1}|\Psi^{d}>$$

Using equation 17, the last term of equation 21 can be replaced with $2(c_o)^2E^{corr}$. Re-insertion of 21 into 18 yields,

22) $$(1-c_o^2)\,V_{ee}^{HF} - \sum_v (c_v)^2\ V_{ee}^v + (1-2\,c_o^{\,2})\,E^{corr} = \ T^{corr} + \ V_n^{corr}$$

which is the same as equation 18, except that the electronic repulsion energy of the CI wavefunction has been approximated as a weighted sum of single determinant electron repulsion energies. Equation 22 may be further simplified by using equation 15 and the gross approximation that $c_o \approx 1$.

23) $$\sum_v (c_v)^2\ V_{ee}^v \ = \ -V_n^{corr}$$

Note that the LHS of (23) contains two-particle energies only, while the RHS of (23) contains one-particle energies only. In atoms, the one-particle nuclear attraction energies are the ones which keep the electron in the atom, while the two-particle repulsion terms contain both electron repulsion terms (J) which raise the total energy and force electrons out of the system, and exchange terms (K), which lower the total energy and help keep the electrons on the atom.

The ground state coefficient can be allowed to be different from 1 by introducing a paramater $\alpha = (1\text{-}c_o^2)$, allowing a less stringent approximation of (22). Again using the virial theorem, equation (22) then becomes

24) $$\alpha V_{ee}^{HF} - \sum_v (c_v)^2\ V_{ee}^v \ = \ V_n^{corr} + 2\alpha T^{corr}$$

From (9), $\alpha = \sum_{v} (c_v)^2$; consequently, V_{ee}^{HF} can be placed inside the summation over $\{ V_{ee}^{v}$'s$\}$ in (24).

25) $$\sum_{v} (c_v)^2 [V_{ee}^{HF} - V_{ee}^{v}] = V_n^{corr} + 2\alpha T^{corr}$$

The total V_n^{corr} and T^{corr} of (25) may be partitioned into the individual occupied/virtual NSO one-particle nuclear attraction energies $\{V_{ni}$, $V_{na}\}$ and kinetic energies $\{T_i, T_a\}$

26) $$V_n^{corr} + 2\alpha T^{corr} = \sum_{i} (n_i - 1) V_{ni}^{corr} + 2\alpha T_i^{corr} + \sum_{a} (n_a) V_{na}^{corr} + 2\alpha T_a^{corr}$$

For convenience, the following abbreviations will be used,

27a) $$V_{ni}^{corr} + 2\alpha T_i^{corr} = U_i ; \qquad V_{na}^{corr} + 2\alpha T_a^{corr} = U_a$$

27b) $$\Delta V_{ee}^{v} = [V_{ee}^{HF} - V_{ee}^{v}]$$

Equation (25) then becomes

28) $$\sum_{v} (c_v)^2 \Delta V_{ee}^{v} = \sum_{i} (n_i - 1) U_i + \sum_{a} (n_a) U_a$$

The two-particle energies on the left hand side of (28) are independent of the summation over {i's} or {a's} on the RHS, therefore it will be replaced by $\beta = \sum_{v} (c_v)^2 \Delta V_{ee}^{v}$ In addition, β could be made to contain the last term in equation (20), lifting the assumption that this was zero. Consequently, all the difficulties of the two electron terms are now contained in the single β term! Equation (28) may be used to isolate expressions for $\{n_i,n_a\}$ in terms of β and $\{U_i,U_a\}$.

29a)
$$n_i - 1 = \frac{\beta}{U_i} - \sum_{j \neq i} (n_j - 1)\frac{U_j}{U_i} - \sum_a n_a \frac{U_a}{U_i}$$

29b)
$$n_a = \frac{\beta}{U_a} - \sum_i (n_j - 1)\frac{U_i}{U_a} - \sum_{b \neq a} n_b \frac{U_b}{U_a}$$

In theory, the summations over {a,b} are infinite, reflecting the full CI limit. However, in practice, the summation is truncated at some finite number N'. Substituting the expressions for n_a and $(n_i\text{-}1)$ from (29) into equation (8) and grouping terms involving β yields,

30)
$$\beta\left(\sum_i^N \frac{1}{U_i} + \sum_a^{N'} \frac{1}{U_a}\right) = \sum_a^{N'}\sum_{b \neq a}^{N'} n_b \frac{U_b}{U_a} + \sum_i^N \sum_{j \neq i}^N (n_j - 1)\frac{U_j}{U_i} + \sum_i^N \sum_a^{N'}\left((n_i - 1)\frac{U_i}{U_a} + n_a \frac{U_a}{U_i}\right)$$

Given a set of approximate NSO's, equations (29) and (30) could be used self-consistently to solve for β and the occupation probabilities. However, obtaining approximate NSO's remains an elusive problem. Stoitsov et. al.[18] found that the occupied NSO's of nuclear fermionic systems closely resembled the orbitals obtained from a single determinant mean-field approximation (MFA). However, the virtual orbitals generated by the MFA differed drastically from the virtual NSO's generated by diagonalization of the multi-determinant one-density matrix. This suggests that the occupied NSO's may be approximated with HF generated MO's, while the virtuals NSO should be approximated by suitable transforming the virtual MO's. Therefore, the total correlated one-density (ρ^{CI}), as given in (4) can be approximated as

31)
$$\rho^{CI} \cong \sum_i^N n_i \phi^*_i \phi_i + \sum_a^{N'} n_a \varphi^*_a \varphi_a$$

where $\{\varphi\}$ are suitably transformed HF virtual orbitals, and $\{\phi\}$ are the HF generated MO's.

9.6 Orbital Eigenvalues and One-Particle Energies:

A RHF MO calculation gives a set of one-particle states which satisfy the following Schrodinger equations for the occupied $\{\phi_i\}$ and virtual $\{\varphi_a\}$ states,[19]

32) $$F|\phi_i> = \varepsilon_i|\phi_i> \;\; ; \;\; F|\varphi_a> = \varepsilon_a|\varphi a>$$

where F is the Fock operator, and $\{\varepsilon_i, \varepsilon_a\}$ are the eigenvalues of the occupied and virtual states respectively. In addition, the orbital eigenvalues can be divided into one and two-particle terms,

33a) $$\varepsilon_i = T_i + V_{ni} + \sum_{j\neq i}^{N} <\phi_i\phi_j||\phi_j\phi_i>$$

and

33b) $$\varepsilon_a = T_a + V_{na} + \sum_{i}^{N} <\varphi_a\phi_j||\phi_j\varphi_a>$$

The virtual eigenstates differ from the occupied eigenstates because they interact with all the electrons in the ground state determinant.

To fractionally occupy, and modify, the Hartree-Fock orbitals to produce an NSO like situation, the Hartree-Fock eigenvalue expressions and total energy expressions are changed to include fractional occupation numbers. The HF generated orbitals will be assumed close if not identical to the NSO's. The part of the CI one-density which contains occupied NSO's will be defined as

34) $$\rho^{occ} = \sum_{i}^{N} n_i \phi_i^* \phi_i$$

The part of the total CI energy which is attributed to the occupied part of the CI one-density is E_{occ},

35) $$E_{occ} = \sum_{i}^{N} n_i (T_i + V_{ni}) + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} n_i n_j <\phi_i\phi_j||\phi_j\phi_i>$$

An analogous expression can be written for the virtual part of the total CI energy,

36) $$E_{virt} = \sum_{a}^{N'} n_a (T_a + V_{na}) + \frac{1}{2}\sum_{a}^{N'}\sum_{b\neq a}^{N'} n_a n_b <\varphi_a\varphi_b||\varphi_b\varphi_a>$$

The interaction energy between the virtual and occupied orbitals (E_{int}) is two-particle in nature, and contains no one-particle terms,

37) $$E_{int} = \sum_{a}^{N'}\sum_{i}^{N} n_a n_i <\varphi_a\phi_i||\phi_i\varphi_a>$$

Since $\{\phi_i\}$ and $\{\varphi_a\}$ are in different sets, there is no factor of 1/2 which normally appears to cancel out overcounting. The HF total energy is

38) $$E^{HF} = \sum_{i}^{N}(T_i + V_{ni}) + \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N} <\phi_i\phi_j||\phi_j\phi_i>$$

Since $E^{corr} = E^{CI} - E^{HF}$, and $E^{CI} = E^{occ} + E^{virt} + E^{int}$, the correlation energy may be rewritten as,

39) $$E^{corr} = \sum_{i}^{N}(n_i - 1)\varepsilon_i + \sum_{a}^{N'} n_a\varepsilon_a + V_{ee}^{corr} + \sum_{i}^{N}\sum_{j}^{N}(1-n_i) <\phi_i\phi_j||\phi_j\phi_i> - \sum_{i}^{N}\sum_{a}^{N'} n_a <\phi_i\varphi_a||\varphi_a\phi_i>$$

The virial theorem can be used to show that $V_{ee}^{corr} = -2T^{corr} - V_n^{corr}$. The last two terms in (39) have opposite signs, and therefore will be approximated as zero, because they will approximately cancel each other. Using $E^{corr} = -T^{corr}$, (39) can be recast as follows;

40) $$\sum_{i}^{N}(n_i - 1)[\varepsilon_i - T_i - V_{ni}] = \sum_{a}^{N'} n_a[\varepsilon_a - T_a - V_{na}]$$

9.7 Numerical Investigations:

The equations derived lead to the expected conclusion that the NSO occupation numbers are dependent upon both the one and two particle energies. In equations (29) and (30), the evils of the two electron energies are contained in β, which may be solved self-consistently. The effect of the one-electron energies on a particular occupation number are explicitly divided into one-orbital terms. However, neither (29) nor (30) separate the one particle energies from the occupation numbers. In equation (40), the eigenvalues and one particle energies of the occupied and virtual orbitals are related to each other through the occupation numbers.

The relationships between $\{n_i, n_a\}$ and the NSO one particle energies was investigated by performing a standard singles/doubles CI calculation on a the closed shell Ne atom, with a 6-311G basis set using the GAMESS[22] package. The GAMESS output consisted of NSO's in the atomic orbital basis, NSO occupation numbers, and the one-electron AO integrals (T_{kl}, Vn_{kl}). The kinetic (T_i) and nuclear attraction (Vn_i) energies of the i^{th} NSO were computed using the following transformations of the GAMESS output.

41) $$T_i = \sum_k c_k^i \sum_l c_l^i T_{kl}$$

42) $$V_{ni} = \sum_k c_k^i \sum_l c_l^i V_{nkl}$$

where the $\{c_k^i\}$ are the coefficients of the k^{th} basis set AO in the i^{th} NSO. The total kinetic and nuclear attraction energies of the system are then

43) $$T^{tot} = \sum_{i}^{N} n_i T_i + \sum_{a}^{N'} n_a T_a$$

44) $$V_n{}^{tot} = \sum_{i}^{N} n_i V_{in} + \sum_{a}^{N'} n_a V_{na}$$

The one-particle MO and NSO energies, as well as the occupation numbers and HF eigenvalues for Ne are in **Table 9.1**. The results show that the occupied NSO's and MO's have very similar kinetic and nuclear attraction energies, reflecting the similarity between the occupied MO's and NSO's. Therefore, approximating the occupied NSO's as HF MO's is a good starting point. In contrast, the virtual NSO's and MO's have quite different kinetic and nuclear attraction energies, reflecting a larger difference between the HF virtual MO's and the NSO's.[18] The total one particle energies {T+V} closely follow the expected stability of the occupied and virtual orbitals, with the occupied orbitals all having lower (T+V) values than the corresponding virtual orbitals.

The numerical accuracy of equations (29) and (30) were investigated by substitution of calculated values of U_i, U_a, n_i, n_a. However, numerical agreement was not acheived because the CI wavefunction and energy did not obey the virial theorem. Typically, the virial ratio <V>/<T> varied between 1.96 and 2.02, with none of the calculations having a perfect ratio of 2. This 1-2% error in the kinetic and potential energies is the same order of magnitude as the correlation energy itself, thereby creating gross errors in the correlation virial relationship, which is fundamental in the derivation of equations (29) and (30). A rough calculation testing equation (40) was performed on Ne and C by using the HF generated eigenvalues and the NSO kinetic and potential energies generated from the SDCI calculation. The results in **Tables 9.1-9.2** show that equation (40) is accurate to within 10% for the atoms in question, indicating that the approximation made in deriving (40) may be applicable in certain situations.

Table 9.1: Neon NSO and MO Orbital Occupancies, One Particle Energies and HF Eigenvalues (a.u.)

Occupied

State #		Occupancy	T_i	$T_i + V_{ni}$	ε_i^{HF}	$\varepsilon_i - T_i - V_{ni}$
1	HF	2.0000	46.269	-49.910	-32.760	17.150
	CI	1.9995	43.961	-49.758		16.998
2		2.0000	5.189	-11.089	-1.919	9.170
		1.9936	5.902	-11.155		9.236
3,4,5		2.0000	4.274	-10.106	-0.842	9.264
		1.9887	4.284	-10.099		9.257

$\Sigma\ (2 - n_i)\ (\varepsilon_i - T_i - V_{ni}) = 0.381$

Virtuals

State #		Occupancy	T_i	$T_i + V_{ni}$	ε_i^{HF}	$\varepsilon_i - T_i - V_{ni}$
6,7,8	HF	0.0000	3.749	-5.792	1.410	7.202
	CI	0.0108	7.369	-7.155		8.565
9		0.0000	6.048	-6.261	1.591	7.852
		0.0070	10.380	-5.987		7.578
10,11,12		0.0000	18.822	-4.669	8.096	12.765
		0.0003	15.182	-3.313		11.409
13		0.0000	207.853	65.122	86.873	21.751
		0.0005	205.118	64.759		22.114

$\Sigma\ n_a\ (\varepsilon_a - T_a - V_{na}) = 0.345$

Table 9.2:
Carbon NSO Orbital Occupancies, One Particle Energies and HF Eigenvalues (a.u.)
Occupied

State #	Occupancy	T_i	$T_i + V_{ni}$	ε_i^{HF}	$\varepsilon_i - T_i - V_{ni}$
1	1.9997	16.329	-17.950	-11.343	6.607
2	1.9504	1.695	-3.970	-0.825	3.145
3	1.0001	1.435	-3.611	-0.434	3.177
4	1.0000	1.422	-3.606	-0.434	3.172

$$\Sigma\ (2 - n_i)\ (\varepsilon_i - T_i - V_{ni}) = 0.158$$

Virtuals

State #	Occupancy	T_i	$T_i + V_{ni}$	ε_i^{HF}	$\varepsilon_i - T_i - V_{ni}$
5	0.0493	1.422	-3.606	0.0347	3.641
6	0.0004	53.611	14.4767	0.598	-13.879
7	0.0001	4.006	-1.110	0.548	1.658

$$\Sigma\ n_a\ (\varepsilon_a - T_a - V_{na}) = 0.174$$

9.7 Conclusions:

The relationships between the CI expansion coefficients, the NSO occupation numbers, and the HF one-particle eigenvalues were investigated, and an approximate relationship between them derived. The virial theorem and Slater-Condon rules allowed derivation of equations relating the one-particle and two-particle parts of the correlation energy. An approximation of the NSO occupation probabilities in terms of other NSO occupation probabilities was developed. A term β which contains the correlation two-electron energies was also introduced. Preliminary sample SDCI calculations performed on Ne and C atoms were encouraging, and verify the accuracy of equation (40) to within 10%. This is surprisingly good, since equation (40) was calculated using a mixture of HF eigenvalues and NSO one-particle energies. In addition, HF calculations in the

GAMESS package can be forced to obey the virial theorem,[20] but no such option exists for CI calculations; the accuracy of equations (29), (30), and (40) should improve if both the HF and CI energies are forced to obey the virial theorem.[23] This relationships derived in this paper are presently being used to develop a computational method which fractionally occupies and suitably transforms HF virtual orbitals in a iterative fashion, approximating the correlation energy without actually performing a CI calculation. Such a method will be an improvement on existing methods.

Further work investigating the effect of extended basis sets and higher CI excitations is underway, along with extensions of the present theory to molecular systems. In addition, a method for using equations (29), (30) and (40) to iteratively generate approximate NSO's from Hartree-Fock generated MO's is currently being developed.

References:

1. I. Shavitt in H.F. Schaefer, (ed.) "*Modern Theroretical Chemistry*", Vol. 3, p 189, (Plenum Press: New York, 1977).
2. J.A. Pople, R. Seeger and R. Krishnan, *Int. J. Quantum Chem. Symp.*, **11**, 149 (1977).
3. S.R. Langhoff and E.R Davidson, *Int. J. Quantum Chem.*, **8**, 61 (1974).
4. C. Moller and M.S. Plesset, *Phys. Rev.* **46**, 618 (1934).
5. J.A. Pople, J.S. Binkley and R. Seeger, *Int. J. Quantum Chem. Symp.* **10**, 1 (1976).
6. R. Krishnan and J.A. Pople, *Int. J. Quantum Chem.* **14**, 91 (1976).
7. J. Cizek, *J. Chem. Phys.* **45**, 4256 (1966).
8. R.J. Bartlett and J. Noga, *Chem. Phys. Lett.*, **134**, 126 (1987).
9. R.J. Bartlett and J. Noga, *J. Chem. Phys.*, **86**, 7041 (1987).
10. C.O. Almbladh and A.C. Pedroza, *Phys. Rev. A.*, **29**, 2322 (1984).
11. R.A. Harris and L.R. Pratt, *J. Chem. Phys.*, **83,** 4024 (1985).
12. J. Cioslowski, *Phys. Rev. A,* **41**, 3458 (1990).
13. J. Keller, and J.L. Gazquez, *Phys. Rev. A*, **20**, 1289 (1979).
14. O. Gunnarsson and B.I. Lundqvist, *Phys, Rev. B,* **13**, 4274 (1976).
15. S.H. Vosko, L. Wilks and M. Nusair, *Can. J. Phys.*, **58,** 1200 (1980).
16. a.) J.P. Perdew, *Chem. Phys. Lett.*, **64**, 127, (1979).

    b.) M.A. Whitehead, "*The Self-Interaction Corrected Local Spin Density Approximationwith Vosko, Wilks and Nusair Correlation and Full Relativistic Corrections*" in "*Recent Advances in Density Functional Methods, Part II*" ed. D.P. Chong (World Scientific: Singapore) 1996.
17. D.S Lewart, V.R. Pandharipande and S.C. Pieper, *Phys. Rev. B.,* **37**, 4950 (1988).
18. M.V. Stoitsov, A.N. Antonov, and S.S. Dimitrova, *Phys. Rev. C.,* **48**, 74 (1993).
19. P.O. Lowdin and H. Shull, Phys. Rev. **101**, 1730 (1956).
20. M. Lehd and F.J. Jensen, J. *Comput. Chem.* **12**, 1089 (1991).

21. A. Szabo and N.S. Ostlund, "*Modern Quantum Chemistry*" (MacMillan: New York) 1982.

22. a) M.W. Schmidt, K.K. Baldridge, J.A. Botz, J.H. Hensen S. Koseki, M.S. Gorden, K.A. Nguyen, T.L. Windus, and S.T. Elbert, *QCPE Bull.*, **10** 52 (1990).

b) M.W. Schmidt, K.K. Baldridge, J.A. Botz, S.T. Elbert, M.S. Gorden, J.H. Hensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis and J.A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).

23. M.S. Gopinathan and M.A. Whitehead, *J. Chem. Phys.*, **65**, 196 (1976).

Chapter 10:

Conclusions, Claims to Original Research and Future Work

10.1 General Conclusions:

Throughout the thesis, computational techniques were successfully used to study the structure and reactivity of a series of nitrogen heterocyclic compounds. Chapter two demonstrated that a simple AAP method can produce predictive results similar to high level *ab initio* calculations. In chapter three, the averaging of *ab initio* nitrogen heterocyclic heats of formation over different basis sets and *isolobal* reaction schemes was shown to successfully cancel basis set and correlation errors, and give heats of formation in agreement with experiment. Comparison of the *ab initio*, semi-empirical and experimental azole, azine, and benzoazole heats of formation demonstrated systematic errors in MNDO and AM1, and led to the determination of semi-empirical heat of formation correction terms. Chapter three also led to the conclusion that PM3 is superior to AM1 and MNDO at reproducing experimental nitrogen heterocyclic heats of formation.

Comparison of MNDO, AM1 and PM3 isocyanate + ethene cycloaddition transition state geometries showed that PM3 transition states bear closer resemblance to the structures and energies of *ab initio* generated transition states. Further proof of the superiority of PM3 at treating nitrogen-rich heterocyclics was presented in chapter five; good correlation between PM3 and *ab initio* generated diazoazole heats of formation led to the first trustworthy estimates for the heats of formation of these compounds. The PM3 and *ab initio* predictions of an unstable singlet spin state 4-azolylidene was used to explain why different products result from thermolysis and photolysis of diazo-1,2,3-triazole.

The trends in the semi-empirical computed activation barriers for the diazoazole + ethyne and diazoazole + ynamine cycloaddition reactions reflect experimental reactivities. The regio-specificity of ynamine cycloaddition with diazoazoles could be

predicted with the MNDO, AM1 and PM3 methods. The lack of diazoazole reactivity with cyanoethynes is also predicted.

These results, together with the results presented in chapters three through five, justified the use of the PM3 method to study structure and reactivity of tetrazepinone compounds. The PM3 method was shown to be superior to the MNDO and AM1 methods at reproducing experimental benzotetrazepinone geometries in chapter seven. In chapter eight, the effect of aryl ring substituents on benzotetrazepinone stabilities and experimental percent yields was successfully modelled and explained with the PM3 method. Thus, the PM3 method should be suited for continued study of the tetrazepinone class of compounds.

The relationships between CI coefficients, NSO occupation numbers, and one-particle energies derived in chapter nine could be developed into a novel approach to calculating molecular electronic structure.

10.2 Claims to Original Research:

Chapter 2: - The first Atom-Atom-Potential treatment of the stability and dynamics of endohedral gas molecules inside fullerene cages was presented. The results conclusively show that the endohedral complexes are energetically feasible, and therefore, possible to synthesize.

Chapter 3: -The first to demonstrate that averaging the heats of formation calculated using a variety of isolobal reaction schemes and *ab initio* total energies from different basis sets, leads to more accurate results than using one reaction scheme or basis set alone.

-The first corrections to PM3 azine heats of formation

-First calculations of indazine, azolotriazine and pyrazepine heats of formation using isolobal reaction schemes.

Chapter 4: -First to show that AM1, MNDO and PM3 give isocyanate geometries in agreement with experiment.

-First to show that PM3 is superior to AM1 and MNDO in predicting isocyanate → β-lactam transition state energies and geometries; The PM3 activation energy barriers and transition state geometries are parallel to *ab initio*, but are obtained with significantly shorter computing time.

Chapter 5: -first study to show the bond localization in diazoazole rings with semi-empirical and *ab initio* calculations.

-first to demonstrate the flexibility of the C=N=N diazo bond angle with semi-empirical and *ab initio* calculations.

-first study to calculate diazoazole heats of formation from *ab initio* total energies.

-first semi-empirical and *ab initio* computational treatments of azolylidene geometries and energies.

-first to show that the triplet spin state is the most stable for azolylidenes.

-first computational explanation of the different reaction products resulting from thermal and irradiated excitation of 4-diazotriazole

Chapter 6: -first semi-empirical treatment of diazoazole cycloaddition reactions with ethyne, ynamine and cyanoethyne.

-first computational prediction of the relative reactivities of ethyne and ynamine with diazoazoles

-first theoretical demonstration of the regiospecificity of ynamine + diazoazole cycloadditions.

-first computational explanation of the lack of reaction between cyanoethyne and diazoazoles.

Chapter 7: -first molecular mechanics, semi-empirical and *ab initio* study of benzotetrazepinones and non-benzofused tetrazepinone ring systems.

-first complete discussion of the unique conformational features of the tetrazepinone ring.

-first computational prediction of tetrazepinone ring flexibility.

-first reported transition state for the unique asymmetric racemization of benzotetrazepinone.

Chapter 8: -first computational study of acid-induced benzotetrazepinone decomposition.

-first quantitative prediction of aryl-ring substituent effects on benzotetrazepinone synthetic yields, and the pH dependence of their acid-catalyzed decompositions.

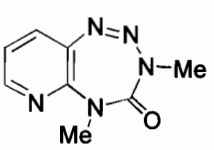
-first proposed mechanism and computational study of the diazophenyl urea ring closure to N-carbamoyl benzotriazole.

Chapter 9: -first to derive the relationship between the Natural Spin Orbital occupation numbers and Configuration Interaction expansion coefficients.

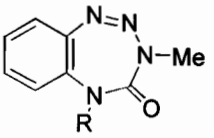
10.3 Future Work

10.3.1. Computer-Aided Molecular Design of Benzotetrazepinone Anti-Cancer Drugs:

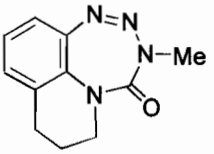
The encouraging results in chapters 7 and 8 suggest that the PM3 method should be used for the computer-aided design and development of tetrazepinones (TZPs)[1,2] **1-3** as anti-cancer agents. Hurdles to TZP development as anti-cancer drugs include *in vivo* instability, unknown decomposition mechanism(s), and a lack of detailed information on their physiological mode(s) of action.



1



2



3

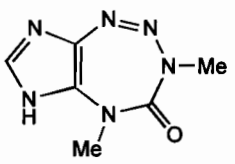
The PM3 method computationally confirmed the acid-catalyzed benzotetrazepinone decomposition mechanism and predicted reasonable pKa's and Hammett reaction constants (ρ and $\rho+$)[5] prior to any experimental work. Computed enthalpy differences semi-quantitatively predict tetrazepinone experimental yields.[5] Continuation of our computational studies could save years of experimental work and hasten tetrazepinone development as anti-tumor drugs by predicting properties and energetics prior to synthesis, and by establishing various decomposition mechanisms through judicially performed calculations.

The lack of a thorough understanding of fundamental tetrazepinone chemistry precludes any comprehensive predictions as to what conditions could optimize benzotetrazepinone anti-tumor behaviour. Once benzotetrazepinone chemistry has been thoroughly established through a combination of computational and experimental means, the focus of the research will broaden towards improving benzotetrazepinone anti-tumor effectiveness. This research has already produced fruitful interdisciplinary collaboration with organic chemists and oncologists. Because of the size and quantity of tetrazepinone compounds to be investigated, semi-empirical methods will be initially preferred over *ab initio* methods in view of time and computer constraints. We expect the semi-empirical methods to give a satisfactory pictures of various possible TZP decomposition

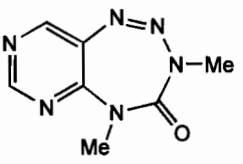
mechanisms, based on the research presented in this thesis. *Ab initio* calculations will be performed for comparison with the semi-empirical results in cases of doubt, and routine *ab initio* studies will be performed as computer time permits.

The following specific subjects should be computationally investigated;

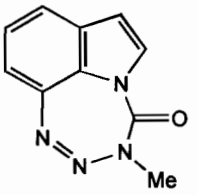
A) THE STABILITIES OF AS-OF-YET UNSYNTHESIZED TETRAZEPINONES AND TETRAZEPINONE ANALOGUES: Semi-empirical computational methods will be used to predict yields, stabilities and geometries of promising tetrazepinone candidates such as **4**, **5** and **6** prior to their synthesis.



4

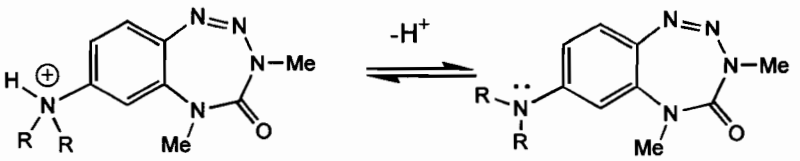


5



6

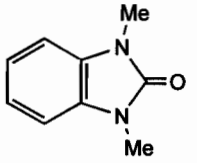
Chapter 8 demonstrated that electron-withdrawing substituents enhance tetrazepinone stability; the amino-tetrazepinone **7** is an example of a tetrazepinone with a "molecular switch" that allowing facile alteration of its stability/activity properties, because protonation alters the electron-donating/electron-withdrawing characteristics of the amino aryl ring substituents. This and other promising tetrazepinones suggested by computational studies are the target compounds in ongoing TZP synthetic efforts.



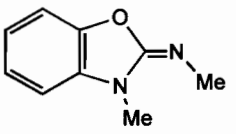
7

B) MECHANISMS OF BENZOTETRAZEPINONE DECOMPOSITION: The acid-catalyzed decomposition of benzotetrazepinones is now well understood due to our computational work,[5] but so far the only proposed alkaline TZP decomposition mechanism[1] has been

refuted by computational and experimental evidence. Other decomposition mechanisms also remain unclear, and some may be important in the anti-tumor mode(s) of action. TZPs do not decompose by mechanisms akin to the basic decompositions of structurally similar imidazotetrazolones anti-cancer agents, which lose CO_2 ; instead TZPs lose N_2 , producing **8** and **9** by unknown mechanisms.

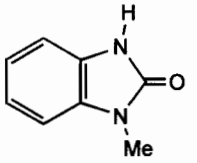


8

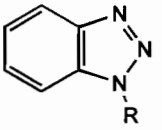


9

Other decompositions such as the loss of diazomethane giving **10**, and the loss of isocyanate giving **11** should be investigated computationally.



10



11

C) PARAMATERIZATIONS OF MOLECULAR MECHANICS FORCE-FIELDS: TZP interactions with biopolymers may be important to anti-tumor activity, but no computational studies of this type have been reported. The size of biopolymers necessitates the use of molecular mechanics methods such as AMBER and OPLS, which are not adequately parameterized for tetrazepinones. In chapter 7, the MM+ method was shown to satisfactorily treats benzotetrazepinones, although MM+ still fails to predict many important benzotetrazepinone structural features. In addition to quantum mechanical calculations, work has begun to optimize tetrazepinone molecular mechanics parameters for the AMBER force-field, using semi-empirical, *ab initio* and X-ray data. The results of this work will enable proper investigation of benzotetrazepinone-biopolymer interactions with molecular mechanics methods.

10.3.2: Development of New Approaches to Molecular Electronic Structure:

The work in chapter nine is a preliminary step in the development of a new theory of molecular electronic structure. The main problem to be solved before these results can be utilized as a coherent computational procedure is the initial estimation of the NSO's; the Hartree Fock occupied MOs are probably a good starting point for estimation of the occupied NSOs; however, the Hartree-Fock virtual orbitals are quite different from the unoccupied NSOs. Transforming the HF virtual MOs with different orbital localization and transformation procedures will be tested, as an approach to solve this problem.

References:

1. B.J. Jean-Claude and G. Just, *J. Chem. Perkin Trans.* **1**, 2525 1991.
2. B.J. Jean-Claude, Ph.D Thesis, McGill University, 1992.
3. B.J. Jean-Claude and C.I. Williams, *Org. Mag. Reson.* accepted, Sept. 1996
4. C.I. Williams, M.A. Whitehead and B.J. Jean-Claude, *J. Mol. Struct. (THEOCHEM)*, accepted , (August 1996)
5. C.I. Williams, M.A. Whitehead and B.J. Jean-Claude, *J. Mol. Struct. (THEOCHEM)*, submitted (October 1996)