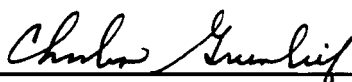


AN ABSTRACT OF THE THESIS OF

Junyoung Kwon for the Master of Science in Physical Science Degree, chemistry emphasis, presented on April 4, 1996.

Applications of Semiempirical Quantum Mechanical Calculations to Simple Molecular System

Abstract approved:



Quantum chemical calculations and molecular modeling have almost achieved the status of being regarded as an "experimental" tool. Selected properties of simple open- and closed-shell molecular systems containing nitrogen, sulfur, and oxygen have been studied using HyperChemTM. Structural and molecular properties are calculated in order to be able to evaluate the quality of results to be expected from this tool. Bond lengths, bond angles, dipole moments, the enthalpy of formation, and the first ionization potential have been calculated via semiempirical methods. Electron density, HOMO, and LUMO molecular orbital maps have been constructed. For the purpose of comparisons, calculations were carried out on a small sample of molecules using CNDO, INDO, MINDO/3, MNDO, and AM1 methods. All other calculations were carried out using the AM1 method.

**APPLICATIONS OF SEMI-EMPIRICAL QUANTUM
MECHANICAL CALCULATIONS TO SIMPLE MOLECULAR SYSTEMS**

A Thesis

Presented to

the Division of Physical Sciences

EMPORIA STATE UNIVERSITY

In Partial Fulfillment


of the Requirements for the Degree


Master of Science

by

Junyoung Kwon

May 1996


Approved for the Major Division


Approved for the Graduate Council

ACKNOWLEDGMENTS

I would like to express my deepest thanks to Professor Charles Greenlief for his patience, concerns, and the guidance he has provided. His help in the completion of this thesis will always be greatly appreciated. Professors Arthur Landis and Jorge Ballester read and critiqued the thesis in its entirety. I am deeply grateful for the time they invested and the suggestions they made. Lastly, I would like to thank my wife, parents, brother, and sister for their prayers and encouragement.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii	
LIST OF TABLES	vii	
LIST OF FIGURES	viii	
CHAPTER		
1	CLASSICAL MECHANICS AND THE QUANTUM REVOLUTION ... 1	
1.1	Background	1
2	QUANTUM CHEMISTRY	8
2.1	Empirical Calculation Methods	8
2.1.1	Molecular Mechanics (Empirical Force Field) Method	8
2.2	Semiempirical Molecular Orbital Theory	12
2.3	Ab Initio Molecular Orbital Theory	15
3	CHOICE OF SEMI-EMPIRICAL APPROXIMATIONS	20
3.1	The CNDO Approximation	20
3.2	The INDO Approximation	20
3.3	The MINDO/3 Approximation	21
3.4	The MNDO Approximation	22
3.5	The AM1 Approximation	23
4	HYPERCHEM	26
4.1	Background	26
4.1.1	Building and Displaying Molecules	26
4.1.2	Optimizing the Structure of Molecules	26

4.1.3	Investigating the Reactivity of Molecules	26
4.1.4	Generating and Viewing Orbitals and Electronic Plots	27
4.1.5	Evaluating Chemical Pathways and Mechanisms	27
4.1.6	Studying the Dynamic Behavior of Molecules	27
4.2	Calculations	28
4.2.1	Molecular Geometry	28
4.2.2	Geometry Optimization	29
4.2.2.1	Steepest Descent	29
4.2.2.2	Conjugate Gradient	30
4.2.2.3	Block Diagonal	30
4.2.3	Selecting Options for a Semi-Empirical Calculation	30
4.2.3.1	Choosing a Semi-Empirical Method	31
4.2.3.2	Convergence Criteria	31
4.2.3.3	Convergence Acceleration	32
4.2.3.4	Charge and Spin Multiplicity	32
4.2.3.5	Unrestricted Hartree-Fock versus Restricted Hartree-Fock	33
4.2.3.6	Log File for Results	34
4.3	Results of Semi-Empirical Calculations	34
4.3.1	Energies of Molecules	34
4.3.2	Heats of Formation	34
4.3.3	Molecular Orbital Energies and Ionization Potentials	35
4.3.4	Molecular Orbital Electron Population	36

4.3.5	Geometries of Molecules	36
4.3.6	Electrostatic Potential(Contour Plots)	36
4.3.7	Dipole Moments	37
4.3.8	Atomic Charges	38
4.3.9	A Sample Log File for SO ₂	38
5	CALCULATIONS AND RESULTS	45
5.1	Enthalpy of Formation	45
5.2	Ionization Potential	45
5.3	Dipole Moment	45
5.4	Geometrical Properties	46
5.5	Frequency Spectra	46
5.6	Selected Properties of NO and N ₂ O ₄	46
6	DISCUSSION	48
6.0	A General Comparison of Semiempirical Calculations and Experimental Results	48
6.1	Enthalpy of Formation	50
6.2	Ionization Potential	53
6.3	Dipole Moments	53
6.4	Molecular Geometries	58
6.5	Frequencies	58
6.6	Selected Properties of NO and N ₂ O ₄	63
6.7	Frequency Spectrum	66

6.8	Orbital Energy Levels for NO	66
6.9	Highest Occupied Molecular Orbital for NO	66
6.10	Lowest Unoccupied Molecular Orbital for NO	67
7	SUMMARY	72
7.1	Conclusion	72
7.2	Thoughts for the Future	73
	REFERENCES	75

LIST OF TABLES

I.	Enthalpy of Formation for Selected Oxides of Nitrogen	51
II.	Enthalpy of Formation for Selected Oxides of Sulfur	52
III.	Ionization Potential for Selected Oxides of Nitrogen	54
IV.	Ionization Potential for Selected Oxides of Sulfur	55
V.	Dipole Moments of Selected Oxides of Nitrogen	56
VI.	Dipole Moments of Selected Sulfur Containing Molecules	57
VII.	Geometrical Properties of Selected Oxides of Nitrogen	59
VIII.	Geometrical Properties of Selected Oxides of Sulfur	60
IX.	Frequency Spectra of Selected Oxides of Nitrogen	61
X.	Frequency Spectra of Selected Oxides of Sulfur	62
XI.	Selected Properties of the NO Molecule	64
XII.	Selected Properties of the N ₂ O ₄ Molecule	65

LIST OF FIGURES

1.	Frequency Spectra of (a) N_2O and (b) SO_2	68
2.	Molecular Orbital Energy Level Diagram for the NO Molecule.	69
3.	Highest Occupied Molecular Orbital of the NO Molecule.	70
4.	Lowest Unoccupied Molecular Orbital of NO the Molecule.	71

Chapter 1

CLASSICAL MECHANICS AND THE QUANTUM REVOLUTION

1.1 *Background*

In 1801 Thomas Young¹ gave convincing experimental evidence for the wave nature of light by showing that light exhibited diffraction and interference when passed through two adjacent pinholes. About 1860 James Clerk Maxwell² developed four equations, known as Maxwell's equations, which unified the laws of electricity and magnetism. Furthermore, Maxwell's equations predicted that an accelerated electric charge would radiate energy in the form of electromagnetic waves consisting of oscillating electric and magnetic fields. The speed predicted by Maxwell's equations for these waves turned out to be the same as the experimentally measured speed of light. Maxwell concluded that light is an electromagnetic wave.

In 1888 Hertz³ detected radio waves produced by accelerated electric charges in a spark as predicted by Maxwell's equations. This convinced scientists that light is indeed an electromagnetic wave. All electromagnetic waves travel at a speed of 2.998×10^8 m/sec in a vacuum. The frequency ν and wavelength λ of a wave are related by $\lambda\nu = c$. Various conventional labels are applied to electromagnetic waves depending on their frequency. In order of increasing frequency one has radiowaves, microwaves, infrared radiation, visible light, ultraviolet radiation, x-rays, and gamma rays. In the late 1800s, physicists measured the intensity of light at various frequencies emitted by a heated black-

body at a fixed temperature. (A black-body absorbs incident light. A good approximation to a black body is a cavity with a tiny hole.) When physicists used the electromagnetic wave model of light and statistical mechanics to predict the intensity-versus-frequency curve for emitted black-body radiation, they found a result in complete disagreement with the high frequency portion of the experimental curves.

In 1900 Max Planck⁴ developed a theory that gave excellent agreement with the observed black-body radiation curves. In his theory Planck assumed that the atoms of the black-body could emit light energy only in amounts given by $h\nu$, where $h = 6.634 \times 10^{-34}$ J·s. The new theory gave curves that agreed well with the experimental black-body curves. Planck's work marks the beginning of quantum mechanics.

Planck's hypothesis that only certain quantities of light energy could be emitted (i.e., that the emission was quantized) was in direct contradiction to all previous ideas of physics. The energy of a wave is related to its amplitude, and the amplitude varies continuously. Moreover, according to Newtonian mechanics, the energy of a material body can vary continuously. Hence physicists expected the energy of an atom to vary continuously. However, only with the hypothesis of quantized energy emission can one obtain the correct black-body radiation curves.

The second application of energy quantization was to the photoelectric effect. In the photoelectric effect, light shining on a metal causes emission of electrons. The energy of a wave is proportional to its intensity (i.e., amplitude) and

is not related to its frequency. Thus, the electromagnetic wave picture of light leads one to expect that the kinetic energy of an emitted photoelectron would increase as the light intensity increases, but would not change as the frequency changes. Instead, one observes that the kinetic energy of an emitted electron is independent of the light's intensity but, increases as the light's frequency increases.

In 1905 Einstein⁵ showed that these observations could be explained by viewing light as particles called photons, with each photon having an energy

$$E = h\nu \quad (1.1)$$

When an electron in a metal absorbs a photon, part of the absorbed energy is used to overcome the forces holding the electron in the metal, and the remainder appears as kinetic energy of the ejected electron. Conservation of energy gives

$$\frac{1}{2} mu^2 \leq h\nu - W \quad (1.2)$$

where u is the velocity of the ejected electron and the work function W is the minimum photon energy required to remove an electron. An increase in the light's frequency increases the photon energy and hence increases the kinetic energy of the emitted electron. An increase in light intensity at fixed frequency increases the rate that photons strike the metal, and hence increases the rate of emission of electrons, but does not change the kinetic energy of the emitted electrons. The photoelectric effect shows that light can exhibit particle-like behavior in addition to the wave-like behavior it shows in diffraction experiments.

In the late nineteenth century, investigations of electric discharge tubes and natural radioactivity showed that atoms and molecules are composed of charged

particles. Electrons have a negative charge. The proton has a positive charge equal in magnitude but opposite in sign to the electron charge, and is 1836 times as massive as the electron. The third constituent of atoms, the neutron (discovered in 1932), is neutral and slightly heavier than the proton. Starting in 1909, Rutherford, Geiger, and Marsden⁶ carried out a series of experiments passing a beam of alpha particles through a thin metal foil and observed the deflections of the particles by allowing them to strike a fluorescent screen. (Alpha particles are helium nuclei.) Rutherford observed that most of the alpha particles passed through the foil essentially undeflected, but, surprisingly, a few underwent large deflections, some being deflected backwards. To get large deflections, one needs a very close approach between the charges, so that the coulombic repulsive force is great. If the positive charge were spread throughout the atom, once the high-energy alpha particle penetrated the atom, the repulsive force would fall off, becoming zero at the center of the atom. Hence, Rutherford concluded that such large deflections could occur only if the positive charge were concentrated in a tiny, heavy nucleus.

According to Rutherford an atom contains a tiny (10^{-13} to 10^{-12} cm radius) heavy nucleus consisting of neutrons and Z protons, where Z is the atomic number. In neutral atoms, outside the nucleus there are Z electrons. The charged particles interact according to Coulomb's law. (The nucleons are held together in the nucleus by strong, short-range nuclear forces).

The chemical properties of atoms and molecules are determined by their

electronic structure, and so the question arises as to the nature of the motions and energies of the electrons. Since the nucleus is much more massive than the electron, we expect the motion of the nucleus to be slight compared to the motion of the electrons.

In 1911 Rutherford⁶ proposed a planetary model of the atom where the electrons revolved about the nucleus in various orbits, similar to the way the planets revolve about the sun. However, there is a fundamental difficulty with this model. According to classical electromagnetic theory, an accelerated charged particle radiates energy in the form of electromagnetic waves (light). An electron circling the nucleus at constant speed is being accelerated, since the direction of its velocity vector is continually changing. Hence, the electrons in the Rutherford model should continually lose energy by radiation, and therefore should spiral in toward the nucleus. Thus, according to classical physics, the Rutherford atom is unstable and would collapse.

One possible way out of this difficulty was proposed by Bohr⁷ in 1913, when he applied the concept of quantization of energy to the hydrogen atom. Bohr assumed that the energy of the electron in a hydrogen atom was quantized, with the electron constrained to move only in allowed orbitals. When an electron makes a transition from one orbit (stationary state) to another, a quantum of light of frequency

$$\nu = \Delta E/h \quad (1.3)$$

is absorbed or emitted, where ΔE is the energy difference between the two energy

states. With the assumption that an electron making a transition from a free (ionized) state to one of the bound orbits emits a photon whose frequency is an integral multiple of one-half the classical frequency of revolution of the electron in the bound orbit, Bohr derived a formula for the hydrogen-atom energy levels. Using Equation (1.3), he obtained agreement with the observed hydrogen spectrum. However, attempts to fit the helium spectrum using the Bohr theory failed. Moreover, the theory could not account for chemical bonds in molecules. The difficulty in the Bohr model arises from its inability to deal with electron-electron interactions. Atomic spectra show discrete frequencies, indicating that only certain energies are allowed. That is, the electronic energy is quantized. However, Newtonian mechanics allows a continuous range of energies. Quantization does occur in wave motion. Hence, de Broglie⁸ suggested in 1923 that the motion of electrons might have a wave aspect. That is, an electron of mass m and speed u would have a wavelength λ associated with it, such that:

$$\lambda = h/mu = h/p \quad (1.4)$$

where p is the linear momentum. De Broglie arrived at Equation (1.4) by using an analogy between electrons and photons. The energy of any particle (including a photon) can be expressed, according to Einstein's special theory of relativity, as $E = mc^2$, where c is the speed of light. Using Equation (1.1), one obtains $mc^2 = hv = hc/\lambda$. Thus, $\lambda = h/mc = h/p$ for a photon traveling at a speed of c . Equation (1.4) is then the corresponding relation for an electron.

In 1927 Davission and Germer⁹ experimentally confirmed de Broglie's

hypothesis by reflecting electrons from metals and observing diffraction effects. In 1932 Stern¹⁰ observed the same effects with helium atoms and hydrogen molecules. These results lead us to the conclusion that the wave effects are not peculiar to electrons, but result from general laws of motion for microscopic particles.

Thus, electrons behave in some respects like particles and in other respects like waves. We are faced with the apparently contradictory "wave-particle duality" of matter (and of light). How can an electron be both a particle, a localized entity, and a wave, a non-localized entity? The answer is that an electron is neither a wave nor a particle, but something else. An accurate pictorial description of an electron's behavior is impossible using the wave or particle concept of classical mechanics. The concepts of classical mechanics have been developed from experience in the macroscopic world and do not provide a proper description of the microscopic world.

Although both protons and electrons show a "duality", they are not the same kinds of entities. Photons always travel at speed c and have zero rest mass. Whereas electrons always have $u < c$ and a nonzero rest mass. Photons must always be treated relativistically, but electrons may be treated nonrelativistically when $u \ll c$.

Chapter 2

QUANTUM CHEMISTRY

2.1 Empirical Calculation Methods

Two general types of empirical methods are used. These are molecular mechanics and semiempirical. Empirical methods have several advantages over *ab initio* (from first principles) methods. First, empirical methods are fast. While speed may not be important for small molecules, it is certainly important for biomolecules. Another advantage of semiempirical methods is accuracy. For specific and well-parameterized molecular systems, an empirical method can be used to calculate values that are closer to experiment than many *ab initio* techniques. The accuracy of an empirical method depends on the database used to parameterize the method. This is true for the type of molecules and the physical and chemical data in the database. Frequently, empirical methods give the best results for a limited class of molecules or phenomena. A disadvantage of empirical methods is that parameters must be available prior to beginning a calculation. Developing parameters is not easy and is very time-consuming.

2.1.1 Molecular Mechanics (Empirical Force Field) Method

The molecular mechanics, MM (also called empirical force field), method is quite different from semi-empirical methods. The MM method is not really a quantum mechanical method, since it does not deal with an electronic Hamiltonian or wave function. Instead, the method begins with a model of a molecule as composed of atoms that interact with each other by simple analytical functions.

The method was developed by Westheimer, Hendrickson, Wiberg, Allinger, Warshel, and others.^{11,12}

The molecular electronic energy U is written as

$$U = E_{\text{str}} + E_{\theta} + E_{\text{vdW}} + E_{\text{W}} \quad (2.1)$$

The first term, the energy of bond stretching E_{str} is taken as a quadratic function of the displacement of each bond length l_i from its expected equilibrium length $l_{i,0}$ and

$$E_{\text{str}} = - \sum_i k_{s,i} (l_i - l_{i,0})^2 \quad (2.2)$$

where $k_{s,i}$ is the force constant for stretching bond l . (For small displacements of the nuclei from their equilibrium positions U is a quadratic function of $l_i - l_{i,0}$). The second term, E_{θ} , is the energy of bond bending and is

$$E_{\theta} = - \sum_i k_{\theta,i} (\theta_i - \theta_{i,0})^2 \quad (2.3)$$

where θ_i , $\theta_{i,0}$, $k_{\theta,i}$ are the angle, the equilibrium angle, and the bending force constant for bond angle l . The expected lengths and angles $l_{i,0}$ and $\theta_{i,0}$ are taken from known equilibrium geometries of small unstrained molecules. The third term, the energy of van der Waals interactions, E_{vdW} , is written as a sum of interactions between pairs of nonbonded atoms. Each pair interaction is taken as the sum of a long-range attraction and a short-range repulsion. The Lennard-Jones 6-12 potential, is generally expressed as

$$U(R) = a/R^{12} - b/R^6 \quad (2.4)$$

(where a and b are constants and R is the interatomic distance). The term E_{W} is present in molecules with internal rotation about single bonds.

If the molecule has one or more small rings, additional terms are added to account for the resulting steric strain. If the molecule has polar groups, terms are added to represent the electrostatic interactions between these groups. For aromatic compounds one also incorporates information such as bond orders taken from quantum mechanical π -electron calculations. If there is intramolecular hydrogen bonding, terms are included to allow for this.

The force constants and other parameters in U are chosen so as to give a good fit to the known geometries, energies, and vibrational spectra of small molecules. After writing down the expression for U , one varies the molecular geometry so as to locate the structure (or structures) that minimizes U . This calculation is done simply and rapidly on a computer. Molecules of one or two hundred atoms are readily handled. The method thus allows one to calculate the geometry of the various possible conformations of a molecule and also gives the energy differences between different conformations. Since the potential-energy function U for molecular vibrations is generated, one can use the MM method to calculate molecular vibrational frequencies.

Note that the zero-level of energy is taken as corresponding to all bond lengths and bond angles having their customary values l_0 and θ_0 , since there are no nonbonded van der Waals interactions or internal-rotation interactions. In such a hypothetical state, one can approximate the molecular bonding energy as the sum of empirical bond energies. To estimate the heat of formation of a gas-phase molecule from its atoms, one adds the MM-calculated equilibrium geometry energy

U_{eq} (the steric energy) to the sum of empirical bond energies for the bonds in the molecule.

To apply the MM method, one needs sufficient data to choose values for the parameters. For this reason most applications have been applied to hydrocarbon systems. Molecules with isolated polar groups are also readily treated. For hydrocarbons the MM method gives very good results. For example, ΔH_f^0 values have typical errors of one or 2 kcal/mol (as compared to an average absolute error of 6.0 kcal/mol for hydrocarbons in a modified neglect of diatomic overlap, MODO, the most accurate of the semiempirical quantum mechanical methods for ΔH_f^0 calculations). Calculated bond lengths are usually within 0.001 Å and bond angles within 2° of experimental values.

A few applications have been made to the calculation of potential-energy surfaces for chemical reactions, but the ability of MM method for kinetics is still uncertain. An important application of the MM method is to intramolecular motions in proteins. Karplus¹³ and co-workers constructed an empirical potential-energy function for a certain protein with fifty-eight amino acid residues. They then used this function to solve the classical mechanical equations of motion for the protein's atoms, using kinetic energies corresponding to a temperature of 295 K. They found that certain positions in the protein molecule underwent large-amplitude "fluid-like" motions governed mainly by collisions between nonbonded atoms. This sort of motion was later confirmed by NMR and X-ray diffraction studies.

2.2 Semiempirical Molecular Orbital Theory

The early semiempirical molecular orbital, MO, methods, including CNDO, INDO, and NDDO were developed by J. A. Pople¹⁴ and his group at a time when computers were able to handle *ab initio* calculations only for the smallest systems. These methods were not intended to reproduce molecular geometries and heats of formation, but rather other electronic properties such as the dipole moment. The simplest method, CNDO (Complete Neglect of Differential Overlap), assumes the atomic orbitals are spherically symmetrical when the electron repulsion integrals are being evaluated. The directionality of p-orbitals is included only via the one-electron resonance integrals; the size depends on the orientations and distances of the orbitals and on a constant assigned to each type of bond. The next stage, the INDO (Intermediate Neglect of Differential Overlap) approximation, included one-center repulsion integrals between atomic orbitals on the same atom. The NDDO (Neglect of Diatomic Differential Overlap) approximation was the first to include the directionality of the atomic orbitals in calculating the repulsion integrals. In this case the three- and four-center integrals where the overlap occurs between atomic orbitals on the same atom were included. More complete descriptions are given by Pople and Beveridge¹⁵ and by Dewar.¹⁶

The semiempirical methods, MINDO/3, MNDO, and AM1, are members of a series of MO techniques developed by M. J. S. Dewar and his group specifically for applications in organic research.¹⁶ The aim of their efforts was to produce a "MO spectrometer" that would eventually be able to give chemically accurate results for

large molecules at a reasonable cost in computer time. The requirements of chemical accuracy and computational economy are not normally compatible, so a number of compromises were made.

A further modification of the INDO method is the MINDO/3 method. Rather than evaluating the one-center repulsion integrals analytically, MINDO/3 uses a set of parameters to approximate them. These parameters, along with the constants used to evaluate the resonance integrals, allow the results to closely approximate the experimental data.

The MINDO/3 method is the last in a series of three MINDO methods and represents a milestone in computational chemistry. This method was the first easy-to-use program with automatic geometry optimization to be made available to a wide range of nonspecialists. MINDO/3 has often been heavily criticized, especially by the Pople school, but its importance in introducing structure and energy calculations to organic chemical research cannot be denied.¹⁷

MNDO is not a more sophisticated version of MINDO/3, but rather an independent method based on the NDDO approximation, and was under development when MINDO/3 was published. The use of NDDO, rather than INDO, was found necessary in order to avoid some systematic MINDO/3 errors. These errors occur for molecules, such as hydrazines or polyfluoroalkanes, where lone pair-lone pair repulsions are important. The directionality of the electron-electron repulsion terms in the NDDO approximation is particularly important in this respect. The advantage of MINDO/3 and MNDO over *ab initio* calculations is not only that

they are several orders of magnitude faster (MNDO is about 1.5 times slower than MINDO/3), but also that calculations for some very large molecules are possible only with the semiempirical methods. The neglect of large numbers of integrals not only saves computer time, but also reduces the core and disk space requirements in comparison with those for an equivalent *ab initio* calculation. Calculations for very much larger systems than can be managed at even the simplest *ab initio* level are therefore possible. The approximations inherent in the neglect of differential overlap methods naturally cause a loss of accuracy; this is largely compensated for by the choice of parameters. As in molecular mechanics methods the parametrization cannot be better than the available experimental data, so results for elements such as beryllium and lithium cannot be expected to be as accurate as those for hydrocarbons. Another problem with the parameterization of self-consistent fields, SCF, methods like MINDO/3 and MNDO is that experimental data include electron-electron correlation.

A better procedure is to choose parameters so that calculations plus a correlation correction will fit the experimental data. This is the philosophy behind MNDO/C.¹⁸ This method does not offer significant improvement over MNDO for ground-state configurations, but does treat excited states considerably better than the standard methods do. Configuration interaction calculations are also available in the MINDO/3 and MNDO programs, but in some ways correlation is then included twice in the calculation: once explicitly and once indirectly via the parametrization. The results are therefore meaningful only when compared with each other, and not

as absolute values, for example, heats of formation.

Many problems with MNDO involve cases where the NDO approximation for electron-electron repulsion is most important. The AM1 method is an improvement over MNDO, even though it uses the same basic approximation. It is generally the most accurate semiempirical method and is the method of choice for most problems. Altering part of the theoretical framework (the function describing repulsion between atomic cores) and assigning new parameters improves the performance of AM1. The method deals with hydrogen bonds properly, produces accurate predictions of activation barriers for many reactions, and predicts heats of formation of molecules with an error that is about 40 percent smaller than with MNDO.¹⁹

2.3 *Ab Initio* Molecular Orbital Theory¹⁹

The term *ab initio* implies a rigorous, nonparameterized molecular orbital treatment derived from first principles. This is not completely true. A number of simplifying assumptions are used in *ab initio* theory as it is currently implemented; but the calculations are more complete, and therefore more expensive with respect to computer time, than those of the empirical or semiempirical methods. It is possible to obtain chemical accuracy via *ab initio* calculations, but the cost in computer time is enormous, and only small systems can be treated this accurately at present. In practice most calculations are performed at lower levels of theory than would be considered definitive, and then the shortcomings are taken into account when higher level calculations are performed. Like the semiempirical

calculations, *ab initio* theory makes use of the Born-Oppenheimer approximation. The nuclei are regarded as fixed on the time scale of electron movement, i.e., the electronic wave function is unaffected by nuclear motion. This is a very good approximation in almost all cases. Only for an extremely flat potential surface, as in some Jahn-Teller systems for instance, does significant coupling exist between the vibrational and electronic wave functions.

In contrast to the NDDO methods, many different possible choices for the sets of atomic orbitals (the *basis set*) are available. Almost all modern *ab initio* calculations employ *Gaussian type orbital* (GTO) basis sets. These basis sets, where each atomic orbital is made up of a number of Gaussian probability functions, have considerable advantages over other types of basis sets for the evaluation of one- and two-electron integrals. They are much faster computationally than equivalent Slater orbitals, for instance. The GAUSSIAN series of programs deal, as the name implies, exclusively with Gaussian-type orbitals and include several optional GTO basis sets of varying size. The use of Gaussians is one of the main advantages of such a widely distributed program system. The methods and basis sets used almost become a standard, and a direct comparison with other literature data is often possible.

The simplest of the optional basis sets is the STO-nG basis set, where STO-3G is the only one to have found wide use, although the basis sets STO-2G to STO-6G were originally tested.²⁰ The STO-nG is an abbreviation for Slater-Type-Orbitals simulated by n Gaussian functions. This means that each atomic orbital

consists of n Gaussian functions added together. The coefficients of the Gaussian functions are selected to give an optimum fit to the corresponding Slater-type orbitals. The STO-3G uses a *minimal basis set*. This means that it has only as many orbitals as are necessary to accommodate the electrons of the neutral atom.

The greatest problem of any minimal basis set is its inability to expand or contract its orbitals to fit the molecular environment because the exponent is fixed. One solution to the problem is to use *split-valence* or *double zeta* basis sets. In these basis sets the atomic orbitals are split into two parts: an inner, compact orbital and an outer, more diffuse one. The coefficients of these two types of orbitals can be varied independently during construction of the molecular orbitals in the self consistent field (SCF) procedure. Thus, the size of the atomic orbital that contributes to the molecular orbital can be varied within the limits set by the inner and outer basis functions. Split-valence basis sets split only the valence orbitals in this way, whereas double zeta basis sets also have split-core orbitals ("double zeta" simply implies two different exponents).

The split-valence basis set most widely used for the early calculations was 4-31G. This nomenclature means that the core orbitals consist of 4 atomic orbitals, and the inner and outer valence orbitals of 3 and 1 Gaussian functions, respectively. The advent of optimization procedures that use an analytical gradient led to the development of split-valence basis sets with fewer primitive Gaussians than 4-31G. The basis set used most commonly for geometry optimizations is now 3-21G. It uses three primitive Gaussians for the core orbitals and a two/one split

for the valence functions.

The next step to improve a basis set is usually the addition of d-orbitals for all heavy (nonhydrogen) atoms. For most organic compounds these do not function as d-orbitals in the normal sense of being involved in bond formation as in transition-metal compounds. Their purpose is to allow a shift of the center of an orbital, a p-orbital for instance, away from the position of the nucleus. Mixing the d-orbital with the p-orbital results in a deformation of the resulting orbital to one side of the atom. This adjustment is particularly important for compounds containing small rings and for compounds of the second-row elements.

The most commonly used *polarization basis set* (i.e., involving d-orbitals) is 6-31G*. This basis set uses six primitive Gaussians for the core orbitals, a three/one split for the s- and p-valence orbitals, and a single set of six d-functions (indicated by the asterisk). Six d-functions (equivalent to five d- and one s-orbital) are used for computational convenience, although the GAUSSIAN programs can also handle basis sets with five real d-orbitals. A further development is the 6-31G** basis set. A set of p-orbitals has been added to each hydrogen in the 6-31G* basis set.

One additional type of basis set, the *diffuse function* augmented basis, is intended for use in calculations on anions or molecules that require very good descriptions of nonbonding electron pairs. These basis sets are obtained by adding a single set of very diffuse s- and p-orbitals to the heavy atoms in a standard basis such as 6-31G*. The purpose of the diffuse functions is to improve the basis set

at large distances from the nucleus, and thus to better describe the high-energy electron pairs associated with anions.

The number of basis functions rises rapidly with an increasing sophistication of the basis set. This is important for two reasons: first, because the number of basis functions the program can handle is limited, and second, because the computer time required is approximately proportional to the fourth power of the number of basis functions.

After selection of an appropriate basis set for the problem is made (usually the largest practical), the type of calculation to be performed is chosen. The normal practice for *ab initio* calculations differ from that used with semiempirical methods. Almost all calculations are performed with full geometry optimization. Because of the cost of *ab initio* calculations and the possibility of improving the basis set, it is often necessary to optimize the geometry with a small basis set and then perform single-point calculations with a better basis set or with a correction for electron correlation. This practice allows energy calculations at levels of theory that are too high for practical full geometry optimization. Usually the results are good approximations to those that would be obtained through a full geometry optimization at the higher level. However, in some cases the inclusion of d-orbitals can lead to large changes in the structure so that single-point calculations on a geometry obtained with a nonpolarization basis set may sometimes be misleading.

Chapter 3

CHOICE OF SEMI-EMPIRICAL APPROXIMATIONS

3.1 The CNDO Approximation¹⁶

The CNDO method is based upon a complete neglect of the differential overlap approximation. This method is the simplest and the least accurate of the ZD) (Zero Differential Overlap) methods. In a ZDO computation, the overlap matrix is assumed to be the unit matrix. In the CNDO equations the Fock matrix, $H_{\mu\nu}$, consists of two elements-diagonal elements, $H_{\mu\mu}$, and off-diagonal elements, $H_{\mu\nu}$. The elements $H_{\mu\mu}$ represent the kinetic energy of an electron in an atomic orbital that is closely related to the valence state ionization energy, and the potential energy of the electron. Electron affinity is included within CNDO/2, which did not appear in the CNDO/1 method. The $H_{\mu\nu}$ represent the dominant effects of bonding, and consists of an overlap integral that goes with a parameter dependent on the two atoms involved in the overlap. As an application one can expect to compute expectation values of observables by using the zero differential overlap approximation. But, in the case of the dipole moment, such a strict ZDO method gives the same value as does the classical computation with two point charges. Therefore, the original CNDO authors¹⁶ included one-center hybridization terms that made significant contributions to the dipole moment.

3.2 The INDO Approximation¹⁶

The INDO (Intermediate Neglect of Differential Overlap) differs from CNDO in the treatment of one-center exchange integrals. The CNDO approximation

retains only the two-electron interactions, that contain the repulsive coulombic force between an electron on atom A and an electron on atom B. But INDO considers not only the repulsive force between two electrons, but also the spin states of electrons that reside in atomic orbitals on the same atom. Two examples are associated with it. One is the exchange phenomena. In order to take at least minimal account of electron exchange, the INDO procedure restores integrals based on monatomic differential overlap. Two electrons in any set of orbitals on the same atom can have their spin either parallel or antiparallel. In the case of parallel spins, the atom is in a triplet state. But with antiparallel spins, the spin state is a singlet. One can understand the difference in energy of two states by use of Hund's rules.

The INDO procedure restores the major exchange interactions neglected in CNDO theory and adds spin interactions. In the diatomic molecule NH, two spectroscopic states result due to the two electrons in degenerate p-orbitals centered on the nitrogen. CNDO gives one energy state because it treats only repulsive coulombic force between the electrons, but INDO yields two different energy states because of spin interactions. INDO calculations have more parameters and are somewhat more complex than CNDO calculations, but they require essentially no extra computation time, and in most situations the results are superior to CNDO calculations.

3.3 The MINDO/3 Approximation²¹

MINDO/3 stands for Modified Intermediate Neglect of Differential Overlap,

version 3; it is a modification of the INDO method. Two elements of the MINDO/3 UHF (Unrestricted Hartree Fock) matrix are off-diagonal elements, $F_{\mu\nu}$, and diagonal elements, $F_{\mu\mu}$. The two-center two-electron, one-center two-electron, two-center one-electron, one-center one-electron, and core-core repulsion integrals involved in the equations for matrix elements are as follows. Two-center two-electron integrals are set equal to an averaged electrostatic repulsion between an electron on atom A and an electron on atom B. The one-center one-electron integral represents the energy that an electron in an atomic orbital would have if all other valence electrons were removed to infinity. This is calculated by adding to the one-electron energy of the atomic orbital in the fully ionized atom the potential due to all the other nuclei in the system. The two-center one-electron integral is approximated by using the overlap integral. This integral includes the ionization potentials of the atomic orbitals on two different atoms.

One-center two-electron integrals are derived from an analysis of atomic spectra, and are written in terms of the Slater-Condon parameters. Core-core repulsion integrals in MINDO/3 take into account not only electron-electron and electron-core coulombic interactions for distances larger than the van der Waals radii, but also the decreased screening of the nucleus by the electrons as the interatomic distance becomes very small.

3.4 The MNDO Approximation ²²

The MNDO method is a Modified Neglect of Diatomic Overlap method based on the NDDO (Neglect of Diatomic Differential Overlap) approximation. The central

assumptions made in the NDDO approximation are the core approximation, and the neglect of electron repulsion integrals involving diatomic differential overlap. The core approximation treats atoms in terms of a core composed of the nucleus and tightly bound inner electrons, together with a valence shell of electrons that are less tightly bound. This of course greatly reduces the number of electrons that must be considered in treating molecules. MNDO was introduced to correct some of the problems with MINDO/3. In MINDO/3, the calculated heats of formation of aromatic hydrocarbons are consistently too positive and those of compounds with triple bonds too negative. Both of these deficiencies are overcome in MNDO.

Bond angles calculated with MINDO/3 are often subject to quite large errors. The MNDO values are much better, particularly for angles centered at carbon or nitrogen. In addition, the ordering of MO's in MNDO agrees much better with that deduced from photoelectron spectroscopy than that from MINDO/3. Therefore, MNDO has been used widely to calculate heats of formation, molecular geometries, dipole moments, ionization energies, electron affinities, and other properties. But it has problems dealing with sterically crowded molecules (too unstable), four-membered rings (too stable), hydrogen bonding (almost nonexistent), and hypervalent compounds (too unstable). For nitrobenzene this method incorrectly yields an out-of-plane nitro group, and a peroxide bond that is too short by about 0.17 Å.

Although AM1 is generally a significant improvement over MNDO, MNDO gives better results for some molecules, such as phosphorus compounds.

3.5 The AM1 Approximation ²³

The AM1 (Austin Model 1) is a modified MNDO method. While MINDO/3 has proved very effective in studies of a wide variety of hydrocarbons, problems arise in the case of molecules containing heteroatoms because of the neglect of one-center differential overlap in the INDO approximation on which MINDO/3 is based.

These problems are avoided in MNDO but other problems are created. In particular MNDO fails to reproduce hydrogen bonds correctly, gives energies that are too positive for crowded molecules (e.g. neopentane) and too negative for ones containing four-membered rings, and gives activation energies that tend to be too large. The error appears because MNDO has a tendency to overestimate the repulsions between atoms when the atoms are at a separation distance approximated by their van der Waals radii.

The obvious way to deal with this is to modify the CRF (Core Repulsion Function) in MNDO. The only difference in the core-core repulsion of the AM1 method from those in the MNDO method is in the last term. The extra terms in the AM1 core-core repulsion define spherical Gaussian functions. Between two and four Gaussian functions per atom are used in AM1. These are the only differences between the MNDO and AM1 functional forms.

The major gains in AM1 over MNDO are the ability to calculate hydrogen bond energies, activation energies for reactions, and heats of formation of molecules with an error that is about 40 percent smaller than with MNDO. Part of the improvement is due to the fact that a better minimum is found, corresponding

in particular to different orbital exponents. This has a large effect on activation barriers and to the ratios of the B parameters for s and p atomic orbitals. This appears to control the bond angles. (The B parameters are s and p atomic orbital one-electron two-center resonance integral terms.) With AM1 the computation time compared with MNDO is not changed significantly.

Problems still exist with AM1. Treatments of phosphorus-oxygen bonds are inaccurate, nitro compounds still have energies that are too positive, and the peroxide bond is still too short.¹⁹

Chapter 4

HYPERCHEM™

4.1 Background

HyperChem™ is a versatile molecular modeling program and a powerful computational program. The program allows one to easily perform many types of molecular and quantum mechanical calculations.

4.1.1 Building and Displaying Molecules

With this program one can draw a two-dimensional (2-D) representation of a molecule, and then generate a three-dimensional (3-D) structure. Manipulation of individual bonds, bond geometries, angles, torsions, and atomic charges is also possible.

4.1.2 Optimizing the Structure of Molecules

A well-defined molecular structure is necessary prior to the calculation of the properties of a molecule. Calculations usually require a structure that can be represented by a minimum on a potential energy surface. The program contains several geometry optimizers to do this. The optimized structure is then used as a starting point for subsequent calculations, such as molecular dynamics simulations.

4.1.3 Investigating the Reactivity of Molecules

HyperChem™ can be used to investigate the reactivity of molecules and their functional groups. One method is to use Frontier Molecular Orbital Theory.²⁴ Molecular orbital energies, coefficients, and nodal properties from single-point quantum mechanical calculations can be used to investigate concepts such as the

relative reactivity of different molecular substituents, regioselectivity of reactions, and site-selectivity of nucleophiles and electrophiles.

4.1.4 Generating and Viewing Orbitals and Electronic Plots

HyperChem™ can be used to plot orbital wave functions resulting from semi-empirical, single-point calculations. It is interesting to view both the nodal properties and the relative size of the wave functions. Orbital wave functions can provide chemical insights. It is also possible to plot the electrostatic potential, the total charge density, or the total spin density determined during a single point semi-empirical calculation. This information is useful in determining reactivity and correlating calculated results with experimental data.

4.1.5 Evaluating Chemical Pathways and Mechanisms

Calculating single-point properties and energies provide information about chemical pathways and mechanisms. Semi-empirical methods can be used to calculate possible reaction pathways for the chemical reactions of molecules. Atomic charges, HOMO electron densities, and LUMO electron densities for the interacting molecules are calculated to choose possible interaction sites. Using this information possible reaction pathways are proposed by the program.

4.1.6 Studying the Dynamic Behavior of Molecules

A molecular system at room temperature is accurately characterized by its motion. Molecular-dynamics simulations calculate the future positions and velocities of atoms based upon their current values. Qualitative and quantitative data can be obtained from HyperChem™ molecular dynamics simulations.

The greatest values of molecular dynamics simulations are that they complement and help to explain existing data, or to design new experiments. These simulations are also useful for structural refinement of models generated from NMR, distance geometry, and X-ray data.

4.2 Calculations

Three steps are used in carrying out any quantum mechanical calculation in HyperChem™. First, one prepares a molecule with an appropriate starting geometry. Second, one chooses a calculation method and its associated options. Third, one chooses the type of calculation (single point, geometry optimization, or molecular dynamics) and the relevant options.

4.2.1 Molecular Geometry

The most important chemical structures are stable, equilibrium molecular geometries and transition intermediates. The equilibrium geometry of a molecule (bond lengths and angles) describes the coordinates of a deep minimum on a potential energy surface. A set of atoms may have a number of potential energy minima, each corresponding to a different isomer of the molecular system. Other, less-deep minima may correspond to chemical reaction intermediates.

Using the coordinates of geometries and minima, together with the nearby values of potential energy, it is possible to calculate spectroscopic properties and macroscopic thermodynamic and kinetic parameters, such as enthalpies, entropies, and rate constants. HyperChem™ can provide the geometries and energy values for many of these calculations.

4.2.2 Geometry Optimization

Through the use of HyperChem™ one can calculate geometry optimizations (minimizations) utilizing either molecular or quantum mechanical methods. Geometry optimizations determine the coordinates of a molecular structure that represent a potential energy minimum. For a potential energy U and Cartesian coordinates r_i , the optimized coordinates satisfy the equation:

$$\partial U / \partial r_i = 0 \quad (4.1)$$

The goals of performing a geometry optimization calculation are as follows: to characterize a potential energy minimum (a geometry optimization results in a new structure at a minimum; examination of atomic coordinates and energy of this structure can be done); to obtain a new stable structure as a starting point for a single point, semi-empirical calculation (this provides a large set of structural and electronic properties.); and to prepare a molecule for a molecular dynamics simulation (HyperChem™ supplies three types of optimizers or algorithms: steepest descent, conjugate gradient, and block diagonal).

4.2.2.1 Steepest Descent

The steepest descent method is a first-order minimizer. It uses the first derivative of the potential energy with respect to the Cartesian coordinates. The method moves down the steepest slope of the interatomic forces on the potential energy surface. The descent is accomplished by adding an increment to the coordinates in the direction of the negative gradient of the potential energy, or the force. In the steepest descent method large forces on atoms are rapidly alleviated.

This is especially useful for eliminating the large nonbonded interactions often found in initial structures. Each step in a steepest descent requires minimal computing time. Its disadvantage is that convergence toward a minimum is very slow.

4.2.2.2 Conjugate Gradient

A conjugate gradient method differs from the steepest descent technique by using both the current gradient and the previous search direction to drive the minimization. A conjugate gradient method is also a first-order minimizer. The advantage of a conjugate gradient minimizer is that it uses the minimization history to calculate the search direction, and converges faster than the steepest descent technique. HyperChem™ provides two versions of the conjugate gradient method: Fletcher-Reeves and Polak-Ribiere. Polak-Ribiere is more refined and is the default choice.

4.2.2.3 Block Diagonal

The Newton-Raphson block diagonal method is a second-order optimizer. It calculates both the first and second derivatives of potential energy with respect to Cartesian coordinates. These derivatives provide information about both the slope and curvature of the potential energy surface. Because of its neglect of off-diagonal blocks, this optimizer can sometimes oscillate and fail to converge.

4.2.3 Selecting Options for a Semi-Empirical Calculation

First, a method for a semi-empirical calculation is chosen. There are two basic versions for the semi-empirical method: one of these is extended Hückel

theory; the others are five self consistent field (SCF) methods. After choosing a semi-empirical method, selecting options associated with the method is necessary to set additional conditions for the calculation. Since extended Hückel theory uses the "one-electron" simplification that assumes no explicit interactions between electrons and is not a self consistent field method, it has no options such as convergence limit, iteration limit, accelerate convergence, and configuration interaction to be determined before performing an SCF calculation.

4.2.3.1 Choosing a Semi-Empirical Method

HyperChem™ supports one independent-electron method (extended Hückel theory), and five semi-empirical SCF methods-CNDO, INDO, MINDO/3, MNDO, and AM1. The choice of the semi-empirical method depends on several factors including experience and user preferences. The same semi-empirical method should be used to compare the results to other studies. Since some methods converge much more quickly than others, a fast method can be used to obtain an approximate solution, and then a more accurate method can be used for the final results.

4.2.3.2 Convergence Criteria

In setting up an optimization calculation, two convergence criteria can be used: the root-mean-square (RMS) gradient or the number of optimization cycles. Suitable default values for ending an optimization calculation are either an RMS gradient of 0.1 kcal/(mol-Å), or a maximum number of cycles that are 15 times the number of atoms involved in the calculation. In general, a gradient limit should be used. For improved precision, a lower gradient limit should be used.

The convergence limit and the iteration limit specify the precision of the calculation. Convergence limit refers to the difference in energy (kcal/mol) between two successive SCF cycles. Iteration limit specifies the maximum number of cycles allowed to reach that goal.

4.2.3.3 Convergence Acceleration

If the calculation exceeds the iteration limit before it reaches the convergence limit, then there is most likely convergence failure. Simply increasing the limit is unlikely to help. Most often, however, the reason for convergence failure lies in the molecular geometry. Some systems converge poorly, particularly those with multiple bonds or weak interactions between open-shell systems. HyperChem™ includes two convergence accelerators. One is the default convergence accelerator, effective in speeding up normally convergent calculations. The other, called the direct inversion in the iterative sub-space method, is available only for MINDO/3, MNDO, and AM1. This accelerator may be helpful in curing convergence problems. It often reduces the number of iteration cycles required to reach convergence.

4.2.3.4 Charge and Spin Multiplicity

Before any semi-empirical calculations can begin, the charge and spin multiplicity of the system under investigation must be specified. HyperChem™ semi-empirical calculations must start with the number of valence electrons (N) and the number with alpha spins (the remaining electrons have beta spins). The alpha electrons have spin of $+\frac{1}{2}$ (spin up) and the beta electrons have a spin of $-\frac{1}{2}$ (spin

down). The program determines this information from the charge and spin multiplicity. N is then computed by counting the electrons associated with each neutral atom and subtracting the charge. The total spin multiplicity is $2S+1$, where S is the total electron spin of the system. Each unpaired electron has a spin vector of $\frac{1}{2}$. Paired electrons have a net spin of zero ($+\frac{1}{2} - \frac{1}{2} = 0$). A closed-shell system has a multiplicity of 1 (singlet). An open-shell system can have a multiplicity of 2, 3, 4 (doublet, triplet, quartet, respectively), or higher.

4.2.3.5 Unrestricted Hartree-Fock versus Restricted Hartree-Fock

Semi-empirical calculations use one of two forms of the wave function: RHF (Restricted Hartree-Fock) or UHF (Unrestricted Hartree-Fock). The RHF wave function is used for singlet electronic states, such as the ground states of stable organic molecules. Spin orbitals are grouped in pairs for an RHF calculation. Each member of the pair differs in its spin function (one alpha and one beta), but both must share the same space function. For N electrons, $N/2$ different molecular orbitals (space functions) are doubly occupied, with one alpha and one beta electron forming a pair.

The UHF wave function is most often used for multiplicities greater than singlets. In this case, the space orbitals for alpha electrons need not be the same as for beta electrons. For open-shell systems, some of the spins are unpaired, and alpha and beta electrons occupy different orbitals. This is because there are more alpha electrons than beta electrons, and the alpha electrons tend to repel each other more than they do beta electrons. The UHF wave function can also apply to

singlet molecules. Usually the results are the same as for faster RHF method.

4.2.3.6 Log File for Results

After the computation method and options are chosen, Start log on the File menu can be used to record calculated results, such as enthalpies of formation (for the MINDO/3, MNDO, and AM1 methods), dipole moments, atomic charges, and ionization potentials. When chemical calculations are completed, Stop log on the File menu turns off the log for recording results in a log file.

4.3 Results of Semi-Empirical Calculations

4.3.1 Energies of Molecules

The total energy in a molecular orbital calculation is the net of electronic kinetic energies and the interactions between all electrons and atomic cores in the system. This interaction energy is the potential energy for nuclear motion in the Born-Oppenheimer approximation. The total energy can also be obtained by taking the sum of isolated atomic energy and binding energy. The stable geometry of a molecule has a minimal total energy. Geometries at different energy minima (local minima plus the global minimum) describe different stable or metastable conformations and isomers of a molecule. A common application of the calculation of molecular energy is the study of organic reaction mechanisms. The energies of different potential intermediates are species that cannot be easily studied by experiment can be investigated. Many such applications include studies of rearrangements, questions of regioselectivity, stereoselectivity, and photochemical reactions.

4.3.2 Heats of Formation

The heats of formation resulting from the semi-empirical calculations use atomic heats of atomization. The first step is to calculate the Hartree-Fock total energy of the system as predicted by HyperChem™. The atomization energy is then computed by subtracting the total energies of the atoms (as predicted by the semi-empirical method being used) in their stoichiometric ratios. The heat of formation, ΔH_f^0 , of the system can then be calculated using the heats of atomization of the atoms and the atomization energy. MINDO/3, MNDO, and AM1 are parameterized by fitting to experimentally determined heats of formation for a set of molecules at 298 K. The energetics of chemical equilibrium can be investigated by comparing the heats of formation of reactants and products. This produces one of the most useful results of a chemical calculation. The accuracy and reliability of the heats of formation depend on the method used.

4.3.3 Molecular Orbital Energies and Ionization Potentials

For multielectron atoms and molecules, Koopmans' theorem¹⁹ states that the energy needed to remove an electron from an orbital of a closed-shell atom or molecule is approximated by the negative of the Hartree-Fock orbital energy. Thus, the first ionization energy of an atom or molecule can be estimated by taking the negative of the energy of the highest occupied molecular orbital. During ionization, the remaining electrons are reorganized, contributing an additional energy term not considered in this theorem. However, Koopmans' theorem does hold for many situations and enables interpretation of photoelectron spectra by molecular orbital

calculations. It is also true that orbital energies can reasonably approximate higher ionization potentials.

4.3.4 Molecular Orbital Electron Population

An orbital population represents the charge density for each molecular orbital in the molecule that is the result of a Hartree-Fock calculation with atomic orbitals. According to a widely used method to analyze LCAO-MO SCF wave functions, the electron probability density for a molecular orbital is the sum of the net population in the atomic orbitals from which the molecular orbital is constructed, and the overlap population from pairs of atomic orbitals.¹⁶

4.3.5 Geometries of Molecules

The geometries obtained from optimization calculations with semi-empirical methods describe the shapes of molecules that have energy minima. The quantum mechanical methods have varying degrees of accuracy and take more time than molecular mechanics methods. The accuracy of the results depends on the molecule. Molecular mechanics force fields have much information built into them and can be accurate for the molecules used in their parameterization. For molecules that were not used in the parameterization, semi-empirical methods are more reliable.

4.3.6 Electrostatic Potential (Contour Plots)

Electron distribution governs the electrostatic potential of molecules. The electrostatic potential describes the interaction energy of the molecular system with a positive point charge. The electrostatic potential is useful for finding sites of

reaction in a molecule: positively charged species tend to attack where the electrostatic potential is strongly negative (electrophilic attack).

HyperChem™ displays the electrostatic potential as a contour plot. Two-dimensional contour plots for a number of variables can be displayed by HyperChem™. To obtain these contour plots, the user needs to specify three things.

The first specification is the plane to be contoured. The plane is specified as a plane parallel to the screen and offset from the center of mass of the current atomic or molecular system by a user-specified number of angstroms. For example, if an offset of 1.0 is selected, then the plane is 1.0 Å closer to the viewer than an offset of 0.0 Å. A negative offset can also be specified that moves the plane away from the viewer.

The second specification is the number of contours to be shown. One can specify the total number of contour lines, n , to be shown by simply stating the number. This number must be equal to or greater than one. The values of the contour lines are normally specified as the default values. In this case the maximum and minimum values on the grid are computed, and then contours are drawn at these values plus $n - 2$ contour lines evenly spaced in between the maximum and minimum value.

The third specification is the number and position of the grid of points on the plane. The size and number of points on the grid should be the same as those specified before computation is exactly the same as those after computing the

values of spatial variables (a function of x, y, and z in three-dimensional Cartesian space).

4.3.7 Dipole Moments

The dipole moment of a molecular system is the sum of all of the multipole moments that exist in the molecule. Therefore, the accurate quantification of the overall distribution of electrons in a molecule is difficult. But the molecular dipole moment represents the charge density in a molecule and can be measured by experiment.

4.3.8 Atomic Charges

A clear relationship exists among energy, geometry, dipole moment, and the electrostatic potential and their experimental values. Calculated atomic charges are a different matter. Atomic charges can be defined in various ways. Mulliken atomic charges are commonly used in molecular orbital theory and are used in HyperChem™. Only an approximate relationship exists between these quantities and experiment. That is, they have varying values according to the basis set and the method of calculation. One of the common uses of Mulliken atomic charges is to indicate chemical reactivity.¹⁶

4.3.9 A Sample Log File for SO₂

A typical log file in HyperChem™ consists of seven parts. The first part (labeled 1 on page 42) shows the options selected for a semi-empirical calculation. The semi-empirical method chosen for calculation was AM1. The Polak-Ribiere method, the default choice in HyperChem™, was selected for the geometry

optimization calculation. The convergence accelerator was not used since the system is relatively simple. As a criteria for terminating the iterations, the convergence limit was set as 0.001 kcal/mol for the difference in binding energy between two successive SCF cycles, and the RMS gradient was selected as 0.01 kcal/(Å mol) for the first derivative of the total energy with respect to displacement of coordinates for each atom. Because SO₂ is a closed-shell system, RHF wave functions were used and the multiplicity was set to one. Since SO₂ is a neutral molecule, the charge on the system was chosen as zero.

Part two (labeled 2 on page 42) shows a portion of the progress of the SCF calculations. The last two lines show that the two convergence criteria were simultaneously met, that is, the difference between two successive calculations of the binding energy is approximately zero (less than 1×10^{-5}). This value is smaller than 0.001 and the difference between the RMS gradients is 0.006. This is smaller than 0.01.

Part three (labeled 3 on page 43) represents several calculated energies, the heat of formation, and the gradient where the SCF MO calculation was terminated. The total energy for SO₂ is expressed in different sets of units. The total energy can be expressed either as the sum of electronic energy and core-core interaction energy, or as the sum of binding energy and isolated atomic energy. Electronic energy is the sum of kinetic energy and potential energy of the electrons. Core-core interaction energy is the sum of electrostatic energies of three repulsive nuclei, i.e., there are one sulfur and two oxygen nuclei. Binding energy is the

energy decrease when forming a stable molecule from its constituent atoms. The heat of formation is calculated by subtracting atomic heats of formation from the total energy.

Part four (labeled 4 on page 43) represents the energy for each of the molecular orbitals, calculated by the SCF method. The ionization potential of a molecule is automatically known by taking the negative of the energy of the HOMO by use of Koopmans' theorem. The total number of molecular orbitals used in the calculation is twelve because the twelve valence atomic orbitals that came from one s- and three p-orbitals for each of the three atoms are used in the LCAO method. The molecular orbitals are arranged in increasing order of energy from the lowest occupied molecular orbital to the highest.

Part five (labeled 5 on page 43) shows the electron density for each of the molecular orbitals. It can be inferred that the electrons are distributed in each molecular orbital and no molecular orbital is occupied by more than two electrons. The fifth and ninth (HOMO) molecular orbitals have almost the same electron populations and have the greatest probability density.

Part six (labeled 6 on page 44) presents the atomic charges and coordinates. It shows that the charge density between oxygen and sulfur atoms was displaced toward the oxygen atoms because the oxygen atoms have a larger electronegativity than the sulfur atom. The atomic coordinates show that SO_2 is a planar, bent molecule. Therefore, SO_2 has a permanent dipole moment.

Part seven (labeled 7 on page 44) shows the dipole moment that was

calculated by the algorithms embedded in HyperChem™. The resultant dipole moment is the sum of the dipole moments calculated separately using a point-charge model and an sp hybrid model. The former regards the atoms as mathematical points that have no volume and idealizes the situation. But the latter deals with the real situation and the electron densities are spread in hybridized atomic orbitals that contribute to the total dipole moment by interacting with three nuclei. The contributions from d atomic orbitals were neglected because there are no valence electrons in these states.

HyperChem™ log start -- Tue Apr 18 11:31:26 1995.

1. Geometry optimization, Semi-empirical, molecule = B:\SO2-AM1.HIN.

AM1

Polak-Ribiere optimizer

Convergence limit = 0.0010000 Iteration limit = 50

Accelerate convergence = NO

Optimization algorithm = Polak-Ribiere

Criterion of RMS gradient = 0.0100 kcal/(A mol) Maximum cycles = 45

RHF Calculation:

Singlet state calculation

Number of electrons = 18

Number of Double Occupied Levels = 9

Charge on the System = 0

Total Orbitals = 12

2. Starting AM1 calculation with 12 orbitals

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=1 Diff=1738.09771]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=2 Diff=18.46257]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=3 Diff=3.09024]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=4 Diff=0.36088]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=5 Diff=0.10200]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=6 Diff=0.13799]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=7 Diff=0.08441]

E=0.0000 Grad=0.000 Conv=NO(0 cycles 0 points) [Iter=8 Diff=0.00006]

E=-232.5508 Grad=0.011 Conv=NO(0 cycles 1 points) [Iter=1 Diff=0.00000]

E=-232.5508 Grad=0.011 Conv=NO(0 cycles 2 points) [Iter=1 Diff=0.00000]

E=-232.5508 Grad=0.012 Conv=NO(0 cycles 3 points) [Iter=1 Diff=0.00000]

E=-232.5508 Grad=0.006 Conv=YES(1 cycles 4 points) [Iter=1 Diff=0.00000]

3. ENERGIES AND GRADIENT

Total Energy -19233.2466830 (kcal/mol)

Total Energy -30.649517602 (a.u.)

Binding Energy -232.5508310 (kcal/mol)

Isolated Atomic Energy -19000.6958520 (kcal/mol)

Electronic Energy -37866.9049709 (kcal/mol)

Core-Core Interaction 18633.6582879 (kcal/mol)

Heat of Formation -47.0328310 (kcal/mol)

Gradient 0.0065248 (kcal/mol/Å)

4. EIGENVALUES(eV)

-37.380886 -37.206879 -19.834658 -17.045744 -16.717564

-16.005463 -12.750371 -12.376399 -10.488334 -1.045624

2.043951 2.153289

5. ATOMIC ORBITAL ELECTRON POPULATIONS

1.714532 0.793066 1.148324 0.875611 1.931731

1.492261 1.748047 1.562195 1.931730 1.513128

1.727181 1.562194

6. NET CHARGES AND COORDINATES

Atom	Z	Charge	Coordinates(Å)		
			x	y	z
1	16	1.468467	-0.30836	0.16045	0.00000
2	8	-0.734233	-1.48125	-0.65622	0.00000
3	8	-0.734233	0.82926	-0.70467	0.00000

7.	Dipole (Debyes)	x	y	z	Total
	Point-Charge	0.124	5.931	0.000	5.932
	sp Hybrid	-0.034	-1.640	0.000	1.640
	pd Hybrid	0.000	0.000	0.000	0.000
	Sum	0.090	4.291	0.000	4.292

HyperChem log stop -- Tue Apr 18 11:32:57 1995.

Chapter 5

CALCULATIONS AND RESULTS

5.1 Enthalpy of Formation

The average error in the enthalpies of formation for the oxides of nitrogen and the oxides of sulfur are ± 16 kcal/mol and ± 19.3 kcal/mol, respectively. Considering these values, the calculated results are not as good as expected. The typical uncertainty for AM1 semi-empirical calculations for all types of neutral closed-shell molecules is ± 9 kcal/mol.

Among the oxides of nitrogen, the enthalpies of formation for the NO, NO₂, and N₂O₄ molecules have the largest differences between the observed and calculated values. Two of these molecules have open-shell configurations. In the case of the oxides of sulfur, SO has the largest difference between the observed and calculated values. Dewar²⁵ recognized in 1985 that AM1 has problems with diatomic molecules. In general these results show that AM1 still has problems with nitrogen and sulfur containing compounds.

5.2 Ionization Potential

The average errors in the ionization potentials for the oxides of nitrogen and the oxides of sulfur are 3.12 eV and 2.88 eV, respectively. In the case of the oxides of nitrogen, the largest deviations are for NO and NO₂, which are open-shell molecules. These differences far exceed the average error of 0.61 eV reported by Stewart²⁶ for 256 compounds.

5.3 Dipole Moment

The average error in the calculated dipole moments for the various oxides of nitrogen is ± 0.73 D, but for the oxides of sulfur the average error is ± 1.91 D. The relative errors for N_2O_5 and SO_2 are the largest. Stewart²⁷ reported an average error of 0.35 D.

5.4 Geometrical Properties

In most cases the average error for bond lengths is $\pm 0.050\text{\AA}$ and for bond angles the average error is 3.3° . In the N_2O_3 molecule with its long N-N bond length the errors are larger. Also, the errors in the bond angles are larger in this molecule. The bond lengths for the oxides of sulfur all have negative deviations from the observed values. The average error in the bond angles for the oxides of sulfur are about two times the norm.

5.5 Frequency Spectra

The average error for the vibrational frequencies is about 7%. *Ab initio* calculations with a STO-3G minimal basis set usually overestimates the frequencies by about 30%. When a 3-21G split valence shell basis set is used, the frequencies are usually overestimated by about 10-15%.

All the frequencies for the oxides of sulfur are larger than the experimental values. In the case of the oxides of nitrogen, frequencies at the low end of the spectrum are smaller than the experimental values and frequencies at the upper end of the spectrum are larger than the experimental values.

5.6 Selected Properties of NO and N₂O₄

The results of calculations for four selected properties for five different semi-empirical calculations are shown in Chapter 6. As the degree of complexity of the theory increases, the results do not always improve. However, the dipole moment tends to be more sensitive to the correct wave function. This statement is only true for dipole moments that are non-zero. On an overall basis, the AM1 method gives superior results for the selected properties.

Chapter 6

DISCUSSION

6.0 A General Comparison of Semi-Empirical Calculations and Experimental Results

J. J. P. Stewart²⁸ has applied semi-empirical methods such as MNDO, AM1, and PM3 to an extensive representative range of compounds. He compared the calculated and experimental results to investigate the validity of the optimized semi-empirical parameters used for predicting molecular properties such as heats of formation, ionization potentials, molecular geometries, vibrational frequencies, and dipole moments.

"MNDO is the oldest of the three methods surveyed, and as a direct result, is the least accurate. MNDO has many advantages over earlier semi-empirical methods. But even as MNDO is such a large improvement over those methods, enumeration of MNDO's good points would be invidious to the other methods. Instead, only the limitations likely to be encountered by users will be mentioned. The principal drawbacks to MNDO are:

- Sterically crowded molecules are too unstable
- Four-membered rings are too stable
- The hydrogen bond is virtually non-existent
- Hypervalent compounds are too unstable
- Activation barriers are generally too high
- Non-classical structures are predicted to be unstable relative to the classical

structure

- Oxygenated substituents on aromatic rings are out-of-plane
- The peroxide bond is systematically too short by about 0.17 Å
- The C-O-C angle in ethers is too large by about 9°. "29

"AM1 is a distinct improvement over MNDO, in that the overall accuracy is considerably improved. Specific improvements are:

- The strength of the hydrogen bond in the water dimer is 5.5 kcal/mol, in accordance with experiment.
- Activation barriers for reaction are markedly better than those of MNDO.
- Hypervalent phosphorus compounds are considerably improved relative to MNDO.
- In general, errors in ΔH_f obtained using AM1 are about 40% less than those given by MNDO. "30

"Unfortunately, with this improvement a few deficiencies were introduced. The most important of these are:

- AM1 phosphorus has a spurious and very sharp potential barrier at 3.0 Å. The effect of this is to distort otherwise symmetric geometries and to introduce spurious activation barriers. A vivid example is given by P_4O_6 , in which the notionally [sic] equal P-P bonds are predicted by AM1 to differ by 0.4 Å. This is by far the most severe limitation of AM1.
- Alkyl groups have a systematic error due to the heat of formation of the CH_2 fragment being too negative by about 2 kcal/mol.

- Nitro compounds, although considerably improved relative to MNDO, are still systematically too positive in energy.
- The peroxide bond is still systematically too short by about 0.17 Å.³¹

6.1 Enthalpy of Formation

Tables I and II compare the values calculated in this work and the experimental values of the heats of formation for the oxides of nitrogen and the oxides of sulfur. While the average errors are greater than the corresponding ones for organic compounds (5.5 kcal/mol), they are about the same as the corresponding errors (14.8 kcal/mol) in MNDO calculations for a number of elements where MNDO has proven good enough to be useful. While these large errors may be due in part to the uncertainty of the thermochemical data, it seems clear that the main problem lies in the neglect of d atomic orbitals and of changes in orbitals with formal charge. There seems to be no adequate way to take these difficulties into account within the MNDO/AM1 formalism. It should, however, be noted that *ab initio* methods give no better results unless a large basis set is used. The cost of the necessary calculations rules out such procedures in most chemical contexts. In the cases of NO and NO₂, which have open-shell configurations, both are predicted to be much too stable. Since AM1 was parameterized by the use of closed-shell species, large errors seem to occur for open-shell configurations. These large errors can be reduced only at the expense of unacceptable compensating errors elsewhere.

**Table I. Enthalpy of Formation for Selected Oxides of Nitrogen
Semi-Empirical AM1**

Species	Calculated kcal/mol	Observed kcal/mol	Obs - Calc kcal/mol
NO	1.2	21.6^{32,33}	20.4
NO₂	-14.9	7.9^{32,33}	22.8
N₂O	28.4	19.6³³	-8.8
N₂O₃	21.8	19.8^{32,33}	-3.0
N₂O₄	24.8	2.2^{32,33}	-22.6
N₂O₅	6.0	2.7^{32,33}	-3.3
HNO₃	-37.5	-32.1^{32,33}	5.4
H₂O	-59.3	-57.8³³	1.5

The superscripts are references given in the reference section beginning on page 75.

**Table II. Enthalpy of Formation for Selected Oxides of Sulfur
Semi-Empirical AM1**

Species	Calculated kcal/mol	Observed kcal/mol	Obs - Calc kcal/mol
SO	-11.4	1.2^{32,33}	12.6
SO₂	-50.8	-71.0³³	-20.2
SO₃	-97.3	-94.6³³	2.7
S₂O	9.9	-12.8³⁴	22.7

6.2 Ionization Potential

Tables III and IV compare the values calculated in this work and the experimental values of the first ionization energies estimated from the HOMO energies, calculated by AM1, using Koopmans' theorem. This theorem neglects electron reorganization and the change in correlation energy upon ionization. The errors occur in molecular orbitals derived largely from the 2s-orbitals of nitrogen, oxygen, and the 3s-orbital of sulfur. It is possible that these errors are due to neglect of 1s-2s interactions in the case of the oxides of nitrogen, and of interactions between core and valence electrons in the oxides of sulfur. The effects of core electrons are not considered in the "unperturbed" core approximation used in AM1. The neglect of d-orbitals may also be partly responsible for the errors.

6.3 Dipole Moments

Calculated and experimental dipole moments are compared in Tables V and VI. The dipole moment is due mainly to the atomic charges: only a small fraction is due to contributions from the lone-pairs of electrons. The resulting calculation of the dipole based on the charge distribution is slightly in error, either due to a smaller charge on atoms than is actually true, or due to some limitation within the semi-empirical methods. More "realistic" charges would be preferable; hopefully they would also accurately reproduce the observed dipole moment. However, the charge distribution cannot be observed directly.

**Table III. Ionization Potentials for Selected Oxides of Nitrogen
Semi-Empirical AM1**

Species	Calculated eV	Observed eV	Obs - Calc eV
NO	4.40	9.26³⁵, 9.54³⁶	4.86, 5.14
NO₂	6.57	11.25³⁷	5.68
N₂O	12.09	12.89^{36,38}	0.80
N₂O₃	11.29	*	*
N₂O₄	11.22	11.40³⁹	0.18
N₂O₅	13.57	12.30⁴⁰	-1.27
HNO₃	12.87	11.95⁴¹	0.92
H₂O	12.46	12.62³³	0.16

* No value found in the literature.

**Table IV. Ionization Potentials for Selected Oxides of Sulfur
Semi-Empirical AM1**

Species	Calculated eV	Observed eV	Obs - Calc eV
SO	8.55	10.29^{42, 43}	1.74
SO₂	10.49	12.31^{33,36}	1.82
SO₃	15.26	11.00³⁶, 12.81⁴⁴	-4.26, -2.45
S₂O	9.90	10.58³⁴	0.68

**Table V. Dipole Moments of Selected Oxides of Nitrogen
Semi-Empirical AM1**

Species	Calculated D	Observed D	Obs - Calc D
NO	0.14	0.15⁴⁵	0.01
NO₂	0.45	0.32⁴⁵, 0.40⁴⁶	-0.13, -0.05
N₂O	0.64	0.17³³, 0.65⁴⁷	-0.47, 0.01
N₂O₃	3.28	2.12⁵²	-1.16
N₂O₄	0.00	0.00⁴⁸	0.00
N₂O₅	0.09	1.39⁴⁹	1.30
HNO₃	2.58	2.17³³	-0.41
H₂O	1.86	1.85^{33,45}	-0.01

**Table VI. Dipole Moments of Selected Sulfur Containing Molecules
Semi-Empirical AM1**

Species	Calculated D	Observed D	Obs - Calc D
SO	2.06	1.55^{45,48}	-0.51
SO₂	4.29	1.63^{45,48}	-2.66
SO₃	0.00	0.00⁴⁵	0.00
S₂O	3.37	1.47⁴⁹	-1.90

6.4 Molecular Geometries

Tables VII and VIII show the optimized structures calculated by AM1 together with experimental values. The experimental values from microwave spectroscopy (r_s) and values from electron diffraction (r_g) differ somewhat from the equilibrium (r_e) values. Discrepancies of less than 0.01 Å in bond lengths and 1° in bond angles are therefore not too meaningful.

Ab initio calculations reproduce the geometries satisfactorily only when extended basis sets including d-orbitals are used. It is therefore surprising that the only significant errors in the AM1 geometries are systematic ones that can be accommodated by applying corrections in the case of bond lengths for the oxides of sulfur. The fact that AM1 gives better results than the *ab initio* SCF treatments reflects the implicit allowance for electron correlation in AM1 via the optimized parameters.

6.5 Frequencies

Tables IX and X compare molecular vibration frequencies calculated by AM1 with experiment. The average error for AM1 is smaller than that for *ab initio* methods. The STO-3G minimal basis set drastically overestimates frequencies in small molecules. The typical error is about 30% but, fortunately, the slightly larger 3-21G split valence set does much better, with most frequencies overestimated by only 10-15%. When electron correlation is included, as with MP2(Moller-Plesset perturbation theory to second order), the error is almost cut in half⁵⁰, even though computed frequencies are still larger than experiment. Although the errors for the

**Table VII. Geometrical Properties of Selected Oxides of Nitrogen
Semi-Empirical AM1**

Species	Calculated Bond Length Å	Observed Bond Length Å	Calculated Bond Angle	Observed Bond Angle
NO	(N-O)1.117	1.150 ⁵¹	-----	-----
NO ₂	(N-O)1.163 (N-O)1.163	1.188 ⁵¹ 1.188 ⁵¹	134.72	134.1 ⁵¹
N ₂ O	(N-N)1.128 (N-O)1.175	1.128 ⁵¹ 1.184 ⁵¹	180.00	180.00 ⁵¹
HNO	(H-N)1.042 (N=O)1.157	1.036 ⁵¹ 1.241 ⁵¹	115.42	116.3 ³⁵
HNO ₃	(O-N)1.193 (N-O)1.333 (O-H)0.982	1.22 ⁵² ,1.206 ⁵¹ 1.44 ⁵² ,1.406 ⁵¹ 1.22 ⁵² ,0.960 ⁵¹	(O-N-O)129.0 (N-O-H)109.7	137.8 ⁵² ,130 ³⁵ 111.1 ⁵² ,102 ³⁵
N ₂ O ₃	(N-O ₁)1.128 (N-O ₂)1.192 (N-O ₃)1.192 (N-N)1.576	1.142 ⁵³ 1.202 ⁵³ 1.217 ⁵³ 1.864 ⁵³	(N-N-O ₁)116.4 (N-N-O ₂)119.3 (N-N-O ₃)124.3	105.1 ⁵³ 112.7 ⁵³ 117.5 ⁵³
N ₂ O ₄	(N-O)1.172 (N-N)1.812	1.180 ³³ 1.750 ³³	(O-N-O)131.8 (O-N-N)114.1	133.7 ^{33,54} 113.2 ³³
N ₂ O ₅	(O-N)1.179 (N-O)1.393	1.18 ⁵¹ 1.3-1.4 ⁵¹	(O-N-O)132.1 (N-O-N)128.2	134 ⁵¹ 95 ⁵¹
N ₂ O ₂	(N-O)1.37 (O-O)1.38 (N-N) 1.27	2.62 ⁵³ 2.18 ⁵³	(N-N-O)92.3 (O-O-N)87.7	101 ⁵³ 79 ⁵³
H ₂ O	(H-O)0.961	0.957 ²⁸	103.53	104.5 ²⁸

**Table VIII. Geometrical Properties of Selected Oxides of Sulfur
Semi-Empirical AM1**

Species	Calculated Bond Length Å	Observed Bond Length Å	Calculated Bond Angle	Observed Bond Angle
SO	(S-O)1.471	1.481⁵¹	-----	-----
SO₂	(S-O)1.429 (S-O)1.429	1.432⁵¹	107.90	119.54⁵¹
SO₃	(S-O)1.350	1.43⁵¹	120.00	120.0⁵¹
S₂O	(S-S)1.808 (S-O)1.428	1.882⁵⁵ 1.462⁵⁵	(S-S-O)126.4	118.1⁵⁵

**Table IX. Frequency Spectrum of Selected Oxides of Nitrogen
Semi-Empirical AM1**

Species	Calculated/cm⁻¹	Observed/cm⁻¹
NO	2410	1904⁴³
N₂O	516(2), 1602, 2684	589(2), 1285, 2224⁵⁶
NO₂	716, 1770, 2235	750, 1318, 1618⁵⁶
N₂O₂	556, 1023, 1131, 1200, 1382, 1938	*
N₂O₄	29, 216, 268, 439, 458, 645, 708, 1683, 1736, 2197, 2235	283, 380, 500, 680, 752, 813, 1265, 1360, 1724, 1749⁵⁷
N₂O₅	104, 232, 461, 512, 659, 695, 826, 992, 1073, 1667, 1727, 2165, 2242	*
HNO	1611, 2200, 3137	1110, 1570, 2854⁵⁵
HNO₃	451, 654, 686, 713, 1319, 1540, 1734, 2172, 3361	456, 579, 647, 762, 879, 1325, 1331, 1708, 3550⁵⁶
H₂O	1885, 3506, 3585	1588, 3650, 3792⁵⁶

* No data found in the literature.

**Table X. Frequency Spectrum of Selected Oxides of Sulfur
Semi-Empirical AM1**

Species	Calculated/cm⁻¹	Observed/cm⁻¹
SO	1007	1124⁵⁸
SO₂	465, 940, 1050	497, 1101, 1318⁵⁶
SO₃	438, 461, 929, 1261	498, 530, 1065, 1391⁵⁶
S₂O	257, 642, 1010	388, 679, 1165⁵⁹

oxides of sulfur are all negative and the errors may be systematic, there is no consistency of the errors between the observed and calculated values for the oxides of nitrogen. This implies that the errors are random. The use of a fixed set of parameters is likely to be the origin of random errors for the oxides of nitrogen and the apparent systematic errors for the oxides of sulfur.

6.6 Selected Properties of NO and N₂O₄

Tables XI and XII compare the calculated and observed values for bond lengths, dipole moments, heats of formation, and ionization potentials for NO and N₂O₄ using five different semi-empirical methods. Among these, CNDO is the simplest and least accurate method, and AM1 is the most complex and accurate. NO and N₂O₄ were selected to represent open-shell and closed-shell systems, respectively.

In the case of N₂O₄, which is a sterically crowded molecule, MNDO predicts the heat of formation to be too positive. The MNDO method tends to predict crowded molecules to be too unstable and the shortcoming seems to be inherent in the basic MNDO approximation. It should also be noted that no molecules with NO bonds were used in the parameterization of MNDO. It is possible that the errors could be reduced by a reparameterization, including such compounds in the reference data. In the case of radicals such as NO, MNDO systematically tends to predict the heats of formation to be somewhat too stable.

The CNDO and INDO methods give very poor results, especially in the calculations of the heats of formation. In the case of NO, a system containing

**Table XI. Selected Properties of the NO Molecule
Semi-Empirical Methods**

Method	Bond Length Å	Dipole Moment D	ΔH_f° kcal/mol	Ionization Energy ev
CNDO	1.154	0.168	-288.6	4.53
INDO	1.161	0.186	-198.9	3.94
MINDO/3	1.167	0.758	21.86	3.58
MNDO	1.124	0.174	-0.16	4.55
AM1	1.117	0.141	1.19	4.40
Experiment	1.150	0.153	21.58	9.26

**Table XII. Selected Properties of the N₂O₄ Molecule
Semi-Empirical Methods**

Method	Bond Length Å	Dipole Moment D	ΔH_f° kcal/mol	Ionization Energy ev
CNDO	(N-O) 1.219 (N-N) 1.366 (O-N-O) 127.1 (O-N-N) 116.5	0.00	-1165.80	15.20
INDO	(N-O) 1.220 (N-N) 1.374 (O-N-O) 126.9 (O-N-N) 116.5	0.00	-931.62	14.61
MINDO/3	(N-O) 1.219 (N-N) 1.513 (O-N-O) 132.5 (O-N-N) 113.7	0.00	-3.95	11.18
MNDO	(N-O) 1.189 (N-N) 1.632 (O-N-O) 128.7 (O-N-N) 115.6	0.00	32.01	11.99
AM1	(N-O) 1.172 (N-N) 1.812 (O-N-O) 131.8 (O-N-N) 114.1	0.00	24.84	11.22
Experiment	(N-O) 1.180 (N-N) 1.750 (O-N-O) 133.7 (O-N-N) 113.2	0.00	2.17	11.40

heteroatoms, the dipole moment improves greatly in MNDO compared to MINDO/3. The charge distribution is predicted much better with MNDO than with MINDO/3. In the case of ionization energy for the N_2O_4 molecule, the agreement with experiment is markedly better for AM1. The improvement obtained by the use of AM1 compared to MNDO is probably due more to having located a better minimum on the energy hyper-surface than to the superiority of the AM1 model. Overall, the AM1 method gives the best results.

6.7 Frequency Spectrum

Figures 1a and 1b show the frequency spectra of N_2O and SO_2 , respectively. In the case of N_2O there is one calculated value that is smaller than the observed value. Two of the frequencies are larger than the experimental values. For SO_2 all calculated values are smaller than the observed values. These results imply that no systematic errors exist and the errors are random.

6.8 Orbital Energy Levels for NO

Figure 2 is the molecular orbital energy level diagram for the NO molecule. The two π bonding molecular orbitals are nearly degenerate, and thus have almost the same energy. But the two π antibonding molecular orbitals, one of which is the HOMO and the other the LUMO, are split in energy. As predicted by the Jahn-Teller theorem the degeneracy is removed because NO is a heteronuclear diatomic molecule. For homonuclear diatomic molecules such as N_2 , the two π antibonding orbitals as well as the two π bonding orbitals are degenerate.

6.9 Highest Occupied Molecular Orbital for NO

The HOMO is the molecular orbital which has the highest energy among molecular orbitals occupied by at least one electron. The HOMO is shown in Figure 3 and it has a nodal plane between the nuclei. That is, the electron density is expelled from the region between the nuclei. Consequently, the HOMO for NO is an antibonding π molecular orbital. The HOMO for NO may be considered as constructed from two p atomic orbitals by a linear combination of atomic orbitals. One of the two p atomic orbitals comes from the nitrogen atom and one comes from the oxygen atom. The two different colors represent different signs in different portions of the molecular orbital. The red color corresponds to a minus sign and the green color corresponds to a plus sign.

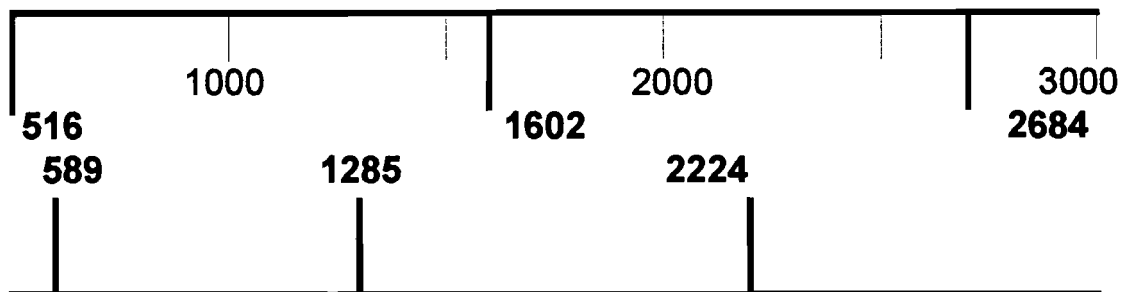
6.10 Lowest Unoccupied Molecular Orbital for NO

Figure 4 represents the LUMO of a NO molecule. (LUMO stands for Lowest Unoccupied Molecular Orbital.) The LUMO is the molecular orbital which has the lowest energy among the unoccupied molecular orbitals. The basic symmetry of the LUMO is the same as that of HOMO. They would be degenerate in a homonuclear molecule such as N_2 . The only visible difference is the change in sign of each lobe of the molecular orbital.

Figure 1. Frequency Spectrum of (a) N_2O and (b) SO_2 .

(a) Frequency Spectrum of N₂O
cm⁻¹

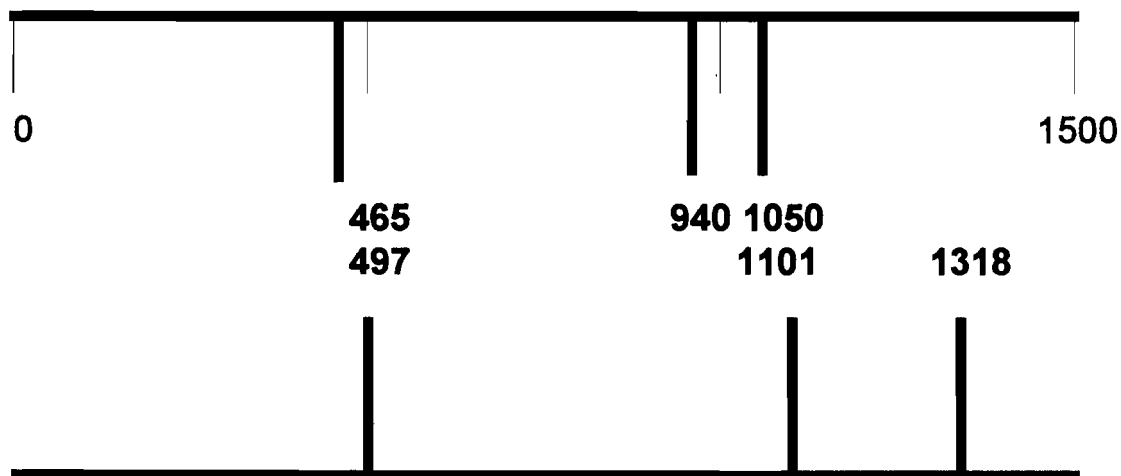
Calculated



Observed

(b) Frequency Spectrum of SO₂
cm⁻¹

Calculated



Observed

Figure 2. Molecular Orbital Energy Level Diagram for the NO Molecule.

Figure 3. Highest Occupied Molecular Orbital of the NO Molecule.

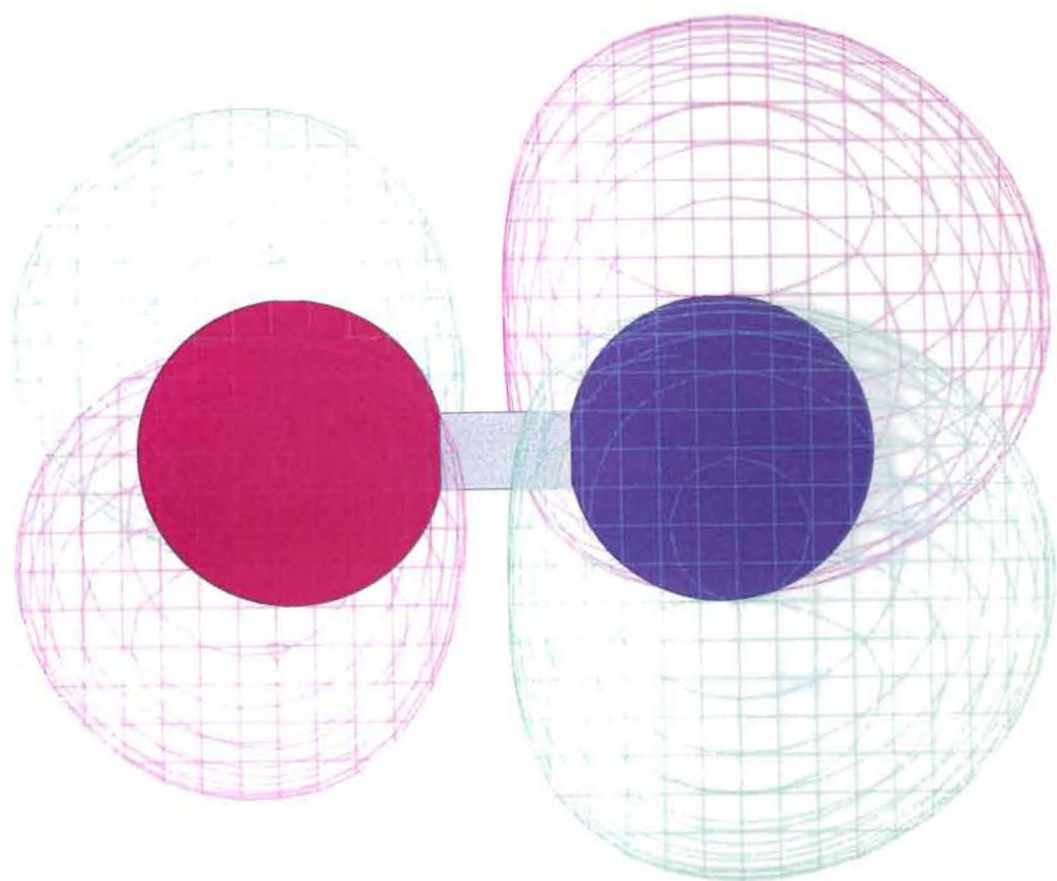
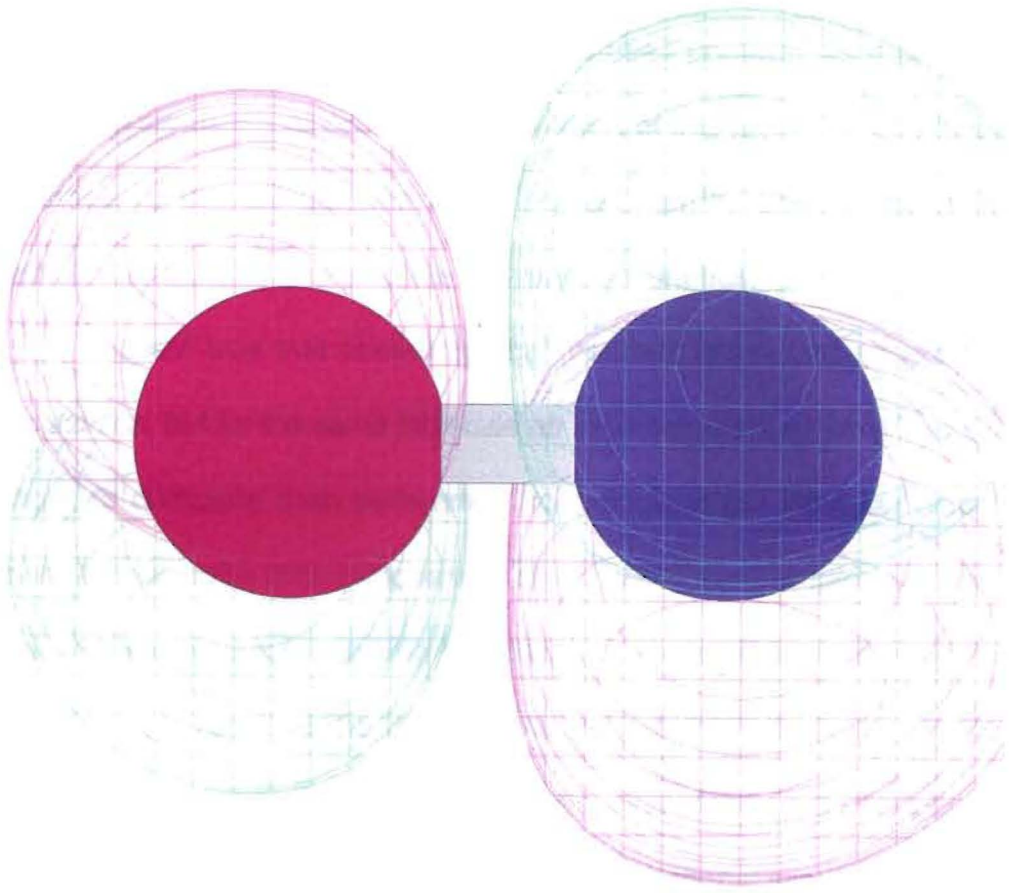


Figure 4. Lowest Unoccupied Molecular Orbital of the NO Molecule.



Chapter 7

SUMMARY

7.1 Conclusion

With modern computers and software packages such as HyperChem™, the time-independent Schrödinger equation can be solved for many molecules of chemical interest. Quantum chemistry should be regarded as one more tool for a chemist to learn how to use. It should be used primarily as a predictive tool, to assist as a screening device. Part of the effort of this research was to evaluate the accuracy of the calculations using HyperChem™, and to assist in choosing methods to apply in future studies of reaction kinetics.

It is still true that semi-empirical methods are not as accurate as *ab initio* calculations, but for the same molecule *ab initio* calculations take about 1000 times longer to complete than semi-empirical calculations. *Ab initio* calculations are limited to molecules containing tens of atoms, while semi-empirical calculations can deal with molecules containing hundreds of atoms. Quantum chemistry has not, however, yet reached the development where the calculated results are more accurate than the best experimental values, even if they are cheaper and quicker.

One set of quantum mechanical calculations simultaneously yields the enthalpies of formation, dipole moments, bond geometries, ionization potentials, and vibrational frequencies. Even unstable or highly reactive species can be studied by quantum mechanics, even when these species can not be studied by experiment. For these reasons the use of software such as HyperChem™

represents a reasonable first approximation to characterize the quantum chemistry of molecules.

7.2 Thoughts for the Future

As limitations on the generality of the semi-empirical methods become apparent, the parameters can more readily be reoptimized in an attempt to remove these limitations. The parameters may be reoptimized as significant amounts of new or improved experimental data become available, or if significant improvements to the algebraic form of the Hamiltonian are developed. Although parameter optimization is now a more straightforward task, it should not be attempted lightly, that is, a proliferation of parameter sets differing only slightly one from another would be undesirable. Rather, only when a significant increase in accuracy could be obtained, such as a decrease in the average error of more than 30%, should a new parameter set be released for general use.

Current experimental data are simply not as accurate as desirable. Unfortunately, determination of the accuracy of any given computational method requires accurate experimental data as standards. Even more unfortunate, little effort appears to have been committed to increasing the accuracy of existing experimental data. With the steady advance in the accuracy of *ab initio* methods, there is a significant probability that within the near future the calculated values will become more accurate than the experimental values, and that semi-empirical methods will be parameterized against high-level *ab initio* calculations rather than experimental values. Even now good *ab initio* calculations of molecular geometries

are of an accuracy comparable with microwave data, and significantly better than many single molecule geometries obtained from X-ray determinations.

Several deficiencies in the predictions of geometric variables have been corrected. Even so, some problems still remain intractable. It is possible that modification of the Hamiltonian, a relatively easy operation now that reparameterization is rapid, will allow correction of these faults.

REFERENCES

1. Serway, R. A. *Physics for Scientists and Engineers*, 3rd ed.; Saunders: Philadelphia, 1992 ; pp 1049-1052.
2. Rohlf, J. W. *Modern Physics from α to Z^0* ; Wiley: New York, 1994 ; pp 7-8.
3. Serway, R. A. *Physics for Scientists and Engineers*, 3rd ed.; Saunders : Philadelphia, 1992 ; pp 957-958.
4. Rohlif, J. A. *Modern Physics from α to Z^0* ; Wiley: New York, 1994 ; pp 66-70.
5. Ibid., pp 76-77.
6. Ibid., pp 162-176.
7. Ibid., pp 82-87.
8. Ibid., pp 136-137.
9. Ibid., pp 140-143.
10. Semat, H. *Introduction to Atomic and Nuclear Physics*; Holt, Rinehart and Winston: New York, 1962; p 202.
11. Burkert, U.; Allinger, N. L. *Molecular Mechanics*. ACS Monograph 177, American Chemical Society: Washington, D.C., 1982
12. Bowen, J. P.; Allinger, N. L. in *Reviews in Computational Chemistry*, Lipkowitz, K. B; Boyd, D. B., Eds., VCH: New York, 1991; pp 81-97.
13. Karplus, M.; Pettitt, B.M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, in: *Advances in Chemical Physics*, Vol.71. Wiley: New York, 1988
14. Pople, J. A.; Segal, G. A. *J.Chem.Phys.* **1966**, *44*, 3289.
15. Pople J. A.;Beveridge D. L. *Approximate Molecular Orbital Theory*, McGraw-Hill: New York, 1970.

16. Dewar, M. J. S. *The Molecular Orbital Theory of Organic Chemistry*, McGraw-Hill: New York, 1969; see also Murrell, J. N.; Harget, A. J. *Semiempirical Self-Consistent Field Molecular Orbital Theory of Molecules*, Wiley-Interscience, London, 1972.
17. Dewar, M. J. S.; Ford, G.P. *J. Am. Chem. Soc.* **1979**, 101, 5558.
18. Thiel, W. *J. Am. Chem. Soc.* **1981**, 103, 1413.
19. Hehre, W. J.; Radom, L.; Schleyer, P. v. R; Pople, J. A. *Ab Initio Molecular Orbital Theory*, Wiley: New York, 1986.
20. Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, 51, 2657; Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, 52, 2769.
21. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975**, 97, 1294, 1298.
22. Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977** 99,4 899.
23. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985** 107, 3902.
24. Fleming, I, *Frontier Orbitals and Organic Chemical Reactions*; Wiley: New York, 1976.
25. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985** 107, 3902.
26. Stewart, J. J. P. *J. Comput. Aided Molecular Design.* **1990** 4, 1.
27. Stewart, J. J. P. *J. Comput. Chem.* **1989** 10, 221.
28. Stewart, J. J. P. *J. Comput. Aided Molecular Design.* **1990** 4, 1.
29. Ibid.
30. Ibid.
31. Ibid.

32. Chase, M. W.; Davies, C. A.; Downey, J. R.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. JANAF Thermochemical Tables, 3rd ed, *J. Phys. Chem. Ref. Data* **1985** 14, Supplements 1 and 2.
33. Stewart, J. J. P. *J. Computer-Aided Molecular Design* **1990** 4, 1.
34. Berkowitz, J.; Eland, J. H. D.; Appelman, E. H. *J. Chem. Phys.* **1977** 66, 2183.
35. White, M. G.; Seaver, M.; Chupka, W. A.; Colson, S. D. *Phys. Rev. Lett.* **1982** 49, 28.
36. Kimura, K.; Katsumata, S.; Achiba, Y.; Yamazaki, T.; Iwata, S.; *Handbook of He I Photoelectron Spectra of Fundamental Organic Molecules*; Halted: New York; 1981.
37. Katsumata, S.; Shiromaur, H.; Mitani, K.; Iwata, S.; Kimura, K. *Chem. Phys.* **1982** 69, 423.
38. Hollas, J. M.; Sutherley, T. A. *Chem. Phys. Letters* **1973** 21, 167.
39. McClelland, B. W.; Gundersen, G; Hedberg, K. *J. Chem. Phys.* **1972** 56, 4541
40. Chong, D. P.; Frost, D. C.; Lau, W. M.; McDowell, C. A. *Chem. Phys. Letters* **1982** 90, 332.
41. Ames, D. L.; Turner, D. W. *Proc. Roy. Soc.* **1976** A348, 175.
42. Lloyd, D. R.; Roberts, P. J.; Hillier, I. H. *J.C.S. Faraday II* **1975** 71, 496.
43. Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*, Van Nostrand- Reinhold: New York, 1979.
44. Alderdice, P. S.; Dixon, R. N. *J. C. S. Faraday II* **1976** 72, 372.
45. Nelson, R. D., et al, *Selected Values of Electric Dipole Moments for Molecules in the Gas Phase*, NSRDS-NBS10: Washington, 1967.
46. Struve, W. S. *Fundamentals of Molecular Spectroscopy*, Wiley: New York, 1989.

47. Smith, D. W.; Hedberg, K. *J. Chem. Phys.* **1956** 25, 1282.
48. Hillier, I. H.; Saunders, V. R. *Mol. Phys.* **1971** 22, 193.
49. Meschi, D. J.; Myers, R. J. *J. Mol. Spectroscopy* **1959** 3, 405.
50. Hout, R. F.; Levi, B. A.; Hehre, W. A. *J. Comput. Chem.* **1982** 3, 234
51. *Tables of Interatomic Distances and Configuration in Molecules and Ions*, Sutton, L. E., Ed.; The Chemical Society: London, 1958.
52. Brittain, A. H; Cox, P. *Trans. Faraday. Soc.* **1963** 65, 1969.
53. Billingsley, J; Callear, A. B. *Trans. Faraday Soc.* **1971** 67, 589
54. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J. Am. Chem. Soc.* **1975** 97, 1302.
55. Harmony, M. D.; Laurie, V. W; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Fafferty, W. J.; Maki, A. G.; *J. Phys. Chem. Ref. Data* **1979** 8, 619.
56. Shimanouchi, T. *J. Phys. Chem. Ref. Data* **1977** 6, 993.
57. Herzberg, G. *Molecular Spectra and Molecular Structure I. Infrared and Raman Spectra*; Van Nostrand: New York, 1945.
58. Herzberg, G. *Molecular Spectra and Molecular Structure II. Spectra of Diatomic Molecules*. Van Nostrand: New York, 1945.
59. Blukis, U.; Meyers, R. J., *J. Phys. Chem.* **1965** 69, 1154.

I, Junyoung Kwon, hereby submit this thesis/report to Emporia State University as partial fulfillment of the requirements for an advanced degree. I agree that the Library of the University may make it available for use in accordance with its regulations governing materials of this type. I further agree that quoting, photocopying, or other reproduction of this document is allowed for private study, scholarship (including teaching) and research purposes of a nonprofit nature. No copying which involves potential financial gain will be allowed without written permission of the author.

Junyoung Kwon

Signature of Author

Apr. 12 1996

Date

Applications of Semi-Empirical Quantum
Mechanical Calculations to Simple Molecular
Systems

Title of Thesis

Dave Cooper

Signature of Graduate Office Staff Member

April 15, 1996

Date Received

original