CHAPTER 3

RESEARCH METHODOLOGY

Computational modelling is a mathematical model to study the behaviour of complex systems such as static and dynamic systems by using computational simulation. Computational chemistry approach can be divided to three main methods molecular mechanics, semi-empirical and quantum mechanics.

Molecular mechanics employs force fields based on classical mechanic principles to investigate the behaviour of atoms and molecules (Kollman et al., 1997). This is accomplished by utilizing a molecular force field, which represents the potential energy as a function of atomic positions, and studying molecular properties while disregarding electron movements. The energy expression is composed of classical equations, including the harmonic oscillator equation, to describe the energy related to bond stretching, angle bending, bond rotation, and intermolecular forces. The typical parameter sets are Merck molecular force field 94 (MMFF94), molecular mechanics 2 (MM2) and molecular mechanics 3 (MM3). As electrons are not explicitly considered for in the classical molecular mechanics, large systems consist of thousands atoms or more such as proteins, drugs and enzymes can be simulated to estimate conformational flexibility and relative stability. Later, molecular mechanics methods are extended to molecular dynamics to simulate the nuclear motion within a molecule at various time and positions by adopting quantum mechanics approximation.

The semi-empirical method is an approach based on quantum mechanics that eliminates electronic integration and incorporates exchange correlation effects and empirical parameters to minimize errors in approximations. It simplifies mathematical equations by combining quantum mechanics techniques with real lab data (Feller, 1996). Semi-empirical method is limited to the study of valence electron behaviour as they interact more readily with other molecules. The calculation of inner orbitals involves adding empirical data obtained from experiments. This method is suitable for studying the behaviour of molecules in a solvent and can calculate molecules with up to 1000 atoms. The semi-empirical method employs three main parameters: The Austin model 1 (AM1) predicts heat of formation, the parametric method 3 (PM3) is well-suited for organic systems, and the intermediate neglect of differential overlap methods (INDO).

Quantum mechanics can be classified into two groups which are Hartree-Fock method and density functional theory (DFT). The simplest wave function method is the Hartree-Fock (HF), in which neglecting the Coulombic electron repulsion explicitly, hence taking the average effect of multi-electron wave function using mean field approximation to calculate the total energy of the molecule (Chuvylkin et al., 2005). The common basis sets used are Slater type orbitals and Gaussian type orbitals. The lack of electron correlation affecting the exact kinetic barriers and description of intramolecular forces such as London dispersion force. Later, the improved quality of HF methods introduced largely extended basis sets such as coupled cluster theory (CC) and configuration interaction (CI) that include electron correlation. However, the advanced calculation of post-HF methods is considerably computationally costly, and their greater approximation makes them more suitable for large systems.

Density functional theory (DFT), in which the total energy is expressed in terms of the total electron density, leads to effective approximations using a Hamiltonian model (Baerends et al., 1997). In general, to study the properties of solid materials, microscopy investigations must be conducted on atoms by analysing electron behaviour. Knowing the position of electrons can describe energy changes in the atom. In quantum mechanics, researchers face difficulties in calculating each electron in an atom due to the high number of interaction between particles. Thus, the most plausible way to analyse electrons is by using the density of particles in a particular region. The common parameter sets used to estimate electron density are local density approximation (LDA), generalized gradient approximation (GGA), and hyper-GGA. DFT calculations give very good qualitative results and can provide increasingly accurate quantitative results in small systems consisting of approximately 100 atoms. The main methods of computational chemistry are summarized in **Table 3.1**.

Table 3.1: Classification of Computational Chemistry Methods

Method	Description	Size of system
Molecular mechanics	Uses classical mechanics force fields to explain and interpret the behaviour of atoms and molecules.	Up to 100, 000 atoms
Semi-empirical	A quantum mechanics-based approach that simplifies equations with the use of empirical parameters.	Up to 1000 atoms
Quantum mechanics	The total energy is expressed in terms of the total electron density.	Up to 100 atoms

The Kohn-Hohenberg and Kohn-Sham are the backbones of DFT development in 1960s (Kohn et al., 1996). Their mathematical theorems are proven to solve the many-body problems successfully (Sholl et al., 2009). Initially, Kohn and Hohenberg postulated that a functional of electron density could represent the ground state energy obtained from the Schrodinger equation (Sholl, 2009). Essentially, the ground state dictates all properties such as energy and wavefunction. However, the theory lacks a comprehensive explanation of the function and requires further elaboration Later, Hohenberg and Kohn developed the idea that the total energy of a system can be minimized by using the exact electron density. (Sholl, 2009). This provided a method for determining the relevant electron density once the actual functional form was established. The functional must be varied until the electron density reaches the minimum energy level, and the true functional is then formed. To express the correct electron density, Kohn and Sham proposed that the solution of the set of equations should be included, where each equation in the set accounts for a single electron (Kohanoff, 2006).

$$\varepsilon_i \psi_i(r) = \frac{h^2}{m} \nabla^2 + V(r) + V_H(r) + V_{XC}(r)] \psi_i(r) \tag{3}$$

ε_i is the energy of the *i*-th molecular orbital, ψ_i (r) is the wavefunction of the *i*-th molecular orbital. h is Planck's constant, m is the mass of an electron, ∇^2 is the Laplacian operator, V(r) is the potential energy of the electron due to its interaction with the nuclei, V_H(r) is the Hartree potential energy of the electron due to its interaction with all other electrons, and V_{XC}(r) is the exchange-correlation potential energy of the electron due to its interaction with all other electrons in the system.

The DFT approach has high accuracy for structural and electronic studies since it solves the exact Schrodinger equation using exchange-correlation functional. However, the selection of the right basis set and exchange correlation are very crucial to determine the accuracy because DFT approach acquires high computational hardware and plenty of time. This matter will be further discussed in next subtopic.

3.1 Exchange-Correlation

The development of accurate functional to represent the complete functional stay as the main topic in the most crucial fields of scientific research. The origin of energy functional emerge by local density approximation (LDA) that assumes the density is the same everywhere hence the calculation becomes simple but not accurate

(Sholl, 2009). To overcome this problem, a generalized gradient approximation (GGA) is constructed that used both the local electron density and the local gradient in the electron density. Later in 1990, rapid theoretical progress produce hybrid functional with the mixture of Hartree-Fock (HF) exchange correlation. The remodelling by hybrid functional improves the estimation of atomic energy, bond length, vibration frequency and others. The application of hybrid functional can be seen through several methods such as Becke, three-parameters, Lee-Yang-Parr (B3LYP) and Minnesota 06 (M06-2X) exchange correlation functional. The hybrid functional which is B3LYP has been recognized as the most popular method for DFT calculations because its less approximation on metal complex compounds and high accuracy are very close to experimental results (Becke, 2014). The Jacob's ladder approach has summarized the DFT functional according to the metaphor of Perdew and Schmidt (Gomes, 2013). As moving from LDA to GGA, metaGGA, and hybrid functional, the accuracy of the DFT calculation increases, but at the cost of greater computational complexity and decreased simplicity.

3.2 Basis Set

In computational modelling, a basis function or basis set is a set of functions used to represent the electronic wave function in DFT or HF codes, in order to convert partial differential equations into algebraic equations for efficient implementation on a computer. The minimal basis sets offered by the Gaussian program is the STO-3G consists of slater-type and Gaussian-type atomic orbitals (Hehre et al., 1969). Since valence orbitals of atoms are more affected by forming a bond than the inner (core) orbitals, split-valence basis sets were introduced to represent more basis functions to describe valence orbitals. Later, more basis functions are assigned to describe polarization and diffusion functions of the whole molecule such as Pople's basis sets. The main basis sets are 3-211+G(d,p), 6-31++G(d,p), and 6-311++G(d,p) widely used to describe the structural, electronic and optical properties of organic compound (Sakthi et al., 2017).

For transition metals in third to higher row elements, effective core potential (ECP) approaches are used to treat inner orbitals electrons as average potential rather than actual particles (Wadt et al., 1985). This pseudopotential effects are very important in describing heavy atoms accurately while saving computational effort. The common basis sets for transition metals is Los Alamos National Laboratory 2 double-zeta (LANL2DZ) that consist of double zeta (two polarization functions) quality and the overall combination of ECP effect (Dunning, 1977).

Based on previous research Pople's basis sets and ECP approaches are more reliable in treating metal complexes (Bahsis et al., 2020). Therefore, this study will conduct preliminary study on several basis sets to identify which functional producing the most stable optimized structure with lowest total energy.

3.3 Polarizable Continuum Model

Solvation is the interaction of a solvent with the dissolved solute, which leads to the stabilization of the solute species in the solution. The physical interaction between solute and solvent molecules which is electrostatic force influences the structure, spectra and other properties of solute (Jelle et al., 2018). Therefore, researchers built a solvation model when modelling reaction to estimate the solvent effects on the behaviours of solute in a realistic environment.

To the present day, the major approaches in computational chemistry regarding solvent effects are:

i- Explicit model: consider molecular details of each solvent molecule

ii- Implicit model: treat solvents as a continuum medium (Zhang et al., 2017)

iii- Hybrid model: apply quantum and classical calculation to different parts of a single molecule

The implicit solvation model has been frequently used over the years to calculate molecular free energy and describes a thermodynamic property which is the dielectric constant of solution accurately with affordable computation cost (Chen et al., 2019). The solvation process begins with the creation of a cavity then, turns on dispersion and repulsion forces and finally turning on electrostatic forces. The solute molecule is embedded in a cavity surrounded by a continuous medium where the solvent is modelled as a continuous mass rather than as discrete particles.

Among the continuum models, the polarizable continuum model (PCM) formulated by (Miertuš et al., 1981) is one of the most established approaches to describe molecular solute. The integral equation formalism of polarizable continuum model (IEFPCM) is the latest version of PCM by accounting for the effect of the outlying charge and it leads the way to correct reaction potential inside the solute cavity (Mennucci, 2012). The IEFPCM method must be specified in software package with the self-consistent reaction field (SCRF) method and the dielectric constant value for accounting the effect of polarizable solvent. **Table 3.2** summarize the solvents used in this research.

Table 3.2: Solvents list and dielectric constant value, ε

Solvent	Dielectric value, ε (Frisch et al., 2016)
Water	78.3553
Acetonitrile	35.688
Methanol	32.613
Chloroform	4.7113
Toluene	2.37
Hexane	1.8819

3.4 Hardware

The computer used in this simulation is hp model: Intel® Xeon® E-22233 Processor (3.6 GHz base frequency, up to 3.9 GHz with Intel® Turbo Boost Technology, 8.25 MB cache, 4 cores), with installed RAM of 16.00 GB and 64-bit of Windows operating system.

3.5 Software

Theoretical simulations of the Pd(II) tetraaza macrocyclic complex are conducted using DFT and its extension, TDDFT, implemented by the Gaussian 16 package. Gaussian input files will be created using GaussView 6 as a graphical interface. The field of theoretical chemistry has made numerous contributions to the understanding of molecular phenomena through computational models and algorithms used by academic and industrial software packages. Density functional theory has been translated into quantum mechanics modelling by many software platforms, such as the Gaussian 16 package.

The Gaussian 16 is an electronic structure program used to investigate chemical problems with a wide range of complexity using modest computer software. Generally, computational time increases with the accuracy of the calculation, size of the system, and hardware specifications. The level of accuracy also depends on the level of exchange-correlation and basis sets used. Therefore, it is crucial to find a balance between the method and equipment to achieve a reasonable level of accuracy within stipulated time frames.

3.6 Ground State Calculation

This section shows the geometry optimization and energy calculation of tetraaza macrocyclic ligand and Pd(II) tetraaza macrocyclic complex by DFT method in gas phase. The calculated bond parameters, MEPS, and NBO analysis provide an explanation for the structural and electronic properties of the tetraaza macrocyclic ligand and its complex. The steps are:

1. Protonated tetraaza macrocyclic ligand is constructed using GaussView.
2. The structure is optimized by applying 'Opt+Freq' job type in order to get lowest energy surface plot and zero imaginary frequency.
3. The computational variables are summarized in **Table 3.3** and the input is saved as *gjf* file.
4. Once calculation is completed, convergence is achieved, the results of bond parameters is saved in *log* file and MEPS visualization is saved in *chk* file.
5. Next, job type 'energy' is selected and keyword 'pop=nbo' is added on optimized structure of the ligand to observe the NBO analysis. The result is saved in *log* file.
6. Steps 1-5 are repeated for Pd(II) complex structure to complete the ground state calculation.

Table 3.3: Gaussian16 Input of Tetraaza Macrocylic Ligand and Its Complex

Compound	Method	Functional	Basis set	Charge
Ligand	DFT/Ground state	B3LYP	STO-3G 3211G ++** 6-311G ++ **	+2
Pd(II) complex	DFT/Ground state	B3LYP	LanL2DZ	+2

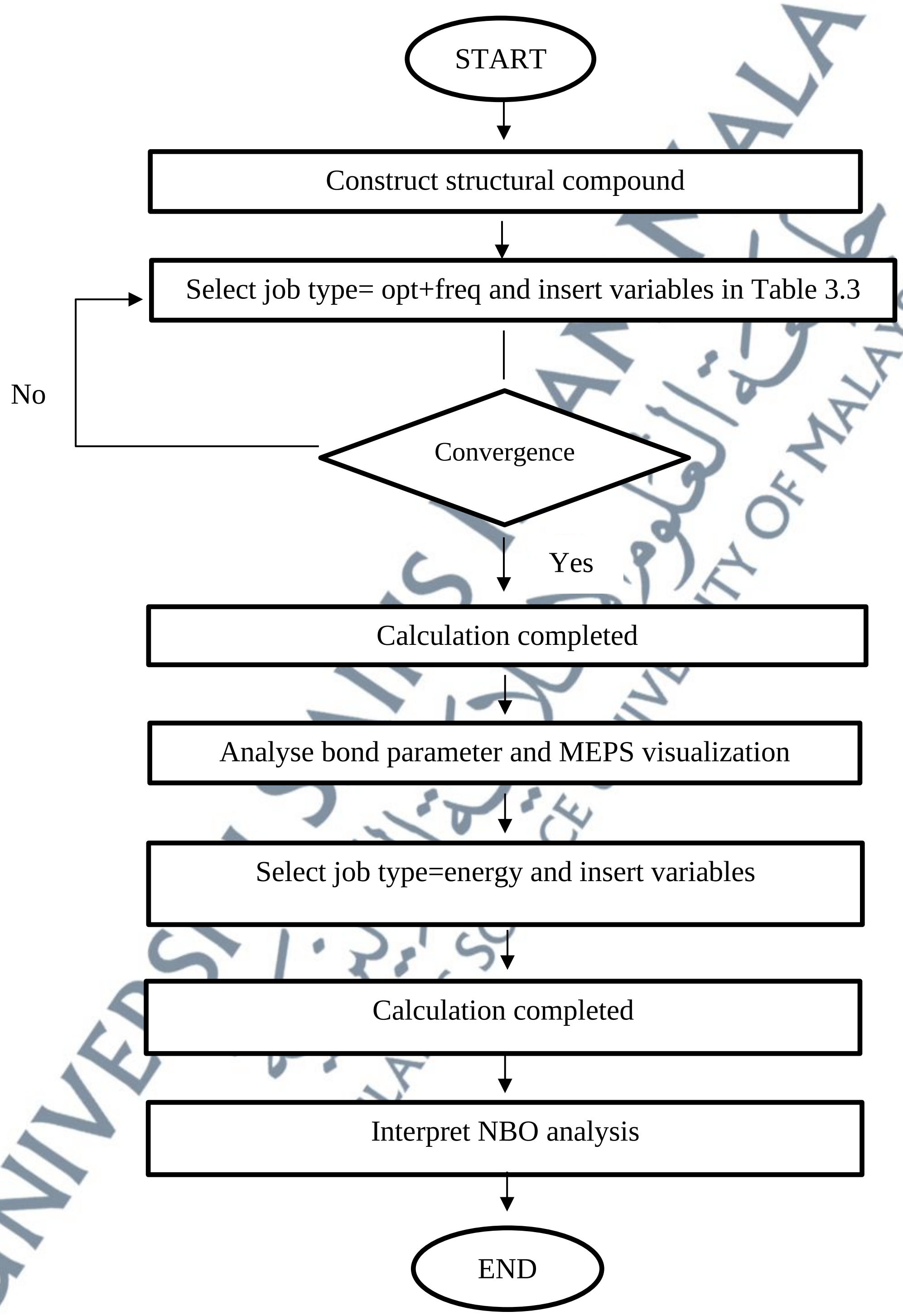

Figure 3.1: Flowchart of the Ground State Calculation

3.7 Excited State Calculation

This section shows the optimization, frequency and energy calculation of Pd(II) tetraaza macrocyclic complex by TD-DFT method in gas phase. From the calculations, the frontier molecular orbital, UV-Vis spectrum, and NLO properties explain the optical properties of the Pd(II) complex. The steps are:

1. Pd(II) complex is constructed using GaussView.
2. The structure is optimized by applying 'Opt' job type in order to get lowest energy surface plot at excited state.
3. The computational variables are summarized in **Table 3.4** and the input is save as *gjf* file.
4. Once calculation is completed, FMO visualization is saved in *chk* file
5. Next, job type 'Frequency' is selected and keyword 'polar' is added to observe NLO activity.
6. Once calculation is completed, NLO properties is saved in *log* file.
7. Next, job type 'Energy' is selected to get UV-vis spectrum. The result is saved in *log* file.

Table 3.4: Gaussian16 Input of Pd(II) Complex At Excited State

Compound	Method	Functional	Basis set	Charge
Pd(II) complex	TD-SCF/DFT	B3LYP	LanL2DZ	+2

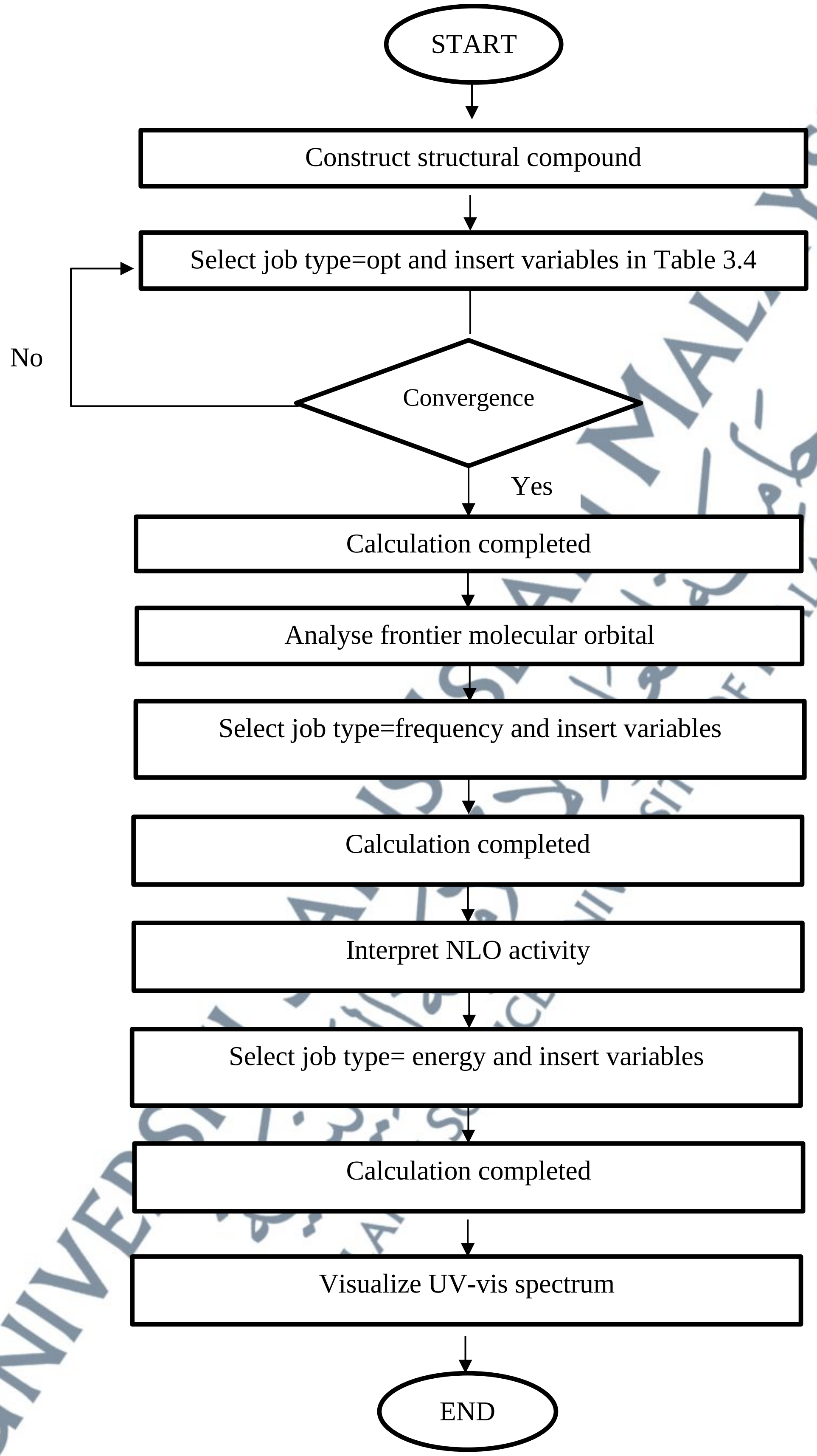

Figure 3.2: Flowchart of the Excited State Calculation

3.8 Solvation Model Calculation

This section presents the optimization, frequency, and energy calculations of the Pd(II) tetraaza macrocyclic complex using the TD-DFT method in a solvent environment. The calculations provide insights into the optical properties of the Pd(II) complex, including its frontier molecular orbital, UV-Vis spectrum, and NLO properties. The steps are:

1. Pd(II) complex is constructed using GaussView.
2. The structure is optimized by applying ‘Opt’ job type in order to get lowest energy surface plot in acetonitrile environment in ground state.
3. The computational variables are summarized in **Table 3.3** and the input is save as *gjf* file.
4. Steps 1-3 are repeated for different solvent which are methanol and water.
5. The structure is optimized by applying ‘Opt’ job type in order to get lowest energy surface plot in acetonitrile environment in excited state.
6. The computational variables are summarized in **Table 3.5** and the input is saved as *gjf* file.
7. Once calculation is completed, FMO visualization is saved in *chk* file
8. Next, job type ‘Frequency’ is selected and keyword ‘polar’ is added to observe NLO activity.
9. Once calculation is completed, NLO properties is saved in *log* file.
10. Next, job type ‘Energy’ is selected to get UV-vis spectrum. The result is saved in *log* file.
11. Steps 1-7 are repeated for different solvent which are hexane, chloroform, methanol and water.

Table 3.5: Gaussian16 Input of Pd(II) Complex in Solvent Phase

Compound	Method	Functional	Basis set	Charge	Model
Pd(II) complex	DFT/ Ground state	B3LYP	LanL2DZ	+2	IEF-PCM
Pd(II) complex	TD-SCF/DFT	B3LYP	LanL2DZ	+2	IEF-PCM

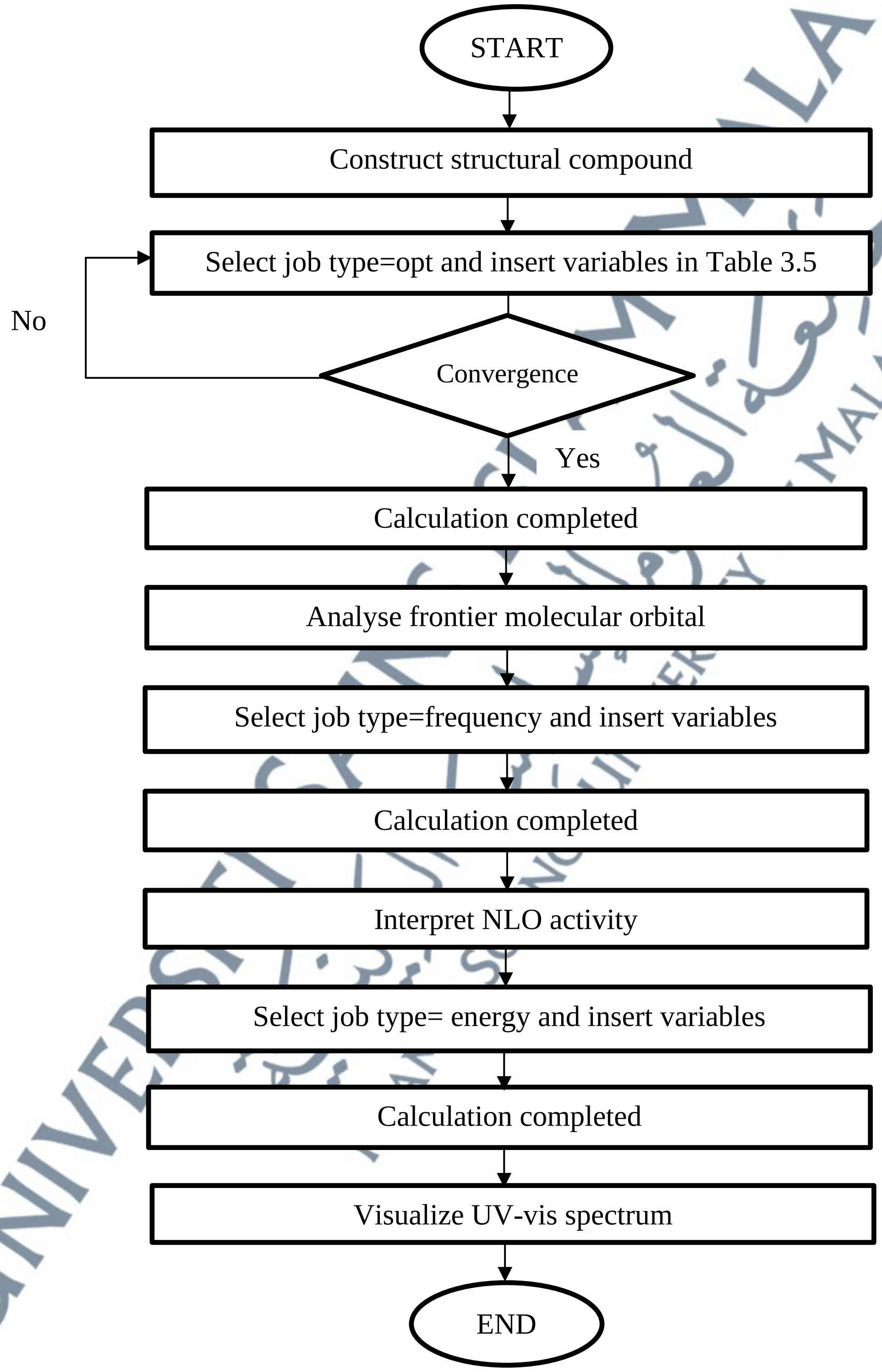

Figure 3.3: Flowchart of the Solvation Model Calculation