THE IMPACT OF COMPUTATIONAL METHODS ON TRANSITION METAL-CONTAINING SPECIES

Jiaqi Wang

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2015

APPROVED:

Angela K. Wilson, Major Professor
Thomas R. Cundari, Committee Member
Martin Schwartz, Committee Member
LeGrande Slaughter, Committee Member
Michael Richmond, Chair of the Department of Chemistry
Costas Tsatsoulis, Interim Dean of the Toulouse Graduate School

Wang, Jiaqi. *The Impact of Computational Methods on Transition Metal-Containing Species*. Doctor of Philosophy (Physical Chemistry), December 2015, 151 pp., 20 tables, 23 figures, references, 110 titles.

Quantum chemistry methodologies can be used to address a wide variety of chemical problems. Key to the success of quantum chemistry methodologies, however, is the selection of suitable methodologies for specific problems of interest, which often requires significant assessment.

To gauge a number of methodologies, the utility of density functionals (BLYP, B97D, TPSS, M06L, PBE0, B3LYP, M06, and TPSSh) in predicting reaction energetics was examined for model studies of C-O bond activation of methoxyethane and methanol. These species provide excellent representative examples of lignin degradation via C-O bond cleavage. PBE0, which performed better than other considered DFT functionals, was used to investigate late 3*d* (Fe, Co, and Ni), 4*d* (Ru, Rh, and Pd), and 5*d* (Re, Os, and Ir) transition metal atom mediated C_β -O bond activation of the β–O–4 linkage of lignin. Additionally, the impact of the choice of DFT functionals, basis sets, implicit solvation models, and layered quantum chemical methods (i.e., ONIOM, Our Own N-layered Integrated molecular Orbital and molecular Mechanics) was investigated for the prediction of pK_a for a set of Ni-group metal hydrides (M = Ni, Pd, and Pt) in acetonitrile. These investigations have provided insight about the utility of a number of theoretical methods in the computation of thermodynamic properties of transition metal hydrides in solution.

As single reference wavefunction methods commonly perform poorly in describing molecular systems that involve bond-breaking and forming or electronic near-degeneracies and are typically best described with computationally costly multireference wavefunction-based methods, it is imperative to *a priori* analyze the multireference character for molecular systems so

that the proper methodology choice is applied. In this work, diagnostic criteria for assessing the multireference character of 4*d* transition metal-containing molecules was investigated. Four diagnostics were considered in this work, including the weight of the leading configuration of the CASSCF wavefunction, C_0^2; T_1, the Frobenius norm of the coupled cluster amplitude vector related to single excitations and D_1, the matrix norm of the coupled cluster amplitude vector arising from coupled cluster calculations; and the percent total atomization energy, %TAE. This work demonstrated the need to have different diagnostic criteria for 4d molecules than for main group molecules.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those people that have helped me on this doctor dissertation and my graduate study at the University of North Texas (UNT).

I would first like to thank my research adviser, Professor Angela K. Wilson, without who nothing would have been possible. During my graduate school, her incredible guidance, encouragement, and patience have enriched my growth not only in my research but also my academic experience. I am also grateful to her for giving me opportunities to present my research at many international and national conference. Without those experience, I would never have been able to discover computational chemistry and appreciate it.

I am also thankful to my amazing coworkers in the Wilson's research group for their enthusiastic sharing of knowledge and techniques throughout my studies. I would like to thank Dr. George Schoendorff and Rebecca Weber for their continuous support and interesting discussions. I would also like to thank Lily Liu for her contributions to the study of C-O bond activation by transition metal atoms, Sivabalan Manivasagam for working on the investigation of diagnostic criteria for 4d transition metal-containing molecules, and Prajay Patel for his contributions to Chapter 5. Special thank goes to Dr. Cong Liu and Dr. Gbenga Oyedepo for helping to introduce me to the field of computational chemistry and their helpful comments on the study of lignin degradation.

I would like to extend my upmost gratitude to my parents, Yinshan and Guiying, for their endless support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Computational chemistry is an important branch of theoretical chemistry in which chemical, mathematical, and computing skills are applied to model the structures, energies, physical and chemical properties, and reactivates of atomic and molecular systems. Computational methods have been employed in many fields, such as physics, chemistry, biology, and geophysics, and are more favorable than traditional laboratory experiments when the materials of interest are too difficult to obtain, too dangerous, or too expensive. Computational chemistry helps chemists gain in-depth understanding about compounds. For instance, molecular bonding information that may not be obtained from experimental methods can be obtained from computational methods. Computational chemistry also assists in experimental chemistry as it can help make predictions beforehand.

Ab initio methods are based on the Schrödinger equation. They are used to solve the many-body Schrödinger equation and provide the energy and wave function that describe the system. The wave function is a mathematical function that can be used to calculate the electron distribution. The Schrödinger equation can only be solved exactly for molecules with one electron. Therefore, several approximations can be made to deal with electron correlation energy that is the energy difference between the real energy and the Hartree-Fock energy. The HF energy generally accounts for 99% of the real energy, and to account for electron correlation energy, the reference wave function must be expanded to more advanced calculations. Correlation effects are usually classified as dynamic or nondynamic correlation.

Density Functional theory (DFT) encompasses another family of computational chemistry methods, which can be used for large systems (hundreds of atoms) at a modest computational cost as compared with *ab initio* methods. The total energy calculated by DFT methods is expressed in terms of the total electron density rather than a wave function. Because of the size of molecules for which DFT can be utilized, DFT methods can be applied to a wide variety of problems. For example, in this work, DFT is used to investigate transition metal catalysis as overviewed below.

Because of the depletion and environmentally unfriendly properties of fossil fuels, renewable resources are needed to replace fossil fuels. Biomass, such as lignin, has been of great interest. Catalytic conversions of biomass to useful chemicals and energies include activation of C-O bonds, hydrogenation, and hydrodeoxygenation. Studies of the intrinsic properties of catalysts that control the reactions would help better elucidate and improve the reaction mechanisms. Since metal centers play important roles in the performance of the TM-based catalysts (both homogenous and heterogeneous), the reactivity of neutral metal atoms can provide a starting point to investigate the intrinsic behavior of TMs in catalysis and guide the design of novel catalysts.

Single reference methods, such as many- body perturbation theory, configuration interaction theory, and coupled cluster theory, may fail badly for systems with significant nondynamic correlation such as for bond-breaking or diradicals. Thus, it is important to *a priori* analyze the nondynamic correlation effect of molecular systems in order to gauge the utility of single reference methods. Over the past few decades, a number of different diagnostic tools have been suggested to aid in assessing dynamic/nondynamic correlation effect for molecule system.

In Chapter 2, a brief review of the theoretical methods most often used in computational studies is given. The Schrödinger equation, Hartree-Fock methods, post Hartree-Fock methods, density functional theory, several diagnostic tools, implicit solvation model, and the ONIOM method (defines two or three layers that are treated by different levels of computational methods within the structure of large molecules) are described. Basis sets associated with the use of these methods are also included in this chapter.

In Chapter 3 a computational investigation of the C_β-O bond cleavage of β–O–4 linkage of lignin by late 3*d* (Fe, Co, and Ni), 4*d* (Ru, Rh, and Pd), and 5*d* (Re, Os, and Ir) transition metal atoms is presented. The PBE0 functional performed better than other considered DFT functionals (BLYP, B97D, TPSS, M06L, B3LYP, M06, TPSSh, and B2PLYP) in predicting reaction energetics for the model studies of C-O bond activation of methoxyethane and methanol. The energetics of the C_β-O bond activation reactions of the β-O-4 linkage show that the 4*d* and 5*d* TM species tend to have lower activation enthalpies than the 3*d* TM species (except for Fe), and besides Os, the 5*d* TM species tended to form more stable products than the 3*d* and 4*d* TM species. Pd would be the most suitable metal TM catalyst as they catalyzed the flattest reaction pathways.

In Chapter 4, four diagnostic criteria have been examined to identify the suitability of single reference wavefunction based quantum chemistry methods for a set of 118 4*d* transition metal species. These diagnostics include the weight of the leading configuration of the CASSCF wavefunction, C_0^2; T_1, the Frobenius norm of the coupled cluster amplitude vector related to single excitations and D_1, the matrix norm of the coupled cluster amplitude vector arising from coupled cluster calculations; and the percent total atomization energy, %TAE, corresponding to a relationship between energies determined with CCSD and CCSD(T) calculations. New criteria, namely $T_1 \geq 0.045$, $D_1 \geq 0.120$ and %TAE ≥ 10% are herein proposed as a gauge for 4*d* transition metal containing molecules to predict the possible need to employ multireference (MR) wavefunction-based methods to describe the energetic and spectroscopic properties of a molecule.

In Chapter 5, three implicit solvation models, CPCM (conductor-like polarizable continuum model), COSMO (conductor-like screening model), and SMD (density-based solvation model), were utilized to calculate pK_a values of a set of Ni-group hydrides in acetonitrile using a direct thermodynamic scheme. The influence of factors such as the SCRF model used, the choice of DFT functionals and basis set within a hybrid two layers QM/QM approaches (ONIOM), atomic radii applied to build a cavity in the solvent, the expansion of size of the high-layer region in ONIOM calculations, and inclusion of Grimme's empirical dispersion correction to DFT functionals on the predictions of pK_a values was examined. The DFT functionals considered in this study include the local spin-density approximation (LSDA), generalized gradient approximation (GGA), meta-GGA, hybrid-GGA, hybrid-meta GGA, and double hybrid functionals. In considering the accuracy of DFT, the MADs of PBE and M06 functionals were investigated by varying the percentage of HF exchange. These investigations help ascertain more efficient and accurate theoretical models for computing the thermodynamic properties of transition metal hydrides.

CHAPTER 2 THEORETICAL BACKGROUND

2.1 Schrödinger Equation

Computational chemistry is a rapidly growing branch of chemistry that investigates chemical problems using various mathematical approximations and scientific computer programs. One of the most important foundations of computational chemistry is the time independent Schrödinger equation,[1–7] Eq. 2.1:

$$\hat{H}\Psi = E\Psi \qquad (2.1)$$

$\hat{H}$ is the shorthand form of Hamilton operator which takes into account five contributions (the kinetic energies of electrons and nuclei, the attraction between electrons and nuclei, and the inter-electronic and inter-nuclear repulsions) to the total energy of the system, E. Ψ is the normalized wavefunction which describes the electronic and nuclear structure of a giving system. For a given molecule, several energy levels (eigenvalues) are allowed by Eq. 2.1. Each allowed energy level has a corresponding wavefunction (eigenfunctions).

The Hamilton operator, $\hat{H}$, has the following form (Eq. 2.2) without external electric or magnetic fields and without taking account of relativistic effects for heavy elements:

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \qquad (2.2)$$

where *i* and *j* represent electrons, *k* and *l* represent nuclei, ħ is the Planck's constant divided by 2π, m_e is the mass of the electron, and m_k is the mass of the nucleus *k*. The distance between electron and nucleus, electrons, and nuclei are represented by r_{ik}, r_{ij}, and r_{kl} respectively. ∇^2 is the Laplacian operator, e is the charge on the electron, and Z is an atomic number.

The wavefunction itself does not have a simple definition or direct physical interpretation. However, the integral of the wavefunction and its complex conjugate ($|\Psi|^2$) over some region of space gives the probability of the electron that can be found within that region of space. The normalized integral of $|\Psi|^2$ over all space must be unity. The wavefunction has to meet several physical requirements so that its physical interpretation is viable. The wavefunction must be quadratically integrable, continuous, and single-valued. Since the Hamiltonian operator is Hermitian, the wavefunctions are orthonormal, i.e., for a one particle system,

$$\int \Psi_i \Psi_j dr = \delta_{ij} \qquad (2.3)$$

where the solutions of Schrödinger equation can be chosen to be orthogonal (the integral in Eq. 2.3 is equal to zero if $i \neq j$) and normalized (the integral in Eq. 2.3 is equal to one if $i = j$). Additionally, the variational principle states that any approximated wavefunction will yield an energy that is greater than or equal to the ground state energy, E_0 (Eq. 2.4).

$$E = \langle \Psi | \hat{\mathrm{H}} | \Psi \rangle \geq E_0 \qquad (2.4)$$

Thus, by varying and adjusting Ψ until the energy is minimized, the “best” solution of Schrödinger equations can be found using the variational principle.

2.1.1 The Born-Oppenheimer Approximation

The wavefunction discussed above is extremely difficult to evaluate for many particle molecular systems due to the correlation of the motions of particles. However, since nuclei are much heavier than electrons, their velocity as compared to that of the electrons are negligible. Therefore, the correlated motions of nuclei and electrons can be assumed to be decoupled, and the energy of electrons can be calculated with the nuclei in fixed position. That is, the nuclei kinetic energy can be set to zero and the inter-nuclear repulsion potential (V_{nn}) becomes an evaluated constant for a given set of fixed nuclear position but still contribute to the potential energy of the system. This assumption is called the Born-Oppenheimer Approximation.[8] In this case, the wavefunction is only dependent on the kinetic energy of electrons (T_e), the attraction potential between electrons and nuclei (V_{ne}), and the inter-electronic repulsion potential (V_{ee}), so the Hamiltonian becomes electronic Hamiltonian, as shown in Eq. 2.5. Application of the electronic Hamiltonian to the wavefunction results in the electronic Schrödinger equation as described in Eq 2.6.

$$\hat{H}_{elec} = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} \qquad (2.5)$$

$$\hat{H}_{elec}\Psi_{elec} = E_{elec}\Psi_{elec} \qquad (2.6)$$

Since the constant terms in the Hamiltonian do not affect the wavefunction, V_{nn} is generally not included for solving the electronic Schrödinger equation but added to the energies result from T_e, V_{ne}, and V_{ee}.

2.1.2 Slater Determinant

The wavefunction Ψ_{elec} is described by the spatial and spin quantum number, $\Psi = \psi(\vec{r})\sigma$, where σ is the spin component and under magnetic conditions can take two values, α (spin up, m_s = 1/2) or β (spin down, m_s = -1/2). The two spin functions are made into an orthonormal set.

$$\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1 \qquad (2.7)$$

$$\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle = 0 \qquad (2.8)$$

Since electrons are indistinguishable, the wavefunction must be anti-symmetric with respect to interchange of spatial or spin coordinates. This demonstrates Pauli's exclusion principle that the same set of quantum numbers can be only assigned to characterized one electron. [9,10] Thus, no more than two electrons can placed in a given molecular orbital. To achieve this mentioned antisymmetric properties, the wavefunction can be constructed by Slater determinant, which is an anti-symmetric product of N one-electron wavefunctions, $\chi_i(\vec{x}_i)$.[11]

$$\Psi_{SD} = \frac{1}{\sqrt{N!}} \begin{bmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \cdots & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & & \chi_N(\vec{x}_2) \\ \vdots & & \ddots & \vdots \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & \cdots & \chi_N(\vec{x}_N) \end{bmatrix} \qquad (2.9)$$

where N is the total number of electrons. The Slater determinant is normalized with factor 1/√N!. The single-electronic wavefunctions are described by the columns and the electron coordination are described by the rows.

2.2 Hartree-Fock Theory (HF)

In Hartree-Fock theory (HF),[12–16] a single Slater determinant is used and the electrons in the system are assumed to interact with each other only through a mean field of all the other electrons so that the electron correlations are neglected effectively. An iterative method, named the Self-Consistent Field method (SCF) is used to get the energy of a system, where an initial guess is made for the electron energies and then the solution is obtained interactively until successive potentials are identical (convergence is reached).

Two types of HF methods can be applied depending on if there are restrictions on the form of the spatial orbitals. If all of the electrons in a system are paired (a closed shell system), the restriction that two paired electrons must have different spins, one with α and other with β spin, is normally required. This situation is known as Restricted HF (RHF). For open shell species, two methods can be considered. If no restrictions is made for the form of the spatial orbitals, Unrestricted HF (UHF) can be used. If the doubly occupied orbitals is forced to have the same spatial parts, the system is solved using Restricted Open-shell HF. As RHF and ROHF put constraints on the spatial orbitals, they always result in higher energies then UHF. The wavefunction of UHF, however, is not an eigenfunction of the S^2 operator and may be contaminated by the contributions from higher lying states, e.g. a singlet state of UHF may also contains contributions from triplet, quintet, or even higher states.

2.3 Electron Correlation

As HF treats electron-electron interaction in an average way and the correlation between motions of electrons is effectively ignored, the total energy obtained from HF is always higher than the real energy because of the overestimated inter-electron repulsion. The electronic correlation energy is the energy difference between the real energy and the HF energy.

$$E_{correlation} = E_{total} - E_{HF} \qquad (2.10)$$

Although the electron correlation energy commonly only constitutes about 1% of the total energy, it can have significant effect on the properties of the system.[47] The electron correlation is usually introduced by taking account of the excited states of a given system by adding additional Slater determinants as the wavefunction expansions use the HF wavefunction as the starting point (Eq. 2.11). The linear combination of these additional excited state determinants with the HF wave function gives a new trial function that should be closer to the real system than the original determinant.

$$\Psi = c_0 \Phi_{HF} + \sum_{i=1} c_i \Phi_i \qquad (2.11)$$

where c_i are the coefficients defining the contribution of each excited state to the wavefunction.

Electron correlation can be loosely categorized as dynamic correlation or non-dynamic correlation.[17] Dynamic correlation refers to the motions due to the inter-electron interactions ignored by HF approximation and c_0 is much larger than any other coefficient. If the coefficients in the wavefunction expansions have similar magnitude as that of the HF wavefunction, the valence and low-lying virtual orbitals have strong correlation and are nearly (or exactly) degenerate. Nondynamic correlation needs to be considered for appropriate treatment of near degeneracy.[18,19] Dynamic correlation emphasizes the error from HF that arises due to the neglect of the correlated motion of the electrons, while nondynamic correlation emphasizes the insufficiency of HF wavefunction to describe near degeneracy situations because of using a single Slater determinant.

An accurate treatment of dynamic and nondynamic correlation effects can be achieved by the full configuration interaction (FCI) solution which includes all possible excited configurations that can be generated from the HF determinant.[20] However, despite the advances in the development of effective computational methods and improvement in capabilities of computational methods, FCI is still limited to model systems containing a small number of atoms and small basis sets due to its steep and unfavorable scaling with the size of molecules. The FCI wavefunction, therefore, needs to be approximated in order to study large size systems, which can be achieved by either single reference (SR) or multireference (MR) methods.

SR electron correlation methods (e.g. the configuration interaction single and double (CISD),[21] second-order Mϕller-Plesset perturbation theory (MP2),[22] coupled cluster with all single and double excitations (CCSD),[23] and coupled cluster method with single and double excitations and a perturbative treatment of triple excitations (CCSD(T))[24,25]) normally use HF wavefunction as the reference wavefunction and are able to recover a large portion of dynamic correlation energy even with truncation of excitations at the double or triple level, but fail badly to describe systems with significant nondynamic correlation energy. They differ in how they calculate the coefficients in front of the wavefunction expansions, c_i in equation 2.10.

2.3.1 Configuration Interaction

Configuration interaction theory is a linear variational method using a linear combination of the ground state HF wavefunction (Φ_0) and other possible excited determinants formed by exciting electrons from occupied spin orbitals to virtual/unoccupied spin orbitals (e.g., single excited Slater determinant, Φ_i^a, of excitation of one electron from *i*th occupied orbital to ath unoccupied orbital, the double excited Slater determinant, Φ_{ij}^{ab}, and the triple excited Slater determinant, Φ_{ijk}^{abc}, etc.), Eq. 2.12 .

$$|\Psi_{CI}\rangle = c_0\Phi_0 + \sum_i^{occ}\sum_a^{virt} c_i^a\,\Phi_i^a + \sum_{ij}^{occ}\sum_{ab}^{virt} c_{ij}^{ab}\,\Phi_{ij}^{ab} + \sum_{ijk}^{occ}\sum_{abc}^{virt} c_{ijk}^{abc}\,\Phi_{ijk}^{abc} + \cdots \tag{2.12}$$

where occupied orbitals are labeled by the indices i, j, k,…, while unoccupied orbitals by a, b, c,…. The coefficients ($c_0, c_i^a, c_{ij}^{ab}, c_{ijk}^{abc}$ ...) in front of the determinants are determined variationally.

Within the choice of basis set, the energy obtained from FCI, which includes excitations of all electrons, is the best possible calculations that can be done. The cost of FCI is related to the binominal coefficient (Eq. 2.13) and computationally intractable.

$$\binom{K}{N} \equiv \frac{(K)!}{(K-N)!N!} \qquad (2.13)$$

where K is the total number of HF orbitals and N is the number of electrons in the system. However, the number of possible excited determinants has to be truncated in order to maintain computational methods that are practical for chemically interesting systems. Truncation of equation 2.12 at zeroth-order is the HF method. Including the single excitations results in the configuration interaction singles (CIS), including both single and double excitations results in the configuration interaction single and double (CISD), and so on: CISDT (configuration interaction with single, double, and triple excitations). One unattractive feature of these mentioned truncated CI methods is that they are problematic for molecular dissociation processes, which is caused by two reasons. First, the truncated CI methods are not size-consistent (size-consistent methods predict the energies of two non-interacting subsystems A and B that are infinitely far apart to be equal to the sum of the energies of the isolated systems). For instance, the energy of an isolated helium atom can be described exactly by CISD as it only has two electrons, while the exact energy of a dissociated helium molecule requires CISDTQ (configuration interaction with single, double, triple, and quadruple excitations). Using only CISD gives an energy of the dissociated helium molecule which is greater than the sum of the energies of two isolated helium atoms.

Second, bonding and anti-bonding orbitals tend to approach each other as bonds are stretched, which makes nondynamic correlation becomes significant for molecular dissociation processors. Since truncated CI methods are SR methods, they cannot describe nondynamic correlation very well.

2.3.2 Coupled Cluster Theory

Coupled cluster (CC) theory[24–27] is one of the most commonly used post-HF quantum chemistry methods to account for electron correlation and includes electron correlation by using the exponential cluster operator to act on the reference wavefunction to generate excited determinants. The CC wavefunction is written in the form described by Eq. 2.14 where $e^{\hat{T}}$ is the exponential operator and can be written as in Eq. 2.15**:**

$$\Psi_{CC} = e^{\hat{T}}\Psi_0 \qquad (2.14)$$

$$e^{\hat{T}} = 1 + \hat{T}_1 + \Big(\hat{T}_2 + \frac{1}{2!}\hat{T}_1{}^2\Big) + \Big(\hat{T}_3 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1{}^3\Big) + (\hat{T}_4 + \hat{T}_3\hat{T}_1 + \frac{1}{2!}\hat{T}_2{}^2 + \frac{1}{2!}\hat{T}_1{}^2\hat{T}_2 + \frac{1}{4!}\hat{T}_1{}^4 + \cdots \qquad (2.15)$$

Where $\hat{T}_i$ is connected operator (acting on an HF reference wavefunction Φ_0 generate all i^{th} excited Slater determinants) and other operators in equation 2.15 are disconnected operators (products of connected operator). Physically, a connected operator such as $\hat{T}_4$ corresponds to simultaneous excitation of four electrons, while disconnected term such as $\hat{T}_2^2$ corresponds to product of two non-interacting excitation of pairs of interacting electrons, double excitation and then another double excitation. The first term in Eq. 2.15 generates the HF method. All single excited states are generated by the second term (Eq. 2.16). From double excited states to higher order excited states, excited Slater determinants are generated by both connected and disconnected operators. For instance, double excited states can be written as Eq. 2.17 and Eq. 2.18, triple excited states can be written as Eq. 2.19, Eq. 2.20, and Eq. 2.21, and so on.

$$\hat{T}_1\Psi_0 = \sum_{i,a} t_i^a \, \Psi_i^a \qquad (2.16)$$

$$\hat{T}_2\Psi_0 = \sum_{i>j,a>b} t_{ij}^{ab} \, \Psi_{ij}^{ab} \qquad (2.17)$$

$$\frac{1}{2}\hat{T}_1{}^2\Psi_0 = \sum_{\substack{i,a \\ j,b}} t_i^a \, t_j^b \Psi_i^a \qquad (2.18)$$

$$\hat{T}_3\Psi_0 = \sum_{i>j>k,a>b>c} t_{ijk}^{abc} \, \Psi_{ijk}^{abc} \qquad (2.19)$$

$$\hat{T}_1\hat{T}_2\Psi_0 = \sum_{\substack{i,a \\ j>k,b>c}} t_i^a \, t_{jk}^{bc} \Psi_{ijk}^{abc} \qquad (2.20)$$

$$\frac{1}{3!}\hat{T}_1{}^3\Psi_0 = \sum_{\substack{i,a \\ j,b \\ k,c}} t_i^a t_j^b t_k^c \, \Psi_i^a \qquad (2.21)$$

The expansion coefficients t_i are equivalent to the c_i coefficients in Eq. 2.12 for CI.

The FCI wavefunction can be achieved by including all possible excitations within equation 2.14, which is computational infeasible for all but the smallest systems. Therefore, as with the CI method, the cluster operator also needs to be truncated at some excitation level to result in practical CC methods, such as CCSD (coupled cluster with all single and double excitations). Unlike truncated CI methods, truncated CC methods are size consistent due to the presence of the disconnected operators that indirectly introduce additional effects of higher order excited states arising from products of connected operators (e.g. product of $\hat{T}_1\hat{T}_2$ generates triple excitation state, even the T operator is truncated at the double operator), but they are not varitaional.[28]

The coupled cluster method including higher excitation states, such as CCSDT (scales as N^8) is generally computationally very expensive computationally and applicable for small systems. Alternatively, the triple excitation can be perturbatively included to the CCSD method (CCSD(T)) developed by Raghavachari et al.[24] Comparing with CCSDT, CCSD(T) considerably saves computational cost as its scale of N^7, but without significant loss in accuracy.[29] CCSD(T) has been used in obtain quantitatively accurate results in quantum chemical calculations.[30] However, CCSD(T) fails to describe the potential energy surfaces for bond breaking reactions.[31–36] The completely renormalized coupled cluster method (CR-CCSD(T)) designed by Piecuch et al. to improve the results of the CCSD(T) calculations in bond breaking regions of molecular potential energy surfaces.[32,33] In a study of H-CH3, H-Cl, Cl-Cl, H-SiH_3, H_3C-CH_3, H_3Si-SiH_3, H_3C-SiH_3, Cl-CH_3, and Cl-SiH_3, CR-CCSD(T) recovered an energy within 1 kcal mol^{-1} of Full-CI for bond-difference between the energy of the equilibrium structure and structure with large bond length (RC-F = 6.5 bohr) of CH_3-F, CR-CCSD(T) calculated the dissociation energy to be the same as determined via MRCI(Q).[36]

2.3.3 MR Character

While the SR methods mentioned above are able to recover a large portion of the dynamic correlation energy, they may be insufficient to describe the molecular systems contains a significant amount of nondynamic correlation energy (e.g., degenerate and quasidegenerate states, excited states, and bond breaking process) that can be recovered by MR methods. However, the application of MR methods are commonly more complicated, more expensive, and much less user-friendly than SR methods, *a priori* analysis of the domination of dynamic (SR character) and nondynamic (MR character) correlation effects of molecular systems is highly recommended so that suitable computational methods can be employed in the studies. Over the past few decades, several diagnostic tools have been suggested to examine the SR/MR character for molecular systems, such as the Frobenius norm (T_1) and matrix 2-norm (D_1) of coupled cluster amplitudes for excitations,[37–39] the weight of leading configuration of a complete active space wave function (C_0^2),[40,41] and the percentage of triples contribution to the atomization energies (%TAE)[42,43] based on CCSD(T) calculations.

The leading coefficient C_0 (or its corresponding weight C_0^2) in a self-consistent-field (SCF) singles and doubles configuration interaction (CISD) wave function is one earlier diagnostic. Since this diagnostic is based on canonical HF orbitals, which might fail for molecular systems where HF determinant is imperfect, a misleadingly large coefficient may be obtained for HF determinant.[44] Alternatively, the C_0 coefficient in full-valence CASSCF[45] (complete active space self-consistent-field) wave function which is based on balanced CASSCF natural orbitals can be used to determine MR character directly.[40,41] CASSCF method is one of the most successful multi-configurational self-consistent-field (MCSCF) variant that is a general method for obtaining qualitatively correct wavefunctions for molecular systems with MR character.[46] Its success can be attributed to inclusion of all possible determinants (with the proper symmetry) that can be formed by partition the electrons and molecular orbitals into active and inactive space. The inactive molecular orbitals are typically either doubly occupied and empty orbitals that are assumed to be chemically uninterested space; while the active molecular orbitals are typically some of the highest valence and some of the lowest virtual orbitals that will change significantly during the chemical processes and are treated with full CI. The accuracy of CASSCF is therefore depended on the selection of the MCSCF active space that has to be decided manually by considering the molecular system at hand and the computational cost.[47] In comparison to CISD, CASSCF has the advantages of size-consistency and ability to describe both excited states and the ground state. $C_0 < 0.095$ ($C_0^2 < 0.90$) indicates systems with significant MR character.[44] However, determination of this C_0^2 diagnostic is not tractable for large molecular systems due to the prohibitive computational cost of full CI calculations for the active electrons and orbitals.

As an alternative, the quality of the HF wave function can also be examined by the T_1 diagnostic of Lee and Taylor based on the coupled cluster wave function so that the computationally complicated and expensive CAS calculations can be avoided. T_1 diagnostic is defined as the Frobenius norm of the single substitution amplitudes vector (t_1) of the closed shell CCSD wave function starting from a restricted HF orbitals divided by the square root of the number of correlated electrons, $T_1=\|t_1\|/N^{1/2}$ to meet the size consistency requirment.[38] The T_1 diagnostic examines the description of the whole molecule by coupled cluster methods but cannot indicate the small problem area of the molecule that cannot be described properly by coupled cluster methods. The SR methods can yield reliable results with T_1 diagnostic smaller than 0.02 suggested by Lee and Taylor.[48] The D_1 diagnostic by Janssen and Nielsen,[49] which is based on the matrix 2-norm of t_1 from single excitations of closed-shell CCSD wave function, indicates the small problem area. A D_1 diagnostic greater than 0.05 was suggested by Lee and Janssen to imply that SR methods may not give reliable description.[39] As Lee pointed out,[48] since T_1 and D_1 diagnostics describe molecules in different ways, these two diagnostics were suggested to be applied in tandem. The T_1/D_1 ratio indicates the homogeneity of the molecular electronic structure. Lee notes that the molecular systems have the T_1/D_1 ratio less than $1/\sqrt{2}$ that is the approximate mathematical relationship for completely homogeneous systems; as the values of ratio becomes further from $1/\sqrt{2}$, the non-homogeneity of the electronic structure becomes greater. If the T_1/D_1 ratio of the molecules becomes much smaller than $1/\sqrt{2}$, the coupled-cluster wave function becomes unreliable for some areas in the molecule but performed well for other areas of the molecules.

An energy based diagnostic, %TAE$_e$[T_4 + T_5], was proposed by Martin and coworkers during the development of W4 theory.[42,43] The %TAE$_e$[T_4 + T_5] diagnostic is the percentages of the total atomization energy (TAE) accounted for by the connected quadruple and quintuple excitations, which can be used as an indicator for the imperfections in CCSD(T) due to the MR character of the molecular systems. However, the application of %TAE$_e$[T_4 + T_5] diagnostic is limited to small size molecular systems due to its immense computational cost. An *a priori* indicator, which can perform as a predictor of %TAE$_e$[T_4 + T_5] but is more computational efficient than %TAE$_e$[T_4 + T_5], would be highly desirable. Of the several variants (also as by-products of W4 theory), such as %TAE$_e$[(T)], %TAE$_e$[SCF], and %TAE$_e$[post-CCSD(T)], %TAE$_e$[(T)] (the total atomization energy resulting from triple excitation) provides strongest correlation with %TAE$_e$[T_4 + T_5]; e.g., R^2 of 0.941 between %TAE$_e$[(T)] and %TAE[(T_4 + T_5)]) over the whole W4-11 dataset of 140 species. Martin and co-workers suggested %TAE$_e$[(T)] < 2% as the cutoff for reliable SR calculations, %TAE$_e$[(T)] between 2-5% for mild nondynamic correlation, %TAE$_e$[(T)] between 5-10% for moderate nondynamic correlation, and %TAE$_e$[(T)] > 10% for significant nondynamic correlation based upon the W4 set of molecules.

2.4 Density Functional Theory

Density functional theory (DFT), built on the Hohenberg-Kohn theorem[50] and Kohn-Sham equations[51], has been used quite extensively in the last few decades because of its reliability and capability to deal with relatively large systems at relatively low computational costs.[52] The basic concept of DFT is that the ground state energy of a system depends only on the electron density.[53] One particular electronic energy can be defined by one and only one particular density and vice versa. Therefore, the electronic states of atoms, molecules, and materials are described in terms of the three-dimensional electronic density, which is a great simplification over wave function theory, which describes electronic properties by calculating or making approximations upon the 3N-dimensional anti-symmetric wave function for a system with N electrons. Thus, the complexity of the electron density is not affected by the size of the system, but is related to the coordinates of the electron density exclusively.

The Hohenberg-Kohn[50] theorem marked the beginning of modern DFT, it yields a poor representation of the kinetic energy of a system; however, the Kohn-Sham (KS) density functional theory,[51] the most widely used theory in quantum chemistry, resolves the problem within Hohenberg-Kohn theorem by introducing an approximate form of kinetic energy by building a fictitious non-interacting system from a set of one-electron functions, where each electron moves in an average repulsion field of surrounding electrons.[54] In the KS equation, the total energy of system is divided in the following parts (Eq. 2.13),

$$E_{DFT[\rho]} = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \qquad (2.13)$$

where $T_s[\rho]$ is the kinetic energy of the hypothetical non-interacting electron system. J is the Coulomb interaction of electrons (or Hartree energy). E_{ne} is an external potential arising from the nuclei, and E_{xc} is the energy which is omitted from the previous terms because of using the idea of a non-interacting electron system, such as electron exchange, correlation energy, the portion of the kinetic energy referring to the differences between the non-interacting and the real system, and self-interaction caused by replacing the exact HF exchange.[55,56] Since the exact form of E_{xc} is unknown, developments of different DFT functionals target to more accurately describe the exchange-correlation energy. Perdew first proposed Jacob's Ladder which classically presents the hierarchy of the functionals with different approximations to the E_{xc}.

The approximation methods that will be discussed here include local density approximation (LDA) (first rung), generalized gradient approximation (GGA) (second rung), meta-GGA (third rung), hybrid-GGA (fourth rung), and double-hybrid-GGA (fifth rung). DFT functionals with dispersion corrections will also be discussed.

2.4.1 Local Density Approximation

LDA[57] depends only on density of a uniform electron gas and is the basis for exchange-correlation functionals. The density of LDA functional is defined as $\rho = \frac{N}{V}$ (the density is equal to the number of electrons *N* divided by the volume of the gas *V*), which is only used in LDA functionals. Although LDA functionals are roughly approximated, the form of its exchange-correlation functionals is known exactly or to a very high accuracy. A more advanced approximation which considers spin polarization in the case of open shell systems, named the local spin density approximation (LSDA), was developed. The electron density, ρ, of LSDA functionals is replaced by the spin electronic densities, ρ_α and ρ_β. LDAs generally describe molecular properties, such as structures and vibrational frequencies, more accurately than the HF method, however, they poorly characterize chemical reaction energetics, such as bond energies and energy barriers. A popular LSDA functional is VWN[58] developed by Vosko, Wilk and Nusair.

2.4.2 Generalized Gradient Approximation

Since the electron density of LDA is a constant value which cannot reflect the rapid variation of densities in a molecule, the generalized gradient approximation (GGA)[59–64] improves the performance of DFT by considering how electron density changes, in other words the gradient of the electron density ($\nabla\rho$), as well as the density itself. GGAs are usually divided into exchange and correlation terms that can be solved individually. For example, B is an exchange functional developed by Becke,[64] P86 is a correlation functional developed by Perdew,[62] and LYP is a correlation functional developed by Lee, Yang, and Parr.[60] Combinations of exchange and correlation functionals result in more complete descriptions of systems. The most widely used combinations include BLYP,[60,64] PBE,[62,63] and BP86.[62,64]

2.4.3 Meta-GGA

Meta-GGA functionals are expansions of the pure GGA functionals, adding the first derivative of the density into the exchange correlation functional, by including the second derivative of the electron densities and/or local kinetic energy densities, $\nabla^2\rho$, in the exchange correlation functional. Common meta-GGA functionals include TPSS,[65] M06L,[66] and BB95.[67] In gas phase studies of molecular properties, these functionals have been shown to offer more accurate results than LDA functionals and GGA functionals with similar computational cost as GGA functionals.

2.4.4 Hybrid-GGA

The self-interaction problem, i.e., the spurious interaction of an electron with itself, causes the exchange part of density functional to be poorly described. On the other hand, the exchange parts in HF are defined exactly so the self-interaction is cancelled. Thus, hybrid-GGA functionals combine the “exact exchange” from HF theory with some conventional treatment of DFT exchange and correlation, such as GGA, to improve the performance of density functionals. The B3LYP functional was the first generally applicable hybrid GGA functional widely used in chemistry.[68–70] Sousa et al.[71] found that 80% of the references included B3LYP in the Web of Science over the years 1999-2006 by analyzing the number of different functional names in article titles and abstract. In B3LYP, three empirical parameters are used to control the combination of the HF exchange and density functional exchange and correlation, as shown below (Eq. 2.14):

$$E_{XC}^{B3LYP} = E_{XC}^{LDA} + a_0 \left(E_X^{HF} - E_X^{LDA}\right) + a_X \left(\underbrace{E_X^{GGA}}_{B88(B)} - E_X^{LDA}\right) + a_C \left(\underbrace{E_C^{GGA}}_{LYP} - E_C^{LDA}\right) \tag{2.14}$$

where $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.8$. This functional has shown great success in predicting geometries of organic molecules in gas phase.[72]

2.4.5 Double-hybrid-GGA

Double-hybrid-GGA functionals introduce nonlocal correlation effects by combining standard hybrid-GGA functionals with a second-order perturbative treatment based on KS orbitals.[73,74] The exchange correlation energy can be expressed as Eq. 2.15:

$$E_{XC} = (1 - a_X)E_X^{DFT} + a_X E_X^{HF} + (1 - a_C)E_C^{DFT} + a_C E_C^{PT2} \tag{2.15}$$

The double-hybrid-GGA functionals are considered in the 5th rung in Perdew's Jacob's ladder, where unoccupied Kohn-Sham orbitals are included in the calculations. The double-hybrid-GGA functionals may provide better description of the kinetic barriers and diffuse orbitals than DFT functionals from lower rungs of Perdew's Jacob's ladder, as pertubational correlation is considered and can correct HF deficiencies while improving the self-interaction error; however, double-hybrid-GGA functionals need more time to meet the convergence due to inclusion of the perturbation. For instance, the run time is roughly twice of a B3LYP job.[75] This functional family includes Grimme's B2PLYP and mPW2PLYP,[76,77] and Truhlar's double hybrids.[73]

2.4.6 Dispersion Correction for Density Functional Theory

Dispersion-corrected DFT functionals were proposed by Grimme to describe long-range electron correlations better in comparison with conventional GGA functionals mentioned above including hybrid-GGA functionals.[78–80] These functionals are not new functionals but a modification of conventional functionals with the empirical correction for dispersion with an add-on energy term. The dispersion correction energy is presented in Eq. 2.16.

$$E_{DFT-D} = E_{KS-DFT} + E_{disp} \qquad (2.16)$$

where $E_{KS\text{-}DFT}$ is usually the Kohn-Sham energy and E_{disp} is an empirical dispersion energy. For instance, B3LYP-D3 denotes a calculation with the usual B3LYP functional plus a D3 dispersion correction energy term. Since the dispersion correction contributes to forces acting on the atoms, dispersion-corrected DFT functionals are expected to lead to a different optimized geometry than arises from the use of the corresponding conventional DFT functionals.

2.5 Implicit Solvation Models

The computational and theoretical consideration presented in the previous sections have been used to investigate thermodynamic properties and chemical reactivity of gas phase molecules enormously due to the simplicity of the gas phase modeling. However, many practical chemical reactions do not occur in the gas phase; much of chemistry and biochemistry takes place within solvents. The interaction between the solvent and the solutes can alter the chemical properties of molecules of interest, such as charge distribution, geometry, vibrational frequency, and chemical reactivity. For instance, the Menschutkin reaction[81] (nucleophilic substitution of a halide by an amine produces an alkylammonium cation and a halide anion) is unfavorable in gas phase because of the difficult separation of opposite charges, while this reaction can take place in the present of water that allows the solvation of anions. Therefore, the effects of solvents molecules around the studied molecules are necessary to considered.

2.5.1 Solvation Models

The methods for solvation computation includes two types: implicit solvation models[82,83] that treats the solvent as a continuum with a dielectric constant ε and explicit solvation models that treats every solvent molecule individually. An explicit treatment of solvent using electronic structure methods would include 100s to 1000s of solvent molecules surrounding the solute molecule explicitly, which is extremely computationally expensive. Efficient simplification of calculations involving solvation effects can be achieved by placing the solute in a suitably shaped cavity in the medium that is defined by a dielectric constant of ε.

The solvation free energy has three main components (Eq. 2.17): the cavitation that is the energy for creating a hole in medium is the destabilization component; dispersion and repulsion forces for the interactions between solute and solvent are the stabilization component; and an electrostatic force that appears as a consequence of polarization between solute and solvent, induced by the charge distribution in the solute, is also a stabilization component. High order couplings among these components may lead to very accurate theoretical models that have not achieved yet in the field of solvation modelling. For polar solutes and solvents, the solvation free energy is dominated by the electrostatic interaction. The contributions from other components that are short range interactions and inherently are of quantum mechanical nature are relatively small in comparison to that from the electrostatic interaction. The electrostatic interaction is long range interaction and can be understood classically.

$$\Delta G_{solvation} = \Delta G_{cavity} + \Delta G_{dispersion} + \Delta_{elec} \quad (2.17)$$

2.5.2 Continuum Solvation Models

The dielectric continuum/implicit solvation model is usually combined with *ab initio* computations where a statistically averaged solvent effect is represented by dielectric constant (ε) for a macroscopic measurement of polarizability. In this case, instead of using 100s-1000s of solvent molecules to qualitatively describe the solvent effect, meaningful results can be obtained by a single calculation.

In principle, Poisson’s equation (Eq. 2.18), which is at the heart of all continuum solvation models, describes the connection between the electrostatic potential $\Phi(r)$, the charge distribution $\rho(r)$, and the dielectric constant ε. For the situation where the medium is a continuum dielectric in a linear relation to the embedding charge, the Poisson equation can be written as:

$$\nabla^2 \Phi(r) = -\frac{4\pi}{\varepsilon}\rho(r) \quad \text{(Eq. 2.18)}$$

For the more common situation that treats the charge distribution inside a cavity, the Poisson equation is properly written as:

$$\nabla\varepsilon(r) \cdot \nabla\phi(r) = -4\pi\rho(r) \quad \text{(Eq. 2.19)}$$

In this case:

$$\nabla\varepsilon(r) \cdot \nabla\phi(r) = -4\pi\rho(r) \qquad r \in V_{in} \qquad \text{(Eq. 2.20)}$$

$$\nabla\varepsilon(r) \cdot \nabla\phi(r) = 0 \qquad r \in V_{out} \qquad \text{(Eq. 2.21)}$$

$$\rho(r) = 0 \qquad r \in V_{out} \qquad \text{(Eq. 2.22)}$$

$$\varepsilon(r) = 1 \qquad r \in V_{in} \qquad \text{(Eq. 2.23)}$$

$$\varepsilon(r) = \varepsilon \qquad r \in V_{out} \qquad \text{(Eq. 2.24)}$$

where V_{in} and V_{out} are the volumes inside and outside the cavity, respectively.

If the Boltzmann distribution of ions instead of a point charge is needed, Poisson's equation can be modified to the Poisson-Boltzmann' equation.

Born[84] first proposed an implicit solvation model with the formula in Eq. 2.25 (q: net charge, a: radius of the cavity, ε: dielectric constant) at the beginning of the 1920s that built upon the idea of a solute having a spherical cavity within the solvent where the interaction between solute and solvent is based upon the net charge.

$$\Delta G_{solvation} = -(1-\frac{1}{\varepsilon})\frac{q^2}{2a} \qquad (2.25)$$

Kirkwood[85,86] improved Born's idea by considering the dipole moment within atoms and molecules instead of the net charge that does not have polarizability (Eq. 2.26). The Born contribution is the first term in the above expansion that includes all possible multiple moments M_{lm}. Perturbation of the expansion to a second term produces the well-known Onsager model[87] (Eq. 2.27) for the solvation energy of dipolar molecules, which results in poor results for molecules in which the dipole moment fails to describe the electron distribution.

$$\Delta G_{solvation} = -\frac{1}{2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(\frac{\varepsilon-1}{\varepsilon+l/(l+1)}\right)\frac{{M_{lm}}^2}{R^{2l+1}} \qquad (2.26)$$

$$\Delta G_{solvation} = -\frac{\varepsilon-1}{2\varepsilon+1}\frac{\mu^2}{a^3} \qquad (2.27)$$

Although, these three models mentioned above can be solved analytically, they only estimate the solvent effect roughly due to the using either the spherical or ellipsoidal cavities that are only realistic for small and symmetric molecules. Therefore, quantitative results commonly cannot be obtained by those models.

Introducing the notion of the self-consistent reaction field (SCRF) allows the Schrödinger equation to be easily applied in solution by applying the dipole moment operator μ (Eq. 2.28). In the SCRF model, the dipole moments of solute within the cavity polarizes the dielectric medium, which induces an electric field in the cavity that in turn alters the dipole moments of solute slightly, etc. This mutual polarization between solute and solvent continues until a self-consistent equilibrium is achieved. The Schrödinger equation for a self-consistent continuum solvation model is:

$$\left\{H - \frac{1}{2}\left[\frac{2(\varepsilon-1)}{(2\varepsilon+1)}\right]\frac{\langle\psi|\mu|\psi\rangle}{a^3}\mu\right\}\psi = E\psi \qquad (2.28)$$

2.6 ONIOM

Despite advances in the development of fast and accurate quantum mechanical (QM) methods for calculating geometries, energies, reactivates, and electronic properties for molecular systems, it is still extremely time-consuming and even impractical to run calculations on large and complex molecular systems. ONIOM (Our Own N-layered Integrated molecular Orbital and molecular Mechanics) scheme,[88–93] a hybrid combination of two or more computational methods, makes the accurate exploration of the chemical properties of very large systems feasible. In the two-layer ONIOM scheme, a large molecule is divided into two fragments. The chemically interesting region of the system, such as bond breaking and formation, is described with the accurate and expensive QM approaches while the remaining part of the system is described with less accurate but more efficient low level approaches, such as QM methods combined with a low level basis set or MM methods. In our research, the two-layer ONIOM (QM/QM) framework is applied, which calculates the total energy of the system as (Eq. 2.29):

$$E^{ONIOM(QM:QM)} = E_{model}^{high} + E_{real}^{low} - E_{model}^{low} \quad (2.29)$$

As described in Equation 2.29, the real system (contains all atoms) is calculated with the low level QM approach, which in our study uses DFT with small basis sets. The model system, containing the chemically important region and the link atoms that are used to cap dangling σ bonds or radicals resulting from cutting covalent bonds between the high level region and low level region, are treated with both low level QM approaches and high level QM approaches (our study uses DFT functionals with large basis set).

2.7 Basis Sets

Basis functions are designed to create the atomic orbitals in the system and are usually expanded as a linear combination of atomic orbitals (LCAO) with the coefficients to be determined. There are two main types of these basis functions: Slater-type orbitals (STOs)[94] and Gaussian-type orbitals (GTOs).[95–97] STOs have the exponential dependence, $e^{-\zeta r}$ (Eq. 2.30):

$$\Psi(r,\theta,\Phi) = NY_l^m(\theta,\Phi)r^{n-l}e^{-\zeta r} \qquad (2.30)$$

where N is a normalization factor, ζ is the exponential parameter, the *n*, *l*, and *m* are classical quantum numbers (principle, angular momentum and magnetic, respectively), the *r*, *θ*, and *φ* are spherical coordinates, and *Y* is the angular momentum part. STOs feature the electron-cusp condition (a cusp at *r*=0) and exponential decay of the function at large values of r (distance from nucleus), and therefore describe the behavior of hydrogenic atomic orbitals very closely, but no analytical solution is available for the general four-center two-electron integrals in HF theory.[98] GTOs, on the other hand, have the exponential dependence, $e^{-\zeta r^2}$, (Eq. 2.31):

$$\Psi(r,\theta,\Phi) = NY_l^m(\theta,\Phi)r^{(2n-2-l)}e^{-\zeta r^2} \quad (2.31)$$

GTOs are computationally more convenient than STOs, which are very time consuming because of the four-center-two-electron integrals. However, GTOs fail to show a cusp at *r*=0 (proper radial shape) and decrease rapidly for large value of *r* because of the squared radius. These specific features diminish the utility of GTOs as basis functions.[99] To solve these problems, a linear combination of GTOs, called contracted GTOs, were introduced to approximate STOs, in which each primitive GTO (the individual GTO that is used to form the contracted GTOs) is defined with a weight coefficient.[47] A set containing the number of basis functions that include the smallest number of contracted functions required to describe each atomic orbital is called a minimum basis set, also known as a single ζ basis set.[4] The STO-3G basis set is the most common single-ζ basis set or a 'minimal' basis set, and is formed by a linear combination of three primitives GTOs for each basis function (contracted GTOs) defined for each type of atomic orbital (core and valence. Using two or more contracted GTOs per atomic orbital can describe the system more accurately. For instance, the double ζ (DZ) basis sets double the basis functions of a minimum basis set for each type of atomic orbital and triple ζ (TZ) basis sets triple the basis functions, and so on.[47]

In addition to multiplying the number of basis functions, polarization functions and/or diffuse functions are used to add more flexibility to the basis set and describe chemical bonding interactions better, respectively. As atoms approach each other to form a bond, the electron density of these atoms will exhibit some distortions, called polarization. To take account of this effect through the basis sets, functions with higher angular momentum are needed to augment basis sets. For example, a set of *p*-orbital can be added to H to allow polarization. In the same way, addition of *d*-functions to the *p*-orbitals of the first row elements, and *f*-functions for *d*-orbitals account for the polarization, etc. The addition of diffuse functions is important when electrons are held at large distance from the nuclei, such as the interaction including anions or the second or third row transition metals.[22] In general, diffuse functions have a smaller exponent than the regular basis functions, so they will extend far from the nuclei.

Valence orbitals are much more strongly affected by chemical bonding than core orbitals, so it is more important to have flexibility in the valence basis functions, so-called "split-valence" basis sets. In such basis sets, the core and valence orbitals are treated differently. Pople and co-workers developed the K-NLMG basis sets, which are the most widely used split-valence basis sets.[100–102] These basis sets include 3-21G, 6-21G, 4-31G, 6-31G, and 6-311G, etc.. Among them, 6-31G uses 6 primitives GTOs to describe one single contracted Gaussian function of the core orbitals, and a combination of two functions for valence orbitals (3 primitives GTOs for inner valence orbitals and 1 primitive GTOs for outer valence orbitals). The polarization functions are represented by "*", and diffuse functions are represented by "+". For instance, 6-311++G(d, p) (same as 6-311++G**) indicates that *d* polarization functions are added to *p*-orbitals, and *s* and *p* diffuse functions are added for heavy elements. Meanwhile, for H or He, *p* polarization functions are added to *s*-orbitals and *s* diffuse functions are added.

Correlation consistent basis sets, developed by Dunning and coworkers,[103–108] aim to systematically recover the correlation energy by increasing the size of the basis set. In this design, the basis sets with similar contributions to correlation energy independent of the function type are included in the same shell. For example, 2*d* and 1*f* functions are added together to a core set of atomic HF functions. Commonly employed correlation consistent basis sets are designated as cc-pV*n*Z, where p for polarization functions, V for Valence, *n* stands for the number of shells the valence functions are split into (*n*= D (3*s*2*p*1*d*), T (4*s*3*p*2*d*1*f*), Q (5*s*4*p*3*d*2*f*1*g*), 5 (6*s*5*p*4*d*3*f*2*g*1*h*), …), and Z for zeta level. The addition of diffusion function with a smaller exponent to every angular momentum is indicted by a prefix "aug". For instance, aug-cc-pVDZ has diffuse *s*, *p*, *d* functions for the C atom.

The properties of transition metal-containing systems are computationally expensive to calculate because a large number of basis functions are needed to describe the electrons. The cost can be reduced by using effective core potentials (ECPs) that describe the potential of the nuclei and core electrons as an average effect.[109–111] In this way, not only is the computational cost reduced, but some relativistic effects can also be included without carrying out the relativistic calculations because these basis functions are generated from relativistic atomic calculations.[112] Both Pople and correlation consistent basis sets have been developed with the pseudopotential version, such as CEP-6-31g(*d*)[113–115] and aug-cc-pV*n*Z(PP)[109–111,116,117]. Other popular pseudopotentials in modern use include the Los Alamos National Laboratory (LANL) ECPs which designated as LANL1DZ (11 electron ECPs)[118] and LANL2DZ (19 electrons ECPs)[119], and the Stuttgart-Dresden ECPs which are designated as SDD.[120,121]

2.8 Reference

(1) Schrödinger, E. *Phys. Rev* **1926**, *28*, 1049.

(2) Schrödinger, E. *Ann. Phys.* **1926**, *79*, 361.

(3) Schrödinger, E. *Ann. Phys.* **1926**, *79*, 489.

(4) Schrödinger, E. *Ann. Phys.* **1926**, *80*, 437.

(5) Schrödinger, E. *Ann. Phys.* **1926**, *79*, 734.

(6) Schrödinger, E. *Ann. Phys.* **1926**, *81*, 109.

(7) Schrödinger, E. *Ann. Phys.* **1926**, *14*, 664.

(8) Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *84*, 457.

(9) Pauli, W. *phys* **1925**, *31*, 373–385.

(10) Pauli, W. *Exclusion Princ. quantum Mech. Nobel Lect. Deliv. Stock. December 13,* **1946**.

(11) Slater, J. C. *Phys. Rev* **1929**, *34*, 1923.

(12) Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 89.

(13) Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 111.

(14) Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 426.

(15) Fock, V. Z. *phys* **1930**, *61*, 126.

(16) Fock, V. Z. *phys* **1930**, *62*, 795.

(17) Singnoglu, O.; Tuan, D. F. *J. Chem. Phys.* **1963**, *38*, 1740–1748.

(18) Bartlett, R. J.; Stanton, J. F. *Rev. Comput. Chem* **1994**, *5*, 65–169.

(19) Bartlett, R. J.; Stanton, J. F. *Rev. Mod. Phys.* **2007**, *79*, 291.

(20) Helgaker, T.; Jørgensen, P.; Olsen, J. *Mol. Electron. Theory; Wiley Chinester, England,* **2000**.

(21) Szabo, A.; Ostlund, N. S. *Mod. Quantum Chem. (Nobel Prize Lect. Geneva, Switz.* **1945**.

(22) Moller, C.; Plesset, M. S. *Phys. Rev* **1934**, *46*, 618–622.

(23) Purvis, C. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(24) Raghavachari, k.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(25) Bartlett, R. J.; Watts, J. D.; Kucharski, S. A.; Noga, J. *Chem. Phys. Lett.* **1990**, *165*, 513–522.

(26) Paldus, J.; Wilson, S.; Diercksen, H. F. *“Coupled Clust. Theory” Methods Comput. Mol. Physics, Plenum Press New York, NY,* **1992**.

(27) Paldus, J.; Li, X.; Prigogine, I.; Rice, S. *"A Crit. Assess. Coupled Clust. Method Quantum Chem. Adv. Chem. Physics, John Wiley Sons, Inc.,* **1999**, *110*, 1.

(28) Barlett, R. J.; Sekino, H.; Purvis |||, G. D. *Chem. Phys. Lett.* **1983**, *98*, 66.

(29) Gustavo, E. S.; Timothy, J. L. *J. Chem. Phys.* **1990**, *93*, 5851.

(30) Dunning, Jr., T. H. *J.Phys. Chem. A* **2000**, *104*, 9062–9080.

(31) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18.

(32) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; McGuire, M. J. *Int. Rev. Phys. Chem* **2002**, *21*, 527.

(33) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; Fan, P. D.; Lodriguito, M.; McGuire, M. J.; Kucharski, S. A.; Kus, T. *Theor. Chem. Acc.* **2004**, *112*, 349.

(34) Piecuch, P.; Kowalski, K. *Comput. Chem. Rev. Curr. Trends, Ed. by J.Leszczyński World Sci. Singapore* **2000**, *Vol. 5*, P. 1.

(35) Piecuch, P.; Kowalski, K. *J. Chem. Phys.* **2000**, *113*, 18.

(36) Ge, Y. B.; Gordon, M. S.; Piecuch, P. *J. Chem. Phys.* **2007**, *127*, 174106.

(37) Lee, T. J.; Taylor, P. R. *Int. J. Quant. Chem. Symp.* **1989**, *S23*, 199.

(38) Lee, T. J.; Rice, J. E.; Scuseria, G. E.; Schaefer, H. F. *Theor. Chim. Acta* **1989**, *75*, 81.

(39) Leininger, M. L.; Nielsen, I. M. B.; Crawford, T. D. . J.; L., C. *Chem. Phys. Lett.* **2000**, *238*, 431.

(40) Sears, J. S.; Sherrill, C. D. *J. Phys. Chem. A* **2008**, *112*, 6741.

(41) Sears, J. S.; Sherrill, C. D. *J. Phys. Chem. A* **2008**, *112*, 3466.

(42) Karton, A.; Daon, S.; Martin, J. M. L. *Chem. Phys. Lett.* **2011**, *510*, 165.

(43) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(44) Lee, T. J.; Taylor, P. R. *Int. J. Quant. Chem. Symp.* **1989**, *S23*, 199–207.

(45) Roos, B. O. *Complet. Act. Sp. Self-Consistent F. Mehod its Appl. Electron. Struct. Calc. Adv. Chem. Physics; Lawley, K. P., Ed.*, **2007**, pp 399.

(46) Ruedenberg, K.; Cheung L. M.; Elbeig, S. T. *Int. J. Quantum Chem.* **1979**, *16*, 1069.

(47) Jensen, F. *Introd. to Comput. Chem. John Wiley Sons Ltd.; Chichester, West Sussex,* **1999**.

(48) Lee, T. J. *Chem. Phys. Lett.* **2003**, *372*, 362.

(49) Janssen, C. L.; Nielsen, I. M. B. *Chem. Phys. Lett.* **1998**, *290*, 423.

(50) Hohenberg, P.; Kohn, W. *Phys. Rev* **1964**, *136*, B864.

(51) Kohn, W.; Sham, L. *Phys. Rev* **1965**, *140*, A1133.

(52) Parr, R. G.; Yang, W. *Density-functional theory atoms Mol. Oxford Univ. Press. New York.* **1989**.

(53) Szabo, A.; Ostlund, N. S. *Mod. Quantum Chem. Introd. to Adv. Electron. Struct. Theory; Mc Graw-Hill, Inc. New York, NY,* **1989**.

(54) Geerlings, P.; DeProft, F.; Langenaeker, W. *Density Funct. Theory A Bridg. Between Chem. Physics. Eds. VUB Unversity Press Brussels,* **1999**.

(55) Koch, W.; Holthausen, M. C. *A Chem. Guid. to Density Funct. Theory. 2 ed.; Wiley-VCH Verlag GmbH Weinheim, Ger.* **2002**.

(56) Fiolhais, C.; Nogueira, F.; Marques, M. *A Prim. Density Funct. Theory. Eds.; Springer-Verlag Berlin,* **2003**.

(57) Polizer, P.; Abu-Awwad, F. *Theor. Chem. Acta.* **1998**, *99*, 83.

(58) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

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

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

(61) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

(62) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(63) Perdew, J. R.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

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

(65) Tao, J.; Perdew, J.; Staroverov, V.; Scuseria, G. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(66) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem.* **2006**, *110*, 5121–5129.

(67) Becke, A. D. **1996**, *104*, 1040–1046.

(68) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(69) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098–3100.

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

(71) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439.

(72) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

(73) Zhao, Y.; Lynch, B.; Truhlar, D. *J.Phys. Chem. A* **2004**, *108*, 4786.

(74) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(75) Schwabe, T.; Grimme, S. *Acc. Chem. Res.* **2008**, *41*, 569.

(76) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398.

(77) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397.

(78) Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787.

(79) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 0341081.

(80) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(81) Menschutkin, N. A. *Z. Phys. Chem.* **1980**, *5*, 589.

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

(83) Tomasi, J.; Menucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.

(84) Born, M. Z. *Physik* **1920**, *1*, 45.

(85) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351.

(86) Kirkwood, J. G.; Westheimer, F. H. *J. Chem. Phys.* **1938**, *6*, 506.

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

(88) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170.

(89) Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62.

(90) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959.

(91) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.

(92) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Theochem-J. Mol. Struct.* **1999**, *462*, 1.

(93) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419.

(94) Levine, I. N. *Quantum Chem. ed.; Prentice Hall,* **2000**.

(95) Boys, S. F. *Proc. Roy. Soc.* **1950**, *A200*, 542.

(96) Taketa, H.; Huzinaga, S.; O-Ohata, K. *J. Phys. Soc. Jpn.* **1966**, *21*, 2313.

(97) Boys, S. F. *Proc. Roy. Soc.* **1950**, *A201*, 125.

(98) Cook, D. B. *Handb. Comput. Quantum Chem. Dover Publ. Inc. Mineola, NY,* **2005**.

(99) Helgaker, T.; Jørgensen, P.; Olsen, J. *Mol. Electron. Theory; John Wiley Sons Ltd. Chichester, West Sussex,* **2000**.

(100) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(101) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(102) Curtiss, L. A.; Raghavachari, K.; Redfern C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(103) Dunning, Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(104) Woon, D. E.; Dunning, Jr., T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(105) Woon, D. E.; Dunning, Jr., T. H. *J. Chem. Phys.* **1994**, *100*, 2975–2988.

(106) Dunning, Jr., T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244–9253.

(107) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, Jr., T. H. *J. Chem. Phys.* **1996**, *388*, 339–349.

(108) Peterson, K. A.; Dunning, Jr., T. H. *J. Chem. Phys.* **2002**, *117*, 10548.

(109) Peterson, K. A. *J. Chem. Phys.* **2003**, *119*, 11099.

(110) Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283.

(111) Peterson, K.A.; Figgen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.

(112) Schwerdtfeger, P. *ChemPhysChem* **2011**, *12*, 3143–3155.

(113) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(114) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026.

(115) Cundari, T. R.; Stevens, W. J. *J. Chem. Phys.* **1993**, *98*, 5555.

(116) Figgen, D.; Peterson, K. A.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2009**, *130*, 164108–164112.

(117) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866.

(118) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270.

(119) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299.

(120) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866.

(121) Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.

CHAPTER 3 DEGRADATION OF β-O-4 LINKAGE OF LIGNIN BY TRANSTION METAL ATOMS

3.1 Introduction

Due to the rapid exhaustion of easily recovered fossil fuels, there has been significant interest in developing sustainable and cleaner alternatives to petroleum-based resources.[1,2] A promising alternative resource is biomass as it has the potential to be processed into fuels, energy, and commodity chemicals.[3] Biomass is more advantageous than traditional fossil fuels because of its abundance, sustainability, and lower yield of unwanted byproducts and damaging gases when degraded.[4] Lignin, one component of lignocellulosic biomass, is a highly branched polymer composed of three primary phenylpropane monomers (p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol) linked together by carbon-oxygen and carbon-carbon bonds.[5,6] The β-O-4 linkage [1-(4-hydroxy-3-methoxyphenyl)-2-(2-methoxyphenoxy)-1,3-propanediol], shown in Figure 1, is the most common linkage in lignin and comprises between 50-60% of all lignin linkages;[7] therefore, it is commonly used as a model compound in lignin studies.

Despite the abundance of lignin, commercial development of lignin lags in comparison to the commercial development of other components of lignocellulosic biomass due to the complex structure of lignin.[3,8] Consequently, several methods have been used to investigate the decomposition of lignin model compounds. Vicuna and coworkers showed that the enzyme in the bacterium *Pseudomonas acidovorans* cleaved the C_β-O bond in a lignin degradation process.[9] However, ligninolytic enzymes are expensive and difficult to remove from reaction systems.[10] Pyrolytic degradation works by heating lignin to high temperatures, where the cleavage of the C_β-O bond initializes the pyrolysis process.[11,12] Another method, supercritical water gasification, converts black liquor, the waste product from the conversion of wood into wood pulp that is mainly composed of lignin, into CO, CH_4, CO_2, C_2, C_4, and H_2 gases.[13] A fragmentation-hydrogenolysis process discovered by Song and coworkers uses nickel-based catalysts and alcohol solvents to break lignin into monomeric phenols[14] while the dehydroxygenation of lignin produces phenols, cyclohexane, benzene, naphthalene, and phenanthrene.[15] However, pyrolysis, gasification, fragmentation-hydrogenolysis, and dehydroxygenation all have low selectivity for lignin bond cleavage.[11–17]

Due to these challenges, it is of interest to study other methods of lignin degradation. Currently, transition metal (TM)-catalyzed depolymerization is a practical decomposition method and is widely used in industries. Metal catalysts do not require extreme temperatures, are highly selective, and reduce the energetic barriers of the reactions, thus allowing for easily controlled reactions.[10] Heterogeneous metal catalysts such as TM surfaces are commonly used in industry because they have high recyclability and are more easily separated from the products than homogeneous catalysts.[18] Although realistic TM catalysis by organometallic complexes (homogenous catalyst) or TM surfaces (heterogeneous catalysts) is relatively complex compared to single TM atoms due to stereoelectronic effects, the TM contained in the catalyst plays an important role in affecting the mechanism and thermodynamic characteristics of the reaction. For instance, in the study of CO_2 activation by Cong et al., the late 3*d* TM (Fe, Co, Ni, and Cu) homogeneous catalysts (first row TM β-diketiminato complexes) and the metal surface catalysts had quite similar mechanisms. Additionally, comparing the reaction energetics, late TMs gave more endergonic reactions than early TMs for both heterogeneous (fcc(100) surfaces) and homogeneous catalysts. The observed similar trends for reaction mechanisms and thermodynamics of CO_2 activation reactions for homogenous and heterogeneous catalysts indicate that the metal dictates the catalytic pathway.[19] Therefore, computational studies that investigate gas phase models can assist experimental investigations by elucidating the catalytic properties of different transition metals and the reaction mechanisms. In our prior work, we considered the activation of the C_α-C_β bond in the β–O–4 linkage of lignin by neutral TM atoms (Ni, Cu, Pd, and Pt)[20] and found that of these metals, Pt is the most thermodynamically favorable for the bond activation reaction. Here, we focus upon the cleavage of the C-O aryl ether bonds of lignin as another decomposition route for lignin.

Although high-level wavefunction theory (WFT) can be more accurate and, in general, more systematically improvable than density functional theory (DFT), the cost of WFT increases more rapidly than DFT as the system size increases. Because of this and the fact that DFT can provide results comparable to experimental data for calculating many properties of molecules, including structural and some relative energetic properties in reaction pathways,[21] DFT is more widely used in computational studies on TM-containing systems.[22] However, the best density functional for the study of TM species is not clearly defined as demonstrated and discussed by prior work that has shown that density functional performance can vary widely for different transition metals, ligands, and properties, and that the performance is not necessarily systematic.[22–29] Functional performance is not known *a priori* and may depend on the type of reaction systems. Thus, it is critical to understand functional performance for different types of reaction systems, and then utilize that knowledge for studies of similar but larger reaction systems. For this study, the C-O bond cleavages of methoxyethane and methanol were investigated first as model systems in order to analyze the performance of different DFT functionals for the C-O activation reactions. Then the functional that performed best to describe the C-O bond activations of methoxyethane and methanol will be applied to study the C_β-O bond cleavage of β-O-4 linkage of lignin.

The reaction pathway for methoxyethane and methanol activations start with binding the metal atom (Fe, Co, Ni, Ru, Rh, and Pd in this study) to the oxygen atom of methoxyethane and methanol to form the adducts, [MO(CH_3CH_2)(CH_3)] and [MO(CH_3)H], respectively. Next, the adducts undergo an oxidative insertion step, forming transition states, $[CH_3CH_2MOCH_3]^{\ddagger}$ and $[CH_3MOH]^{\ddagger}$. Lastly, the C-O bonds break, and the products, [(CH_3CH_2)M(OCH_3)] and [CH_3MOH], are formed. Since this work is focused on the central C-O bond cleavage, hydrogen migration is not investigated, and the activated intermediate is considered as the product in this paper. Three different types of enthalpy changes, binding enthalpies (enthalpy differences between reactants and adducts), activation enthalpies (enthalpy differences between adducts and transition states), and reaction enthalpies (enthalpy differences between reactants and products), happen throughout the reaction. The performance of the considered DFT functionals, which include the pure-GGA functionals BLYP[30,31] and B97D,[32] the meta-GGA functionals TPSS[33] and M06L,[24] the hybrid-GGA functionals B3LYP[30,31,34] and PBE0,[35–37] and the hybrid-meta-GGA functionals M06[24], and TPSSh[33,38], is examined in terms of different types of enthalpy changes throughout the reaction pathway. The completely renormalized coupled cluster method with singles, doubles, and perturbative triples [CR-CCSD(T)] is a useful gauge for the performance of the DFT functionals. CR-CCSD(T) has been shown to be effective in describing single-bond breaking reactions.[39–42] In a study of H-CH_3, H-Cl, Cl-Cl, H-SiH_3, H_3C-CH_3, H_3Si-SiH_3, H_3C-SiH_3, Cl-CH_3, and Cl-SiH_3, CR-CCSD(T) recovered a bond dissociation energy within 1 kcal mol^{-1} of Full CI for each molecule.[43] For CH_3-F, CR-CCSD(T) resulted in the same dissociation energy as was predicted by MRCI(Q).[44] The DFT functional that was in best agreement with CR-CCSD(T) for the degradation of methoxyethane was then applied to study the activation of the β-

O-4 linkage by the considered TM atoms. The results from this study will aid in metal catalyst design for lignin decomposition and provide a greater understanding of suitable density functionals for studies of similar molecular species.

3.2 Computational Details and Methodology

DFT calculations were performed using the Gaussian 09 program,[45] while CR-CCSD(T)[39,42] calculations were performed using the NWChem[46] program. Thermal energy corrections were conducted at 298.15 K and 1 atm in the gas phase, and the calculated energies are reported in kcal mol^{-1}.

The geometries of the ground and transition states of the molecules were optimized using the B3LYP density functional because it has been shown to provide reliable structure for the ground-state geometries of many TM compounds.[22] Single-point energy calculations were then performed on the optimized structures using CR-CCSD(T) and nine density functionals, including BLYP, B97D, TPSS, M06L, B3LYP, PBE0, M06, and TPSSh (Table 3.1). Thermal corrections obtained from B3LYP were then added to the single point energies to estimate the enthalpies of the molecules. The transition states are indicated by exactly one imaginary frequency.

The correlation consistent polarized triple-ζ quality (cc-pVTZ)[47,48] basis sets were used for main group and 3*d* TM atoms, while cc-pVTZ basis sets with relativistic pseudopotentials (cc-pVTZ-PP)[49-52] were used for 4*d* and 5*d* TM atoms.

3.3 Results and Discussion

The energies determined using each DFT functional are compared with the energies calculated using CR-CCSD(T)/cc-pVTZ, and mean absolute deviations (MADs) are provided to consider which functional best models the reaction pathway.

3.3.1 Performance of DFT Functionals for Model Systems

The MADs of each of the considered DFT functionals for the binding enthalpy, activation enthalpy, and reaction enthalpy with respect to CR-CCSD(T)/cc-pVTZ energies were determined for methoxyethane and methanol species.

For methoxyethane species containing 3*d* TMs (Figure 3.2), all considered density functionals performed well in reproducing the accuracy of CR-CCSD(T)/cc-pVTZ for the binding enthalpies with most of the MAD values less than 5 kcal/mol. Notably, PBE0, B3LYP, and TPSSh had the lowest MAD values of 1.3 kcal/mol, 1.9 kcal/mol, and 1.6 kcal/mol, respectively, for binding enthalpies. All of the considered density functionals performed poorly in predicting the activation enthalpies accurately with MAD more than 5 kcal/mol, except for B3LYP (2.2 kcal mol^{-1}), PBE0 (3.7 kcal mol^{-1}), and M06 (4.9 kcal mol^{-1}) that resulted in the lowest MAD values for activation enthalpies. Although BLYP, B97D, and TPSS had no significant error in predicting the binding enthalpies with an acceptable MAD value of 3.9 kcal/mol, 2.8 kcal/mol, and 3.3 kcal/mol, respectively, they failed to predict activation enthalpies and reaction enthalpies accurately with large MAD values. PBE0 and B3LYP, on the other hand, performed the best in producing the activation enthalpies and reaction enthalpies close to the CR-CCSD(T)/cc-pVTZ. Overall, the considered density functionals resulted in lower MAD values and performed better in describing binding enthalpies than activation and reaction enthalpies. When considering all three enthalpy terms for the 3*d* species, PBE0 and B3LYP, which are hybrid-GGA functionals, performed the best with MAD values of 2.7 kcal/mol and 3.1 kcal/mol respectively. For methoxyethane species containing 4*d* TMs (Figure 3.3), the MAD values of the binding enthalpies were not always the lowest of all three enthalpy terms, such as B97D, TPSSh, and M06 resulted in lowest MAD values for activation enthalpies. All the functionals performed worst in predicting the reaction enthalpies. None of the functionals had acceptable MAD values in predicting reaction enthalpies. B3LYP performed the best in predicting the most accurate values of activation and reaction enthalpies with MAD values of 3.1 kcal/mol and 9.1 kcal/mol, respectively, and also produced the second lowest MAD value of 1.7 kcal/mol for binding enthalpies. For binding enthalpies, PBE0 performed the best with MAD value of 1.7 kcal/mol. In considering the performance of different types of density functionals, the density functionals with HF exchange (except TPSSh) for all three reaction terms were in the best agreement with those of the CR-CCSD(T)/cc-pVTZ calculations, with overall MAD values of 4.6 kcal/mol from B3LYP, 5.7 kcal/mol from PBE0, and 7.1 kcal/mol from M06. When considering all three enthalpy terms for the 4*d* species, PBE0 and B3LYP still performed the best with MAD values of 5.7 kcal/mol and 4.6 kcal/mol, respectively. Compared to the performance of different types of density functionals for the 3*d* species, all considered density functionals resulted in larger MAD values and performed worse in predicting the thermodynamic properties of the 4*d* species. PBE0 produced the lowest MADs for binding enthalpies of both 3*d* and 4*d* species, while B3LYP resulted in the lowest MADs for activation enthalpies of both 3*d* and 4*d* species.

For methanol species (Figure 3.4-3.5), all of the considered density functionals performed best in describing the binding enthalpies (followed by activation enthalpies), except M06 (produced lower MAD value for activation enthalpies than binding enthalpies for 3*d* species) and

PBE0 (produced lower MAD value for reaction enthalpies than activation enthalpies for 4*d* species).The density functionals without HF exchange and TPSSh performed worse in prediction activation and reaction enthalpies; most notably B97D, which had an extremely large MAD value of 30.2 kcal/mol for reaction enthalpies of 4*d* species. B3LYP gave the most accurate values of reaction enthalpy for 3*d* species, and binding and activation enthalpies for 4*d* species, while PBE0 had the lowest MAD value of 3.6 kcal/mol for reaction enthalpy for 4*d* species. In considering overall performance of the considered density functionals for all three types of enthalpies examined in this study, M06 performed the best in predicting the thermodynamic properties of 3*d* species with MAD value of 5.6 kcal/mol, followed by PBE0 (6.9 kcal/mol) and then B3LYP (7.2 kcal/mol), while PBE0 resulted in the lowest MAD value of 4.3 kcal/mol for 4*d* species, followed by B3LYP (7.2 kcal/mol) and then M06 (7.3 kcal/mol).

PBE0 had the lowest MAD values for both 3*d* and 4*d* species (Figure 3.6) when considering the enthalpies changes of C-O bond activation reactions of methoxyethane and methanol together, which means that PBE0 performed the best among all considered density functionals. Therefore, PBE0 will be applied in the study of C_β-O bond activation of β-O-4 linkage of lignin. B3LYP and M06 also produced similar MAD values as PBE0 for 3*d* species. Interestingly, these three functionals all have percentages of HF exchange between 20% and 30%, suggesting that the amount of HF exchange significantly affects the accuracy of the density functionals in predicting the enthalpies for the considered reaction. In two prior studies on the performance of DFT functionals in predicting enthalpies of formation for 3*d* and 4*d* transition metal-ligand bonded species, the hybrid-GGAs with HF exchange percentages between 20% to 27% provided lower MADs than the hybrid-GGAs with higher HF exchange.[25,26] The optimal percentage of HF exchange was averaged to be 26% for the 3*d* species and 34% for the 4*d* species for predicting the reaction energetics in a study of C-O bond activations of dimethyl ether by late 3*d* and 4*d* TM ions.[19] Conversely, the local GGA functionals (BLYP, B97D, and TPSS) without any HF exchange resulted in the poorest performance in most cases. B97D consistently resulted in the largest MAD values for both the reaction type and the metal considered.

3.3.2 C-O Bond Activation of the β-O-4 Linkage by Late 3*d*, 4*d*, and 5*d* TMs Using PBE0

As shown in Figure 3.7, the activation of the C_β-O bond of the entire β-O-4 linkage catalyzed by a TM atom (Fe, Co, Ni, Ru, Rh, Pd, Os, Ir, and Pt) was considered using PBE0 that was selected due to its most similar results with the results from CR-CCSD(T) in comparison with other considered density functionals for study of the oxidative insertion of TM atoms into methoxyethane and methanol, as discussed earlier. The binding enthalpies (Figure 3.8) are the enthalpy differences between the reactants and the adducts shown in Figure 3.7. The TM atoms first approached the oxygen atom of β-O-4 linkage and formed adducts. The orbital overlap between the metal atom center and the O oxygen atom accompanies with the electron donation. Therefore, the metal atom centers need to have vacancies on either the n*d* or (n+1)*s* orbitals in order to accept incoming electron(s) from the oxygen atoms. The electron configurations of the TM atoms ground state are: Fe: [Ar] $3d^7$ $4s^1$, Co: [Ar] $3d^8$ $4s^1$, Ni: [Ar] $3d^9$ $4s^1$, Ru: [Kr] $4d^7$ $5s^1$, Rh: [Kr] $4d^8$ $5s^1$, Pd: [Kr] $4d^{10}$ $5s^0$, Os: [Xe] $4f^{14}$ $5d^6$ $6s^2$, Ir: [Xe] $4f^{14}$ $5d^8$ $6s^1$, Pt: [Xe] $4f^{14}$ $5d^9$ $6s^1$. Thus all the TM atoms can accept extra electron(s) and have an exothermic binding enthalpies varying from -2.9 to -13.1 kcal/mol, as shown in Figure 3.8. The magnitudes of binding enthalpies of 3*d* species are in the decreasing order of Co, Fe, and Ni, that of 4*d* species are in the decreasing order of Rh, Pd, and Ru, and that of 5*d* species are in the decreasing order of Pt, Ir and Os. Most of the 3*d* and 4*d* species had lower binding enthalpies than the 5d species. The TM atoms in the Co Group resulted in the strongest bonding with β-O-4 linkage except for Ir (-5.4 kcal/mol) that had a lower binding enthalpy than Pt (-11.1 kcal/mol). The activation enthalpies are the enthalpy differences between the adducts and the transition states shown in Figure 3.8. Positive activation enthalpies were produced by all of the considered TM atoms. The activation enthalpies (Figure 3.8) of 3*d* species followed the trend Fe < Co < Ni, 4*d* species followed the trend Rh < Ru < Pd, and 5*d* species followed the trend Pt < Os < Ir. These trends show that earlier TM species tend to have lower activation enthalpies than later TM species for both 3*d* and 4*d* species. Most of the 3*d* TM species had higher activation enthalpies than the 4*d* and 5*d* species. The reaction enthalpies (Figure 3.8) are the enthalpy differences between the reactants and the products. All considered TM atoms produced exothermic reactions. The reactions catalyzed by earlier TM atoms had more exothermic reactions than those catalyzed by later TM atoms for 3*d* and 4*d* species. Most of 3*d* and 5*d* species formed a more stable products than 4*d* species. Fe, Ru, and Ir showed the most exothermicity among 3*d*, 4*d*, and 5*d* species, respectively. Os, belonging to the same group as Fe and Ru, formed a less stable products than Ir and Pt. This trend can be explained by the electronic configurations of the TM atoms that have been shown to have large effects on the determination of enthalpy formations.[20] Recall that the ground state electron configurations of the Os atoms is Os: [Xe] $4f^{14}$ $5d^{6}$ $6s^{2}$. The doubly occupied *s* orbital that is more diffuse than the *d* orbitals on Os may cause greater repulsion to decrease the orbital overlap between Os and C-O bond; therefore, the product formed by Os is less stable than the products formed by Pt and Ir. Overall, Pd would be the most suitable TM catalysts as they catalyzed the flatter reaction pathways in comparison with other TM atoms, which means that low energies are required to overcome the reaction barriers by Pd, and are also easily removed from the reaction system due to the formation of less stable product.

This study is also interested in finding out the selectivity of C_α-C_β and C_β-O bond cleavage of β–O–4 linkage by TM atoms because cleavages of these two bonds are the most important steps in the decomposition of lignin and compete with each other. The thermodynamic properties (calculated with rp-ccCA-ONIOM(3) method) for C_α-C_β bond activation were obtained from the previous study by Oyedepo and coworkers[20] that was focused on the C_α-C_β bond cleavage of the β–O–4 linkage of lignin using Ni, Cu, Pd, and Pt neutral atoms in gas phase. The comparison will focus on Ni, Pd, and Pt that were applied both in Oyedepo's study and our study. As shown in Table 3.2, the activation enthalpies produced by the considered TM atoms for C_α-C_β bond activation are in the order of Ni < Pt < Pd, which is in the opposite order for C_β-O bond activation. The magnitudes of reaction enthalpies of TM species are in the order of Pd < Ni < Pt for both C_β-O and C_α-C_β bond activations. Pd and Pt atoms results in higher reaction barriers for the C_α-C_β bond activation than C_β-O bond activation, which makes sense since the dissociation energy of C_β-O bond is lower than that of C_α-C_β bond. Thus, the C_β-O bond cleavage can be expected to be more favorable than the C_α-C_β bond cleavage in decomposition of the β–O–4 linkage of lignin with the considered Pd and Pt atoms. The degradation of β–O–4 linkage of lignin is therefore predicted to start with the C_α-C_β bond cleavage. The C_β-O bond activation reactions are more exothermic than C_α-C_β bond activation reactions.[20] Therefore, the products formed from C_α-C_β bond activation are expected to be easier to be separated from the catalytic systems than C_α-C_β bond activation.

Since methoxyethane includes the main section of the linkage involved in the C_β-O bond activation of β–O-4 linkage. Investigation of the thermodynamic properties change from methoxyethane species to β–O-4 linkage species will be useful to examine if methoxyethane is an effective representative model to study the activation reactions of β–O-4 linkage. Figure 3.8 shows the reaction energetics of the methoxyethane species calculated by PBE0 and by CR-CCSD(T) and the reaction energetics of the β-O-4 linkage species calculated by PBE0. For the binding enthalpies produced by the 3*d* and 5*d* TMs, methoxyethane species with PBE0, methoxyethane species with CR-CCSD(T), and β-O-4 linkage species with PBE0 followed the trend of Co < Fe < Ni and Pt < Ir < Os. Among the 4*d* TMs, Rh produced the lowest binding enthalpies in comparison to other TMs for both methoxyethane species with CR-CCSD(T) and β-O-4 linkage species with PBE0. However, for methoxyethane species with PBE0, Pd results in the lowest binding enthalpies. The activation enthalpies produced by 3*d* TMs increased in the order of Fe < Co < Ni for methoxyethane species with PBE0, methoxyethane species with CR-CCSD(T), and β-O-4 linkage species with PBE0. Pd and Ir were predicted to produce the largest activation enthalpies of the 4*d* and 5*d* TMs for methoxyethane species with PBE0, methoxyethane species with CR-CCSD(T), and β-O-4 linkage species with PBE0. For the reaction enthalpies produced by the 3*d* and 5*d* TMs, methoxyethane species with PBE0, methoxyethane species with CR-CCSD(T), and β-O-4 linkage species with PBE0 follow the same general trend of Fe < Co < Ni and Os < Ir < Pt. For the 4*d* TMs, Ru was predicted to produce the lowest reaction enthalpies for the β-O-4 linkage species with PBE0 and the methoxyethane species with CR-CCSD(T), while Rh was predicted to produce the lowest reaction enthalpies for the methoxyethane species with PBE0. In terms of the binding enthalpies and activation enthalpies, activation reactions of methoxyethane and β-O-4 linkage had similar energetics except Ni, Ru, and Ir, which gave ~10 kcal mol^{-1} differences between the β-O-4 linkage and methoxyethane species. The β-O-4 linkage species had much lower reaction enthalpies than the methoxyethane species, which may be due to the resonance effect of the benzene rings. Furthermore, comparison between the C-O bond activation of methoxyethane and the β-O-4 linkage using PBE0 indicated that a similar thermodynamic trend occurs for both reactions: for 3*d* and 4*d* species, from earlier metals to later metals, the reactions tend to have greater activation enthalpies and to be less exothermic. Therefore, methoxyethane can be used as a representative model to investigate the impact of TM catalysts on the degradation of β-O-4 linkage qualitatively but not quantitatively.

3.4 Conclusions

CR-CCSD(T) and DFT calculations were performed to investigate the usefulness of different density functionals in predictions of the energetics of the degradation of methoxyethane by late TMs. The analysis of the performance of the density functionals considered in this study can serve as a guide for the selection of the optimal density functional in future work. An evaluation of a number of density functionals suggested that PBE0, B3LYP, and M06 resulted in the lowest overall MADs, relative to CR-CCSD(T), for the prediction of the reaction energetics (binding, activation, and reaction enthalpies) for C-O bond activation of methoxyethane and methanol by several late 3*d* and 4*d* TM atoms; therefore, they are recommended for similar investigations. In contrast, pure-GGA and meta-GGA functionals performed the poorest overall, particularly the B97D functional. Prompted by these results, the reaction enthalpies of the TM atom-catalyzed degradation of the β-O-4 linkage were examined using the PBE0 density functional. For the C-O bond activation of β-O-4 linkage, the 5*d* TM species resulted in lower activation enthalpies and more stable products than the 3*d* TM species. The Pd in particular was found to provide the flattest reaction pathway. In comparison with the C_α-C_β bond activation reaction, the C_β-O bond activation reactions have lower reaction barriers and are more exothermic. Considering the effectiveness of the transition metals for breaking the C-O bond, the reaction energetics for the activation of β-O-4 linkage species followed similar patterns as observed for the activation of the methoxyethane species. The thermodynamic data obtained in this study can provide insight into the intrinsic catalytic properties of the evaluated transition metals, as well as serve as a guide for the selection of the most advantageous transition metals for the design of catalysts for lignin decomposition.

3.5 References

1. C. Yang, H.; Yan, R.; Chen, H.; Lee, D. H.; Zheng, *Fuel*, 2007, **86**, 1781.

2. M. P. Pandey and C. S. Kim, *Chem. Eng. Technol.*, 2011, **34**, 29–41.

3. J. Zakzeski, P. C. A. Bruijnincx, A. L. Jongerius and B. M. Weckhuysen, *Chem. Rev.*, 2010, **110**, 3552–3599.

4. A. Beste and A. C. Buchanan, *Chem. Phys. Lett.*, 2012, **550**, 19–24.

5. M. Stöcker, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 9200–9211.

6. F. S. Chakar and A. J. Ragauskas, *Ind. Crops Prod.*, 2004, **20**, 131–141.

7. D. R. Dimmel, L. F. Bovee and B. N. Brogdon, *wood chem. Technol.*, 1993, **14**, 1–15.

8. E. Furimsky, *Appl. Catal. A Gen.*, 2000, **199**, 147–190.

9. K. R. Vicuña, B. González, M. D. Mozuch and T. K. Kirk, *Appl. Environ. Microbiol.*, 1987, **53**, 2605–2609.

10. G. Banerjee, J. S. Scott-Craig and J. D. Walton, *Bioenergy Res.*, 2010, **3**, 82–92.

11. D. Ferdous, A. K. Dalai, S. K. Bej and R. W. Thring, *Energy and Fuels*, 2002, **16**, 1405–1412.

12. T. Faravelli, A. Frassoldati, G. Migliavacca and E. Ranzi, *Biomass and Bioenergy*, 2010, **34**, 290–301.

13. V. Sricharoenchaikul, *Bioresour. Technol.*, 2009, **100**, 638–643.

14. Q. Song, F. Wang, J. Cai, Y. Wang, J. Zhang, W. Yu and J. Xu, *Energy Environ. Sci.*, 2013, **6**, 994–1007.

15. G. W. Huber, S. Iborra and A. Corma, *Chem. Rev.*, 2006, **106**, 4044–4098.

16. H. Wang, M. Tucker and Y. Ji, *J. Appl. Chem.*, 2013, **68**, 1–9.

17. J. R. Gasson, D. Forchheim, T. Sutter, U. Hornung, A. Kruse and T. Barth, *Ind. Eng. Chem. Res.*, 2012, **51**, 10595–10606.

18. Y. -C. Lin and G. W. Huber, *Energy Environ. Sci.*, 2009, **2**, 68.

19. C. Liu, T. R. Cundari and A. K. Wilson, *J. Phys. Chem. C*, 2012, **116**, 5681–5688.

20. G. A. Oyedepo and A. K. Wilson, *ChemPhysChem*, 2011, **12**, 3320.

21. M. Orio, D. A. Pantazis and F. Neese, *Photosynth. Res.*, 2009, **102**, 443–453.

22. C. J. Cramer and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10757–10816.

23. Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157.

24. Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 194101.

25. S. M. Tekarli, M. L. Drummond, T. G. Williams, T. R. Cundari and A. K. Wilson, *J. Phys. Chem. A*, 2009, **113**, 8607–14.

26. M. L. Laury and A. K. Wilson, *J. Chem. Theory Comput.*, 2013, **9**, 3939–3946.

27. W. Jiang, M. L. Laury, M. Powell and A. K. Wilson, *J. Chem. Theory Comput.*, 2012, **8**, 4102–4111.

28. T. R. Cundari, H. Arturo Ruiz Leza, T. Grimes, G. Steyl, A. Waters and A. K. Wilson, *Chem. Phys. Lett.*, 2005, **401**, 58–61.

29. K. E. Riley and K. M. Merz, *J. Phys. Chem. A*, 2007, **112**, 6044.

30. A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.

31. R. G. Lee, C.; Yang, W.; Parr, *Phys. Rev. B*, 1998, **37**, 785.

32. S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.

33. J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.

34. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.

35. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

36. C. Adamo, M. Cossi and V. Barone, *J. Mol. Struct.*, 1999, **493**, 195.

37. M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.*, 1999, **110**, 5029.

38. V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2003, **119**,

12129–12137.

39. S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.

40. K. Kowalski and P. Piecuch, *J. Chem. Phys.*, 2000, **113**, 18.

41. P. Piecuch, K. Kowalski, I. S. O. Pimienta and M. J. McGuire, *Int. Rev. Phys. Chem*, 2002, **21**, 527.

42. P. Piecuch, K. Kowalski, I. S. O. Pimienta, P. D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski andT. Kus, *Theor. Chem. Acc.*, 2004, **112**, 349.

43. P. Piecuch and K. Kowalski, *Comput. Chem. Rev. Curr. Trends, Ed. by J. Leszczyński World Sci. Singapore, 2000, Vol. 5, p. 1.*

44. P. Piecuch and K. Kowalski, *J. Chem. Phys.*, 2000, **113**, 18.

45. Y. Ge, M. S. Gordon, and P. Piecuch, *J. Chem. Phys.*, 2007, **127**, 174106.

46. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O¨. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, GAUSSIAN 09 (Revision A.02), Gaussian, Inc., Wallingford CT, 2009.

47. M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477.

48. N.B. Balabanov and K.A. Peterson, *J. Chem. Phys.*, 2005, **123**, 064107.

49. T. H. Dunning, Jr. *J. Chem. Phys.*, 1989, **90**, 1007–1023.

50. K.A. Peterson and C. Puzzarini, *Theor. Chem. Acc.*, 2005, **114**, 283.

51. K. A. Peterson, D. Figuregen, M. Dolg and H. Stoll, *J. Chem. Phys.*, 2007, **126**, 124101.

52. D. Figuregen, K.A. Peterson, M. Dolg, and H. Stoll, *J. Chem. Phys.*, 2009, **130**, 164108.

3.6 Figures and Tables

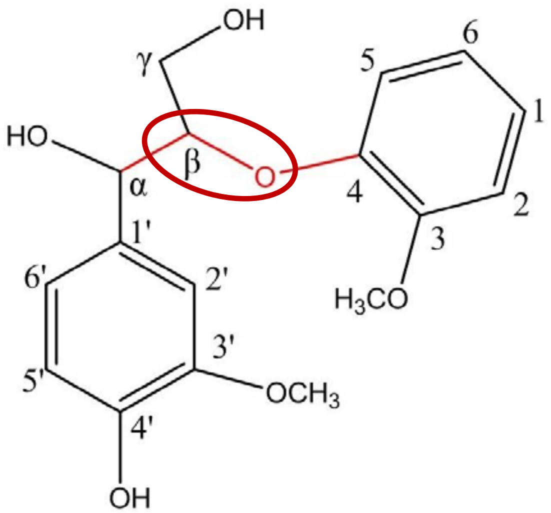


Figure 3.1 The β-O-4 linkage in lignin. The C_β-O bond, the bond of interest in this study, is circled in red.

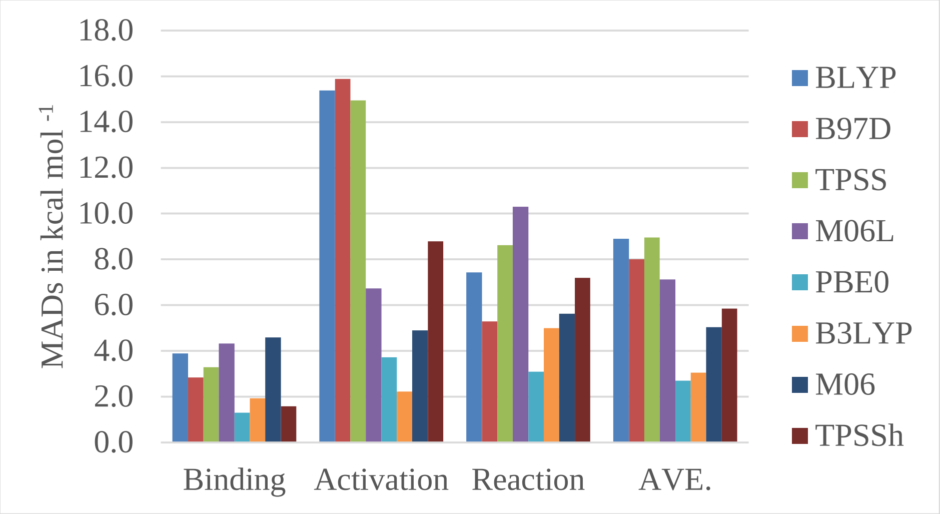


Figure 3.2 MADs of each functional for binding enthalpy, activation enthalpy, and reaction enthalpy of methoxyethane species with 3*d* TMs, relative to CR-CCSD(T).

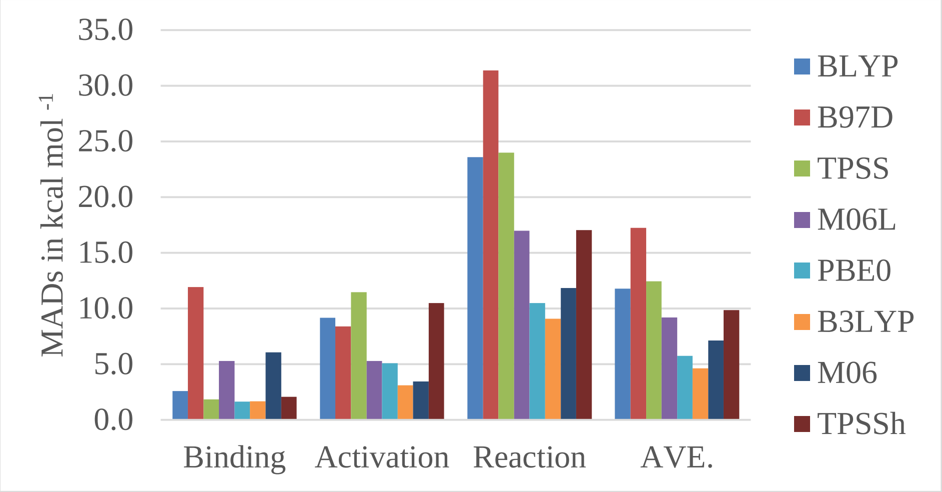


Figure 3.3 MADs of each functional for binding enthalpy, activation enthalpy, and reaction enthalpy of methoxyethane species with 4*d* TMs, relative to CR-CCSD(T).

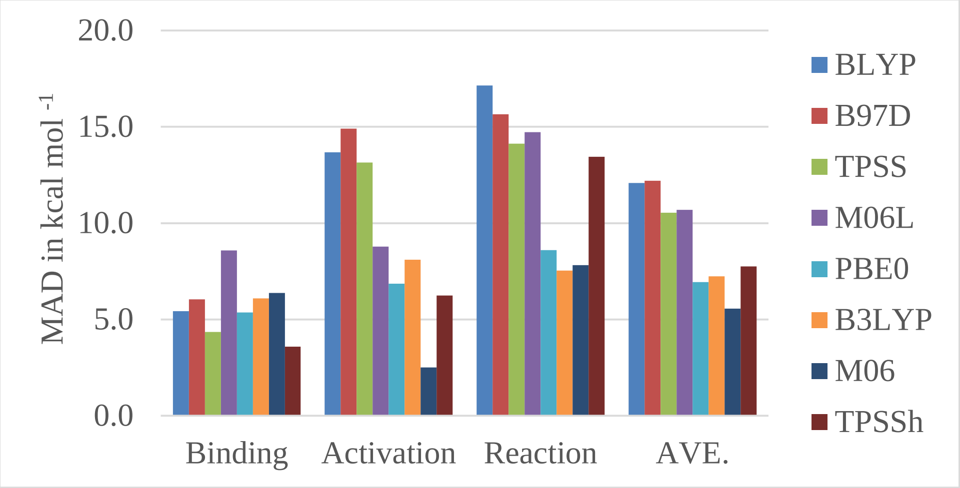


Figure 3.4 MADs of each functional for binding enthalpy, activation enthalpy, and reaction enthalpy of methanol species with 3*d* TMs, relative to CR-CCSD(T).

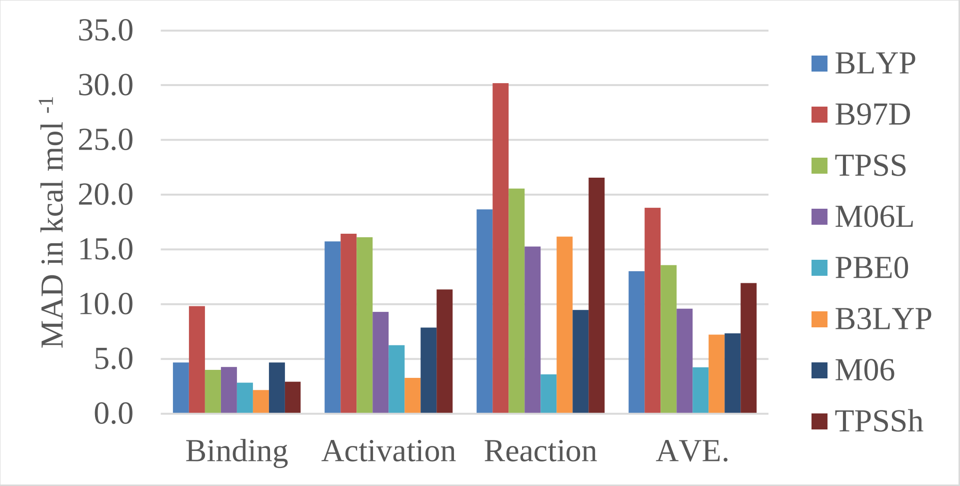


Figure 3.5 MADs of each functional for binding enthalpy, activation enthalpy, and reaction enthalpy of methanol species with 4*d* TMs, relative to CR-CCSD(T).

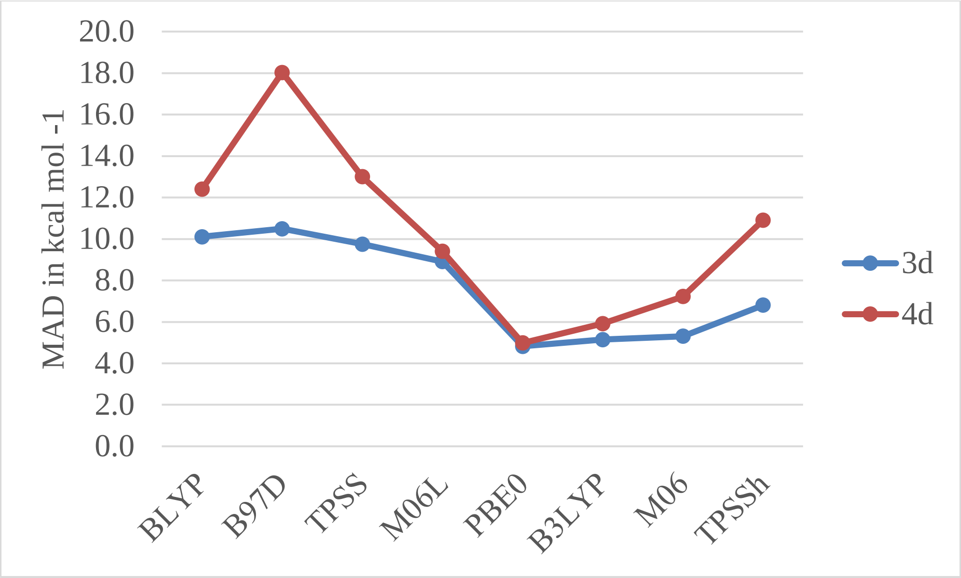


Figure 3.6 MADs of the overall 3*d* and 4*d* metal species, relative to CR-CCSD(T).

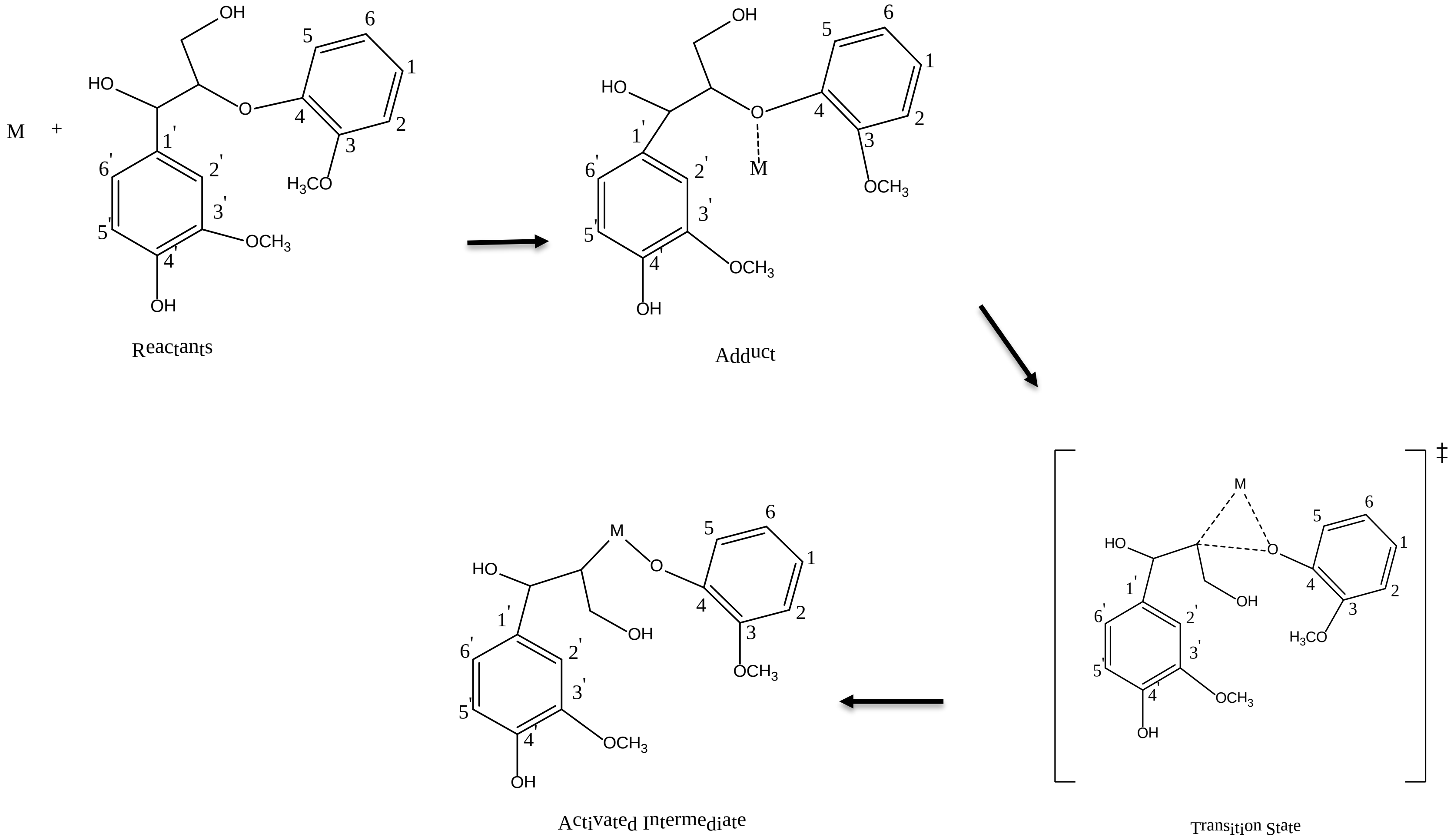


Figure 3.7 Reaction pathway of the oxidative cleavage of the C_β-O bond in the β-O-4 linkage.

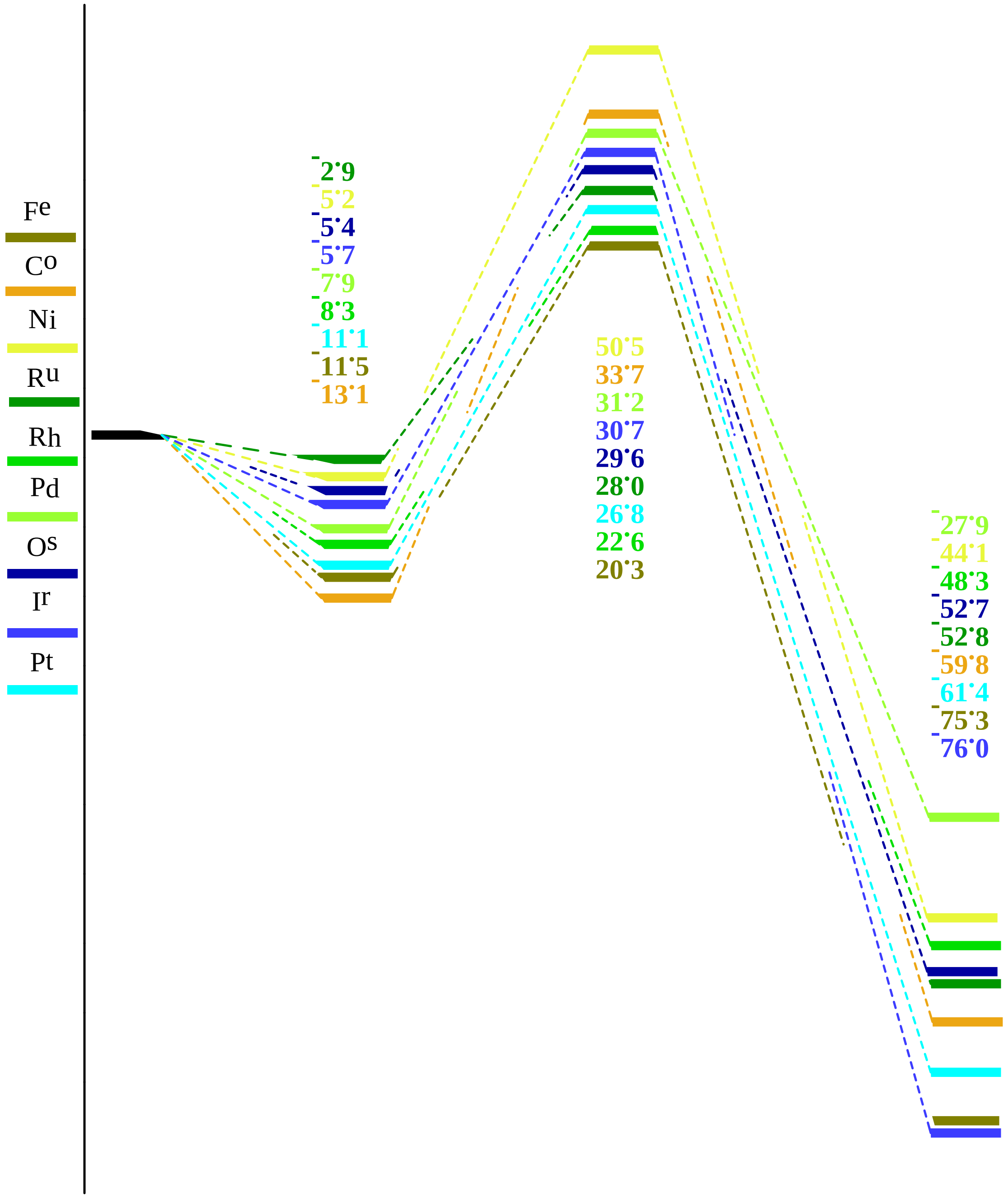


Figure. 3.8 Energetics of the C_β-O bond activation reactions of β-O-4 linkage system catalyzed by each transition metal atom in kcal mol^{-1}, using PBE0.

Table 3.1 Overview of the density functionals applied in this study

Functional	X[a]	Type	Exchange Functional	Correlation Functional
BLYP[33,34]	0	P-GGA[b]	Becke88	Lee-Yang-Parr
B97D[35]	0	P-GGA	B97-D	B97-D
TPSS[36]	0	M-GGA[c]	Tao-Perdew-Staroverov-Scuseria	Tao-Perdew-Staroverov-Scuseria
M06L[27]	0	M-GGA	M06L	M06L
B3LYP[33,34,37]	20	H-GGA[d]	Becke88	Lee-Yang-Parr
PBE0[38-40]	25	H-GGA	Perdew-Burke-Ernzerhof	Perdew-Burke-Ernzerhof
M06[27]	27	HM-GGA[e]	M06	M06
TPSSh[36,41]	10	HM-GGA	Tao-Perdew-Staroverov-Scuseria	Tao-Perdew-Staroverov-Scuseria
B2PLYP[42]	53	DH-GGA[f]	Becke88	Lee-Yang-Parr

a. X = Hartree-Fock exchange percentage

b. P-GGA = Pure-GGA

c. M-GGA = Meta-GGA

d. H-GGA = Hybrid-GGA

e. HM-GGA = Hybrid-Meta-GGA

f. DH-GGA = Double-Hybrid-GGA

Table 3.2 Comparison of activation enthalpies and reaction enthalpies between C_α-C_β and C_β-O bond activations of β-O-4 linkage of lignin

	Activation Enthalpy[g]		Reaction Enthalpy	
	C_α-C_β[h]	C_β-O	C_α-C_β	C_β-O
Ni	17.3	45.3	-17.4	-44.1
Pd	31.2	23.3	-3.7	-27.9
Pt	24.3	15.7	-47.8	-61.4

g. The energy difference between reactants and transition states.

h. The C_α-C_β bond activation reactions were calculated with rp-ccCA-ONIOM method.

[1]CHAPTER 4 MULTIREFERENCE CHARACTER for 4*d* TRANSITION METAL-CONTAINING MOLECULES

4.1 Introduction

Ab initio electron correlation methods are used to describe the interactions of electrons resulting from electronic interactions beyond the mean-field approximation of Hartree-Fork (HF) for atoms and molecules.[1] Electron correlation is commonly classified as two types: (a) dynamic correlation which is related to the movements of electrons with respect to each other, and (b) nondynamic (or static) correlation which arises from near-degeneracies of occupied and virtual orbitals. Single reference (SR) methods, such as truncated configuration interaction (CI), coupled cluster (CC), and Møller-Plesset (MP2) perturbation theory are able to account for a large portion of the dynamic correlation energy, but may not accurately describe molecular species with significant nondynamic correlation. For instance, coupled cluster with single, double, and perturbative triple excitations, CCSD(T),[2,3] with extrapolation of energetic properties to the complete basis set (CBS) limit, (CCSD(T)/CBS),[4–7] can result in thermodynamic energies within chemical accuracy (1 kcal mol^{-1} from experiment), on average, for small, main group molecules that are dominated by dynamic correlation.[8–15] However, for systems with significant nondynamic correlation, SR methods can result in significant errors. For example, the enthalpy of formation determined using CCSD(T)/CBS for O_3 results in a deviation of ~3.0 kcal mol^{-1} in comparison to experiment (34.10 ± 0.4 kcal mol^{-1}),[16,17] as a single-reference Hartree-Fock wavefunction reference, upon which CCSD(T) is built, provides an inadequate description of the

---

[1] This chapter is presented from J. Wang, S. Manivasagam, A. K. Wilson, "Multireference Character for 4d Transition Metal-Containing Molecules" *J. Chem. Theory Comput.* **2015**, DOI: 10.1021/acs.jctc.5b00861 with permission from American Chemical Society.

molecular ground state. Multiconfiguration or multireference (MR) methods, such as MCSCF,[18,19] MRCI,[20–22] and CASPT2,[23] are better suited than SR methods to describe degenerate and quasidegenerate states.

MR methods, however, are generally unable to address molecules as large as those that can be addressed by SR methods, since MR methods can quickly become computationally intractable due to the inclusion of all chemically relevant electrons and orbitals within an active space. In fact, an active space with ~32 active orbitals is near the present limit for most computer architectures and *ab initio* software packages, meaning that molecules such as VBr_4 or $Cr(CO)_6$ with a full valence CASSCF have reached the number of active orbitals that can be correlated utilizing internally contracted CASPT2.[24,25] An additional challenge in terms of MR methods is that their application, through the utilization of a restricted active space (i.e. restricted active space (RAS), generalized active space (GAS), occupation restricted multiple active space (ORMAS)) typically requires much more chemical intuition[26] than is required for SR methods since user selection of active space for MR methods can significantly affect the predicted properties (e.g. Ref. 27–29).

Before engaging in potentially complicated MR calculations, multireference diagnostic criteria can be used for *a priori* analysis of the SR or MR character for molecular systems. These diagnostics can help gauge suitable approaches (i.e., single or multireference wavefunction) to be used in calculations.

Several diagnostic criteria that can be used to assess the quality of the reference wavefunction have been developed to aid in determining the SR/MR character of molecules. One common diagnostic is the square of the CI coefficient, C_0, the leading configuration in CISD or CASSCF calculations.[30–33] Considering this diagnostic, molecules with significant MR character typically are indicated by a $C_0 \leq 0.95$ ($C_0^2 \leq 0.90$).[30] However, the determination of C_0^2 is not practical for large molecules due to the computational cost of large full-valence CASSCF calculations. For 3*d* TM-containing systems considered in a study by Jiang et al.,[25] C_0^2 was not used for molecules of more than a few atoms, due to the computational impracticality of using large full valence CASSCF to establish criteria.

As an alternative, the T_1 and D_1 diagnostics from CCSD calculations are widely used.[31,34–38] The T_1 diagnostic is defined as the Frobenius norm of the single substitution amplitudes vector (t_1) of the closed shell CCSD wavefunction (utilizing restricted HF orbitals), divided by the square root of the number of correlated electrons to address size consistency concerns:

$$T_1 = \|t_1\|/\sqrt{N} \qquad (1)$$

SR methods typically can perform well for molecules with a T_1 diagnostic smaller than 0.02, as suggested by Lee and Taylor in a study of 23 main group species. [31,34] The D_1 diagnostic by Janssen and Nielsen,[36]

$$D_1 = \|t_1\|_2 \qquad (2)$$

is based on the matrix 2-norm of t_1 from single excitations of the closed-shell CCSD wavefunction, and is closely related to the value of the largest single excitation amplitude. Molecules are commonly dominated by dynamic correlation if the D_1 diagnostic is smaller than 0.05.[41] Both T_1 and D_1 diagnostics have been extended to open shell CCSD wavefunctions.[37,38]

Lee performed a correlation analysis of T_1 and D_1 diagnostics and determined a squared correlation coefficient (R^2) of 0.96 based upon 10 open shell main group species and of 0.90 based upon 29 closed shell main group molecules.[38] Although this high degree of correlation between the T_1 and D_1 diagnostics shows that the diagnostics may provide similar predictions of the SR/MR character (as both diagnostics are related by the t_1 amplitudes), the T_1 diagnostic is an average value for the whole molecule and may fail to indicate problems for small regions of a molecule which can be indicated by a large D_1 diagnostic. For instance, para-benzyne has a T_1 of 0.0189, which means that the majority of the molecule can be described successfully with SR methods. However, D_1 (0.0646) of para-benzyne is greater than 0.05, which means that SR methods may fail to describe the electronic structure of a small part of this molecule properly.[36] Therefore, these two diagnostics were suggested to be used together as a diagnostic.[38]

Martin *et al*.[39,40] proposed an energy based diagnostic, %$TAE_e[(T_4 + T_5)]$ (the percentage of the connected quadruple and quintuple excitations contribute to total atomization energy (TAE)), during the development of W4 theory to indicate the imperfections of CCSD(T) due to nondynamic correlation effects in the system. Even though %$TAE_e[(T_4 + T_5)]$ provides the best indicator for nondynamic correlation effects, it is obviously not useful for practical applications due to its immense computational cost of CCSDTQ5 calculations. Therefore, this *a priori* diagnostic, %$TAE_e[(T)]$ (the percentage of the TAE resulting from triple excitations),

$$\%TAE_e[(T)] = 100 \times (TAE_e[CCSD(T)] - TEA_e[CCSD])/TAE_e[CCSD(T)] \tag{3}$$

is derived from the original rule of %$TAE_e[(T_4 + T_5)]$, but is more favorable than %$TAE_e[(T_4 + T_5)]$ as an indicator for nondynamic correlation in terms of CPU time, since no more than a CCSD(T)/cc-pVTZ single point calculation is required to calculate %$TAE_e[(T)]$. Additionally, %$TAE_e[(T)]$ can provide a prediction that strongly correlates with the %$TAE[(T_4 + T_5)]$ diagnostic (e.g., R^2 of 0.941 between %$TAE_e[(T)]$ and %$TAE[(T_4 + T_5)]$) for 20 mostly main group diatomic molecules).[39] Therefore, %$TAE_e[(T)]$ is not only much less expensive than %$TAE_e[(T_4 + T_5)]$, but also provides information similar to that of %$TAE_e[(T_4 + T_5)]$ and can be a reliable energy-based diagnostic for the performance of coupled cluster theory and the impact of MR character. Martin and co-workers suggested %TAE_e[(T)] ≤ 2% as the cutoff for proper SR calculations, %TAE_e[(T)] between 2-5% for mild nondynamic correlation, %TAE_e[(T)] between 5-10% for moderate nondynamic correlation, and %TAE_e[(T)] ≥ 10% for significant nondynamic correlation based upon the W4 set of molecules (main group molecules).[40]

The historical interpretation of the T_1, D_1 and %TAE_e[(T)] diagnostics was developed based upon small, well-behaved main group species. Due to partially filled *d* orbitals and the similar spatial extent and energy of the (n+1)*s* and n*d* orbitals, many transition metal (TM)-containing species are open shell systems that may have degenerate or low-lying nearly degenerate electronic states which may necessitate a multireference wavefunction-based approach to properly describe the near degeneracies.[41–43] As such, the computational study of TM-containing species can become quite challenging due to the need to address these nondynamical electron correlation effects.[44–49] Since TM species tend to have more degenerate states and a smaller atomic energy gap between n*d* and (n+1)*s* orbitals than main group species, and may have greater MR character as a result, the historical diagnostic criteria ($T_1 \leq 0.02$, $D_1 \leq 0.05$ and %TAE ≤ 10%) may be not appropriate for TM-containing molecules.

As the proper methodology choice (e.g., single or multireference wavefunction-based approach) is imperative in the description of molecules and diagnostics can serve as an important aid in this choice, investigations of suitable diagnostic criteria for TM-containing molecules are needed. A study by Jiang *et al.* focused on the analysis of several diagnostics (T_1, D_1, %TAE, and spin contamination) for the ccCA-TM/11all set of 225 3*d* TM-containing species. Jiang showed that 3*d* TM-containing species do need different diagnostic criteria than the diagnostic criteria, $T_1 \leq 0.02$, $D_1 \leq 0.05$, developed for main group species, and proposed $T_1 \leq 0.05$, $D_1 \leq 0.15$ as diagnostic criteria for SR methods for 3*d* TM-containing species. Jiang also found a moderate correlation between T_1 and D_1 diagnostics with a R^2 of 0.73 which is much smaller than the R^2 of 0.96 and 0.90, mentioned earlier, for small main group species.[25] Therefore, T_1 and D_1 yield less similar information about properties of the electronic structure for 3*d* TM-containing species than for main group species, which reaffirms that using T_1 and D_1 together rather than using a single diagnostic may provide a more reliable prediction of multireference character as molecules increase in size.[38]

Jiang's study also suggested that $|\%\mathrm{TAE_e}[(\mathrm{T})]| \leq 10\%$ can be used as a diagnostic criterion for the reliability of SR methods. However, using %TAE alone is not recommended as the sole criterion for molecules with small magnitude of $\mathrm{TAE_e}[(\mathrm{T})]$ values, as in this case slight changes of $\mathrm{TAE_e}[\mathrm{CCSD(T)}]$ and/or $\mathrm{TAE_e}[\mathrm{CCSD}]$ have a significant impact on $\%\mathrm{TAE_e}[(\mathrm{T})]$.[25]

Although, various diagnostics for 3*d* TM-containing molecules have been analyzed in Jiang's previous study, analysis of diagnostics on the basis of 4*d* TM-containing molecules is needed to provide more comprehensive insight into the MR character of *d*-block molecules. Since 4*d* TMs have different atomic energy gaps between the 4*d* and 5*s* orbitals than the 3*d* TMs have between the 3*d* and 4*s* orbitals,[50] 4*d* TM-containing species may exhibit a different degree of nondynamic correlation than 3*d* TM-containing species. Therefore, the diagnostic criteria for indicating MR character in 3*d* TM-containing species may not be ideal for 4*d* TM-containing species. Thus, in this study, the T_1, D_1, and $\%\mathrm{TAE_e}[(\mathrm{T})]$ diagnostic criteria are reexamined for 118 4*d* TM-containing species. These molecules investigated are from the 4*d*Hf-210 data set,[51] a set of molecules that includes a variety of binding types including hydrides, chalcogenides, halides, metal dimers, and coordination complexes.

4.2 Computational Methods

In this study, four diagnostics, C_0^2, T_1, D_1, and %TAE_e[(T)], were applied for 4*d* TM-containing species. (Hereafter, for simplicity, %TAE will be used to represent %TAE_e[(T)] where "e" represents the equilibrium geometry).

The optimized geometries for the 118 structures were determined in earlier work,[51] using B3LYP[52–54] in combination with the cc-pVTZ basis set and small core pseudopotentials were utilized for elements gallium through krypton and 4*d* TMs.[55–60] CASSCF, CCSD, and CCSD(T) single point calculations with cc-pVTZ-DK basis sets[55,59–62] (tight-*d* correlation consistent basis sets were used for Si, S, and Cl)[60,61] based upon a restricted open shell HF (ROHF) reference wavefunction were performed. Scalar relativistic effects were included using the second order one-particle Douglas-Kroll-Hess Hamiltonian.[63,64] The C_0^2 diagnostic was obtained by employing a CASSCF calculation with a full valence active space. All calculations were carried out using MOLPRO.[65]

4.3 Results and Discussions

4.3.1 4*d* Transition Metal Species.

The four diagnostics considered (C_0^2, T_1, D_1, and %TAE) were examined for a set of 118 4*d* TM-containing molecules, including hydrides, chalcogenides, halides, and metal dimers. The square of the leading coefficient, C_0^2, in the full valence CASSCF wavefunction directly shows the extent to which the configuration state functions (CSFs) correlate within the CAS. For molecules with C_0^2 greater than 0.9, nondynamic correlation (MR character) is significant.

Calculations for C_0^2 are only considered for molecules with three or less main group elements (83 molecules) due to the unfavorable cost of large full valence CASSCF calculations. The predictions from T_1, D_1, and %TAE diagnostics are to be calibrated by the predictions from the C_0^2 coefficient.

4.3.1.1 Hydrides.

The T_1, D_1, C_0^2, and %TAE diagnostics are presented for hydride systems in Table 4.1. The values of C_0^2 (greater than 0.9) for all hydrides except RuH (0.483) suggest that these hydrides do not have significant MR character. The corresponding T_1 and D_1 values of ZrH (0.042, 0.079), MoH (0.028, 0.061), TcH (0.032, 0.083), RuH (0.037, 0.100), and PdH (0.025, 0.057) are less than 0.045 and 0.120 and are within the 3*d* TM criteria ($T_1 \leq 0.05$ and $D_1 \leq 0.15$), but outside of the criteria of the main group thresholds ($T_1 \leq 0.02$, $D_1 \leq 0.05$). The %TAE values for all of the hydrides are less than %10. The %TAE is only positive for RuH. The negative values of %TAE for all other hydrides imply that the binding energies of these hydrides may be overestimated by the CCSD method.

4.3.1.2 Chalcogenides.

The T_1 and D_1 diagnostics together with C_0^2 and %TAE are presented for chalcogenides in Table 4.2. All values for C_0^2 are greater than 0.9 for the mono-chalcogenides of all early TMs (Y, Zr. Nb, and Mo), which indicates that the nondynamic correlation may not play an important role for these early TM molecules. The values of T_1, D_1, and %TAE for these early metal mono-chalcogenides species are less than 0.045, 0.120 and 10%, respectively. For late mono-chalcogenides, values of C_0^2 are less than 0.9 for RuO, RuS, and RhO, which suggests that these three molecules have significant nondynamic correlation. These three molecules also have large values of T_1, D_1, and %TAE (greater than 0.045, 0.120, and 10% respectively). Other late mono-chalcogenides (Tc and Pd species) have C_0^2 greater than 0.9, which implies that dynamic correlation is dominant for these molecules. Except for TcSe (T_1 = 0.056 and D_1 = 0.131), the late mono-chalcogenides predicted to be dominated by the dynamic correlation by C_0^2 have T_1 less than 0.045 and D_1 less than 0.120. The %TAE values for Tc and Pd species vary significantly from -2.2% (PdO) to 17.7% (TcO). For the three dioxides (ZrO_2, NbO_2, and MoO_2), the values of C_0^2 were less than 0.9, which suggests possible nondynamic correlation. The corresponding T_1, D_1, and %TAE of MoO_2 are greater than 0.045, 0.120, and 10%, respectively.

4.3.1.3 Halides.

The T_1, D_1, C_0^2, and %TAE diagnostics are presented for halides in Table 4.3. The C_0^2 values were only calculated for mono- and di-halides because of the large full valence CASSCF space requirements for larger molecules. The values of C_0^2 are greater than 0.9 for all mono-halides, which implies that the dynamic correlation is the dominant effect in the mono-halides. The T_1 diagnostics for all of the mono-halides are less than 0.045. All of the mono-halides have D_1 less than 0.12, except for RhCl (D_1=0.140). The predictions from %TAE diagnostics may not be consistent with predictions based upon C_0^2 diagnostics for several TM mono-halides. For instance, C_0^2 values imply SR character for RuBr (0.997), PdCl (0.992), and PdBr (0.992), yet these three molecules have very large absolute %TAE values (more than 19%), which suggests that MR methods may need to be utilized for calculations of their energetic properties. The D_1

diagnostic increases from fluorine to bromine for all of the metal mono-halides. Additionally, the values of the T_1 diagnostic increase from fluorine to bromine for the considered TM species, with the exception of Mo and Pd species. Overall, the nondynamic correlation is increasingly important for heavier halides. The same trends for the T_1 and D_1 diagnostics have been found for the 3*d* TM mono-halides in previous work.[25] The values of %TAEs decrease from fluorine to bromine for Zr, Mo, and Tc species.

For the dihalides, the C_0^2 values are consistently greater than 0.90, which implies SR character for these dihalides. All of the considered dihalides have T_1 less than 0.45 and D_1 less than 0.120, except $NbBr_2$ (T_1=0.046, D_1=0.136). Most of the SR dihalides have %TAE less than 10, except for $ZrBr_2$, $RhCl_2$, $RhBr_2$, $PdCl_2$, and $PdBr_2$. The inconsistent predictions by %TAE diagnostic from C_0^2 occur for $ZrBr_2$, $RhCl_2$, $RhBr_2$, $PdCl_2$, and $PdBr_2$, which are suggested to have SR character by C_0^2, but to have MR character by larger values of %TAE (more than 10); therefore, using %TAE alone may not make reliable predictions of nondynamic correlation. Similar to the trends found for mono-halides, from fluorine to bromine, the values of T_1 and D_1 diagnostics increase for Nb, Mo, Tc and Pd, but decrease for Zr and Ru.

As mentioned earlier, the C_0^2 was not determined for most trihalides and tetrahalides due to the impracticality of necessary large full valence CASSCF calculations. For trihalides of the early TMs, the T_1 diagnostic values are between 0.010-0.034, the D_1 diagnostic values are between 0.022-0.068, and the %TAE values are between 0.3-4.9% for Y, Zr, Nb, and Mo. On the basis of all of the considered diagnostics, nondynamic correlation may not be severe for these molecules. The values of all considered diagnostics increase from the late TM trihalides to the early TM trihalides, which suggests that nondynamic correlation effects have increasing prominence. Although $RuCl_3$ and $RhCl_3$ have T_1 values less than 0.045, these two molecule have very large %TAE values (18.9% and 28.9% respectively), which imply that the triples correction to the CCSD results is considerable and CCSD is not recommended for calculation of their bonding energies. The tetrahalides of Zr and Nb have T_1 less than 0.045, D_1 less than 0.12, and %TAE less than 10%, which suggests that a single reference treatment may be suitable. Although relatively small values of the T_1 and %TAE (except for TcF_4, which has a large %TAE value of 20.1%) suggest that dynamical correlation is important in the tetrahalides Mo and Tc, the D_1 values, which are between 0.08 and 0.15, suggest mild nondynamical correlation for these tetrahalides. Overall, nondynamic correlation may not be significant for most of the early transition meal trihalides and tetrahalides with small values of D_1 (less than 0.045), T_1 (less than 0.12), and %TAE (less than 10%).

4.3.1.4 Metal Dimers.

The T_1, D_1, C_0^2, and %TAE diagnostics are presented for TM dimers in Table 4.4. The values of C_0^2 (less than 0.8) imply that all TM dimers are dominated by significant nondynamic correlation effects. All of the dimers have the T_1 diagnostics greater than or near 0.045 (except for Pd_2), and D_1 diagnostics much greater than 0.120 (except for Zr_2, Mo_2, and Pd_2). The values of %TAE for the 4*d* TM dimers are much greater than 10% (except Nb_2), indicating CCSD is not reliable for calculating bonding energies of the 4*d* TM dimers. After further investigation of TAE values, a large difference was found between the TAE (= D_e for diatomic molecules) of CCSD(T) calculations and the D_0 of the experimental data of 4*d* TM dimers. Also, the negative values of CCSD(T) TAEs predict Tc_2, Ru_2, Rh_2, and Pd_2 are dissociative, while the positive experimental D_0 data show that extra energies are needed to dissociate all of the 4*d* TM dimers.

Therefore, coupled cluster calculations should not be applied to obtain qualitatively correct bonding energies for 4*d* TM dimers.

4.3.2 Discussion

The average values of diagnostics for different molecule classifications (e.g., halides, dihalides, and chalcogenides) are summarized in Table 4.5. The average values of the T_1, D_1, and %TAE diagnostics are 0.030, 0.076, and 4.7% respectively for the 110 TM molecules (TM dimers are not included), which suggests that the historical $T_1 \leq 0.02$, $D_1 \leq 0.05$ for main group species is not transferrable for 4*d* TM species. The average T_1 of 0.030, D_1 of 0.076, and %TAE of 4.7% for 4*d* TM-containing molecules system are smaller than the average T_1 of 0.044, D_1 of 0.127, and %TAE of 7.5% for 3*d* TM system, respectively.[25] Since the average T_1 and D_1 values of chalcogenides species are greater than that of other considered species (except for TM-dimers), the nondynamic correlation is more important for chalcogenides species. As shown in section 4.3.1, despite the wide range of metal-ligand bonding in the systems analyzed, similar cutoffs for SR calculations are found for T_1, D_1 and %TAE based on calibration from C_0^2. Most molecules, which are suggested to be dominated by nondynamic correlation with $C_0^2 \geq 0.9$, have $T_1 \geq 0.045$, $D_1 \geq 0.120$, and %TAE ≥ 10. Therefore, the general criteria is useful for reliable prediction of MR character of 4*d* TM-containing species. By considering the consistency of the predictions from the considered diagnostics (T_1, D_1, and %TAE) with predictions from C_0^2 for 83 4*d* TM-containing molecules, and the average values of the considered diagnostics for 110 4*d* TM-containing molecules, $T_1 \leq 0.045$, $D_1 \leq 0.120$, and %TAE $\leq 10\%$ are suggested as the diagnostic criteria for reliable calculations by SR methods for 4*d* TM-containing molecules.

Most molecules that are predicted to be dominated by dynamic correlation with C_0^2 more than 0.9 are also suggested to have SR character using the proposed diagnostic criteria. As shown in Table 4.6, 13 molecules have $T_1 \geq 0.045$, 12 molecules have $D_1 \geq 0.120$, and 22 molecules have %TAE ≥ 10%. Combining all three criteria, 7 molecules (RuO, RuS, RhO, MoO_2, RhC, RuGe, and RhSi) were found with $T_1 \geq 0.045$, $D_1 \geq 0.120$, and %TAE ≥ 10%. These molecules are predicted to be dominated by nondynamic correlation.

T_1, D_1, and %TAE diagnostics based on coupled cluster calculations are more accessible for large molecules than C_0^2 based on large full valence CASSCF calculations. Each diagnostic (T_1, D_1, and %TAE) has been compared to predictions made using the C_0^2 coefficient alone, since the square of the leading coefficient, C_0^2 in the large full valence CASSCF wavefunction, directly shows the extent to which the configuration state functions (CSFs) correlate within the CAS. If a strong linear correlation exists between each diagnostic (T_1, D_1, and %TAE) and the C_0^2 coefficient, then each diagnostic alone has the ability to indicate likely MR character. The C_0^2 values have been determined for 83 4*d* TM-containing molecules comprised of four or fewer atoms. In considering these species, the small squared correlation coefficient values (R^2 less than 0.2) in the relationships T_1 vs. C_0^2, D_1 vs. C_0^2, and %TAE vs. C_0^2 (Figures 4.1-4.3) suggest that there is very little correlation between each diagnostic and the C_0^2 coefficient.

The results of T_1, D_1, %TAE, and C_0^2 diagnostics do not always agree with one another; one diagnostic may indicate possible MR character while another does not suggest this as shown in section 4.3.1. Therefore, utilization of only one diagnostic is often not sufficient to indicate a system's possible MR character.[66–68] The combined use of various diagnostics can provide a more reliable prediction of MR character.[38,69] Thus, how the diagnostics should be combined to provide a useful indication of MR character has been considered. The relationships between the

T_1, D_1, and %TAE diagnostics are considered in this study. The correlation between T_1 and D_1 diagnostics for 110 4*d* TM-containing molecules without TM dimers shows a moderate linear correlation with an R^2 value of 0.655 (Figure 4.4). The suggested diagnostics criteria of $T_1 \leq 0.045$, and $D_1 \leq 0.120$ for reliable SR calculations can meet the correlation presented by R^2 of 0.655in Figure 4.4. Both T_1 and D_1 diagnostics have essentially no correlation with the %TAE diagnostic, with R^2 values of 0.062 and 0.081 respectively (Figure 4.5-4.6). Based on the W4 set of molecules, no good correlations were identified between the %TAE$_e$[(T_4 + T_5)] and the T_1 or D_1 diagnostics (R^2 of 0.38 and 0.33 respectively) either.[40]

Table 4.6 lists the 4*d* TM-containing molecules likely to exhibit significant nondynamic correlation effects (TM dimers are not included), based on different diagnostics. For molecules with three or less atoms, nine molecules have significant nondynamic correlation effects according to their C_0^2 coefficient from CASSCF calculations, they can be predicted to be MR systems by only applying the T_1 diagnostic criteria. A similar analysis identifies seven molecules when solely applying the D_1 diagnostic criteria and nine molecules when applying %TAE diagnostics criteria. As shown in Table 4.6, the %TAE criterion has more molecules consistent with C_0^2 than D_1 criterion. However, the %TAE criterion has more molecules that are inconsistent with C_0^2 than the D_1 criterion has, which is caused by high sensitivity of the %TAE diagnostic to minor changes in TAE$_e$[CCSD(T)] when the TAE is small.

Seven TM molecules and all of the TM dimers are gauged to have severe nondynamical correlation when using the three diagnostic criteria (T_1, D_1, and %TAE) together. Several molecules, which do not show severe nondynamical correlation effects by C_0^2 alone, are shown by other diagnostic criteria to possibly have severe nondynamical correlation, and are included in Table 4.6. For these species, the spin contamination ($\langle S^2-S_z^2-S_z \rangle$) can also be used to consider the reliability of the diagnostics.[25] Since strong nondynamic correlation effects may lead to a mixture of strongly correlated high spin states with the ground states, significant spin contamination, which has been eliminated by the ROHF reference wave function, may reappear with the utilization of UCCSD calculations. To consider this, the values of $\langle S^2\text{-}S_z^2\text{-}S_z \rangle$ in UCCSD calculations were examined for four molecules (TcSe, $NbBr_2$, $RuBr_3$, and TcF_4) that were predicted to have severe nondynamic correlation effects by at least two diagnostics of T_1, D_1, and %TAE but were predicted to be dominated by dynamic correlation effects by the C_0^2 in Table 4.6. The value of spin contamination for each of the molecules is less than 0.1 (TcSe =0.0068, $NbBr_2$ = 0.011, $RuBr_3$ = 0.007, and TcF_4 = 0.007), a factor that is used to gauge nondynamical correlation effects;[25] thus, nondynamical correlation may not be critical for these molecules despite the presence of large T_1, D_1, or %TAE.

4.4 Conclusion

In this study, four widely used diagnostics (T_1, D_1, C_0^2, and %TAE) were examined to consider their utility in indicating the nondynamic correlation of 4*d* TM-containing molecules with representative bonding types (hydrides, chalcogenides, halides, metal dimers, and several other TM species).

The historical criteria $T_1 \leq 0.02$ and $D_1 \leq 0.05$ for main group species is not practical for 4*d* TM-containing species. Thus, instead new T_1, D_1, and %TAE diagnostic criteria are suggested for 4*d* TM-containing species. While the recent criteria of $T_1 \leq 0.05$, $D_1 \leq 0.15$, and %TAE $\leq$ 10% for 3*d* TM-containing species can be utilized for most 4*d* TM-containing species, the criteria $T_1 \geq 0.045$, $D_1 \geq 0.120$, and %TAE $\geq$ 10% is proposed as a more fine-tuned set of criteria to aid in identifying the significance of nondynamic correlation for 4*d* TM-containing species.

4.5 Reference

(1) Friesner, R. A.; Murphy, R. B.; Beachy, M. D.; Ringnalda, M. N.; Pollard, W. T.; Dunietz, B. D.; Cao, Y. *J. Phys. Chem. A* **1999**, *103*, 1913–1928.

(2) Watts, J. D.; Bartlett, R. J. *Int. J. Quantum Chem.* **1993**, *48*, 51.

(3) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

(4) Dunning, Jr. T. H. *J. Phys. Chem. A* **2000**, *104*, 9062.

(5) Peterson, K. A.; Figuregen, D.; Dixon, D. A. *Theor. Chem. Acc.* **2012**, *131*, 1.

(6) Feller, D.; Peterson, K. A.; Dixon, D. A. *Mol. Phys.* **2012**, *110*, 2381.

(7) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chem. Rev.* **2012**, *112*, 4–74.

(8) Deleuze, M. S. *J. Am. Chem. Soc.* **2000**, *122*, 1130.

(9) Okamoto, K. M. *J. Phys. Chem. B* **2001**, *105*, 2001.

(10) Frankcombe, T. J.; Bhatia, S. K.; Smith, S. C. *Carbon N. Y.* **2002**, *40*, 2341.

(11) Bultinck, P. *Farady Discuss.* **2007**, *135*, 347.

(12) Wu, J. S.; Pisula, W.; Mullen, K. *Chem. Rev.* **2007**, *107*, 718.

(13) Hajgató, B.; Deleuze, M. S.; Tozer, D. J.; Proft, F. *J. Chem. Phys.* **2008**, *129*, 084308.

(14) Koshino, M.; Ando, T. *Solid State Commun.* **2009**, *149*, 1123.

(15) Yurtsever, E. *J. Phys. Chem. A* **2009**, *113*, 924.

(16) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Phys. Chem.* **2004**, *120*, 4129.

(17) Morgon, N. H. *Int. J. Quantum Chem.* **2012**, *112*, 3256–3260.

(18) Hinze, J. *J. Chem. Phys.* **1973**, *56*, 6424.

(19) Werner, H. J.; Meyer, W. *J. Chem. Phys.* **1980**, *73*, 2342.

(20) Buenker, R. J. P. *J. Chem. Phys.* **1970**, *53*, 1368.

(21) Shepard, R. *Int. J. Quantum Chem.* **1987**, *31*, 33.

(22) Hanrath, M.; Engel, B. *J. Chem. Phys.* **1997**, *225*, 197.

(23) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.

(24) Jiang, W.; DeYonker, N. J.; Determan, J. J.; Wilson, A. K. *J. Phys. Chem. A* **2012**, *116*, 870–885.

(25) Jiang, W.; DeYonker, N. J.; Wilson, A. K. *J. Chem. Theory Comput.* **2012**, *8*, 460–468.

(26) Peterson, K. A. *J. Chem. Phys.* **1995**, *102*, 262.

(27) Camacho, C.; Yamamoto, S.; Witek, H. A. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5128–5134.

(28) Ma, D.; Li Manni, G.; Gagliardi, L. *J. Chem. Phys.* **2011**, *135*, 044128.

(29) Slavícek, P.; Martínez, T. J. *J. Chem. Phys.* **2010**, *132*, 234102.

(30) Langhoff, S. R.; Davidson, E. R. *Int. J. Quantum Chem.* **1974**, *8*, 61.

(31) Lee, T. J.; Taylor, P. R. *Int. J. Quant. Chem. Symp.* **1989**, *S23*, 199.

(32) Sears, J. S.; Sherrill, C. D. *J. Phys. Chem. A* **2008**, *112*, 6741.

(33) Sears, J. S.; Sherrill, C. D. *J. Phys. Chem. A* **2008**, *112*, 3466.

(34) Lee, T. J.; Rice, J. E.; Scuseria, G. E.; Schaefer, H. F. *Theor. Chim. Acta* **1989**, *75*, 81.

(35) Jayatilaka, D.; Lee, T. J. *J. Chem. Phys.* **1993**, *98*, 9734.

(36) Janssen, C. L.; Nielsen, I. M. B. *Chem. Phys. Lett.* **1998**, *290*, 423.

(37) Leininger, M. L.; Nielsen, I. M. B.; Crawford, T. D.; Janssen,C. L. *Chem. Phys. Lett.* **2000**, *238*, 431.

(38) Lee, T. J. *Chem. Phys. Lett.* **2003**, *372*, 362.

(39) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(40) Karton, A.; Daon, S.; Martin, J. M. L. *Chem. Phys. Lett.* **2011**, *510*, 165.

(41) Langhoff, S. R.; Bauschlicher, C. W. *Annu. Rev. Phys. Chem.* **1988**, *39*, 181–121.

(42) Moore, C. E. *At. energy levels. Washington, DC US Natl. Bur. Stand circ. no. 467*.

(43) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757–10816.

(44) Silverstone, H. J.; Sinanoglu, O. *J. Chem. Phys.* **1966**, *44*, 1899.

(45) Schmidt, M. W.; Gordon, M. S. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233.

(46) Roos, B. O.; Andersson, K.; Fulscher, M. P.; Malmqvist, P. -A.; Serrano-Andres, L.; Pierloot, K.; Merchan, M. *Adv. Chem. Phys.* **1996**, *93*, 219.

(47) Siegbahn, P. E. M. *Adv. Chem. Phys.* **1996**, *93*, 333.

(48) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2006**, *125*, 074110.

(49) Li, X.; Paldus, J. *J. Chem. Phys.* **2007**, *126*, 234303.

(50) Hay, P. J.; Wadt, W. R. *J. Chem. Phys*. **1985**, *82*, 1.

(51) Manivasagam, S.; Laury, M. L.; Wilson, A. K. *J. Phys. Chem. A* **2015**, DOI: 10.1021/acs.jpca.5b02433.

(52) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098–3100.

(53) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(54) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(55) Dunning, Jr. T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(56) Dunning, Jr. T. H. *J. Chem. Phys.* **1994**, *100*, 2975.

(57) Peterson, K. A. *J. Chem. Phys.* **2003**, *119*, 11099.

(58) Peterson, K. A.; Puzzarini, C. *Theor. Chem. Acc.* **2005**, *114*, 283.

(59) Peterson, K. A.; Figuregen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.

(60) Woon, D. E.; Dunning, Jr. T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(61) Dunning, Jr. T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.

(62) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, Jr. T. H. *J. Chem. Phys.* **1999**, *110*, 7667.

(63) Peterson, K. A; Figuregen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.

(64) Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 64107.

(65) MOLPRO, version 2009. 1. http://www. molpro. ne. (accessed; November 1, 2011).

(66) Fogueri, U.; Kozuch, S.; Karton, A. Martin, J. L. *Theor. Chem. Acc.* **2012**, *132*, 1.

(67) Jiang, W.; Jeffrey, C. C.; Wilson, A. K. *J. Phys. Chem. A* **2012**, *116*, 9969.

(68) Zhao, Y.; Tishchenko, O.; Gour, J. R.; Li, W.; Lutz, J. J.; Piecuch, P.; Truhlar, D. G. *J. Phys. Chem. A* **2009**, *113*.

(69) Nielsen, I. M. B.;Janssen, C. L. *Chem. Phys. Lett.* **1999**, *310*, 568.

(70) Laury, M. L.; DeYonker, N. J.; Jiang, W.; Wilson, A. K. *J. Chem. Phys.* **2011**, *135*, 214103.

(71) Verhaegen, G.; Smoes, S.; Drowart, J. *J. Chem. Phys*. **1964**, 40, 239.

(72) Arrington, C. A.; Blume, T.; Morse, M. D.; Doverstaal, M.; Sassenberg, U. *J. Phys. Chem.* **1994**, *124*, 194301.

(73) Gupta, S. K.; Pelino, M.; Gingerich, K. A. *J. Phys. Chem*. **1979**, *83*, 2335.

(74) Hopkins, J. B.; Langridge-Smith, P. R. R.; Morse, M. D.; Smalley, R. E. *J. Chem. Phys*. **1983**, *78*, 1627.

(75) Brewer, L.; Winn, J. S. *Faraday, Symp. Chem. Soc*. **1980**, *14*, 126.

(76) Cocke, D. L.; Gingerich, K. A. *J. Chem. Phys*. **1974**, *60*, 1958.

(77) Shim, L.; Gingerich, K. A. *J. Chem. Phys*. **1984**, *80*, 5107.

4.6 Figures and Tables

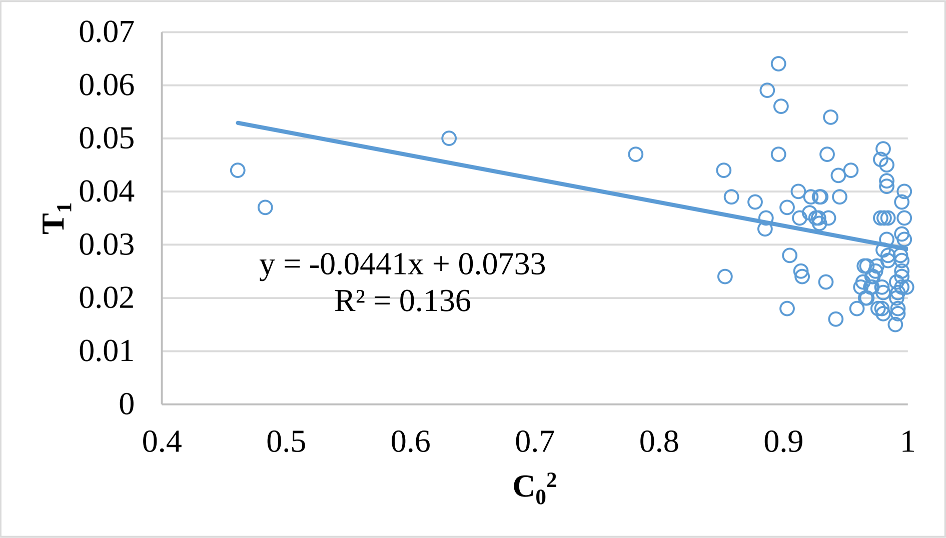


Figure 4.1. Scattering plot of C_0^2 and T_1 diagnostics for 83 4*d* TM-containing molecules.

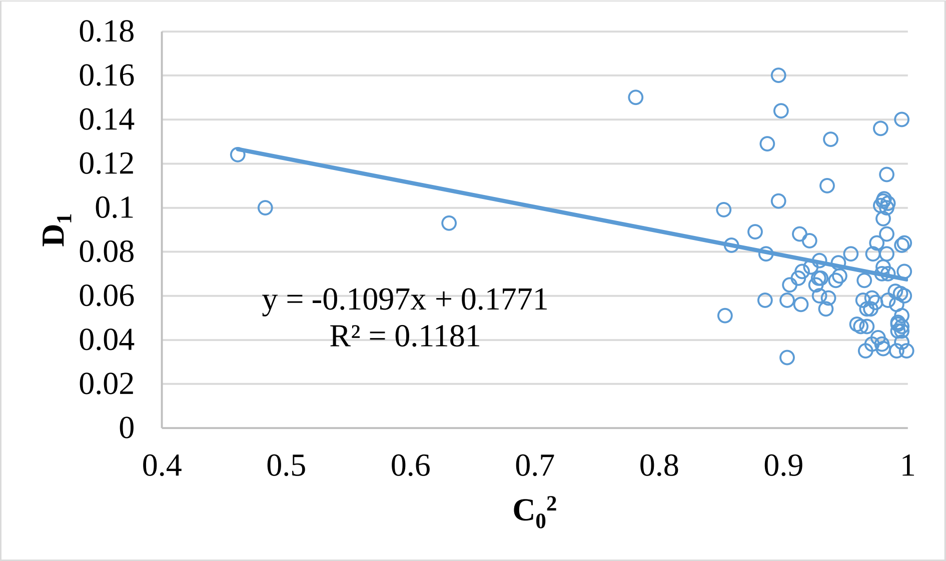


Figure 4.2 Scattering plot of C_0^2 and D_1 diagnostics for 83 4*d* TM-containing molecules.

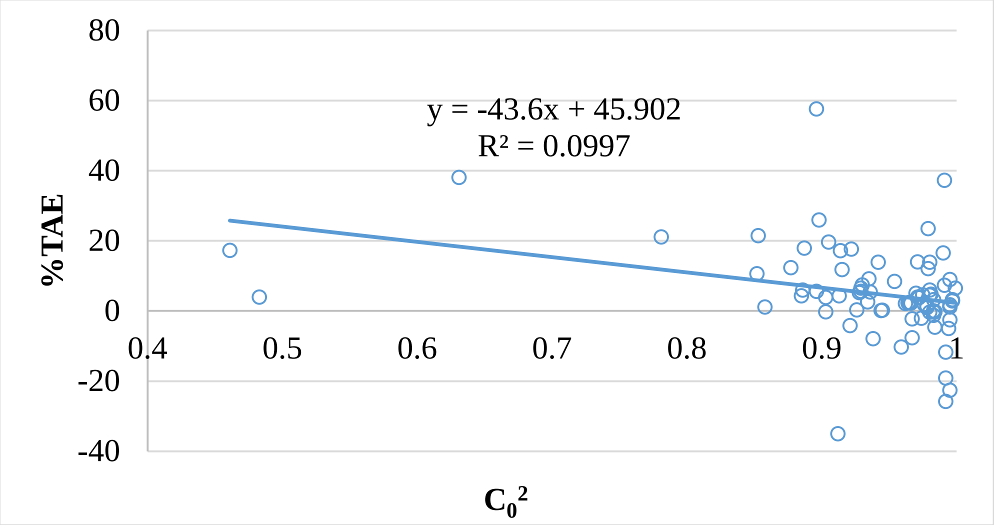


Figure 4.3 Scattering plot of C_0^2 and %TAE diagnostics for the 4*d* TM-containing molecules, excluding metal dimers.

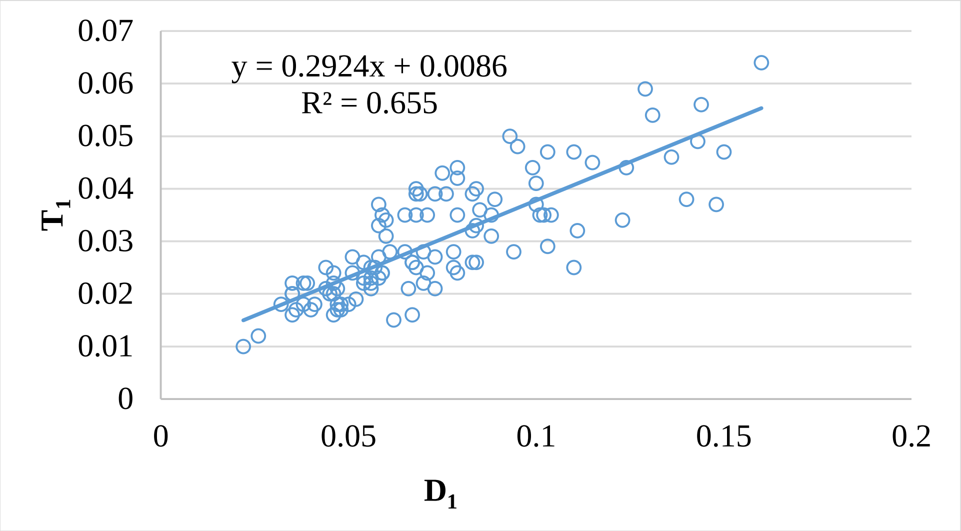


Figure 4.4 Scattering plot of T_1 and D_1 diagnostics for 110 4*d* TM-containing molecules without TM dimers.

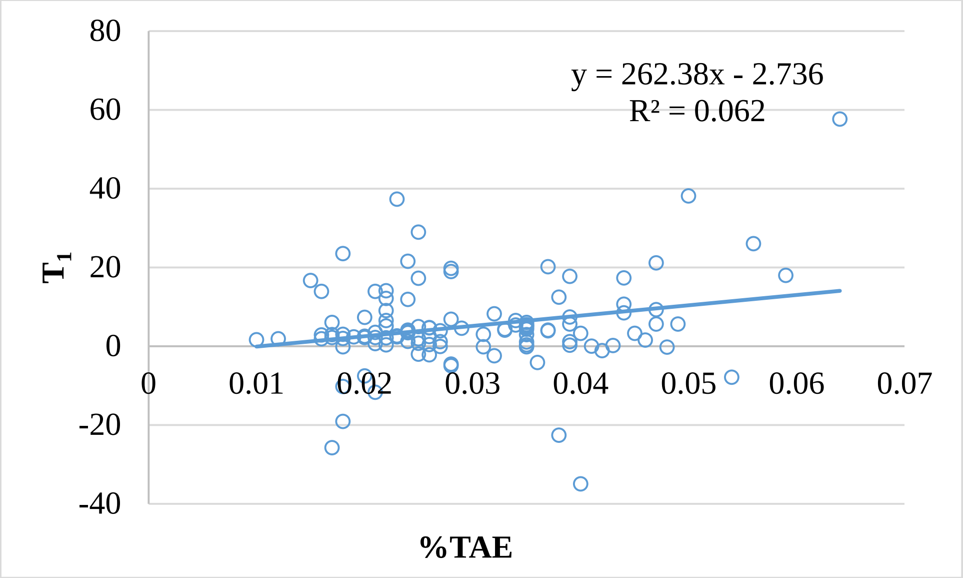


Figure 4.5 Scattering plot of %TAE versus T_1 for 110 4*d* TM-containing molecules.

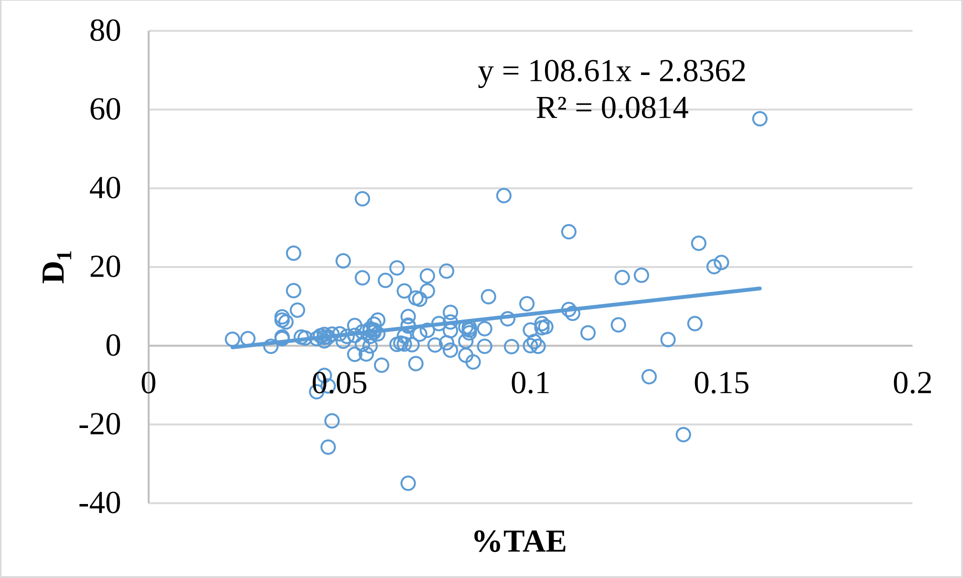


Figure 4.6 Scattering plot of %TAE versus D_1 for 110 4*d* TM-containing molecules.

Table 4.1 Diagnostics for metal hydrides calculated by CCSD, CCSD(T), or CASSCF with the cc-pVTZ-DK basis set

Molecule	T_1	D_1	C_0	C_0^2	%TAE
YH	0.018	0.032	0.950	0.903	-0.2
ZrH	0.042	0.079	0.991	0.983	-1.2
NbH	0.048	0.095	0.990	0.980	-0.3
MoH	0.028	0.061	0.997	0.994	-5.0
TcH	0.032	0.083	0.998	0.995	-2.5
RuH	0.037	0.100	0.695	0.483	4.0
RhH	0.020	0.046	0.983	0.967	-7.6
PdH	0.025	0.057	0.987	0.974	-2.1

Table 4.2 Diagnostics for metal chalcogenides calculated by CCSD, CCSD(T), or CASSCF with the cc-pVTZ-DK basis set

Molecule	T_1	D_1	C_0	C_0^2	%TAE
YO	0.035	0.079	0.967	0.936	5.4
YS	0.035	0.068	0.964	0.928	5.2
YSe	0.039	0.076	0.964	0.929	5.6
ZrO	0.034	0.060	0.964	0.929	6.5
ZrS	0.039	0.068	0.965	0.930	7.4
ZrSe	0.044	0.079	0.977	0.954	8.4
NbO	0.035	0.065	0.962	0.926	0.3
NbS	0.039	0.069	0.972	0.945	0.2
NbSe	0.043	0.075	0.972	0.944	0.1
MoO	0.040	0.068	0.955	0.912	-35.0
TcO	0.039	0.073	0.960	0.922	17.7
TcS	0.044	0.110	0.967	0.935	9.2
TcSe	0.054	0.131	0.968	0.938	-7.9
RuO	0.056	0.144	0.948	0.898	26.0
RuS	0.064	0.160	0.946	0.896	57.6
RhO	0.072	0.182	0.679	0.461	17.3
PdO	0.026	0.054	0.983	0.967	-2.2
PdS	0.018	0.047	0.979	0.959	-10.3
ZrO_2	0.035	0.079	0.941	0.886	6.0
NbO_2	0.039	0.083	0.926	0.858	1.1

MoO_2	0.045	0.121	0.923	0.852	10.6

Table 4.3 Diagnostics for metal halides calculated by CCSD, CCSD(T), or CASSCF in combination with the cc-pVTZ-DK basis set

Molecule	T_1	D_1	C_0	C_0^2	%TAE
YF	0.020	0.035	0.983	0.966	2.2
ZrF	0.022	0.038	0.985	0.971	14.0
ZrCl	0.022	0.054	0.985	0.970	5.1
ZrBr	0.024	0.059	0.985	0.971	4.0
NbF	0.027	0.058	0.992	0.984	-0.1
NbCl	0.031	0.088	0.992	0.983	-0.2
NbBr	0.035	0.102	0.992	0.984	-0.2
MoF	0.022	0.035	0.999	0.999	6.5
MoCl	0.020	0.035	0.996	0.991	7.3
MoBr	0.022	0.039	0.997	0.995	9.0
TcF	0.025	0.044	0.998	0.995	1.8
TcCl	0.024	0.046	0.998	0.995	1.2
TcBr	0.027	0.051	0.998	0.995	1.1
RuF	0.031	0.060	0.999	0.997	3.0
RuCl	0.035	0.071	0.999	0.997	2.9
RuBr	0.040	0.084	0.999	0.997	3.2
RhCl	0.038	0.140	0.997	0.995	-22.6
PdF	0.021	0.044	0.996	0.992	-11.7
PdCl	0.017	0.047	0.996	0.992	-25.8
PdBr	0.018	0.048	0.996	0.992	-19.1
YF_2	0.018	0.041	0.988	0.976	1.9
ZrF_2	0.022	0.046	0.981	0.962	2.1
$ZrCl_2$	0.017	0.036	0.990	0.980	6.0
$ZrBr_2$	0.018	0.038	0.989	0.979	23.5
$NbCl_2$	0.035	0.101	0.989	0.978	1.0
$NbBr_2$	0.046	0.136	0.989	0.978	1.5
MoF_2	0.024	0.079	0.986	0.972	3.8
$MoCl_2$	0.026	0.084	0.987	0.975	4.6
$MoBr_2$	0.029	0.103	0.990	0.980	4.5
TcF_2	0.023	0.054	0.966	0.934	2.6
$TcCl_2$	0.023	0.058	0.982	0.964	2.3
$TcBr_2$	0.026	0.067	0.982	0.965	2.4

RuF_2	0.028	0.070	0.992	0.984	-4.6
$RuCl_2$	0.045	0.115	0.991	0.983	3.2
$RuBr_2$	0.041	0.100	0.991	0.983	0.0
RhF_2	0.035	0.104	0.990	0.981	4.8
$RhCl_2$	0.037	0.109	0.990	0.980	13.9
$RhBr_2$	0.022	0.070	0.990	0.979	12.1
PdF_2	0.023	0.056	0.996	0.991	7.3
$PdCl_2$	0.015	0.062	0.995	0.990	16.6
$PdBr_2$	0.029	0.107	0.970	0.942	13.9
YF_3	0.035	0.059	-	-	1.8
YCl_3	0.010	0.022	-	-	1.6
YBr_3	0.012	0.026	-	-	1.8
ZrF_3	0.021	0.047	-	-	2.2
$ZrCl_3$	0.017	0.040	-	-	2.2
$ZrBr_3$	0.020	0.045	-	-	2.5
$NbCl_3$	0.022	0.056	-	-	0.3
$NbBr_3$	0.026	0.067	-	-	0.4
$MoCl_3$	0.025	0.068	-	-	4.9
TcF_3	0.034	0.123	-	-	5.3
RuF_3	0.033	0.084	-	-	4.0
$RuCl_3$	0.028	0.078	-	-	18.9
$RuBr_3$	0.049	0.143	-	-	5.6
$RhCl_3$	0.024	0.093	-	-	28.9
ZrF_4	0.019	0.052	-	-	2.3
$ZrCl_2Br_2$	0.017	0.048	-	-	2.9
$ZrClBr_3$	0.018	0.050	-	-	3.0
$ZrCl_3Br$	0.016	0.046	-	-	2.8
$NbBr_4$	0.025	0.078	-	-	0.7
$NbCl_4$	0.021	0.066	-	-	0.6
$MoCl_4$	0.028	0.094	-	-	6.8
MoF_4	0.026	0.083	-	-	4.7
$MoBr_4$	0.032	0.111	-	-	8.2
TcF_4	0.037	0.148	-	-	20.1

Table 4.4 Diagnostics for metal dimers calculated by CCSD, CCSD(T), or CASSCF in combination with the cc-PVTZ-DK basis set

Molecule	T_1	D_1	C_0	C_0^2	TAE CCSD	TAE CCSD(T)	exptl. D_0[a]	%TAE
Y_2	0.098	0.202	0.883	0.780	14.824	20.669	37.36	28.3
Zr_2	0.047	0.095	0.831	0.690	-2.221	11.666	70.38	119.0
Nb_2	0.150	0.438	0.520	0.271	4.289	4.661	126.37	8.0
Mo_2	0.043	0.106	0.810	0.656	-0.598	-0.133	98.93	-349.6
Tc_2	0.043	0.135	0.860	0.739	-43.173	-30.568	-	-41.2
Ru_2	0.156	0.586	0.856	0.733	0.418	12.807	73.56	96.7
Rh_2	0.109	0.406	0.741	0.549	-30.505	-42.325	67.34	27.9
Pd_2	0.034	0.097	0.790	0.624	-22.018	-40.857	23.75	46.1

[a] Experimental data from ref 71 - 77.

Table 4.5 Average diagnostic values for different molecule classifications

classification	T_1	D_1	%TAE
Hydrides (8)	0.031	0.069	-1.9
Chalcogenides (21)	0.042	0.090	6.2
Mono-halides (20)	0.026	0.059	-0.9
Dihalides (20)	0.028	0.079	7.3
Trihalides (14)	0.025	0.068	5.7
Tetrahalides (10)	0.024	0.078	5.2
TM-Dimers (8)	0.085	0.259	-8.1
Other small coordination complexes (17)	0.036	0.086	11.5
Overall (118)	0.037	0.098	3.1
Overall without TM-Dimers (110)	0.030	0.076	4.7

Table 4.6 TM-containing molecules that may be dominated by nondynamic correlation effects (TM dimers are not included), as gauged by noted diagnostic criteria

$C_0^2 \leq 0.9$[b]		$T_1 \geq 0.045$		$D_1 \geq 0.12$		%TAE ≥ 10		
RuH	YC_2	NbH	YB_2	TcSe	TcF_4	TcO	$RhBr_2$	RuGe
RuO	YN	TcSe	YN	RuO	RhC	RuO	$PdCl_2$	RhSi
RuS	RhC	RuO	RhC	RuS	RuGe	RuS	$PdBr_2$	RhGe
RhO	RuGe	RuS	RuGe	RhO	RhSi	RhO	$RuCl_3$	PdSi
ZrO_2	RhSi	RhO	RhSi	MoO_2		MoO2	$RhCl_3$	PdGe
NbO_2	RhGe	MoO_2		$NbBr_2$		ZrF	TcF_4	PdCO
MoO_2		$NbBr_2$		TcF_3		$ZrBr_2$	YB_2	
YB_2		$RuBr_3$		$RuBr_3$		$RhCl_2$	RhC	

[b] C_0^2 is only considered for molecules with three or less atoms.

Table 4.7 Diagnostics for other metal species calculated by CCSD, CCSD(T), or CASSCF in combination with the cc-pVTZ-DK basis set

Molecule	T_1	D_1	C_0	C_0^2	%TAE
YB_2	0.050	0.093	0.794	0.631	38.1
YC_2	0.033	0.058	0.941	0.885	4.3
YN	0.047	0.103	0.946	0.896	5.6
YC_4	0.027	0.073	-	-	3.9
ZrN	0.037	0.058	0.950	0.903	3.9
ZrOF	0.035	0.088	0.955	0.913	4.3
$ZrOF_2$	0.024	0.059	-	-	3.4
$ZrOCl_2$	0.021	0.056	-	-	3.5
RhC	0.072	0.182	0.936	0.877	12.4
RuSi	0.036	0.085	0.960	0.921	-4.2
RuGe	0.047	0.150	0.884	0.781	21.1
RhSi	0.059	0.129	0.942	0.887	17.9
RhGe	0.024	0.051	0.924	0.853	21.5
PdSi	0.025	0.056	0.956	0.914	17.2
PdGe	0.028	0.065	0.952	0.905	19.7
PdCO	0.024	0.071	0.956	0.915	11.8

CHAPTER 5 THE IMPACT OF COMPUTATIONAL METHODS ON PREDICITON OF PKA OF TRANSITION METAL HYDRIDES VIA A QM/QM APPROACH

5.1 Introduction

Over the past decade, density functional theory (DFT) has become a preferred tool to study molecular properties, such as geometries, spectroscopic constants, and energetic properties,[1–8] of transition metal (TM) containing systems due to its low computational cost and comparable accuracy to account for electron correlation in comparison to wave function theory (WFT). While the development of various DFT functionals has motivated their wide applications, care is still needed for the selection of the right methodology prior to the investigation of a particular scientific problem at hand. In 2001, J. Perdew proposed the famous Jacob’s ladder to classify DFT functionals into five rungs that present the hierarchy of density approximations.[9] The local density approximation (LDA), also known as the local spin density approximation (LSDA) is on the first rung. Second rung functionals, generalized gradient approximation (GGA), improve on the LDA functionals by including the dependence on the spin-density gradients, but introduce self-interaction errors.[10,11] On the third rung is the meta-generalized gradient approximation (M-GGA) functionals that decrease the amount of self-interaction by considering the kinetic energy density as an additional variable. The hybrid generalized gradient approximation (H-GGA) and the hybrid meta-generalized gradient approximation (HM-GGA) functionals on the fourth rung explicitly include a percentage of Hartree-Fock (HF) exchange that addresses the self-exchange shortcoming of DFT functionals. Finally on the fifth rung are the double hybrid generalized gradient approximation (DH-GGA) functionals that add correlation from second order perturbation theory to the correlation functional.[12,13]

Based on Perdew's Jacob's ladder, each rung contains more factors that are added to the functionals of the rung below. As a result, the inherent complexity of the functional increases as you progress up Jacob's ladder which presents an assumption of greater accuracy pertaining to the factors now included in the higher rung functionals that are not present in the lower rung functionals. Unfortunately, the actual reliability for DFT functionals does not always systematically agree with the expectation from Perdew's Jacob's ladder. Therefore, the rational choice of DFT functionals should be determined by carefully considering the calibration of DFT functionals with experiments or high accuracy WFT methods for a particular application.

There are many studies focused on gauging the utility of DFT functionals in the field of TM chemistry in the gas phase. Zhao and Truhlar developed and evaluated the M05 and M06 suite of functionals that were parameterized for the atomic and molecular properties of a robust database that included both main group and TM species.[6,7,14–17] They found that M06 and M06L performed the best based upon the 3*d* TM reaction energies of TMRE48 database.[6,7] They also found that an increase in the percentage of HF exchange included in functionals improved the accuracy of calculated reaction barrier heights but adversely affect the calculated atomization energies for TM species. Schultz and coworkers[18,19] calculated bond dissociation energies for 21 TM species, metal-metal and metal-ligand bond lengths for 13 TM species, and binding and atomization energies for 8 TM dimers using 57 functionals with double- and triple- ξ basis sets. The best results were obtained with G96LYP, followed closely by MPWLYP1M, XLYP, BLYP, and MOHLYP. BP86 and TPSS outperformed the other tested functionals in the study of dissociation energies, bond lengths, and vibrational frequencies of 3*d* TM systems including TM dimers, monohydrides, monoxides, mononitrides, and monofluorides with triple- and occasionally quadruple- ξ basis sets.[20] Riley and Merz[21] has examined the performance of 12 functionals with 6-31G* and TZVP basis sets for the calculation of enthalpies of formation of 94 TM species. They reported that the performance of DFT functionals was affected by the utilization of the TM, the number of ligands, the type of ligands, and the basis sets. Overall, TPSS1KCIS with TZVP basis set provided the lowest mean absolute deviation of 9.1 kcal mol^{-1}.

Tekarli et al.[22] have calculated the gas-phase enthalpies of formation of 19 3d TM-containing species to assess the performance of 44 density functionals paired with cc-pVTZ and cc-pVQZ basis sets. Among the considered functionals, B97-1, PBE1KCIS, TPSS1KCIS, B97-2, and B98 produced the lowest mean absolute deviations relative to experiment of 6.9, 8.1, 9.6, 9.7, and 10.7 kcal mol^{-1}, respectively. They also found that including HF exchange in the functionals lead to more accurate results while functionals with HF exchange greater than 40% of resulted in decline in accuracy. The enthalpies of formation of 30 4*d* species were calculated by 22 DFT functionals combined cc-pVTZ and cc-pVQZ basis sets (small core, energy-consistent relativistic pseudopotentials were included).[23] Of the functional utilized, the functionals that yielded the lowest mean absolute deviation from experiment are B2GP-PLYP (4.25 kcal mol^{-1}) and mPW2-PLYP (5.19 kcal mol^{-1}). The cc-pVQZ basis set improved on the cc-pVTZ basis set by 0.5 kcal mol^{-1}. The inclusion of HF exchange did not lower the deviation of calculated enthalpies of formation by DFT functionals from experiment. Jiang et al.[24] assessed the performance of both single and double hybrid density functionals for the calculation of enthalpies of formation of 193 molecules in ccCA-TM/11 set. Overall, mPW2-PLYP provided results closest to experimental data than the other considered DFT functionals with a mean absolute deviation of 3.0 kcal mol^{-1}. They also found that the accuracy of DFT functionals showed dependence on the choice of species. For instance, open shell molecules and bigger molecules generally had higher mean absolute deviations from experiment than closed shell molecules and smaller molecules, respectively. Additionally, B97-1 was recommended for V, Cr, Mn, Fe, and Ni species, while Sc and Zn, Co, and Cu species preferred TPSS1KCIS, PBE1KCIS, and ωB97X, respectively.

Commonly, chemical reactions including organometallic complexes are carried out in solvated phase instead of gas phase. The studies of certain properties of TM-containing molecules may not be described appropriately in gas phase in spite of its simplicity unless one can be reasonably confident that solvent effects have no influence on the particular properties being studied. However, compared with the number of gas phase studies for the performance of density functionals with TM chemistry, publications on the solvent effects of TM compounds received less attention with respect to the performance of density functionals. In order to obtain a more comprehensive view of the utility of DFT functionals in computation, the investigation of the ability of DFT functionals to describe chemical properties in solvation phase is of great interest in this study.

The accuracy of DFT, however, is commonly not transferable between different TM systems. Additionally, it is not practical to include all chemical properties in one study. For the solvated phase, pK_as exhibit the strongest effects of solvation relative to their gas phase analogs due to the charge separation of the species involved. Therefore, this study focused on investigating the ability of density functionals to predict the acidity (pK_a) of TM hydrides that is significantly related to the reactivity of TM hydrides. TM hydrides are important intermediates in many important catalytic and stoichiometric process such as hydrogenation and hydroformylation.[25–33] Although the pK_a values of various TM hydrides have been measured experimentally, it is worth noting that experimental characterization of pK_a is not accessible for all TM hydrides even in the far future. As an alternative to experimental measurement, reliable DFT functionals may become necessary to predict the pK_a of TM hydrides so that their reactivity can be predicted quantitatively under a given set of conditions.

However, well-defined TM hydrides that are soluble in organic solvents usually carry bulky ligands. To balance the cost and accuracy of theoretical methods for these massive molecules, the use of the ONIOM scheme[34–42] is preferred for investigation of the properties of TM hydrides, such as pK_a, rather than using expensive theoretical methods for entire TM hydrides. In the two layer ONIOM scheme that was applied in this study, the investigated molecular system is defined in two regions: the small model system that is chemically important (e.g. bond formation and breaking) is treated with a more accurate and expensive method (the high-level method) and the real system (consists of the entire molecule) is treated with a less accurate and expensive method (the low-level method).

In this case, the investigation of optimal combinations of DFT functionals and basis sets for the improvement on the accuracy of ONIOM scheme on prediction of pK_a of TM hydrides is important. In the study by Qi et al.,[43] a two-layer ONIOM method with DFT(B3LYP, B3PW91, and B3P86)/LANL2DZ+p as high-level methods and HF/LANL2MB as the low-level method resulted in the deviation of 1.5 pK_a units from experiment for a variety of TM hydrides in acetonitrile. However, they targeted on developing a new *ab initio* protocol to calculate the pK_a of TM hydrides and only considered three DFT functionals. To assess the appropriateness of DFT functionals within the ONIOM scheme for TM hydrides in solvated phase, comprehensive studies must be carried out where a wide variety of functional should be tested.

For this investigation, the utility of the two layer ONIOM scheme with a variety of density functionals and several basis sets to accurately calculate pK_a values of Group 10 TM hydrides ($[HNi(depe)_2]^+$, $[HNi(depp)_2]^+$, $[HNi(PNP)_2]^+$, $[HPd(depe)_2]^+$, $[HPd(depp)_2]^+$, $[HPd(PNP)_2]^{+,}$ $[HPt(depe)_2]^+$, $[HPt(depp)_2]^+$, $[HPt(PNP)_2]^+$) relative to the experimental data[44–47] was evaluated. Additionally, the influence of the addition of HF exchange and dispersion corrections was considered. As shown in the above examples[22–24] and several other studies,[48–50] the choice of basis set had an impact on the performance of DFT functionals. Therefore, an understanding of the basis set influence was also assessed by using several basis sets. Since the performance of DFT functionals also depended on the size of the molecules,[21,24] this study also examined the influence of the size of the model system of the TM hydrides within the ONIOM scheme on the performance of DFT functionals.

For practical computations of bulk TM species, theoretical methods are often used in combination with implicit solvation models. In this study, the solvent effect was included implicitly with three popular solvation models: SMD, COSMO, and C-PCM. The choice of cavity model, which defines the shape and size of the cavity occupied by a solute species in the solvent, has shown to have impact on the prediction of pK_a of organic acids using DFT funcionals.[51–56] For instance, in the studies of the aqueous solvation free energies of 10 organic species calculated with seven cavities (UAKS, UAHF, UAHF, Bondi, Pauling, UA0, and UFF) using the B3LYP/6-31+G(d) method with the C-PCM solvation model, UAKS and UAHF resulted in the lowest MADs relative to experiment in comparison to the other considered cavity models. Additionally, the Pauling cavity performed best for anion species but gave the worst agreement with experiment for neutral species.[51] Sadlej-Sosnowska studied the solvation free energies and pK_a values of nine neutral organic compounds and their anions with DPCM, C-

PCM, and IEFPCM solvation models. The deviations of solvation free energy with UA0 cavity from the experimental data for anions are greater than that with UAHF cavity by 2-3 kcal mol^{-1}.[52] A systematic study of solvation free energy and pK_a values of monoprotic, diprotic, and triprotic acids based on DFT(B3LYP, PBE, BVP86, and M05-2X)/aug-cc-pVTZ methods combined with the C-PCM and SMD solvation models showed that the Pauling cavity in combination with M05-2X showed the best performance among the UFF, UAKS, Pauling, and Klamt cavity models.[56] Despite the evidence that accurate prediction of pK_a values has been shown to be related to the choice of cavity models, the performance of DFT functionals has not been assessed in terms of the choice of cavity models yet for TM-containing species. Therefore, the degree to which the several considered cavity models including UA0, UAKS, Pauling, Bondi, UFF, Klamt, and Coulomb-SMD can accurately calculate the pK_a values of TM hydrides was evaluated.

5.2 Theoretical Methods

5.2.1 Geometry Optimization and Frequency Calculations.

All calculations were performed using the GAUSSIAN 09 software package.[57] For all considered TM hydrides, complete geometry optimization and frequency calculations (vibrational ZPE scaled by 0.9890)[58] were performed using B3LYP/cc-pVTZ in both the gas phase and with acetonitrile solvent (all considered TM hydrides showed no significant structural changes for gas and solvation phases, and the structures are included in Supporting Information). Acetonitrile solvent systems were treated using C-PCM continuum solvation model.[59–62] All stationary points were verified to be true minima by no imaginary frequencies. The thermochemical corrections from B3LYP/cc-pVTZ frequency calculations were added to the single point energies to obtain gas phase and solvation free energies at 298 K.

5.2.2 Single-Point Calculations.

Subsequently, single-point energies were performed with the two-layer ONIOM method.[34–42] The total energy of two-layer ONIOM is defined as

$$E(ONIOM) = E(low, real) + E(high, model) - E(low, model) \quad (5.1)$$

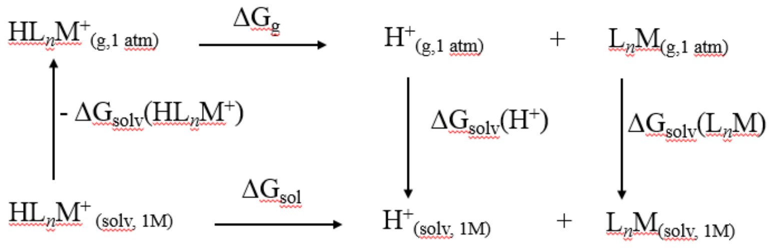
where E(low, real) is the energy of the real system with a low-level method, E(high, model) is the energy of the model system with a high-level method, and E(low, model) is the energy of the model system with a low-level method. Since the selection of the separate spaces is arbitrary in ONIOM calculations, the effect of expanding the boundary of the model system is desirable to investigate (Fig. 1). Thus, we successively expanded the model region from including the metal atom and 4 phosphorous atoms, then all the atoms within the rings and finally to include all atoms except for the very outside methyl group as shown in Figures 5.1a, 5.1b and 5.1c, respectively. The ensuing results from these expansions are defined as ONIOM-1, ONIOM-2, and ONIOM-3, respectively. The ONIOM-1 scheme is used for the majority of the DFT calculations in this study due to computational cost and the inclusion of the atoms that contain the largest portion of the electron density.

To evaluate the impact of the selection of DFT methods on the accuracy of the ONIOM scheme in predicting pK_as of the TM hydrides, the following DFT methods were utilized (summarized in Table 5.1): BLYP[63,64], PBE[65], and B97-D[66] (GGA), M06L[16], BB95[67], and TPSS[68] (M-GGA), PBE0[65,69,70], B3LYP[63,64,71], and B3P86[64,72] (H-GGA), M06[16], M05-2X[14], M06-2X[16], M06HF[17,73] (HM-GGA), and B2PLYP[74] (DH-GGA). Additionally, Grimme's empirical dispersion correction (D3)[75] was added to several DFT functionals selected from GGA, M-GGA, H-GGA, and HM-GGA functionals, to evaluate the effect of a dispersion correction on the accuracy of predictions of pK_as of the TM hydrides. For the PBE0 functional, the impact of the percentage of the Hartree-Fock (HF) exchange was evaluated by varying percentages from 0% to 80%.

The basis sets listed in Table 5.2 were considered in this study to assess the influence of basis set on the calculated pK_a values. The relativistic effective core potential (ECP) and valence double-ξ basis set of Hay and Wadt (LANL2DZ)[76–78] and the Stuttgart/Dresden (SDD)[79–81] relativistic ECP and valence triple-ξ basis set (For these two basis sets, 10-core-electrons were adopted for Ni, 28-core-electrons for Pd, and 60-core-electrons for Pt) were used paired with DFT functionals as the low-level method. The high-level method using DFT functionals was paired with several of Dunning's all electron basis sets (cc-pVDZ, cc-pVTZ, aug-cc-pVDZ, and aug-cc-pVTZ).[82–85] For Ni species, Dunning's correlation consistent basis sets with the one-particle Douglas-Kroll-Hess Hamiltonian for scalar relativistic effects were applied (e.g., aug-cc-pVTZ-DK).[86] For Pd and Pt species, the small-core relativistic pseudopotential basis sets (e.g., aug-cc-pVTZ-PP) were used.[84,85] In the following sections, the terms DK for Ni and PP for Pd and Pt are dropped for clarity from the selected basis sets.

Three implicit solvation models, C-PCM,[59,60,62] SMD,[87] and COSMO,[61] were employed to include solvent effects in single point calculations. The effect of the cavity choice on the prediction of the pK_a of the TM hydrides was also evaluated in this study. The UA0, UAKS, Pauling, Bondi and default cavities (UFF for C-PCM, Klamt for COSMO, and Coulomb-SMD for SMD) were applied.

Scheme 1. Direct Thermodynamic Cycle



The direct thermodynamic scheme for calculating p*K*$_a$s of unknown acids shown in Scheme 1 has been used mainly due to its simplicity[88–99] and was used in this study with the value -4.39 kcal mol^{-1} for the gas phase free energy of a proton, ΔG$_{gas}$(H$^+$), derived using the Sackur-Tetrode equation.[100] For the value of the experimental solvation phase free energy of the proton in acetonitrile, ΔG$_{solv}$(H$^+$), from 1 atm to 1 M standard state, -260.2 kcal mol^{-1} has been indicated as the best result by several studies[101–107] and therefore was also used in this study. Thus the solvation free energy (ΔG_{sol}) can be calculated using the following equations (Eq. 5.2-5.4).

$$\Delta G_{sol} = \Delta G_g + \Delta\Delta G_s \quad (5.2)$$

where

$$\Delta G_g = G_g(L_nM) + G_g(H^+) - G_g(HL_nM^+) \quad (5.3)$$

$$\Delta\Delta G_s = \Delta G_s(L_nM) + \Delta G_s(H^+) - \Delta G_s(HL_nM^+) \quad (5.4)$$

The p*K*$_a$ values related to free energies of solvation were calculated as

$$pK_a = \frac{\Delta G_{solv}}{2.303RT} \quad (5.5)$$

All the calculated gas phase free energies in units atm were converted to molar units and the solvation phase free energies were calculated using [(E_{soln} + G_{nes}) − E_{gas}], as defined in the parametrization of continuum solvent models.[108,109] An error of 1.36 kcal mol^{-1} in ΔG_{solv} produces a deviation of 1 pK_a unit. Ho and Coote reported that a direct thermodynamic cycle can be expected to depart from experiment by 3.5 pK_a units.[110]

5.3 Results and Discussions

The considered molecules are grouped based on central TM atoms (Ni, Pd, and Pt) and the ligands (depe, depp, and PNP) in order to evaluate the impacts of the selected DFT functionals, basis sets, cavities, solvation models, and the expansion of size of the high-layer region on the calculated pK_as of TM hydrides. All the error analyses are based on mean absolute deviations (MADs) with respect to experimental data.

5.3.1 Utility of DFT Functionals in the Real System.

Table 5.3 summarizes the combination of DFT functionals within the ONIOM schemes that were designed to assess the influence of the fourteen considered DFT functionals paired with the LANL2DZ basis set in the description of the real system of the TM hydrides (Figure 5.1). The model systems (Figure 5.1) were treated with four DFT functionals (PBE, M06L, B3LYP, and M06), including GGA, M-GGA, H-GGA, and HM-GGA, combined with the aug-cc-pVTZ basis set. The MADs (an average of results from the four selected high-layer methods) of each low-level method for the calculating pK_a values of the TM hydrides relative to experiment with C-PCM, COSMO, and SMD are reported in Figure 5.2. Among the fourteen density functionals considered in this section, the B97-D performed best with MADs of 5.5, 2.7, and 2.3 pK_a units for C-PCM, COSMO, and SMD respectively, followed by B3LYP (6.3, 3.4, 2.9 pK_a units), and M06-L (7.2, 4.5, 3.8 pK_a units). Except for the three mentioned outperformed functionals, all other considered GGA, M-GGA and H-GGA functionals performed similarly with MAD values of about 7.9, 5.0, and 4.3 pK_a units for C-PCM, COSMO, and SMD, respectively. The functional with the poorest performance is M06-2X with MADs of 10.5, 7.7, and 7.1 pK_a units for C-PCM, COSMO, and SMD, respectively.

In some cases, including HF exchange is a disadvantage for description of the TM hydrides, as the performance of M06-L with zero percent HF exchange is better than other considered functionals from the Minnesota functionals that include HF exchange between all three solvation models. Additionally, DFT functionals with higher percent HF exchange do not always provide results closer to the experimental data. For instance, M06-2X (54% HF exchange), which includes double the percent HF exchange as M06 (27% HF exchange), resulted in the larger MADs as compared to M06 for all three solvation models. However, with the exclusion of M06-2X, the considered non-local Minnesota functionals yielded lower MADs while including a higher percentage of HF exchange for C-PCM and COSMO.

There are also some other cases showing that including HF exchange has no influence on the performance of DFT functionals. For instance, for the pairs of functionals that differ only with respect to the inclusion of HF exchange (BLYP and B3LYP, PBE, and PBE0) the MADs yielded by PBE (0% HF exchange) were similar to MADs from PBE0 (25% HF exchange) for all three solvation models, while B3LYP, also with 25% HF exchange, resulted in lower MADs than the local BLYP functional for all three solvation models. All non-local Minnesota functionals (M06, M05-2X, and M06-HF), except for M06-2X, yielded similar MADs of about

5.5 pK_a units with the SMD. Therefore, it is difficult to determine the effect of HF exchange on the performance of DFT functionals to describe the TM hydride pK_as. Among the three selected solvation models, SMD provided the best comparison with experimental pK_a data while C-PCM yielded the highest MADs for all of the fourteen considered DFT functionals.

The performance of the different types of DFT functionals at modeling the pK_as is shown in Figure 5.3. The GGA functionals produced similar MADs from experiment (7.1, 4.2, and 3.6 pK_a units) as H-GGA functionals (7.3, 4.4, and 3.8 pK_a units) for C-PCM, COSMO, and SMD, respectively, and performed better than all the other types of functionals regardless of solvation method. In contrast, DH-GGAs performed the worst with MADs of 9.8, 6.9, and 6.0 pK_a units for C-PCM, COSMO, and SMD, respectively, which indicates that the addition of a fraction of the MP2 correlation energy is disadvantage for description of TM hydride pK_as. Compared against HM-GGAs, M-GGA functionals without HF exchange yielded lower MADs for all three solvation models. Therefore, HF exchange is not necessary for accurate description the TM hydrides. COSMO and SMD performed similarly and resulted in MADs significantly lower than that from C-PCM.

The comparison of different types of DFT functionals is considered with respect to central TM atoms (Table 5.4) and ligand systems (Table 5.5) with the three solvation models. For the Ni species, as the functionals become more complex (from first rung to fifth rung of Jacob's ladder), larger MADs were yielded, with exception of DH-GGAs, for all three solvation models. For Pd and Pt species, H-GGAs yielded the lowest MADs in comparison to other types of functionals while DH-GGAs always performed the worst with all three solvation models. Moving down the periodic table from Ni to Pt, the MADs of non-local exchange functionals (H-GGA, HM-GGA, and DH-GGA) decrease, which indicates that non-local exchange in functionals can describe TM hydrides with heavier central TM atoms better than those with lighter central TM atoms. Considering the overall MADs of different types of functionals, the increase in MADs upon inclusion of HF exchange is more significant for M-GGA functionals (HM-GGA) than it is for the GGA functionals (H-GGA). As shown in Figure 5.1, the size of the considered ligands increases in the order of depe, depp, and PNP. Similar performance was found for each type of functional between all three solvation models as the size of the ligand increased.

5.3.2 Utility of DFT Functionals in the Model System.

The utility of the DFT functionals in describing real system of ONIOM scheme, evaluated in las section, is not the only factor impacts the accuracy of the calculated pK_a values. The utility of the DFT functionals for the model system also plays an important role in calculating the pK_a values; therefore, the section is focus on the utility of the DFT functionals for model system. Table 5.6 summarizes the combination of DFT functionals as ONIOM schemes designed to measure the influence of the fourteen considered DFT functionals combined with aug-cc-pVTZ basis set in the description of model systems of the TM hydrides (Figure 5.1). The real systems (Figure 5.1) were treated with three DFT functionals (B97-D, M06L, and B3LYP) paired with the LANL2DZ basis set, which were selected based on their better performance as low-level methods shown in Section 5.3.1. The MADs (an average of results from the three selected real system methods) for each high-level method are based upon deviations of the calculated pK_a values of the TM hydrides from experimental data for each functional using the C-PCM, COSMO, and SMD are reported in Figure 5.4. For C-PCM and COSMO, B3LYP, M05-2X, and M06-HF are the three functionals that performed better than the other considered functionals, where B3LYP and M06-HF resulted in the lowest MADs with C-PCM and COSMO, respectively. For SMD, B97-D, TPSS, and M05-2X yielded the same MAD value of 2.1 pK_a units. Therefore, unlike the consistency for DFT functionals that were found to perform best in describing the real systems among different solvation models, the performance of DFT functionals in describing the model system were highly related to the selection of the solvation model.

The most accurate pK_a values were yielded by different DFT functionals with different solvation models (MAD of 2.0 pK_a units by B3LYP with C-PCM, 1.9 pK_a units by M06-HF with COSMO, and 2.1 pK_a units by B97-D, TPSS, and M05-2X with SMD). PBE provided the worst comparison with experimental data with MADs of 6.7, 6.5 and 5.5 pK_a units for C-PCM, COSMO, and SMD, respectively. BB95 and M06 also performed considerably worse than other considered functionals (except PBE), which resulted in the same MADs of 6.2 pK_a units for C-PCM and 5.0 pK_a units for SMD, and similar MADs of about 5.6 pK_a units for COSMO.

There are also conflicts regarding the conclusion on the impact of including HF exchange on the performance of DFT functionals to describe the model system of the TM hydrides. For instance, in comparing the MADs from PBE0 and B3LYP with 25% HF exchange, the MADs yielded by the local PBE and BLYP functionals were almost doubled. However, the data for the Minnesota functionals suggested that the additional parameter of HF exchange did not necessarily improve the performance of DFT functionals, since larger MADs were observed for M06 (27% HF exchange) and M06-2X (54% HF exchange) than for local M06-L with all three solvation models. DFT functionals with SMD always yielded lower MADs than C-PCM and COSMO, except for M06-HF, where COSMO resulted in the lowest MAD of 1.9 pK_a units.

The performance of different types of DFT functionals is shown with the three solvation models in Figure 5.5. The present data showed that H-GGA functionals provided the results, which were most accurate relative to the experimental data with MADs of 3.2, 2.8, and 2.5 pK_a units, while the DH-GGA functionals resulted in the highest MADs of 5.2, 4.6, and 4.0 pK_a units for C-PCM, COSMO, and SMD, respectively. The poor performance of DH-GGAs infers that the addition of a fraction of the MP2 correlation energy should not be considered for the accurate description of the model system of TM hydrides. For all three solvation models, the local exchange types of functionals (GGA and M-GGA) yielded larger MADs than their respective non-local exchange types of functionals (H-GGA and HM-GGA), which indicates that inclusion of HF exchange is necessary to describe the model system of TM hydrides more appropriately. This improvement of the DFT functionals by including HF exchange in SMD was less obvious then that in C-PCM and COSMO.

The different types of DFT functionals were compared with respect to central TM atoms in Table 5.7 to assess if their ability to describe the model system was determined by their performance on the description of metal center. Based on the overall MADs shown in Table 5.7, the MADs for all types of functionals decrease from lighter to heavier metal with all three solvation models. For Ni species, M-GGAs yielded the lowest MADs of 3.7, 3.4, and 2.6 pK_a units with C-PCM, COSMO, and SMD, respectively. The H-GGA functionals performed the best for C-PCM and COSMO with MADs of 3.2 and 2.9 pK_a units, respectively. The GGA, M-GGA, and H-GGA functionals resulted in similar MADs of about 2.5 pK_a units with SMD for Pd species. For Pt species, the HM-GGA functionals produced comparable MADs of about 1.8 pK_a units for COSMO and SMD that were lower than for other types of functionals. The DH-GGA functional resulted in the largest MADs for all considered metal species with all three solvation models. The model system is described better by H-GGA functionals than GGA functionals. Ergo, following the same conclusion based on the overall performance of different types of functionals, the reduction in MADs for H-GGA functionals from GGA functionals is more significant for TM hydrides with lighter central TM atoms than for those with heavier central TM atoms.

5.3.3 Impact of HF Exchange on the Accuracy of DFT Functionals.

Although there was no systematic trend found between the percentage of HF exchange and the accuracy of Minnesota functionals for the prediction of the pK_as of TM hydrides, H-GGA and HM-GGA functionals showed improvement in predicting more accurate pK_a values than the local exchange GGA and M-GGA types of functionals. Therefore, some light might be still shed on the impact of HF exchange by investigating if the performance of other functionals can be systematically improved as a function of the percentage of HF exchange. PBE0 with 25% HF exchange did improve the accuracy of the local PBE without HF exchange. Additionally, PBE includes no empirical parameters that may affect the performance of DFT.

Therefore, using the PBE functional to examine the impact of HF exchange on the calculation of pK_as for TM hydrides with DFT functionals can avoid interference from other empirical parameters. The percentage of HF exchange considered was from 0 to 80% in intervals of 5%. The MADs with respect to different central TM atoms and different size of ligands of TM hydrides were taken into account with the ONIOM(PBE/aug-cc-pVTZ:B97-D/LANL2DZ) scheme and SMD. B97-D was selected due to its most comparable results to the experimental data and the SMD solvation model was used since it resulted in lower MADs than either C-PCM or COSMO. As shown in Figure 5.6, for the TM hydrides with all considered ligands (depe, depp, and PNP), 50% HF exchange was preferred. The MAD curves were smooth with the exception of the point with 35% HF exchange for TM hydrides with depp and PNP ligands. For Ni species, the minima all laid at 50%. The MAD curves of Pd and Pt species were significantly flatter than those for the Ni species. For the Pd species, all values between 40 and 80% yielded a roughly comparable performance with the greatest deviation being 0.6 pK_a units. The Pt species had a "double-well" shaped curve with two comparable minima at 30% and 40%. For the overall MADs of the considered species, the minimum can be found at 50% HF exchange.

5.3.4 Impact of Adding the Grimme's Empirical Dispersion Correction on the Accuracy of DFT Functionals.

The results in Section 3.1 and 3.2 indicated that the dispersion-corrected functional, B97-D, was the best choice to describe both the real and model systems in the QM/QM scheme for TM hydrides with the SMD due to having the lowest MADs in reference to experimental pK_a values. Therefore, it is of interest to evaluate the influence of adding the Grimme's empirical (D3) dispersion correction on both low- and high-level methods in ONIOM(DFT/aug-cc-pVTZ:B97-D/LANL2DZ) schemes. B97-D/LANL2DZ method was applied in this section as the low-level method due to its superior performance relative to other methods as shown in Table 5.3 with the SMD. The non-dispersion-corrected DFT functionals were selected from four types of DFT functionals (BLYP and PBE from the GGAs, M06L and TPSS from the M-GGAs, PBE0 and B3LYP from the H-GGAs, and M05-2X and M06-2X from the HM-GGAs) to describe the model system of the TM hydrides. The results are shown in Figure 5.7. It was found that most of the DFT functionals with the dispersion correction gave more accurate results relative to experimental pK_a values with the exception of B3LYP and M05-2X. In order to determine whether the impact of adding the dispersion correction will be consistent on different species, the overall performance of the DFT functionals with and without the dispersion correction was considered with respect to different central TM atoms as well as different ligand sizes in the TM hydrides (Figure 5.8). Although DFT-D3 methods resulted in lower MADs for all considered species, the improvement by adding dispersion correction varies. The reductions in the MADs were more significant for the lighter central TM atoms and for TM hydrides with larger sized ligands than for heavier central TM atoms and for TM hydrides with smaller sized ligands. The comparison of different types of DFT functionals with and without the dispersion correction is shown in Table 5.8. For the non-local exchange functionals, H-GGAs and HM-GGAs, the addition of Grimme's dispersion correction reduced the MADs more significantly than for local exchange functionals, GGAs and M-GGAs.

5.3.5 Impact of the Choice of Basis Set.

It is well-known that the chosen basis set will also affect the accuracy of calculated properties in addition to the selected DFT functional. Therefore, the influence of the basis set on the accuracy of calculated pK_as was assessed with two double-ξ and two triple-ξ quality correlation consistent basis sets with DFT functionals as the high-level method, and LANL2DZ and SDD with DFT functionals as the low-level method. The aug-cc-pVTZ basis set was utilized for high-level methods when comparing the performance of LANL2DZ and SDD basis sets for low-level methods. For the comparison of the considered correlation consistent basis sets for high-level methods, LANL2DZ was applied for low-level methods.

The DFT functionals for high-level method included B97-D, TPSS, B3LYP, and M05-2X that yielded similar MADs of about 2.3 pK_a units and performed better than the other considered

DFT functionals with respect to experimental pK_a values. B97-D was applied for low-level methods since it yielded a lower MAD than other considered functionals as shown in Section 3.1. All calculations in this section used SMD based on the solvation model's performance in previous sections.

For the high-level methods, the dependence of the four functionals, selected to describe the model system of the TM hydrides, upon the quality of correlation consistent basis set is shown in Table 5.9. Only a small reduction in MAD of 0.2 pK_a units was found when the basis set quality was increased from aug-cc-pVDZ to aug-cc-pVTZ for B97-D and B3LYP. In contrast, the MADs of B97-D and B3LYP showed considerably reduction of MADs (more than 0.5 pK_a units) upon improving the basis set from cc-pVDZ to cc-pVTZ. TPSS has a different dependence on the quality of basis set as cc-pVDZ (1.4 pK_a units) unexpectedly provided the lower MAD than cc-pVTZ (3.3 pK_a units).

As shown in Figure 5.9, the accuracy of the basis set displayed a dependence on the central TM atoms of the TM hydrides, where cc-pVDZ and cc-pVTZ had similar performance for Ni and Pt species while cc-pVDZ performed better than cc-pVTZ for Pd. Similarly, aug-cc-pVDZ outperformed aug-cc-pVTZ for Pd species but yielded higher MADs than aug-cc-pVTZ for Pt species. In contrast, the accuracy of the basis sets were not affected by the ligand sizes of the TM hydrides, as both double-ξ and triple-ξ basis sets, with or without the diffuse functions, consistently resulted in similar MADs. Both of the considered double- and triple-ξ correlation consistent basis sets provided a more accurate description of the model system of the TM hydrides by including diffuse functions, except for the Ni species.

For the low-level methods, SDD performed better than LANL2DZ with respect to central TM atoms and ligand sizes of the TM hydrides, except for Ni species (Figure 5.9). The performance of both LANL2DZ and SDD improved as the central TM atoms of TM hydrides becomes heavier.

5.3.6 Impact of Cavity Models of Implicit Solvation Models.

The calculated pK_a values were also compared to the experimental data from the viewpoint of the cavities used in computing the C-PCM, COSMO, and SMD reaction fields with the ONIOM(B3LYP/aug-cc-pVTZ:B97D/LANL2DZ) scheme. Five cavity models, Pauling, Bondi, UA0, UAKS, and the default cavity for each solvation model within the Gaussian 09 package (UFF for C-PCM, Klamt for COSMO, and SMD-Coulomb for SMD) were applied to determine the effect of the atomic radii used to build a cavity in the solvent (acetonitrile) on the predicted pK_a values of the TM hydrides. As shown in Table 5.10, for C-PCM, the Pauling cavity generated the lowest MAD of 0.9 pK_a units in comparison to the other cavity models, while UA0 resulted in the largest MAD of 3.3 pK_a units. The default cavity of C-PCM gave comparable accuracy with the Pauling cavity for the TM hydrides with Pd and Pt as central atoms while much larger MADs were yielded by the default cavity then by the Pauling cavity for the TM hydrides for Ni(depe)$_2$ and Ni(depp)$_2$. The Bondi and UAKS cavities resulted in the same MADs of 1.1 pK_a units and performed slightly worse than Pauling cavity but showed reduction of MADs of 2.2 pK_a units relative to UA0. For both COSMO and SMD, the calculated pK_a values of TM hydrides with Ni and Pd as central atoms and depp and PNP ligands from the default cavity depart slightly less than those obtained using the Pauling, Bondi, and UAKS cavities. Among the five cavities used for COSMO and SMD, the default cavity only performed better than the UA0 cavity for TM hydrides with Pt as the central atom. Similar to C-PCM, COSMO, and SMD also provided the poorest performance with UA0 cavity.

5.3.7 Impact of the Expansion of the Size of Model system.

In order to test the influence of size of the model system on the performance of DFT functionals to predict pK_a values of TM (Pd and Pt) hydrides, four functionals, B97-D, TPSS, B3LYP, and M05-2X were used due to their closer results to experimental data when used within the high-level methods. As shown in Table 5.11, among the four functionals, only B97-D showed improvement when the size of model system was expanded from ONIOM-1 to ONIOM-3, while the MADs of the other three functionals increased. The largest deviation of the MADs between the four functionals are 0.5, 1.8, and 2.7 pK_a units for ONIOM-1, ONIOM2, and ONIOM-3, respectively. The accuracy of the calculated pK_a values of the TM hydrides showed a larger dependence on the selection of DFT functionals when a larger sized model system was utilized.

5.4 Conclusion

In this study, the performance of fourteen density functionals has been investigated with respect to their accuracy when used within the high- and low- level methods of the QM/QM ONIOM scheme in predicting pK_a values of nine Group 10 TM hydrides with the C-PCM, COSMO, and SMD solvation models. The calculated pK_a values were compared to the experimental data.

For the performance of the functionals in describing the real system of the TM hydrides, B97-D, regardless of solvation model, yielded the lowest MADs while B2PLYP performed the worst at reproducing experimental pK_a values. For different types of DFT functionals, GGA and

H-GGA functionals provided a more accurate description of the real system than M-GGA, HM-GGA and DH-GGA functionals. Inclusion of HF exchange in the low-level method did not lead to a reduction in the deviation relative to experiment.

When considering the performance of the functionals for describing the model system of the TM hydrides, the selection of solvation models had a noticeable influence. The results from SMD were always better than those from COSMO and C-PCM using the same functional. B3LYP and M06-HF resulted in lowest MADs with C-PCM and COSMO, while B97-D, TPSS, and M05-2X yielded the same MADs and performed better than other functionals for SMD. Overall, the functionals with HF exchange produced results closest to experimental data in comparison with other types of functionals.

The impact of inclusion of HF exchange and Grimme's empirical dispersion correction on the accuracy of the pK_as was also investigated. The minimum MAD can be found at 50% HF exchange using PBE within the high-level method. The improvement obtained by adding the dispersion correction varies. Comparing with the reduction in MADs based on metal size and ligand size on the inclusion of the dispersion correction, the TM hydrides with lighter central TM atoms and larger sized ligands had their MADs decrease more significantly.

The effect of the selection of the basis set was also examined. Among four high-level methods within the ONIOM schemes, the B97-D and B3LYP functionals showed a reduction of their MADs upon improving the quality of the correlation consistent basis set from double-ξ to triple-ξ.

The performance of five selected cavity models (Pauling, Bondi, UA0, UAKS, and the default cavity for each solvation model within Gaussian 09 package (UFF for C-PCM, Klamt for

COSMO, and SMD-Coulomb for SMD)) were investigated for the three different implicit solvation models, C-PCM, COSMO, and SMD. UA0 always resulted in the highest MADs for C-PCM, COSMO, and SMD relative to experiment. C-PCM resulted in lowest MADs when using the Pauling cavity instead of the default cavity, while for COSMO and SMD, the default cavities are suggested to be applied.

The expansion of the size of model system had positive impact on the accuracy of pK_a calculations of the TM hydrides only when using B97-D within the high-level method.

5.5 References

(1) Hyla-Kryspin, I.; Grimme, S. *Organometallics* **2004**, *23,* 5581.

(2) Song, P.; Guan, W.; Yao, C.; Su, Z. M.; Wu, Z. J.; Feng, J. D.; Yan, L. K. *Acc.* **2007**, *117*, 407.

(3) Gutsev, G. L.; Mochena, M. D.; Jena, P.; Bauschlicher, C. W., Jr;. Partridge, H., III. *J. Chem. Phys.* **2004**, *121*, 6785.

(4) Yao, C.; Guan, W.; Song, P.; Su, Z. M.; Feng, J. D.; Yan, L. K.; Wu, Z. J. *Theor. Chem. Acc.* **2007**, *117*, 115.

(5) Quintal, M. M.; Karton, A.; Iron, M. A.; Boese, A. D.; Martin, J. M. L. *J. Phys. Chem. A* **2006**, *110*, 709.

(6) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(7) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(8) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757.

(9) Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I. *J. Chem. Phys.* **2005**, *123*, 062201.

(10) Mori-Sánchez, P.; Cohen, A. J. *J. Chem. Phys.* **2006**, *125*, 201102.

(11) Perdew, J. P.; Zunger, A. *Phys. Rev. B.* **1981**, *23*, 5048.

(12) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(13) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8,* 4398.

(14) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2,* 364–382.

(15) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2005**, *123*, 161103.

(16) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(17) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem.* **2006**, *110*, 5121–5129.

(18) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 11127-11143.

(19) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4388-4403.

(20) Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 44103-44130.

(21) Riley, K. E.; Merz, K. M., Jr. *J. Phys. Chem.* **2007**, *111*, 6044-6053.

(22) Tekarli, S. M.; Drummond, M. L.; Williams, T. G.; Cundari, T. R.; Wilson, A. K. *J. Phys. Chem. A* **2009**, *113*, 8607-8614.

(23) Laury, M. L.; Wilson, A. K. *J. Chem. Theory Comput.* **2013**, *9*, 3939-3946.

(24) Jiang, W.; Laury, M. L.; Powell, M.; Wilson, A. K. *J. Chem. Theory Comput.* **2012**, *8*, 4102-4111.

(25) Wang, W. H.; Muckerman, J. T.; Fujita, E.; Himeda, Y. *ACS catal.* **2013**, *3*, 856-960.

(26) Stewart, M. P.; Ho, M. H.; Wiese, S.; Lindstrom, M. L.; Thogerson, C. E.; Raugei, S.; Bullock, R. M.; Helm, M. L. *J. Am. Chem. Soc.* **2013**, *135*, 6033-6046.

(27) Liu, T.; Dubois, D. L.; Bullock, R. M. *Nat. Chem.* **2013**, *5*, 228.

(28) Luca, O. R.; Blakemore, J. D.; Konezny, S. J.; Praetorius, J. M.; Schmeier, T. J.; Hunsinger, G. B.; Batista, V. S.; Brudvig, G. W.; Hazari, N.; Crabtree, R. H. *Inorg. Chem.* **2012**, *51*, 8704-8709.

(29) Espino, G.; Caballero, A.; Manzano, B. R.; Santos, L.; Pérez-Manrique,. M.; Moreno, M.; Jalón, F. A. *Organometallics* **2012**, *31*, 3087-3100.

(30) Crabtree, R. H. *Organomet. Chem. Met. Wiley New York,* **2001**.

(31) Bäckvall, J. E. *J. Organomet. Chem.* **2002**, *625*, 105-111.

(32) Hoskin, A. J.; Stephan, D. W. *Coord. Chem. Rev.* **2002**, *233*, 107-129.

(33) Andrews, L. *Chem. Soc. Rev.* **2004**, *33*, 123-132.

(34) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. THEOCHEM* **1999**, *462*, 1-21.

(35) Hopkins, B. W.; Tschumper, G. S. *J. Comput. Chem.* **2003**, *24*, 1563.

(36) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959.

(37) Karadakov, P.B.; Morokuma, K. *Chem. Phys. Lett.* **2000**, *317*, 589-596.

(38) Rega, N.; Iyengar, S. S.; Voth, G. A.; Schlegel, H. B.; Vreven, T.; Frisch, M. J. *J. Phys. Chem. B* **2004**, *108*, 4210-4220.

(39) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.

(40) Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62.

(41) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432.

(42) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125-132.

(43) Qi, X.; liu, L.; Fu, Y.; Guo, Q. *Organometallics* **2006**, *25*, 5879-5886.

(44) Berning, D. E.; Noll, B. C.; DuBois, D. L. *J. Am. Chem. Soc.* **1999**, *121*, 11432-11447.

(45) Curtis, C. J.; Miedaner, A.; Ellis, W. W.; DuBois, D. L. *J. Am. Chem. Soc.* **2002**, *124*, 1918-1925.

(46) Curtis, C. J.; Miedaner, A.; Raebiger, J. W.; DuBois, D. L. *Organometallics* **2004**, *23*, 511-516.

(47) Raebiger, J. W.; Miedaner, A.; Curtis, C.; J.; Miller, S. M.; Anderson, O. P.; DuBois, D. L. *J. Am. Chem. Soc.* **2004**, *126*, 5502-5514.

(48) Wang, N. X.; Wilson, A. K. *J. Phys. Chem. A* **2003**, *107*, 6720-6724.

(49) Wang, N. X.; Wilson, A. K. *J. Phys. Chem. A* **2005**, *109*, 7187-7196.

(50) Yockel, S.; Mintz, B.; Wilson, A. K. *J. Chem. Phys.* **2004**, *121*, 60.

(51) Takano, Y.; Nouk, K. N. *J. Chem. Theory Comput.* **2005**, *1*, 70-77.

(52) Sadlej-Sosnowska, N. *Theor. Chem. Acc.* **2007**, *118*, 281-293.

(53) Krol, M;. Wrona, M.; Page, C. S.; Bates, P. A. *J. Chem. Theory Comput.* **2006**, *2*, 1520.

(54) Lide, D. R. *CRC Handb. Chem. Physics, CRC, 86th edn,* **2005**.

(55) Fernández, M. I.; Canel, M.; Carcia, M. V.; Santaballa, J. A. *Chem. Phys. Lett.* **2010**, *490*, 159-164.

(56) Lee, T. B.; McKee, M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10258-10269.

(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.;Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; , and Fox, D. J., *Gaussian 09, Revision B.01*; Gaussian, Inc., Wallingford, CT, 2010.

(58) DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. *Mol. Phys* **2009**, *107*, 1107.

(59) Barone, V.; Cossi, M. *J. Phys. Chem.* **1998**, *102*, 1995–2001.

(60) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. comput. Chem.* **2003**, *24*, 669–681.

(61) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, *5*, 799–805.

(62) Klamt, A. *J. Chem. Phys.* **1995**, *103*, 9312–9320.

(63) Lee, C.; Hill, C.; Carolina, N. *Phys. Rev. B Condens. Matter. Phyes.* **1988**, *37*, 785–789.

(64) Becke, A. D. *Phys. Rev. A.* **1988**, *38*, 3098–3100.

(65) Perdew, J.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(66) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(67) Becke, A. D. **1996**, *104*, 1040–1046.

(68) Tao, J.; Perdew, J.; Staroverov, V.; Scuseria, G. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(69) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.

(70) Adamo, C.; Cossi, M.; Barone, V. *J Mol Struc-Theochem*. **1999**, *493*, 145–157.

(71) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(72) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822-8824.

(73) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

(74) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(75) Crimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(76) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270–283.

(77) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 284–298.

(78) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299–310.

(79) Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chem. Acc.* **1990**, *77*, 123–141.

(80) Andrae, D.; Häußermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866–872.

(81) Igel-Mann, G.; Stoll, H.; Preuss, H. *Mol. Phys.* **1988**, *65*, 1321–1328.

(82) Dunning, Jr., T. H. *J. Chem. Phys.* **1989**, *90*, 1007-1023.

(83) Balabanov, N.B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, 064107.

(84) Peterson, K.A.; Figuregen, D.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2007**, *126*, 124101.

(85) Figuregen, D.; Peterson, K. A.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2009**, *130*, 164108.

(86) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89.

(87) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(88) Kallies, B.; Mitzner, R. *J. Phys. Chem. B* **1997**, *101*, 2959-2967.

(89) Shapley, W. A.; Backsay, G. B.; Warr, G. G. *J. Phys. Chem. B* **1998**, *102*, 1938-1944.

(90) Topol, I. A.; Tawa, G. J.; Caldwell, R. A.; Eissenstat M. A.; Burt, S. K. *J. Phys. Chem. A* **2000**, *104*, 9619-9624.

(91) Liptak M. D., Shields, G. C. *J. Am. Chem. Soc.* **2001**, *123*, 7314-7319.

(92) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413-7422.

(93) Klicic, J. J.; Friesner R. A.; Liu S. Y.;Guida, W. C. *J. Phys. Chem. A* **2002**, *106*, 1327-1335.

(94) Magill, A. M.; Cavell, K. J.; Yates, B. F. *J. Am. Chem. Soc.* **2004**, *126*, 8717-8724.

(95) Murlowska, K.; Sadlej-Sosnowska, N. *J. Phys. Chem. A* **2005**, *109*, 5590-5595.

(96) Pliego, Jr., J. R. *Chem. Phys. Lett*. **2003**, *367*, 145-149 .

(97) Lu, H.; Chen, X.; Zhan, C. G. *J. Phys. Chem. B* **2007**, *111*, 10599-10605.

(98) Zimmermann, M. D.; Tossell, J. A. *J. Phys. Chem. A* **2009**, *113*, 5105-5111.

(99) Riojas, A. G.; Wilson, A. K. *J. Chem. Theory Comput.* **2014**, *10*, 1500-1510.

(100) McQuarrie, D. A. *Stat. Mech. Stat. Mech. Univ. Sci. Books Sausalito, CA*, **2000**.

(101) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, Jr., T. R. *J. Phys. Chem. A* **1998**, *102*, 7787-7794.

(102) Zhan, C. -G.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534-11540.

(103) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **1999**, *103*, 11194-11199.

(104) da Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. *J. Phys. Chem. A* **2000**, *104*, 2402-2409.

(105) Pliego, J. R. *Chem. Phys. Lett.* **2003**, *381*, 244-245.

(106) Palascak, M. W.; Shields, G. C. *J. Phys. Chem. A* **2004**, *108*, 3692-3694.

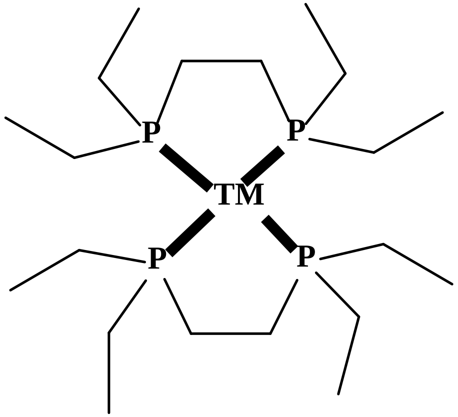
(107) Camaioni, D. M.; Schwerdtfeger, C. A. *J. Phys. Chem. A* **2005**, *109*, 10795-10797.

(108) Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields G. C. *J. Am. Chem. Soc.* **2002**, *124*, 6421.

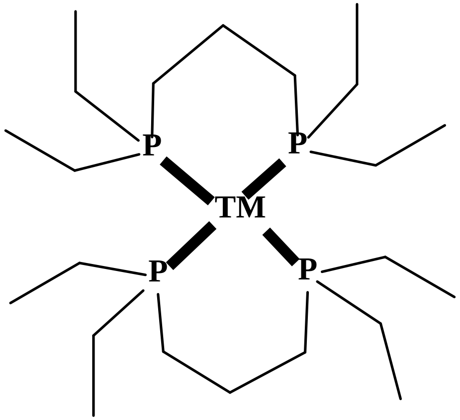
(109) Hodgson, J. L.; Roskop, L. B.; Gordon, M. S.; Lin, C. Y.; Coote .M. L. *J. Phys. Chem. A* **2010**, *114*, 10458-10466.

(110) Ho, J.; Coote, M. *Theor. Chem. Acc.* **2010**, *125*, 3-21.

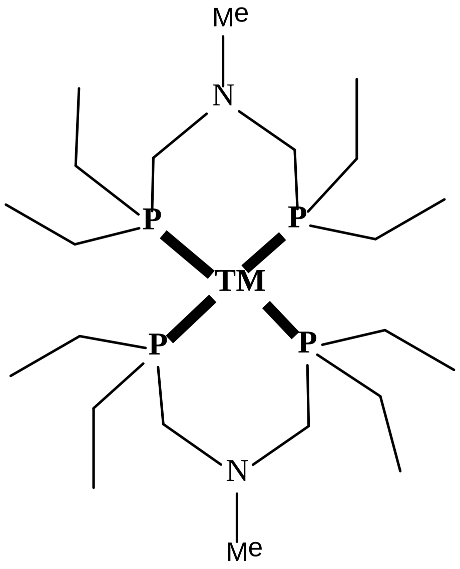
5.6 Figures and Tables



TM(depe)$_2$

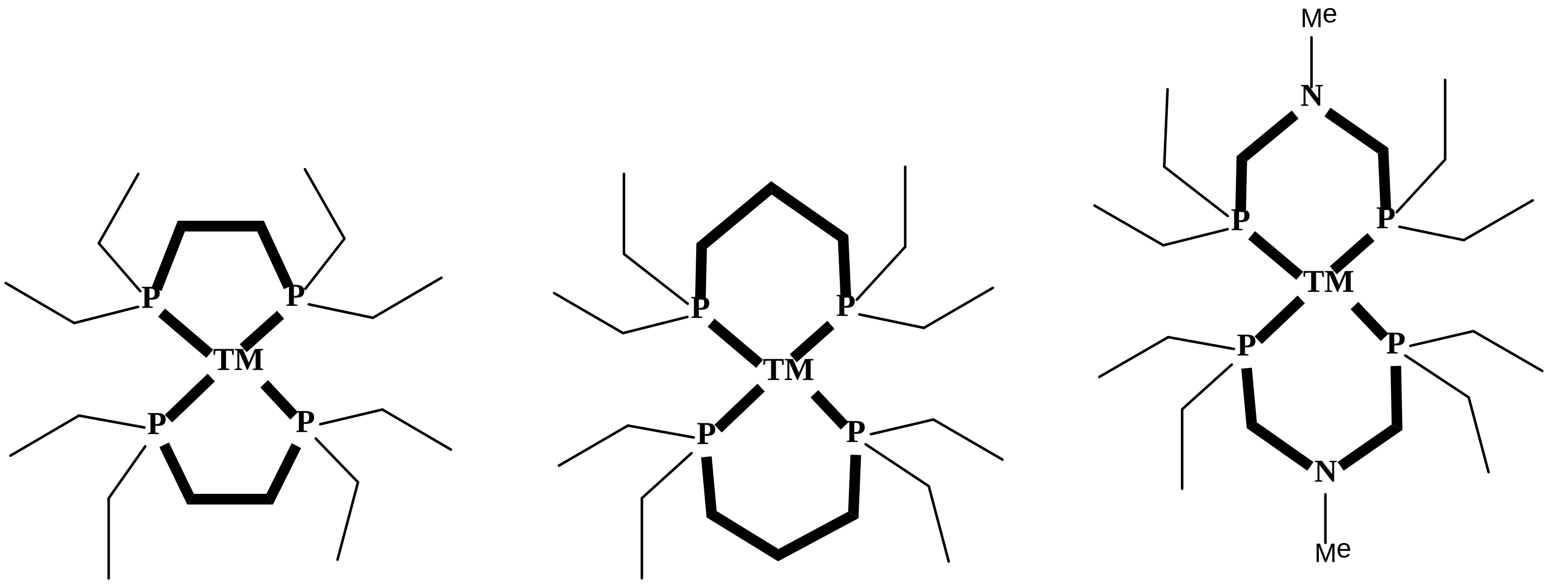


TM(depp)$_2$



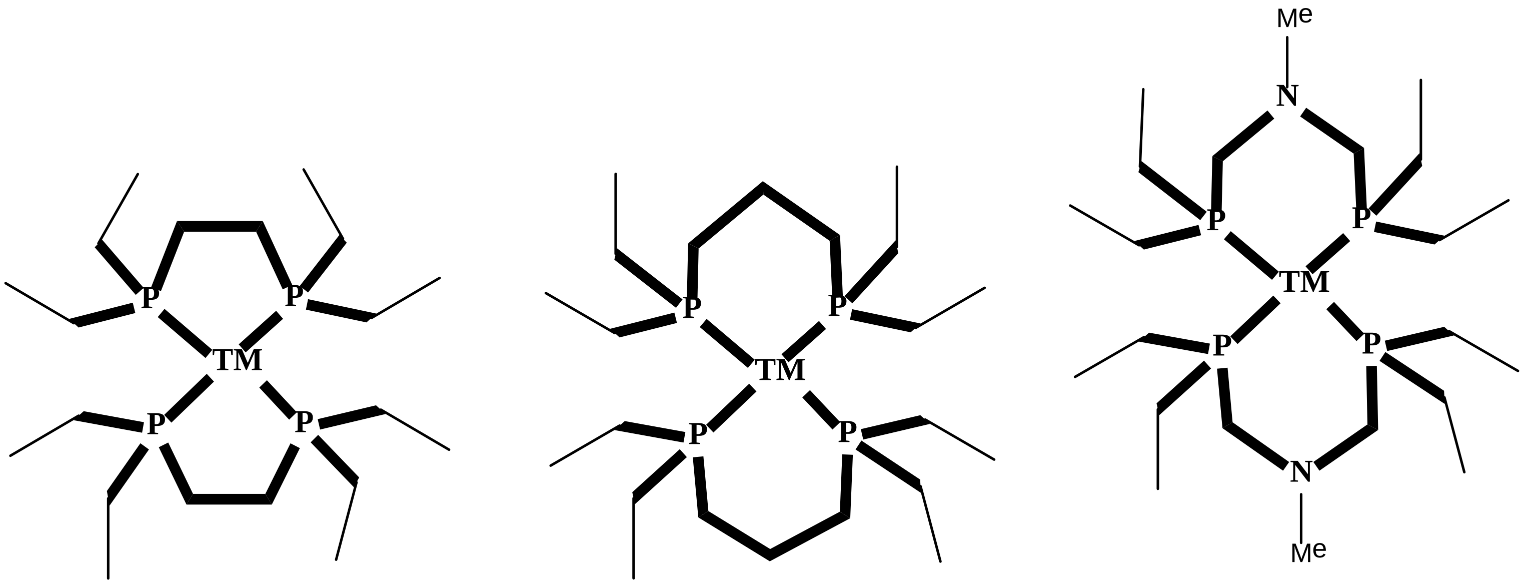
TM(PNP)$_2$

a. ONIOM-1: The QM/QM partitioning scheme for TM hydrides with the TM atom (Ni, Pd, and Pt) and four phosphorous atoms in the layer using the high-level method (in bold).



TM(depe)$_2$ TM(depp)$_2$ TM(PNP)$_2$

b. ONIOM-2: The QM/QM partitioning scheme for TM hydrides with all the atoms within the rings in the layer using the high-level method (in bold).



TM(depe)$_2$ TM(depp)$_2$ TM(PNP)$_2$

c. ONIOM3: The QM/QM partitioning scheme for TM hydrides with all except for the very outside methyl group in the layer using the high level-method (in bold).

Figure 5.1 Expansions of the model system within the QM/QM partitioning scheme.

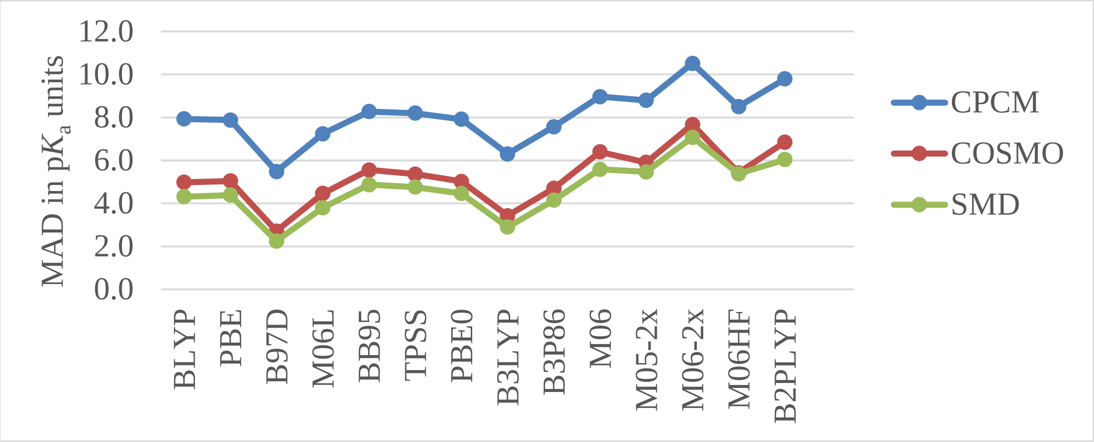


Figure 5.2 Mean absolute deviation (MAD) in pK_a values for fourteen GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA functionals within low-level methods relative to experiment. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ).

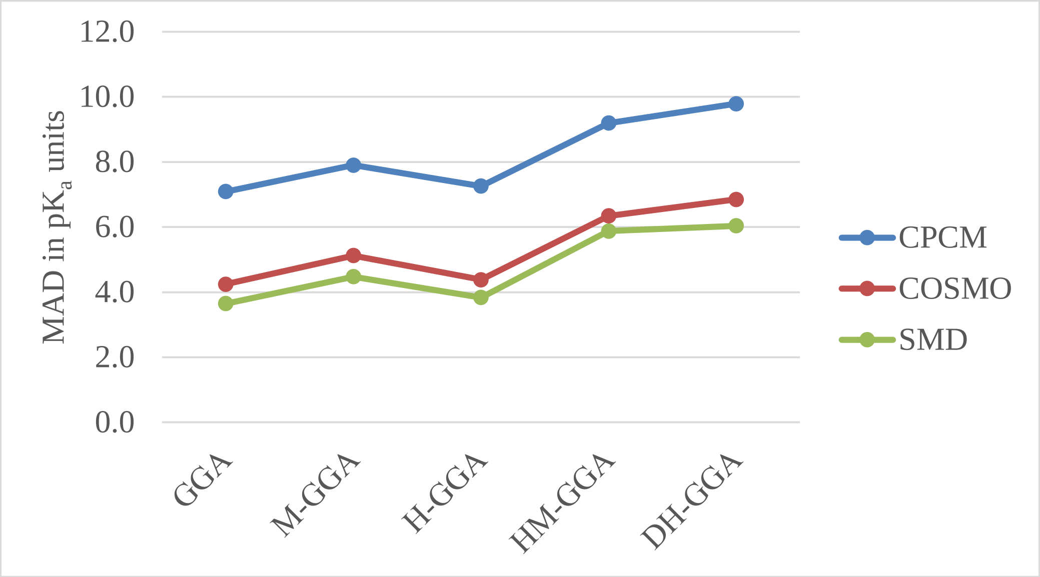


Figure 5.3 Mean absolute deviation (MAD) in pK_a values for five different types of density functionals, GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA functionals, within low-level methods relative to experiment. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ).

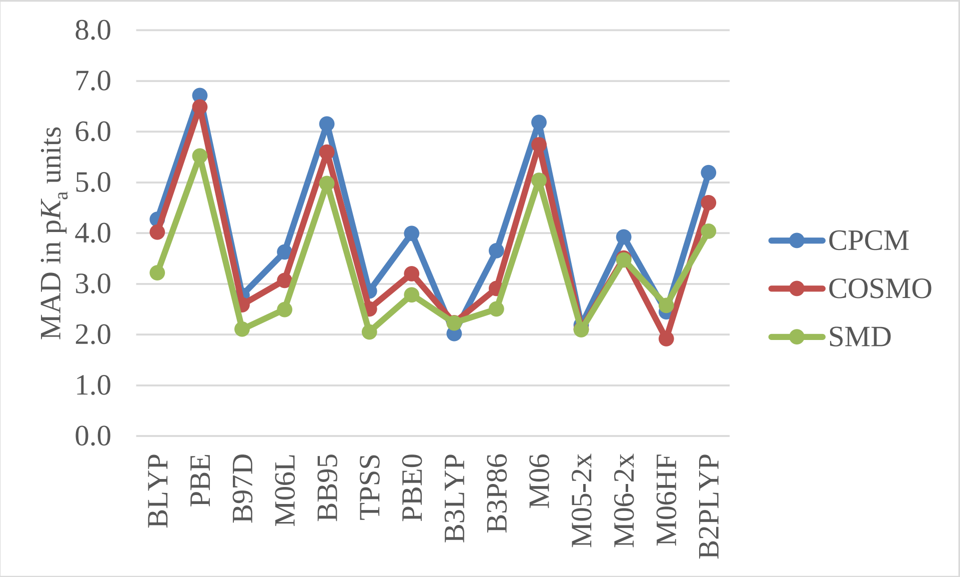


Figure 5.4 Mean absolute deviation (MAD) in pK_a values for fourteen GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA functionals within high-level methods relative to experiment. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ).

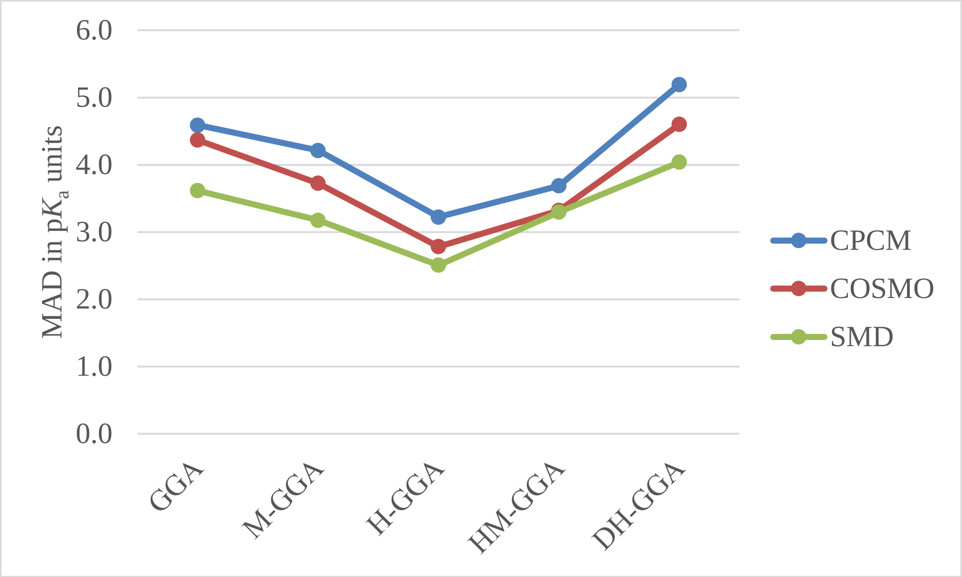


Figure 5.5 Mean absolute deviation (MAD) in pK_a values for five different types of density functionals, GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA functionals, within high-level methods relative to experiment. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ).

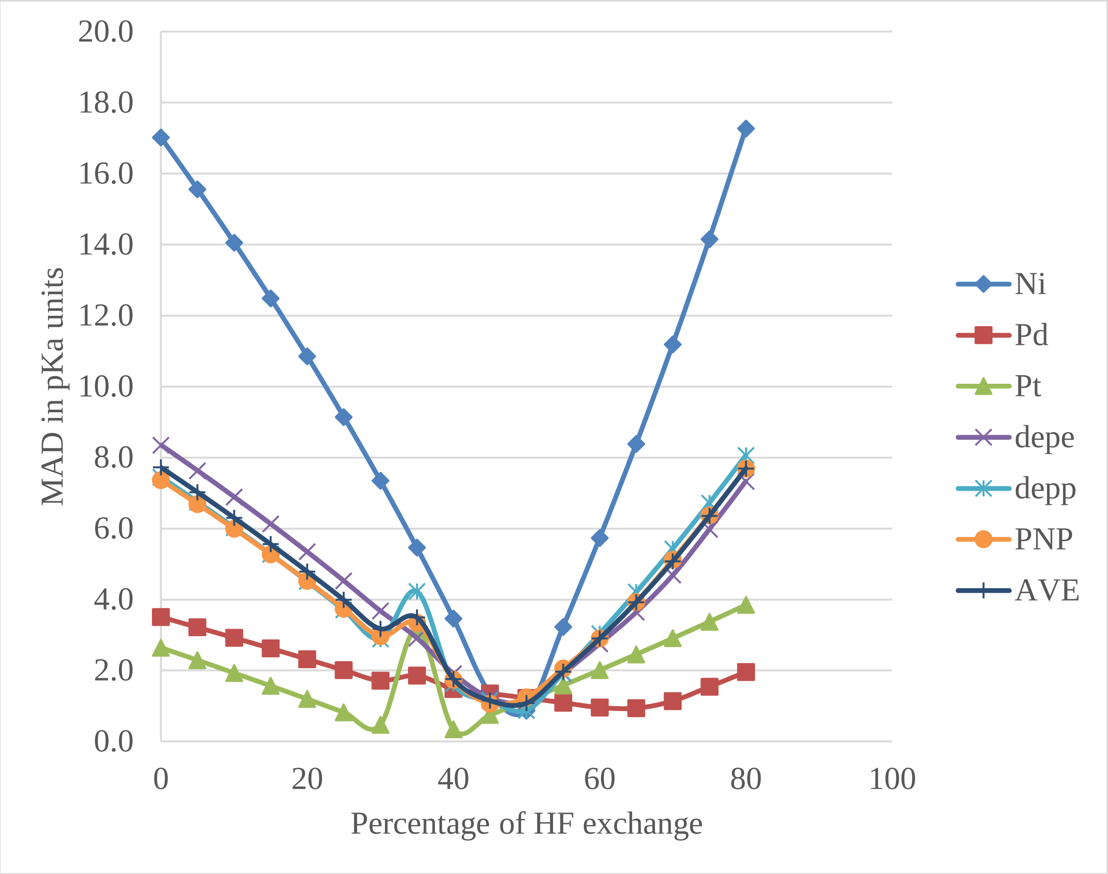


Figure 5.6 MADs of PBE0 *vs.* percentage of HF exchange.

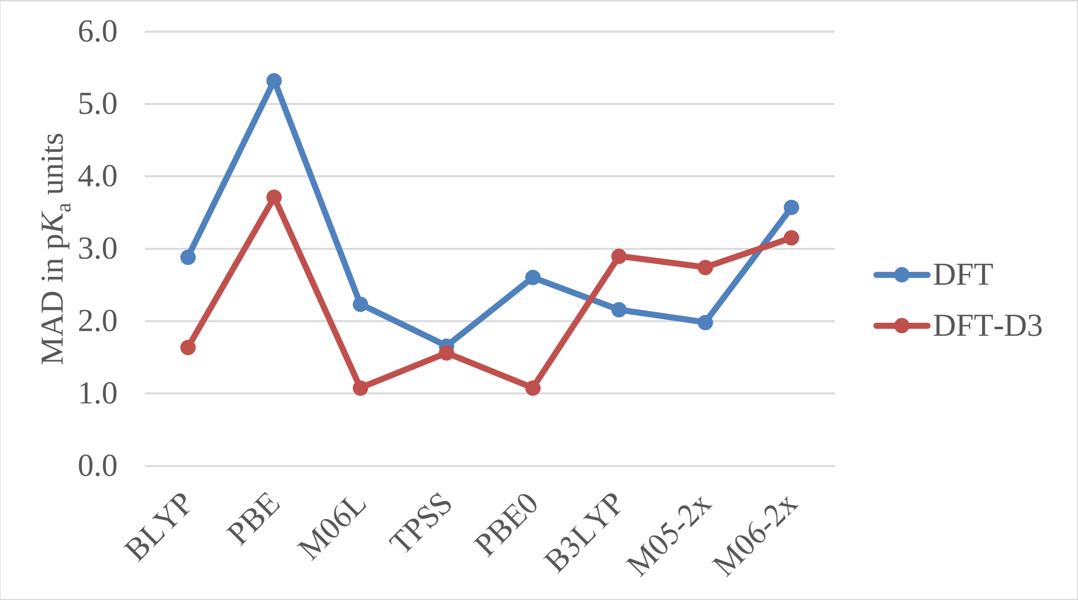


Figure 5.7 MADs of DFT *vs.*DFT-D3 with SMD.

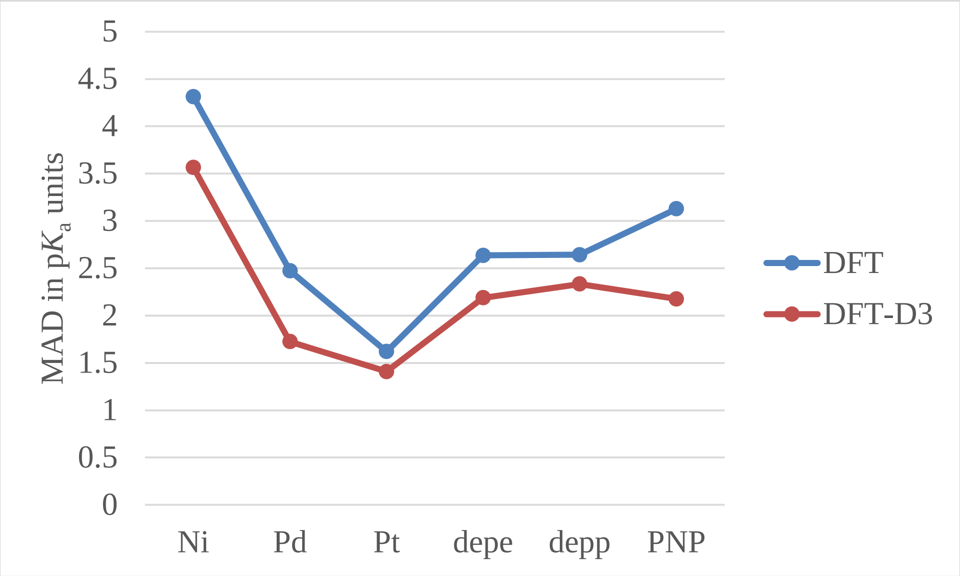


Figure 5.8 Mean absolute deviation (MAD) in pK_a values of DFT and DFT-D3 with the SMD relative to experiment, with respect to different central TM atoms and ligand size of TM hydrides.

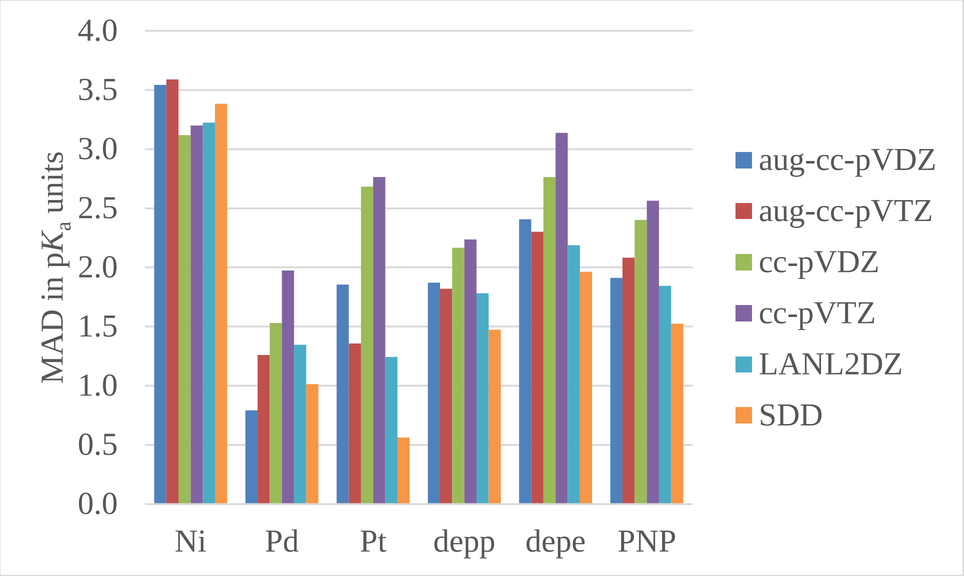


Figure 5.9 Mean absolute deviation (MAD) in pK_a values of different basis set relative to experiment, with respect to different central TM atoms and ligand size of TM hydrides.

Table 5.1 Summary of the DFT functionals utilized in this work

	Type	%HF	Exchange/ Correlation
BLYP[63,64]	GGA	0	Becke88/Perdew86/Lee-Yang-Parr
PBE[65]	GGA	0	Perdew-Burke-Ernzerhof/ Perdew-Burke-Ernzerhof
B97D[66]	GGA	0	B97D/B97D
M06L[16]	M-GGA	0	M06L/M06L
BB95[67]	M-GGA	0	Becke88/Perdew86/Becke95
TPSS[68]	M-GGA	0	Tao-Perdew-Staroverov-Scuseria/ Tao-Perdew-Staroverov-Scuseria
PBE0[65,69,70]	H-GGA	25%	Perdew-Burke-Ernzerhof/ Perdew-Burke-Ernzerhof
B3LYP[63,64,71]	H-GGA	20%	Becke88/Perdew86/Lee-Yang-Parr
B3P86[64,72]	H-GGA	20%	Becke88/Perdew86
M06[16]	HM-GGA	27%	M06/M06
M05-2X[14]	HM-GGA	52%	M05-2X/M05-2X
M06-2X[16]	HM-GGA	54%	M06-2X/M06-2X
M06HF[17,73]	HM-GGA	100%	M06HF/M06HF
B2PLYP[74]	DH-GGA	50%	Becke88/Perdew86/Lee-Yang-Parr

GGA stands for generalized-gradient approximation, M-GGA stands for meta GGA, H-GGA stands for hybrid GGA, HM-GGA stands for hybrid meta GGA, and DH-GGA stands for double hybrid GGA.

Table 5.2 Summary of the basis sets utilized

	Ni Species	Pd and Pt Species
Real system	SDD LANL2DZ	SDD LANL2DZ
Model system	cc-pVDZ-DK aug-cc-pVDZ-DK cc-pVTZ-DK aug-cc-pVTZ-DK	cc-pVDZ-PP aug-cc-pVDZ-PP cc-pVTZ-PP aug-cc-pVTZ-PP

Table 5.3 Theoretical methods for the description of real and model systems within the two-layer ONIOM scheme

Method[d]	Model system[a]	Real system[c]
1	PBE	DFT[b]
2	M06L	DFT[b]
3	B3LYP	DFT[b]
4	M06	DFT[b]

[a] Combined with the aug-cc-pVTZ basis set on main group atoms while aug-cc-pVTZ basis set on Ni using the one-particle Douglas-Kroll-Hess Hamiltonian for scalar relativistic effects and a small core relativistic pseudopotential basis set (aug-cc-pVTZ-PP) on Pd and Pt.

[b] DFT functionals include three GGA functionals (BLYP, PBE, and B97D), three M-GGA functionals (M06L, BB95, TPSS), three H-GGA (PBE0, B3LYP, B3P86), four HM-GGA functionals (M06, M05-2X, M06-2X, M06HF), and one DH-GGA (B2PLYP).

[c] Combined with the LANL2DZ basis set

[d] Each method was examined with three C-PCM, COSMO, and SMD

Table 5.4 Mean absolute deviation (MAD) in pK_a values of GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA types of functionals within low-level methods with different solvation models relative to experiment, with respect to central TM atoms of the TM hydrides. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ)

	Central TM Atom	GGA	M-GGA	H-GGA	HM-GGA	DH-GGA
C-PCM	Ni	7.2	8.0	8.2	10.7	10.6
	Pd	8.0	8.8	7.7	9.7	10.3
	Pt	6.1	6.9	5.9	7.6	8.5
COSMO	Ni	4.5	5.5	5.5	8.1	7.8
	Pd	5.2	6.1	4.9	6.5	7.5
	Pt	3.1	3.8	2.8	4.5	5.2
SMD	Ni	3.8	4.8	4.9	7.5	6.9
	Pd	4.2	5.1	4.0	5.6	6.4
	Pt	2.9	3.6	2.7	4.5	4.9
Overall-MAD[f]	Ni	5.2	6.1	6.2	8.8	8.4
	Pd	5.8	6.7	5.5	7.3	8.1
	Pt	4.0	4.8	3.8	5.5	6.2

[f] Average results of C-PCM, COSMO, and SMD solvation models

Table 5.5 Mean absolute deviation (MAD) in pK_a values of GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA types of functionals within low-level methods with different solvation models relative to experiment, with respect to ligands of the TM hydrides. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ:DFT/LANL2DZ)

	Ligand	GGA	M-GGA	H-GGA	HM-GGA	DH-GGA
C-PCM	depe	7.2	7.7	7.4	9.2	10.0
	depp	7.1	8.1	7.3	9.3	9.9
	PNP	7.0	7.9	7.1	9.5	9.5
COSMO	depe	4.1	4.7	4.2	6.1	6.8
	depp	4.1	5.2	4.3	6.3	6.8
	PNP	4.5	5.5	4.6	6.6	7.0
SMD	depe	3.8	4.3	3.9	5.9	6.3
	depp	3.5	4.6	3.8	5.9	6.0
	PNP	3.7	4.6	3.8	5.8	5.9
overall-MAD[f]	depe	5.0	5.6	5.2	7.1	7.7
	depp	4.9	6.0	5.1	7.2	7.6
	PNP	5.1	6.0	5.2	7.3	7.5

[f] Average results of C-PCM, COSMO, and SMD

Table 5.6 Theoretical methods for the description of real and model systems within the QM/QM scheme

Method[d]	High-level[a]	Low-level[c]
1	B97D	DFT[b]
2	M06L	DFT[b]
3	B3LYP	DFT[b]

[a] Combined with the aug-cc-pVTZ basis set on main group atoms while aug-cc-pVTZ basis set on Ni using the one-particle Douglas-Kroll-Hess Hamiltonian for scalar relativistic effects and a small core relativistic pseudopotential basis set (aug-cc-pVTZ-PP) on Pd and Pt.

[b] DFT functionals include three GGA functionals (BLYP, PBE, and B97D), three M-GGA functionals (M06L, BB95, TPSS), three H-GGA (PBE0, B3LYP, B3P86), four HM-GGA functionals (M06, M05-2X, M06-2X, M06HF), and one DH-GGA (B2PLYP).

[c] Combined with the LANL2DZ basis set

[d] Each method was examined with C-PCM, COSMO, and SMD

Table 5.7 Mean absolute deviation (MAD) in pK_a values for GGA, M-GGA, H-GGA, HM-GGA, and DH-GGA functionals as low-layer methods relative to experiment, with respect to TM species. All of the results are from calculations with ONIOM(DFT/aug-cc-pVTZ(DK on Ni, PP on Pd, Pt):DFT/LANL2DZ)

	Central TM Atom	GGA	M-GGA	H-GGA	HM-GGA	DH-GGA
C-PCM	Ni	7.0	3.7	3.9	4.8	8.9
	Pd	3.7	3.7	3.2	4.1	6.2
	Pt	3.0	2.9	2.6	2.8	4.2
COSMO	Ni	6.8	3.4	3.5	4.9	7.5
	Pd	3.7	3.4	2.9	4.1	5.9
	Pt	2.6	2.2	2.1	1.7	3.3
SMD	Ni	5.9	2.6	2.9	5.1	6.7
	Pd	2.7	2.6	2.5	3.5	4.9
	Pt	2.2	2.1	2.2	1.9	3.2
overall-MAD[f]	Ni	6.6	3.2	3.4	4.9	7.7
	Pd	3.4	3.2	2.9	3.9	5.7
	Pt	2.6	2.4	2.3	2.1	3.6

[f] Average results of C-PCM, COSMO, and SMD

Table 5.8 Mean absolute deviation (MAD) in pK_a values of GGA, M-GGA, H-GGA, and HM-GGA types of functionals for comparison of DFT and DFT-D3 relative to experiment with the SMD

	GGA	M-GGA	H-GGA	HM-GGA
DFT	4.1	1.9	2.4	2.8
DFT-D3	2.7	1.3	2.0	2.9

Table 5.9 Mean absolute deviation (MAD) in pK_a values relative to experiment for four functionals when combing with different basis sets as high-level methods

	aug-cc-pVDZ	aug-cc-pVTZ	cc-pVDZ	cc-pVTZ
B97D	3.0	2.8	4.0	2.7
TPSS	1.5	1.5	1.4	3.3
B3LYP	1.1	0.9	1.6	1.1
M05-2X	1.8	1.9	1.9	1.9
One-particle Douglas-Kroll-Hess Hamiltonian for scalar relativistic effects was considered for Ni; A small core relativistic pseudopotential basis set (aug-cc-pVTZ-PP) on Pd and Pt				

Table 5.10 Mean absolute deviation (MAD) of five cavity models in pK_a values relative to experiment using ONIOM(B3LYP/aug-cc-pVTZ:B97D/LANL2DZ)

C-PCM					
	Pauling	Bondi	UA0	UAKS	Default
Ni	0.9	1.2	3.6	1.2	3.4
Pd	1.5	1.6	4.3	1.7	1.7
Pt	0.3	0.4	2.2	0.6	0.6
depe	0.7	0.8	3.5	1.4	2.0
depp	0.8	1.0	3.6	0.9	1.8
PNP	1.2	1.4	2.9	1.1	1.9
Overall	0.9	1.1	3.3	1.1	1.9
COSMO					
	Pauling	Bondi	UA0	UAKS	Default
Ni	0.9	1.2	3.6	1.2	0.6
Pd	1.5	1.6	4.3	1.7	1.3
Pt	0.3	0.4	2.2	0.6	0.8
depe	0.7	0.8	3.6	1.4	0.9
depp	0.8	1.0	3.6	0.9	0.6
PNP	1.3	1.4	2.9	1.1	1.2

Overall	0.9	1.1	3.4	1.1	0.9
SMD					
	Pauling	Bondi	UA0	UAKS	Default
Ni	0.6	0.9	3.6	0.9	0.5
Pd	1.0	1.1	4.0	1.1	0.9
Pt	0.4	0.3	2.0	0.4	0.9
depe	0.7	0.7	3.4	1.0	0.8
depp	1.0	0.6	3.5	0.6	0.7
PNP	0.9	1.0	2.6	0.9	0.8
Overall	0.8	0.8	3.2	0.8	0.8

Table 5.11 Mean absolute deviation (MAD) in pK_a values relative to experiment of three expansions of model system of TM hydrides with the SMD

ONIOM Scheme	ONIOM-1	ONIOM-2	ONIOM-3
B97D/aug-cc-pVTZ:B97D/LANL2DZ	1.4	2.8	3.2
TPSS/aug-cc-pVTZ:B97D/LANL2DZ	1.2	0.9	0.5
B3LYP/aug-cc-pVTZ:B97D/LANL2DZ	0.9	1.2	1.5
M05-2X/aug-cc-pVTZ:B97D/LANL2DZ	1.0	1.2	1.3

CHAPTER 6 CLOSING REMARKS AND FUTURE OUTLOOK

My dissertation focuses on two aspects. The first aspect is the examination of utilizing a variety of density functionals for calculating the thermodynamic properties of TM containing reactions in both gas phase and solvation phase. For gas phase studies, the energetics of C-O bond activation of methoxyethane and methanol by late 3*d* and 4*d* TM atoms were calculated by a set of density functionals. The functional that reproduced the results from accurate *ab initio* (CR-CCSD(T)) method in the model reactions by methoxyethane and methanol, was applied to investigate the ability of TM atoms to break the C_β-O bond of β-O-4 linkage of lignin. For the solvation phase study, the performance of density functional were assessed for predicting the pK_a of 9 Group 10 TM hydrides. The second aspect is the determination of diagnostic criteria for analysis of multireference character of the molecules containing 4d TMs.

6.1 C-O Bond Cleavage Using Transition Metal Atoms: Inconsideration of the Utility of Density Functionals and Intrinsic Catalytic Properties of TMs

Due to the rapid exhaustion of easily recovered fossil fuels, the degradation of lignin, one of the most abundant biomass on the earth, is of considerable interest as its potential to be used as renewable energy source. However, the investigated methods for the decomposition of lignin, such as ligninolytic enzymes and pyrolytic degradation, are either expensive or poor at selectivity of the products. As TM-catalyzed depolymerization has been widely used in industry, the ability of TMs to activate β-O-4 linkage of lignin, which comprises about 50% of all the linkages in a typical lignin structure, has been of increasing interest. Although density functionals have been applied in the field of TM containing chemical reactions due to its good balance of computational cost and accuracy, the accuracy of density functionals are highly system dependent. In order to produce reliable energetics of the decomposition of the β-O-4 linkage of lignin by TMs so that the results can guide the design of practical TM catalysts, the appropriate density functional for this particular investigation should be used instead of randomly choosing a density functional based on its outperformance in other systems. My study focused on the C_β-O bond of β-O-4 linkage of lignin by TMs; therefore, density functionals that can reliably describe C-O bond activation by TMs should be used. However, using the β-O-4 linkage of lignin to assess the performance of density functionals is computationally expensive. The late TM mediated C-O bond activations of methoxyethane and methanol were applied to assess the utility of density functionals for this type of reactions. For the C-O bond activation of both model molecules, PBE0 resulted in the lowest deviation from CR-CCSD(T) calculations that was used to gauge the utility of density functionals. Therefore, the C_β-O bond cleavage of β-O-4 linkage of lignin using late TM atoms was investigated with PBE0, which found that Pd produced flatter reaction pathway than other considered TM atoms and were suggested to be used in more practical catalysts, such as organometallic complexes and metal surfaces for the degradation of the β-O-4 linkage of lignin.

Another important perspective is to compare the performance of density functionals in solvation phase. As mentioned above, there are no density functionals that can provide reliable results universally. The present interest is approaching the best density functional for calculation of pK_a of TM hydrides using implicit solvation models that are important intermediates for many catalytic processes. Since well-defined TM hydrides are commonly sizable with bulky ligands, density functional was always combined with a low level basis set, LANL2DZ that cannot describe the behavior of TM sufficiently. Therefore, the ONIOM method, which utilizes the combinations of density functionals with more complex basis sets as the high level method to portray the chemically important part of molecules as well as with smaller basis sets as the low level for the whole structure, is preferred. The main purpose of this study is to obtain the optimal combination of density functionals for using the ONIOM method to predict the pK_a of TM hydrides. Among 14 density functionals, using B97D within low level method resulted the lowest MADs relative to experiment. For high level methods, B3LYP and M06HF resulted in lowest MADs with C-PCM and COSMO, while B97D, TPSS, and M05-2X yielded the same MADs that are lower than MADs from other functionals with SMD. Addition of the Grimme's empirical dispersion correction did improve the accuracy of density functionals. Additionally, the HF exchange played an important role in the performance of density functionals using within high level methods. 50% HF exchange, in general, yielded the lowest errors.

6.2 Diagnostic Criteria for 4d TM-Containing Molecules

Due to partially filled *d* orbitals and the similar spatial extent and energy of the (n+1)*s* and n*d* orbitals, many transition metal (TM)-containing species are open shell systems that may have degenerate or low-lying nearly degenerate electronic states that may necessitate a multireference (MR) wavefunction-based approach to properly describe the near degeneracies. Therefore, a prior analysis of the MR character is necessary for TM molecules so that the proper theoretical methods can be used in calculations. Diagnostics can serve as an important aid in the analysis of the MR character. In our study, investigations of suitable diagnostic criteria for 4*d* TM molecules were assessed for three diagnostics, T_1, D_1, and %TAE based on various categories of 110 4*d* TM molecules. The criteria $T_1 > 0.045$, $D_1 > 0.120$, and %TAE > 10 is proposed as a set of criteria to aid in identifying the significance of MR character for 4*d* TM-containing species. Additionally, the combined use of various diagnostics can provide a more reliable prediction of MR character than using a diagnostic alone.

6.3 Future Interests

Different abilities of TMs on breaking C_α-C_β and C_β-O bonds of the β-O-4 linkage of lignin affect the reaction mechanism of decomposition of lignin. The results obtained from the above study on TM atoms (Chapter 3) mediated C_β-O bond activations of the β-O-4 linkage of lignin as well as C_α-C_β bond activation from a previous study by our group have provided an useful insight into the intrinsic catalytic properties of different TMs, which can guide our investigation of the reaction mechanism in the degradation of the β-O-4 linkage of lignin with more practical TM catalysts than single TM atoms. Organometallic complexes and TM clusters will be considered to activate the β-O-4 linkage of lignin. Since catalysis reactions that include organometallic complexes commonly happen in solvent, the impact of the choices of solvent on the reaction energetics of TM mediated lignin degradation is also of interests. The study in Chapter 5 that accessed the utility of computational methods for reactions with large TM containing species in solvation phase can aside the choice of appreciated computational methods for this mentioned investigations.

In Chapter 5, the thermodynamic cycle that was used for calculating the pK_a of TM hydrides is a direct cycle. While a direct cycle is the most commonly implemented cycle, it does not always result in smaller errors than other thermodynamic cycles. Therefore, in future work, we aim to test the performance of various thermodynamic cycles, such as proton exchange and semi-direct cycles, for calculating the pK_a of TM containing species. Since the implicit solvation models used in Chapter 5 are not able to explicitly account for hydrogen-bonding, inclusion of discrete solvent molecules can sometimes improve the modeling of solute-solvent interactions. Therefore, we are also interested in investigating the impact of including solvent molecules explicitly in calculated pK_a of TM containing species.