CHAPTER 2

METHODOLOGY

- Quantum Chemical Parameters
- Description of Software Used
  - MOPAC
  - PCMODEL Software
  - CACHe Software
- Geometry Optimization
METHODOLOGY:

The study materials of this thesis are halides of Co (II) and Ni (II) and their aminothiazole complexes. The 3D modeling and geometry optimization of the halide will be done by CAChE software using molecular mechanics with EHT option. Eigenvalues, eigenvectors and overlap matrix values will be obtained with the same software, using the same option. With the help of these values, eigenvector analysis, magnitude of contribution of atomic orbital in MO formation and population analysis will be made and discussed. The method adopted for various calculations is based on the following principles.

Thanks to the pioneering contribution of J. A. Pople [64] and his colleagues[64-67] whose work in computational quantum chemistry has revolutionized research in theoretical chemistry since the 1960s. Much of the modern computational quantum chemistry uses Gaussian orbitals as a basis set. Pople has developed algorithms for the ab initio calculation of molecular properties based on Gaussian orbitals. The computer programs developed by Pople and his many collaborators have been packaged as a commonly available program called GAUSSIAN 94. This program, released in 1994, was developed at Carnegie-Mellon University in the U.S.A. In the ab initio calculations (which use the exact Hamiltonian), the task is greatly simplified by expressing the AOs used in the LCAO as linear combinations of Gaussian orbitals. A Gaussian type orbital (GTO) is a function of the form $e^{-ar^2}$. However, in the semi-empirical calculations, the AOs have been used as Slater-type orbitals as basis sets. A Slater-type orbital (STO) is a function of the form $e^{-ar}$. (Hydrogenic orbitals as basis sets are not all satisfactory.) We may recall that the set of AOs used to construct LCAO-MOs is called the basis set. In the case of diatomic molecules, the hydrogenic orbitals from the basis set. For example, the 1s orbitals of the two atoms constitute the basis for the MO
σ_g (1s). For polyatomic molecules, the first basis set used in large-scale computational studies consisted of STOs:

\[
S_{nlm}(r, \theta, \phi) = N_n r^{n-1} e^{-\xi r} Y_{lm}(\theta, \phi)
\]  

(1)

where \( N_n \) is the normalization constant defined as

\[
N_n = \frac{(2\xi)^{n+1/2}}{[(2n)!]^{1/2}}
\]

where \( \xi \) is the orbital exponent (zeta), \( n \) and \( l \) are the principal and orbital momentum (azimuthal) quantum numbers and \( Y_{lm}(\theta, \phi) \) are the angularly dependent spherical harmonics. The STOs differ from the hydrogenic orbitals in that the STOs have no nodes and the orbital exponent is not necessarily equal to \( Z/n \). In principle, the orbital exponents should be chosen to minimize the energy but this selection is still a formidable task even with modern computers. Slater has given a set of empirical rules for determining \( \xi \) for orbitals in free atoms. \( \xi \) increases with increasing atomic number, reflecting the contraction of the orbital as the nuclear charge increases.

The STOs are not used directly in modern theoretical research because the integral in the resulting secular determinants are difficult to evaluate. In particular, integrals involving more than one nuclear center, called multicentre integrals, are awkward to evaluate using STOs. If we use GTOs instead of STOs, however, all the multicentre integrals are very easy to evaluate. As a matter of fact, the four-center integrals can become two-center integrals which can be easily computed. The use of GTOs was
suggested in 1950 by S. F. Boys. The advantage of GTOs is that the product of two Gaussian functions is itself a Gaussian function that lies between the centers of the two contributing functions. The GTOs may be represented by

\[ G_{nlm}(r, \theta, \phi) = N_n r^{n-1} e^{-ar^2} Y_{l,m}(\theta, \phi) \]  

(2)

for the basis set in MO calculations.

The GTO does a reasonably good job of describing the STO for values of \( r \) greater than \( a_0 \) (the Bohr radius) but it underestimates its magnitude for values of \( r \) less than \( a_0 \). These discrepancies turn out to be very significant in molecular calculations. To overcome this difficulty quantum theorists prefer the use of curve fit STOs to sums of Gaussian functions; the curve fit improving with \( N \), the number of Gaussian functions used. If, for instance, \( N=3 \), i.e., we use a sum of three Gaussian functions to represent one STO; such a basis set is called the **STO-3G basis set**. In the STO-3G basis set, all AOs are described by a sum of three Gaussian functions. Although this procedure leads to a proliferation of integrals to evaluate, each one is relatively easy. Thus, the overall procedure is quite efficient. However, there are some disadvantages, too, in using GTOs since many GTOs are needed to obtain for atoms the same degree of accuracy as the STOs.

It will be pertinent to define here the so-called **valence state ionization potentials**. Consider the electronic configuration of carbon atom, viz., \( 1s^2 \chi_1^1 \chi_2^1 \chi_3^1 2p_3^1 \) where \( \chi_i \) (\( i=1, 2, 3 \)) are the three triply hybridized orbitals in the sp\(^2\) hybridized state, the hybrids being formed from 2s, 2p\(_x\) and 2p\(_y\) AOs. Here the carbon atom is said to be promoted to a **valence state** obtained through electronic rearrangement prior to the bond formation. This electronic configuration may be considered to be
equivalent to the configuration $1s^2 \ 2s^1 \ 2p_x^1 \ 2p_y^1 \ 2p_z^1$. The valence state ionization potential of the $2p_z$ orbital in the carbon atom is given by the difference in energy between neutral carbon atoms in the configuration $1s^2 \ 2s^1 \ 2p_x^1 \ 2p_y^1$. Similarly, the valence state ionization potential of the $2s$ orbital in carbon atom is given by the energy difference between the neutral atom configurations $1s^2 \ 2s^1 \ 2p_x^1 \ 2p_y^1 \ 2p_z^1$ and the C$^+$ in the configuration $1s^2 \ 2p_x^1 \ 2p_y^1 \ 2p_z^1$. Typical values of the valence state ionization potentials are $-13.60$ eV (for hydrogen 1s), $-21.4$ eV (for carbon 2s), $-11.4$ eV (for carbon 2p) and $-26.0$ eV (for nitrogen 2s).

A few remarks are in order for three important terms, viz., correlation energy, electron correlation and configuration interaction, which are often involved in both the Hartree-Fock SCF theory of multielectron atoms and Hartree-Fock-Roothaan SCF theory of polyatomic molecules. It may be recalled that for helium atom, the use of a hydrogenic 1s orbital with variable orbital exponent gives a ground state energy of $-77.5$ eV, compared with the true value of $-79.0$ eV. The Hartree-Fock wave function for the helium ground state gives an energy of $-77.9$ eV, which is still in error by 1.1 eV. The energy error of the Hartree-Fock wave function is called the correlation energy since it results from the fact that the Hartree-Fock wave functions neglects the instantaneous correlations in the motions of electrons. Electrons repel one another and correlate their motions to avoid being close together; this phenomenon is called electron correlation.

Let us now consider configuration interaction. A method used to improve the Hartree-Fock wave function is configuration interaction. When a Hartree-Fock ground state wave function of an atom or a molecule is calculated, one also obtains expressions for unoccupied excited state orbitals. It is possible to show that the set of functions obtained by making all possible assignments of electrons to the available
orbitals is a complete set. Hence, the true wave function of the ground state can be expressed as

$$\psi = \sum_j c_j \psi_{orb,j}$$  \hspace{1cm} (3)

where the $\psi_{orb,j}$ s are approximate orbital functions that differ in the assignments of electrons to orbitals. Each $\psi_{orb,j}$ is a Slater determinant of spin orbitals. The functions $\psi_{orb,j}$ are called configuration functions (or configurations). One uses a variation procedure to find the values of the coefficients $c_j$ that minimize the variational integral. This type of calculation is called configuration interaction[68].

**POPULATION ANALYSIS:**

The molecular orbitals are formed by the linear combination of basis functions. Most molecular quantum-mechanical methods (such as SCF, CI etc.) begin the calculation with the choice of a basis functions $\chi_r$, which are used to express the MOs $\phi_i$ as $\phi_i = \Sigma_i c_{ri} \chi_r$ ($c =$ coefficient of $\chi$, $r =$ number of atomic orbital, $i =$ molecular orbital number). The use of an adequate basis set is an essential requirement for the calculation. The basis functions are usually taken as AOs. Each AO can be represented as a linear combination of one or more Slater-type orbitals (STOs)[69-71]. Each molecular orbital $\phi_i$ is expressed as

$$\phi_i = \Sigma_i c_{ri} \chi_r$$  \hspace{1cm} (4)

where, the $\chi_r$ s are the STO basis functions. Here we use the STO–6G basis set (which is contracted Gaussian)[72-75] for the SCF calculation.

The coefficients in linear combination for each molecular orbital being found by solution of the Roothaan equation[76]. The most efficient way to solve the Roothaan equation is to use matrix–algebra methods. In matrix–algebra methods, the matrix elements are computed,[77] and the secular equation is solved to give the set of orbital energies (i.e.
eigenvalues). These orbital energies[78-80] are used to solve Roothaan equations for the set of coefficients (i.e. eigenvectors) giving a set of molecular orbitals. The calculations are done using a computer.

By the above calculation, the values of orbital energies (eigenvalues) and eigenvectors (coefficients) have been calculated.

A widely used method to analyze SCF wave function is population analysis, introduced by Mulliken[81,82]. He proposed a method that apports the electrons of an n–electron molecule into net populations \( n_r \) in the basis functions \( \chi_r \) and overlap populations \( n_{r,s} \) for all possible pairs of basis functions.

For the set of basis functions \( \chi_1, \chi_2, \ldots, \chi_b \), each molecular orbital \( \phi_i \)
has the form \( \phi_i = \sum c_{si} \chi_s = c_{1i} \chi_1 + c_{2i} \chi_2 + \ldots + c_{bi} \chi_b \). For simplicity, we shall assume that the \( c_{si} \)'s and \( \chi_s \)'s are real. The probability density associated with one electron in \( \phi_i \) is (s and b are the number of the atomic orbital other than r)

\[
|\phi_i|^2 = c_{1i}^2 \chi_1^2 + c_{2i}^2 \chi_2^2 + \ldots + 2c_{1i}c_{2i} \chi_1 \chi_2 + 2c_{1i}c_{3i} \chi_1 \chi_3 + 2c_{2i}c_{3i} \chi_2 \chi_3 + \ldots
\]

Integrating this equation over three-dimensional space and using the fact that \( \phi_i \) and the \( \chi_s \)'s are normalized, we get

\[
1 = c_{1i}^2 + c_{2i}^2 + \ldots + 2c_{1i}c_{2i} S_{12} + 2c_{1i}c_{3i} S_{13} + 2c_{2i}c_{3i} S_{23} + \ldots
\]  

(5)

Where the S's are overlap integrals: \( S_{12} = \int \chi_1 \chi_2 dv_1 dv_2 \), etc. Mulliken proposed that the terms in Eq.5 be apportioned as follows. One electron in the molecular orbital \( \phi_i \) contributes \( c_{1i}^2 \) to the net population in \( \chi_1 \), \( c_{2i}^2 \) to the net population in \( \chi_2 \), etc., and contributes \( 2c_{1i}c_{2i} S_{12} \) to the overlap population between \( \chi_1 \) and \( \chi_2 \), \( 2c_{1i}c_{3i} S_{13} \) to the overlap population between \( \chi_1 \) and \( \chi_3 \), etc.

Let there be \( n_i \) electrons in the molecular orbital \( \phi_i \) (\( n_i = 0,1,2 \)) and let \( n_{r,i} \) and \( n_{r,s,i} \) symbolize the contributions of electrons in the molecular
orbital $\phi_i$ to the net population in $\chi_r$ and to the overlap population between $\chi_r$ and $\chi_s$, respectively. We have

$$n_{ri} = n_i c_{ri}^2$$  \hspace{1cm} (6)

$$n_{rs,i} = n_i (2c_{ri} c_{si} S_{rs})$$  \hspace{1cm} (7)

Based on the above principle, the contributions of electrons in each occupied molecular orbital has been calculated with the help of eigenvector values. The overlap population analysis has also been made with the help of the values of eigenvector and overlap matrix for distinguishing the bonding, nonbonding and antibonding nature of molecular orbital.

**DESCRIPTION OF SOFTWARE USED:**

1- MOPAC

2- CAChe

3- PCMODEL Software

**1. MOPAC**

MOPAC [83-85] is a general-purpose semi-empirical molecular orbital package for the study of solid state and molecular structures and reactions.

The semi-empirical Hamiltonians MNDO,[86] MINDO/3,[87] AM1,[88] PM3,[89] and MNDO-d[90,91] are used in the electronic part of the calculation to obtain molecular orbitals, the heat of formation and its derivative with respect to molecular geometry. Using these results MOPAC calculates the vibrational spectra, thermodynamic quantities, isotopic substitution effects and force constants for molecules, radicals, ions, and polymers. For studying chemical reactions, a transition state location routine[92] and two transition state optimizing routines[93-95] are available. For users to get the most out of the program, they must
understand how the program works, how to enter data, how to interpret the results, and what to do when things go wrong.

1. **CACHe**

CACHe is a computer-aided molecular design (CAMD) modeling tool for the Microsoft® Windows operating systems. CACHe enables us to draw and model molecules and perform calculations on a molecule to discover molecular properties and energy values.

CACHe uses computational chemistry as an essential part of computer-aided chemistry. By applying computational molecular models derived from mathematical equations to a chemical sample we create, we can calculate molecular properties such as the position and behavior of electrons, and nuclei in certain conditions.

The computational chemistry tools that CACHe uses are derived from classical mechanics and quantum mechanics, and are applied to the chemical sample by a number of computational applications that perform calculations based on specified parameters.

3. **PCMODEL Software**

PCMODEL Software has been obtained from Indian Institute of Science, Bangalore, by the courtesy of Prof. Chandrasekharan and has been installed in a P.C. in our Computer Laboratory. This software is used for molecular modeling and for geometry optimization to create data file for MOPAC input.

**Geometry Optimization**

Many systematic mathematical procedures (algorithms) exist to find a local minimum of a function of several variables. These procedures will find a local minimum in $E$ in the neighborhood of initially assumed geometry. The process of finding such a minimum is called Geometry Optimization or energy minimization. For a molecule with several
conformations, one must repeat the local minimum search procedure for each possible confirmation, so as to locate the Global Minimum. For a large molecule, there may be too many confirmations for all of them to be examined. Moreover, the true-global minimum equilibrium geometry might correspond to a highly unconventional structure that might not think to be considered.