5-2022

The Exploration of Small Molecules, Lanthanide Complexes, and Catalysis using Electronic Structure Theory, Dynamics, and Machine Learning

Gavin McCarver
University of Tennessee, Knoxville, gmccarve@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Gavin McCarver entitled "The Exploration of Small Molecules, Lanthanide Complexes, and Catalysis using Electronic Structure Theory, Dynamics, and Machine Learning." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Chemistry.

Robert J. Hinde, Major Professor

We have read this dissertation and recommend its acceptance:

Konstantinos D. Vogiatzis, Craig E. Barnes, Kate Jones

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

The Exploration of Small Molecules, Lanthanide Complexes, and Catalysis using Electronic Structure Theory, Dynamics, and Machine Learning

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Gavin Alexander McCarver

May 2022

Dedication

This dissertation is dedicated to all those in my life who have been there for me over the last two decades of my scholastic career. First and foremost, my parents Richard and Toni McCarver, who have supported me as a both a budding scientist interested in electromagnets and soap to the chemist I am today. Next to my siblings, Miranda, Garrett, and Jared McCarver, who have helped me in too many ways to list. To Gordon Tolman, Robert Tolman, and Ian Tolman, who could not be here to see this moment but have inspired me in their own ways. To my lab mates both past and present: Kiran Dhah, Andrea Becker, Justin Kirkland, Grier Jones, and Brett Smith for helping from the beginning to the end of grad school. And finally to my wonderful girlfriend and best friend Rhianna Moore, who has let me vent, bounce ideas off her, and has helped me destress and survive in the last year of grad school like no one else could have.

Acknowledgements

I would like to thank both Dr. Robert J. Hinde and Dr. Konstantinos (Kostas) D. Vogiatzis for their invaluable guidance, mentorship, and help over the last five years. Dr. Hinde helped me to become the physical and computational chemist that I wanted to become and Kostas helped me reach new heights as a chemist in general. I would also like to thank the National Institute for Computational Sciences (NICS) at the University of Tennessee, Knoxville for their computational resources which allowed the work detailed in this dissertation to progress. Finally, I would like to thank Dr. Craig Barnes and Dr. Kate Jones for serving on my doctoral committee and providing insight from their own areas of expertise.

Abstract

With the ever increasing availability of computational resources, more challenging chemical systems can be studied. Among these challenges are the rotational and vibrational spectra of diatomic molecules within spectroscopic accuracy, the environmental perturbations induced on a rotating water molecule, the prediction of free binding energies of lanthanide complexes using machine learning, and the study of catalytic mechanisms through a theoretical framework. High levels of electronic structure theory were combined with a rigorous treatment of either the anharmonic vibrational wave functions to study diatomic molecules or the rotational wave functions to study H_2O-pH_2 interactions. The former was initially applied to the CF^+ cation and excellent agreement was observed between theoretical and experimental spectroscopic constants. Likewise, the H_2O-pH_2 interactions were utilized to identify satellite peaks in the infrared spectra of a H_2O-doped, pH_2 crystal lattice. These peaks most likely occur due to a vacancy site directly around the H_2O molecule. The study of lanthanide complexes is challenging due to their unique electronic structure. Specifically, lanthanide-tris-β-diketone complexes were studied to calculate their respective free binding energies. Machine learning was utilized in this instance to act as the function which mapped the structure of the β-diketone ligands to the free binding energies. Predictions were made and several β-diketone ligands were identified which maximized the separation between lanthanide and lutetium. Finally, the study of catalytic mechanisms using theoretical methods is not without challenge due to the complex electronic structure of such systems. The hydrogen evolution reaction, the dehalogenation of CH_2Cl_2, the hydrogenation of small, unsaturated hydrocarbons, and the hydroformylation reaction were studied using either molecular electrocatalysts or transmetalated forms of the HKUST-1 metal-organic framework.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

- BTC – Benzene Tricarboxylate
- CASSCF – Complete Active Space Self Consistent Field
- CBS – Complete Basis Set
- CC – Coupled Cluster
- CCSD(T) – Coupled Cluster with single, double, and perturbative triple excitations
- DFT – Density Functional Theory
- DFTB – Density Functional Tight Binding
- DKH – Douglas-Kroll-Hess
- EST – Electronic Structure Theory
- FCI – Full Configuration Interaction
- HCP – Hexagonal Close-Packed
- HF – Hartree-Fock
- HO – Harmonic Oscillator
- ISM – Interstellar Medium
- metaD – Metadynamics
- MM – Molecular Mechanics
- MOF – Metal-Organic Framework
- MP2 – 2^{nd} order, Møller–Plesset Perturbation Theory
- MRCI – Multireference Configuration Interaction
- PEC – Potential Energy Curve
- PES – Potential Energy Surface
- pH_2 – Parahydrogen
- SCF – Self Consistent Field
- TS – Transition State
- VPT2 – Vibrational, 2^{nd} Order Perturbation Theory
- ZORA – Zeroth Order Relativistic Approximation

Chapter 1

Introduction

1.1 Challenges in Computational Chemistry

Through the course of the twentieth century, science saw the rise of many new technologies which pushed the field of chemistry forward. Some of these technologies helped to advance traditional, benchtop chemistry while others helped on the theoretical side of chemistry. Since the derivation of the Schrödinger[3] and Dirac[4] equations in the early twentieth century, the subfield of theoretical chemistry has grown and progressed and through a collaborative approach with computational sciences, has spawned a new subfield of chemistry often referred to as computational chemistry. In this subfield, a combination of theory and computation come together to benefit many areas of science including chemistry, materials science, and even biology. In its early stages, computational chemistry was considered more of a supplemental tool for experimental work but now sets the stage in some instances to not only be vital for but also to supplant experimental work.

Just as traditional, synthetic chemistry has its limitations, so too does computational chemistry. A limitation that exists for this subfield of chemistry is often due to the scaling of different computational methods as a function of the size of the system of interest. Generally, as higher accuracy is desired, the size of the system that may be studied decreases in a nonlinear fashion. This scaling problem is different depending on the approximations included in the underlying theoretical framework. For example, full configuration interaction (FCI) may scale as N! (where N is the size of the system) and can be useful for diatomics and triatomics, coupled cluster with single, double, and perturbative triple excitations (CCSD(T)) scales as N^7 and is more appropriate for systems less than 10-20 atoms, and density functional theory (DFT) scales as N^3. The latter of which is the most commonly applied method for larger, molecular complexes. Beyond

DFT, even larger systems can be examined utilizing methods such as density functional tight binding (DFTB) or molecular mechanics (MM) which in some instances, scale linearly.[5] Thus, each method is applicable for different chemical systems and may solve different challenges and problems in computational chemistry.

In recent years, many computational chemists have put forth what they believe to be the current challenges in computational chemistry or challenges which may arise in the future for computational chemistry. Dr. Thomas Hofer wrote in a 2013[6] that one of the most prominent challenges in chemistry was the protein folding problem.[7] This problem is quite interdisciplinary in nature and links fields such as biology and biochemistry with essentially every field of chemistry: inorganic, organic, analytical, physical, and theoretical chemistry. The role of theoretical and computational chemistry for this problem lies in the ability to study proteins on an atomistic scale. Hofer continues to discuss the merits of the proper inclusion of solvation effects which play a very large role in problems such as protein folding.[8–12] Solvent effects may either be introduced implicitly through the use of an external potential or explicitly through the inclusion of solvent molecules.[13–16] The latter method achieves a higher level of accuracy but at a greater computational cost. A final challenge that Hofer discusses is that of nuclear quantum effects which are strongly linked to systems which involve hydrogen transfer reactions such as those of certain proteins.[17–23]

The three challenges put forth by Hofer are by no means the only grand challenges in computational chemistry, however. Krylov *et al*. focused on three different challenges in computational chemistry: catalyst design, long-range charge and excitation transfer, and intrinsically disordered proteins.[24] The role catalysts play in modern society are incredibly far reaching; generating the petroleum products[25] that are so prevalent in every aspect of life as well leading to the synthesis of any number of pharmaceutical compounds.[26] In designing new catalysts, higher activity and selectivity are often sought after as these improve the yield of any given reaction, reduces the amount of waste produced, and reduce the number of side reactions which can make unwanted side-products.[27] By using computational methods and approaches, a better understanding of the underlying physics and chemistry of different catalysts can be achieved. This gives computational chemists the ability to suggest new catalysts to experimentalists prior to the need for costly and time-consuming synthetic studies.[28] Krylov *et al*. follow up their discussion on

catalyst design by introducing another area of chemistry which deals with long-range transfer reactions such as those observed in light harvesting[29] and cellular respiration. These materials are usually quite large, such as one decaheme cytochrome which contains 200k atoms. To study such a large system, a combination of computational approaches is required in which parts of the cytochrome are treated classically using traditional force fields while other parts are treated quantum mechanically. The latter of which usually includes active sites and the area surrounding said active sites. Finally, Krylov *et al* draws similar connections to the first challenge discussed by Dr. Hofer: intrinsically disordered proteins.[30,31] Such a problem is challenging from a computational perspective due to the large size of such proteins. Because of the large number of atoms in these proteins, many thousands of conformations are possible for a protein and identifying the exact experimental geometry is often difficult or impossible to achieve.

Three challenges in computational chemistry which are by no means grand challenges are examined and discussed in the following chapters of this dissertation: the reproduction of high accuracy, rotational, vibrational, and rovibrational spectra of diatomics and triatomics, the study of lanthanides and their separation in solution, and the study of catalysis within a computational framework. A brief discussion about the current challenges and limitations of each of these is discussed below.

1.1.1 High Accuracy, ab initio Spectroscopy

Rotational spectroscopy plays a large and important role in chemical disciplines such as atmospheric and astronomical chemistry.[32] The spectra that are collected from scanning other planets' atmospheres as well as from the interstellar medium (ISM) usually contain a multitude of chemical signatures.[33] These signatures contain information not only of the chemical identity but also yield information regarding the rotational excitation that led to that spectral peak. To identify the molecule which gave rise to said spectral peaks, it is important to have accurate values for different spectroscopic constants which is an area of research which may be examined using computational chemistry. To this end, the CCSD(T) method is used as the primary level of theory to approach the accuracy desired to calculate the spectroscopic constants of certain small- to

medium-sized molecules.[34,35] Once calculated, the spectroscopic constants can be used to help identify some of the spectral peaks collected.

The challenge of the calculation of the spectroscopic constants of interest for rotational spectroscopy lies in the inclusion of higher order effects. These effects account for core-valence interactions, the complete basis-set (CBS) limit, higher order terms within the coupled cluster expansion such as full triple excitations, full quadruple excitations, and so on, relativistic effects, as well as vibrational corrections.[36] A composite scheme is used to develop force constants[37] for the molecule of interest which can then be used in conjunction with the vibrational, second-order perturbation (VPT2)[38] method to calculate the spectroscopic parameters. As the name suggests, VPT2 may also be applied for the calculation of vibrational spectra. As the size of the molecule increases, the calculation of the spectroscopic parameters increases in complexity in two ways: 1) the number of calculations required to develop the necessary force constants increases and 2) the complexity of the individual calculations increases due to more electrons, atoms, and basis functions. Thus, the number and size of the calculations necessary to determine theoretical spectroscopic constants increases with system size in a non-linear fashion. For a more complete explanation of how computational chemistry is applied for the study of rotational spectroscopy, we refer the reader to work by Dr. Ryan Fortenberry and Dr. Cristina Puzzarini.[32,34,39]

In determining the spectroscopic constants of small- to medium-sized molecules, a collaborative effort may be made between computational chemistry and astrochemistry. According to Dr. Christina Puzzarini, several grand challenges of astrochemistry exist which may be aided through computational chemistry.[40] Primarily, computational chemistry may be used in the identification of new molecules in the ISM. As more observational studies are performed, more spectra are being generated which may include yet unidentified chemical species. Two approaches exist which may aid in the identification of these new species: 1) gas phase studies of likely candidates and 2) computational studies of likely candidates. By using both gas phase and computational studies, a more robust identification may be made. The next grand challenge put forward by Dr. Puzzarini is that of our understanding of the chemical reactivity within the ISM. This includes improving the knowledge of the pathways molecules may take as they interact with one another in the ISM. This is a large endeavor, however, and is not without limitations, as mentioned by Dr. Puzzarini. Finally, based upon the first two grand challenges, a third challenge

may be put forth which deals with the origins of life in the universe.[41] Life arose through a complex series of events and by furthering our understanding of the chemistry of the ISM and the reaction mechanisms possible therein, possible explanations of where life came from may be possible and computational chemistry is one such pillar of science which may help to achieve this.

1.1.2 Chemistry of the f-elements

The study of lanthanides using computational methods is quite different when compared to either transition metals or actinides. The 4f orbitals of the lanthanides are buried beneath the already filled 5s and 5p orbitals which leads to the lanthanides favoring interactions more strongly based on steric effects rather than electronic effects. This is different from transition metal complexes which favor interactions based on electronic effects while actinides are somewhere in between the two extremes.[42] A consequence of this is the high-spin nature of lanthanide atoms, cations, and complexes which actually leads to calculations including them being less challenging. This is because high accuracy calculations involving lanthanides do not need to resort to methods such as the complete active space SCF (CASSCF) method.[43–45] Transition metal and actinide complexes should utilize this method due to the near degeneracy among lower lying excited states.[46] The challenge of calculations involving lanthanides is therefore different from transition metal complexes and usually involves either a larger number of calculations needing to be performed or the proper inclusion of relativistic effects. As there are fourteen naturally occurring lanthanides (no stable isotopes exist of promethium), any study which hopes to examine the lanthanide series requires fourteen set of more calculations. Thus, if work were to be performed which hoped to identify the strongest lanthanide-ligand interaction, fourteen calculations must be performed which, depending upon the size of the lanthanide-ligand system, may take several hours to days to perform. If this work were to be extended to multiple ligands, then the number of calculations required would begin to grow even larger. An example of such work is the study of lanthanide binding as it relates to lanthanide separation. For a more detailed explanation of the work that goes into lanthanide (and actinide) separation for heterocyclic N-donors in solutions, we refer the reader to a recent review article.[47]

Additionally, the proper inclusion of certain physical aspects of lanthanides must be included to compute the most theoretically correct values. Lanthanides are generally considered to be heavy elements which have different physics from lighter elements such as oxygen or fluorine. Thus, the inclusion of relativistic effects for any study involving lanthanides is a necessity.[48,49] These effects can be included in the most complete manner utilizing four-component methods or approximations can be included which leads to more tractable calculations at the cost of some amount of accuracy. The zeroth-order, relativistic approximation (ZORA)[50,51] and Douglas-Kroll-Hess (DKH)[52,53] method both approach relativistic corrections in different manners but their inclusion in any calculations allows for a more chemically and physically meaningful result

1.1.3 Computational Catalysis

The third and final challenge in computational chemistry discussed here is that of computational catalysis. Specifically, computational catalysis as performed using molecular complexes and not surfaces. For the surfaces, we refer the reader to recent literature.[54–57] Catalysis as a field of chemistry has been in existence for many centuries and is now responsible for much of our modern world.[58] The production of ammonia through the Haber-Bosch process[59] helps to feed the world and the catalytic cracking of long chain hydrocarbons[25] allows for the production of valuable materials such as gasoline and polymeric materials are two prominent examples of the impact that catalysis has on the modern world. Catalysts work by lowering the activation barrier of reactions which then may proceed through alternative mechanistic pathways. This is often achieved through the use of transition metal complexes but can also be facilitated using p-block metal complexes as well

The challenge in studying catalysis through a computational framework is two-fold: 1) many possible pathways may exist for even simple reactions and 2) the theoretical framework itself may change which pathway is the most energetically favorable. With regard to the first point, if one considers the hydrogen evolution reaction using an Sb-salen catalysis, then many possible pathways are feasible due to the non-innocence of the ligand. It may be more favorable to protonate two different atoms on the salen ligand which complicates the supposed two-electron/two-proton

reaction mechanism as more than two protons may be involved. In addition to this, different theoretical frameworks such as different density functionals or wave-function based methods may arrive at different conclusions given the different natures of their formulations. Because of this, the study of catalytic systems and mechanisms using computational methods must begin with a benchmarking study to identify the most appropriate theoretical framework to be used but must also examine enough of the possible reaction mechanisms to as to arrive at the most accurate answer. For a recent review on the literature of the study of computational catalysis on the surface of metal-organic frameworks (MOFs), we refer the reader to reference 60 and for a more general approach to computational catalysis, we refer the reader to reference 28.[28,60]

1.2 Dissertation Overview

This dissertation is outlined in the following manner. In Chapter 2, a discussion on the low temperature, rotational and vibrational excitations of two small molecules (CF^+ and H_2O) is presented. Chapter 3 details work performed on the study of lanthanide compounds which range in size from simple diatomics to much larger, molecular complexes. Chapter 4 then outlines several collaborative works performed which explore computational catalysis using molecular complexes. Finally, Chapter 5 draws conclusions from the work discussed and connections among the different areas studied.

Chapter 2

Rotational, Vibrational, and Rovibrational Excitations of Isolated CF^+ and Matrix Isolated H_2O

The following chapter examines how electronic structure theory (EST) calculations can be used to gain a deeper understanding of the rotational, vibrational, and/or the rovibrational nature of small molecules isolated in the gas-phase or in parahydrogen matrices. In the former case, we will consider diatomic molecules, while in the latter case we will consider a triatomic molecule. Due to the small number of atoms in the molecules of interest, a thorough examination utilizing high accuracy EST methods can be performed. Each section will begin with the explanation of the context of the problem followed by an exploration of previous work. Following this, a detailed analysis of the calculations performed for each system will be presented. All of the work presented herein corresponds to published work or work near being published where the author's (Gavin McCarver) contribution was to perform the electronic structure calculations and develop programs to utilize said calculations.

2.1 Introduction

Small molecules such as diatomics and triatomics allow for a very thorough examination using EST methods due to the relatively few numbers of electrons, small number of basis functions, and few degrees of freedom within the potential energy surfaces. Likewise, when high

accuracy experimental work is done on such small systems, near quantitative comparisons can be made. Low-temperature, rare-gas matrices as well as the interstellar medium (ISM) are two environments which give rise to essentially isolated molecules. Herein we refer to molecules in either environment as isolated even though matrix isolation involves dopant-matrix interactions. From these two environments, rotational, vibrational, and rovibrational spectra may be collected which, due to the low temperatures, are often easy to analyze.

The study of the rich chemical environment of the ISM began with the identification of the inversion transition of ammonia in 1968.[61] Thus far, more than two hundred molecules have been identified in the ISM by their rotational, vibrational, or electronic excitations.[62] In the case of homonuclear diatomics such as H_2, their presence in the ISM has only been identified using electronic excitations. The chemistry of this environment is one of high-energy photons, incredibly low densities, and very long timeframes between molecular collisions. Because of this, many exotic species that are not usually found in terrestrial environments have been identified such as C_2,[63,64] H_3^+,[65] C_5H,[66] and even fullerenes.[67] The list of identified diatomic molecules ranges from the simple H_2,[68] to the strange SiS,[69] to the heavy and biologically relevant FeO.[70] These molecules and others paint a picture of the unique and strange environment that arises from the ISM and thus, it is of great chemical importance to better understand this environment for multiple reasons. The simplest molecules to study in this case are diatomics such as those mentioned previously. The potential energy surfaces of diatomic molecules are one-dimensional and as a consequence, the vibrational, rotational, and rovibrational states of these molecules can be examined with very little computational effort.

A similar environment that allows for high resolution, low-temperature studies of the rotational and vibrational nature of small and isolated molecules is that of rare-gas matrices.[71] These environments are composed of spherical, rare-gas elements (the host) which act as inert material used to isolate certain guest molecules which may be highly reactive in other environments. The infrared and radio spectra obtained from these experiments lead to information not only about the isolated guest molecules but also the guest-host and host-host interactions. In addition to rare-gas elements, the unique and highly quantum material para-hydrogen (pH_2) can also be used to isolate guest molecules. pH_2 is spherically symmetric and so may act in the same

manner as rare-gas atoms while at the same time leading to more chemistry upon the cleavage of the H-H bond and thus the potential for hydrogen-tunneling to occur.[72] Such low-temperature, matrix isolated studies have been performed on carbon monoxide,[73] formic acid,[73,74] water,[74,75] ammonia,[76–78] Li-acetylene complexes,[79] and others.[80,81] Just as with studies involving diatomics identified in the ISM, a better understanding of the chemistry and physics of matrix isolated systems at low-temperatures is invaluable for the broader chemical body of knowledge.

Herein is discussed two systems relevant to low-temperature studies examined from a computational perspective: the gas-phase, CF^+ molecular cation which has been observed in the ISM and the study of H_2O in a pH_2 matrix as it relates to rotational perturbations to the antisymmetric stretch of H_2O.

2.2 Theoretical Rovibrational Line Lists of the CF^+ Molecular Cation Using High Accuracy Electronic Structure Theory

2.2.1 Abstract

The CF^+ molecule has been examined using high levels of electronic structure theory. Using coupled cluster (CC) with single, double, and perturbative triple excitations (CCSD(T)) at the complete basis set limit with corrections from full triple, quadruple, quintuple, and hextuple excitations within a finite-basis CC wavefunction, full configuration interaction (FCI), and relativistic effects, a potential energy curve (PEC) has been developed. Significant multireference character was identified in the electronic structure of CF^+ which prompted the need for such high-level corrections to the PEC. This curve has been used to populate a vibrational Hamiltonian matrix to examine the vibrational properties of CF^+ and with the inclusion of a centrifugal barrier, rotational and rovibrational properties were studied as well. The equilibrium bond length and vibrational constant were reproduced to within 0.00014 Å and 0.1526 cm^{-1}, respectively, of the experimental values. Experimental photon energies from rovibrational spectra were reproduced with an error no larger than 0.5 cm^{-1} which is excellent for such a fit. The triplet excited state ($a^3\Pi$)

was found to lie 4.8075 eV (38774.9 cm^{-1}) higher in energy with a slightly elongated bond length of 1.2106911 Å. High accuracy rovibrational line lists for the ^{12}C and ^{13}C isotopologues for both the $X^1\Sigma^+$ and the excited $a^3\Pi$ state were generated. Such a procedure allows for the reproduction of the spectroscopic constants and line lists of vibrational, rotational, and rovibrational excitations for other diatomic species.

2.2.2 Introduction

The observation and identification of spectral peaks collected from the interstellar medium (ISM) via radiospectroscopy is remarkably interesting as it allows researchers a view into a chemical environment unattainable on Earth. This environment encompasses very low temperatures and extremely low densities and as a consequence chemical reactions in the ISM only occur on very long time scales.[82] To date, close to 250 molecules[83] have been identified in the ISM ranging from diatomic molecules to fullerenes[84]. Of these, most are composed of hydrogen, carbon, nitrogen, oxygen, silicon, and/or sulfur[83] while other, more exotic atoms have been identified such as aluminum[85], phosphorus,[86] and iron.[87] Fluorine has been identified in three different molecular forms in the ISM and is most often found bound to hydrogen as HF. CF^+ (the second most common fluorine containing molecule found in the ISM) was first observed in 2006 by Neufeld *et al*.[88] and is thought to be produced via a reaction between HF and a C^+ cation. While the CF^+ molecule may play a small role in the chemistry of the ISM, it and other fluorocarbons have shown usefulness in chemical vapor deposition,[89] refrigerants,[90] and electronics manufacturing[91] and thus, a better understanding of their properties and fundamental chemistry may aid certain technological advances. Additionally, a theoretical understanding of CF^+ may allow for a better understanding of the different environments of the ISM which give rise to such exotic chemical species.

CF^+ has been widely studied both from experimental[2,92–96] and theoretical standpoints.[94,97–100] However, previous theoretical work has been limited in either scope or method. Peterson *et al*.[97] examined the spectroscopic properties of the ground state singlet ($X^1\Sigma^+$) and excited triplet state ($a^3\Pi$) of CF^+, SiF^+, and CCl^+ using MRCI and a relatively small basis set. Similar work done

by Petsalakis *et al.* [100] used MRDCI and a larger basis set to specifically study the electronic states of CF^+. More recent work done by Cazzoli *et al.*[94] examined the rotational spectra and the hyperfine splitting of $^{12}CF^+$ and $^{13}CF^+$ with very high levels of quantum chemistry while others have examined the vibrational nature[93], rovibrational nature,[99] and the hyperfine structure in the rotational spectrum of CF^+.[98] To date, we have not identified current literature which details rotational, vibrational, rovibrational, and electronic properties of the CF^+ cation together using high levels of electronic structure theory. Other studies have been performed on similar systems, but they have been focused on the vibrational dipole moments of XH^+ molecular ions (X = ^{24}Mg, ^{40}Ca, ^{64}Zn, ^{88}Sr, ^{114}Cd, ^{138}Ba, ^{174}Yb, and ^{202}Hg).[101] This work may serve as a guide for others who wish to study such diatomic molecules and ascertain detailed information regarding their rotational, vibrational, and rovibrational states. In addition, a limited study involving the ground and first excited electronic states of CF^+ is also presented.

In this work, we detail the creation and utilization of a vibrational Hamiltonian matrix which serves as the basis of all calculations. It is used to quantify the energy levels as well as the relative intensities of all excitations within a subset of the vibrational energies. Corrections due to centrifugal distortions have allowed us to examine the rotational and rovibrational states of CF^+ for both ^{12}C and ^{13}C isotopes for the singlet and triplet electronic states. This section is outlined in the following manner: mathematical and computational details are discussed first followed by results where a discussion is made regarding the electronic structure and anharmonic nature of the CF^+ molecule. This is then followed by simulated rotational line lists. A discussion of said results follows and then conclusions are made.

2.2.3 Computational and Mathematical Details

Potential Energy and Dipole Moment Curves

The coupled cluster (CC) method as implemented in MRCC[102,103] forms the basis of the calculations performed in the construction of the potential energy and dipole moment curves. We use the CC method from single and double excitations (CCSD) up to six-fold excitations (CCSDTQP6) using different basis sets depending upon the number of excitations included in the

CC expansion. At the core of our calculations, we utilized the CC method with single, double, and perturbative triple excitations (CCSD(T)) [104–106] with the cc-pCVnZ[107] (n=T, Q, 5) core-valence basis set extrapolated to the complete basis set limit according to Eq. 2.1 and 2.2 for the HF[108] and correlation energies,[109,110] respectively. The CBS extrapolation was performed using the first formulas of Eq. 2.1 and 2.2 for the ground state singlet while the second formulas were utilized for the excited triplet state due to an increase in computational cost. Corrections to this method include the effect of triple excitations from perturbative triples using a pentuple-ζ quality basis set (ΔT/5Z), full quadruples from full triples using a triple- ζ quality basis set (ΔQ/TZ), full pentuples from full quadruples using a double- ζ quality basis set (ΔP/DZ), and full hextuples from full pentuples using the 6-31G basis set (Δ6/6-31G). In addition, corrections from full configuration interaction (ΔFCI/6-31G) and relativistic effects using the second order Douglas-Kroll-Hess approximation (ΔDK2/QZ) as implemented in ORCA 5.0[111] were included. Each of these corrections is performed using the largest basis set given available computational resources. This composite scheme for the electronic energy is used to obtain the most accurate representation of the equilibrium structure and is best characterized as

$$\text{CCSD(T)/cc-pCV}\infty\text{Z} + \Delta\text{T/5Z} + \Delta\text{Q/TZ} + \Delta\text{P/DZ} + \Delta\text{6/6-31G} + \Delta\text{FCI/6-31G} + \Delta\text{DK2/QZ}.$$

In addition to the potential energy curve (PEC) for the electronic energy, an applied electric field was used to numerically calculate the dipole moment as a function of the internuclear distance. These calculations were performed with applied electric fields of 0.002 and -0.002 a.u. To calculate the dipole moment for each point on the PEC, the difference in the total electronic energy using the two electric field values is calculated and then divided by twice value of the field itself. This value is then converted from a.u. to D using a factor of 2.541. Both the diatomic molecule and the electric field were aligned with the z-axis. The dipole moment curve is calculated at the same level of theory as the potential energy curve except the relativistic corrections were not included as the inclusion of a finite electric field within a relativistic Hamiltonian cannot be performed in ORCA 5.0. The relativistic effects were small and on the order to 0.1% of the total energy. The center of mass of the CF^+ diatomic molecule was set as the origin for all calculations.

$$E_{HF}(CBS) = \frac{E(QZ)e^{1.62A} - E(5Z)e^{1.62B}}{e^{1.62A} - e^{1.62B}}, A = 3.87, B = 5.07$$

$$E_{HF}(CBS) = \frac{E(TZ)e^{1.62A} - E(QZ)e^{1.62B}}{e^{1.62A} - e^{1.62B}}, A = 2.96, B = 3.87 \quad (2.1)$$

$$E_{corr}(CBS) = E_{corr}(5Z) + 0.0078\frac{E(5Z)}{C}, C = 4.71$$

$$E_{corr}(CBS) = E_{corr}(QZ) + 0.0078\frac{E(QZ)}{C}, C = 3.68 \quad (2.2)$$

Dunham Coefficients and Rovibrational Constants

The PEC of a diatomic molecule can be fit in a Dunham-type manner[112] which takes the form

$$V(R) = hca_0(1 + a_1\xi + a_2\xi^2 + a_3\xi^3 + \cdots + a_n\xi^n), \xi = \frac{(R - R_e)}{R_e} \quad (2.3)$$

where h is the Planck constant, c is the speed of light in cm/s, $a_0 = \omega_e^2/4B_e$, ω_e is the vibrational constant in cm^{-1}, and B_e is the rotational constant in cm^{-1} and is equal to $h/[8\pi^2\mu R_e^2 c]$ (μ being the reduced mass and R_e being the equilibrium bond length). The coefficients for the polynomial fit (a_1, a_2, …a_n) can then be used to calculate Dunham parameters ($Y_{l,m}$) from which many spectroscopic constants of interest for the rovibrational states of a diatomic can be calculated.[112] No connection is made between traditional spherical harmonics and the $Y_{l,m}$ terms, however. The energy levels of the rovibrational states can be calculated using Eq. 2.4, below,

$$F_{vJ} = \sum_{l,m} Y_{l,m}\left(v + \frac{1}{2}\right)^{l}[J(J+1)]^{m} \quad (2.4)$$

where J and *v* represent the rotational and vibrational quantum numbers, respectively. The connection between the different Dunham coefficients and ordinary band spectrum constants is outlined in Table A2.1.[112] This methodology offers a simple way to compute many spectroscopic constants and compare directly with experimental values.

Vibrational Hamiltonian Matrix

To construct a vibrational Hamiltonian matrix, the potential energy curve is refit to Eq. 2.5,

$$V(x) = c_n x^n + c_{n-1}x^{n-1} \ldots + c_2 x^2, x = R - R_e \quad (2.5)$$

where R_e is the equilibrium bond length. From this polynomial fit, the c_2 coefficient can be used to calculate the harmonic force constant and from that, the vibrational constant. An anharmonic, vibrational Hamiltonian matrix can be populated given a specific c_n (n > 2) coefficient by evaluating Eq. 2.6 for every combination of *v* and *v*`. This matrix has dimensions *v* x *v* where *v*-1 is the maximum value used for the vibrational quantum number.

$$H_{anharm}[v, v`] = \sum_{2}^{n} c_n \langle \Psi_v | x^n | \Psi_{v'} \rangle \quad (2.6)$$

Here, c_n are the coefficients of the polynomial fit (Eq. 2.5), and Ψ_v and $\Psi_{v`}$ are the initial and final harmonic oscillator wave functions, respectively. The anharmonic, vibrational Hamiltonian matrix

of Eq. 2.6 is then added to a harmonic, vibrational Hamiltonian matrix constructed according to Eq. 2.7,

$$H_{harm}[v, v] = \frac{(2v + 1)\,\hbar\,\omega}{2} \tag{2.7}$$

where ω is vibrational constant (calculated as $\sqrt{\frac{2c_2}{\mu}} * (2\pi c)^{-1}$) and $\hbar$ is the reduced Planck constant. H_{harm} is only populated on the diagonal. The total, vibrational Hamiltonian matrix that is constructed as a sum of the anharmonic and harmonic matrices can then be diagonalized where the eigenvalues correspond to the vibrational energy levels of the non-rotating, diatomic molecule and the eigenvectors yield information on the contribution of each harmonic wavefunction to the total anharmonic wavefunction.

Centrifugal Barrier

To correct for the rigid rotor approximation for a diatomic molecule, a centrifugal barrier must be included to properly account for the perturbations due to the rotational motion. This is accomplished through the addition of a slightly modified form of Eq. 2.6 where a $1/R^2$ term is introduced (Eq. 2.8).[113]

$$H^J_{centrifugal}[v, v`] = \frac{\hbar^2\,J(J+1)}{2\mu}\left\langle \Psi_v \middle| \frac{1}{(x + R_e)^2} \middle| \Psi_{v'} \right\rangle \tag{2.8}$$

Here, *J* corresponds to the rotational quantum number. This matrix can then be added to those created using Eqs. 2.6 and 2.7 to yield a Hamiltonian matrix capable of capturing the pure vibrational, pure rotational, and coupled rovibrational nature of a diatomic molecule. The integrals calculated in Eq. 2.8 are truncated between $x = \pm 0.9R_e$.

Transition Dipole Moments

To calculate the transition dipole moment matrix elements of any transition $v`, j` \leftarrow v, j$, the eigenvectors discussed previously are used in conjunction with a polynomial fit (Eq. 2.9) for the dipole moment. The dipole moment was fit using a standard polynomial function to a high enough degree so as to minimize residual errors. A transition dipole moment matrix is constructed according to Eq. 2.10, below,

$$f(y) = a_0 + a_1 y^1 \ldots + a_n y^n \tag{2.9}$$

$$M[v, v`] = \sum_{\gamma=0}^{\gamma_{max}} a_\gamma \left\langle \Psi_v^J \middle| x^\gamma \middle| \Psi_{v'}^{J'} \right\rangle \tag{2.10}$$

where γ_{max} is defined as the highest polynomial term included in the dipole moment function, a_γ is the γth coefficient of the dipole moment function (Eq. 2.9). The transition dipole moment of an excitation is given according to Eq. 2.11, below, where ψ_v^J and $\psi_{v'}^{J'}$ are the anharmonic wave functions from Eq. 2.6 - 2.8.

$$T_{v'J'vj} = \left\langle \Psi_v^J \middle| M[v, v`] \middle| \Psi_{v'}^{J'} \right\rangle \tag{2.11}$$

The intensity (I) of a transition is related to the square of the transition dipole moment and may be given according to the Einstein A-coefficient (Eq. 2.12), below,

$$A_{v'J'vJ} = \frac{64\,\pi^4 \omega^3}{3h^4 c^3} \frac{S_J^{J'}\, T_{v'J'vJ}^2}{(4\pi\varepsilon_0)\, 2J' + 1} \tag{2.12}$$

where ω is the photon energy in J, ε_0 is the permittivity of free space, and $S_J^{J'}$ is the Hönl-London Factor.[114]

2.2.4 Results and Discussion

Electronic Structure of CF^+

In our investigation of the CF^+ molecule, we began by examining how the T1 diagnostic changes as the molecule is stretched. From an empirical standpoint, if a molecule shows a T1 diagnostic value that is fairly large (> 0.02), then it is said to contain significant nondynamical correlation effects.[115] This means that the molecule may show significant multireference character and a single-reference method such as traditional coupled cluster may fail to accurately describe the electronic structure of the molecule. We stretched the C-F bond from 0.9 Å to 2.0 Å in increments of 0.01 Å at the CCSD(T)/cc-pVQZ level of theory and examined the T1 diagnostic over this range of bond lengths. In the repulsive range of the potential energy curve (R_{C-F} < 1.15 Å), the T1 diagnostic smoothly transitions from 0.01 to 0.02. From this point on (R_{C-F} > 1.15 Å), the T1 values continue to increase until a maximum of 0.068 is reached at 1.87 Å. This indicates that a multireference method should be utilized for the CF^+ molecule everywhere beyond the equilibrium geometry.

Because of the large T1 diagnostic values we observed when stretching the CF^+ molecule, we chose to further examine the electronic structure by looking at the frontier molecular orbitals at the Hartree Fock level of theory and the determinantal contributions to the total electronic energy at the configuration interaction singles and doubles (CISD) level of theory. The occupied molecular orbitals (MOs) at the HF/cc-pVQZ level of theory were examined at six internuclear distances (R_{C-F} = 0.70, 1.00, 1.15, 1.35, 1.50, and 2.00 Å) so that we could visualize how the electronic structure changes at different points along the PEC. The frontier molecular orbitals (MOs) at R_{C-F} = 1.00, 1.15, and 1.35 Å are shown in Figure 2.1. At equilibrium (1.15 Å), the electronic structure of the CF^+ molecule is $1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^2 1\pi^4 5\sigma^2$ where the first two σ orbitals

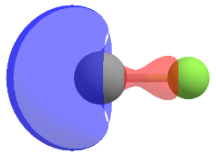
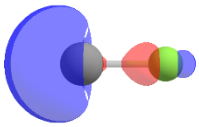
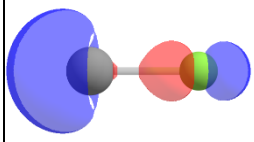
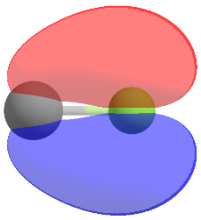
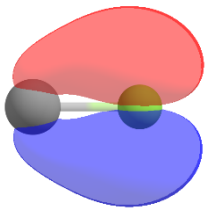
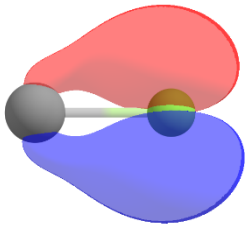
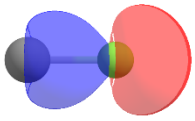
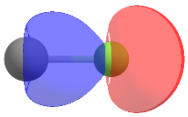
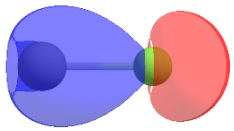
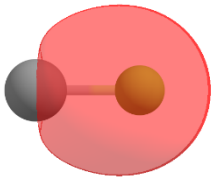
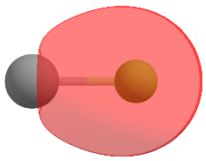
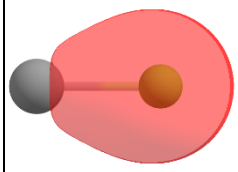
MO	R_{C-F} (Å)		
	1.00	1.15	1.35
$5\sigma^2$			
$1\pi^4$			
$4\sigma^2$			
$3\sigma^2$			

Figure 2.1 Frontier molecular orbitals (MO) for the CF^+ molecular cation as a function of the internuclear separation. Color code: C – black, F – green.

($1\sigma^2 2\sigma^2$) correspond to the F_{1s} and the C_{1s} atomic orbitals, respectively. These MOs are followed by an orbital ($3\sigma^2$) which changes slightly when the bond is stretched. At short distances, it is composed of a C_{sp} hybridized orbital and the F_{2s} atomic orbital and gradually becomes more polarized toward the F_{2s} orbital at large distances. In a similar fashion, the $4\sigma^2$ orbital is composed of the C_{2pz} and F_{2pz} atomic orbitals at short distances which persist until $R_{C\text{-}F}$ = 1.35 Å after which the orbital becomes almost entirely localized as the C_{2s} orbital at 2.0 Å. The next two orbitals ($1\pi^4$) exist as degenerate, π orbitals at 0.7 Å which appear to be bonding but change to be nonbonding orbitals composed of the F_{2px} and F_{2py} atomic orbitals. The HOMO of the CF^+ molecule ($5\sigma^2$) changes significantly as the bond is stretched and may be one of the orbitals responsible for the multireference character of the molecule. At $R_{C\text{-}F}$ = 0.70 Å, the $5\sigma^2$ orbital is composed of a C_{sp} hybridized atomic orbital that persists far past the equilibrium bond distance. From $R_{C\text{-}F}$ = 1.50 Å onwards, the sp character of the MO begins to resemble a combination of a $p_z - p_z$ orbital between C and F. The contribution from each atomic orbital is nearly equal and the behavior indicates that at even longer bond distances, this MO should become even more localized as the F_{2pz} atomic orbital. Examination of the electron density indicates that at $R_{C\text{-}F}$ = 2.00 Å, there is a node where no bonding between the C and F atoms exists. From this information, we can now ascertain that this MO will not become more localized on the F atom as the 2_{pz} atomic orbital and will instead persist as a combination of the F_{2pz} and C_{2pz} atomic orbitals at and beyond dissociation. If the C-F bond is cleaved homolytically, then this orbital would contain a single electron on each of the atoms. This orbital may then be either an antiferromagnetically coupled singlet where opposite spins are found on each nucleus or a paramagnetic triplet. This behavior may be due to an intersystem crossing where the triplet state becomes lower in energy than the singlet state.

With CISD, we observed that the contribution from the HF reference determinant decreases as the bond is stretched. The contributions from the HF determinant decrease from 95.5% at $R_{C\text{-}F}$ = 0.7 Å to 88.0% at $R_{C\text{-}F}$ = 2.0 Å which further indicates an increasing multireference character. As such, the inclusion of many higher order excitations (triple, quadruple, pentuple, and hextuple) as well as corrections from FCI should help to mitigate the limitations of using a single-reference method to describe the electronic structure of CF^+.

CF^+ Potential Energy Curve & Spectroscopic Constants

To construct the PEC of the CF^+ molecule, the C-F bond was stretched from 0.70 Å to 0.90 Å in increments of 0.05 Å, 0.95 Å to 1.50 Å in increments of 0.01 Å, and 1.55 Å to 2.00 Å in increments of 0.05 Å to give a total of seventy-one points on the curve. The curve for both the ground state singlet ($X^1\Sigma^+$) and excited triplet states ($a^3\Pi$) are shown in Figure 2.2, below. An 11^{th} degree polynomial was initially utilized to fit the ab initio PEC. The equilibrium bond length at the composite level of theory is 1.1543954 Å and differs from the experimental value by 0.00014 Å which is excellent for such a fit. Likewise, the harmonic stretching frequency (1792.8180 cm^{-1}) differs from the experimental value[2] by 0.1526 cm^{-1} when the polynomial fit based on Eq. 2.5 is used in conjunction with the equilibrium bond length. Thus, we feel confident that the PEC that has been developed can be utilized to examine various properties of the CF^+ molecule. For comparison, Table 2.1 lists both experimental and theoretical work done for CF^+, specifically, measurements or calculations of the equilibrium bond length, the vibrational constant, and the rotational constant. Previous experimental work performed by Gruebele *et al.*[2] and Cazzoli *et al*.[94] fit experimental rovibrational spectra to the so-called Dunham expression (Eq. 2.3) in order to determine Dunham parameters for the CF^+ molecule. The Dunham parameters[112] can be connected to ordinary band spectrum constants (Table A2.1) such as the harmonic frequency (ω_e) and the anharmonic corrections ($\omega_e x_e$, $\omega_e y_e$, $\omega_e z_e$,) and so direct comparisons between experimental and theoretical results can be performed with ease. It must be noted, however, that the work of Gruebele *et al.* and Cazzoli *et al.* calculated the Dunham parameters in different ways: Gruebele *et al*. fit rovibrational spectra to a Dunham expansion in terms of the equilibrium rotational constant B_e, the harmonic frequency, and the first six potential coefficients (a_1, a_2, … a_6) while Cazzoli *et al.* fit their transition energies directly to Eq. 2.4. For the current work, the PEC was fit to Eq. 2.3 in the same manner as Gruebele *et al.* which was then used to calculate the Dunham coefficients.

The PEC is quite exhaustive and thus should match experimental values quite well (see Table 2.2). To best compare against the work of Gruebele *et al.*, the PEC was truncated between $R_{C\text{-}F}$ = 1.00 and 1.40 Å as this range of bond lengths is sufficient to describe vibrational levels up to v = 7, which are the levels measured in the experimental spectrum obtained by Gruebele *et al*.

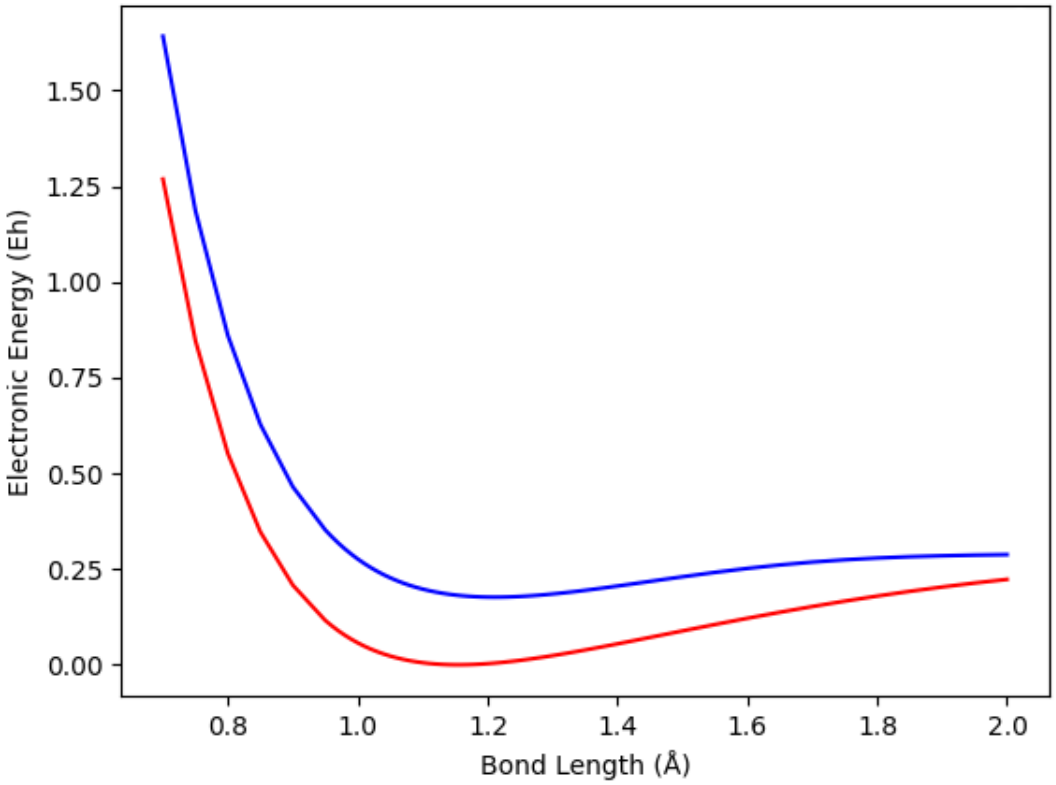


Figure 2.2 Potential energy curve (PEC) for the CF^+ molecule. The ground state singlet ($X^1\Sigma^+$) is shown in red and the excited triplet state ($a^3\Pi$) is shown in blue.

Table 2.1 Comparison of the vibrational constant (ω_e, cm^{-1}), rotational constant (B_e, cm^{-1}), and equilibrium bond length (R_e, Å) for CF^+ against experimental and other theorical work.

	Gruebele[2]	Cazzoli[94]	Peterson[97]	Inostroza[99]	CCSD[93]	Ne Matrix[93]	This Work
ω_e	1792.6654	-	1788.8	1789.5	1826.3	1759.99, 1762.9	1792.8180
B_e	1.7204176	1.72091207	1.7145	1.7101	1.714	-	1.720110
R_e	1.1542551	1.154089	1.1563	1.159	1.156	-	1.1543954

Table 2.2 Polynomial and Dunham coefficients based on Eq. 2.3 for the $^{12}CF^{+}$ isotopologue. Values given in cm^{-1}. The PEC for the current work was truncated between R = 1.00 Å and R = 1.40 Å.

	Cazzoli[94]	Inostroza[99]	Gruebele[2]	This Work
a_0	-	-	466986.800	466852.173
a_1	-	-	-2.90001	-2.89986
a_2	-	-	5.3862	5.4536
a_3	-	-	-7.754	-7.951
a_4	-	-	9.87	7.66
a_5	-	-	-14.93	-1.103
a_6	-	-	30.15	-1.437
B_e	1.72091207	1.7101358	1.7204176	1.720113
ω_e	-	1789.5	1792.6654	1792.8180
$Y_{1,0}$	1792.6712	1816.096	1792.6718	1792.077
$Y_{2,0}$	13.22958	13.73885	13.22968	13.047
$Y_{3,0}$	0.045260	-0.98131	0.045279	0.031704
$Y_{4,0}$	-0.000114	-	0.000115	-0.000696
$Y_{0,1}$	1.7203888	1.7101358	1.7204186	1.719822
$Y_{1,1}$	-0.0188210	-0.041985	0.0188228	0.0189564
$Y_{2,1}$	0.0000408	0.01285	0.0000407	0.0000537
$Y_{3,1}$	-	-0.002735	4.90E-07	-1.7395E-06
$Y_{0,2}$	-6.34E-06	-0.00247	6.34E-06	6.34E-06
$Y_{1,2}$	9.67E-09	0.00564	5.47E-09	1.043E-09
$Y_{2,2}$	-	-0.00346	5.82E-10	-5.917E-10
$Y_{0,3}$	-	2.30E-4	2.33E-12	2.337E-12
$Y_{1,3}$	-	-5.639E-4	6.92E-12	-3.710E-17

With regards to the polynomial coefficients, the a_1 term agrees with experiment to almost an exact degree. Likewise, the a_2 and a_3 terms agree quite well (within ~0.2 cm^{-1}) while the a_4 term differs by 2.0 cm^{-1}. The a_5 and a_6 terms do not show agreement, however, and differ by a significant degree. These terms may be responsible for the deviation between the theoretical and experimental work discussed below. The spectroscopic constants show very good agreement. The equilibrium rotational constant B_e agrees within ~0.0003 cm^{-1} while the harmonic frequency agrees within 0.153 cm^{-1} which is excellent for the theoretical calculation of spectroscopic constants. Of the Dunham Y coefficients, all agree with experiment within 0.60 cm^{-1}. Of those coefficients that are not exceedingly small ($Y_{1,0}$, $Y_{2,0}$, and $Y_{0,1}$), the difference between the experimental values and those calculated theoretically is at most 1.4% which is again in excellent agreement. These results indicate the quality of the PEC as it relates to the calculation of properties for the CF^+ molecule is excellent. As a final showcase of the quality of this PEC, Figure 2.3 shows the PEC as a function of the unitless stretching parameter ξ (See Eq. 2.3) for the current fit as well as the difference between the current fit and the fit of Gruebele *et al*.. Near equilibrium, the difference is essentially zero but at both the largest and smallest bond distances, the difference becomes more significant. At the repulsive region of the PEC, the difference is up 40 cm^{-1} which while small, is still significant for such high accuracy calculations but amounts to an error of 0.31%. Likewise for the attractive portion of the PEC, the difference is slightly smaller, on the order of 20 cm^{-1} (0.25%). At this region of the potential energy curve, effects not captured in the EST calculations may become significant enough to cause such deviation from experimental values. These effects could be due to the finite basis sets chosen, the lack of a multireference ansatz, an absence of rotational corrections, or the experimental values at these regions of the PEC may not be well defined. In the regions of the PEC far from equilibrium, anharmonic effects become more prevalent; these are discussed below and help to detail the peculiar nature of the CF^+ molecule at large bond distances. Shown in Table 2.3 is this same comparison of the current work for the different spin states and isotopologues. This shows how the isotopic substitution does little to change the spectroscopic constants while the excited triplet shows a large deviation between it and the ground state.

Another way the currently constructed PEC can be compared to experimental and other theoretical work is in the rovibrational energy levels and transition energies between levels. Shown

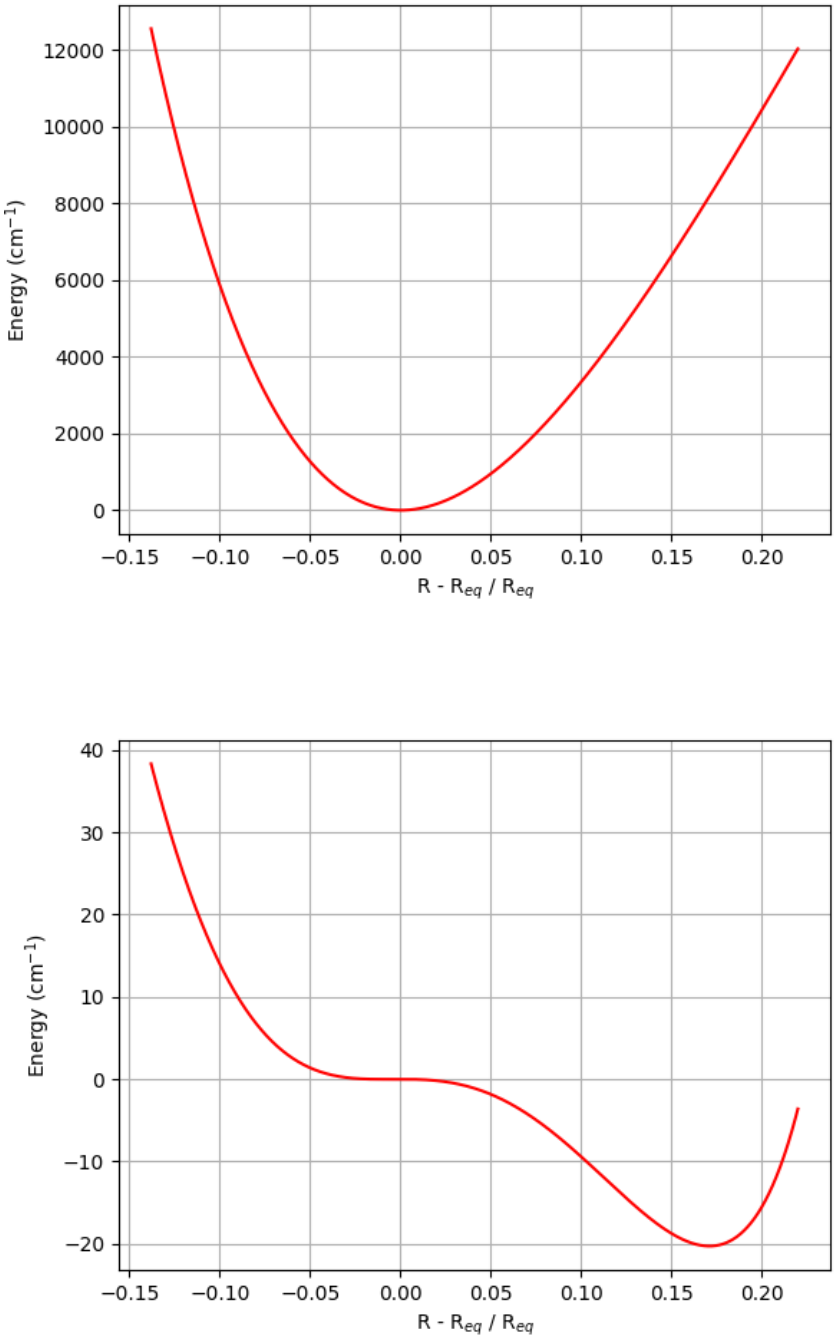


Figure 2.3 (Top) Potential energy curve (PEC) for the $^{12}CF^{+}$ molecule between R = 1.00 Å and R = 1.40 Å. (Bottom) The difference between the PEC computed here and the work of Gruebele *et al.*[2]

Table 2.3 Polynomial and Dunham coefficients based on Eq. 2.2.3 for the different spin states and isotopologues of CF^+. Values given in cm^{-1}.

	$^{12}CF^+$		$^{13}CF^+$	
	$X^1\Sigma^+$	$a^3\Pi$	$X^1\Sigma^+$	$a^3\Pi$
a_0	466852.173	415256.67751	466852.173483	415256.67751
a_1	-2.89986	-3.039681	-2.899861	-3.039681
a_2	5.4536	5.205759	5.453577	5.205759
a_3	-7.951	-7.727011	-7.951653	-7.727009
a_4	7.66	12.028556	7.655127	12.028596
a_5	-1.103	-11.583268	-1.09475	-11.583323
a_6	-1.437	-29.568424	-1.465556	-29.569402
B_e	1.720113	1.563835	1.629984	1.481214
ω_e	1792.8180	1611.467543	1749.912716	1572.902258
$Y_{1,0}$	1792.077	1611.430	1749.902	1572.867
$Y_{2,0}$	13.047	14.995780	12.319460	14.286620
$Y_{3,0}$	0.031704	-0.119588	0.002681	-0.111206
$Y_{4,0}$	-0.000696	-0.002089	0.000545	-0.001896
$Y_{0,1}$	1.719822	1.563740	1.638653	1.489790
$Y_{1,1}$	0.0189564	0.018621	0.017315	0.017310
$Y_{2,1}$	0.0000537	-1.68E-04	5.87E-06	-1.52E-04
$Y_{3,1}$	-1.7395E-06	-5.92E-06	-1.73E-07	-5.24E-06
$Y_{0,2}$	6.34E-06	5.89E-06	5.75E-06	5.35E-06
$Y_{1,2}$	1.043E-09	-5.51E-08	5.92E-09	-4.88E-08
$Y_{2,2}$	-5.917E-10	-5.10E-09	-7.99E-10	-4.41E-09
$Y_{0,3}$	2.337E-12	-6.62E-13	2.67E-12	-5.72E-13
$Y_{1,3}$	-3.710E-17	-1.48E-12	-2.83E-13	-1.25E-12

in Table 2.4 are the rotational energy levels up to J = 30 on the v = 0 surface compared to the experimental work of Gruebele[2] and the theoretical work of Inostroza.[99] The lowest lying energy level at J = 0 shows a deviation of ~0.1 cm^{-1} which decreases in magnitude until a value of J = 19 is reached. From there, the error becomes more negative and increases in magnitude up to J = 30. The error never exceeds a value of -0.14 cm^{-1} however, which indicates the high quality of the PEC. For the vibrational energy levels in Table 2.5, the error between the computed and experimental ground state energy is quite small. This error increases substantially, however, as higher energy levels are examined. This is due to the incompleteness of the vibrational Hamiltonian Matrix (Eqs. 2.5 – 2.8) which was truncated up to a value of v = 37. Values beyond this point cannot be evaluated due to a large amount of numerical noise present that arises from the evaluation of integrals (Eq. 2.6). The comparison between the transition energies between rovibrational states is a final indication of the high quality of the PEC. In Table 2.6, the largest error between the rovibrational transition energies of Gruebele and the current work is 0.447 cm^{-1}. The average error for the seventeen chosen rovibrational transition energies is 0.146 cm^{-1}. Thus, the possible transition energies that could be observed for a gas-phase CF^+ diatomic molecule could be reproduced with minimal error up to a vibrational state of v = 7.

An additional facet that is available due to our examination of many different rotational and vibrational states is how the spectroscopic constants change for different states. Shown in Figures 2.4a and 2.4b are plots of the vibrational constant (ω_e) and the first anharmonic correction term ($\omega_e x_e$) for J = 0 to J = 10, respectively, for both isotopologues and spin states. Likewise, Figures 2.4c and 2.4d show the rotational constant (B_e) and the first centrifugal distortion term (D_e) for v = 0 to v = 10, respectively, again for both isotopologues and spin states. Both the vibrational constant and the first anharmonic correction term show non-linear behavior in regard to the value as a function of the rotational quantum number J. Both values increase in magnitude as a function of J which indicates that the potential is becoming narrower (which leads the vibrational constant to increase) and more anharmonic (which increases the first anharmonic term). Near identical behavior is observed between the isotopologues and the different spin states for the vibrational constant but differences are observed for the anharmonic term. Of the four species, the $^{12}CF^+$ isotopologue on the $X^1\Sigma^+$ ground state surface shows the steepest increase in the anharmonic

Table 2.4 Rotational energy levels on the ground vibrational surface (v = 0). O-C indicates the difference between the observed values of Gruebele and the calculated values in the current work using Eq. 2.5. All values given in cm^{-1}.

J	Gruebele[2]	Inostroza[99]	This work (Eq. 2.5)	O-C
0	893.344	903.6	893.451	0.107
1	896.766	907.0	896.872	0.106
2	903.609	913.7	903.715	0.106
3	913.874	923.9	913.978	0.104
4	927.56	937.4	927.662	0.102
5	944.667	954.3	944.766	0.099
6	965.192	974.6	965.288	0.096
7	989.137	998.2	989.229	0.092
8	1016.499	1025.3	1016.587	0.088
9	1047.278	1055.7	1047.361	0.083
10	1081.471	-	1081.549	0.078
11	1119.078	-	1119.150	0.072
12	1160.097	-	1160.163	0.066
13	1204.526	-	1204.585	0.059
14	1252.363	-	1252.415	0.052
15	1303.606	-	1303.650	0.044
16	1358.253	-	1358.288	0.035
17	1416.3	-	1416.326	0.026
18	1477.747	-	1477.763	0.016
19	1542.589	-	1542.596	0.006
20	1610.824	-	1610.820	-0.004
21	1682.449	-	1682.434	-0.015
22	1757.461	-	1757.435	-0.026
23	1835.856	-	1835.818	-0.038
24	1917.632	-	1917.581	-0.051
25	2002.783	-	2002.719	-0.064
26	2091.307	-	2091.230	-0.077
27	2183.2	-	2183.108	-0.092
28	2278.457	-	2278.351	-0.106
29	2377.074	-	2376.953	-0.121
30	2479.047	-	2478.910	-0.137

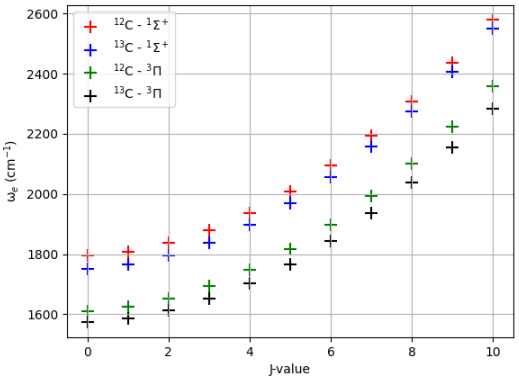
Table 2.5 Vibrational energy levels on the ground rotational surface (J = 0). O-C indicates the difference between the observed values of Gruebele and the calculated values in the current work using Eq. 2.5. All values given in cm^{-1}.

v	Gruebele[2]	Inostroza[99]	This work (Eq. 2.5)	O-C
0	893.344	903.6	893.451	0.107
1	2659.702	2693.1	2659.918	0.216
2	4400.005	4451.5	4400.252	0.247
3	6114.519	6179.0	6114.800	0.281
4	7803.506	7876.3	7803.886	0.380
5	9467.228	9543.5	9467.810	0.582
6	11105.941	11181.2	11106.850	0.909
7	12719.901	127898	12721.261	1.360
8	14309.36	14369.8	14311.277	1.917
9	15874.568	15921.6	15877.116	2.548
10	17415.769	-	17418.983	3.214
11	18933.209	-	18937.108	3.899
12	20427.128	-	20432.103	4.975
13	21897.764	-	21907.296	9.532
14	23345.353	-	23377.226	31.873
15	24770.126	-	24878.836	108.710
16	26172.314	-	26461.015	288.701
17	27552.144	-	28155.696	603.552
18	28909.838	-	29971.661	1061.823
19	30245.62	-	31907.408	1661.788
20	31559.706	-	33960.289	2400.583

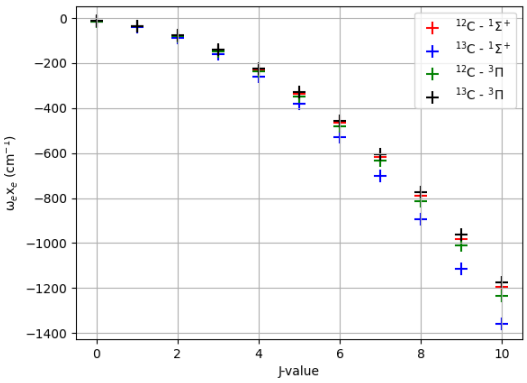
Table 2.6 Selected rovibrational excitations. O-C indicates the difference between the observed values of Gruebele and the calculated values in the current work using Eq. 2.5. All values given in cm^{-1}.

v	v'	J	J’	Inostroza[99]	Gruebele[2]	This Work (Eq. 2.5)	O-C
0	17	1	16	-	1703.21	1703.318	0.108
0	12	1	11	-	1722.863	1722.974	0.111
0	15	1	16	-	1815.911	1815.995	0.084
0	19	1	20	-	1826.728	1826.801	0.073
1	13	2	12	-	1693.453	1693.488	0.035
1	8	2	7	1730.6	1712.195	1712.231	0.036
1	16	2	17	-	1792.012	1792.020	0.008
2	6	3	5	1707.3	1693.882	1693.918	0.036
2	11	3	12	-	1751.74	1751.766	0.026
3	19	4	18	-	1619.954	1620.075	0.121
3	14	4	13	-	1639.357	1639.468	0.111
3	22	4	23	-	1754.62	1754.715	0.095
4	13	5	14	-	1705.62	1705.819	0.199
5	12	6	11	-	1597.509	1597.849	0.340
5	8	6	7	1611.1	1611.813	1612.146	0.333
5	6	6	5	1617.9	1618.757	1619.084	0.327
6	1	7	2	1614.7	1620.253	1620.700	0.447

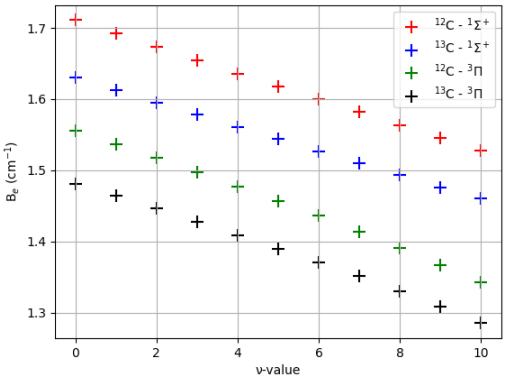
A



B



C



D

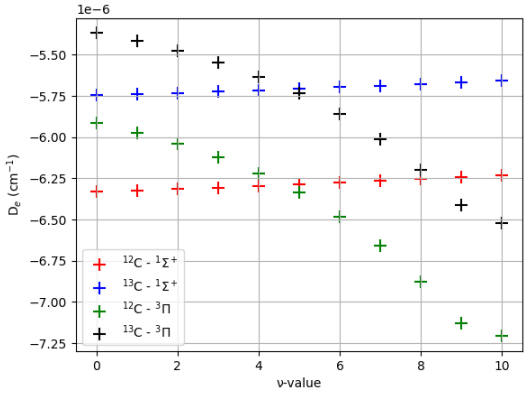


Figure 2.4 Plots of the vibrational constant (A) and first anharmonic correction term (B) as a function of rotational quantum number J and plots of the rotational constant (C) and first centrifugal distortion term (D) as a function of the vibrational quantum number *v*.

term while the $^{13}CF^+$ isotopologue on the $a^3\Pi^+$ excited state surface shows the smallest increase in the same term. This behavior may arise from an interplay between the shape of each respective PEC and the small differences observed in the isotopic labelling. The rotational constant shows a nearly linear behavior as a function of the chosen vibrational state which arises mainly due to the longer bond distances associated with higher vibrational states. The first centrifugal distortion terms show interesting behavior not observed with the other spectroscopic constants. The D_e values on the $X^1\Sigma^+$ ground state surface show parallel behavior where they decrease in magnitude in a more linear fashion as compared to the $a^3\Pi^+$ excited state surfaces. The latter increase in magnitude in non-linear fashions. Because the isotopologues change in parallel manners between the two spin state surfaces, the unique behavior of the D_e values is most likely an effect of the PEC itself. The unique curvature of the $X^1\Sigma^+$ and $a^3\Pi^+$ states result in vastly different behaviors for the D_e terms.

Vibrational Energy Levels and the Effect of Anharmonicity

Using the PEC derived from Eq. 2.5 up to n = 10, we were able to generate a vibrational Hamiltonian matrix and examine it with respect to the contribution from each harmonic wave function to the first eleven anharmonic states (Figure 2.5). The ground state energy level (left-most column of Figure 2.5) includes contributions almost entirely from the v = 0 harmonic oscillator (HO) wave function which is expected. The HO approximation is a good approximation only near the equilibrium geometry at the bottom of the potential. Further from the equilibrium geometry, anharmonic effects become more pronounced and the HO approximation breaks down and corrections must be included such as those discussed previously. Moving to higher CF+ vibrational levels (successive columns to the right) in Figure 2.5, the diagonal terms (the HO terms) become lighter in color as more and more contributions are included from other HO wave functions. Some symmetry between states can be observed where some wave functions include mostly even (or odd) states but this is not perfectly conserved in the anharmonic wave functions. The energy level diagram of the CF^+ molecule is plotted alongside the PEC in Figure 2.6. The vibrational Hamiltonian matrix is converged up to the 13^{th} level within ~1.0 cm^{-1} and so this is the highest level that should be examined with only minimal error in the approximation. Beyond this point, the energy levels have an error associated with the finite size of the vibrational Hamiltonian matrix.

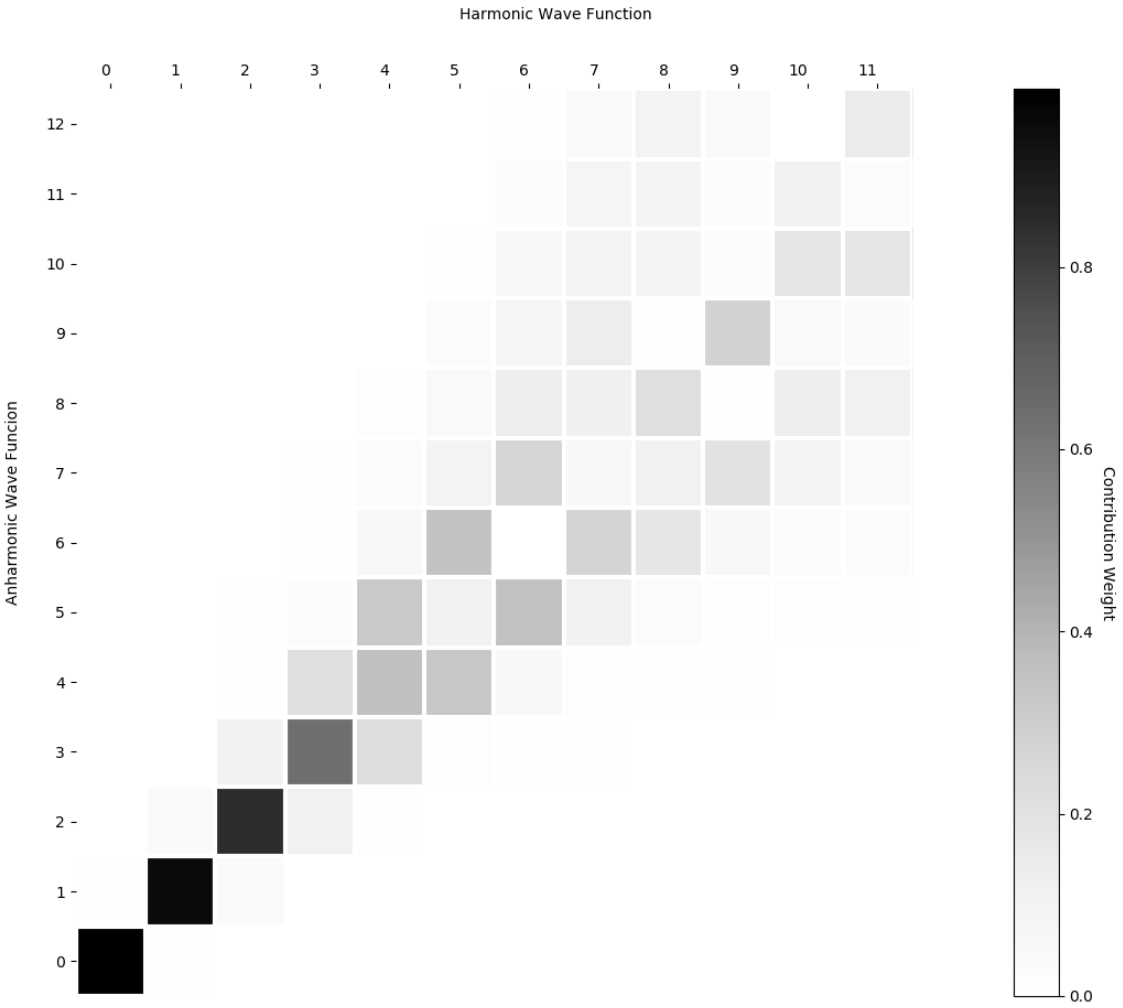


Figure 2.5 Anharmonic effects of the $^{12}CF^{+}$ molecule. Each anharmonic wave function is shown as a given row and each entry in that row is the contribution from a given harmonic oscillator wave function.

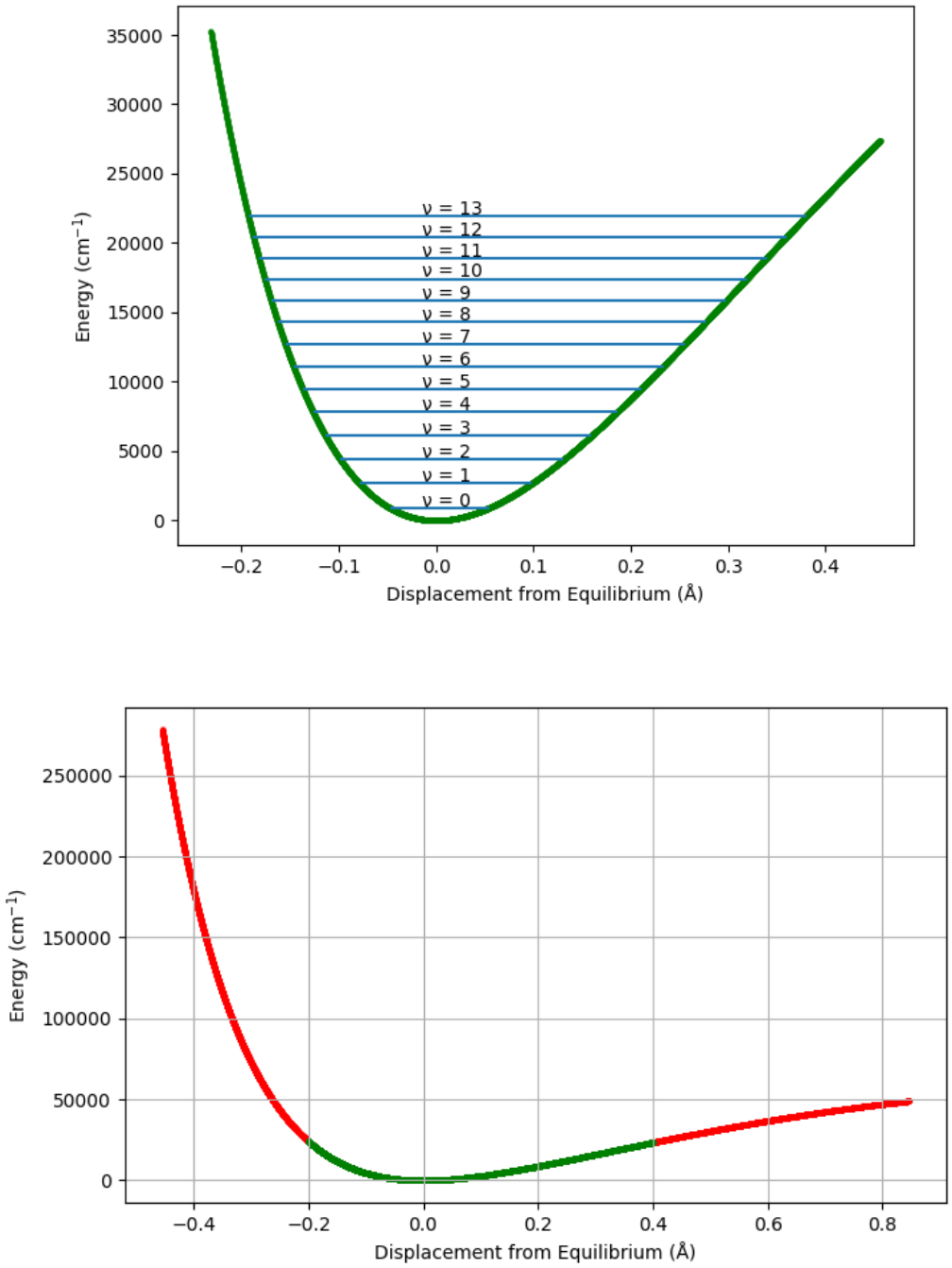


Figure 2.6 (Top) The PEC as a function of displacement up to the 13th excited state. (Bottom) The PEC in red with the region included in the upper plot in green.

Simulated Intensities - Einstein A Coefficients

Following the generation of the fully anharmonic, vibrational Hamiltonian matrix (Eqs. 2.6 – 2.7), the polynomial fit for the dipole moment function (Figure 2.7), as well as the transition dipole moment matrix (Eq. 2.10), transition dipole moments can be calculated for any vibrational, rotational, or rovibrational excitation. Each transition dipole moment is calculated according to Eq. 2.11 which can be combined with Eq. 2.12 to yield the Einstein A coefficient for spontaneous emission for a given de-excitation. Work by Peterson *et al.*[97] calculated the Einstein A coefficients for the singlet ground state and the triplet excited state of CF^+ using single and double excitations within the MRCI level of theory and that work may serve as a comparison of the current results. Shown in Table 2.7 are values obtained by Peterson *et al*. and values from the current work for both the $^{12}CF^+$ and $^{13}CF^+$ isotopologues for the first five $v \leftarrow v + 1$ transitions. Likewise, Table 2.8 shows the first five $v \leftarrow v + 2$ transitions and Table 2.9 shows the first five $v \leftarrow v + 3$ transitions. Good agreement is made between the work of Peterson *et al*. and the current work for both the $X^1\Sigma^+$ and the excited $a^3\Pi$ state. Differences between the two may be due to the different underlying PECs and the different EST methods which were used to construct them or it may be due to differences in the generation of the anharmonic Hamiltonian matrix as Peterson *et al*. utilized the method of Cooley.[116] An interesting thing to note is the effects that the size of the vibrational constants and the magnitude of the dipole moment functions (Figure 2.7) have with regards to the magnitude of the intensity of each vibrational excitation. The $X^1\Sigma^+$ state has a larger vibrational constant and thus larger transition energies, which are cubed in Eq. 2.12 when calculating the Einstein A coefficients, but also shows a dipole moment function that is generally smaller in magnitude than the excited $a^3\Pi$ state. Thus, the magnitude of the vibrational excitations will be sensitive to these factors which are evident in Tables 2.7 - 2.9. The $X^1\Sigma^+$ state shows larger Einstein A coefficients for the $v \leftarrow v + 1$ transitions while the excited $a^3\Pi$ state shows larger coefficient values for the $v \leftarrow v + 2$ and $v \leftarrow v + 3$ transitions. While the latter transitions are much smaller in intensity, this behavior is still telling of how the different electronic structures of the ground and excited state would result in different vibrational spectra.

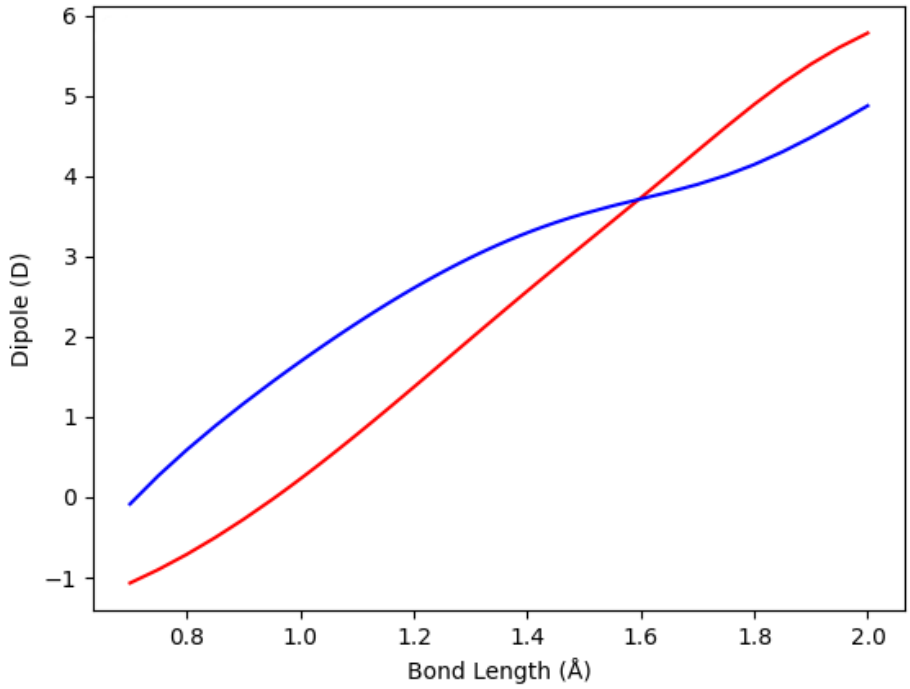


Figure 2.7 Dipole moment curves for the $X^1\Sigma^+$ ground state (red) and the excited $a^3\Pi$ state (blue) of $^{12}CF^+$.

Table 2.7 Calculated Einstein A coefficients of spontaneous emission (in s^{-1}) for the first five $v \leftarrow v + 1$ transitions.

v	A_v^{v+1}					
	$X^1\Sigma^+$			$a^3\Pi$		
	$^{12}CF^+$		$^{13}CF^+$	$^{12}CF^+$		$^{13}CF^+$
	Peterson	This Work	This Work	Peterson	This Work	This Work
0	75.05	77.57	72.69	29.08	28.13	26.78
1	144.6	149.5	140.2	53.76	51.85	49.50
2	208.7	215.9	202.7	74.48	71.35	68.34
3	267.7	277.0	260.3	91.43	86.87	83.49
4	321.9	333.0	313.2	104.8	98.62	95.15
5	371.2	384.1	361.6	114.7	106.8	103.5

Table 2.8 Calculated Einstein A coefficients of spontaneous emission (in s^{-1}) for the first five $v \leftarrow v + 2$ transitions.

v	A_v^{v+2}					
	$X^1\Sigma^+$			$a^3\Pi$		
	$^{12}CF^+$		$^{13}CF^+$	$^{12}CF^+$		$^{13}CF^+$
	Peterson	This Work	This Work	Peterson	This Work	This Work
0	1.729	1.839	1.690	2.528	2.509	2.292
1	5.286	5.450	5.006	7.234	7.271	6.649
2	10.57	10.75	9.874	13.78	14.00	12.82
3	17.31	17.64	16.20	21.88	22.39	20.53
4	25.37	26.00	23.90	31.24	32.07	29.46
5	34.71	35.73	3.284	41.55	42.63	39.26

Table 2.9 Calculated Einstein A coefficients of spontaneous emission (in s^{-1}) for the first five $v \leftarrow v + 3$ transitions.

v	A_v^{v+3}					
	$X^1\Sigma^+$			$a^3\Pi$		
	$^{12}CF^+$		$^{13}CF^+$	$^{12}CF^+$		$^{13}CF^+$
	Peterson	This Work	This Work	Peterson	This Work	This Work
0	0.0233	0.02704	0.02475	0.1032	0.1010	0.0899
1	0.1036	0.1108	0.1012	0.4204	0.4037	0.359
2	0.2682	0.2835	0.2584	1.052	1.0123	0.9004
3	0.5331	0.5808	0.5281	2.099	2.036	1.810
4	0.8988	1.041	0.9441	3.650	3.593	3.191
5	1.384	1.704	1.543	5.787	5.808	5.154

Simulated Rotational Spectra

Through a combination of the Einstein A coefficients of spontaneous emission and statistical mechanics, simulated pure rotational spectra at any given temperature can be generated. The partition function (q) for the rigid rotor approximation is given by Eq. 2.13, below,

$$q_{rot} = \frac{1}{\sigma}\sum_{J=0}^{J=\infty}(2J+1)e^{-yJ(J+1)}\ , y = {\hbar^2}/{2I_e kT} \tag{2.13}$$

where σ is the symmetry number (1 for heteronuclear and 2 for homonuclear diatomic molecules), J is the rotational constant, $\hbar$ is the reduced Planck's constant, I_e is the moment of inertia, k is the Boltzmann constant, and T is the temperature in kelvin. If y is small, this sum can be approximated as an integral given in Eq. 2.14 which can be further simplified to Eq. 2.15.

$$q_{rot} = \frac{1}{\sigma}\int_{J=0}^{J=\infty}(2J+1)e^{-yJ(J+1)}dJ = \frac{1}{\sigma y} \tag{2.14}$$

$$q_{rot} = \frac{1}{\sigma}\frac{T}{1.4388\ B_e} \tag{2.15}$$

The intensity of a transition at a given temperature is then given as the product of the population of the initial rotational state and the Einstein A coefficient.

Simulated rotational spectra at 30K (a representative temperature of the ISM) and 650K (the rotational temperature determined by Gruebele *et al.*[2]) are shown in Figure 2.8. Here, the

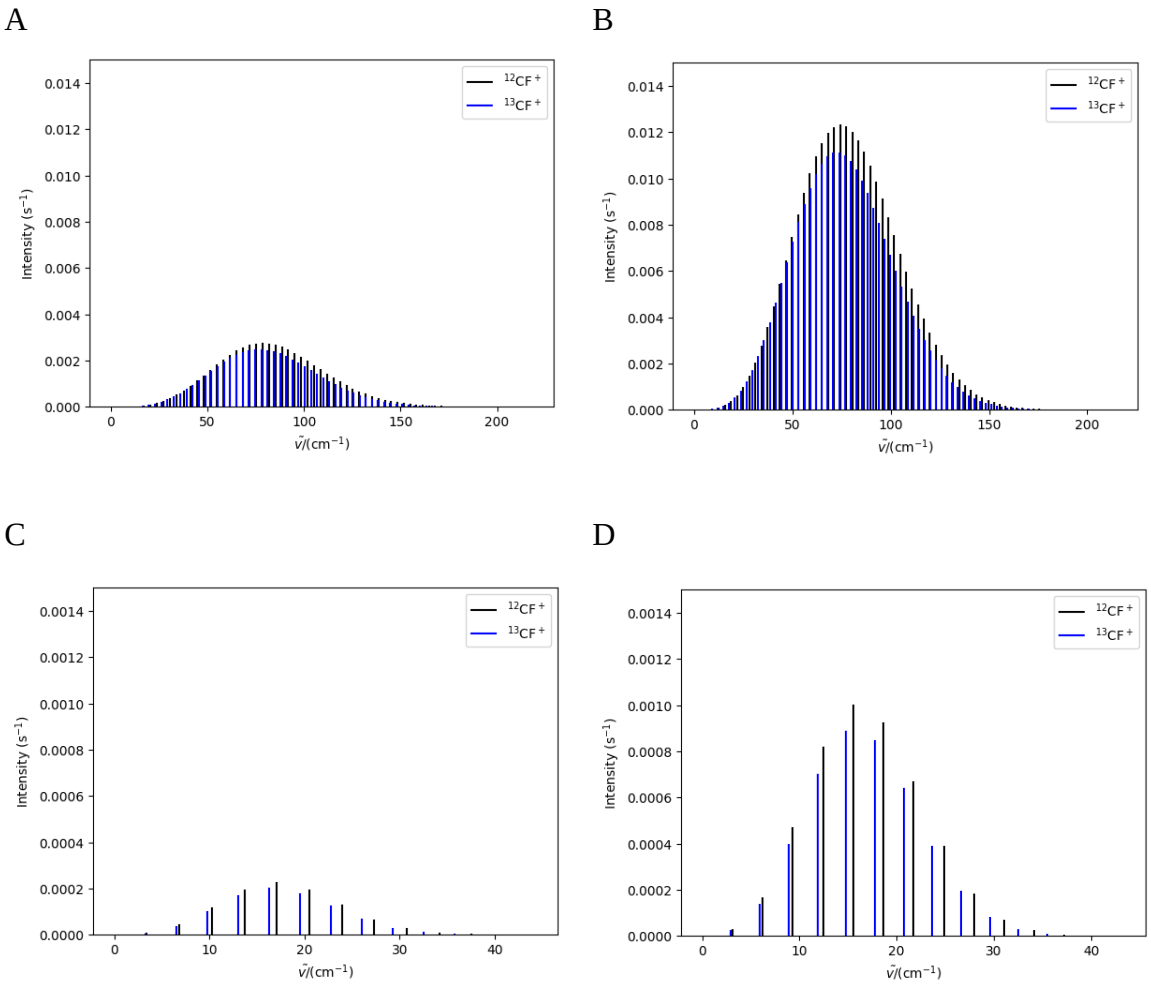


Figure 2.8 Simulated rotational spectra for the $X^1\Sigma^+$ state at A) 650K and C) 30K and the excited $a^3\Pi$ state at B) 650K and D) 30K for both the $^{12}CF^+$ (black lines) and $^{13}CF^+$ (blue lines) isotopologues.

excited triplet state results in higher intensities than the ground state singlet. Likewise, the $^{12}CF^+$ isotopologue shows slightly higher intensities than $^{13}CF^+$ at both simulated temperatures.

2.2.5 Conclusions

Quantum chemical calculations were utilized in order to generate a potential energy curve (PEC) in the gas phase and under the perturbation of applied electric fields in order to calculate the dipole moment function for the CF^+ cation. These functions were then used as the starting point in generating a vibrational Hamiltonian matrix both with and without the addition of a centrifugal barrier to account for rotational distortions. Analysis of the resulting vibrational energy levels showed very good agreement with previous experimental literature. The PEC results in an equilibrium bond length and fundamental stretching frequency which differ from experimental values by 0.00014 Å and 0.1526 cm^{-1}, respectively. Comparison of calculated vibrational Einstein A coefficients of spontaneous emission with those of previous theoretical work show very good agreement with errors less than 0.5 s^{-1}. Simulated rotational spectra were generated at 30K and 650K for both the ground state singlet ($X^1\Sigma^+$) and excited triplet states ($a^3\Pi$) for both the $^{12}CF^+$ and $^{13}CF^+$ isotopologues. This work may serve as a starting point for others who wish to study other diatomics to a high degree of accuracy.

2.3 Rotational Excitations of H_2O in a pH_2 Matrix

2.3.1 Abstract

High-resolution, infrared spectra of water molecules in low-temperature, parahydrogen (pH_2) matrices showed two satellite peaks directly around the antisymmetric stretch of water shifted by 1-2 cm^{-1}. Through the construction of three potential energy curves and two potential energy surfaces, dynamics simulations of the rovibrational nature of a H_2O molecule in different environments was used to identify these satellite peaks. Three chemical environments were examined: a pristine, hexagonal close-packed (HCP) solid composed of twelve pH_2 molecules, an

HCP solid in which one site has been replaced with a H-atom, and an HCP solid in which one site has been left vacant. These three environments have been used to quantify each of their respective effects on the antisymmetric stretch of H_2O accompanied by a rotational excitation. Dynamical simulations indicate that the effect of a H-atom substitution in the HCP solid directly around the H_2O molecule is not large enough to lead to rotational perturbations of 1-2 cm^{-1}. Upon the removal of one of the pH_2 sites, however, two perturbations are observed which correspond closely to the observed, satellite peaks.

2.3.2 Introduction

Low-temperature, rare-gas matrices (hosts) are a unique environment that allow for the isolation of guest molecules from one another, in a manner similar to the gas-phase, while also imposing weak perturbations on said guest molecules.[71] High resolution infrared spectra may be taken of the guest molecules embedded in the matrix and host-host and host-guest interactions may be quantified. The host-guest interactions are of great interest as the host may perturb the nature of the guest molecule in ways that are not possible to replicate through other means. An example of this are satellite peaks that are present in the infrared spectra of certain guest-doped parahydrogen (pH_2) matrices.

Kufeld *et al.*[75] examined a water-doped, pH_2 matrix and used high resolution infrared spectroscopy to analyze the interaction. Their work began with the UV irradiation of formic acid in the pH_2 matrix which, through photodissociation, leads to the formation of radical species such as oxygen atoms which in turn can react with the surrounding pH_2 molecules. This then leads to a cascade reaction where hydrogen atoms may travel through matrix via hydrogen tunneling, ultimately leading to the formation of isolated H_2O in the matrix. High resolution infrared spectra are taken during the entire process to evaluate the changes in the matrix as a function of time. The antisymmetric stretch of H_2O (at 3765.49 cm^{-1}) is initially accompanied by two satellite peaks: one blue-shifted by 2.03 cm^{-1} and one red-shifted by 1.21 cm^{-1}. These peaks show the greatest intensities immediately following photodissociation and decrease in intensity as time progresses. Kufeld *et al*. hypothesized that these peaks are a result of defects in the pH_2 matrix immediately

around the H_2O molecule; these defects are thought to perturb the rovibrational states of H_2O slightly. The proposed defects may either be a vacancy in which the H_2O molecule is only surrounded by eleven pH_2 molecules in the hexagonal closed-packed solid or a substitution for one of the pH_2 molecules such as a H-atom. Using a combination of electronic structure theory and dynamics, one can examine the rovibrational excitations of H_2O in each of these three environments to determine where the satellite peaks may arise from.

2.3.3 Potential Energy Curves and Surfaces

In order to examine the rovibrational excitations of a water molecule in a pH_2 matrix, two potential energy surfaces (PES) and three potential energy curves (PEC) must be constructed. The H_2O-H and H_2O-pH_2 surfaces are needed during the dynamical calculations themselves while the H_2, H-pH_2, and pH_2-pH_2 curves are needed prior to but not explicitly during the dynamical calculations. The surfaces are constructed in three-dimensional space which describes the exact position of either a H-atom or a pH_2 molecule relative to a H_2O molecule whereas the three curves are constructed in a one-dimensional space due to the spherical symmetry of both the H-atom and the pH_2 molecule. The details for each surface and curve are described below. MRCC[102,117] (version 2020-02-22) is used to construct the three curves while ORCA[118] (version 4.2) is used for the two surfaces.

1-Dimensional Curves

H_2 Curve

The H-H curve of molecular hydrogen is the simplest curve constructed for this work. As there are only two electrons in this system, the FCI limit can be reached using the coupled cluster (CC) approach including single and double excitations (CCSD). Likewise, an expansion in the one electron basis can be accomplished to approach the complete basis set (CBS) limit using very large basis sets. The H-H bond was stretched from 0.4 Å to 2.0 Å in increments of 0.005 Å using both the aug-cc-pVQZ (ACCQ) and aug-cc-pV5Z (ACC5) basis sets. The CBS potential energy curve

was constructed using an extrapolation scheme for both the Hartree Fock (HF)[119] and correlation energies[120] at the CCSD level of theory using Eqs. 2.16 and 2.17, respectively. The CBS curve results in an equilibrium bond distance of 0.7417642 Å and a vibrational constant of 4393.80995 cm^{-1}. A comparison of theoretical values and experimental values for H_2 are shown in Table 2.10. An excellent reproduction when compared to experiment is achieved and going forward, calculations which utilize an H_2 molecule will be performed with a bond length of 0.74176 Å.

$$E_{CBS}^{HF} = E_{5Z}^{HF} + A\left[E_{5Z}^{HF} - E_{qZ}^{HF}\right], A = 0.165995 \quad (2.16)$$

$$E_{CBS}^{corr} = E_{5Z}^{corr} + B\left[E_{5Z}^{corr} - E_{QZ}^{corr}\right], B = 0.888601 \quad (2.17)$$

H-pH_2 Curve

The construction of the potential energy curve for the H-pH_2 van der Waals dimer is made both more and less complicated due to the spherical symmetry of the pH_2 molecule. A similar van der Waals dimer (He-pH_2) has been examined both experimentally and theoretically.[121] In contrast to that work, an approximation is made whereby discrete H_2 orientations in space are utilized and then spherically averaged to best replicate the unique, spherical symmetry of a pH_2 molecule. A Lebedev quadrature scheme was utilized with fourteen points on the sphere.[122] These fourteen points correspond to seven unique orientations of the H_2 molecule in space. Due to the symmetry of the H-pH_2 interaction, these seven orientations can be further reduced to three unique orientations which correspond to the H-H bond pointing directly towards the H-atom, the H-H bond being found perpendicular to the H-atom, and the H-H bond being found skewed by 45°. These orientations are shown in Figure 2.9. In order to spherically average the H_2 molecule, the total electronic energy of each respective orientation is multiplied by 2/15, 4/15, and 9/15, according to the Lebedev quadrature scheme. Previous work has been done examining the H-H_2 interaction but was performed using orthohydrogen and not parahydrogen.[123]

Just as with the H_2 PEC, there are very few electrons in this system and so the FCI limit can be reached using the CCSD(T) and CCSDT levels of theory. A composite scheme is used

Table 2.10 Experimental and theoretical spectroscopic constants for the H_2 molecule at the CCSD/CBS level of theory.

	Experiment[124,125]	This Work
Equilibrium Bond Length (Å)	0.74144	0.7417642
Dissociation Energy (D_e, eV)	4.4777	4.748321
Zero Point Energy (cm^{-1})	2179.307	2176.9195
Stretching Frequency (ω_e, cm^{-1})	4401.21	4939.80995
Anharmonic Term ($\omega_e x_e$, cm^{-1})	121.336	130.5285
Rotational Constant (B_e, cm^{-1})	60.853	60.35994

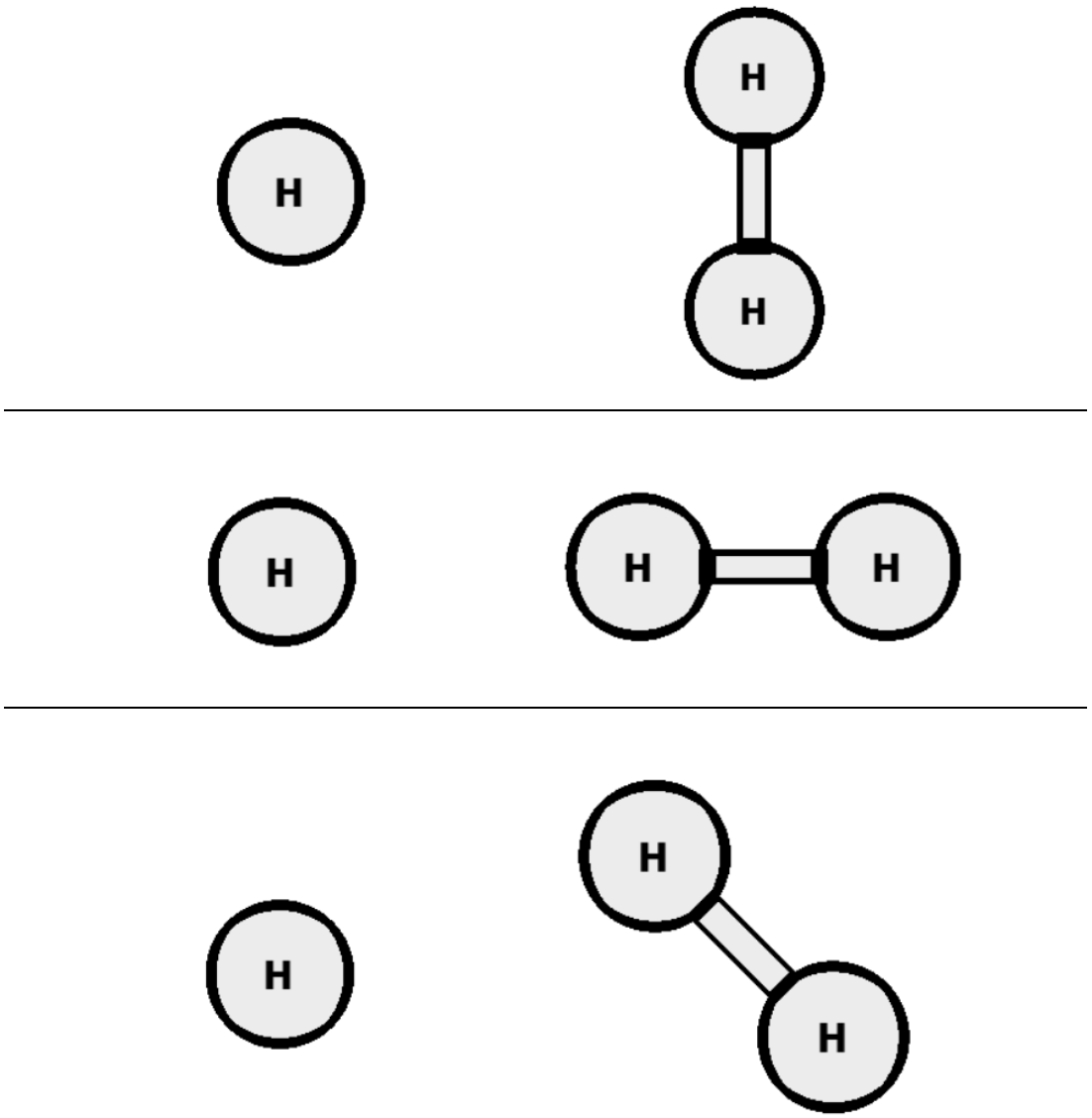


Figure 2.9 Representations of the three unique H-pH_2 orientations used in the construction of the PEC.

whereby CCSD(T) is utilized with the ACCT and ACCQ basis sets which are then extrapolated to the CBS limit using Eq. 2.18 and Eq. 2.19 for the HF and correlation energies, respectively. A correction term is then added to the potential curve which accounts for the difference between perturbative triples (CCSD(T)) and full triples (CCSDT) using the ACCQ basis set. This composite scheme is shown explicitly in Eq. 2.20, below. The distance between the H-atom and the center of mass of the H_2 molecule was examined between R = 2.000 and 10.000 Å in increments of 0.081 Å, thus capturing the highly repulsive wall of the curve and the low energy, attractive portion of the curve. The H_2 bond length was set to 0.74176 Å. Counterpoise corrections[126] were used to account for the basis set superposition error; the CCSD energies for H_2 were based on a restricted Hartree-Fock reference wavefunction while the CCSD(T) and CCSDT energies for H-pH_2 were based on an unrestricted Hartree-Fock wave function. Standard values are used for the SCF and CC convergence. At the CBS limit, the minimum energy distance is at 3.4501 Å with an interaction energy of 16.0177 cm^{-1}. Because of the weak interaction, no bound vibrational state was found for the H-pH_2 van der Waals dimer.. Thus, any H-pH_2 interactions that would be observed in the pH_2 matrix are purely a result of the H-atom being physically unable to escape the crystal.

$$E_{CBS}^{HF} = E_{QZ}^{HF} + A\left[E_{QZ}^{HF} - E_{TZ}^{HF}\right], A = 0.273973 \tag{2.18}$$

$$E_{CBS}^{corr} = E_{QZ}^{corr} + B\left[E_{QZ}^{corr} - E_{TZ}^{corr}\right], B = 0.651174 \tag{2.19}$$

$$E = E_{CBS}^{CCSD(T)} + \Delta T_{ACCQ} \tag{2.20}$$

pH_2-pH_2 Curve

Just as with the H-pH_2 PEC, the pH_2-pH_2 curve is made both more and less difficult due to the spherical nature of the pH_2 molecule. Previous work on the H_2-H_2 van der Waals dimer has been completed both experimentally[127–129] and theoretically[130–132] which showed that the system is vibrationally bound by a single state which indicates that the pH_2 crystal lattice is only stable at very low temperatures, as observed by Kufeld *et al.*[133] The same approximation is made here in which discrete orientations of the H_2 molecule are used and then spherically averaged. In this case,

seven unique orientations of the H_2 molecule are used which results in forty-nine orientations being used for the pH_2-pH_2 interaction. Due to symmetry, these forty-nine orientations are reduced to nine unique orientations, shown in Figure 2.10. Just as with the H_2 and H-pH_2 curves, there are few electrons to correlate and so a high level of electronic structure theory can be utilized. A similar composite scheme to the H-pH_2 van der Waals dimer is utilized except corrections from the difference in energy between CCSD(T)/ACCT and CCSDTQ/ACCT are included. The minimum energy distance between the dimers was found to be at 3.4858 Å which is very similar to the H-pH_2 dimer. The interaction energy of the pH_2 dimer was calculated to be 22.2579 cm^{-1} with a vibrational constant of 47.253 cm^{-1}. The ZPE of this system is 19.710 cm^{-1} and is 2.55 cm^{-1} lower than the interaction energy which indicates that this system is vibrationally bound but only at very low temperatures. Previous theoretical work showed that the binding energy of the pH_2-pH_2 van der Waals dimer was 2.895 cm^{-1} which is in very good agreement with the current results.[130] This result helps to illustrate why the pH_2 solid is stable at only cryogenic temperatures as higher temperatures would cause the vibrational motion of the dimer to overcome the interaction energy and thus cause the crystal to melt. The CBS interaction energies for both the H-pH_2 and pH_2-pH_2 van der Waals dimers as a function of intermolecular distance are shown in Figure 2.11, below.

Zero-point Vibrational Motion

As both the pH_2 molecules and H-atoms are very light, their zero-point vibrational motion in the solid state has a large amplitude as compared to other, heavier atoms. Thus, when they are found in a crystal lattice such as that synthesized experimentally by Kufeld *et al*,[133] this motion must be accounted for as it controls many of the properties of the crystal lattice. Other work done on simulating such an environment includes the work of Kühn *et al*. which examined an anharmonic extension of the Einstein model for the pH_2 crystal lattice.[134] The HCP lattice that is formed upon the cooling of pH_2 molecules shows a lattice constant of 3.79 Å[135] which is larger than the minimum energy distance found when constructing the pH_2-pH_2 potential energy curve. The zero-point vibrational motion is approximately Gaussian in nature and so to correct the pH_2-pH_2 PEC, we used a method known as Gaussian smearing. This method places three Gaussian probability distributions (one for each cartesian axis) on top of the PEC and takes the weighted

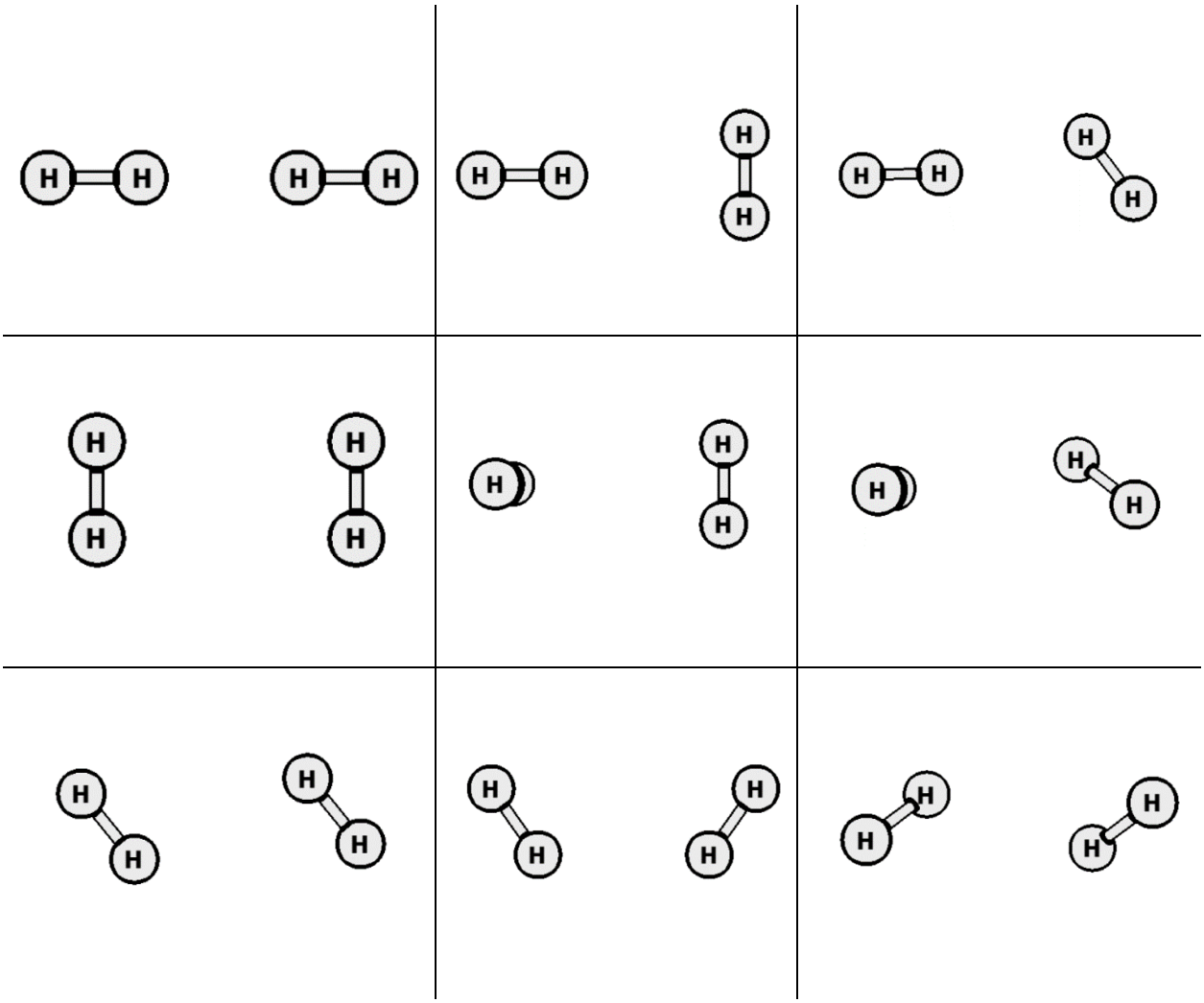


Figure 2.10 Pictorial representations of the nine unique pH_2-pH_2 orientations used in the construction of the PEC.

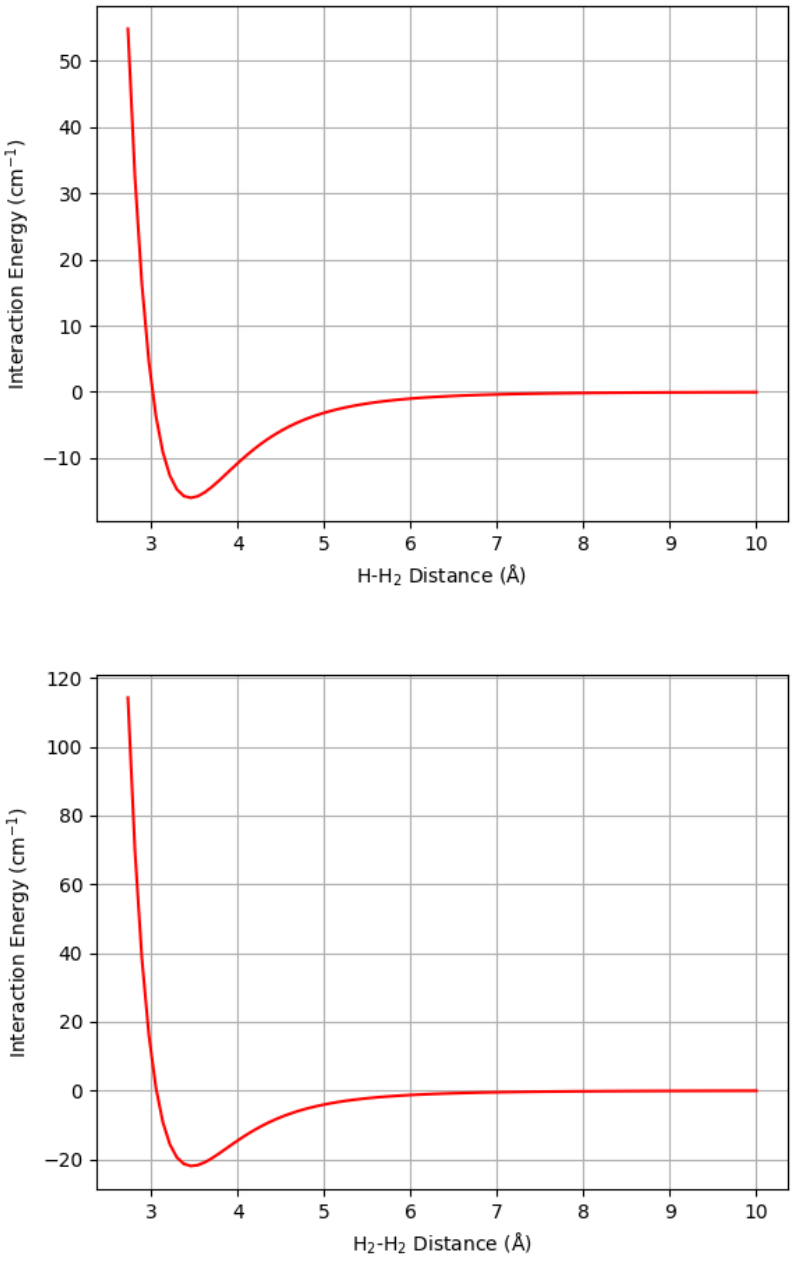


Figure 2.11 Potential energy curves for the H-pH_2 (top) and pH_2-pH_2 (bottom) van der Waals dimers as a function of intermolecular distance.

average of those three distributions to arrive at an interaction energy that now implicitly includes the zero-point motion of the system. To perform simulations using this methodology, a numerical form of the Gaussian probability distributions must be used and two aspects of this methodology must be tested: the number of points used to approximate the Gaussian probability distribution and the σ value used to control the width of the Gaussian. Previous work has looked at using a similar methodology to simulate solid pH_2 and determined that a σ value of 0.30 Å reproduces the lattice constant of solid pH_2 (3.79 Å) using the pair potential of Silvera-Goldman.[136] As the current pH_2-pH_2 potential energy curve is of a different form, we have tested the PEC in order to determine the σ value which reproduces the lattice constant of solid pH_2. Shown in Figure 2.12 is the PEC with different σ values for the Gaussian probability distributions. As σ increases, the minimum of the PEC is brought up and is pushed to larger distances as more points from the repulsive wall of the curve are included which are higher in energy than the minimum. It was found that for the current PEC, a σ value of 0.4875265 Å reproduces the lattice constant of pH_2 (see Figure 2.13). This work used five points along each of the Gaussian probability distributions to discretize the function. Convergence was checked for this factor and it was found that given a rotating water molecule at the center of an HCP lattice composed of twelve pH_2 molecules, the interaction energies of the system show little change when more than three points are included in the discretization of each of the Gaussian probability distributions (Table 2.11, Figure 2.14). A similar study was done for the H-pH_2 interaction in order to calculate the σ value for an H atom and it was found that a value of 0.5563 Å reproduces the previously mentioned H_2 lattice constant of 3.79 Å. While this experimental value is not relative to the H-atom, we have chosen to use it for this case so as to closely reproduce the effects that an H-atom may experience while acting as a substitution in a pH_2 HCP lattice.

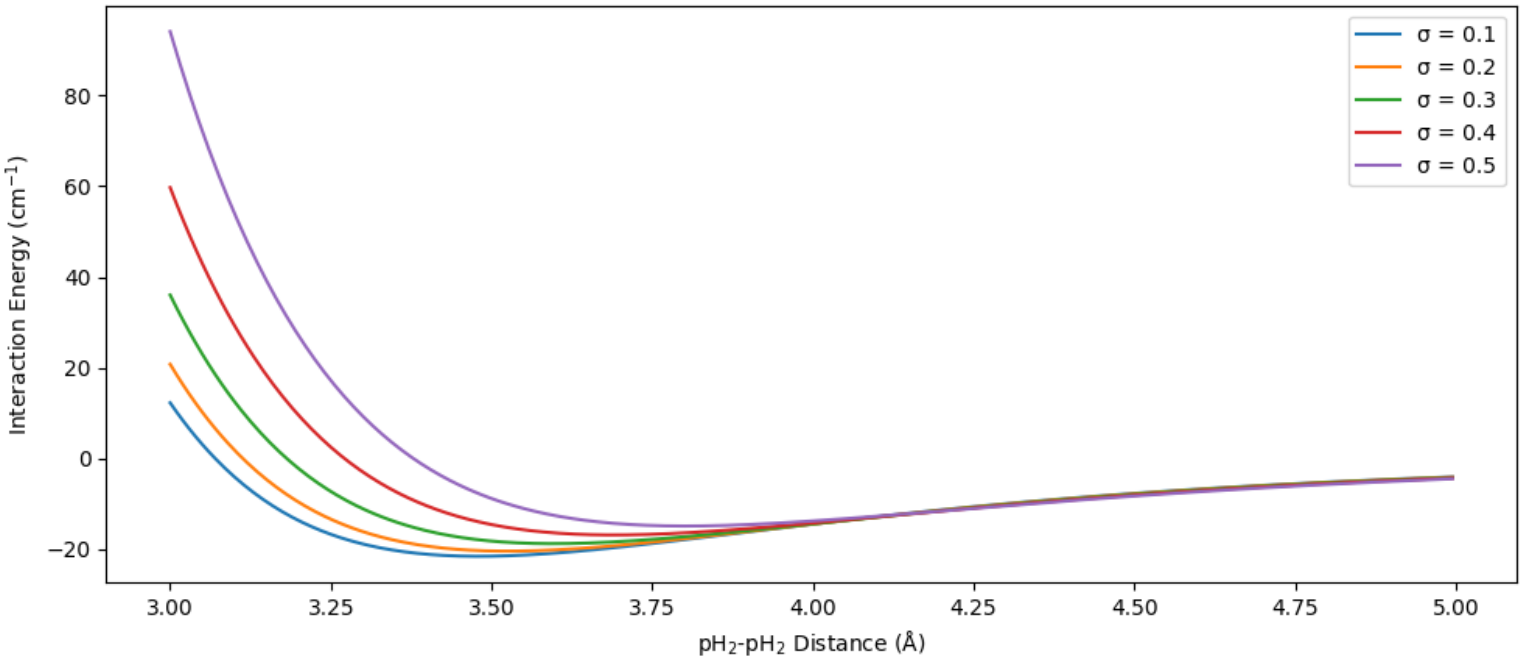


Figure 2.12 Potential energy curves for the pH_2-pH_2 van der Waals dimer using Gaussian smearing with varying σ values.

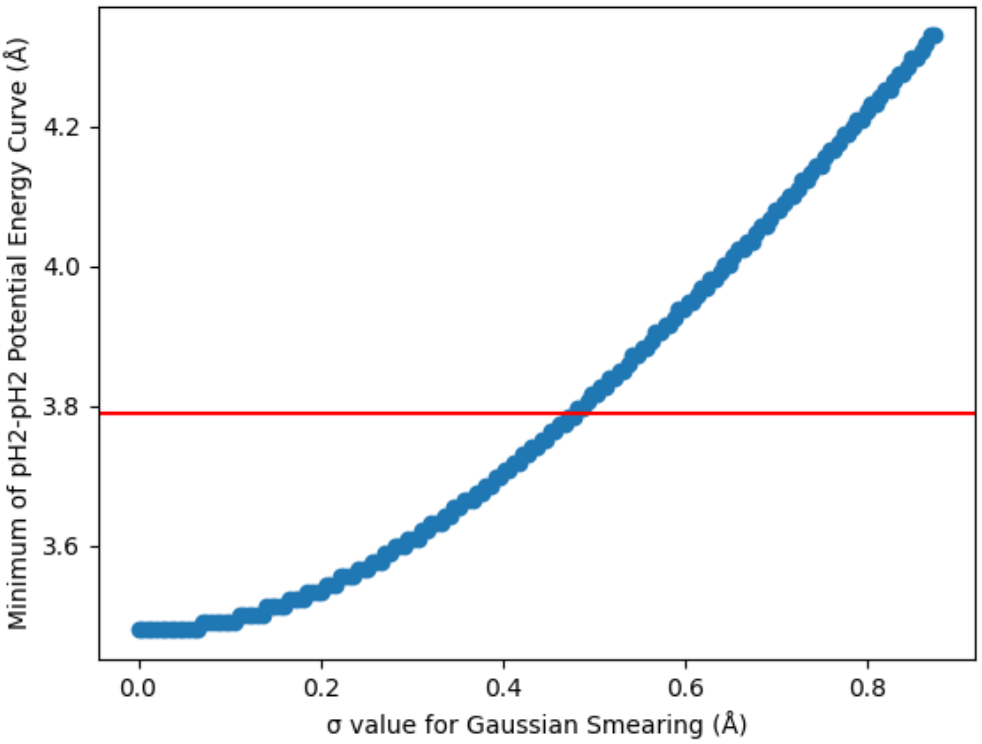


Figure 2.13 Scatter plot showing the effect that the σ value for the Gaussian smearing has regarding the minimum of the theoretical pH_2-pH_2 PEC. The red line shows the value of the lattice constant of pH_2 (3.79 Å).

Table 2.11 Effect of the number of points included in the discretization of the Gaussian probability distribution for the H_2O-pH_2 interaction. Number of discretization points is the cube of the number of points used to discretize a single Gaussian probability distribution. Values shown correspond to the interaction energies (in cm^{-1}) of a H_2O-$(pH_2)_{12}$ system in an HCP lattice geometry. 216 unique orientations of the rotating H_2O molecule are used. MSD and MAD are relative to the previous number of discretization points.

Number of Discretization Points	Average Interaction Energy	Standard Deviation	Minimum Interaction Energy	Maximum Interaction Energy	MSD	MAD
1	-54.183	8.021	-68.434	-42.693	-	-
8	-52.502	6.865	-64.757	-41.750	1.680	2.076
27	-52.355	6.793	-64.397	-41.605	0.148	0.214
64	-52.338	6.799	-64.389	-41.602	0.017	0.019
125	-52.341	6.797	-64.389	-41.602	-0.003	0.005
216	-52.340	6.797	-64.389	-41.602	0.001	0.001

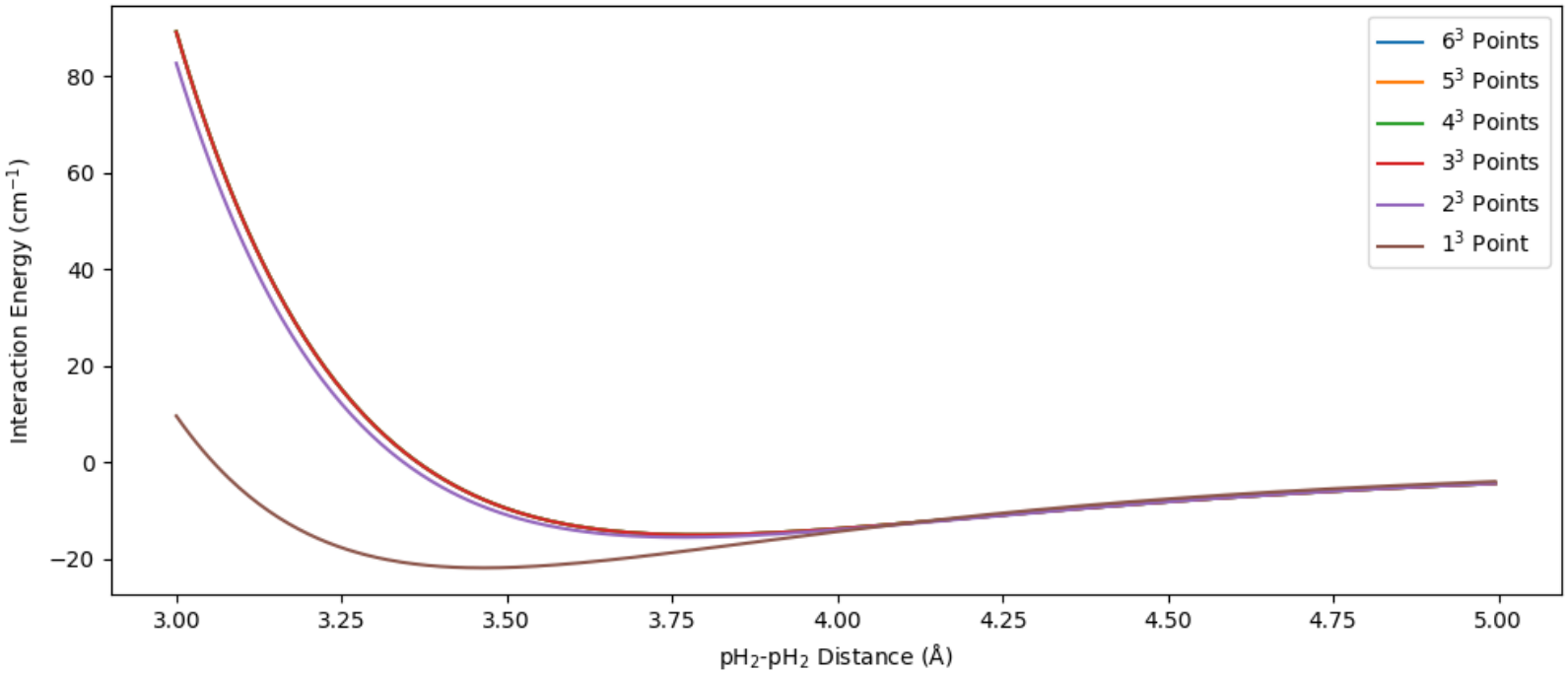


Figure 2.14 Plot showing the effect that the number of discretization points for the Gaussian smearing has regarding the pH_2-pH_2 PEC. Note that the 4^3, 5^3, and 6^3 points lines lie almost on top of each other due to little change observed when more than 3^3 points in the discretization is used.

3-Dimesional Surfaces

Disclosure

The following subsection has been taken, with permission, from a journal article published in *The Journal of Chemical Physics* with only minor modifications.[137] This work has been authored by Dr. RJ Hinde and myself and contains no edits or modifications from others.

H_2O-H Surface

The construction of the H_2O-H potential energy surface is far more complex than the previously mentioned potential energy curves due to the addition of two degrees of freedom. When the H_2O molecule is held at a fixed geometry, the potential energy surface is a function of three coordinates that specify the position of the H atom relative to the H_2O molecule. Thus, a function of higher complexity must be used to determine the interaction energy of the H_2O-H van der Waals dimer at any given point in space. Many such interpolation schemes exist which are able to evaluate such a surface including hyperspherical expansions,[138] a bond-bond method,[139] quadrature and spline-based methods,[140] and more recently, machine learning (ML) based approaches.[141,142] In addition to interpolation along an entire surface, numerous methods exist for interpolation on the surface of a sphere[122,143,144]. Here we report a quadrature-based interpolation method which can reproduce the H_2O-H potential energy surface with very high accuracy.

Orca 4.2[118] was used to calculate the interaction energy of the H_2O-H system, using the CC approach including single, double, and perturbative triple excitations, or CCSD(T), according to the following protocol. The H_2O molecule was held fixed at its equilibrium geometry as determined from spectroscopic data: the OH bond length was fixed at 1.80965034 a_0 and the HOH bond angle was fixed at 1.82404493 radians.[145] Scalar relativistic effects were accounted for by utilizing the second-order Douglas-Kroll-Hess (DKH)[52,53] Hamiltonian using the DK-recontracted aug-cc-pVnZ-DK basis sets (n = T, D, Q).[146] For brevity, these basis sets will henceforth be referred to as aug-cc-pVnZ, or simply as nZ. Basis set extrapolation to the complete basis set (CBS) limit was also carried out using the TZ and QZ energies. The Hartree-Fock energy E^{HF} was extrapolated using the method of Karton *et al.*[119] (Eq. 2.21) while the correlation energy E^{corr} was extrapolated following Huh *et al.*[120] (Eq 2.22).

$$E_{CBS}^{HF} = E_{QZ}^{HF} + A\left[E_{QZ}^{HF} - E_{TZ}^{HF}\right], A = 0.273973 \quad (2.21)$$

$$E_{CBS}^{corr} = E_{QZ}^{corr} + B\left[E_{QZ}^{corr} - E_{TZ}^{corr}\right], B = 0.651174 \quad (2.22)$$

Counterpoise corrections were used to account for the basis set superposition error;[126] the CCSD(T) energies for H_2O-H were based on an unrestricted Hartree-Fock reference wavefunction, while the CCSD(T) energies for H_2O were based on a restricted Hartree-Fock wavefunction. The SCF energy convergence cutoff was set to 10^{-9} E_h, a linear dependence threshold of 10^{-6} was used, and the CC residual threshold was set to 10^{-8}. These cutoffs and thresholds are chosen to reduce the possibility of near-linear dependences and to increase the accuracy of the interaction energy. Because the H_2O-H interaction energies are on the order of 10 to 100 cm^{-1}, the H_2O-H, H_2O, and H energies must be computed to high precision.

The H atom position relative to the H_2O molecule was defined as follows. The H_2O center of mass was set as the origin of the Cartesian coordinate system, with the H_2O molecule in the (x,y) plane and the C_2 axis of the H_2O molecule oriented along the x-axis; the oxygen atom is located on the negative x-axis (Figure 2.15). The position of the H atom is specified by spherical polar coordinates (R, θ, φ) with respect to the Cartesian axes defined by the H_2O molecule, with the H atom's Cartesian coordinates defined as $x = R\sin(\theta)\cos(\varphi)$, $y = R\sin(\theta)\sin(\varphi)$, $z = R\cos(\theta)$. The angular coordinates (θ, φ) are chosen to reside on an angular grid defined by the Lebedev quadrature rule.[122] Other grids such as a Gaussian product grid were considered but due to the widespread use, accuracy, and efficiency of Lebedev grids in other electronic structure methods such as Density Functional Theory (DFT), we chose to use Lebedev grids.[147] Due to the C_{2v} symmetry of the water molecule, the number of unique points in the grid is less than the total number of points which reduces the number of calculations required (discussed below). For the spherical coordinates, θ ranges from 0 to π and φ ranges from 0 to 2π. For the radial coordinate R,

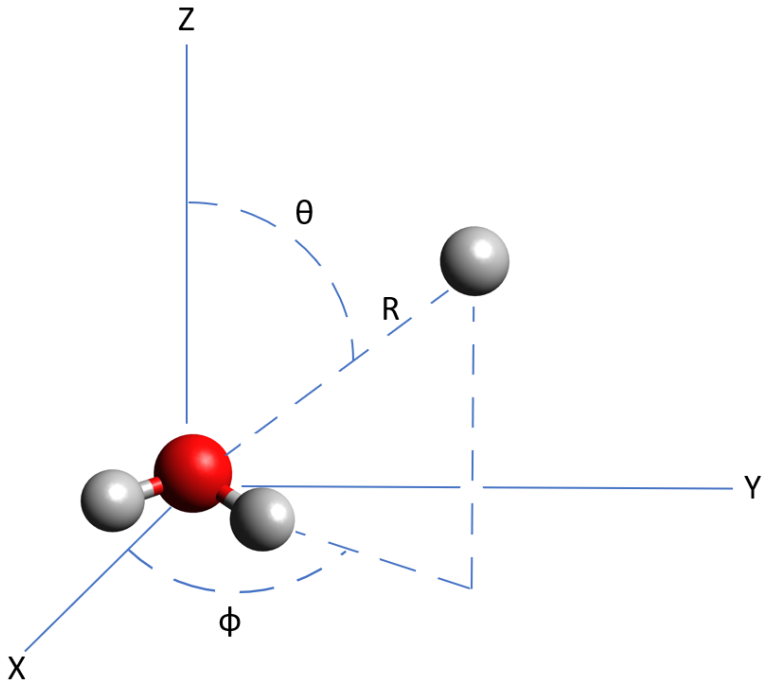


Figure 2.15 Coordinate system for the H_2O-H dimer in both cartesian and spherical polar coordinates

twenty-seven evenly spaced values were chosen, ranging from 2.5 to 9.0 Å in increments of 0.25 Å, with four additional R values of 3.79, 3.84, 7.58, and 7.63 Å included to provide more data for H_2O-H distances that are close to the nearest- and 6th-nearest-neighbor distances for solid parahydrogen (3.79 and 7.58 Å).[135] This yields 3,410 points at which the counterpoise-corrected CCSD(T) interaction energies are calculated. Of these points, 1,054 represent symmetry-unique geometries of the H_2O-H dimer.

At a fixed value of R, the PES can be approximated as a truncated expansion in spherical harmonics ($Y_{jm}(\theta, \varphi)$):

$$V(R, \theta, \varphi) = \sum_{j=0}^{j_{max}} \sum_{m=-j}^{j} C_{jm}(R) Y_{jm}(\theta, \varphi) \tag{2.23}$$

Because the angular grid points were chosen according to a Lebedev quadrature scheme, at each value of R the spherical harmonic expansion coefficients C_{jm} can be estimated using the Lebedev quadrature rule. Eq 2.24 shows how these coefficients are evaluated using spherical harmonics, the potential energies $E(\theta, \varphi)$ on the sphere, and the weights $w(\theta, \varphi)$ of each Lebedev quadrature point.

$$C_{jm} = \sum_{\theta,\varphi} Y_{jm}^{*}(\theta, \varphi) E(\theta, \varphi) w(\theta, \varphi) \tag{2.24}$$

Due to the C_{2v} symmetry of the water molecule, the unique number of points on the angular grid is reduced. For fixed values of φ and θ, only θ values less than or equal to π/2 must be evaluated due to the mirror plane of the water molecule in the xy-plane. Likewise, only φ values between 0 and π must be evaluated due to the mirror plane in the xz-plane. This reduces the number of unique points in the angular grid by up to 75% depending upon the size of the grid. Likewise, the symmetry of the water molecule may simplify the spherical harmonic expansion of Eq. 2.23 by utilizing the parity of (j,m) pairs. This fact is discussed in further detail near the end of this subsection.

The Lebedev quadrature scheme has been designed for numerous angular grid sizes ranging from six points up to several thousand points. Because of this, we chose to examine how the accuracy of the interpolation at a fixed value of R changes with increases in the angular grid size. Grids ranging from 6 points (Lebedev-3) to 266 points (Lebedev-27) at R = 3.0 and 3.79 Å were chosen with H_2O-H interaction energies being calculated at each of those points using the CCSD(T)/aug-cc-pVQZ level of theory. These interaction energy values were used to calculate the C_{jm} coefficients (Eq 2.24) up to a value of j_{max} determined by the size of the Lebedev grid. The C_{jm} coefficients converge smoothly up to a Lebedev-17 grid after which very little change is observed. The C_{jm} coefficients were then used to reconstruct the potential energy values at the Lebedev grid points using Eq. 2.23. The mean absolute error (MAE, Eq 2.25), root mean squared error (RMSE, Eq 2.26), and coefficient of determination (COD, Eq 2.27) were calculated in reference to the reproduced potential energy values of each individual grid (shown in Table 2.12). For example, the error of the Lebedev-17 grid is calculated based on the difference between the actual *ab initio* interaction energies and the energies calculated using the truncated spherical harmonic expansion of Eq. 2.24. As the Lebedev grid becomes larger and the spherical harmonic expansion becomes more complete, the reconstruction errors approach 0.000 cm^{-1} and the COD values approach 1.000. For the Lebedev-17 grid at 3.79 Å, the MAE and RMSE approach 9 nE_h (9 x 10^{-9} E_h) and the COD is 0.9999997 which is sufficiently close to 1.0 to consider the actual and interpolated values to be perfectly correlated. This same grid at 3.0 Å shows similar accuracy with MAE and RMSE of 36-46 nE_h which is still a very small error considering the much larger range of interaction energies at this distance. Thus, we consider the H_2O-H system at R = 3.0 and 3.79 Å to be sufficiently mapped using the Lebedev-17 grid. To illustrate the potential energy surface for these spheres, Figure 2.16 shows two-dimensional plots of the PES as a function of θ and φ at these two R values where the behaviors of the repulsive and attractive regions of the PES change with a change in R. At $\theta = \pi/2$, for example, the hydrogen atom is moving in the plane of the H_2O molecule and so the effect of the hydrogens of the H_2O molecule are evident in the shifting repulsive-attractive behavior of the PES.

Table 2.12 Effect of Lebedev Grid size on the accuracy of reconstructing the potential energy surface at the Lebedev grid points on spheres with R = 3.00 Å and 3.79 Å. Values calculated at the CCSD(T)/QZ level of theory.

Lebedev Scheme	J_{max}	Number of Grid Points	Symmetry Reduced Grid Points	MAE (cm^{-1})		RMSE (cm^{-1})		COD	
				3.00 Å	3.79 Å	3.00 Å	3.79 Å	3.00 Å	3.79 Å
3	1	6	4	0.088	3.339	0.095	3.788	0.9997157	0.1444688
5	2	14	6	3.004	0.660	3.674	0.819	0.8760921	0.9484573
7	3	26	11	0.796	0.802	0.931	0.986	0.9935833	0.9547674
9	4	38	16	0.112	0.254	0.148	0.306	0.9998134	0.9955806
11	5	50	17	0.069	0.059	0.091	0.073	0.9999329	0.9997437
13	6	74	27	0.050	0.019	0.067	0.022	0.9999640	0.9999735
15	7	86	28	0.020	0.005	0.025	0.007	0.9999950	0.9999980
17	8	110	34	0.008	0.002	0.010	0.002	0.9999992	0.9999997
19	9	146	41	0.004	0.001	0.005	0.001	0.9999998	1.0000000
21	10	170	51	0.002	0.000	0.002	0.000	1.0000000	1.0000000
23	11	194	57	0.001	0.000	0.001	0.000	1.0000000	1.0000000
25	12	230	68	0.000	0.001	0.000	0.001	1.0000000	0.9999999
27	13	266	75	0.001	0.000	0.002	0.000	1.0000000	1.0000000

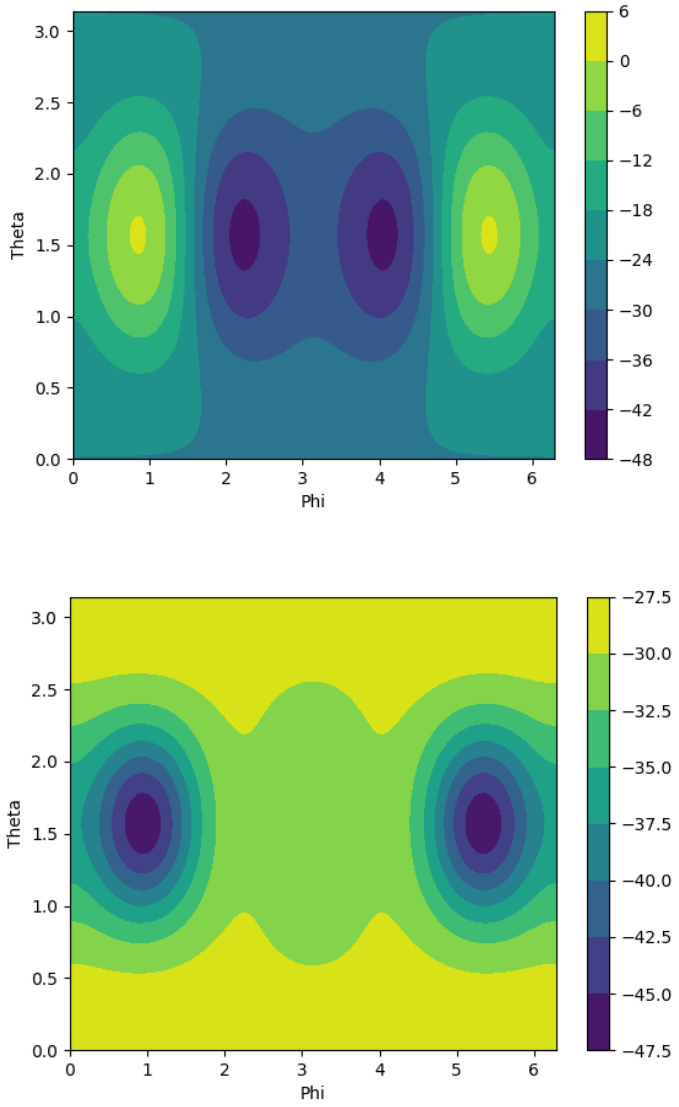


Figure 2.16 Potential energy surface at R = 3.00 Å (top) and R = 3.79 Å (bottom). Energy values are in cm^{-1}. Values are calculated at the CCSD(T)/QZ level of theory.

$$\frac{1}{N}\sum_{i=1}^{N}|E_i^{Lebedev} - E_i^{ab\ initio}| \tag{2.25}$$

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(E_i^{Lebedev} - E_i^{ab\ initio})^2} \tag{2.26}$$

$$1 - \frac{\sum_{i=1}^{N}\left(E_i^{Lebedev} - E_i^{ab\ initio}\right)^2}{\sum_{i=1}^{N}\left(\bar{E} - E_i^{ab\ initio}\right)^2}, \bar{E} = \frac{1}{N}\sum_{j=1}^{N} E_j^{ab\ initio} \tag{2.27}$$

As the PES is constructed using an expansion in spherical harmonics (Eq 2.23), we can examine the effect that the maximum j-value has on the accuracy of the PES. For the Lebedev-17 grid, the maximum j-value for which the C_{jm} coefficients can be accurately estimated is $j_{max} = 8$, and a truncated expansion (Eq 2.23) using higher values of j_{max} can introduce errors. To demonstrate this, we use Eq 2.23 and the potential energies at the Lebedev-17 grid points to evaluate the C_{jm} coefficients at R = 3.0 and 3.79 Å for j_{max} values from 1 to 15, and then use the C_{jm} coefficients to reconstruct the PES at the Lebedev-17 grid points. The errors in the reconstructed PES decrease with increasing j_{max} until $j_{max} = 9$, at which point the errors begin to increase (Table 2.13). According to this data, a value of either $j_{max} = 7$ or $j_{max} = 8$ could be used with similar accuracy. When the C_{jm} coefficients are constructed using a j_{max} value of 8, however, the MAE and RMSE is half that when a j_{max} value of 7 is used, indicating that the expansion should be performed with the larger j_{max} value. If efficiency is more important, however, then the j_{max} value of 7 could be used. For $j_{max} = 8$, the PES must be evaluated using eighty-one coefficients which is ~25% more than for $j_{max} = 7$, while showing similar (small) errors. For all future comparisons, however, $j_{max} = 8$ is chosen as it is the more complete expansion and leads to a lower error. To further show the convergence of the expansion, Table 2.14 lists five different C_{jm} coefficients as a function of the size of the Lebedev grid at both 3.0 and 3.79 Å. At the Lebedev-17 grid, each of the selected C_{jm} coefficients has converged to within 0.002 cm^{-1} or better. Both

Table 2.13 Effects of Spherical Harmonic expansion at R = 3.00 Å and 3.79 Å based on spherical harmonic coefficients calculated using potential energies obtained at the Lebedev-17 grid points. These results are for potential energies calculated at the CCSD(T)/QZ level.

j_{max}	Number of Components in Expansion	MAE (cm^{-1})		RMSE (cm^{-1})		COD	
		3.00 Å	3.79 Å	3.00 Å	3.79 Å	3.00 Å	3.79 Å
1	4	4.414	2.824	5.459	3.691	0.7577507	0.3520117
2	9	4.438	1.830	5.442	2.259	0.7593280	0.7572594
3	16	0.576	0.582	0.714	0.752	0.9958607	0.9731325
4	25	0.104	0.180	0.132	0.228	0.9998588	0.9975188
5	36	0.071	0.065	0.103	0.080	0.9999144	0.9996975
6	49	0.041	0.016	0.060	0.021	0.9999703	0.9999786
7	64	0.015	0.005	0.021	0.007	0.9999965	0.9999977
8	81	0.008	0.002	0.010	0.002	0.9999992	0.9999997
9	100	0.006	0.002	0.007	0.002	0.9999996	0.9999998
10	121	0.008	0.004	0.011	0.006	0.9999990	0.9999985
11	144	0.025	0.014	0.033	0.018	0.9999914	0.9999841
12	169	0.055	0.055	0.071	0.067	0.9999586	0.9997878
13	196	0.117	0.200	0.147	0.251	0.9998240	0.9969951
14	225	0.491	0.617	0.633	0.803	0.9967444	0.9693140
15	256	4.334	1.758	5.837	2.521	0.7230651	0.6977164

Table 2.14 Effect of Lebedev grid on selected expansion coefficients (C_{jm}) at R = 3.00 Å and 3.79 Å. These results are for potential energies calculated at the CCSD(T)/aug-cc-pVQZ level. The C_{jm} coefficients are given in cm^{-1}.

J	*M*	*R (Å)*	*Lebedev Grid*										
			5	*7*	*9*	*11*	*13*	*15*	*17*	*19*	*21*	*23*	*25*
2	*0*	*3.00*	-0.1382561	-1.5175302	-1.538435	-1.5146077	-1.5148204	-1.5138681	-1.5136146	-1.5138526	-1.5138365	-1.5138403	-1.5139878
		3.79	7.9391414	9.94633897	9.88431409	9.87132706	9.87101428	9.87018931	9.87023835	9.87050729	9.87024825	9.8704959	9.87067259
3	*3*	*3.00*		12.6718373	12.798867	12.7471	12.7499775	12.7529765	12.7531476	12.7533225	12.7532919	12.7532656	12.7532283
		3.79		-4.5748205	-4.6851923	-4.66169	-4.6633849	-4.6630834	-4.6627114	-4.6627128	-4.6628189	-4.6627845	-4.662813
5	*3*	*3.00*			0.240841	0.19390499	0.20597973	0.20442573	0.20439085	0.20435819	0.20438456	0.20444334	0.20185312
		3.79			0.47585326	0.47617487	0.47538064	0.47533898	0.47520307	0.47526575	0.47532144	0.47519578	0.47523551
8	*0*	*3.00*							0.02037476	0.01944378	0.01960607	0.01949899	0.01918443
		3.79							-0.0136968	-0.013792	-0.0137879	-0.0137626	-0.0141914
8	*4*	*3.00*							0.0061272	0.0047173	0.00482023	0.00472871	0.00478654
		3.79							0.00480384	0.00556992	0.00544272	0.00550028	0.0052009

the large coefficients (C_{20}, C_{33}, C_{53}) and small coefficients (C_{80} and C_{84}) show this convergent behavior indicating the near-completeness of the Lebedev-17 grid.

To move from the interpolation on a single sphere to interpolation in 3-dimensional space, the Lebedev quadrature must be carried out at multiple values of R. For the thirty-one radial distances, a Lebedev-17 grid was constructed according to the computational details discussed previously. The C_{jm} coefficients for each of these radial spheres were then calculated (Eq 2.23) and each of the eighty-one individual coefficients was fit to a function of R between 2.5 and 9.0 Å. This function acts as the third-dimension in the PES and allows for the calculation of the C_{jm} coefficients at any given R. For the fitting function, we examined both polynomials in R (Eq 2.28, using n values ranging from 1 to 15) as well as a Morse potential.

$$C_{jm}(R) = a_{n,jm}R^n + a_{n-1,jm}R^{n-1} + \cdots + a_{2,jm}R^2 + a_{1,jm}R + a_{0,jm} \qquad (2.28)$$

As the order of the polynomial increases, the MAE for the replication of the entire PES approaches a constant value of ~0.01 cm^{-1} and the COD approaches 1.0. With a 12^{th} degree polynomial, the fit is essentially converged with a MAE of 0.013 cm^{-1} and an RMSE of 0.049 cm^{-1}. Near the repulsive wall (R between 2.5 and 3.0 Å), the MAE of the reproduced PES is quite large compared to the attractive region ($R \geq 3.0$ Å); values of 0.21 and 0.02 cm^{-1} were calculated for each respective region. This fact does not change the conclusions drawn about the excellent reproduction of the PES, however. This is because the interaction energies near the repulsive wall are much larger in magnitude than in the attractive region and the larger MAE near the repulsive wall simply reflects the larger scale of energies there. Furthermore, simulations at cryogenic temperatures are unlikely to examine this region of the PES and so the larger error in this region would not have a measurable effect. The Morse potential performs as well as a 9^{th} degree polynomial (~1 cm^{-1} MAE) and has less than half the number of fitting parameters.

The error in reproducing the original PES due to fitting the C_{jm} coefficients with a polynomial is negligible. Figure 2.17 shows that when we compare the PES that has been

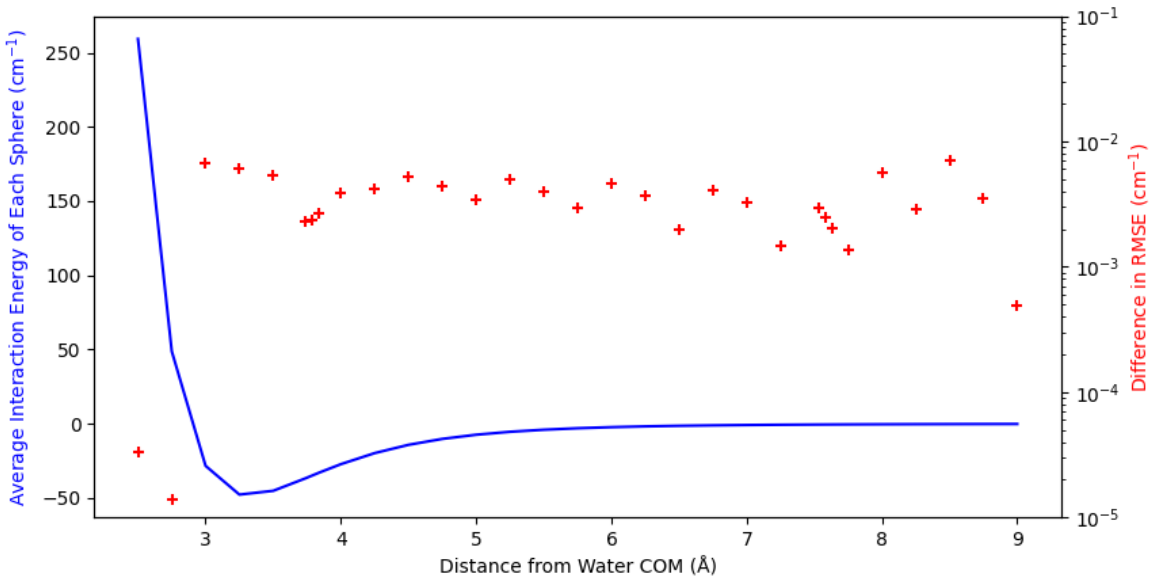


Figure 2.17 The blue line (left vertical axis) shows the average interaction energy in cm^{-1} on each sphere used to construct the PES. The red crosses (right vertical axis) show how the RMSE values for each sphere differ for the PES constructed using spherical harmonic coefficients obtained from the R-dependent fits and using spherical harmonic coefficients computed directly from the ab initio results for each sphere. Values shown in this figure are for the PES computed using the CCSD(T)/QZ basis set.

reproduced using spherical harmonic coefficients from the R-dependent polynomials with the PES computed on each sphere using the spherical harmonic coefficients obtained from that sphere's *ab initio* data, the difference in RMSE never exceeds a value of 0.007 cm^{-1} (shown as red crosses, right y-axis). This figure also shows that even for larger interaction energies (near the repulsive wall and the potential minimum), the surface is well reproduced. Thus, we feel confident that a polynomial expansion in R can be used to create a three-dimensional potential energy surface with minimal loss in accuracy.

To examine the applicability of the PES discussed thus far, we have created several testing datasets that are composed of points that do not sit on the Lebedev grid points. These include spherical datasets at R = 3.0, 3.79, and 4.58 Å, and a dataset composed of 500 points randomly distributed in space between R = 2.5 and 9.0 Å. The last dataset is split into a mostly-repulsive dataset composed of 40 points between R = 2.5 and 3.0 Å and an attractive dataset composed of points between R = 3.0 and 9.0 Å. For these datasets, we calculate the interaction energy at each point using CCSD(T) and the aug-cc-pVQZ basis set. The characteristics for each dataset (number of points, minimum, maximum, and average values) are shown in Table 2.15, below alongside the MAE, RMSE, and COD values for each dataset. This data leads to some important observations regarding the PES. The error is largest before the minimum of the PES is reached (R < 3.5 Å) most likely due to the larger magnitudes of the interaction energies in this region but the error is below 0.01 cm^{-1} in every other region. The COD for the datasets is ~0.99999 which shows a good correlation between the actual and predicted values which, for a PES, is excellent. Thus, C_{jm} coefficients evaluated up to j_{max} = 8 using the Lebedev-17 grid, then expanded in R using 12^{th} degree polynomials, provide an excellent reproduction of the H_2O-H PES between R = 2.5 and 9.0 Å.

In the creation and subsequent reproduction of the H_2O-H potential energy surface, we were able to scan the entire surface to locate the global minimum energy orientation of the H-atom with respect to the H_2O molecule. Due to the interactions of the H-atom with either the H-atoms or the lone pairs of the H_2O molecule, the minimum energy position for the H atom is in the plane of the H_2O molecule, almost colinear with either of the O-H bonds, at an angle of

Table 2.15 Statistical information and results of the interpolation using the different testing datasets.

Dataset	Number of Points	Minimum (cm^{-1})	Maximum (cm^{-1})	Average (cm^{-1})	MAE (cm^{-1})	RMSE (cm^{-1})	COD
R = 3.00 Å Sphere	1000	-43.497	0.424	-22.763	0.018	0.023	0.9999957
R = 3.79 Å Sphere	3197	-46.732	-28.432	-31.914	0.004	0.005	0.9999985
R = 4.58 Å Sphere	1000	-16.795	-10.917	-12.501	0.007	0.009	0.9999696
Repulsive	40	-26.761	238.729	68.610	0.069	0.096	0.9999981
Attractive	460	-53.905	-0.169	-9.424	0.005	0.007	0.9999997

122.25° with respect to the x-axis and at R = 3.41 Å. This location minimizes all total interactions and results in a CBS binding energy of 61.297 cm^{-1}.

As this PES will subsequently be used in dynamical calculations, an efficient implementation of the PES is highly desired and so we examined a reduced form of the spherical harmonic expansion based on symmetry. Examination of the expansion coefficients showed that only (j, m) pairs of the same parity (even-even or odd-odd) result in coefficients of sufficient size to be additive to the sum. This is due to the underlying C_{2v} symmetry of the H_2O molecule. Thus, the eighty-one terms in Eq 2.23 for $j_{max} = 8$ can be reduced to forty-five terms which results in an efficiency boost of ~1.8. In reproducing the entire H_2O-H PES using this reduced form, the accuracy remains essentially unchanged with differences in the MAE values appearing in the 9^{th} and 10^{th} decimal places and so the speedup can be easily implemented to make dynamical calculations far more efficient.

H_2O-pH_2 Surface

The surface used to examine the H_2O-pH_2 interaction is more complicated than the H_2O-H surface simply due to the need to express the H_2 molecule as a spherically symmetric molecule. The work done to construct the H-pH_2 and pH_2-pH_2 curves is combined with the methodology of the H_2O-H surface to construct the H_2O-pH_2 surface. The H_2 molecule is assumed to be spherically averaged over seven different orientations which correspond to the fourteen points on a Lebedev-5 grid. Work done by Zeng *et al*. constructed a similar PES for the H_2O-pH_2 van der Waals dimer where the H_2 molecule wasn't spherically averaged but instead was treated as an adiabatic-hindered-rotor.[148] The points in space which determine the pH_2 position are chosen to lie at the same points as that of the H_2O-H surface: a Lebedev-17 grid is used with 110 points per radial sphere (34 symmetry unique points) and thirty-one radial distances are used which lie between R = 2.5-9.0 Å. The same interpolation scheme is used as that of the H_2O-H surface whereby an expansion of spherical harmonics is used to approximate all points on a given sphere of radius R and each expansion coefficient (see Eq. 2.23) is approximated using a polynomial function (Eq. 2.28) to allow for an expansion to three dimensions. The MAE (Eq 2.25), RMSE

(Eq 2.26), and COD (Eq 2.27) for some of the individual radial spheres are shown in Table 2.15, below. The results indicate that the interpolation of each sphere is nearly perfect with MAE values no greater than 0.05 cm^{-1} and COD values of 0.9999 in the regions of the PES of interest (3.5-4.0 Å) for the current dynamical calculations. The minimum energy position of the pH_2 molecule with respect to the rigid H_2O was found to lie at 3.2059 Å in the plane of the H_2O molecule. This position forms an angle of 101.4° with respect to the x-axis and results in a CBS interaction energy of 88.523 cm^{-1}. This position is slightly different from the minimum energy position of the H_2O-H system and may arise from the larger size of the pH_2 molecule compared to the H-atom.

Vibrational Perturbations

As the infrared spectra that are recorded by Kufeld *et al*[133] include both rotational and vibrational excitations, the H_2O-H and H_2O-pH_2 surfaces must account for the vibrational motion of the H_2O molecule within the pH_2 matrix. To achieve this, four surfaces were constructed which include different amounts of perturbations due to the vibrational modes of H_2O. One surface will assume the H_2O molecule is in the ground state for each of the three vibrational modes while the other three surfaces assume the H_2O molecule is in the first excited state for one of the normal modes. The three normal modes of the H_2O molecule were found by constructing a Hessian matrix that had been populated using an experimental potential energy surface for water.[149] This decision was made as all EST calculations were performed with the H_2O molecule in an experimentally defined geometry. Each normal mode of the H_2O molecule was stretched by factors of $\pm\sqrt{5q}, \pm\sqrt{3q}, \pm\sqrt{1q},$ and 0q where q is the dimensionless vibrational normal mode coordinate. The 0q stretch corresponds to the static, experimental geometry for water. For a given surface, the H_2O molecule was stretched by the seven values of q and a quadratic function (F(q)) was fit to the interaction energy at each of those points. From this function, the effect of each stretch on the interaction can be included to the surface. Using bra-ket notation, the effect of the ground state vibrational (Ψ_0) motion is accounted for using Eq.

Table 2.16 Mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (COD) for the interpolation of the H_2O-pH_2 interaction as a function of radial sphere.

R (Å)	MAE (cm^{-1})	RMSE (cm^{-1})	COD
2.5	0.05317	0.067539	0.999914
2.75	0.03326	0.04254	0.999885
3	0.016137	0.020537	0.999915
3.25	0.006548	0.009588	0.999953
3.5	0.003059	0.004161	0.999977
3.74	0.003011	0.003736	0.999949
3.79	0.003064	0.00377	0.999935
3.84	0.002915	0.003595	0.999924
4	0.002337	0.003012	0.999883
5	0.00044	0.000548	0.999928
6	0.000136	0.000173	0.999972
7	0.000169	0.000223	0.999099
8	0.00024	0.000298	0.983595
9	7.15E-05	0.000104	0.987888

2.29 while the effect of the first excited state (Ψ_1) is accounted for using Eq. 2.30. These integrals correspond to the expectation value for the perturbation of the stretch of the water molecule.

$$\langle \Psi_0 | F(q) | \Psi_0 \rangle \tag{2.29}$$

$$\langle \Psi_1 | F(q) | \Psi_1 \rangle \tag{2.30}$$

To include these perturbed interaction energies and construct the four potential energy surfaces, some amount of each normal mode must be included. For the PES which has each normal mode in the ground vibrational state, the values calculated using Eq. 2.29 for each stretch are added to the interaction energy at the static geometry. Likewise, for one of the other surfaces, the values calculated using Eq. 2.30 for one stretch are added to the interaction energy at the static geometry alongside the values calculated using Eq. 2.29 for the other two stretches. This was then repeated for each of the respective stretches to construct four surfaces each for the H_2O-H and H_2O-pH_2 interactions. For the H_2O-H surfaces, the EST calculations were performed at the CCSD(T)/CBS/DK level theory, as discussed above. For the H_2O-pH_2 surfaces, however, a composite scheme was utilized as the addition of the extra H-atom (and subsequent basis functions) and an additional electron make the calculations more computationally intensive. Likewise, the requirement of seven unique H_2O-H_2 orientations per point in space to approximate the spherical nature of the pH_2 molecule make these calculations more demanding. This composite scheme combines the coupled cluster approach and second-order, Møller-Plesset (MP2) perturbation theory.[150] The static geometry was calculated at the CCSD(T)/CBS/DK level of theory while the calculations which examined the vibrational perturbations were calculated at the MP2/CBS/DK level of theory with corrections at the CCSD(T)/TZ/DK level of theory. The eight surfaces generated were then tested using the Lebedev quadrature scheme to ensure that interpolation could still be achieved and COD values greater than 0.999 were achieved at all surfaces which indicates excellent reproduction.

2.3.4 Dynamical Calculations

Program Scheme

In order to utilize the potential energy surfaces and curves discussed previously to examine the rovibrational nature of a H_2O in different pH_2 HCP environments, dynamical simulations must be performed. As this type of simulation has not been performed for this type of environment, a program must be developed to perform said simulations. The schematic for the developed program is outlined below in five steps.

1 – Initialize the HCP lattice

By initializing the HCP lattice, several aspects of the calculation are controlled such as the lattice constant (3.79 Å for pH_2), the size of the lattice (including either the first, second, or third nearest shells of the HCP lattice around the H_2O molecule), and the location of an H-atom substitution or a vacancy.

2 – Precalculate the interaction energies given a specific HCP lattice and rotational quadrature

To examine the rotational dynamics of a H_2O molecule in a simulated pH_2 matrix, the H_2O molecule must rotate within the lattice. This is done numerically using Euler angles which control the orientations of the H_2O molecule in space. The three Euler angles (α, β, γ) define the H_2O molecule's orientation within a space-fixed frame. Evaluation of the H_2O rotational wave functions and the H_2O molecule's interaction with the matrix are carried out via rotational quadrature. This rotation is systematic throughout the course of a calculation (the Euler angles for each quadrature node are kept constant for a specific calculation) and so all interaction energies as a function of the Euler angles can be precomputed for computational efficiency. The total number of interaction energies to be computed is a product of the Gaussian smearing for the H-atoms and pH_2 molecules, the number of sites in the lattice, and the number of points utilized for the rotational quadrature. Thus, as the size of both the lattice and the rotational quadrature are increased, so too does the computational complexity increase. These interaction

energies are stored alongside the corresponding weights for the specific Euler angles to be used later to populate the potential energy terms of the Hamiltonian matrix (see below). The quadrature nodes and weights for the α and γ Euler angles are based on evenly spaced nodes between [0, 2π). The quadrature nodes and weights for the β Euler angle are based on Legendre polynomials and utilize a Gauss-Legendre Quadrature scheme of order N.

3 – Populate the potential energy terms of the Hamiltonian matrix

Each entry in the Hamiltonian matrix is a function of several different quadrature schemes used to account for not only the rotational nature of the H_2O molecule but also the integration of the wave functions. When a rotation operator is applied to an asymmetric top molecule such as water, the elements of the Hamiltonian matrix are known as Wigner D functions.[151,152] If the rotation of said asymmetric top molecule is specified using Euler angles (as mentioned previously), the normalized Wigner D functions can be decomposed into three components shown in Eq. 2.31, below, where J, K, and M are the rotational quantum numbers. Here, J corresponds to the total angular momentum, K to the molecule-fixed component of J, and M to the space-fixed component of J.

$$\begin{aligned} D^J_{MK} &= \sqrt{\frac{2J+1}{8\pi^2}}\left\langle JM\middle|e^{-i\alpha\hat{J}_z}e^{-i\beta\hat{J}_y}e^{-i\gamma\hat{J}_z}\middle|JK\right\rangle, \\ &= e^{-i(M\alpha+K\gamma)}d^J_{MK}(\theta), \\ d^J_{MK}(\theta) &= \left\langle JM\middle|e^{-i\beta\hat{J}_y}\middle|JK\right\rangle. \end{aligned} \tag{2.31}$$

The third line of Eq. 2.30 represents the small-d portion of the Wigner D function and its explicit form is given as

$$d_{MK}^{J}(\theta) = \sum_{n=n_{min}}^{n_{max}} (-1)^n W_n^{JMK}(\theta) \tag{2.32}$$

where n_{min} and n_{max} are given as zeros or positive integers according to Eq. 2.33 and 2.34

$$n_{min} = max(0, K - M), \tag{2.33}$$

$$n_{max} = min(J - M, J + K) \tag{2.34}$$

and

$$W_n^{JKM} = w_n^{JMK}(\cos\frac{\theta}{2})^{2J+K-M-2N}(-sin\frac{\theta}{2})^{M-K+2N}, \tag{2.35}$$

$$w_n^{JMK} = \frac{\sqrt{(J+M)!\,(J-M)!\,(J+K)!\,(J-K)!}}{(J-M-n)!\,(J+K-n)!\,(n+M-K)!\,n!}. \tag{2.36}$$

Given this formulation for the Wigner D functions, the construction of each entry in the potential energy portion of the Hamiltonian matrix is straightforward. Utilizing bra-ket or Dirac notation, a specific entry in the Hamiltonian matrix can be described as a summation over the Euler angles and as a function of the rotational quantum numbers J, K, and M for the bra and J`, K`, and M` for the ket. The form of each entry is given according to Eq. 2.37, below,

$$H_{JMK,J`M`K`}^{Potential} = \sum_{\alpha,\beta,\gamma} \langle \Psi_{JMK} | H_R | \Psi_{J`M`K`} \rangle, \tag{2.37}$$

where the rotational Hamiltonian (H_R) is simply a product of the total interaction energy and the weight for that given set of Euler angles and

$$\langle \Psi_{JMK} | = D_{mk}^{J}, \tag{2.38}$$

is given as the normalized Wigner D function (Eq. 2.31). This summation is then evaluated for all entries in the Hamiltonian matrix to arrive at the now complete potential energy portion of the Hamiltonian matrix.

4 – Populate the kinetic energy terms of the Hamiltonian matrix

The kinetic energy terms of the Hamiltonian matrix are easier to compute than those of the potential energy terms and implicitly account for the vibrational stretches of water. Each term is given according to either Eq. 2.39 or Eq. 2.40 if the K-terms of the entry are identical or differ by ±2, respectively.

$$H_{JMK,JMK}^{Kinetic} = \frac{1}{2}(B+C)[J(J+1) - K^2] + AK^2 \tag{2.39}$$

$$H_{JMK,JM(K`=K\pm2)}^{Kinetic} = \frac{1}{4}(B-C)\sqrt{J(J+1) - K(K\pm1)}\sqrt{J(J+1) - (K\pm1)(K\pm2)} \tag{2.40}$$

Here, A, B, and C correspond to the rotational constants for a given vibrational state. As mentioned previously, the H_2O-H and H_2O-pH_2 potential energy surfaces were extended four-fold to implicitly account for the vibrational modes of the H_2O molecule in different vibrational states. A H_2O molecule will have different rotational constants depending upon its vibrational state . If the molecule is in the ground vibrational state for all three vibrational modes, then the pure rotational constants may be used. If, however, the molecule is in a vibrationally excited state, then the rotational constants must be recalculated for said state. The rotational constants for a H_2O molecule in the ground state and each of the first excited states for each vibrational mode are shown in Table 2.17 alongside the vibrational constant (ω) for each vibrational mode.[153]

5 – Analyze Eigenvectors and Eigenvalues

Once the potential and kinetic energy terms have been evaluated for a given Hamiltonian matrix, the terms can be added together to yield a total matrix which can then be diagonalized and analyzed. The values of interest are those of the lowest four energy states which correspond to the rotational ground state and the first three rotational excited states for the H_2O molecule in a given environment and a given vibrational state.

Results and Discussion

Stark Effect

As a way to ensure that the implementation of the Wigner D functions was done correctly, the effective rotational Hamiltonian (H_R) in Eq. 2.37 can be modified to reproduce the Stark Effect.[154] This effect arises due to an applied electric field which breaks the degeneracy of certain states. The Stark effect is observable as a shift in the energy levels which is either linearly proportional or quadratically proportional to the magnitude of the applied electric field. Eq. 2.41 shows the explicit form of the new effective Hamiltonian where E is the applied electric field

Table 2.17 Rotational constants and vibrational constants for a H_2O molecule in different vibrational states. All values given in cm^{-1}.

	State (v_1, v_2, v_3)			
	(0, 0, 0)	(1, 0, 0)	(0, 1, 0)	(0, 0, 1)
A	27.857256	27.138860	31.084730	26.630360
B	14.514488	14.299100	14.674810	14.422540
C	9.279863	9.101500	9.136060	9.141830
ω	0.0000000	3657.053210	1594.746300	3755.928700

and μ is the dipole moment of water (1.85 D). Fig. 2.18 shows how an applied electric field perturbs the degeneracy of the first ten states of a rotating H_2O molecule. The first excited state of a rotating water molecule in the absence of an external field is three-fold degenerate. With an applied electric field, the degeneracy of these states is broken and the energy gap begins to widen. A similar effect is observed for the higher excited states. The behavior observed in Figure 2.18 is similar to that of work done by Andreev *et al*. which indicates to us that our mathematical implementation is correct.[155]

$$H_R = \mu E \sin(\beta) \cos(\gamma) \tag{2.41}$$

Effect of Rotational Quadrature

As with grids in density functional theory calculations, the choice in the number of points examined for a rotating H_2O molecule must be examined for quality and convergent behavior. This will be referred to as the rotational quadrature of the system. This quadrature is based upon Euler angles (as mentioned previously) which play a direct role in the calculation of the Wigner D functions for a rotating asymmetric top. For large rotational quantum numbers (J), the calculation of the rotational integrals requires the use of highly oscillatory functions to approximate the Wigner D function to a high enough degree of accuracy. Thus, as the Hamiltonian matrix is expanded with larger J-values, a larger rotational quadrature must be utilized. This is akin to requiring more points on a curve in order to fit a higher order polynomial to said curve.

We initially tested the rotational quadrature to ensure completeness could be achieved. Here, completeness refers to the quadrature scheme resulting in a perfect reproduction of the identity matrix given no external perturbation. We tested rotational quadrature values from one to twenty-five for J_{max} values up to eight. At a given J-value, the completeness of the quadrature scheme was achieved using rotational quadrature values greater than or equal to 2J+1. For example, with a J-value of one, the minimum rotational quadrature to be used should be three and for a J-value of two, a rotational quadrature of five should be used, and so on. Thus, this expression serves as the baseline minimum for all dynamics simulations.

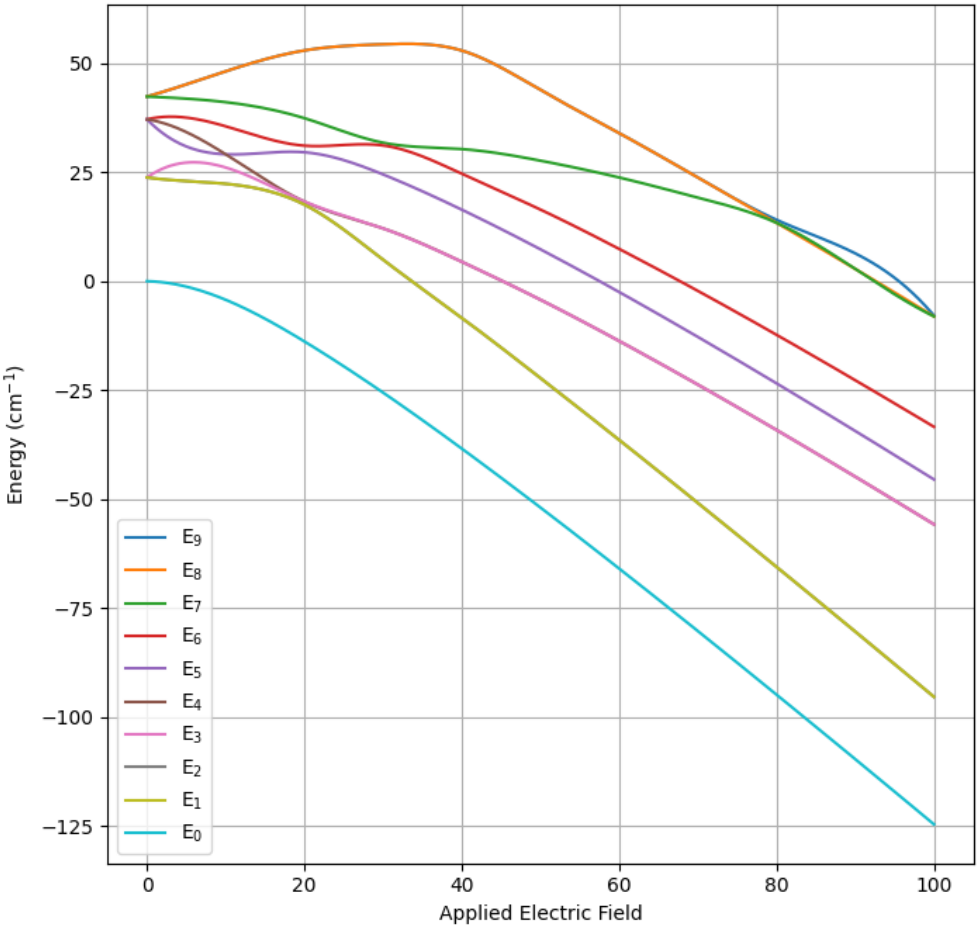


Figure 2.18 Energy level diagram for the first 10 states of a rotating H_2O molecule with a varying applied electric field. A rotational quadrature of 10 and J-value of 1 were utilized.

Likewise, convergence must be checked with the applied, effective Hamiltonian. Here, convergence was checked by examining the eigenvalues of the resulting Hamiltonian matrix to ensure that their values do not change to a great degree after increasing the rotational quadrature. This behavior must be examined for every J-value in order to establish a baseline minimum rotational quadrature for a given Hamiltonian matrix in addition to the minimum previously discussed. This convergent behavior is shown in Figure 2.19 where a single shell of the pH_2 HCP lattice was utilized and a J-value of 1 was chosen. It was found that for J-values up to J = 7, a rotational quadrature of 15 leads to sufficient convergence regarding the lowest four energy states. Thus, all subsequent calculations utilizing a single shell of the pH_2 HCP lattice will be based upon a minimum rotational quadrature value of 15.

Effect of the Expansion in Rotational States

Just as with the rotational quadrature, the size of the Hamiltonian matrix must be checked for quality and convergence. This is best understood as an analogue to the size of the basis set for standard *ab initio* calculations. As more one-electron basis functions are included in each calculation such as a Hartree Fock calculation, the total electronic energy decreases in an exponential fashion and given a large enough number of basis functions, the complete basis set (CBS) limit can be approximated. A similar effect can be observed with regards to the rotational Hamiltonian where increasing the maximum J-value is akin to increasing the size of the chosen basis set. As more states are included in a calculation, a better approximation to the ground state and the first three excited states can be made. Thus, a J-value of one represents a minimum basis set and any J-value greater than one would lead to a better approximation to the lowest four states. As the difference in energy between the ground state and the first three excited states are of great interest, these values are shown in Tables 2.18, 2.19, and 2.20 for rotational quantum numbers of fifteen, twenty, and twenty-five, respectively, as functions of J for a rotating H_2O molecule in a HCP lattice of pH_2 molecules. Herein, these three energy gaps will collectively be referred to as the transition energies of the rotating H_2O molecule.

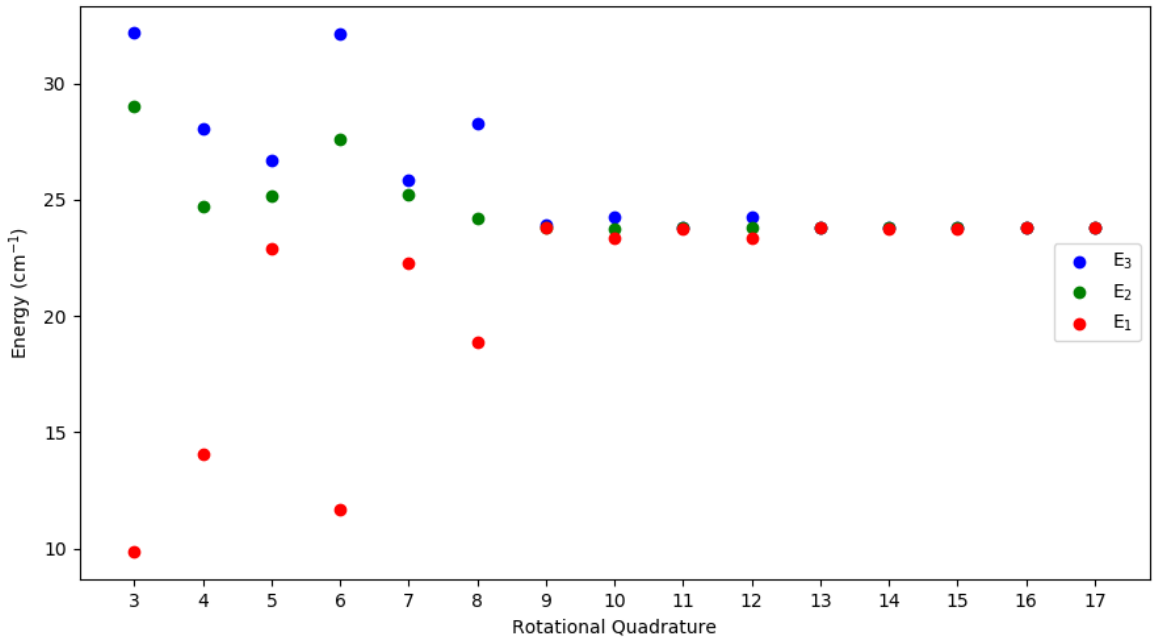


Figure 2.19 The first three transition energies (E_1, E_2, E_3) of a H_2O molecule rotating in a single-shell, HCP lattice of pH_2 molecules as a function of the rotational quadrature. A J-value of one is used.

Table 2.18 Transition energies between the ground state and the first three excited states (E_1, E_2, and E_3) for J-values up to J = 7 using a single shell of pH_2 HCP lattice sites. A rotational quadrature of fifteen is used. Values shown in cm^{-1}.

J_{max}	E_1	E_2	E_3	Average	Standard Deviation
1	23.825	23.787	23.771	23.794	0.028
2	23.825	23.787	23.771	23.794	0.028
3	23.819	23.776	23.752	23.782	0.034
4	23.833	23.789	23.766	23.796	0.034
5	23.828	23.686	23.643	23.719	0.097
6	23.894	23.814	23.771	23.826	0.063
7	23.878	23.763	23.688	23.776	0.095

Table 2.19 Transition energies between the ground state and the first three excited states (E_1, E_2, and E_3) for J-values up to J = 8 using a single shell of pH_2 HCP lattice sites. A rotational quadrature of twenty is used. Values shown in cm^{-1}.

J_{max}	E_1	E_2	E_3	Average	Standard Deviation
1	23.801	23.798	23.784	23.794	0.009
2	23.801	23.798	23.784	23.794	0.009
3	23.794	23.780	23.772	23.782	0.011
4	23.808	23.794	23.786	23.796	0.011
5	23.804	23.715	23.638	23.719	0.083
6	23.869	23.844	23.767	23.827	0.053
7	23.853	23.793	23.683	23.776	0.086
8	23.854	23.799	23.690	23.781	0.083

Table 2.20 Transition energies between the ground state and the first three excited states (E_1, E_2, and E_3) for J-values up to J = 8 using a single shell of pH_2 HCP lattice sites. A rotational quadrature of twenty-five is used. Values shown in cm^{-1}.

J_{max}	E_1	E_2	E_3	Average	Standard Deviation
1	23.798	23.796	23.788	23.794	0.005
2	23.798	23.796	23.788	23.794	0.005
3	23.791	23.785	23.770	23.782	0.011
4	23.805	23.799	23.784	23.796	0.011
5	23.801	23.705	23.652	23.719	0.076
6	23.866	23.834	23.781	23.827	0.043
7	23.849	23.782	23.697	23.776	0.076
8	23.851	23.788	23.705	23.781	0.073

With a rotational quadrature of fifteen, J-values of one, two, three, and four appear to show good agreement with regards to the degeneracy of the photon energies with standard deviations not exceeding 0.034 cm^{-1}. J-values between five and seven, however, show standard deviations up to 0.097 cm^{-1} which indicates that convergence in the transition energies has not been achieved and thus, higher rotational quadrature values are required for these J-values. When the rotational quadrature is increased to twenty, J-values of one, two, three, and four show even lower standard deviations among the transition energies while only a slight improvement is observed for higher J-values. This behavior is again seen with a rotational quadrature of twenty-five. Thus, we believe that convergence is achieved up to a J-value of four for a rotational quadrature of twenty-five with regards to the three-fold degeneracy of the transition energies. Higher J-values would require even higher rotational quadrature values.

Effect of the Size of the Simulated HCP Lattice

As written, the dynamical simulations are able to algorithmically generate an HCP lattice of arbitrary size, both in the choice of the lattice constant and number of sites. Thus, we can examine the effect that including more sites in a calculation will have on the calculated transition energies. We performed the same analysis as before with two shells of the pH_2 HCP lattice and examined how the transition energies for a given rotational quadrature and J-value changed with the inclusion of more pH_2 sites. The second shell of pH_2 sites reside between 3.79 Å and 7.58 Å which are still near the non-spherically symmetric region of the potential energy surface. If calculations are required to be performed with even more sites beyond the shell, then a spherical potential may be used to approximate the effects of those sites. However, this was not necessary based upon results utilizing a second shell of pH_2 sites. Transition energies up to a J-value of seven for a rotational quadrature value of fifteen are shown in Table 2.20 and up to a J-value of eight for a rotational quadrature value of twenty in Table 2.21. These results show similar trends as in the single-shell case (Tables 2.17-2.19) except the standard deviations are slightly larger in most cases. Shown in Table 2.22 is the average transition energy for J-values up to eight for both a single and double shell of the HCP lattice. Based upon the comparison detailed in Table 2.22, the inclusion of the second shell of the pH_2 HCP lattice does little to change the average transition

Table 2.21 Transition energies between the ground state and the first three excited states (E_1, E_2, and E_3) for J-values up to J = 7 using a double shell of pH_2 HCP lattice sites. A rotational quadrature of 15 is used. Values shown in cm^{-1}.

J_{max}	E_1	E_2	E_3	Average	Standard Deviation
1	23.844	23.822	23.717	23.794	0.068
2	23.844	23.822	23.717	23.794	0.068
3	23.834	23.817	23.697	23.783	0.074
4	23.847	23.831	23.710	23.796	0.074
5	23.826	23.709	23.636	23.724	0.096
6	23.888	23.828	23.756	23.824	0.066
7	23.873	23.748	23.710	23.777	0.086

Table 2.22 Transition energies between the ground state and the first three excited states (E_1, E_2, and E_3) for J-values up to J=8 using a double shell of pH_2 HCP lattice sites. A rotational quadrature of 20 is used. Values shown in cm^{-1}.

J_{max}	E_1	E_2	E_3	Average	Standard Deviation
1	23.840	23.798	23.745	23.794	0.048
2	23.840	23.798	23.745	23.794	0.048
3	23.830	23.793	23.725	23.783	0.053
4	23.843	23.806	23.738	23.796	0.053
5	23.801	23.704	23.665	23.723	0.070
6	23.863	23.825	23.787	23.825	0.038
7	23.848	23.743	23.740	23.777	0.062
8	23.849	23.751	23.746	23.782	0.058

Table 2.23 Average Transition energies for rotational quadrature values of 15 and 20 for both a double shell of the pH_2 HCP lattice up a J-value of eight. Values shown in cm^{-1}

J_{max}	Rotational Quadrature - 15		Rotational Quadrature - 20	
	Single Shell	Double Shell	Single Shell	Single Shell
1	23.794	23.794	23.794	23.794
2	23.794	23.794	23.794	23.794
3	23.782	23.783	23.782	23.783
4	23.796	23.796	23.796	23.796
5	23.719	23.724	23.719	23.723
6	23.826	23.824	23.827	23.825
7	23.776	23.777	23.776	23.777
8	-	-	23.781	23.782

energies for any given J-value up to eight and for either rotational quadrature values. In addition, a second shell results in a larger standard deviation for the transition energies which does not systematically decrease with an increase in the rotational quadrature value. Thus, a single shell of the pH_2 HCP lattice is sufficient to study the transition energies for a rotating H_2O molecule given a high enough rotational quadrature and J-value.

Effect of H-atom Substitutions and Vacancies

As hypothesized by Kufeld *et al*,[133] the two satellite peaks observed around the antisymmetric stretch of H_2O in a pH_2 lattice may arise from either a vacancy or a substitution in the HCP lattice directly around the H_2O molecule. Thus, an in-depth analysis of the effect of a vacancy or an H-atom substitution on the transition energies of the rotating H_2O molecule must be performed. The HCP lattice around the H_2O molecule is essentially spherically symmetric and so a perturbation from either a vacancy or a substitution must be nearly the same regardless of which of the twelve HCP sites around the H_2O molecule is involved. This means that a substitution or vacancy at any site in the first shell of the pH_2 HCP lattice should result in identical transition energies. In addition to checking for the convergence with regards to the J-value and minimum rotational quadrature as was done with the pristine HCP lattice, the symmetry of the substitution sites must also be checked. To examine this, each of the twelve sites in the first shell of the pH_2 HCP lattice are replaced with either an H-atom substitution or a vacancy and the first four energy levels are evaluated. The standard deviation in the ground state energy level of the twelve sites is used as the metric of spherical symmetry and these values are shown in Table 2.24 for rotational quadrature values of fifteen, twenty, and twenty-five for J-values up to eight for both H-atom substitutions and vacancies. Immediately apparent is the larger standard deviation values for the vacancies as opposed to the H-atom substitutions which may arise due to the larger perturbation that a vacancy provides. Very little change is observed with regards to the size of the Hamiltonian matrix, however. As the J-value is increased from one to eight, the standard deviation values for the H-atom substitutions roughly doubles for each of the three rotational quadrature values. This value remains small however, below 0.03 cm^{-1}.

Table 2.24 Standard deviations of the ground state energies over the twelve sites of a single shell of a pH_2 HCP lattice for both H-atom substitutions and vacancies for rotational quadrature values of 15, 20, and 25. Values shown in cm^{-1}.

J_{max}	Rotational Quadrature - 15		Rotational Quadrature - 20		Rotational Quadrature - 25	
	H-Atom	Vacancy	H-Atom	Vacancy	H-Atom	Vacancy
1	0.015	0.167	0.014	0.164	0.013	0.105
2	0.027	0.168	0.025	0.164	0.026	0.165
3	0.027	0.172	0.025	0.168	0.026	0.165
4	0.029	0.172	0.028	0.169	0.028	0.169
5	0.029	0.172	0.028	0.168	0.028	0.170
6	0.029	0.172	0.028	0.169	0.028	0.169
7	0.029	0.173	0.028	0.169	0.028	0.170
8	-	-	0.028	0.169	0.028	0.170

Likewise for the vacancy sites except the standard deviation values are larger in magnitude but still relatively small (~0.1 – 0.17 cm^{-1}). Based upon the results shown in Table 2.23, very little change is observed in the choice of substitution site and thus, the choice of rotational quadrature and J-value may be made based upon the results gathered from the single-shell, pristine HCP lattice.

Rovibrational Excitations

Based upon the results of the previous analysis, a J-value of four is used in conjunction with a rotational quadrature of twenty-five as this combination led to convergence with regards to the lowest four states of the unperturbed pH_2 HCP lattice. In addition, this combination resulted in low standard deviations the for the H-atom substitution and vacancy sites. The transition energies for the perturbed systems will be averaged over the twelve substitution sites, however, to ensure the spherical symmetry of the HCP lattice is properly accounted for.

In order to compare the unperturbed HCP lattice, the H-atom substituted lattice, and the lattice with a vacancy site, multiple dynamical simulations were performed. First, each of the four surfaces were examined with regards to the implicit vibrational modes of the H_2O molecule. This was done as the R(0) peak of the IR spectra of Kufeld *et al.*[133] is composed of a vibrational and rotational excitation. Thus, to accurately capture the underlying physics within our simulations, we must examine the effect of the different (implicit) vibrational modes on the rotational states. Next, we calculated the transition energies for the different lattice environments for each vibrational state (shown in Table 2.25). The antisymmetric stretch of water is shown in the second set of rows of Table 2.25 as vibrational state (0,0,1) and is the mode which corresponds to R(0) from the spectra obtained by Kufeld. The transition energies for the HCP lattice geometry show the expected degeneracy as the excitation should be anisotropic given the symmetry of the lattice. The two defect geometries, however, do not show the same degeneracy which is expected given the change in symmetry.

The E_1, E_2, and E_3 peaks correspond to excitations from the 0_{00} rotational state to the 1_{00} state which should correspond exactly to the R(0) peak in the HCP lattice geometry with the

Table 2.25 Photon energies between the ground state and the first three excited states (E_1, E_2, and E_3) for the ground vibrational state of H_2O. Values shown in cm^{-1}.

State (v_1, v_2, v_3)		Environment		
		$H_2O[pH_2]_{12}$	$H_2O[pH_2]_{11}H$	$H_2O[pH_2]_{11}$
(0,0,0)	E_3	23.81	24.66	25.73
	E_2	23.80	23.81	23.87
	E_1	23.78	22.95	22.00
(0,0,1)	E_3	23.41	24.30	25.51
	E_2	23.41	23.42	23.52
	E_1	23.39	22.49	21.26
(1,0,0)	E_3	23.82	24.67	25.55
	E_2	23.81	23.83	23.86
	E_1	23.80	22.99	22.20
(0,1,0)	E_3	23.59	24.58	25.86
	E_2	23.58	23.59	23.65
	E_1	23.54	22.64	20.98

addition of a vibrational excitation. The two satellite peaks observed by Kufeld *et al.* around the R(0) peak should correspond to slightly perturbed $1_{00} \leftarrow 0_{00}$ excitations. For the asymmetric stretch of water (0,0,1), the 1_{00} peak is blue-shifted by 0.81 and 1.60 cm^{-1} for the H-atom substitution and vacancy environments, respectively and red-shifted by 0.85 and 1.73 cm^{-1} for the same environments. Based upon these shifts, we believe that the satellite peaks observed by Kufeld *et al*. arise from vacancy sites directly around the H_2O molecule and not H-atom substitutions for two reasons: 1) the shift in the peaks is larger in magnitude for the vacancy environment and 2) there is a greater asymmetry between the two shifts when compared to the H-atom substituted peaks. The satellite peaks observed by Kufeld *et al*. are shifted by -1.21 and 2.03 cm^{-1} relative to the R(0) peaks which more closely matches the behavior in the peaks from the vacancy geometry.. The presence of an H atom directly around a H_2O molecule in the pH_2 lattice may have some contribution to the change in the IR spectra but only slightly. The calculated shift in the 1_{00} peak in the H atom substitution geometry is small and thus may only act to slightly broaden the R(0) peak and not contribute to the two satellite. Likewise, the lifetime of an H atom found directly around an H_2O molecule in the pH_2 lattice may be very short and thus may contribute very little to the overall IR spectra. This short lifespan is an effect of the diffusion of the H-atom through the pH_2 lattice which is facilitated via a tunneling reaction ($H + H_2 \rightarrow H_2 + H$).[156–159] The H-atom substitutions are less likely to occur around the H_2O dopant molecule as the diffusion process is affected by the energy of the initial and final states of a single diffusion step. If the initial and final states differ in energy due to different environmental conditions, such as the presence of an H_2O molecule, the diffusion process is much less likely to occur.[160] Thus, once an H_2O molecule has formed in the lattice, the chances of it being near an H-atom are quite low, further limiting the effect that such an interaction would have on the antisymmetric stretch of H_2O. In addition, due to the stronger interaction between an H_2O and pH_2 molecule, it is more thermodynamically favorable for an H_2O molecule to be completely surrounded by pH_2 molecules as opposed to having a single H-atom substitution directly around the H_2O molecule.

2.3.5 Conclusions

Through an extensive development of potential energy surfaces and potential energy curves as well as dynamical simulations, two satellite peaks observed in infrared spectra of a pH_2, HCP lattice doped with H_2O molecules may be identified. Both the H-pH_2 and pH_2-pH_2 interactions were examined and it was found that the minimum energy distance between the monomers was at 3.4501 and 3.4858 Å, respectively with interaction energies equal to 16.02 and 22.26 cm^{-1}. No vibrationally bound state was identified for the H-pH_2 van der Waals dimer as the interaction energy was found to be too small. Thus, this dimer would not persist at any temperature. Conversely, the pH_2-pH_2 van der Waals dimer was found to be bound by 2.55 cm^{-1}. To account for the large amplitude, zero-point vibrational motion of both the H-atom and the pH_2 molecule in the solid state, a Gaussian probability distribution was utilized in the construction of the potential energy surfaces of the H_2O-H and H_2O-pH_2 systems. The minimum energy geometry for these two van der Waals dimers was found to lie in the plane of the H_2O molecule at an angle of 122.25° and 101.4° with respect to the x-axis, respectively, at distances of 3.41 and 3.2059 Å with interaction energies of 61.297 and 88.523 cm^{-1}. In the development of both 3-dimensional surfaces, we included perturbations to account for the vibrational modes of the H_2O molecules. These surfaces were then used for dynamical simulations to examine the rotational excitations of a H_2O molecule in different environments: an HCP lattice of twelve pH_2 molecules, a perturbed HCP lattice of eleven pH_2 molecules and one H-atom, and a perturbed HCP lattice with only eleven pH_2 molecules. These environments and the different vibrational states of H_2O were then examined to see how the $1_{00} \leftarrow 0_{00}$ rotational excitation peak of H_2O changed as a function of environment. It was found that an H-atom substation in the HCP lattice resulted in peak shifts that were too small and too symmetric. Conversely, a vacancy site in the HCP lattice resulted in larger peak shifts that were asymmetric in nature. The results of the latter environment correlate well with the spectrum of Kufeld *et al.* who observed both a red- and blue-shifted peak around the antisymmetric stretch of water that is quite large (1-2 cm^{-1}) and asymmetric. Based upon our results, we believe that the presence of a vacancy directly adjacent to the H_2O molecule is responsible for the satellite peaks observed by Kufeld *et al*.

2.4 Conclusions

Here we have demonstrated that properly developed potential energy surfaces can be reliable and useful for the prediction of spectroscopic properties of small molecules and molecular systems. With regards to diatomic molecules, the 1-dimensional curve of the CF^+ cation was examined using very high levels of electronic structure theory which resulted in errors of 0.00014 Å for the equilibrium bond length and 0.2254 cm^{-1} for the vibrational constant. Using this curve, we were able to generate line lists for many rotational, vibrational, and rovibrational excitations for the ^{12}C and ^{13}C isotopes for both the singlet ground state and the first excited triplet state. The mathematical formulations used for the generation of the line lists for CF^+ is generalizable to other diatomic molecules. Currently, this formulation is being used to examine if there is a correlation between the T_1 diagnostic values for diatomics which have been identified in the interstellar medium (including CF^+) and their respective spectroscopic constants at the CCSD(T)/CBS level of theory and to make predictions for the spectroscopic constants of lanthanide diatomic molecules. For larger more complex potential energy surfaces such as the H_2O-H and H_2O-pH_2 van der Waals dimers, their use lies in simulations. Binding energies and minimum energy geometries may be obtained for these systems but the numerical solution to the interaction as a function of the position of the dimers has allowed us to perform dynamical simulations to better understand the rotational excitations of a H_2O molecule in an HCP lattice of pH_2 molecules. The threefold splitting of the first rotational excitation of a H_2O molecule in an HCP lattice with a single vacancy site was found to correlate well with experimental infrared spectra of a H_2O-doped, pH_2 cryogenic matrix. As an extension of this work, similar potential energy surfaces may be developed for NH_3 to see if the satellite peaks for the NH_3 doped crystal lattice are also a product of a vacancy site around the dopant molecule.

Chapter 3

The Computational Study of Lanthanides from Diatomic Molecules to Larger, Molecular Complexes

The following chapter examines how electronic structure theory calculations and machine learning can be used to better understand the chemistry of the lanthanide series. A benchmarking study is presented first where many combinations of basis set, density functional, and relativistic approximation are used to determine the optimal method by which lanthanides can be studied using computational methods. Following this is work in which electronic structure theory is used in tandem with machine learning to predict the free binding energies of lanthanide-tris-β-diketone complexes, which can lead to the development of machine learning models for the prediction of energies of thousands of lanthanide complexes.

3.1 Introduction

The electronic structure of f-block elements is quite unique when compared to either d-block or heavy p-block elements. Because of this, many modern technologies such as catalytic cracking[161] and bioimaging[162] have been developed and/or improved upon through the inclusion of f-block elements. The lanthanides in particular show unique magnetic properties due to their propensity towards high spin states[163] across the series and while the actinides do not exactly share this same behavior, they are also used extensively due to their unique nuclear structure.[164] Just as with d-block and certain p-block metals, f-block elements are also used as catalysts such as those

found in hydrocarbon cracking,[161] in the Baeyer-Villiger oxidation reaction,[165] and other chemical transformations.[166–168] While lanthanides have not been studied to the same degree as the transition metals, a better understanding of the unique electronic structure along the 4f series would nevertheless allow for an improvement for certain properties and thus an improvement in some modern technologies.

Any technological application requires very pure starting materials. For example, for transition metals such a nickel, its separation and extraction is straightforward as nickel carbonyl boils at a relatively low temperature (43 °C) and thus, it can be easily separated from other metal carbonyls.[169]On the contrary, lanthanides are more difficult to be separated and purified due to the very similar physicochemical properties of adjacent elements in the series. The current technique that allows for the most efficient separation among the lanthanides involves a series of subsequent liquid extraction steps.[170] During these steps, an aqueous solution of different lanthanide cations is mixed with an immiscible, organic-phase solution which often contains a chelating ligand. Once mixed, the two phases temporarily allow for the chelating agent to interact with the lanthanide cations which is often a thermodynamically favorable process. Afterwards, the two phases are allowed to separate and because of the favorable chelating interaction, some of the lanthanide cations will be found in the organic-phase solution. The chelating interaction is not identical across the lanthanide series, however, and a distribution is formed whereby more of some lanthanides are found in the organic phase than others. This process is then repeated numerous times to further improve the purity of the lanthanide-based solutions.

Research on lanthanide separation is often focused on the refinement of the interaction with the chelating agent. Different families of chelating agents have historically been examined for the separation of lanthanide cations and includes organophosphorus acids such as bis(2-ethylhexyl)phosphate (HDEHP), carboxylic acids such as different versatic acids, and amines such as quaternary ammonium salts.[171] Modern research focuses on different families of chelating agents such as diglycolamides,[172–180] phosphine oxides,[181–188] and phenanthrolines[189–194] which have all been shown to effectively separate adjacent lanthanide cations both from experimental and computational standpoints. Through a combination of experimental and computational methods, a synergistic approach can be taken whereby many different ligand scaffolds are

examined using computational methods. Computations can reveal ligands with enhanced properties and desirable separation performance between different lanthanide cations, while experimental work can either verify or disprove the theoretical findings. This hybrid approach allows for the examination of many hundreds if not thousands of ligand scaffolds in a manner more efficient than can be achieved purely through experiment. For the exploration of this multitude of ligand scaffolds, an efficient and accurate computational protocol is required. Such a protocol is examined in the following section.

3.2 Selecting Quantum-Chemical Methods for Lanthanide-Containing Molecules: A Balance between Accuracy and Efficiency

Disclosure

The following subsection contains an adapted reproduction of a published manuscript in *Inorganic Chemistry*.[195] Gavin McCarver performed all work in this manuscript under the guidance of Dr. Konstantinos D. Vogiatzis and Dr. Robert J. Hinde.

3.2.1 Abstract

An analysis of how different density functionals, basis sets, and relativistic approximations affect the computed properties of lanthanide-containing molecules allows one to determine which method provides the highest accuracy. Historically, many different density functional methods have been employed to perform calculations on lanthanide complexes and so herein is a detailed analysis of how different methodological combinations change the computed properties of three different families of lanthanide-containing molecules: lanthanide diatomic molecules (fluorides and oxides) and their dissociation energies; larger, molecular complexes and their geometries; and lanthanide bis(2-ethylhexyl)phosphate structures and their separation free energies among the lanthanide series. The B3LYP/Sapporo/Douglas−Kroll−Hess (DKH) method was shown to most accurately reproduce dissociation energies calculated at the CCSDT(Q) level of theory with a mean

absolute deviation of 1.3 kcal/mol. For the calculations of larger, molecular complexes, the TPSSh/Sapporo/DKH method led to the smallest deviation from experimentally refined crystal structures. Finally, this same method led to calculated separation factors for lanthanide bis(2-ethylhexyl)phosphate structures that matched very closely with experimental values.

3.2.2 Introduction

The separation of rare-earth ions in solution is an important process because of the numerous technological, industrial, and medical uses that exist for lanthanide metals and complexes. These uses include, but are not limited to, fracking catalysts, permanent magnets, contrasting agents in magnetic resonance imaging, and renewable energy technology.[171] Because of their wide use, there is a large demand for high-purity lanthanide compounds; this demand resulted in 210,000 tons of rare earths (lanthanides, yttrium, and scandium) being mined in 2019, with most of the recent production coming from China.[196] The size of this market and the criticality of some of the lanthanides in modern technology demand efficient separation techniques. However, because lanthanide cations have similar physicochemical properties, their separation becomes a very challenging task.[197,198] Solvent extraction is the most common separation process in which the slight differences in the properties along the lanthanide series are exploited to allow preferential extraction of specific lanthanides from a mixture using organic-based ligand extractants.[174,199–204]

Computational methods are often utilized for the design of new ligands for lanthanide separation which provide synthetic chemists with targets that show promise at a theoretical level. The study of such lanthanide-containing molecules is primarily performed using density functional theory (DFT)[205] because of its computational efficiency.[42,206] Previous studies have demonstrated the applicability of DFT to theoretical studies of lanthanide separation; numerous classes of ligand extractants have been examined including diglycolamides,[175,179,180] bis(phosphine) oxides,[207] bis(lactam) phenanthrolines,[208] and others.[209] This area of research is often overshadowed by other studies that focus on the separation of lanthanides from minor actinides; thus, the field of

lanthanide−lanthanide separation using computational methods is not as mature in its study as lanthanide-actinide and actinide-actinide separation.

Previous benchmarking studies showed that for lanthanide diatomic molecules (oxides and fluorides), the choice of basis sets and methodologies that account for relativistic effects affected the accuracy for each density functional studied.[210,211] For hydrated lanthanide species, the TPSS or B3LYP functionals provided higher accuracy when they were used in conjunction with the Stuttgart basis set and effective core potentials (ECPs).[212] Another study examined how using different basis sets and density functionals affects computed lanthanide extraction selectivities and found that a combination of the B3LYP functional and a large-core ECP best matched experimental results when it came to the selectivity and binding energies of four different ligands.[213] These benchmarking studies generally indicated superior performance is achieved by using either a meta-GGA (generalized gradient approximation) or a hybrid density functional, but there are numerous studies that show no clear indication of which single method best represents certain properties of specific lanthanide-containing molecules.[214–217]

Here, a comprehensive benchmarking study on lanthanide-containing molecules is presented. A total of 198 combinations of density functionals, basis sets, and relativistic approximations are examined across three different molecular properties: the bond dissociation energies of lanthanide diatomic molecules, geometric structures of larger lanthanide complexes, and separation factors of the bis(2-ethylhexyl)phosphate (HDEHP) ligand as it relates to early, middle, and late lanthanides (La, Gd, and Lu). The examination of dissociation energies of diatomic molecules allows us to study a wide array of methods with small structures that are computationally cheap. From those results, we can then further probe those methods that were proven to reproduce dissociation energies accurately on structures that more closely resemble molecules seen during the lanthanide separation process. Finally, once we identify the methods that were proven to be accurate for the study of the larger, more chemically relevant structures, these methods can then be applied to a system where separation factors can be calculated and compared to experimental values. Thus, the hierarchy of these calculations leads us to better understand which electronic structure methods best model lanthanide complexes. The last section concludes our results.

3.2.3 Computational Details

All calculations were performed using the ORCA software package (version 4.1).[118] The following basis sets were tested for the lanthanide atoms: SARC,[218] SARC2,[219] Sapporo-TZP,[220] Ahlrichs (def2-TZVP),[221] Stuttgart relativistic small core (RSC),[222] ANO-RCC-VTZP,[223] and cc-pVDZ-DK3.[206] The Ahlrichs and Stuttgart sets include their respective ECPs. Other such basis sets and pseudopotentials have been developed for lanthanide atoms, such as those of Lu *et al*., that aim to explore more complex environments of lanthanide chemistry, but these were not considered for this work.[224] Relativistic effects are very important for compounds containing heavier elements such as lanthanides, and so any electronic structure calculation on complexes containing such elements must account for these effects. While both scalar relativistic effects and spin−orbit coupling are important for lanthanide compounds, only the former are considered for this work because corrections from spin−orbit coupling were shown to be relatively small for the systems studied in the first portion of this work.[225] Three common methodologies for introducing scalar relativistic effects were included: the Douglas−Kroll−Hess (DKH) scheme,[52,53] the zeroth-order regular approximation (ZORA),[50,51,226] and ECPs.[227] Both DKH and ZORA reduce the four-component formalism for the expression of the relativistic Dirac Hamiltonian to the two-component form and, consequently, to the one-component, scalar relativistic variant. DKH approximates the Dirac Hamiltonian by applying a series of unitary transformations, whereas ZORA uses an expansion with respect to a perturbation parameter.[228] For an in-depth analysis of the differences between DKH and ZORA as they relate to f-block element diatomic molecules, we refer the reader to Hong *et al*.[229] The SARC and SARC2 basis sets were developed by considering both the DKH and ZORA relativistic corrections, whereas the Sapporo set was developed only with DKH. Therefore, the Sapporo basis set will only be combined with the DKH approximation and not with ZORA. Basis sets that are not included in ORCA were obtained from the Basis Set Exchange database.[230] A wide range of density functionals were tested in this study: the local density approximation (LDA) SVWN-5 functional,[205,231,232] six GGA functionals (BP86,[233,234] PW91,[235] BLYP,[233,236] PBE,[237] OLYP,[236,238] and OPBE[237,239]), two meta-GGA functionals (TPSS[235] and M06- L[240]), six hybrid functionals (B3LYP,[231,236,241] O3LYP,[239] PBE0,[242] M06,[243] M06-2X,[243] and TPSSh[244]), two long-range hybrid functionals (ωB97-X[245] and CAM-B3LYP[246]),

and a double hybrid functional (B2PLYP[247]). Some of these basis sets (ANO-DK3 and cc-pVDZ-DK3) and density functionals (ωB97-X, CAM-B3LYP, and B2PLYP) are only considered for a portion of the molecules studied in the first section of this work. For all calculations, the high-spin configurations of the lanthanide ions were assumed because those are the ground- state configurations for the lanthanide(III) ions.[206]

For lanthanide oxide (LnO) and fluoride (LnF) molecules, different basis sets were used for the oxygen and fluorine atoms, depending on the basis set chosen for the lanthanide as not all of the basis sets used for the lanthanide atoms were available for the lighter atoms. This also allowed us to reproduce results from previous studies.[215] SARC and SARC2 basis sets were used for the lanthanides in combination with the def2-TZVP basis sets for oxygen and fluorine, while the Stuttgart lanthanide basis sets were paired with the aug-cc-pVDZ basis set for the oxygen and fluorine atoms (see the upper part of Table 3.1). When the def2-TZVP and Sapporo sets were used for the lanthanides, then the same basis sets were used for the oxygen and fluorine atoms as well. The resolution-of-identity (RI)[248] approximation, as implemented in ORCA, was used for all calculations with the corresponding auxiliary basis sets.[249] Dissociation energies (D_e) were calculated by determining the difference in the electronic energy between the diatomic at its optimized geometry and the energy of the two infinitely separated substituent, neutral atoms. The ground-state electronic configurations of the lanthanide ions were taken from Peterson *et al.*[206] (assuming the Ln^+ charge state for the fluorides and Ln^{2+} for the oxides). In addition, frequency calculations were performed on optimized geometries to acquire harmonic zero-point vibrational energies (ZPVEs) so that experimental dissociation energies (D_0) could be converted to values directly comparable to the theoretical values. For all diatomics, the ZPVEs were about 1 kcal/mol.

Two different triple-ζ quality basis sets were used for the lanthanide atoms in the larger molecular complexes considered in this study. A triple-ζ basis set was used for all atoms in the first coordination sphere of the lanthanide and a double-ζ basis set for all other atoms (lower part of Table 3.1). The RI approximation was again used for these calculations. Geometry optimizations were performed for all structures, and frequency calculations ensured that a minimum in the potential energy surface was found.

Table 3.1: Basis sets used for all calculations performed for the LnO and LnF diatomic species (upper part) and lanthanide molecular complexes (lower part). Included in parentheses in the upper part are the contraction schemes for the lanthanide basis sets.

Diatomic Molecules

Lanthanide	**Contraction Scheme**	**Oxygen, Fluorine**
SARC/SARC2	18s12p9d3f/18s12p9d4f	def2-TZVP
Sapporo-TZP	11s9p7d5f3g1h	Sapporo-TZP
def2-TZVP	6s4p3d1f	def2-TZVP
Stuttgart (RSC)	10s8p5d4f3g	aug-cc-pVDZ
ANO-DK3	6s4p2d1f	ANO-DK3
cc-pVDZ-DK3	8s7p5d3f1g	cc-pVDZ-DK

Lanthanide Complexes

Lanthanide	**1st coordination sphere**	**All other atoms**
SARC/SARC2	def2-TZVP	def2-SV(P)
Sapporo-TZP	Sapporo-TZP	Sapporo-DZP

3.2.4 Results and Discussion

Diatomic Molecules

Because of the large number of methods considered in this study, diatomic molecules were employed for a prescreening of different levels of theory. The test set includes certain lanthanide diatomic molecules (oxides and fluorides) with experimentally measured dissociation energies.[210,211,250,251] The dissociation energies for each of the 25 diatomics (with ZPVE corrections) for each method are reported in the Appendix 3. To gauge the performance of the methods under consideration, the mean deviation and mean absolute deviation (MAD) are discussed. These two values are defined as the average of the signed differences and average of the absolute differences from the reference values, respectively.

We have considered two different sets of dissociation energies (D_0) as reference values. The first set of reference values consists of a combination of experimentally refined data together with estimated values for specific diatomic molecules.[210,211] Compared with these reference values, the most accurate method (TPSSh/SARC2-QZV/ZORA) had a MAD of 8.76 kcal/mol, while the majority of the methods considered had MAD values between 10.0 and 20.0 kcal/mol. MAD values of this magnitude are uncommon for DFT calculations. The second set of reference values was obtained from the work of Solomonik and Smirnov.[225] In that article, the authors reported highly accurate dissociation energies (D_e) computed with the coupled-cluster single, double, and triple excitations with a perturbative quadruples [CCSDT(Q)] method at the complete basis set limit for 17 lanthanide-containing diatomic molecules. A comparison between the experimentally estimated values, the CCSDT(Q) values, and our DFT results for the seven molecules that were included in each of these three independent studies (LaF, LaO, EuF, EuO, YbF, LuF, and LuO) and for three of these cases (LaO, EuO, and EuF), the difference between the two sets of reference data was less than 2 kcal/ mol. That was not the case for the remaining four diatomic molecules (LaF, YbF, LuO, and LuF), where deviations between the two sets of reference data were exceeding 15 kcal/mol. For example, for lutetium fluoride (LuF), the "estimated experimental" dissociation energy is 135 ± 10 kcal/mol, while the CCSDT(Q) energy is 170.7 kcal/mol. DFT results for the LuF molecule are in very good agreement with the highly accurate CCSDT(Q)

results, with dissociation energies (D_0) varying between 165 and 175 kcal/mol, depending upon the choice of the combination of functional/basis set/relativistic approximations. Because of this agreement and because the large MAD errors between the DFT and "estimated experimental" dissociation energies due to the inconsistent reference data, we concluded that the CCSDT(Q) results are more trustworthy. For that reason, we compare in the next paragraph De values between DFT and CCSDT(Q).

The MAD values from the 198 combinations of functional/basis set/relativistic correction considered in this study are shown in Figure 3.1. The most accurate methods (i.e., with the lowest MAD values) are shown as dark blue in the heat map. It becomes evident that the SVWN-5 (LDA) density functional is outperformed by all other functionals, as was expected. All MAD values obtained with the def2-TZVP/ECP combination were larger than 7.0 kcal/mol. Similarly, some functionals consistently exhibit large deviations from the reference values, independent of the basis set/relativistic correction combination. For example, PW91, M06, M06-2X, and ωB97-X showed deviations that exceeded 8 kcal/mol. The two most accurate combinations have a MAD value of less than 2 kcal/mol. Those are B3LYP/Sapporo/DKH (1.3 kcal/mol) and PBE0/SARC2/DKH (1.8 kcal/mol). A second group of functionals that showed reasonable agreement with the CCSDT(Q) dissociation energies (MAD values below 3 kcal/mol) are CAM-B3LYP/SARC2-QZV/DKH (2.1 kcal/mol), PBE0/Sapporo-TZP/DKH (2.2 kcal/mol), PBE0/SARC2-QZV/ ZORA (2.3 kcal/mol), B3LYP/SARC2-QZV/DKH (2.5 kcal/mol), PBE0/SARC-TZVP/ZORA (2.7 kcal/mol), and CAM-B3LYP/ANO-RCC-VTZP/DKH (2.8 kcal/mol).

A more careful examination of the results presented in Figure 3.1 allows us to draw some general conclusions about the performance of the methods under consideration. First and foremost, the poor performance of the def2-TZVP/ECP basis set may indicate the need for the explicit inclusion of core electrons in calculations involving lanthanide species. For GGA functionals, ZORA provided higher accuracy than DKH, while for hybrid functionals, DKH and ZORA showed similar accuracies. However, more testing should be performed to better understand the differences between these methods, especially when they are applied on molecular lanthanide complexes. To further understand the effect that different density functionals, basis sets, and relativistic approximations have on lanthanide complexes, the top performing methodologies were

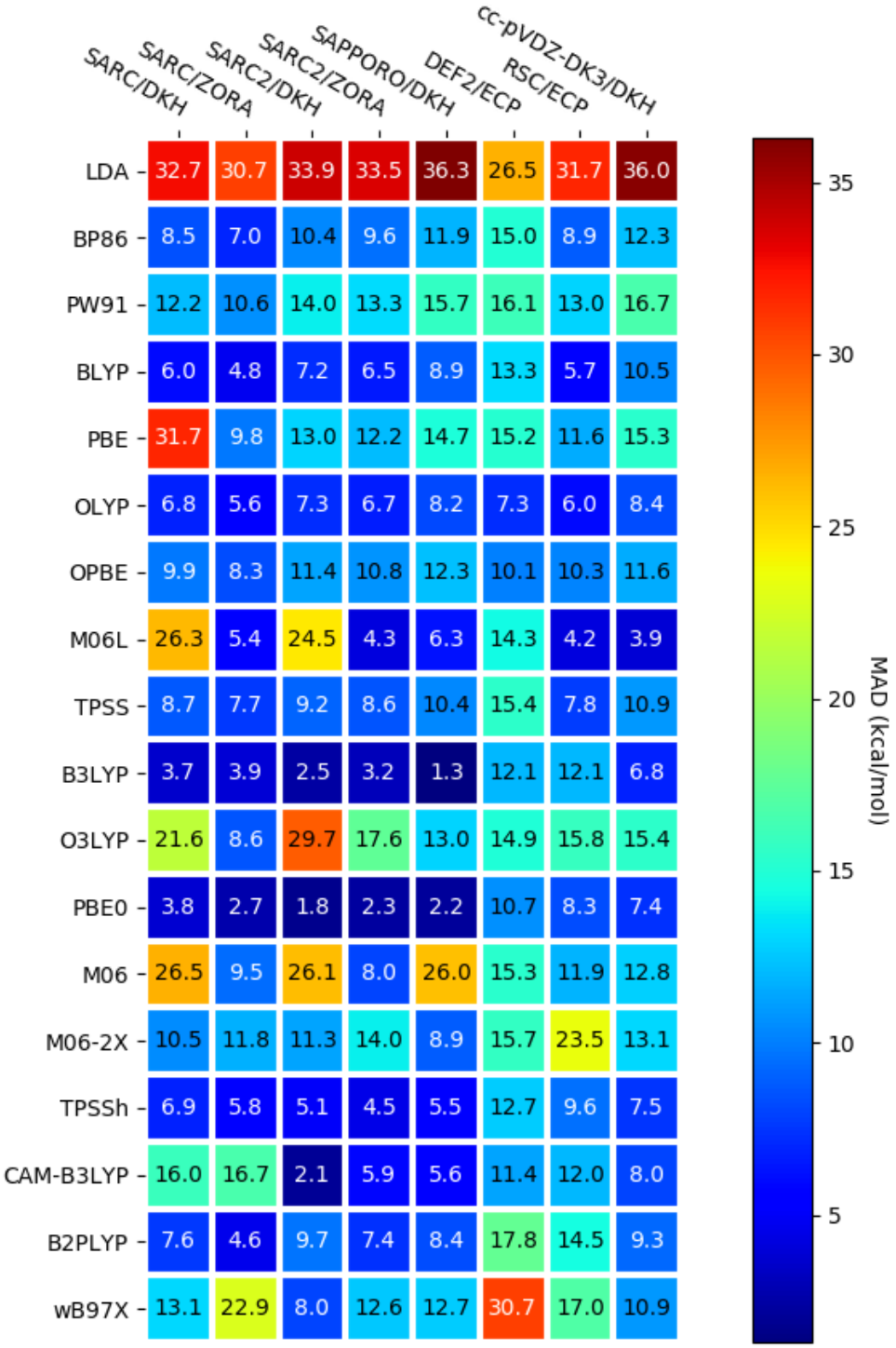


Figure 3.1 Heatmap showing the MAD from computationally calculated dissociation energies (D_e) for seven lanthanide diatomics.

tested on larger lanthanide-containing complexes. The Sapporo-TZP/DKH methodology showed high accuracy when combined with hybrid density functionals and so the three most accurate methods in this group will be applied to larger molecular complexes: B3LYP/Sapporo-TZP/DKH, PBE0/Sapporo-TZP/DKH, and TPSSh/Sapporo-TZP/DKH. In addition, the SARC-TZVP/ZORA methodology showed the highest accuracy when combined with GGA density functionals and so three of those methods will likewise be included for the next step of this study: BLYP/SARC-TZVP/ZORA, OLYP/SARC-TZVP/ZORA, and BP86/SARC-TZVP/ZORA

Lanthanide-Containing Molecular Complexes

The most accurate methods from the previous study of the dissociation energies of lanthanide diatomic molecules have been tested on larger molecular complexes to gain a better understanding of how they perform on geometry optimizations of polynuclear, lanthanide-containing molecules. To do this, we have tested three GGA functionals (BLYP, OLYP, and BP86) with the SARC-TZVP basis set and ZORA and three hybrid functionals (TPSSh, PBE0, and B3LYP) with the Sapporo-TZP basis sets and DKH. The SARC2-QZV basis sets were not considered because they were shown to lead to only a marginal increase in accuracy in general compared to the SARC basis sets. In addition, SARC2 are of quadruple-ζ quality, which will make them cost-prohibitive for large-scale computations of polynuclear complexes.

All computations were performed on molecular complexes, which include 18 experimentally refined crystal structures. These structures were chosen for three reasons: they include a majority of the naturally occurring lanthanides (9 of the possible 14), they encompass three different families of structures, whereby one family consists of many crystal structures that share a similar ligand environment but contain different lanthanides, and they resemble complexes that could be important in ligand-based lanthanide separation processes. These three families include (1,2-dimethoxyex-thane-O,O′)-tris(1,1,1,5,5,5-hexafluoroacetylacetonato-O,O′)-lanthanide,[252] triaquatris(2-hydroxybenzoato)lanthanide trihydrate,[253] and (N,N′-bis(2-pyridylmethylene)ethane-1,2-diamine)tris(nitrato-O,O′)lanthanide,[254] (Figure 3.2).

A B C

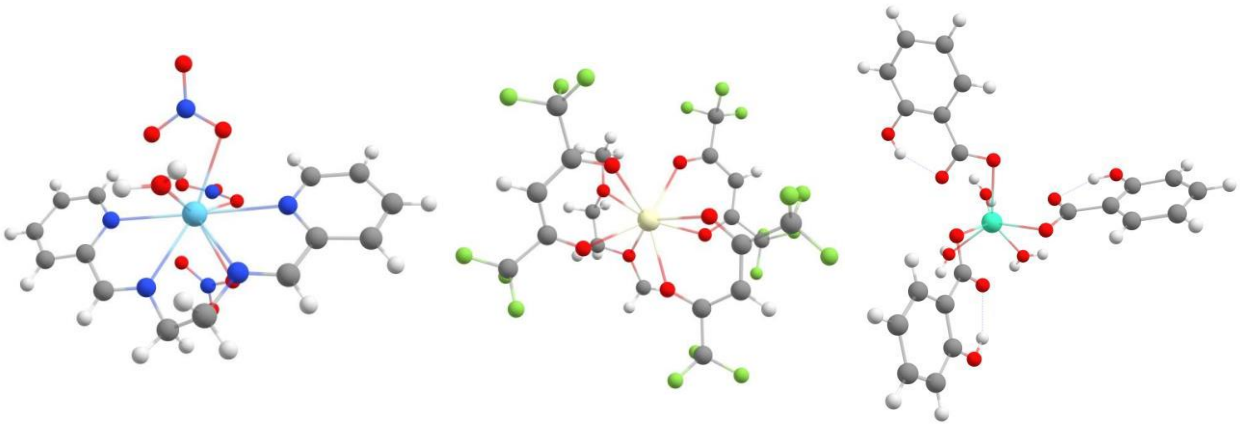


Figure 3.2 The three different families of Ln complexes considered at the second part of the benchmark study (a) (N,N'-bis(2-pyridylmethylene)ethane-1,2-diamine)-tris(nitrato-O,O')-Ln monohydrate, (b) (1,2-Dimethoxyexthane-O,O')-tris(1,1,1,5,5,5-hexafluoroacetylacetonato-O,O')-Ln, and (c) triaqua-tris(2-hydroxybenzoato)-Ln trihydrate. H, C, N, O, and F atoms are shown in white, black, blue, red, and green respectively. The light blue, yellow, and teal atoms in the center of the complexes correspond to the lanthanide atoms.

To test the performance of the chosen methods, geometry optimizations of each complex were performed, and each method's accuracy was assessed based on the MAD values of the bond lengths of the first coordination sphere from the experimentally refined structures (Figure 3.3). The TPSSh/ Sapporo-TZP/DKH method shows a MAD value of 0.069 Å, with the second most accurate method (PBE0/Sapporo-TZP/ DKH) showing an error 15% higher (0.079 Å). From the GGA functionals, the BP86/SARC-TZVP/ZORA level of theory provided similar accuracy (0.085 Å), while OLYP/SARC-TZVP/ZORA and BLYP/SARC-TZVP/ZORA had MAD values of 0.181 and 0.106 Å, respectively. Finally, B3LYP/Sapporo-TZP/DKH had the largest deviations among the methods tested in this study that utilized a hybrid functional (0.162 Å). In order to test these methods and their applicability to lanthanide separation, the most accurate method overall (TPSSh/Sapporo-TZP/DKH) and the most accurate GGA-based method (BP86/SARC-TZVP/ZORA) will be applied to the final part of this study.

Determination of Lanthanide Selectivity using HDEHP

To further validate the conclusions that have been reached in this study, a final set of calculations were performed using the TPSSh/Sapporo-TZP/DKH method. In addition, the most accurate method that utilized a GGA functional (BP86/SARC-TZVP/ZORA) was also tested to evaluate the performance of a computationally efficient density functional. These calculations involved determination of the separation energies between lanthanide pairs using a ligand that is often employed in solvent-based cation separation: bis(2-ethylhexyl)phosphate, also known as HDEHP (Figure 3.4).[171]

These specific calculations were chosen with the premise of furthering the field of ligand design with respect to solvent- based lanthanide extraction and separation. The HDEHP ligand was first suggested for use in lanthanide separation in 1957; since then, an abundance of experimental data have been made available showing that it is able to quite adept at separating most pairs of lanthanides in the series.[171] This experimental data can be used to gauge the performance of the two methods as they are used to mimic ligand exchange during the solvent-extraction process. For that purpose, calculations involving lanthanum, gadolinium, and lutetium

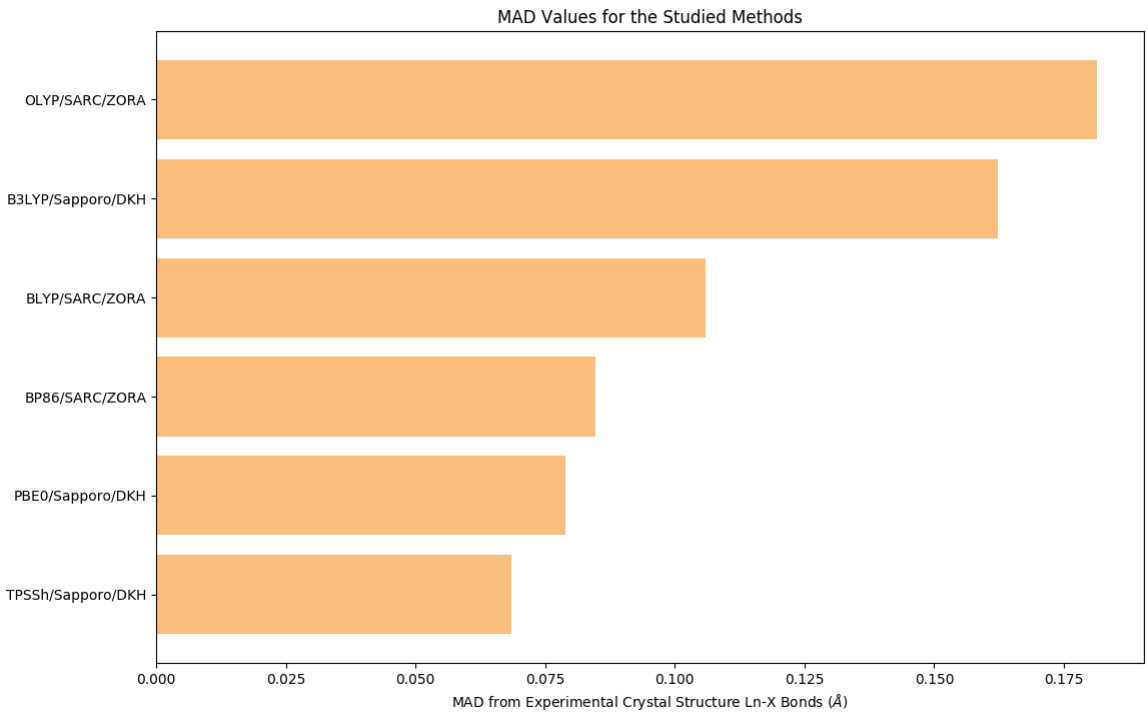


Figure 3.3 MAD of computed Ln-X bond lengths compared to experimentally refined crystal structures of the six methods examined.

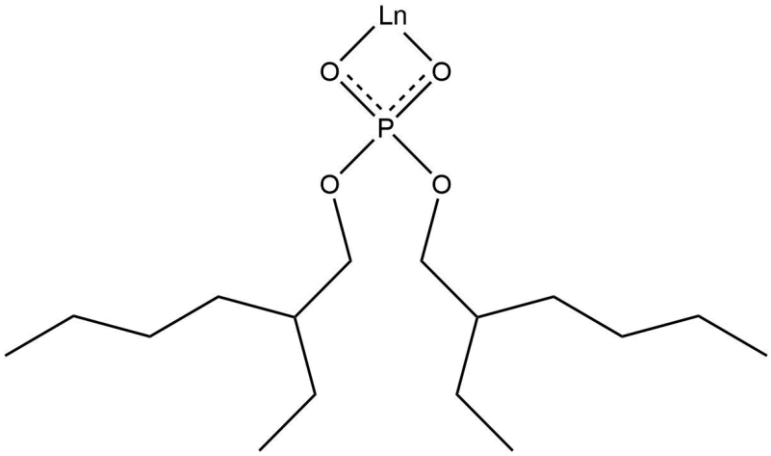


Figure 3.4 HDEHP ligand binding to a lanthanide atom.

were performed to evaluate the theoretical methods' performance on light, middle, and heavy lanthanides, respectively. The $Ln(HDEHP)_3$ complex is expected to be found in the nonpolar solvent following extraction, but such a complex is too large to be examined with the TPSSh/Sapporo/DKH method. Thus, the calculations involved two different ligand environments, $Ln(NO_3)_3(H_2O)_3$ and $Ln(NO_3)_2(H_2O)_3(HDEHP)$, as a first-step approximation to the $Ln(HDEHP)_3$ complexes. Reaction free energies ΔG_{eq}(Ln1,Ln2) between two lanthanides Ln1 and Ln2 were computed by the addition of the corresponding thermal corrections and harmonic zero-point energies. These values were then used to determine the energy difference between the left- and right-handed side of Eq. 3.1.

$$Ln1(NO_3)_3(H_2O)_3 + Ln2(NO_3)_2(H_2O)_3(HDEHP) \leftrightharpoons Ln1(NO_3)_2(H_2O)_3(HDEHP) + Ln2(NO_3)_3(H_2O)_3 \quad (3.1)$$

A positive ΔG^{eq}(Ln1,Ln2) value indicates that the reaction equilibrium will favor the reactants (left-hand side of Eq. 3.1), and Ln2 will be in the organic phase, coordinated to HDEHP. Because the separation of lanthanide ions in solution is often performed experimentally in nitric acid, we have chosen nitrates to fill the first coordination sphere of the lanthanides. Recent work has showcased the importance of accurately modeling lanthanide (and actinide) systems with bound nitrates.[255,256] Thus, the ligand extractant is competing with the nitrate ions, which often results in nitrates being bound to the final lanthanide−ligand complex. The computed reaction free energy differences ΔG^{eq}(Ln1,Ln2) for La/Gd, Gd/Lu, and La/Lu are 2.14, 1.63, and 3.77 kcal/mol, respectively, for the TPSSh/Sapporo-TZP/DKH method and 1.60, 6.11, and 7.71 kcal/mol, respectively, for the BP86/ SARC-TZVP/ZORA method. Both methods agree that (1) for every pair, the heavier lanthanide (Ln2) will remain in the organic phase because all computed ΔG^{eq}(Ln1,Ln2) have positive values and (2) the separation between lanthanum and gadolinium is easier than that between gadolinium and lutetium, in agreement with the experimental findings.[171]

In order to compare the calculated ΔG_{eq} (Ln1, Ln2) with the experimental data, we have computed separation factors $SF_{Ln1/Ln2}$ between a lanthanide pair Ln1/Ln2 by using the following expression:[257]

$$\Delta G^{eq}(Ln1,Ln2) = -RT \ln(SF_{Ln1/Ln2}) \quad (3.2)$$

From the computed free energies obtained at the TPSSh/ Sapporo-TZP/DKH level, the computed separation factors for the La/Gd, Gd/Lu, and La/Lu pairs are 36.75, 15.67, and 575.74, respectively. These values are in very good agreement with the experimental values of 44.6, 9.55, and 425, respectively.[171] Note that the positive and negative values for ΔG^{eq} (forward or reverse reaction for Eq. 3.2) both lead to the same separation factors. Thus, we have taken the inverse of the explicitly calculated separation factors to better match the experimental values. For the BP86/SARC-TZVP/ZORA method, however, the separation factors do not follow the experimental trends, nor do they have the proper magnitude (the Gd/Lu and La/Lu separation factors are larger than 30000). As a possible correction to this discrepancy, implicit solvation was examined using the conductor-like polarizable continuum model (CPCM) available in ORCA with water as the solvent.[258] The calculated separation factors follow the same order for La/Gd, Gd/Lu, and La/Lu pairs, in agreement to the experimental trends. When the solvent was changed to hexane, however, opposing trends were observed, whereby the separation factor for La/Gd was smaller than that for Gd/Lu. This discontinuity between the different solvents may be explained by understanding which ligands would be bound to the lanthanide ion in the respective solutions. In aqueous solution, the different lanthanide ions would be bound by nitrate ions and water molecules, whereas organic solutions would lead to the formation of $Ln(HDEHP)_3$ complexes. Because the $Ln\text{-}(NO_3)_2(H_2O)_3(HDEHP)$ complexes are between what is expected in either solvent, the inclusion of solvent effects may lead to erroneous results. In order to best approximate the experimental conditions, one would need to perform calculations with the $Ln(HDEHP)_3$ complexes, which, while feasible, are computationally demanding for the TPSSh/Sapporo-TZP/DKH method. While these computed free energy differences may seem small and within the error of the DFT method, calculation of the separation factors utilizes very small physicochemical differences among the series, which results in small differences in the energetics of the system when examined computationally, in agreement with previous theoretical studies.[208,213,259]

3.2.5 Conclusions

In this section, we have examined the accuracy of several commonly used density functionals for computations on lanthanide diatomic molecules and complexes. Our work is organized in three stages. In the first stage, a total of 198 combinations of different functionals, basis sets, and relativistic approximations were tested on the accuracy on the dissociation energies of lanthanide diatomic molecules (oxides and fluorides). It was found that the B3LYP density functional along with the Sapporo basis set and DKH relativistic approximation showed the lowest MAD value (1.3 kcal/mol) from highly accurate CCSDT(Q) dissociation energies. Other notably accurate methods include PBE0/SARC2-QZV/DKH, PBE0/Sapporo-TZP/DKH, and PBE0/SARC2-QZV/ZORA. In the second stage, a subgroup of the initial 198 methods was further tested on lanthanide-containing molecular complexes. The TPSSh density functional in conjunction with the Sapporo-TZP basis set and DKH method exhibited the smallest deviation in the Ln−X bond lengths of the first coordination sphere of 18 different crystal structures with a MAD value of 0.069 Å. To further validate this result, in the last stage of this study, the TPSSH/Sapporo-TZP/DKH method was applied to the study of lanthanide separation via solvent extraction using the HDEHP ligand. The computed separation factors agreed with known experimental results, showing a larger separation energy difference for the La/Gd pair than for the Gd/Lu pair. The excellent performance of the TPSSh/Sapporo-TZP/DKH method may be explained by examining individually the density functional, basis set, and relativistic method. The TPSSh density functional was developed as a nonempirical method and thus has no bias toward or against any specific group of molecules, which leads to a good performance in many different areas of transition-metal and lanthanide chemistry.[260–262] With regards to the Sapporo basis set, its excellent performance stems from the extra polarization functions (3g1h) and from the additional five f functions which makes it larger than a standard triple-ζ basis set. Further testing is needed to identify the contribution of each function type to their performance. With regard to the DKH relativistic method, there are inherent differences in how it and ZORA approximate the Dirac equation, but work done by Hong *et al.*[229] has shown that these two methods lead to very similar molecular properties for lanthanide and actinide diatomic molecules such as those studied in the first section of this work. Thus, one would expect similar results when using either relativistic

approximation, but only the DKH method is studied here because the Sapporo basis set was only constructed with DKH and not ZORA. These results indicate that a larger basis set, specifically one with high angular momentum basis functions, paired with a nonempirical hybrid density functional provides a high-quality computational model for the electronic structure and geometries of lanthanide molecules and can be used for the computation of separation factors and other properties. In addition, the BP86/SARC-TZVP/ZORA method may be used as a first step for the optimization of such lanthanide-containing molecules because it was shown to be nearly as accurate as the TPSSh/Sapporo-TZP/DKH method during the second stage of this work but at a much lower computational cost. The examination of alternative organic ligands for lanthanide separations is an ongoing topic of research in our group, where we use DFT data in combination with machine learning for screening large molecular databases. The current study is used to identify accurate and computationally efficient levels of theory for the generation of reliable computational data.

3.3 Ligand Engineering for Lanthanide Capture using Machine Learning: Ln-Tris-β-diketone Complexes (Ln = La, Lu) as a Test Case

3.3.1 Abstract

Electronic structure theory when combined with machine learning may offer an alternative procedure on the fast and reliable examination of the properties of a large number of molecules and materials. The free binding energies of lanthanide-tris-β-diketone (Lnβ_3) complexes are such properties which can be used to identify ligands (β-diketones) which show high separability between La and Lu. The computation of these values may expedite the process by examining many ligands in a shorter timeframe than may be achieved synthetically. A wide examination of different machine learning protocols was performed, and models which may be used for predicting free binding energies of Lnβ_3 complexes were identified.

3.3.2 Introduction

The increasing demand for green energy technologies and other modern electronics places an increased demand for rare earth elements such as the lanthanides.[196,197] These elements are usually found in low concentrations in ore deposits and because of this, an efficient separation technique is required.[171,198] Solvent extraction via a chelating ligand is the standard method to separate lanthanide elements in large quantities but this process is made more difficult due to the similar physicochemical properties of adjacent lanthanides. The similarities in the electronic structure along the series results from the nature of the 4f orbitals which are deeply buried by the fully filled 5s and 5p orbitals. The consequence of this is that lanthanide complexes favor molecular geometries based on steric effects and not on electronic effects as is observed in many transition metal complexes. Therefore, these small, physicochemical changes across the lanthanides series complicate their separation using chelating agents.[263] The small difference in the ligation energy between different lanthanides must be compounded through the use of multiple separation steps.[171]

Current research has focused on several different families of chelating agents such as diglycolamides,[172–180] phosphine oxides,[181–188] and phenanthrolines[189–194] which have all been shown to effectively separate lanthanide cations both from experimental and theoretical standpoints. The multidentiticity of these ligands allows for favorable interactions with the large, lanthanide cations found in solution and by modifying the scaffold around the chelation sites, more favorable interactions may be achieved. One excellent example of this is the work of Hasegawa *et al*. who utilized a synergistic extraction using modified β-diketone and phenanthroline carboxamide ligands.[190] Their approach resulted in high selectivity for samarium over neodymium and europium.

For the design of new ligands and the improvement of the efficiency of specific separation steps, a vast number of ligand scaffolds that may be considered, synthesized, tested, and verified leads to a very large chemical space.[264,265] Computational methods allow for the examination of many of these ligand scaffolds prior to the need for synthesis[184] This in turn allows for the elucidation of the underlying physical properties that may lead to a better understanding of what

affects the separation process. It is often useful to identify patterns within the dataset of ligand scaffolds which may lead to the quantification of structure-function relations. This would allow for a sense of interpretability and thus use a more chemical mindset in designing new ligands. Machine learning (ML) is one such way to identify said patterns in large datasets such as those which would be used for lanthanide separation studies; this is commonly referred to as high-throughput screening. The high throughput screening of different ligand scaffolds has been examined previously in a few cases with respect to the separation of lanthanide cations. The work of Solov'ev *et al*. was the first of its kind which explicitly combined elements of electronic structure theory and ML for the prediction of something akin to the separation of lanthanide cations.[266] Their work focused on the prediction of HOMO and LUMO energies of different N-donor heterocycles using PM7 and density functional theory (DFT) calculations. A dataset of 388 ligands was constructed and cross-validated to then make predictions on external datasets which consisted of forty-five similar ligands and 1,546 vastly different ligands. The ML models were able to reproduce the HOMO and LUMO energies of the training data set to within 0.01 eV and within 0.26 eV for the two testing data sets. More recent work by Chaube *et al*. applied ML for the prediction of lanthanide-ligand binding energies from a large, experimental dataset.[267] Many different ML algorithms were utilized and following cross-validation, the authors were able to make predictions on 71 million lanthanide containing molecules.

This work aims to be the first of its kind wherein electronic structure theory (EST) and ML are used together to better understand the structure-function relationships inherent in ligand-lanthanide systems. Such a high-throughput approach would allow for DFT level results at a reduced computational cost. A Recent review has shown the efficacy of such an approach for numerous applications such as learning force fields, retrosynthetic technologies, catalysis, and drug design.[268] In addition, recent work in our own group has shown how ML can be used for the acceleration of EST methods such as coupled cluster[269] or for the prediction of CO_2 binding on small organic molecules.[270] As such, we have chosen β-diketone ligands as the ligand scaffold for lanthanide separation due to previous experimental work using them[190,271,272] and because of their simple structure which allows for easy functionalization (see Figure 3.5). This work serves as an initial study to better understand how different machine learning architectures can be used for the

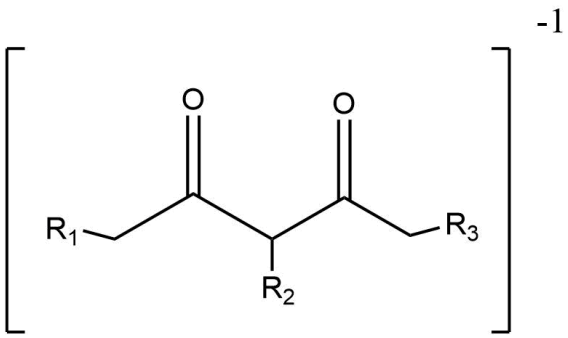


Figure 3.5 Chemical structure of a pre-functionalized, β-diketonate ligand.

prediction of properties for lanthanide-ligand complexes. The section is outlined in the following manner: we first begin with a description of the generation of our β-diketone ligand database followed by the computational protocol we used to calculate the binding free energies of the lanthanide-tris-β-diketone systems. Next, we describe the protocol we followed with regards to the ML application. Our results are then discussed in three sections: initial cross-validation of many thousands of ML models, active learning using three models which showed promise, and then a discussion regarding the predictions made using this model.

3.3.3 Computational Details

Database Generation

The molSimplify program[273] was utilized for the generation of the molecular structure database. We began first by designing a β-diketonate anionic ligand (Figure 3.5) in which the acidic hydrogen of a 1,3-diketone (herein referred to as a β-diketone or simply as β) has been removed. This ligand serves as a chemical base where functionalizations may be performed on any hydrogen. Using the molSimplify program, we functionalized the β ligands at different positions using different functional groups. Our database consisted of 1,224 unique β ligands where we have performed single- and double-functionalizations using twenty-four functional groups (see Table 3.2). The single functionalizations were performed at two sites of the β ligands: either on the methyl group which is attached at the 1-position (R_1 in Figure 3.5) or the 2-position (where the acidic hydrogen has already been removed, R_2). This resulted in forty-eight β ligands being generated. The double functionalizations were performed at three combinations of sites: two functionalizations on the R_1 site of the β ligand which generated 300 structures due to symmetry, a functionalization at R_1 and R_3 (300 generated structures), and a functionalization at R_1 and R_2 which generated 576 structures. The database consists of a wide array of β ligands with varying degrees of steric and electronic characteristics which affect the binding free energies for the Lnβ_3 complexes. The molSimplify program was used to modify both the β ligands themselves and the Lnβ_3 complexes. The generated structures were initially optimized using empirical forcefields and

Table 3.2 List of functional groups used for the generation of the *β*-diketonate database.

Halide	*-F* *-Cl* *-Br*
Alkane	*-CH_3* *-CH_2CH_3* *-$CH_2CH_2CH_3$* *-$CH_2CH_2CH_2CH_3$* *-$CH_2CH_2CH_2CH_2CH_3$*
Amine	*-NH_2* *-$NHCH_3$* *-$NHCH_2CH_3$* *-$NHCH_2CH_2CH_3$* *-$NHCH_2CH_2CH_2CH_3$* *-$N(CH_3)_2$* *-$N(CH_2CH_3)_2$* *-$N(CH_2CH_2CH_3)_2$* *-$N(CH_2CH_2CH_2CH_3)_2$*
Alkoxy	*-OH* *-OCH_3* *-OCH_2CH_3* *-$OCH_2CH_2CH_3$* *-$OCH_2CH_2CH_2CH_3$*
Trifluoromethyl	*-CF_3*
Phenyl	*-C_6H_5*

then reoptimized using EST methods (*vide infra*). No symmetry considerations for *mer* and *fac* isomers of the Lnβ_3 complexes were taken.

Conformational Search of the β-diketone Ligands

The β ligands were optimized in a two-step process first using metadynamics (metaD) to sample the conformational space of the ligand and second using density functional theory (DFT) to further refine the metaD optimized structure. This methodology was chosen as the β ligands are relatively small and may show a great degree of conformational freedom due to the functionalizations performed previously. To calculate the free binding energies of a Lnβ_3 complex, an accurate value for the electronic energy of the free β ligands must be calculated which is itself a function of the chosen conformation. The free β ligands that were generated using molSimplify were used as input structures for the CREST program[274] which samples a large volume of the potential energy surface of a given molecule or system using density functional tight binding (DFTB) and performs systematic optimizations of said conformational space to arrive at structures lower in energy than may be found using traditional geometry optimization algorithms. CREST generates a sorted list of the identified conformers which may be used as further input for other calculations. From the list of generated conformers, only the ten lowest energy conformers for a given β ligand are chosen to be further refined using DFT. The lowest energy structure of these ten DFT optimized structures was used for all future calculations.

DFT Optimizations

DFT optimizations were performed with the BP86 density functional, the SARC-TZVP basis set, and the zeroth-order relativistic approximation (ZORA) for the calculation of the free binding energies for Lnβ_3 complexes using. The SARC-ZORA-TZVP[218] basis set was applied to the La and Lu atoms, the ZORA-def2-TZVP basis set for the first coordination sphere of each lanthanide, and the ZORA-def2-SV(P) basis set for all other atoms.[275] The resolution of identity (RI) approximation[248] was utilized to accelerate the calculation of the four-index integrals with the

SARC/J auxiliary basis set being used for all atoms.[218] The ZORA relativistic approximation is used for all calculations.[50,51,226] The BP86/SARC-TZVP/ZORA level of theory was used based on our previous benchmarking study that included a variety of lanthanide molecules such as diatomic molecules and larger, molecular complexes.[195] Analytical frequency calculations were performed to ensure all structures are minima within their respective potential energy surfaces. Tight optimization and SCF convergence criteria are utilized and dispersion effects were accounted for by using Grimme's D3 semiempirical method[276] with the Beck-Johnson damping function.[277] The geometries obtained using molSimplify for the β ligands and the LnX_3 complexes (where X is either NO_3^- or the β ligands) were used as input structures without modification. The justification to explicitly include nitrated complexes comes from recent work by Peterson *et al.* in which they examined the binding of $[Ln(NO_3)]^{2+}$ complexes (with Ln = La to Lu).[255] As mentioned above, the β ligands themselves were first optimized using the forcefields in molSimplify and then their conformational space was examined using the CREST program. From the list of generated conformers, the ten lowest energy structures of each β ligand were then used as input structures for the subsequent DFT geometry optimization. Following this, the BP86/SARC-TZVP/ZORA level of theory is used to optimize the β ligands and the $Ln\beta_3$ complexes. All DFT calculations were performed with the ORCA 5.0 software package.[111]

To calculate the free binding energies (ΔG^{298}), a composite scheme is used which includes the total electronic energy of the system and corrections from thermal, thermodynamic, and solvation effects. Thermal and thermodynamic corrections are accounted for using an analytical frequency calculation which is performed at the optimized geometry. Likewise, solvation effects are accounted for using the conductor-like polarizable continuum model (CPCM)[258] of ORCA using water as a solvent. The difference in energy between the electronic energy and the Gibbs free energy accounts for all thermal and thermodynamic effects while the difference in energy between the solvated and the non-solvated system accounts for all solvation effects. This composite scheme is used for all structures including the β ligands themselves, the $Ln\beta_3$ complexes, the nitrate ions, and the $Ln(NO_3)_3$ complexes.

The equation to calculate the free binding energy given a specific lanthanide (Ln = La, Lu) and a functionalized β-diketone (β) is shown below (Eq. 3.3). The energy of the fully nitrated

lanthanide complex and three free β-diketones is subtracted from the energy of the Lnβ_3 complex and three free nitrate ions to arrive at the total free binding energy of the Lnβ_3 complex (ΔG_{Ln}). This numerical value is then used as the output value to be predicted by the ML models. To achieve the highest separation of La and Lu, the largest possible difference between their respective free binding energies is desired. This difference of a difference ($\Delta\Delta G$) is the metric which is to be maximized using ML by either learning said $\Delta\Delta G$ value or learning both ΔG values and then calculating $\Delta\Delta G$ from those values. The prediction of both output values are examined and discussed below.

$$Ln(NO_3)_3 + 3\beta^- \rightarrow Ln\beta_3 + 3NO_3^- \qquad (3.3)$$

Machine Learning Algorithms

Nine different ML algorithms (also referred to as *learners*) are examined for the prediction of free binding energies of Lnβ_3 systems. K-nearest neighbors (KNN), decision trees (DT), random forest (RF), kernel ridge regression (KRR),[278] Gaussian process regression (GPR), Bayesian ridge regression (BRR), support vector machines (SVM), AdaBoost (ADA),[279,280] and XGBoost (XGB).[281] These algorithms encompass simple methods (KNN and DT), methods which map data to higher dimensional spaces (KRR, GPR, GPR, and SVM), and ensemble methods (RF, XGB, ADA). The Scikit-learn[282] program package is used for all of the ML algorithms. Each learner was trained using ten-fold cross validation to determine the optimal hyperparameters which resulted in the highest accuracy (defined as the highest r^2 value). Examination of how these seven learners interpret, learn, and predict on the data sets will allow for a determination of which one is optimal for the prediction of free binding energies.

Molecular Representations

The methodology used for encoding chemical structural information into a form which a learner can utilize is crucial for the development of reliable, accurate, and transferable models. This representation encodes the geometric data in a format that is invariant to the translation and rotation of a molecule/system and to the permutation of the atomic indices. As a real system does not change energy upon translation, rotation, or the permutation of atomic indices, so too must the molecular representation. Popular molecular representations with broad applicability in data-driven chemical studies are the Coulomb matrices (CM)[283] and the Bag-of-Bonds (BoB)[284] methods while more complex representations include Persistence Images (PI),[270] Smooth Overlap of Atomic Positions (SOAP),[285] and the Many-Body Tensor Representation (MBTR).[286] The QML software package is used to generate the CM and BoB representations, DScribe[287] is used to generate the SOAP and MBTR representations, and the Ripser[288] library is used for the creation of the persistence diagrams and from them the PIs.[289] The molecular representation input vectors for CM, BoB, MBTR, and SOAP were padded with zeros so as to retain a consistent input vector size.

The CM molecular representation was developed by Rupp *et al*. and is a square atom-by-atom matrix where each diagonal element is an approximate potential energy of the free atom and the off-diagonal terms are pair Coulombic potentials. As an improvement to CM, the Bag-of-Bonds method by Hansen *et al*. placed each atomic pair in a specific vector (bag) based off the elemental pairs and then sorted this vector by value. By calculating the local density of all atoms in a given system, the SOAP molecular representation includes more chemical information but at a greater computational cost as compared to CM or BoB. MBTR are similar to the SOAP method in that more atomic interactions are included but in this case, MBTR includes many n-body terms to describe the interactions of all elemental combinations. These n-body terms are then broadened, summed, and weighted so that the influence of atom pairs far from each other is minimized. Finally, the PI molecular representation is an application of persistent homology to chemical systems in which the connectedness (bonding information), proximity, and empty space among atoms is stored in a vectorized form.

3.3.4 Results and Discussion

Data Analysis

Of the 1,224 β ligands which encompass the entire dataset, 100 of those were initially tested using the MetaD/DFT protocol to calculate the free binding energies. Random selection was used to choose the first 100 instances. Statistical information regarding the initial dataset is shown in Table 3.3 and histograms for the initial dataset are shown in Figure 3.6. The free binding energies for Lu are larger in magnitude than those of La in every case which would indicate that the chelating agent would preferentially bind to the Lu cations. The range and standard deviation are similar for both La and Lu whereas they are much smaller for the ΔΔG values. This requires that the error for the predicted ΔΔG values must be much smaller than the error for the predicted ΔG_{Ln} values. This results from an error of 5 kcal/mol for the ΔG_{Ln} values which is within 0.5 standard deviations from the average whereas the same error for the ΔΔG values is more than two standard deviation values. Thus, the prediction of the ΔΔG values are to be held to a higher, statistical standard.

We performed ten-fold, cross-validation using a large variety of different input (molecular representation), learner, and output to determine the optimal model to predict either the ΔG_{Ln} or ΔΔG values. As input, we used either the geometry of the β ligands as generated from molSimplify, the DFT optimized geometry for the β ligands, or the geometry of the Lnβ_3 complexes as generated from molSimplify. We utilized the five molecular representations mentioned previously. For the PI molecular representations, we used resolutions of 20x20, 50x50, and 100x100 to examine how the resolution affects the performance, if at all. Each of the nine learners were examined and predictions were made on the ΔG_{La}, ΔG_{Lu}, and ΔΔG values. The geometries of the β ligands were used to predict all three output values whereas the geometries of the Lnβ_3 complexes were only used to predict the respective ΔG_{Ln} values. The ΔG_{Ln} values were predicted individually with different models (and therefore different hyperparameters) as well as predicted together but only for a subset of learners (KNN, DT, RF, KRR, SVM, GPR). For each combination of input and output, the model which resulted in the highest accuracy (defined as the largest coefficient of determination, COD, as calculated with the Scikit-learn) was further determined. These ten

Table 3.3 Statical information for the initial, 100 instance dataset. All values shown in kcal/mol.

	ΔG_{La}	ΔG_{Lu}	$\Delta\Delta G$
Minimum	-102.68	-121.11	-6.16
Maximum	-51.35	-67.18	-18.43
Range	51.34	53.93	12.27
Average	-78.74	-92.64	-13.90
Standard Deviation	9.26	9.45	2.18

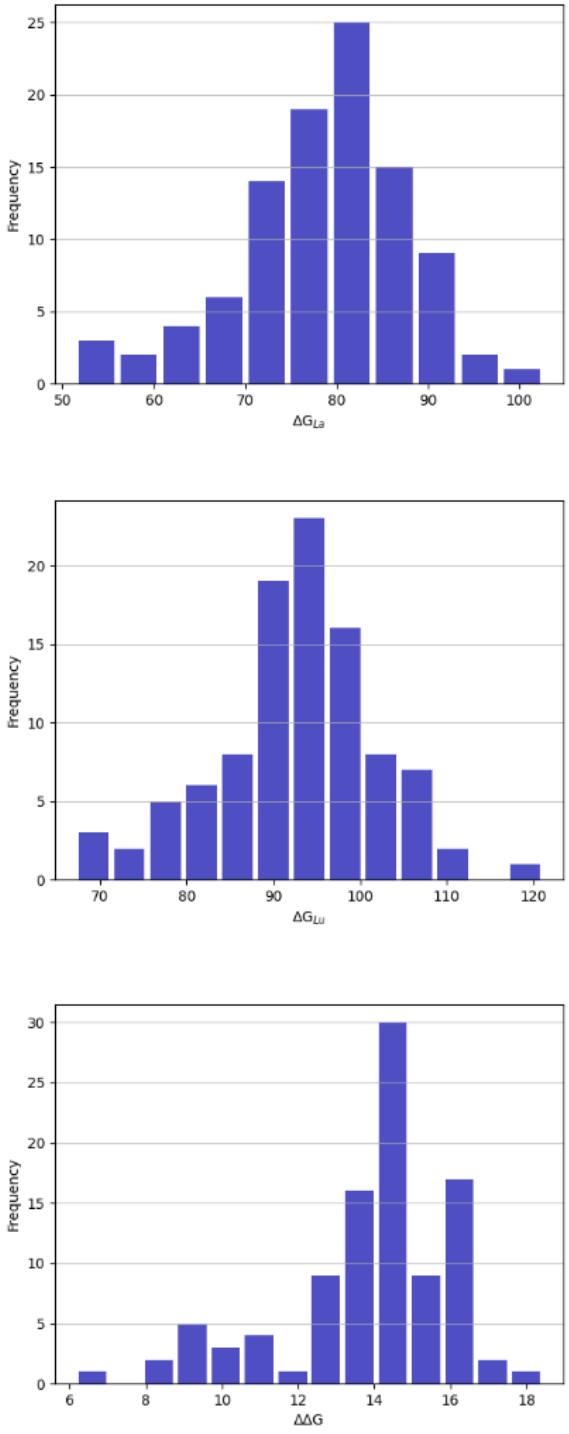


Figure 3.6 Histograms for the initial, 100 instance dataset. Free binding energies for La (left), Lu (middle), and the ΔΔG values (right). All values shown in kcal/mol.

specific models are shown in Table 3.4. Immediately apparent is the high performance achieved using the PI molecular representation. Likewise, only three of the nine learners are included in Table 3.4 which demonstrates their high usability for the prediction of ΔG_{Ln} and $\Delta\Delta G$ values for Lnβ_3 complexes. Each of these ten models was then further applied to examine how well each of them learned the respective output values in terms of learning curves.

Model Selection

Figure 3.7 includes the learning curves for each of the ten models outlined in Table 3.4. The learning curves were constructed by randomly selecting 10% of the data to be used as the testing dataset while different incremental amounts of the remaining data are used for training. The first three models predict ΔG_{La} values and show similar behavior. As the size of the training set increases, the testing error decreases. This is seen as an increase in the COD value (blue curves, left axis) and a decrease in the root mean squared error (RMSE, red curves, right axis). None of the three models have converged with respect to the training set size as both the training and testing errors should converge upon a similar value when a large and well-balanced dataset is used. Thus, we believe that a larger dataset is required for models **1**-**3** to reach convergence. Similar behaviors and trends are observed for models **4**-**6** which predict ΔG_{Lu}. The training curves for the prediction of ΔG_{La} and ΔG_{Lu} values show that extrapolated errors of ~3 kcal/mol (3%) could be achieved.

Models **7** and **8** aim to predict both ΔG_{La} and ΔG_{Lu} values simultaneously. Both models show a maximum COD (minimum RMSE) at about 65 and 75% of the total data, respectively. Trends similar to the previously discussed models are seen. The COD in these cases exceeds 0.8 with a RMSE of less than 3 kcal/mol. This behavior indicates that when both ΔG values are predicted simultaneously using the same model, a better prediction can be made. This may be due to the two values sharing commonalities within the model and thus these portions of the model may be reinforced by having two output values. Because models **7** and **8** show similar learning capabilities, we have chosen to select model **7** as the better model simply because it requires more computational resources to optimize the β ligands using DFT for model **8** as opposed to simply generating them from molSimplify for model **7**.

Table 3.4 Best performing models based upon cross-validated combinations of input and output.

Model Number	Input	Molecular Representation	Learner	Output	r^2
1	β – molSimplify	PI (100x100)	ADA	ΔG_{La}	0.540
2	β – DFT	PI (20x20)	RF		0.570
3	Laβ_3 – molSimplify	PI (20x20)	ADA		0.548
4	β – molSimplify	PI (20x20)	ADA	ΔG_{Lu}	0.556
5	β – DFT	PI (50x50)	RF		0.537
6	Luβ_3 – molSimplify	PI (100x100)	ADA		0.582
7	β – molSimplify	PI (20x20)	RF	ΔG_{Ln}	0.531
8	β – DFT	PI (20x20)	RF		0.565
9	β – molSimplify	PI (20x20)	XGB	$\Delta\Delta G$	0.407
10	β – DFT	PI (50x50)	XGB		0.466

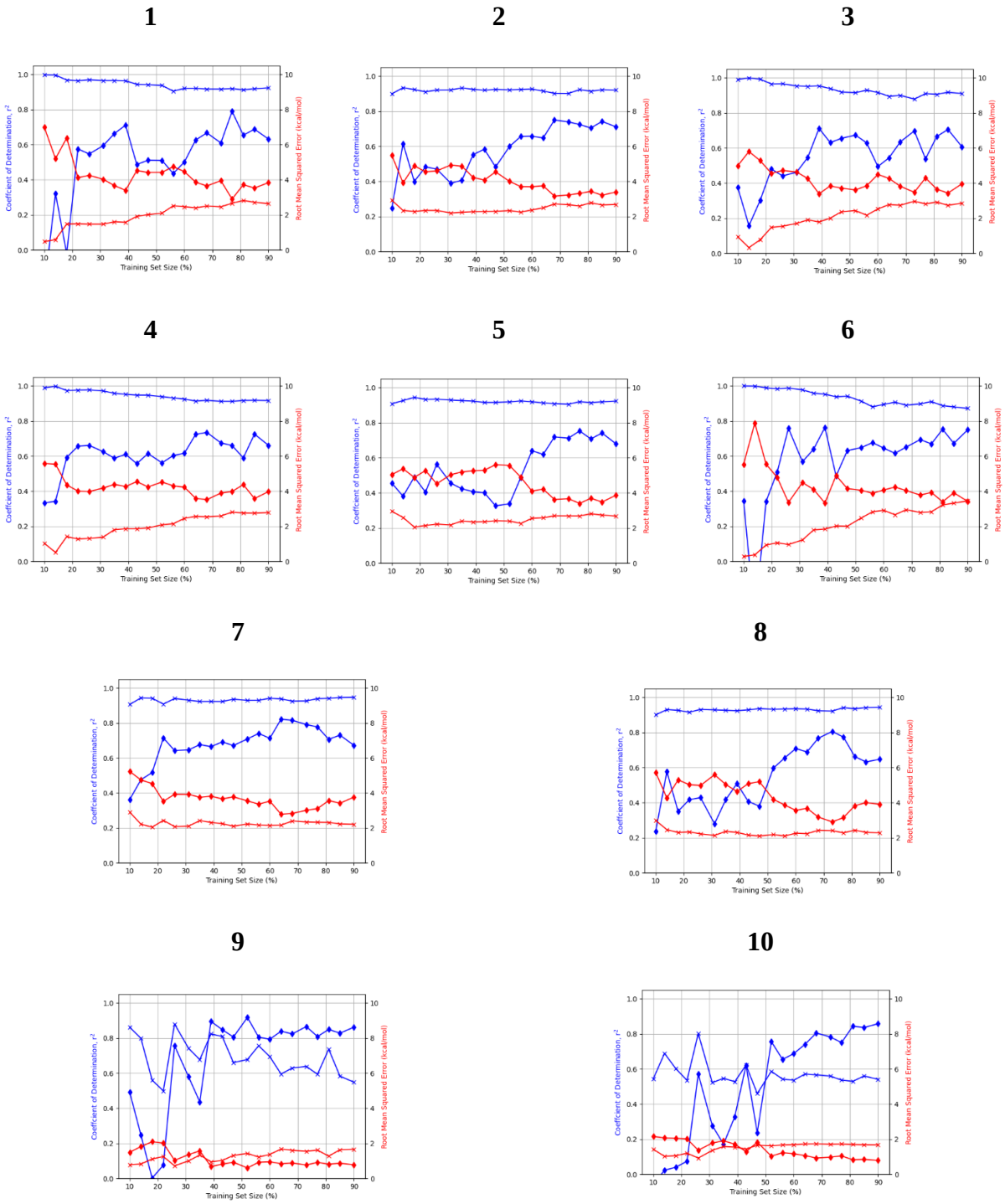


Figure 3.7 Learning curves for the ten top-performing models. Each model number corresponds to the same model as outlined in Table 3.4. Blue lines show coefficient of determination values (r^2) while red lines show root mean squared errors. Training data values are shown with crosses (✕) while testing data are shown in diamonds (◆).

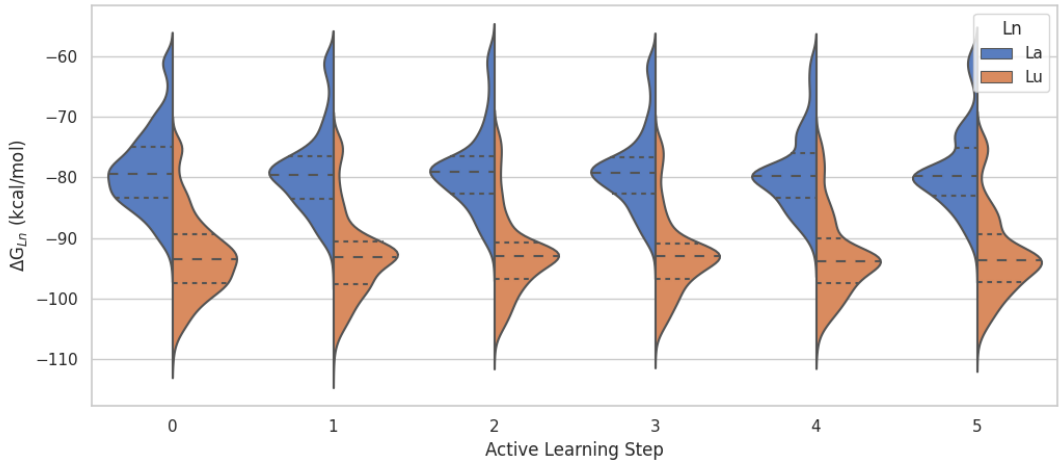
Finally, models **9** and **10** show an oscillatory behavior that indicates low predictability. Thus, a function that maps the geometry of the β ligands directly to the $\Delta\Delta G$ values does not appear to be possible within the dataset of this study. The COD values for both models fluctuate and in some cases, the training COD is lower than the testing COD which indicates that the model is unable to learn given the training data set. The RMSE curves look promising, however, but the standard deviation for the $\Delta\Delta G$ dataset is 2.18 kcal/mol which is very similar to the final RMSE value of each curve (~1.7 kcal/mol). Given this poor performance, no prediction will be made directly for the $\Delta\Delta G$ values and instead, the individual free energies must be predicted (either together or separately) and then the $\Delta\Delta G$ value must be calculated as the difference between them.

Model Evaluation

Given this analysis, we sought to perform active learning steps to increase the size of our dataset and evaluate how well each model can interpolate within the larger dataset of the 1,124 remaining β ligands. To do this, we chose two sets of models to make predictions on the remaining dataset: model **A** (same as model **7)** predicts both ΔG_{La} and ΔG_{Lu} values simultaneously and model **B** which includes models **3** and **6** to predict ΔG_{La} and ΔG_{Lu} values separately.

The active learning process is composed of five steps: (i) perform ten-fold, cross-validation to identify the hyper-parameters which result in the lowest error of the given model, (ii) train the model given these hyper-parameters, (iii) make predictions of the remaining β ligands, (iv) perform DFT calculations to explicitly calculate the ΔG_{La} and ΔG_{Lu}, and (v) evaluate the predicted free energies with respect to the actual DFT calculated free energies. This process is repeated five times. For each step, the twenty-five β ligands with the largest predicted $\Delta\Delta G$ values in magnitude were selected.

The top of Figure 3.8 shows the violin plots for the predicted ΔG_{La} and ΔG_{Lu} values for model **A** as a function of the active learning step. Note that these values are the predicted values of the remaining β ligands which have not been examined using DFT to explicitly calculate their respective ΔG_{Ln} and $\Delta\Delta G$ values. During the first step, you can see a Gaussian-like distribution of the predicted ΔG_{Ln} values that begins to cluster around the average values as the active learning



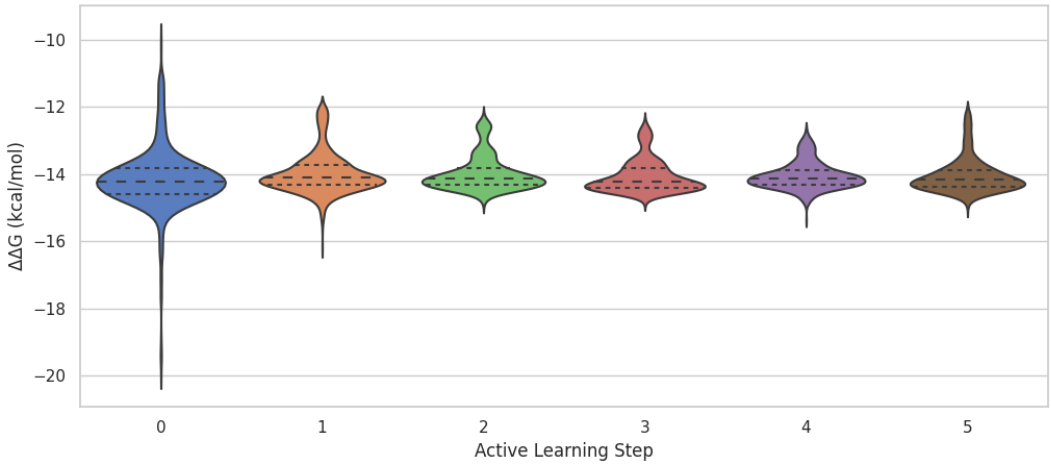


Figure 3.8 Violin plots for Model **A**. ΔG_{Ln} values (top) and $\Delta\Delta G$ values (bottom)

steps progress. The result of this is that the model simply predicts near the average ΔG_{Ln} values which results in a lack of variation in the predicted $\Delta\Delta G$ values (bottom of Figure 3.8). This also results in errors that are oscillatory in nature as a function of the active learning step; errors decrease following certain active learning steps and then increase in other steps. This behavior may be a result of the simple learner utilized for this model (random forest) which may find that predicting near the average ΔG_{Ln} values results in the lowest overall error; known as regressing to the mean. Given a large enough dataset and a relatively simple regression model, the values that are predicted usually fall within the realm of the average value as more extreme values (near the maximum or minimum values in the dataset) are more rare occurrences and thus not well-represented in the dataset. This leads to a model, such as model **A**, being more likely to predict values near the mean when the dataset set is expanded. Thus, while a model may show promise initially (as in the learning curves of Figure 3.7), subsequent analysis and use of the model identifies flaws with its use and so model **A** is a poor choice for the prediction of free binding energies for Lnβ_3 complexes.

Model **B** utilizes a more sophisticated learner (AdaBoost) and thus should be more well-behaved as a function of active learning steps. As the active learning steps progress, the highest peak in the violin plots (top of Figure 3.9) does not correspond to the average ΔG_{Ln} value for either La or Lu which is promising as it indicates that the model is not simply predicting the average value as seen in model **A**. The lack of a Gaussian shape and the bimodal distribution for the ΔG_{Ln} values may indicate a poor representations of the original dataset, however. The violin plots for the $\Delta\Delta G$ values show variation more akin to the original dataset (Table 3.3). Certain outliers are present which show positive $\Delta\Delta G$ values, however, but as we are only interested in maximizing the $\Delta\Delta G$ values, we did not investigate these outliers further. While model **B** appears to perform better than model **A**, it has limited predictability which may stem from its inability to properly reproduce the shape of the original ΔG_{Ln} histograms. When comparing the predicted and actual ΔG_{Ln} values, the errors do not converge in a smooth manner as more active learning steps are taken. Figure 3.10 shows that for both the ΔG_{Ln} and $\Delta\Delta G$ values, the addition of more data in the form of more active learning steps does nothing to lower the error of the predicted values. Thus, model **B** also fails to be of use for the prediction of free binding energies of Lnβ_3 complexes.

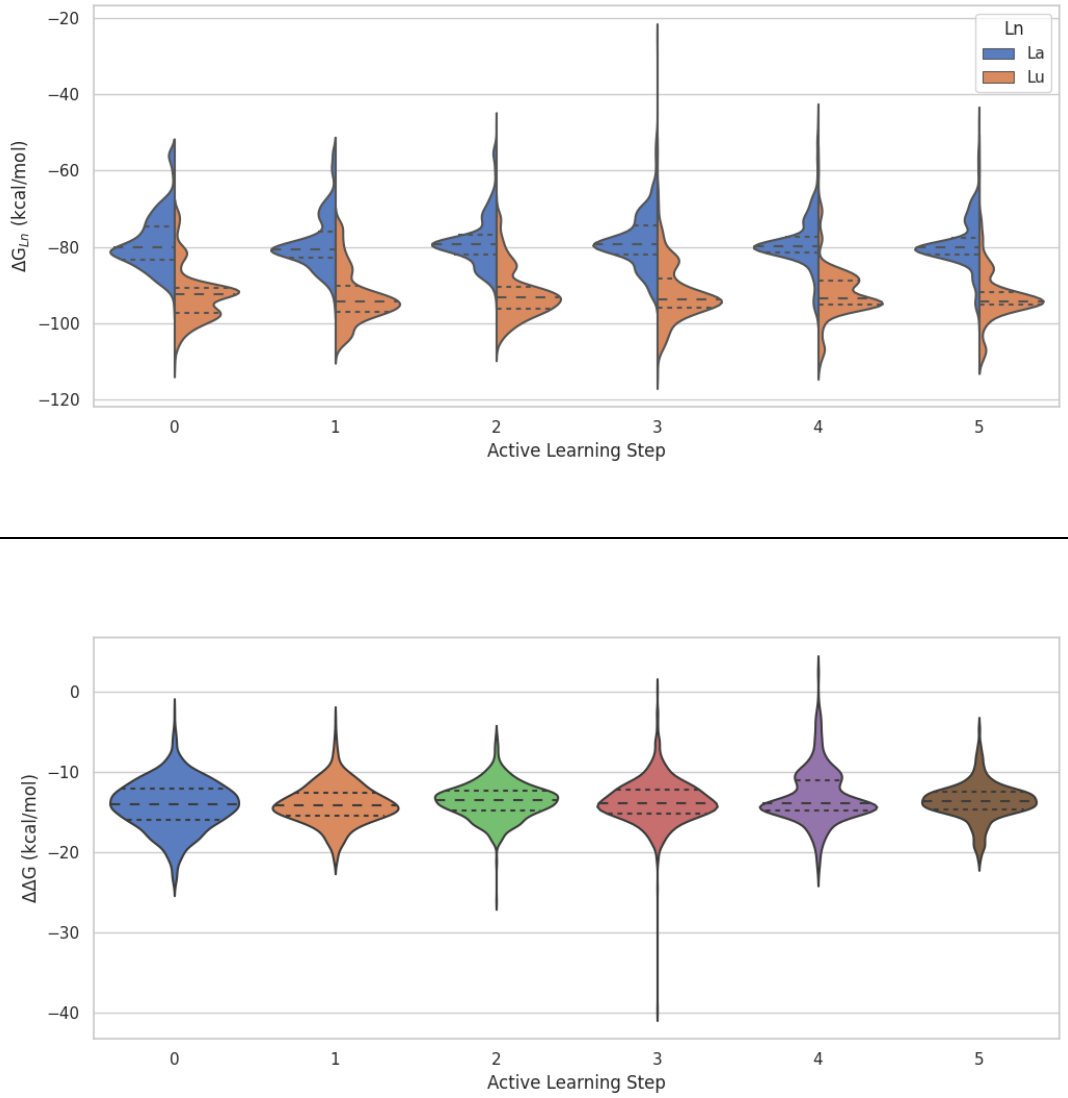
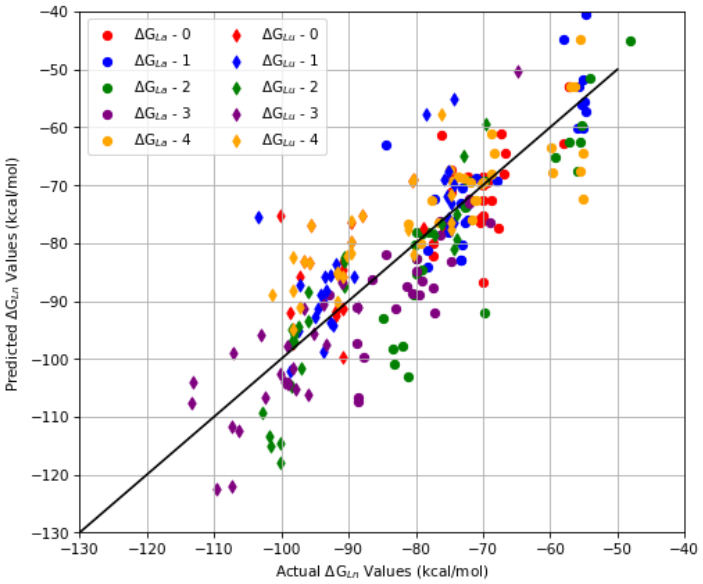


Figure 3.9 Violin plots for Model **B**. ΔG_{Ln} values (top) and $\Delta\Delta G$ values (bottom)



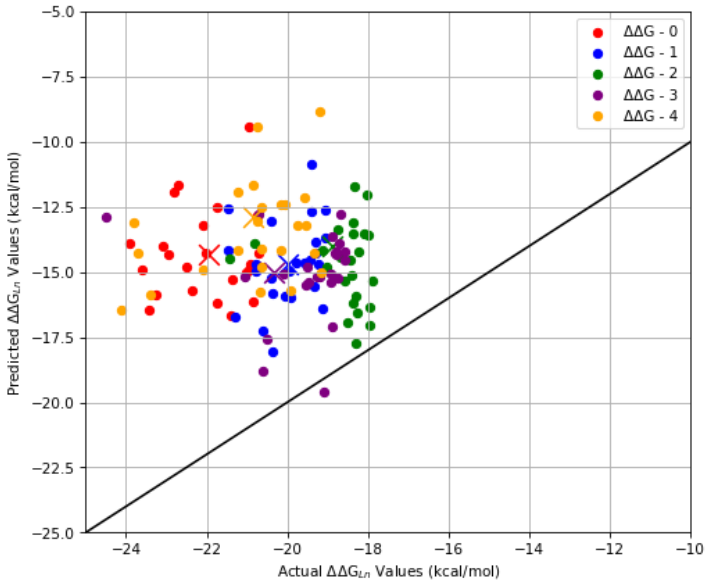


Figure 3.10 Plots showing actual and predicted ΔG_{Ln} values (left) and $\Delta\Delta G$ values (right) for Model **B**. The larger points in each plot correspond to the average of each set of data.

Following this, we investigated the predictive capabilities of a different ML architecture to better understand why models **A** and **B** fail to accurately predict the free binding energies of the Lnβ_3 complexes. This model is referred to as model **C** and was chosen based upon our initial cross-validation study. We chose a learner with a different functional form from those used for models **A** and **B**. The kernel ridge regression (KRR) learner has shown extensive use in high-dimensionality problems and may show promise for the current study. We chose the model with the highest accuracy from our cross-validation study which utilized the KRR learner and proceeded to perform five iterative active learning steps just as with models **A** and **B**. Figure 3.11 shows violin plots for both the ΔG_{Ln} and $\Delta\Delta G$ values. The shape of the ΔG_{Ln} violin plots is much smoother than those seen for model **B** while at the same time does not predict that the maximum value is the average value, as was seen with model **A**. Based upon these observations, model **C** may lead to better predictability than either model **A** or **B**. Shown in Figure 3.12 is a plot of the DFT calculated $\Delta\Delta G$ values versus the ML predicted $\Delta\Delta G$ values as a function of the active learning steps. The first active learning step is similar to that seen with model **B** but as the steps progress, the predictive capabilities of the KRR learner become more apparent. The predicted $\Delta\Delta G$ values using the initial, 100-instance dataset show a large error compared to the actual values with a mean absolute error (MAE) of 4.5 kcal/mol. This error is similar to the MAE observed with model **B**. Once the dataset is expanded with the new values, however, the error drops to between 1.0 – 1.7 kcal/mol for the next four active learning steps. This indicates that the functional form of the KRR learner is more appropriate for the prediction of free binding energies for Lnβ_3 complexes.

The differences in performance among model **A** (using the random forest learner), model **B** (using the AdaBoost learner), and model **C** (using the kernel ridge regression learner) may lie in the dimensionality inherent in the prediction of free binding energies for Lnβ_3 complexes. Such predictions rely on high dimensional input vectors from the molecular representation and from the small physicochemical changes along the lanthanide series. These two factors result in a regression problem with significant complexity. Therefore, a complex function which maps the input vector (the molecular representation of the β ligand) to the output value (ΔG_{Ln} or $\Delta\Delta G$) is required. The learners used in models **A** and **B** are not as complex as the learner used for model C and thus may not be appropriate for the prediction of such free binding energies. For similar studies such as those

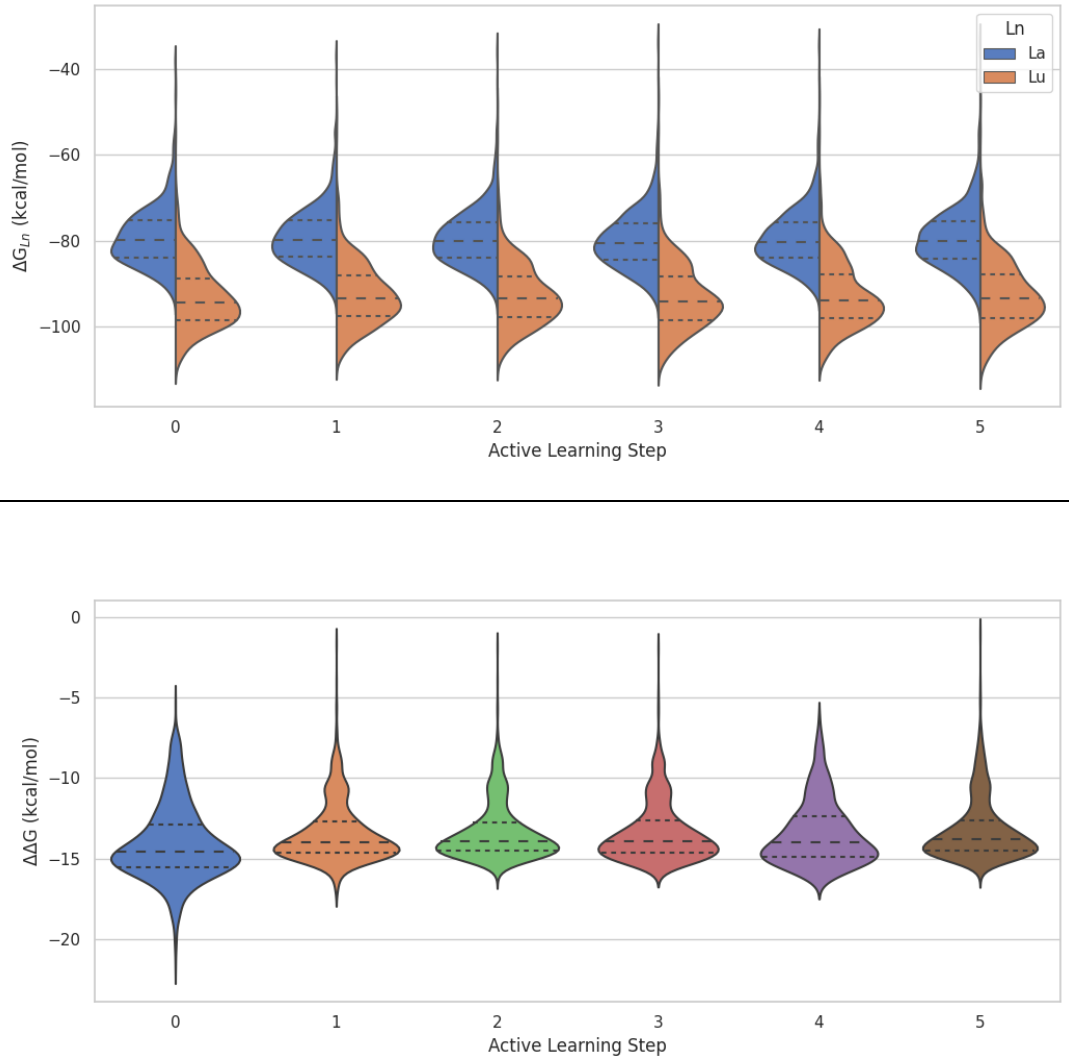


Figure 3.11 Violin plots for model which utilizes the KRR learners, PIs as molecular representations, and predicts the individual ΔG_{Ln} values. ΔG_{Ln} values (top) and $\Delta\Delta G$ values (bottom)

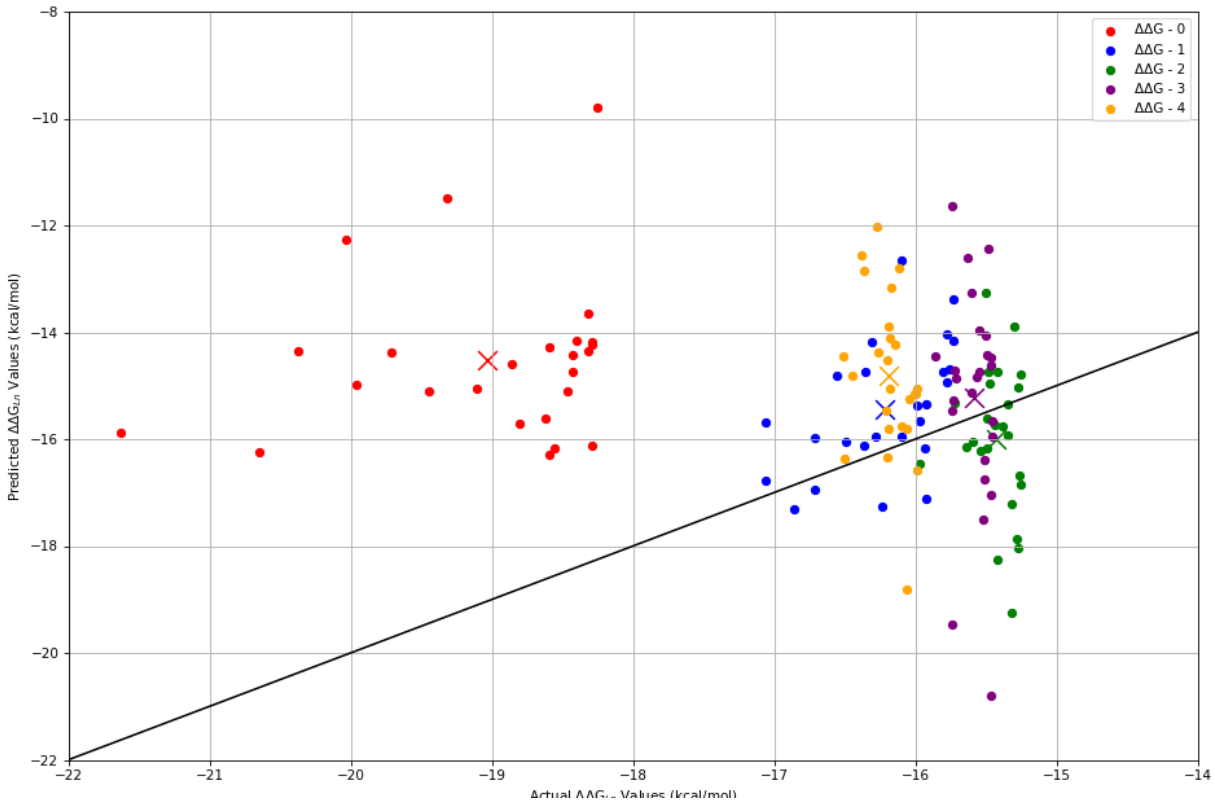


Figure 3.12 Plot showing actual and predicted ΔΔG values for Model **C**. The X's correspond to the average ΔΔG values for each active learning step.

which may use different chelating ligands or which may utilize different metals such as the actinides, a kernel method is recommended based on the work presented herein.

3.3.5 Conclusions

Based upon the information presented here, we have concluded that the most appropriate a machine learning model to predict ΔG_{Ln} values for Lnβ_3 complexes involves a learner which is able to map data to higher dimensional spaces (such as KRR, SVM, BRR, or GPR). Other learners such as random forest and AdaBoost were able to learn within a given dataset quite well (with r^2 values between 0.7 – 0.9) but showed limited predictability due to either regressing to the mean (random forest) or simply not improving when more data is introduced (AdaBoost). The lack of predictability for these two learners may be due to the complexity of the dataset which may arise from the unique interactions of the β ligands and the lanthanide atoms. This interaction results from small physicochemical changes between lanthanum and lutetium which may be too complex for learners such as random forest or AdaBoost. A learner that is much more appropriate for higher-dimensional datasets was shown to not result in these same behaviors. The kernel ridge regression learner resulted in similar training accuracies while also being able to make accurate predictions beyond the given dataset. Following five active learning steps, the error in the predicted $\Delta\Delta G$ values was below 2 kcal/mol. This work can be extended to other ligand systems such diglycolamides or bis-phosphine oxides or to other metal separation systems such as the actinides.

3.4 Conclusions

Computational methods are an invaluable aspect of scientific research due to the ability to study systems on the atomic scale. Lanthanide complexes in particular are important to study as they are found so ubiquitously in modern technologies such as green energy and electronic devices. The high prevalence of lanthanide-based materials thus requires very pure starting materials which is a difficult task to achieve due to the similar physicochemical properties along the lanthanide series. Through the use of computational methods, a high-throughput screening approach can be

used to analyze many hundreds or thousands of lanthanide complexes in order to better understand the lanthanide separation process. To begin this screening approach, a computational method was identified which resulted in accurate predictions of thermodynamics and geometric values for different lanthanide complexes. This method involved the BP86 density functional, the zeroth-order, relativistic approximation (ZORA), and the SARC basis set. Following this, a wide breadth of study was undertaken in order to identify a machine learning model which could accurately make predictions regarding the free binding energies of lanthanide-tris-β-diketone complexes. A combination of the kernel ridge regression (KRR) learner and persistent images resulted in errors between 1 – 1.7 kcal/mol compared to the DFT computed $\Delta\Delta G$ values.

Chapter 4

Computational Catalysis: Molecular Complexes and Metal-Organic Frameworks

The following chapter examines how electronic structure theory calculations can be applied for the understanding of catalytic reaction mechanisms. The dehalogenation of CH_2Cl_2, the hydrogen evolution reaction (HER), and the hydrogenation of light alkenes including ethyne, ethene, and propene are representative examples that are discussed in this chapter. Each section begins first with an explanation of the context of the particular project. Following this, experimental results are outlined and a detailed analysis of the calculations performed for each system and reaction(s) will be presented. Most of the work presented herein corresponds to published work or work near being published where the author's (Gavin McCarver) contribution was to perform said electronic structure calculations.

4.1 Introduction

Catalysis is central to many manufacturing processes vital to the economic growth of modern societies. Processes such as the production of ammonia to be used as fertilizer (utilizing the Haber-Bosch process)[59], the catalytic cracking of long chain hydrocarbons to produce value added products such as gasoline,[25] as well as the catalytic process which produces sulfuric acid to be used in any number of products all would be impossible on the current scale without the use of effective catalysts. Catalysts are categorized as either heterogeneous, whereby the catalyst is generally in the solid state and interacts with either liquid or gaseous reactants,[290] or homogeneous

where the catalyst and reactants are in the same phase (usually in solution).[291] The role of a catalyst is to lower the activation barrier and allow the reactions to proceed at a higher rate and oftentimes with a much more efficient conversion of reactants to products (atom economy). An uncatalyzed reaction may require either higher temperatures or pressures or may proceed at extremely low rates. Catalysts most often employ transition metal complexes and can exist in a wide array of structural motifs such as single atoms sites, two-dimensional materials, metal clusters, or other nanomaterials.[292] In addition, if a single-atom site or a metallic cluster is bound to a substrate, then catalysis can occur there as well, oftentimes utilizing the unique properties of those substrates.

The large number of possible catalytic reactions leads to an intractable number of combinations of reactants, products, and catalysts.[265] This chemical space is quite vast and so not every reaction can be examined using every catalyst. Thus, most catalysts are designed usually with one reaction in mind in the hopes of selectively producing a single product such as the hydrogen evolution reaction (HER)[293] or the hydrogenation of small alkenes.[294] Reactions such as these involve small, gaseous molecules which may be utilized in either homo- or heterogenous systems. In the case of hydrogenation, it is preferable to pass a feedstock of gaseous alkenes and H_2 over a heterogeneous catalyst so as to maximize the number of interactions while minimizing the amount of catalyst required.[295] Other reactions such as the production of H_2 in aqueous media require the use of a homogeneous catalyst due to the reactant species being H^+ and e^-.[296] Both previously mentioned reactions would benefit from a synergistic synthesis whereby the catalyst encompasses both homogeneous and heterogeneous properties.[297,298]

Metal-organic frameworks (MOFs) and heterogenized-homogenous catalysts are two such classes of materials which include the previously mentioned properties. MOFs involve highly ordered materials which can exist (usually) in the solid state and interact with gaseous and liquid reactants. Likewise, homogenous catalysts which have been heterogenized to some material (and no longer exist explicitly in the aqueous phase) exhibit a combination of the desired properties of both heterogeneous and homogeneous catalysts. MOFs have been shown to be excellent catalysts for reactions such as hydrogenation[294,299–302] and dehydrogenation,[303–306] dimerization[307–309] and oligomerization,[310] hydrolysis[311–317] and dehydration,[318–320] alkane oxidation,[309,321–341] alkane epoxidation,[342–344] CO oxidation,[345–349] sulfur oxidation,[350,351] water oxidation,[352,353] and reduction

reactions.[354–357] Each of the previously mentioned reactions have been studied through a combination of experimental and theoretical chemistry.[60,293,358] Likewise, the types of reactions which have been examined with heterogenized, homogeneous catalysts include HER,[359] dehalogenation,[360] hydrogenation,[361] and carbonylation.[362]

Herein is described three main catalytic reactions which include the dechlorination of dichloromethane, the production of H_2 in aqueous medium, and the hydrogenation of small unsaturated hydrocarbons. The latter reaction includes a brief discussion on hydroformylation while the second reaction is examined using two different catalysts. The dechlorination reaction is examined solely using one catalytic complex. Each reaction is studied with a material that shares properties with both homogeneous and heterogeneous catalysts.

4.2 Electrocatalytic Dechlorination of Dichloromethane Using a Molecular Copper Complex

Disclosure

The following subsection contains a modified subsection of a published manuscript in *Inorganic Chemistry*.[363] Gavin McCarver performed all density functional theory calculations for this study. Experimental work was performed by Caroline K. Williams and Amir Lashgari (University of Cincinnati). Manuscript preparation was done under the guidance of Dr. Konstantinos D. Vogiatzis and Dr. Jianbing "Jimmy" Jiang (University of Cincinnati). Reprinted with permission from *Inorganic Chemistry*. Copyright 2021 American Chemistry Society.

4.2.1 Abstract

Density functional theory (DFT) calculations were performed to elucidate the mechanism by which the Cu(I) triazole catalyst converts CH_2Cl_2 into more benign compounds such as CH_4, C_2H_4, C_2H_6, and H_2. Experimental methods showed that the Cu(I) triazole catalyst displayed high Faradaic efficiency towards CH_2Cl_2 dehalogenation which would indicate low barriers towards the

breaking of the two C-Cl bonds and the subsequent formation of two C-H bonds. Barriers of 0.15 and 0.45 eV were calculated for the two bond-breaking and bond-formation steps which are exceptionally low for such strong C-Cl bonds. The formation of all products was found to be exergonic by as much as -1.2 eV. The Cu(I) triazole complex shows excellent use as a dehalogenation catalyst and may be utilized and studied for similar reactions.

4.2.2 Introduction

Halogenated organic compounds often pose risks to both human and animal life in numerous ways that can affect the organism directly (in the form of either acute or chronic exposure) or affecting the environment in which the organism is found (as in the thinning of the ozone layer).[364,365] The remediation of such compounds is often performed by converting the halogenated species into more benign molecules. One such example is dichloromethane (CH_2Cl_2, often referred to as DCM) which is a common solvent used in organic synthesis. The release of DCM into water sources raises concerns as the compound has a half-life of over a year when in groundwater sources. Current techniques for the removal of DCM from ground and wastewater include pervaporation, photodegradation, and heterogeneous catalytic methods.[366] In recent decades, electrocatalysis has been utilized for the decomposition (and subsequent removal) of DCM from water sources.[367,368] Materials such as metal oxides,[360,369] nanomaterials,[370] metal species deposited on carbon materials,[371] fullerenes,[372] and even the heme enzyme Cytochrome P450[373] have all been utilized with either high efficiency or reactivity towards the dehalogenation of certain halogenated organic compounds. Molecular electrocatalysts can show improvements over other materials as they can be tuned for certain properties via ligand functionalization as well as the easy inclusion of different metal species via transmetalation. Common ligands for such systems include porphyrinoids,[358] fullerenes,[374] diphosphines,[375] dimethylglyoxime,[376] and thiosemicarbazones[377] with metals such as Fe,[358] Co,[378] Ni,[375] and Mn.[375]

Experimental work using a Cu(I) triazole (**CuT2**, shown in Figure 4.1) molecular electrocatalyst for the dehalogenation of DCM to form CH_4 was recently performed. The complex was synthesized via a one-step azide-alkyne cycloaddition using $CuSO_4{\cdot}5H_2O$. Several studies

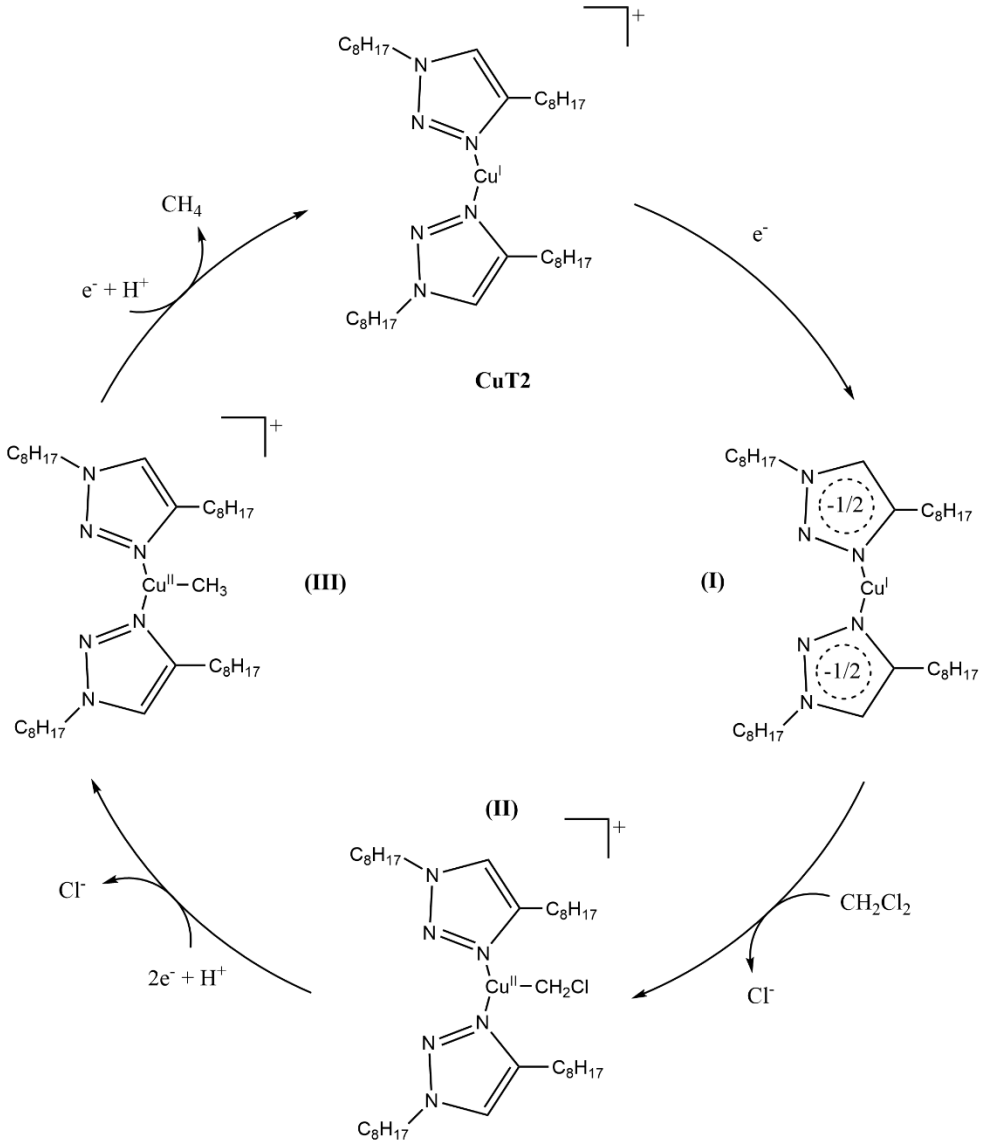


Figure 4.1 Proposed four-electron-two-proton catalytic mechanism for the production of CH_4 from CH_2Cl_2 using **CuT2**.

were performed to elucidate the chemical and electrocatalytic properties of the **CuT2** catalyst. By using linear sweep voltammetry (a method to investigate current as a function of voltage), the catalytic current of the **CuT2** complex was shown to drastically increase when DCM was added to the solution thus indicating high activity. Following this, controlled potential electrolysis was performed in order determine the product distribution and Faradaic efficiency towards dehalogenation. As Cu catalysts have been shown to be quite effective towards C-C coupling to form C2 products, several products were expected using the **CuT2** catalyst. The major product was found to be CH_4 (70% Faradaic efficiency) with less than 5% of the products being C_2H_6 and C_2H_4. The other major product was molecular hydrogen (H_2) which accounted for 25% of the produced material during electrolysis. Control experiments were likewise performed to ensure that the **CuT2** compound was the catalytically active material. From these results, it was determined that the **CuT2** compound was indeed catalytically active towards the dehalogenation of CH_2Cl_2 in order to form more chemically benign products.

4.2.3 Computational Details

Calculations were performed with the ORCA 4.2 software package.[118] A computational procedure similar to Jiang *et al.*[359] was followed where the B3LYP[232,233,236] density functional was utilized. Standard convergence criteria were used for both the SCF cycles and geometry optimization steps. The resolution of identity (RI)[248] approximation with the Chain of Spheres algorithm (COSX)[379] was used to accelerate the computation of the four index integrals with the def2/J auxiliary basis sets.[380] A singlet state was assumed for all structures with a Cu(I) site and a doublet state for those with a Cu(II) site. Each structure was optimized with the Def2-SVP basis set for all main group elements (C, H, N) and the def2-TZVP basis set for Cu.[275] Dispersion effects were accounted for by using Grimme's D3 semiempirical method[276] with the Becke-Johnson damping function.[277] Following the geometry optimization, the energy for each structure was corrected for by including various effects including thermal, thermodynamic, and solvation effects and through the use of larger basis sets for each atom. Thermal effects correct for the zero-point energy of a system according to the quasi-harmonic oscillator approximation while

thermodynamic effects correct for enthalpic and entropic effects. Both effects are accounted for by performing a numerical frequency calculation which also ensures that the geometries of all reactants, products, and reaction intermediates are all located on stationary points on the potential energy surface. Implicit solvation effects correct for calculations performed in the gas phase. This effect is included using the conductor-like polarizable continuum model (CPCM)[258] using water as the solvent for all intermediates and transition states. Lastly, a single point calculation was performed using basis sets larger than those utilized during the optimization steps (the def2-TZVPP basis set for Cu and def2-TZVP basis set for all other atoms). The total energy of each structure is evaluated according to Eq. 4.1,

$$G_{total} = E^{\mathrm{SVP}} + \delta E^{\mathrm{SVP}}_{\mathrm{Thermal}} + \delta E^{\mathrm{SVP}}_{\mathrm{Thermodynamic}} + \delta E^{\mathrm{SVP}}_{\mathrm{Solvent}} + \delta E^{\mathrm{TZVP}} \quad (4.1)$$

where E^{SVP} corresponds to the total electronic energy at the optimized geometry with the SVP basis set while $\delta E^{\mathrm{SVP}}_{\mathrm{Thermal}}$, $\delta E^{\mathrm{SVP}}_{\mathrm{Thermodynamic}}$, $\delta E^{\mathrm{SVP}}_{\mathrm{Solvent}}$, and δE^{TZVP} correspond to corrections from thermal, thermodynamic, and solvent effects as well as the effects of utilizing a larger basis set, respectively.

4.2.4 Results

A plausible four-electron/two-proton mechanism was proposed for the dechlorination of CH_2Cl_2 by the **CuT2** catalyst for the generation of CH_4 (Figure 4.2). Figure 4.2.2 shows the free energy profile for the studied mechanism. All reaction free energies reported were calculated at 298 K (ΔG^{298}). The potential of any incoming electrons is referenced relative to the half reaction of ½ $H_2 \longrightarrow H^+ + e^-$ and has a ΔG^{298} = +0.63 eV. Following an initial reduction to form intermediate **I** which is exergonic by -1.65 eV, nucleophilic attack via CH_2Cl_2 leads to the formation of intermediate **II** and the subsequent loss of Cl^-. The binding of CH_2Cl_2 to the reduced catalyst was shown to be thermodynamically unfavorable by 0.13 eV (3.0 kcal/mol). This step proceeds

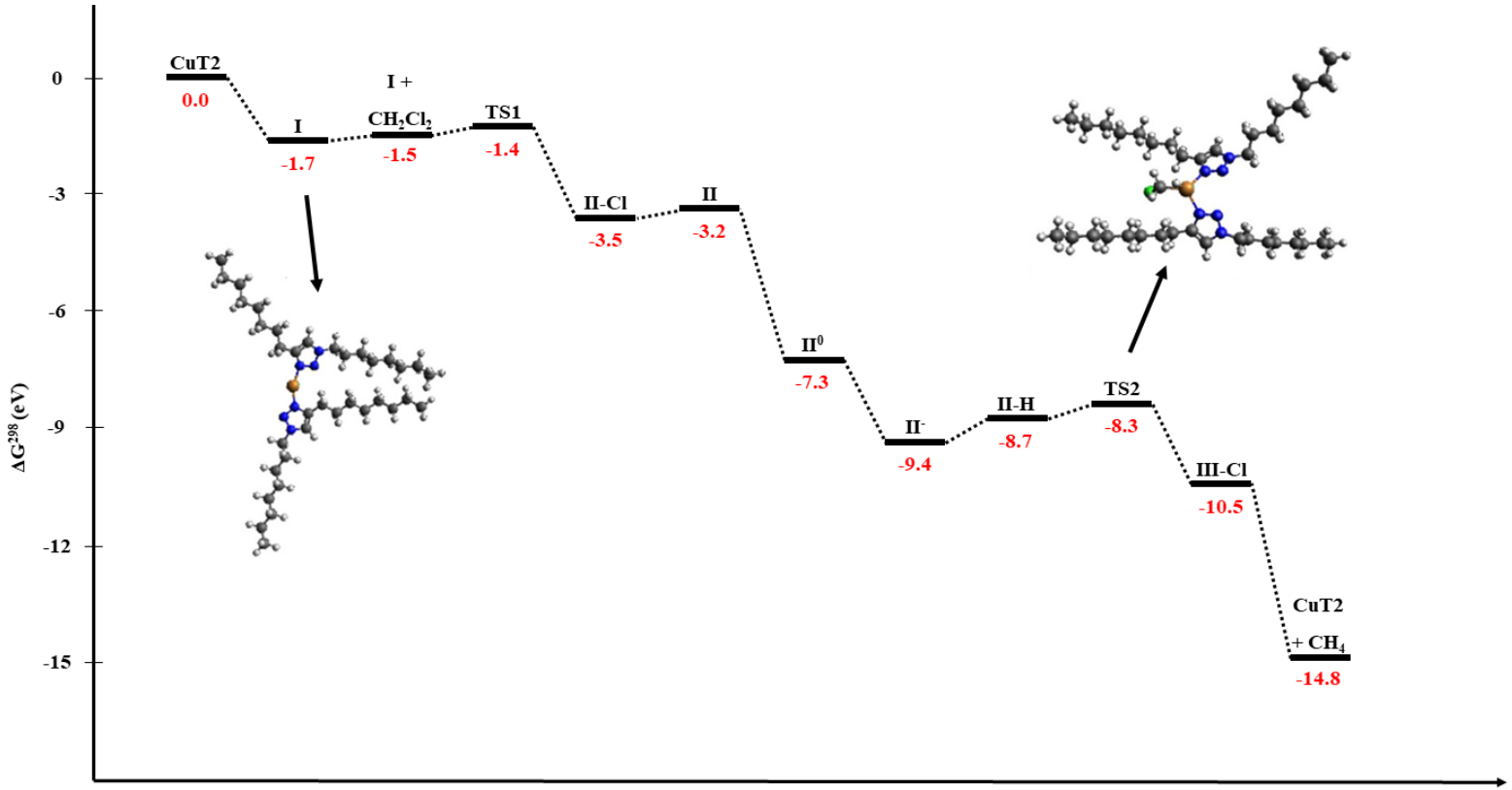


Figure 4.2 Free energy profile of the dechlorination reaction together with representative examples of (near)-linear Cu(I) and pseudo-tetrahedral Cu(II) complexes. Color Code: Cu – gold, N – blue, C – grey, Cl – green, H – white.

following the breaking of the first C-Cl bond and the formation of a C-Cu bond which has a small transition barrier of 0.15 eV which is significantly lower than the C-Cl bond strength of dichloromethane (2.46 eV). The decoordination of Cl^- from the Cu site has a free energy difference of 0.28 eV and leads to the formation of intermediate **II** which must now undergo five steps: (i) two one-electron reductions, (ii) a protonation via nucleophilic attack, (iii) the activation and (iv) dissociation of the second C-Cl bond, and (v) the subsequent removal of a second Cl^-. Calculations indicate that intermediate **II** proceeds following the two one-electron reductions to form intermediates $\mathbf{II}^0$ and $\mathbf{II}^-$, respectively. As no mono-chlorinated species (CH_3Cl) were observed experimentally, protonation should precede the fourth (and final) one-electron reduction step to form intermediate **II**-H. The breaking of the second C-Cl bond and the initial proton transfer step happen simultaneously and show a kinetic barrier of 0.45 eV which again is much lower than the computed C-Cl bond strength of 3.40 eV for CH_3Cl. Finally, the formation of intermediate **III**, the fourth one-electron reduction, and the last protonation steps are energetically favorable, which lead to the formation of CH_4 and the regeneration of the **CuT2** electrocatalyst.

The formation of C_2 products was also considered from intermediate **III** during the last step of the catalytic reaction. The following free energies of formation were computed:

$$\text{III} + H^+ \longrightarrow \mathbf{CuT2} + CH_4 \qquad \Delta G = -4.4 \text{ eV} \qquad (4.2)$$

$$2\ \text{III} \longrightarrow 2\ \mathbf{CuT2} + C_2H_6 \qquad \Delta G = -3.0 \text{ eV} \qquad (4.3)$$

$$2\ \text{III} \longrightarrow 2\ \mathbf{CuT2} + C_2H_4 + H_2 \qquad \Delta G = -1.2 \text{ eV} \qquad (4.4)$$

Even if the formation of the three products is exergonic, it is evident that the CH_4 formation (–4.4 eV) is more thermodynamically favorable than C_2H_6 and C_2H_4 (–3.0 eV and –1.2 eV, respectively). Alternatively, C_2H_4 formation might proceed through the $CClH_2$ radical intermediate (intermediate **II**):

$$2\ CH_2Cl + 2\ e^- \longrightarrow C_2H_4 + 2\ Cl^- \qquad \Delta G = +1.8 \text{ eV} \qquad (4.5)$$

However, this process is unfavorable as it involves the direct dissociation of two C-Cl bonds. Thus, the formation of CH_4, C_2H_6 and C_2H_4 should procced via the same reaction intermediate (**III**).

4.2.5 Conclusions

Using computational methods, a better understanding of the mechanism through which the **CuT2** molecular electrocatalyst works was achieved. The conversion of CH_2Cl_2 to form CH_4 through a four-electron/two-proton mechanism was shown to be highly efficient with a Faradaic efficiency of 70%. This high efficiency indicates the reaction proceeds with small activation barriers, the largest of which was calculated to be 0.45 eV (10.4 kcal/mol) for the simultaneous breaking of the second C-Cl bond and the formation of the second C-H bond. Likewise, the formation of C2 products (C_2H_4 and C_2H_6) were shown to be thermodynamically favorable as well. The **CuT2** catalyst thus serves as an excellent material for the conversion of CH_2Cl_2 into more benign products such as CH_4, C_2H_4, C_2H_6, and H_2.

4.3 Electrocatalytic Proton Reduction using p-Block Metal Catalysts

Disclosure

The following section includes work with minor modifications which has been taken with permission from the following two manuscripts:

(1) A manuscript that has been submitted for publication in *Chemistry – A European Journal* which is currently in revision. Gavin McCarver performed all DFT calculations for this study. Experimental work was performed by Caroline K. Williams, Ashwin Chaturvedi, Sourmalya Sinha, and Marcus Ang (University of Cincinnati). Manuscript preparation was done under the guidance of Dr. Konstantinos D. Vogiatzis and Dr. Jianbing “Jimmy” Jiang (University of Cincinnati).

(2) A manuscript that has been submitted for publication. Gavin McCarver performed all DFT calculations for this study. Experimental work was performed by Ashwin Chaturvedi and Soumalya Sinha (University of Cincinnati). Manuscript preparation was done under the guidance of Dr. Konstantinos D. Vogiatzis and Dr. Jianbing "Jimmy" Jiang (University of Cincinnati).

4.3.1 Abstract

The electrocatalytic production of molecular hydrogen (H_2) is one pathway through which a carbon-neutral energy source may be obtained. Expensive transition metals such as platinum and palladium increase the production cost of H_2 and thus alternative and cheaper metal catalysts are being considered. Here, two such alternative metal catalysts which utilize more abundant, main-group elements have been used for the hydrogen evolution reaction (HER) to produce H_2 via electrolysis. The antimony N,N'-ethylenebis(salicylimine) (**SbSalen**) and a tin porphyrin (**SnPEGP**) catalysts were both synthesized using metalation of the respective free-base ligand. The SnPEGP catalyst was synthesized with the inclusion of a polyethylene glycol (PEG) unit to enhance solubility. Faradaic efficiencies of 100% and 94% were calculated experimentally for the respective catalysts. Computational methods helped to elucidate the mechanisms through which each catalytic surface proceeded in the production of H_2. It was found that the ECEC mechanism (which alternates reduction steps, E, and protonation steps, C) was most favorable for both the **SbSalen** and **SnPEGP** catalysts.

4.3.2 Introduction

The production of molecular hydrogen (H_2) from aqueous sources is a topic of active research due its possible use as an alternative to carbon-based fuel sources. A large portion of the H_2 production comes from fossil fuel-based sources, however, which precludes the H_2 being considered carbon-neutral. Excellent catalysts have been developed which can produce molecular hydrogen with high efficiency but often utilize expensive, transition metals such as platinum. P-block metal catalysts are a possible avenue towards making the production of H_2 more cost

effective and more environmentally friendly. Molecular electrocatalysts can further improve the production of H_2 by 1) allowing for functionalization through the ligand, 2) allowing for functionalization by choosing a different metal center, and 3) allowing for a homogeneous catalyst to be heterogenized, thus allowing for a synergistic system that is both homo- and heterogeneous in nature. Current molecular electrocatalysts that have been used for the hydrogen evolution reaction (HER) include a nickel phosphine complex,[381] porphyrinoids and polypyridyl species,[359,382,383] and N,N'-ethylenebis(salicylimine) (salen) complexes.[384] By utilizing different p-block metals and ligand scaffolds, the production of H_2 via electrolysis can be tuned to achieve the highest Faradaic efficiency. Two such examples include an antimony salen (**SbSalen**) complex and a tin porphyrin (**SnPEPG**) complex. Both were shown to be electrocatalytically active towards the production of H_2.

4.3.3 Hydrogen Evolution using a Molecular Antimony Complex

Overview and Experimental Results

The antimony N,N'-ethylenebis(salicylimine) complex was synthesized via antimony metalation of a salen ligand. In order to confirm the catalytical activity of the **SbSalen** complex, control experiments were performed whereby acid in the form of H_2SO_4 was added to a blank system (one in which no **SbSalen** is present). This solution showed a much lower current than one in which the **SbSalen** complex was present which indicates that the complex is indeed catalytically active. When the complex was heterogenized on a carbon paper working electrode, a Faradaic efficiency of 100% was measured to produce molecular hydrogen. This same carbon paper was then used for a second control experiment to test the longevity of the catalyst which again showed a Faradaic efficiency of 100%, indicating the electrocatalyst is stable and reusable for at least two cycles. A turnover frequency of 43.4 s^{-1} was calculated for the complex, which was found to be comparable to other, p-block metal based molecular catalysts.[381,385,386] X-ray photoelectron spectroscopy was used to detect any possible demetalated Sb on the carbon paper surface and none was found. Three control experiments were performed: one using a blank carbon paper, one using the antimony trichloride salt ($SbCl_3$) without the ligand, and the other using the metal-free salen ligand (**H2Salen**). These experiments indicated that the carbon paper is not catalytically active

while both the antimony salt and the metal free ligand showed some catalytic activity (Faradaic efficiencies of 93% and 74%, respectively) which demonstrates that the heterogenized **SbSalen** complex is very active towards the production of molecular hydrogen.

Computational Details

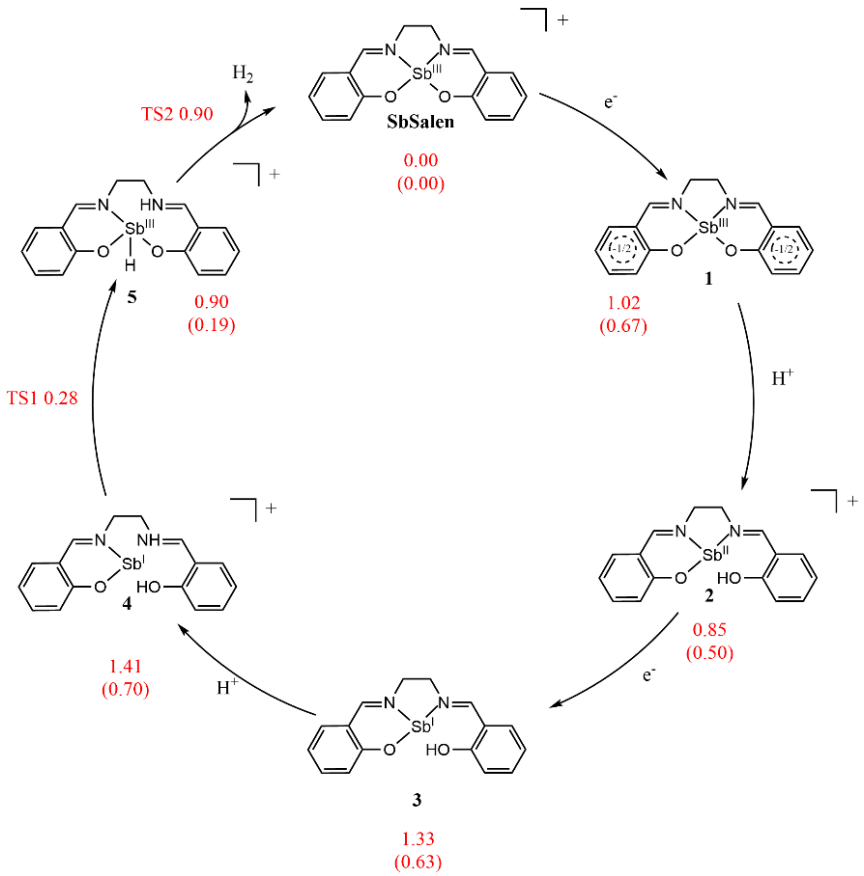
Theoretical calculations were performed by following a similar approach as in our previous study.[363] We utilized the Orca 5.0 software package[111] with the B3LYP density functional.[236,241] All structures were optimized using an unrestricted wave function and the def2-TZVPP basis set for Sb and def2-TZVP basis set for all other atoms.[275] Tight convergence criteria were used for both the SCF cycles and geometry optimization steps. The resolution of identity (RI) approximation[248] was applied to accelerate the computation of the four index integrals with the def2/J auxiliary basis sets.[380] Analytical frequency calculations were performed to ensure the minima of the potential energy surface were found for the reactants, products, and reaction intermediates, as well as to characterize transition states by locating a single imaginary frequency along the reaction coordinate. The two lowest spin states were considered for all intermediates and it was found that all but one structure preferred the lowest spin state; a singlet spin state for all structures with an even number of electrons and a doublet spin state for the reduced structures with an odd number of electrons. Single-point calculations with the conductor-like polarizable continuum model (CPCM)[258] implicit solvation model were performed for all intermediates and transitions states using water as the solvent. To properly account for the chemical potential of the H^+/e^- pairs, half of the total energy of H_2 was added following the addition of each pair.[387] For the individual reduction steps, a value of 4.22 eV was added as the Ag/AgCl electrode used during the CV experiments shows a E^0 value 0.22 eV higher than the standard hydrogen electrode (SHE) in water (4.44 ev).[388] Likewise, to properly account for the energy of H^+, a value of 11.78 eV was added at each protonation step. This value was chosen as the total energy of H_2 at the current level of theory equates to -31.99 eV. The two e^- would contribute 8.44 eV of this total energy which requires then that each additional H^+ would contribute 11.78 eV to the total energy of H_2.

Mechanistic Insights

The catalytic mechanism of HER systems has been extensively studied for transition-metal-based molecular systems. However, it has been underexplored for main-group-element-based systems. The HER is a two-electron/two-proton process. According to previous studies,[387,389–393] these electron transfer (E) and chemical reaction (C) steps can occur via several pathways: CCEE, CECE, CEEC, EECC, ECEC, and ECCE.[394] These steps of the reaction mechanism can occur either on the Sb center itself or on the salen ligand, given that the ligand shows non-innocent behavior. Although hydrogen evolution traditionally is carried out via a metal-centered pathway, where a transition metal utilizes multiple oxidation states for the two-electron/two-proton transfer steps, redox non-innocent ligands can also participate in HER.[395] Two categories of reactivity are ligand-centered reactivity, where the electron and proton transfer are both performed at the ligand, and metal-assisted ligand-centered reactivity, where the metal center of the catalyst acts as an electron reservoir for the catalytic reaction occurring at the ligand.[396] Berben's aluminum(III) NNN type pincer system showed ligand-centered HER, where the aluminum center's role is to tune the reduction potential of the molecular catalyst.[393] In contrast, Grapperhaus's copper(II) thiosemicarbazone system uses the NNSS-type ligand to perform HER while being supplied electrons from the Cu center, resulting in a metal-assisted ligand-centered reaction.[396] When the metal center is changed in the thiosemicarbazone complex, the available d-orbitals of that transition metal alter the reduction pathway, shifting the HER reactivity to be completely ligand-centered. Both of these examples show that it is possible to completely forgo the need of a metal-hydride pathway for HER with the use of non-innocent ligand.

DFT calculations were performed to elucidate which of these possible reaction mechanisms is more probable for the HER on the **SbSalen** complex. We have followed a computational procedure similar to our previous work on the electrocatalytic dechlorination of dichloromethane.[363] The proposed reaction mechanism is shown in Figure 4.3A and its corresponding reaction profile in Figure 4.3B. A detailed plot with the energy levels of all feasible reaction intermediates is provided in the Figure 4.4. Competitive reaction mechanisms are discussed below and also shown in Figures 4.5 and 4.6.

A



B

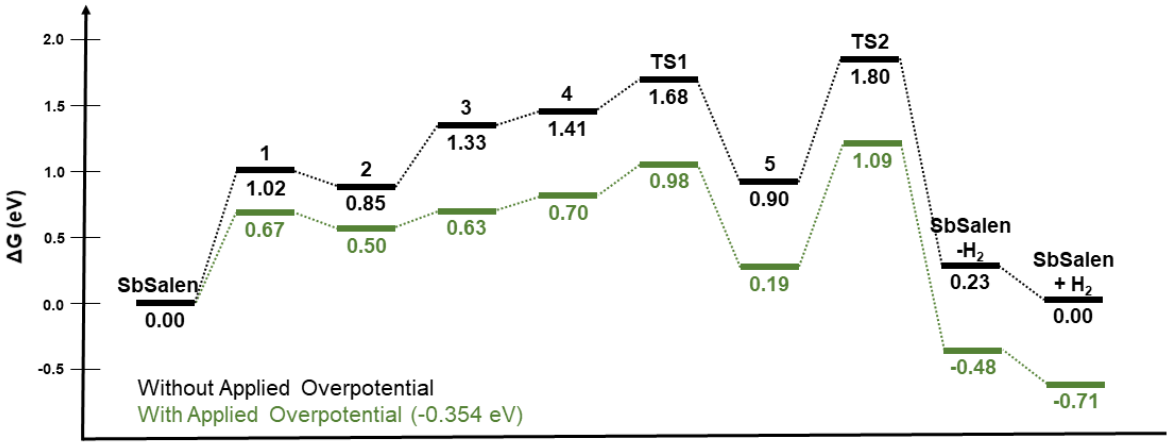


Figure 4.3 Proposed mechanistic cycle (A) and free energy profile (B) of **SbSalen** for HER. Values shown in eV. Computed ΔGs of each intermediate relative to the **SbSalen** catalyst shown in red. Values in parentheses represent the ΔGs with an applied overpotential (0.354 eV), which lowers the relative energy of each reduced species.

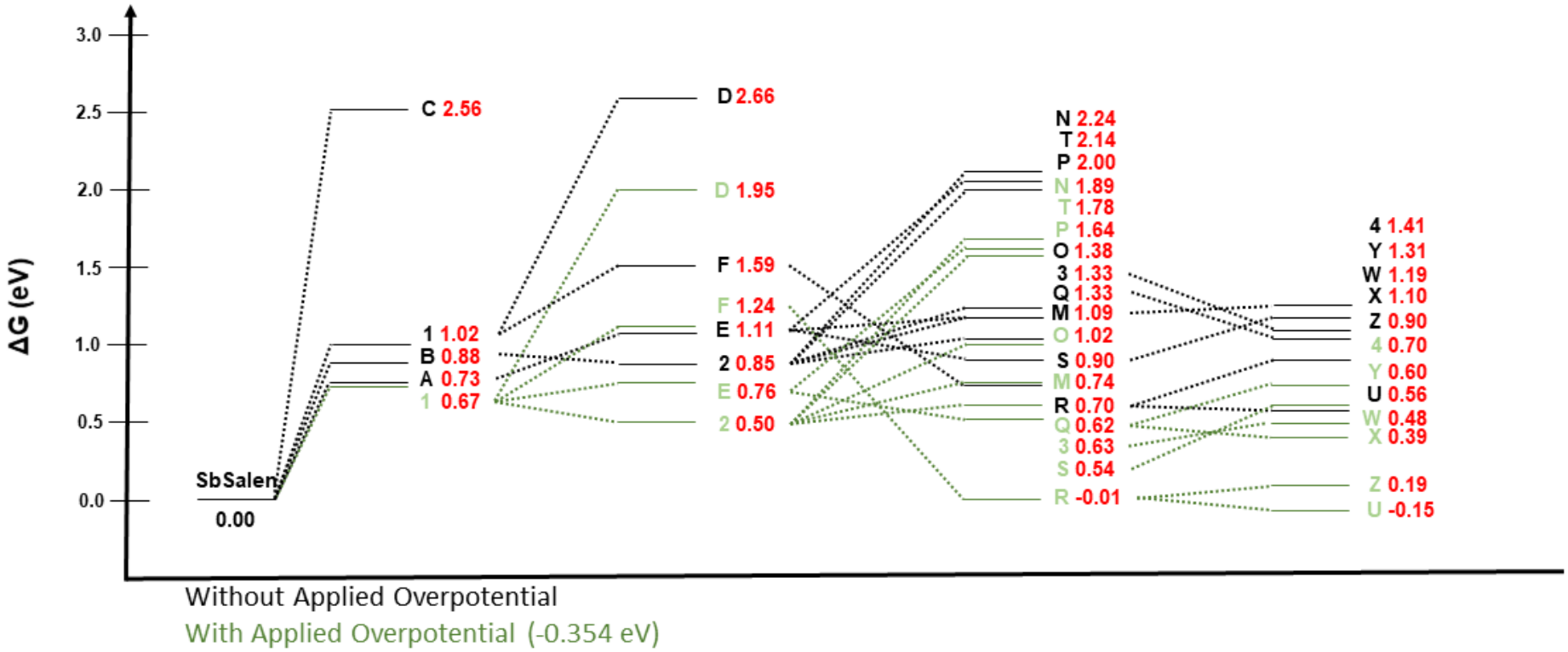
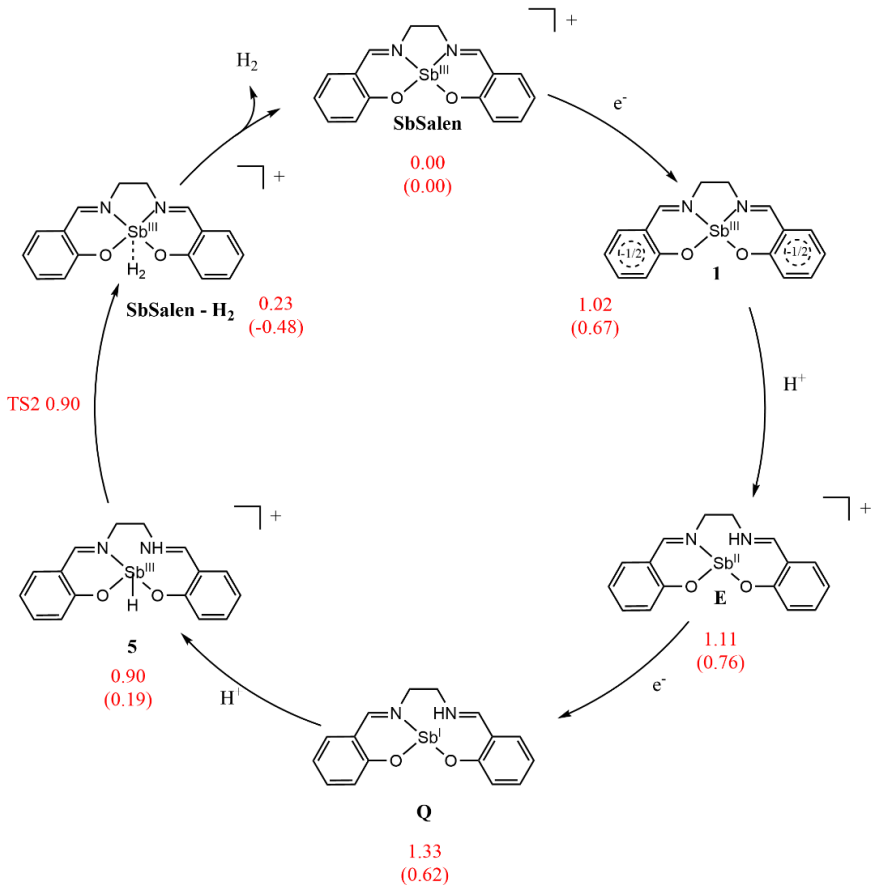


Figure 4.4 Proposed free energy profile for the two-electron/two-proton addition steps on **SbSalen** for HER. Values shown in eV. Each intermediate is shown connected via the lowest energy pathway to form that intermediate.

A



B

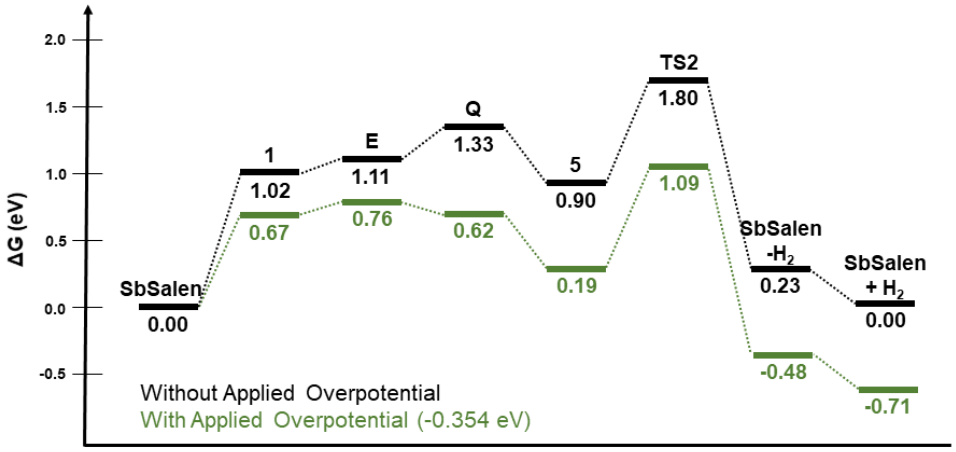
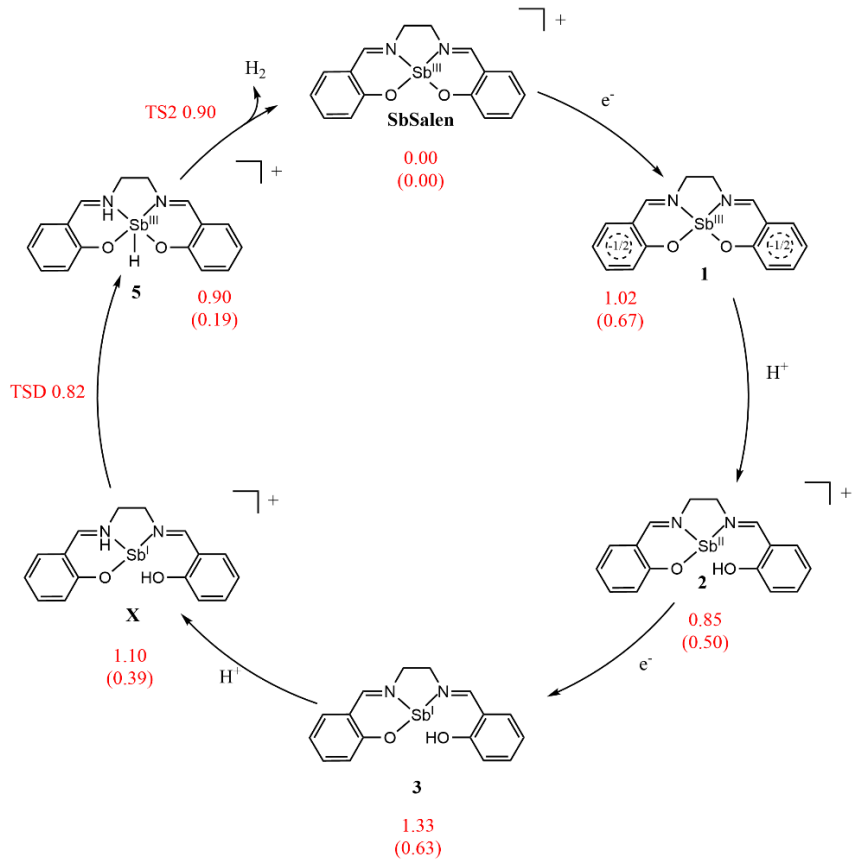


Figure 4.5 Proposed mechanistic cycle (A) and free energy profile (B) of **SbSalen** for HER following the formation of an Sb-H and N-H bond. Values shown in eV. Computed ΔGs of each intermediate relative to the **SbSalen** catalyst shown in red. Values in parentheses represent the ΔGs with an applied overpotential (0.354 eV), which lowers the relative energy of each reduced species.

A



B

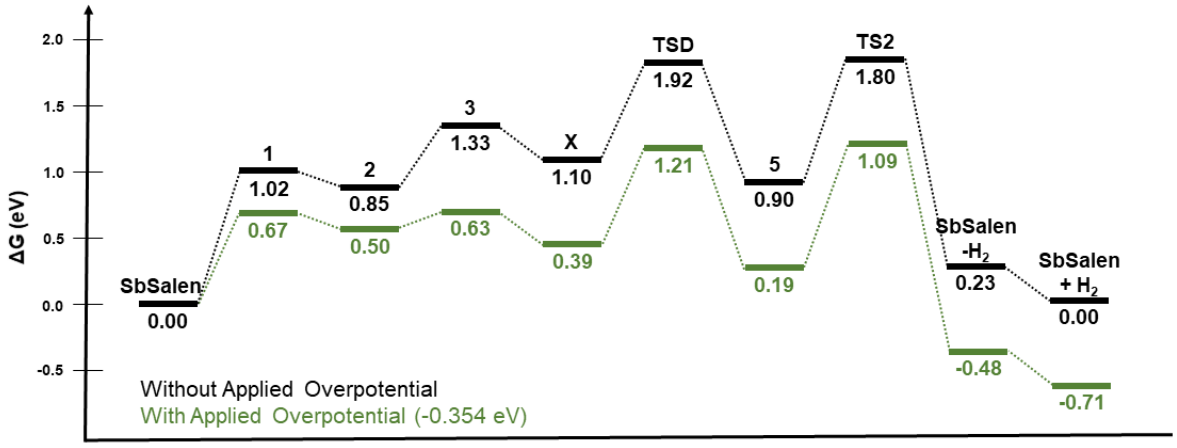


Figure 4.6 Proposed mechanistic cycle (A) and free energy profile (B) of **SbSalen** for HER following the formation of an O-H and N-H bond on opposite sides of the Sb atom. Values shown in eV. Computed ΔGs of each intermediate relative to the **SbSalen** catalyst shown in red. Values in parentheses represent the ΔGs with an applied overpotential (0.354 eV), which lowers the relative energy of each reduced species.

The first step of the reaction involves either a reduction or the protonation of an Sb, N, or O atom. The free energy difference (ΔG^{298}) between the starting **SbSalen** catalyst and the reduced species was found to be 1.02 eV (0.67 with the applied overpotential of 354 meV,[397] see intermediate **1** in Figure 4.3A). With the application of the overpotential, the reduction step is lower in energy than either an O- or N-protonated **SbSalen** complex; 0.88 (intermediate **B)** and 0.73 eV (**A**), respectively. The protonation of the Sb site was found to be less thermodynamically favorable with a ΔG^{298} value of 2.56 eV (**C**). Thus, the reaction most likely is initialized with the reduction of the **SbSalen** complex but protonation of a N or an O atom lie close in energy.

From intermediate **1**, the reaction can proceed with either the protonation of an Sb, O, or N atom or a second reduction step. The protonation of an O atom was found to be the most thermodynamically exergonic step with a free energy difference of –0.17 eV (intermediate **2**). The protonation of an N atom was found to lie slightly higher with a ΔG^{298} value of 0.09 eV (**E**) while the protonation of the Sb atom was found to be higher in energy at a value of 0.57 eV (**F**). A second reduction (**D**) was found to be the most unfavorable step at 1.63 eV (1.28 with the applied overpotential) higher in energy than the singly reduced species and it was found that the triplet spin state was the most favorable state by 0.17 eV. Based upon these thermodynamic considerations, the first two steps for HER on the **SbSalen** complex most likely occur in an EC fashion through O atom protonation, but several competing reaction channels involving N or Sb protonation lie close in energy. In addition, mechanisms whereby two subsequent protonation steps occur prior to reduction were examined and each case resulted in free energy differences larger than 1.58 eV.

The spin density of the singly reduced species shows mostly delocalization along the salen ligand which is conserved when protonation of the Sb atom occurs. Conversely, if protonation occurs on the O or N atoms, the spin density is localized mostly on the Sb atom, indicating that the Sb atom has been reduced from Sb^{III} to Sb^{II}. This change in the electronic structure of the **SbSalen** complex helps to showcase the unique, synergistic nature of the Sb-Salen interaction whereby either the ligand or the metal may be involved in the reduction and protonation steps.

The third step can proceed through different reaction channels due to a variety of possible intermediates that lie close in energy. These reaction channels involve a second reduction step, the

protonation of the unprotonated O atom, the protonation of the Sb atom, or the protonation of either of the N atoms, which are not identical due to the lower symmetry of intermediate **2**. The most thermodynamically favorable step changes upon the addition of the overpotential. Without it, the protonation of the N atom nearest the protonated O atom is the most favorable step with a free energy difference of 0.24 eV (intermediate **M**) while the application of the overpotential renders a second reduction the most favorable step (intermediate **3**). The protonation of the opposite N, the second O, and the Sb atom were found to occur with free energies of 1.39 eV (intermediate **N**), 0.52 eV (intermediate **O**), and 1.15 eV (intermediate **P**), respectively, therefore, these steps are less favored. Following the initial reduction and protonation steps, it was found that the Sb atom is reduced from Sb^{III} to Sb^{II}. Likewise, the second reduction step here involves a further reduction of Sb^{II} to Sb^{I}.

The last step of the reaction corresponds to the second proton addition in one of the four available sites: the second O atom, either of the N atoms, or the Sb atom. It was found that the protonation of the Sb atom to form a hydride is more favorable than any other site with a free energy difference of -0.77 eV (intermediate **U**). The computed reaction barrier that involves the dissociation of the Sb-H and O-H bonds is significantly high (1.47 eV) and thus, it cannot lead to viable H_2 formation. For that reason, intermediate **U** is discarded. For the complex with two protonated O atoms, the second protonation step proceeds with a free energy difference of –0.14 eV (**W**) while the complexes with a protonated N and O atom proceed with free energies of –0.23 eV when the protonated N and O atoms are on different sides of the salen ligand (**X**) and 0.07 eV for complex with the protonated N and O atoms on the same side (intermediate **4**). As these steps either proceed slightly thermodynamically uphill (in the case of the protonation of the N atom closest to the already protonated O atom) or downhill, there is a possibility that the final protonation could occur at any of the four sites. Because of this, we have chosen to study the kinetics of the H_2 evolution steps which must proceed with the transfer of either one or two protons following an ECEC two-electron/two-proton transfer. We have considered that the Sb atom most likely facilitates the proton transfer as an intermediate step. Thus, an initial transition state must be identified which involves the X-H bond dissociation (X = O or N) and a Sb-H bond formation which brings the second H atom in close proximity to the first H atom for the H_2 formation step.

The first two complexes (intermediates X and 4) to discuss involve the formation of both an O-H and an N-H bond. These bonds can occur near adjacent ligand atoms or on opposite sides of the Sb ligand sphere. For the first two cases where the O-H and N-H bonds are formed near each other, either the N-H bond or the O-H bond can be dissociated first, leading to two possible transition barriers. The cleavage of the O-H bond and subsequent formation of an Sb-H bond (**TS1**) proceeds with a small barrier of 0.28 eV. Similarly, N-H bond dissociation and formation of Sb-H bond proceeds with a slightly larger barrier of 0.49 eV. From these two intermediates, the kinetic barriers which to lead to the generation of H_2 are 0.90 eV (**TS2**) and 1.47 eV, respectively. The latter limits the viability of this reaction channel due to the high barrier. For the case where the N-H and O-H bonds are formed across the Sb atom, the observed barriers are 1.31 eV for the N-H bond and 0.82 eV for the O-H bond dissociation (**TSD**). The latter has a second transition barrier of 0.90 eV (**TS2**). Based upon these results, three favorable pathways can lead to the formation of H_2 on the **SbSalen** complex: the first considers an N-H and an Sb-H bond formation (with a barrier of 0.90 eV), the second an O-H and an N-H bond formation between adjacent ligand atoms where the O-H bond is cleaved first (with barriers of 0.28 and 0.90 eV), and a third in which an O-H and N-H bond are formed on the different sides of the **SbSalen** complex where the O-H bond is cleaved first (with barriers of 0.82 and 0.90 eV). The activation and thermodynamic barriers for each of these three mechanisms are comparable which indicates that any of the three mechanisms could lead to H_2 formation. Alternative reaction channels were also studied such as the protonation of both N atoms which encounters a kinetic barrier of 2.68 eV for the initial N-H bond dissociation and the Sb-H bond formation. If two O atoms are protonated, the first barrier associated with the O-H bond dissociation can be overcome as it shows a transition barrier of 0.72 eV, but the second barrier of 1.47 eV limits the feasibility of this reaction channel.

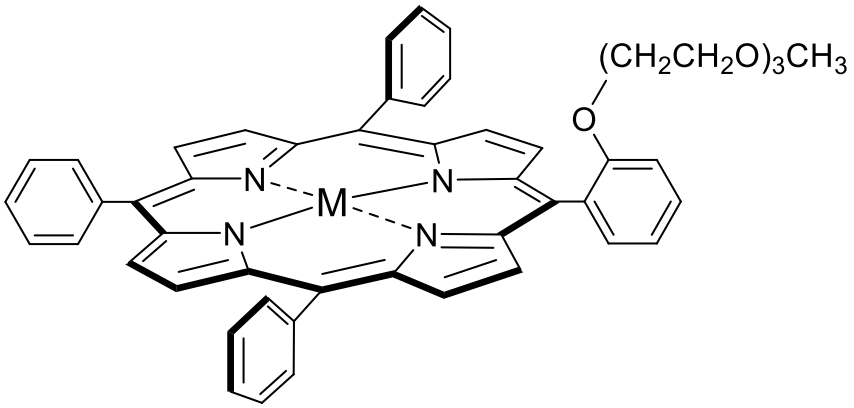
Based upon our results, we can draw some general conclusions regarding HER on the **SbSalen** catalyst. Firstly, as many reaction channels are possible, some amount of the catalyst may be found as stable intermediates from which no forward catalysis can occur. As the reaction proceeds, many intermediates are produced which are no longer catalytically active such as those with large, forward kinetic barriers. Thus, while the mechanism to produce H_2 proceeds with low enough barriers, many **SbSalen** complexes may be trapped as stable intermediates which may

lower the number of available catalytic sites. A remedy to this effect would be to run the electrocatalytic experiments at elevated temperatures so there is enough thermal energy to overcome the many barriers associated with HER on the **SbSalen** complex. Secondly, as the reaction proceeds, certain intermediates show conformational changes as a function of protonation or reduction steps. These changes may lead to the formation of other thermodynamically stable structures which would require elevated temperatures to change the geometry to a less stable conformation that is further along the reaction. Thus, many of the **SbSalen** active sites may be found near large kinetic or thermodynamic barriers which lowers the overall efficiency of the catalyst.

4.3.4 Hydrogen Evolution using a Molecular Tin Complex

Overview and Experimental Results

A tin porphyrin complex was synthesized via tin metalation of a free-base porphyrin ligand which has had a poly(ethylene glycol) unit (PEG) attached to enhance solubility in organic solvents (**SnPEGP,** Figure 4.7). Cyclic voltammetry (CV) was used to determine the HER activity of the **SnPEGP** catalyst using trifluoroacetic acid (TFA) as the proton source. When 100 equivalents of TFA were added, a dramatic increase in current was measured which suggests electrocatalytic proton reduction. A control experiment was performed whereby CV was ran with no catalyst and 100 equivalents of TFA. No catalytic activity was observed in this case which indicates that the **SnPEGP** complex is indeed the active catalyst. Likewise, very little activity was observed when the porphyrin ligand was utilized without tin metalation. A turnover frequency (TOF) of 5300 s^{-1} was measured which is comparable to other p-block metal HER electrocatalysts.[389,398] A Faradaic efficiency of up to 94% was observed. UV/Vis spectroscopy showed minor changes in the intensity and location of the Soret and Q-bands associated with the porphyrin ligand during the reaction.



H2PEGP: M = H,H ; **SnPEGP:** M = Sn^{4+}, 2 Cl^{-}

Figure 4.7 Structure of the **SnPEGP** catalyst.

Computational Details

Theoretical calculations were performed by following a similar approach as in our previous study.[363] We utilized the Orca 5.0 software package[111] with the B3LYP density functional.[236,241] All structures were optimized using an unrestricted wave function. The def2-TZVPP basis set was used for Sn, def2-TZVP basis set for the N atoms of the porphyrin ring and the H atoms involved in H_2 production, and all other atoms were assigned the def2-SV(P) basis set.[275] The resolution of identity (RI) approximation[248] was applied to accelerate the computation of the four index integrals with the Def2/J auxiliary basis sets.[380] Analytical frequency calculations were performed to ensure the minima of the potential energy surface were found for the reactants, products, and reaction intermediates, as well as to characterize transition states by locating a single imaginary frequency along the reaction coordinate. The two lowest spin states were considered for all intermediates: a singlet spin state for all structures with an even number of electrons and a doublet spin state for the reduced structures with an odd number of electrons. Single-point calculations with the conductor-like polarizable continuum model (CPCM)[258] implicit solvation model were performed for all intermediates and transitions states using acetonitrile as the solvent. To properly account for the chemical potential of the H^+/e^- pairs, half of the total energy of H_2 was added following the addition of each pair.[387] For the individual reduction steps, a value of 4.19 eV was added as the Ag/AgCl electrode used during the CV experiments shows a E^0 value 0.25 eV higher than the standard hydrogen electrode (SHE) in water (4.44 ev).[388] Likewise, to properly account for the energy of H^+, a value of 11.80 eV was added at each protonation step. This value was chosen as the total energy of H_2 at the current level of theory equates to -31.99 eV. The two e^- would contribute 8.38 eV of this total energy which requires then that each additional H^+ would contribute 11.80 eV to the total energy of H_2. TDDFT calculations were performed using the CAM-B3LYP[246] density functional with the Chain of Spheres RI methodology mentioned above. 50 roots were requested for the calculation and a maximum dimension of five was used for the Davidson expansion. The same basis sets were utilized as above, and the geometries optimized at the B3LYP level of theory were used without modification.

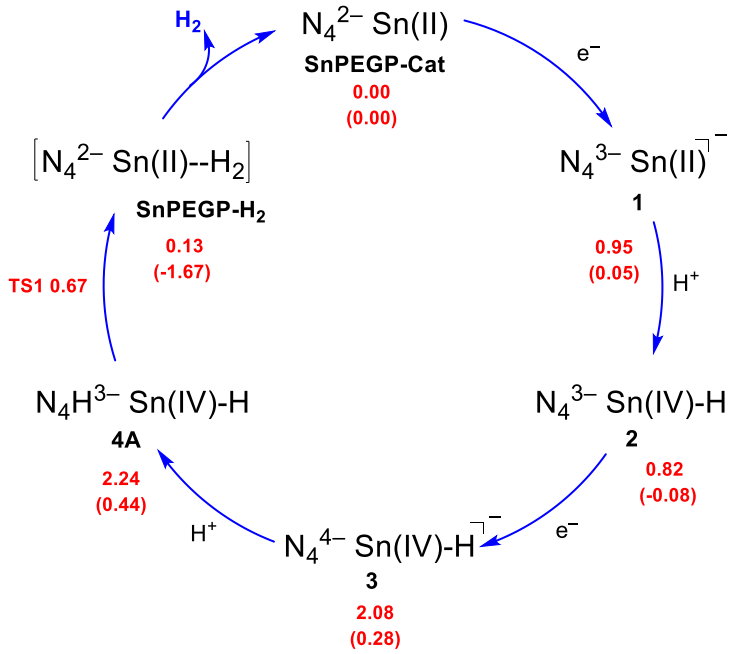
Mechanistic Insights

DFT calculations were performed in order to better understand the HER mechanism catalyzed by **SnPEGP**. We considered the interactions of the reactants (H^+ and e^-) both with the Sn center as well as the porphyrin ligand. Although HER traditionally occurs via metal-centered pathways, redox non-innocent ligands can also participate in the reaction.[395] In such a synergistic system, HER can occur via ligand-centered reactivity, where the H^+/e^- transfer steps are facilitated only on the ligand, or metal-assisted ligand-centered reactivity, where the metal center is more directly involved in the reaction.[399] Examples of such systems include Berben's aluminum(III) NNN type pincer[393] and Grapperhaus's copper(II) thisemicarbazone system which uses a NNSS-type ligand to perform HER.[396] Thus, an examination of different reaction pathways is necessary in order to best understand the HER catalyzed using **SnPEGP**. The proposed reaction mechanism is shown in Figure 4.7A and its corresponding reaction profile in Figure 4.7B.

The two-electron/two-proton transfer steps for HER can occur in several pathways depending on the the order of the chemical steps (C, protonation) and electron transfer steps (E): CCEE, CECE, CEEC, EECC, ECEC, and ECCE.[394] The first step of the reaction involves a reduction of the system or the protonation of an Sn or N atom. Because of the lower symmetry due to the PEG unit, there are two unique N atoms that can be protonated. The free energy difference (ΔG^{298}) of the reduced **SnPEGP** catalyst (intemediate **1** in Figure 4.8A) was found to be 0.95 eV (0.05 with the addition of an applied overpotential)[397] which is lower in energy than any of the protonated complexes (0.20 – 0.66 eV). Thus an initial reduction step is favored. Following this, it was found that protonation was most likely to occur as compared to a second reduction. The second reduction was found to have a free energy difference of 2.08 eV (1.18 eV with the applied overpotential) which is higher than that of any of the protonated complexes. The protonation of the Sn atom (intemediate **2**) was the only step which was found to be exergonic by –0.13 eV while the protonation of the N atoms was found to be higher in energy (0.21 and 0.23 eV).

Following the formation of **2**, either protonation or reduction can occur on the **SnPEGP** surface. It was found that a second protonation at either of the avaiable N atoms proceeded with free energy differences of ~0.90 eV while a second reduction (intermediate **3**) proceeded with a free energy difference of 1.26 eV which is reduced to 0.36 eV with the applied overpotential. Thus,

A



B

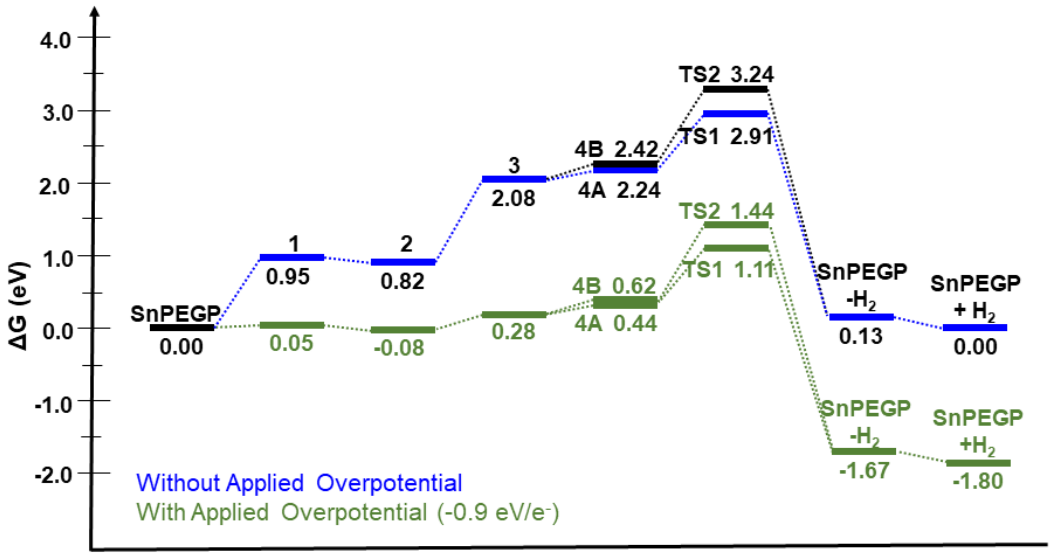


Figure 4.8 (A) Proposed mechanistic cycle for the HER over **SnPEGP** via the ECEC mechanism. (B) Free energy profile for the HER over **SnPEGP**. Values calculated at the B3LYP level of theory

the third step of the reaction most likely involves a reduction which means that HER on the **SnPEGP** catalyst favorably occurs in an ECEC pathway whereby reduction is followed by protonation which is then repeated. The final step of the reaction involves the protonation of either of the N atoms which proceeds with similar free energy differences. The intermediate in which protonation has occured on the N atom closest to the PEG unit (intermediate **4A**) shows a free energy difference of 0.16 eV while the protontion of the other N atom (**4B**) proceeds with a free energy difference of 0.34 eV. Because of this small energy difference, the final protonation step could occur at either atom. Intermediate **4B** is the only intermediate discussed thus far which resulted in a lower energy for the triplet spin state as compared to the singlet spin state.

Following the two-proton/two-electron transfer steps, we investigated the kinetics associated with the bond breaking steps. Due to the small energy difference observed for the final protonation step, the formation of H_2 could occur on two different surfaces. The kinetic barrier associated with the breaking of the N-H and Sn-H bonds to form H_2 was found to be 0.67 eV (15.4 kcal/mol) for **4A** and 0.82 eV (18.9 kcal/mol) for **4B**. These barriers are similar in magnitude but because the former is smaller, we believe that HER most likely proceeds following the formation of intermediate **4A.** Following this, we found that H_2 is only weakly bound (0.13 eV) to the **SnPEGP** catalyst which allows for catalyst regeneration.

4.3.5 Conclusions

The evaluation of two molecular electrocatalysts and their activity towards the electrocatalytic production of molecular hydrogen (H_2) was studied using computational methods. Both molecular catalysts, once heterogenized, showed activity towards the hydrogen evolution reaction (HER). The antimony N,N'-ethylenebis(salicylimine) (**SbSalen**) catalyst showed a Faradaic efficiency of 100% while the pegylated, tin porphyrin (**SnPEGP**) catalyst showed a Faradaic efficiency of 94%. It was found that both catalysts proceed through similar mechanisms to produce an initial reduction followed a protonation which is then repeated. HER on the **SbSalen** catalyst proceeds through a mostly ligand-centered pathway where the role of the Sb atom occurs later in the reaction mechanism in which it acts as an anchor point to produce a metal-hydride

which can then couple with the nearby proton to produce H_2. Conversely, the SnPEGP catalyst produces H_2 mechanism which directly involves the protonation of the Sn site to ultimately yield a metal-hydride.

4.4 Computational Catalysis Utilizing Metal-Organic Frameworks

Disclosure

The following section includes minor modifications from work which has been taken with permission from the following three manuscripts:

(1) A review article on computational catalysis for metal-organic frameworks co-authored by Gavin McCarver, Dr. Rajeshkumar Thayalan, and Dr. Konstantinos D. Vogiatzis, and published in the *Coordination Chemistry Reviews* journal.[60]

(2) An article published in *The Journal of Physical Chemistry Letters.*[400] This work has been authored by Dr. Konstantinos D. Vogiatzis, Dr. Rajeshkumar Thayalan, and myself in collaboration with Kamolrat Metavarayuth, Otega Ejegbavwo, Michael L. Myrick, Thomas M. Makris, Olivia M. Manley, Juan D. Jimeniz, Kexun Chen, Natalia B. Shustova and Donna A. Chen (University of South Carolina), Sanjaya D. Senanayake, Anatoly I. Frenkel, and Sooyeon Hwang (Brookhaven National Laboratory), and Amani M. Ebrahim (Stony Brook University). All experimental experimental work was performed by the Shustova and Chen groups while all computational details and results are the work of myself, Dr. Konstantinos D. Vogiatzis, and Dr. Thayalan Rajeshkumar.

(3) A manuscript that has been submitted for publication. This work has been authored by Dr. Konstantinos D. Vogiatzis, Dr. Rajeshkumar Thayalan, and myself in collaboration with Donna A. Chen, Julian P. Stetzler, Deependra M. Shakya, Abhijai Mathur, Natalia M. Shustova, Michael L. Myrick, Kamolrat Metavarayuth, Michael Royko, Jochen Lauterbach, Gregory Tate, John R. Monnier (University of South Carolina), and Juan D. Jimenez and Sanjaya Senanayake (Brookhaven National Laboratory).

4.4.1 Abstract

Metal-organic frameworks (MOFs), a family of porous hybrid organic/inorganic materials, have shown great promise for many challenging chemical applications including gas separations, catalysis, and sensors. Numerous recent articles explore the field of catalytically active MOFs at both an experimental and computational level. The stability, porosity, and periodic nature of MOFs allow them to act as both the active catalyst as well as supporting material of single atoms or small metal clusters deposited on their surface. The HKUST-1 MOF, composed of Cu(II) metal cations and benzene-tricarboxylate organic linkers, was examined as a possible hydrogenation catalyst following partial transmetalation of the Cu(II) sites. Computational studies showed that this complex, when transmetalated with Rh(II) to form $(Cu_xRh_{1-x})_3BTC_2$ (abbreviated as CuRhBTC, $0 \leq x \leq 1$), was very active towards the hydrogenation of propene. Further studies were performed to gauge the ability of this complex towards hydroformylation as well the hydrogenation of other small hydrocarbons such as ethyne and ethene. Numerous spectroscopic studies were also performed to further our understanding of the CuRhBTC MOF.

4.4.2 Introduction

Metal-organic frameworks (MOFs) are a class of one-, two-, or three-dimensional coordination polymers composed of inorganic nodes (single metal atoms or metal clusters) connected via organic linkers for the formation of open networks that contain potential voids.[401] Solvent removal upon heating typically leads to (near-permanent) porosity and large surface area. A large variety of inorganic nodes and/or organic linkers can be used for their synthesis, which has led to more than 14,000 known structures.[402] In addition to the ability to synthesize a large number of MOFs, there also exists a multitude of alternative synthetic pathways (such as microwave heating or mechanochemistry) and final reaction products (such as nanoparticles, thin films, or membranes) which allow for MOFs to be used in many areas of chemistry. Due to their unique properties, MOFs with significant stability have been suggested for many chemical applications such as separations,[403] sensors,[404] energy storage,[405] drug delivery,[406] and catalysis.[407]

Numerous studies have demonstrated the applicability of MOFs as catalysts which capitalize on intraframework exposed metal sites, usually formed upon solvent removal and activation of the material, or extraframework sites (single atoms or metal clusters) deposited in the MOF.[408,409] The topologies of MOFs allow them to bridge the gap as both heterogeneous catalysts (by being in the solid state and interacting with gaseous/liquid reactants and products) and homogeneous catalysts (having a uniform and predictable active site). In addition, they have also shown to be capable of acting as either a basic catalyst or acidic catalyst, allowing them to be used in any number of catalytic cycles.[410,411] Lastly, due to their unique abilities, MOFs can be involved in oxidation, reduction, electrocatalytic, and photocatalytic reactions. Reactions such as the hydrogenation of alkenes, the electrooxidation of ethanol, water-splitting reactions, and C-H functionalization have all been performed using MOFs both from an experimental and a computational standpoint.[60]

Computational methods, in conjunction with experimental techniques, have been extensively used for the characterization and the elucidation of the properties of MOFs. In our recent review, computational studies on catalysis with MOFs were presented.[60] A detailed report of all the numerous experimental studies as well those which focus on the adsorption of small molecules using computational methods was beyond the scope of this article. Our aim was to demonstrate how computational chemistry has been used recently to obatain a deeper understanding of reaction mechanisms, the examination of competitive reaction channels, and the elucidation of the electronic structure of key reaction intermediates and transition states. Extensive and thorough examination of many MOFs has the potential to provide structure-function relations. Such relations can be used by either synthetic chemists for the targeted synthesis of new materials, or by computational chemists for the optimization of specific properties by high-throughput computational screening studies, which can lead to the *in-silico* discovery of new functional MOFs. Herein is a detailed examination of the HKUST-1 MOF, which was examined for hydrogenation catalysis.

4.4.3 Spectroscopic Studies for the Direct Identification of Mixed-Metal Centers in Metal-Organic Frameworks

Overview and Experimental Results

The unique nature of the HKUST-1 MOF allows for the close proximity of two metal centers not usually seen in other complexes. The interaction between these centers is often small, showing very little if any bonding character between them. The traditional HKUST-1 MOF is composed entirely of Cu(II) centers but through transmetalation, many of these centers may be replaced with other transition metals. To positively identify these transmetalated sites, different spectroscopic techniques may be used including Raman spectroscopy and electron paramagnetic resonance (EPR).

Upon transmetalation with Rh(II), the HKUST-1 MOF is chemically altered to form $(Cu_xRh_{1-x})_3BTC_2$ (abbreviated CuRhBTC, $0 \leq x \leq 1$) which was shown to be catalytically active towards the hydrogenation of propene (*vida infra*). Raman spectroscopy experiments were performed using the CuBTC, RhBTC, and CuRhBTC MOFs to identify the respective metal-metal stretches of the materials. The ν(Cu-Cu) stretch was identified at 227 cm^{-1} and showed high intensity relative to the other materials which suggests strong resonance enhancement while the ν(Rh-Rh) stretch was found at 331 cm^{-1} and the ν(Cu-Rh) stretch was found at 285 cm^{-1}. All three peaks are identified in the CuRhBTC material while only the Cu-Cu and Rh-Rh stretches are found in the CuBTC and RhBTC materials, respectively. This indicates the inclusion of Rh(II) sites into the CuBTC material. Likewise, EPR experiments were performed which identify a broad isotropic signal centered at g = 2.17 for CuBTC which changes signal location to g = 2.13 for the transmetalated, CuRhBTC MOF which indicates heteronuclear Cu(II) sites due to substitution by Rh(II).

Computational Details

Calculations were performed with the ORCA 4.2 software package.[118] The PBE[237] and B3LYP[232,233,236] density functionals were utilized as they have been shown to be reliable for such

spectroscopic studies. For the transition metals, the def2-TZVPP basis set was used (with the corresponding effective core potential for Rh), while the def2-TZVP basis set was used for all other atoms.[221] Tight convergence criteria were employed for both the SCF cycles and geometry optimization steps. Grimme's D3 semiempirical correction[276] was included with the Becke-Johnson damping function.[277,412,413] The resolution of identity approximation was used to accelerate the computation of the four index integrals.[248] All of the Raman spectra were computed numerically with the harmonic approximation, and each mode was corrected for by using an empirical scaling factor of 1.0306 for values calculated using PBE and 1.0044 for values calculated using B3LYP.[1] Two truncated molecular models were used that consisted of a dimer paddlewheel structure capped with the benzenetricarboxylate linker where the acidic hydrogens were replaced with Li^+ cations (Figure 4.9A) and a further truncated model capped with acetate ligands (Figure 4.9B). These models are referred to as *large* and *small*, respectively. The carboxylate groups of the *large* model were terminated with Li^+ cations following previous work which showed that such a system gave a better representation of the cationic, inorganic building blocks of certain MOFs.[414–419] The ground spin states of the nodes considered in this study are a closed-shell singlet for the RhBTC complex (based upon previous experimental work),[420] a triplet for CuRhBTC (based upon previous work in our group)[294] while an antiferromagnetically coupled singlet for CuBTC was utilized based upon previous experimental work.[323,421] TDDFT calculations were performed to elucidate the electronic structure of the CuBTC node in the Raman experiments. The CAM-B3LYP[246] density functional and the smaller molecular model consisting of the paddlewheel node capped with acetate groups (Figure 4.9B) were used to calculate the excited states of the CuBTC structure around 18,800 cm^{-1}, which corresponds to the excitation wavelength of 532 nm used during the Raman experiments.

Results and Discussions

Calculations were performed to simulate the Raman spectra of the CuBTC, CuRhBTC, and RhBTC MOFs. The predicted metal-metal stretches for the different MOFs using different levels of theory are displayed in Table 4.1, below. The predicted *ν*(Cu-Rh) mode using the PBE density functional and the larger BTC model is consistent with the experimental stretch. Likewise, the

A **B**

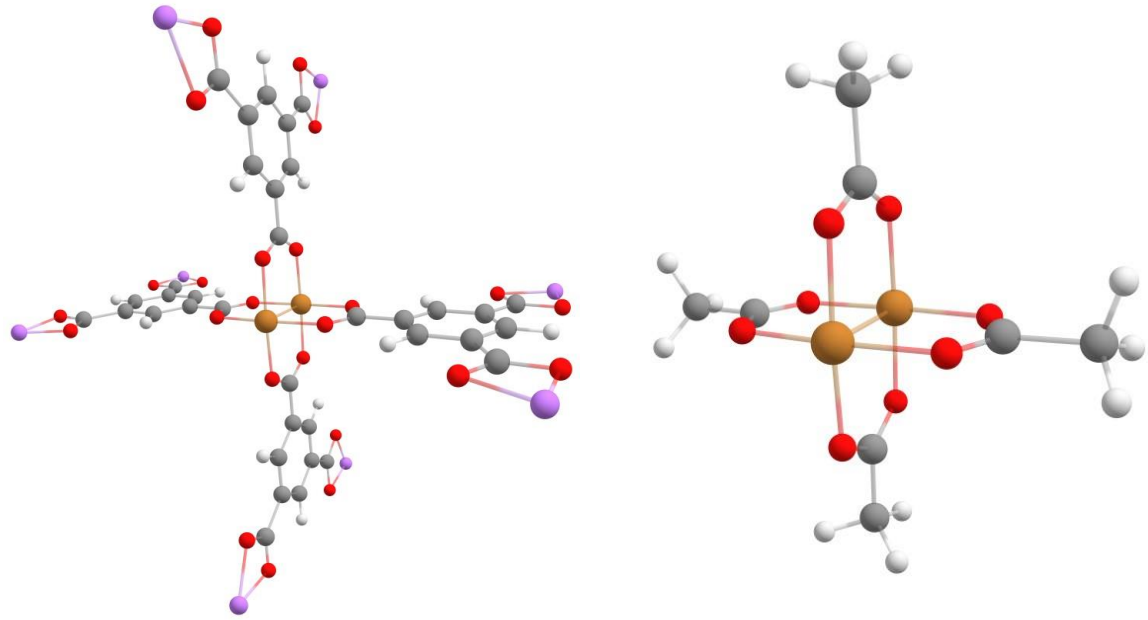


Figure 4.9 Dinuclear paddle-wheel molecular models with: A) benzene tricarboxylate ligands and B) acetate groups. Color code: Cu - gold, C - black, O - red, H - white, Li - purple.

Table 4.1 Metal-Metal stretches (in cm^{-1}) for the CuBTC, CuRhBTC, and RhBTC complexes for each of the three levels of theory. Each value has been scaled according to the above-mentioned scaling factors. Here, *Small* refers to the model shown in Figure 4.9B and *Large* refers to the model shown in Figure 4.9A. Each value is scaled by a factor of 1.0306 for PBE based results and 1.0044 for B3LYP. [1]

	PBE-Small	PBE-Large	B3LYP-Small	Experimental
CuBTC	232.6	241.8	226.7	227
CuRhBTC	263.9	300.2	225.3	285
RhBTC	363.2	357.0	360.1	331

ν(Cu-Cu) and *ν*(Rh-Rh) modes occur within 15-26 cm^{-1} of the experimental values at 227 and 331 cm^{-1}, respectively. These values are consistent with the stretching frequencies reported in the literature for CuBTC[422] and RhBTC.[423] Furthermore, TDDFT calculations predict that the high intensity of the *ν*(Cu−Cu) stretch relative to the *ν*(Cu−Rh) and *ν*(Rh−Rh) stretches in the Raman spectra may be attributed to resonance enhancement. Four d to σ* excitations were identified for the antiferromagnetically coupled singlet of the CuBTC node with energies corresponding to the 532 nm excitation wavelength whereas no excitations were observed close to the same energy for the CuRhBTC and RhBTC nodes.

Additional calculations were performed to examine the interaction of small molecules with the three different MOF models and how the metal-metal stretching modes change upon adsorption. The first of which involved examining how the *ν*(Rh-Rh) mode changes upon the binding of either hydrogen, propene, or following the formation of a Rh-H species. These results are shown in Table 4.2, below. Upon the addition of one H_2 molecule, the *ν*(Rh-Rh) stretching frequency is redshifted by 20-40 cm^{-1} depending on the chosen model. Likewise, the effect of a second H_2 molecule further redshifts the stretching mode by an additional 5-15 cm^{-1}. The *ν*(Rh-Rh) values are redshifted by 50 and 90 cm^{-1} for one and two propene molecules, respectively, for the small model and 30-60 cm^{-1} for the large model. Experimental Raman spectra showed a *ν*(Rh-Rh) peak that has been redshifted by 20-30 cm^{-1} and based upon these computational results, this shift may be due to the presence of H_2 bound to the RhBTC MOF.

Likewise, experimental Raman spectra for the CuBTC MOF during the catalytic cycle showed redshifted values for the *ν*(Cu-Cu) mode by ~20 cm^{-1}. Upon the adsorption of either one or two H_2 molecules, a shift of 10-12 cm^{-1} was predicted which does not correspond to the same magnitude in the observed, experimental shift (Table 4.3). When propene is adsorbed to the CuBTC MOF structure, larger shifts are predicted. Values for the *ν*(Cu-Cu) stretching mode are 198.3 and 213.9 cm^{-1} for the adsorption of one and two propene molecules, respectively, which more closely correspond to the ~20 cm^{-1} shift observed experimentally.

Final calculations were performed on the CuRhBTC MOF model to see if the adsorption of H_2 or propene causes new peaks in the Raman spectra to appear in the 180-200 cm^{-1} region. It was found that the H_2 molecule preferentially adsorbed to the Rh site but no new peaks appeared

Table 4.2 Effects of adsorbed species on the *v*(Rh-Rh) mode. Here, *Small* refers to the model shown in Figure 4.3.1B and *Large* refers to the model shown in Figure 4.3.1A. Each value is scaled by a factor of 1.0306.[1] No *v*(Rh-Rh) was observed for the RhBTC-H species using the small model.

Structure	*v*(Rh-Rh) (cm^{-1})	
	Small Model	Large Model
RhBTC	363.2	357.0
RhBTC-H_2	320.2	339.5
RhBTC-$(H_2)_2$	315.2	326.1
RhBTC-C_3H_6	316.7	325.9
RhBTC-$(C_3H_6)_2$	273.3	295.0
RhBTC-H	-	297.6

Table 4.3 Effects of adsorbed species on the *v*(Cu-Cu) mode. Here, *Small* refers to the model shown in Figure 4.3.1B and *Large* refers to the model shown in Figure 4.3.1A. Each value is scaled by a factor of 1.0306.[1]

Structure	*v*(Cu-Cu) (cm^{-1})	
	Small Model	Large Model
CuBTC	232.6	241.8
CuBTC-H_2	225.3	235.5
CuBTC-$(H_2)_2$	219.3	229.2
CuBTC-C_3H_6	213.9	-
CuBTC-$(C_3H_6)_2$	198.3	-

between 180-200 cm^{-1} in the simulated Raman spectra for either the small or large model. When a single propene molecule is bound to the Rh site, two new peaks are observed at 183.0 and 191.5 cm^{-1} for the small model and a single peak is observed at 185 cm^{-1} for the larger model. These vibrational stretches correspond primarily to methyl rotations on the adsorbed propene molecule which may be the cause of the new peaks observed experimentally.

4.4.4 Small Molecule Interactions with the $(Cu_xRh_{1-x})_3(BTC)_2$ MOF

Overview and Experimental Results

The CuRhBTC MOF was recently shown to be an excellent catalyst for the hydrogenation of propene to propane.[294] Upon further testing, however, it was found that the introduction of ethyne to the system led to the deactivation of the MOF. This was observed experimentally through a color change and a lack of catalytic activity. It was thought that this deactivation may arise through a stronger interaction between the CuRhBTC MOF and ethyne as compared to propene. In order to help verify or disprove this hypothesis, computational methods may be applied in order to calculate the binding energy between the CuRhBTC MOF and propene, ethyne, and hydrogen as well as ethene. Such a study should be performed using high levels of electronic structure theory such as the coupled cluster (CC) or the Complete Active Space Self-Consistent Field (CASSCF) method. Presented herein is a study which examines the interaction of the four previously mentioned small molecules on the CuRhBTC MOF surface using the Domain-based Local Pair Natural Orbital (DLPNO-CCSD-(T)) method. In addition, two different kinds of basis sets are examined where one includes additional, diffuse basis functions. This study is performed as a first step in improving our understanding of the interactions occurring on the atomistic scale of the CuRhBTC MOF.

Computational Details

A single node of the CuRhBTC MOF (Figure 4.9) was utilized for all calculations which were performed with the ORCA 5.0 software package.[111] The model complex (A) was truncated

with acetate groups (B) in order to make the calculations more tractable. Geometry optimizations were performed using the PBE0[242] density functional with the def2-TZVPP basis set for the Cu and Rh atoms (with the corresponding def2-ECP for Rh) while the def2-TZVP basis set was used for all other atoms.[275] Tight convergence criteria were used for both the SCF iterations and the geometry optimization steps. Analytical frequency calculations were performed in order to ensure that minimum of the potential energy surface is located. The resolution of identity (RI) approximation using the Chain of Spheres method (RIJ-COSX) was used to accelerate the computation of all four-index integrals.[379] The def2/J auxiliary basis set was used for all atoms.[380] A triplet spin state was assumed for all calculations based upon previous theoretical benchmarking which showed the triplet spin state was the most thermodynamically favorable.[294] Dispersion effects were accounted for by using Grimme's D3 semiempirical method[276] with the Beck-Johnson damping function during the geometry optimization steps.[277] Using the optimized geometries, single point calculations were performed using the DLPNO-CCSD-(T) method, with tight PNO settings. Perturbative triples were accounted for using standard, non-DLPNO methods. These calculations were performed both with and without the addition of diffuse basis functions for each atom using the def2-TZVPPD set for Cu and Rh atoms and the def2-TZVPD set for all other atoms.

Results and Discussions

The respective binding energies of hydrogen, ethyne, ethene, and propene on the Rh(II) site of the CuRhBTC MOF were found to be 2.60, 9.80, 10.94, and 13.15 kcal/mol using non-diffuse basis sets and 2.81, 10.58, 11.85, and 14.25 kcal/mol using the diffuse basis sets. In contrast to the expected results, both sets of calculations indicate that a stronger interaction would be observed between the CuRhBTC MOF and propene as compared to ethyne. Thus, the experimental effects seen upon the addition of ethyne to the MOF is more complex than the interaction of the π-orbitals of ethyne with the d-orbitals of the Rh(II) center.

The change in color and subsequent loss of activity observed experimentally following the addition of ethyne to the CuRhBTC MOF may be due to poisoning. Additional calculations were performed on species in which either a $-C{=}CH_2$ or $-C{\equiv}CH$ functional group was attached to the

Rh(II) site as these species may be similar to those observed during coking but these species were higher in energy than the π-bound system and thus, would not be favorable to form. Previous experimental work indicated that Rh-ethyne interactions may lead to polymerization[424] or to the formation of carbon species deposits[425] which may then lead to poisoning of the metal complex. Such effects would be difficult to examine from a theoretical perspective, however, and thus were not tested. Based upon these results, the deactivation of the CuRhBTC MOF upon the addition of ethyne is likely due to more than the interaction of the small hydrocarbon with the MOF itself. As a further examination of this system, we have tested the hydrogenation of ethyne, ethene, and propene using the CuRhBTC MOF which is discussed in the following section.

4.4.5 Hydrogenation Reactions utilizing the $(Cu_xRh_{1-x})_3(BTC)_2$ MOF

Overview and Previous Results

The ability to catalytically hydrogenate and dehydrogenate chemical systems is a fundamental process in organic synthesis to produce both fine and bulk chemicals. The hydrogenation reaction requires a catalyst that can activate and insert H_2 into an unsaturated bond while dehydrogenation often requires an oxidizing agent to form the unsaturated bond. Ru, Rh, and Ir complexes as well as other transition metals that have been deposited on inert, bulky material such as Ru on activated carbon or Pt on alumina are active catalysts for hydrogenation reactions.[294,302] These catalysts utilize expensive transition metals which raise the cost of hydrogenation and finding new catalysts based on more earth-abundant elements is an active area of research. This is showcased in dehydrogenation studies which widely examine the use of first row transition metal species to facilitate the production of unsaturated products. Recent studies have examined MOFs for the hydrogenation of alkenes (and alkynes) as the MOFs themselves offer the ability to fine tune the activity of reactions as well as offer a support system for other active sites.[299–301] Three MOFs in particular (HKUST-1, NU-1000, and UiO-67) have been shown to catalytically add H_2 to alkenes and alkynes for the production of their saturated analogues.

Previous experimental work by Shakya *et al.* showed that a transmetalated form of the HKUST-1 MOF is an active catalyst towards the hydrogenation of propene to form propane.[294]

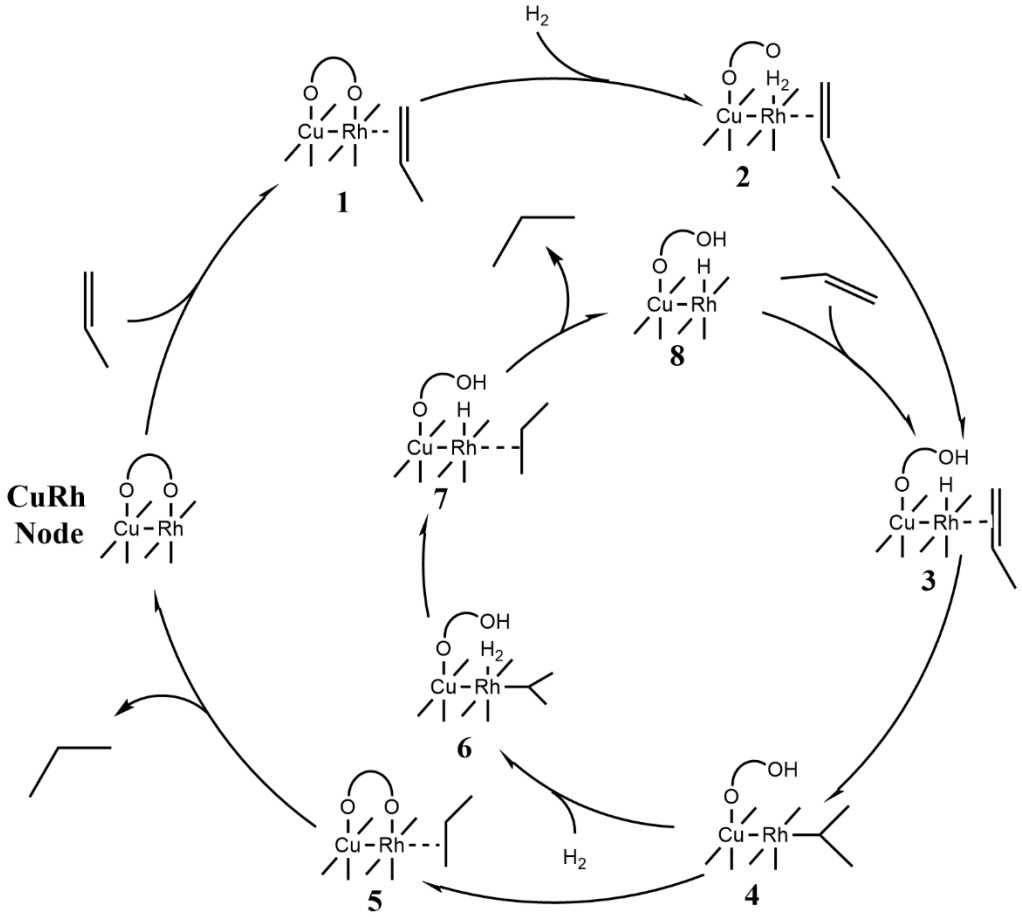


Figure 4.10 Schematic outline of the reaction pathway for the hydrogenation of propene using the CuRhBTC MOF.

They were able to replace different amounts of the Cu(II) sites of HKUST-1 to form different CuMBTC (M = Cu, Rh, Ru, Co, Ir, Ni) complexes. These complexes were then tested for their ability to catalytically add H_2 to propene and it was found that CuRhBTC was the most catalytically active combination of those tested. DFT calculations were then performed which led to a catalytic mechanism that showed hydrogenation may occur in the following steps: (i) adsorption of propene (intermediate 1 in Figure 4.10, **Int** 1), (ii) simultaneous adsorption of H_2 and the decoordination of one of the oxygen atoms of the acetate ligand to the Rh(II) site, (iii) H_2 cleavage to yield a Rh-H moiety and the protonation of the decoordinated oxygen of one of the acetate ligands, (iv) protonation of the 1-position of propene, (v) protonation of the 2-position of propene, and (vi) desorption. The theoretical activation barrier (15.4 kcal/mol) and the experimental barrier (6.3 kcal/mol) disagreed and so an alternative pathway was hypothesized which involved an outer mechanism composed of the previously mentioned steps and a new, inner mechanism. The inner mechanism begins with the cleavage of H_2 and the first protonation of the propene (**Int** 4). From here, it differs from the outer mechanism as the acetate ligand remains protonated throughout the rest of the mechanism. Thus, a second H_2 molecule is required which binds to the Rh(II) site following the initial protonation of propene (**Int** 6). This H_2 is then cleaved to fully protonate propene (**Int** 7) and leads to the formation of a Rh-H bond. Previous work in our group showed similar binding energies of propylene on the Rh(II) site of both structures in Figure 4.9 and so the truncated structure in Figure 4.9A was used throughout the reaction mechanisms. An extension of this work is presented here where the hydrogenation of propene, ethyne, and ethene is examined using the CuRhBTC MOF.

Computational Details

A single node of the CuRhBTC MOF (Figure 4.9) was utilized for all calculations which were performed with the ORCA 5.0 software package.[111] The model complex (A) was truncated with acetate groups (B) for all calculations. Geometry optimizations were performed using the M06L[426] density functional with the def2-TZVPP basis set for the Cu and Rh atoms (with the corresponding def2-ECP for Rh) while the def2-TZVP basis set was used for all other atoms.[275] Tight convergence criteria were used for both the SCF iterations and the geometry optimization

steps. The RI approximation was used to accelerate the computation of all four-index integrals.[248] The def2/J auxiliary basis set was used for all atoms.[380] A triplet spin state was assumed for all calculations.[294] Numerical frequency calculations were performed to ensure that minima of the potential energy surface for all structures were obtained as well as to characterize transition states by locating a single imaginary frequency along the reaction coordinate.

Results and Discussions

Tables A4.1 – A4.2 include the reaction energies, enthalpies, and key bond distances for the three hydrogenation reactions: propene, ethene, and ethyne, respectively. The reaction profiles for propene, ethene, and ethyne hydrogenation are shown in Figures 4.11 - 4.13, respectively. The reaction mechanism begins with the adsorption of the unsaturated hydrocarbon as the binding of H_2 was found to be less favorable. Propene, ethene, ethyne, and H_2 were found to bind with enthalpies of -14.7, -14.1, -10.4, and -2.0 kcal/mol, respectively. The second step was found to involve the insertion of H_2 which requires an open site on the Rh(II) atom. This is achieved through the breaking of a Rh-O bond (**TS1**) forming a detached oxygen from one of the acetate ligands. The reaction barrier for this step varies given the choice of unsaturated hydrocarbon but ranges from 12.6 kcal/mol for ethyne to 16.4 kcal/mol for propene. The H_2 molecule shows no activation during TS1 ($R_{H\text{-}H}$ = 0.760 - 0.769 Å) but the distance between it and the Rh(II) varies from 2.068 Å with ethyne, 2.210 Å with propene, and 2.393 Å with ethene. This trend may be due to either the electronic structure of the double or triple bond or may be due to steric effects from the different hydrocarbons. **TS1** is also accompanied by a significant elongation in the Cu-Rh atomic distance by ~10.5% (0.256 Å).

Intermediate **2** results in slight activation of the H_2 molecule with bond distances of ~0.865 Å for the three cases. Conversely, the Cu-Rh distance decreases slightly following **TS1**. The barrier for **TS2** is much lower (essentially barrierless) for propene (ΔH^{298} = 0.1 kcal/mol), ethene (1.0 kcal/mol), and ethyne (2.6 kcal/mol). During this step, the H_2 bond is cleaved to form a Rh-H bond and to protonate the decoordinated oxygen atom. The distance between the two hydrogen atoms was found to range from 0.962 Å to 0.981 Å for **TS2**. Intermediate **3** shows the complete cleavage

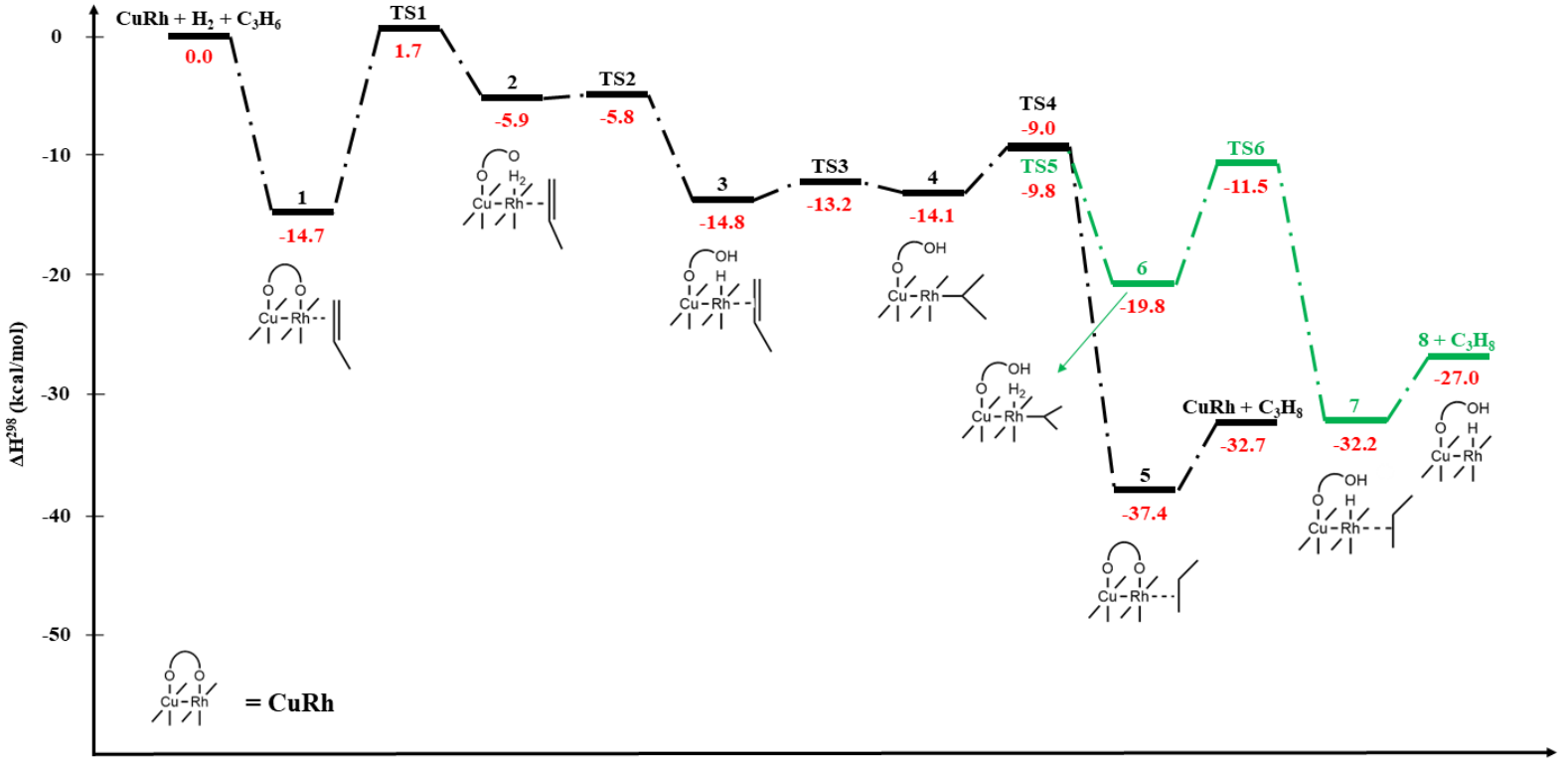


Figure 4.11 Reaction profile for propene hydrogenation catalyzed at the bimetallic CuRh node of CuRhBTC. Values in kcal/mol. Red values indicate energy relative to starting materials in kcal/mol. Green lines indicate the reaction profile of the inner mechanism (see Figure 4.9).

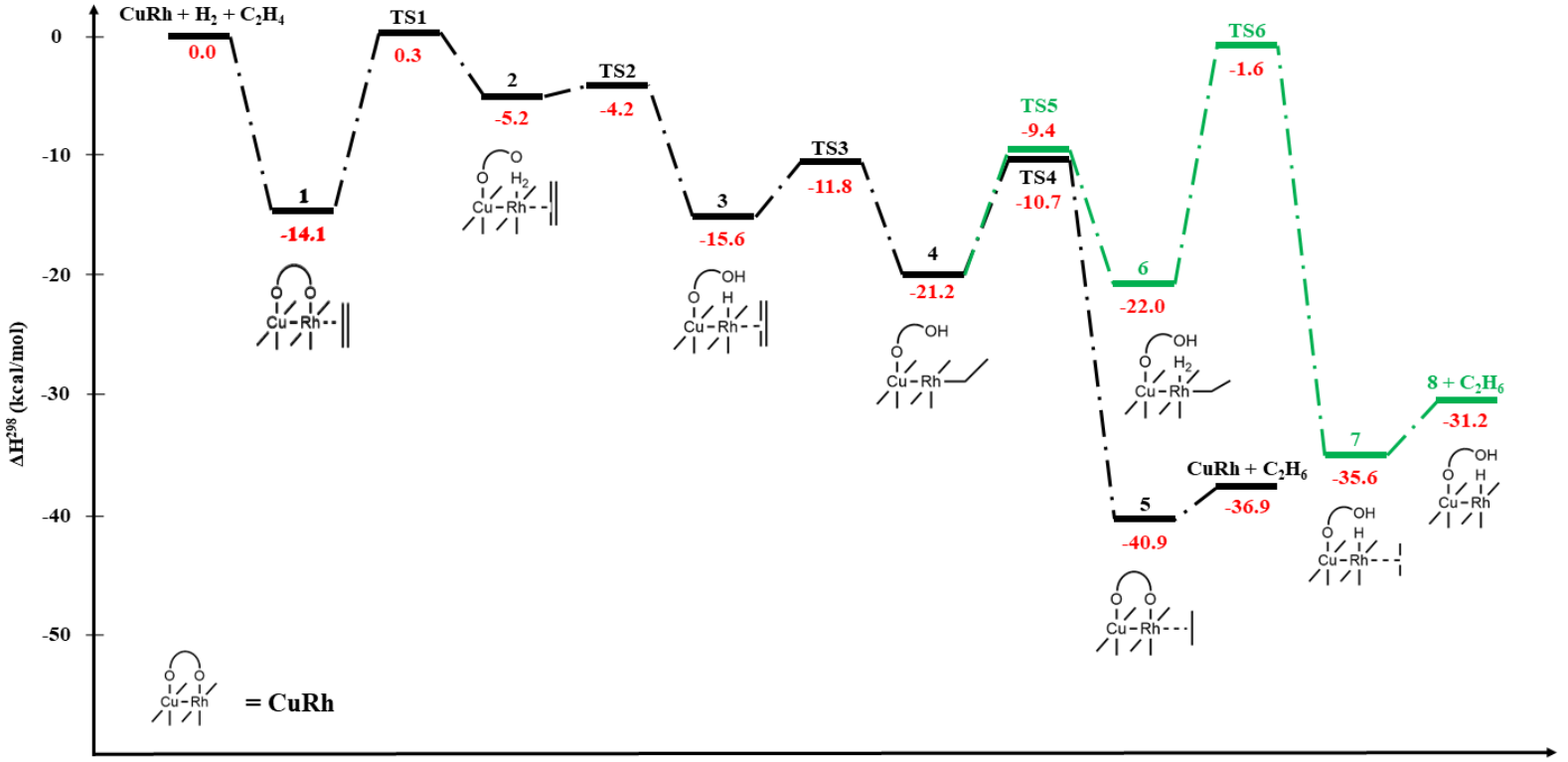


Figure 4.12 Reaction profile for ethene hydrogenation catalyzed at the bimetallic CuRh node of CuRhBTC. Values in kcal/mol. Red values indicate energy relative to starting materials in kcal/mol. Green lines indicate the reaction profile of the inner mechanism (see Figure 4.9).

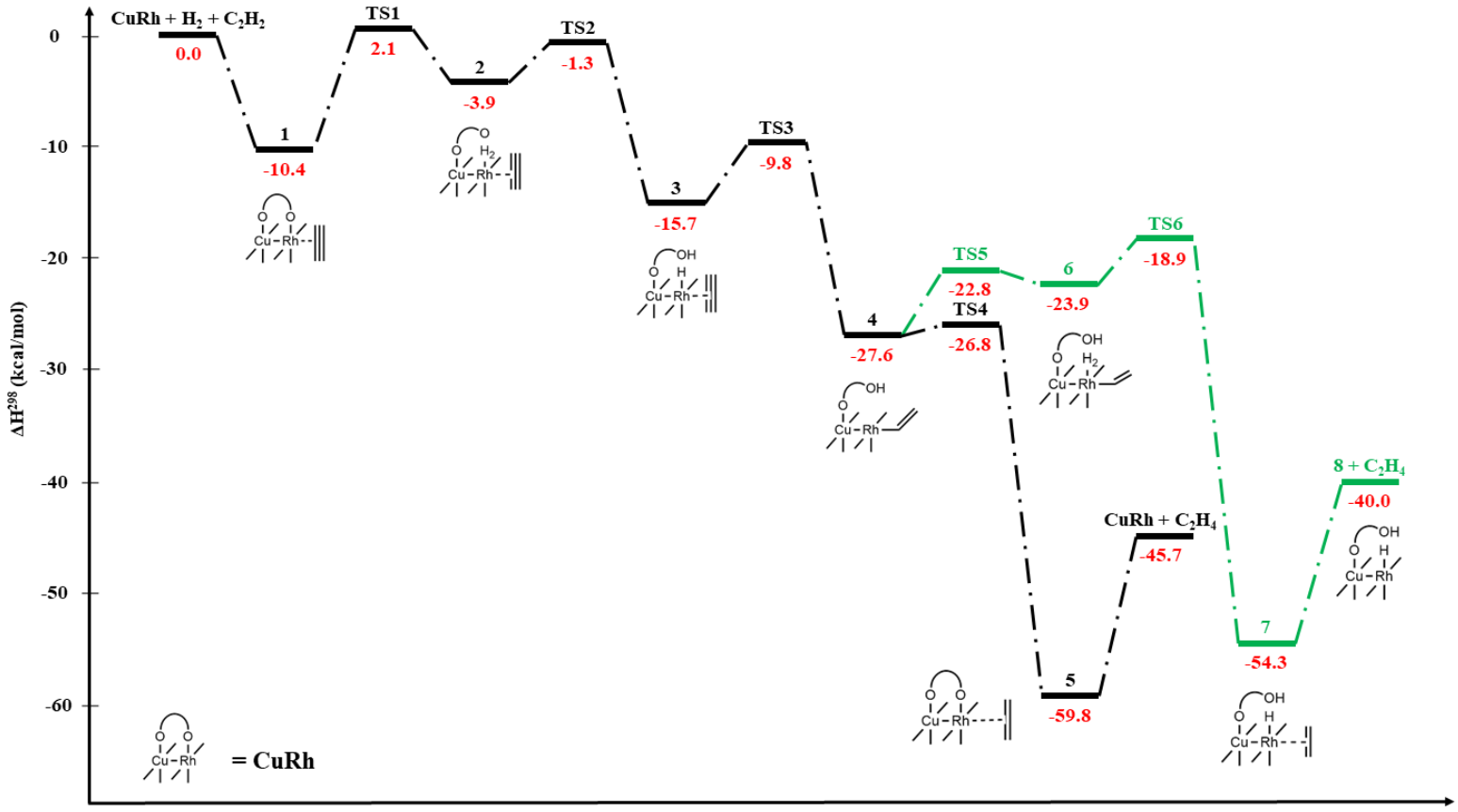


Figure 4.13 Reaction profile for ethyne hydrogenation catalyzed at the bimetallic CuRh node of CuRhBTC. Values in kcal/mol. Red values indicate energy relative to starting materials in kcal/mol. Green lines indicate the reaction profile of the inner mechanism (see Figure 4.9).

of the H_2 molecule with H-H distances of 2.497 Å for propene and 2.724 and 2.833 Å for ethyne and ethene, respectively. This step also results in an even longer Cu-Rh distance which ranges from 2.826 Å for ethyne to 3.101 Å for ethene. **TS3** involves the first protonation of the hydrocarbon which causes a slight elongation of the unsaturated bond which is further elongated at intermediate **4**. The barrier for this transition step is quite small for all three hydrocarbons with a maximum value of 5.9 kcal/mol for ethyne.

Following the formation of intermediate **4**, the inner and outer loop pathways diverge. If the proton found on the detached oxygen of the acetate ligand is allowed to protonate the hydrocarbon directly (**TS4**), then this leads to the formation of the final product and the reformation of the original CuRhBTC catalyst. This step results in activation barriers which show no clear trend in the choice of hydrocarbon. **TS4** with ethyne is essentially barrierless with a ΔH^{298} value of 0.8 kcal/mol while ethene and propene show larger barriers; ΔH^{298} = 10.5 and 5.1 kcal/mol, respectively. Following **TS4**, the formation of intermediate **5** is very exothermic with ΔH^{298} values of -33.0, -30.2, and -28.3 kcal/mol for the formation of ethene, ethane, and propane. Likewise, the still unsaturated ethene requires 14.1 kcal/mol to desorb from the Rh(II) site while the fully saturated hydrocarbons require less than 5.0 kcal/mol for desorption. For the outer loop mechanisms, the largest barriers observed all occur during **TS1** where the H_2 molecule is inserted to the Rh(II) site while a Rh-O bond is simultaneously cleaved.

The inner loop mechanism begins with intermediate **4** and involves the introduction of a second H_2 molecule to the Rh(II) site (**TS5**). This step does not require the cleavage of a Rh-O bond, however, and results in lower ΔH^{298} values than those seen in **TS2**. The H_2 molecule is slightly activated during TS5 with $R_{H\text{-}H}$ distances ranging from 0.754 – 0.853 Å. **TS6** results in varied transition barriers ranging from 5.0 and 8.3 kcal/mol for ethyne and propene to 20.4 kcal/mol for ethene which may be a result of a severe deformation of the structure. The formation of intermediate **7** and the desorption of the more saturated products follows very closely to what was seen for the outer loop where formation is highly favorable and desorption requires very little energy except in the case of ethyne which forms ethene and thus would show a favorable interaction of the π- and d-orbitals.

These results indicate that the hydrogenation of propene, ethyne, and ethene should be theoretically possible using the CuRhBTC MOF. Based upon the activation barriers of the inner loop mechanism, ethyne hydrogenation to form ethene should result in the highest reaction rate but may be hindered if the hydrogenation continues for another cycle using the ethene product. The thermodynamic stability of intermediate **6** using ethene may be the cause of the poisoning of the CuRhBTC MOF upon the introduction of ethyne. This hypothesis may be validated following the introduction of ethene to the catalyst and observing if poisoning is still present. If not, then an alternative hypothesis must be examined to better understand the poisoning of the catalyst in the presence of ethyne.

4.4.6 Hydroformylation on Bimetallic Metal-Organic Frameworks with Paddlewheel-type nodes

Overview and Experimental Results

As the CuRhBTC MOF was shown to be catalytically active towards the hydrogenation of propene, further experimental studies were performed to identify other reactions that may be catalyzed by this material. Initial studies showed that the CuRhBTC MOF may be catalytically active towards hydroformylation but it was found that upon the addition of CO at high temperatures, nearly all of the Rh sites are reduced from Rh(II) to metallic Rh(0). At lower temperatures, however, a smaller percentage of Rh sites are reduced. This effect was not observed using the pure RhBTC MOF, however, where no catalytic activity was observed. It was hypothesized that CO poisons the MOF catalyst in some form, either through deformation of the paddlewheel structure or through a binding interaction that is stronger than small hydrocarbons such as ethene, ethyne, and propene. If CO were to bind too strongly to the Rh active sites, then catalysis could no longer take place and thus the catalyst would be deactivated. To understand how CO may lead to poisoning of the CuRhBTC catalyst and to understand why, upon the addition of CO, metallic Rh is formed, we performed a study which examines how CO alters the electronic structure of different BTC MOF complexes.

Computational Details

A single node of the CuRhBTC MOF (Figure 4.9) was utilized for all calculations which were performed with the ORCA 5.0 software package.[111] The model complex (A) was truncated with acetate groups (B) for all calculations. Geometry optimizations were performed using the M06L[426] density functional with the def2-TZVPP basis set for the transition metals (with the corresponding def2-ECP for Ru, Rh, and Ir) while the def2-TZVP basis set was used for all other atoms.[275] Tight convergence criteria were used for both the SCF iterations and the geometry optimization steps. The RI approximation was used to accelerate the computation of all four-index integrals.[248] The def2/J auxiliary basis set was used for all atoms.[380] A triplet spin state was assumed for all calculations which involved the CuRh node,[294] an antiferromagnetically coupled singlet for the CuCu node,[421] and a singlet spin state for the RhRh node.[420] For all other nodes, the lowest three spin states were examined using both the bare complex as well as the CO-adsorbed complex. Numerical frequency calculations were performed to ensure that minima of the potential energy surface for all structures were obtained.

Results and Discussions

The strongest binding enthalpy (ΔH^{298}) of CO was found to be 42.9 kcal/mol for the CuRuBTC MOF with the CuIr and RhRu (CO bound to Ru) nodes showing similar binding enthalpies of 41.9 and 40.6 kcal/mol, respectively (Table 4.4). Such strong binding would limit the feasibility of catalysis due to strong chemisorption. The lowest observed binding energies were seen with CuCu, CuMn, CuFe, and CuNi with enthalpies of 7.8, 13.0, 14.0, and 14.8 kcal/mol, respectively. The latter three values are within a range that is not too small to prefer desorption but not too large either to prefer chemisorption (as is the case with CuRu, CuIr, and RhRu).

We may examine the change in the stretching frequency of CO as a function of the choice of metal in the node to better understand changes in the electronic structure of the system upon binding.[427] A shift of -141.8 cm^{-1} was observed for the CuRuBTC MOF which is accompanied by a strong CuRu-CO interaction and thus the poor activity that may be observed with regards to

Table 4.4 Calculated binding enthalpies (kcal/mol) for CO based upon the lowest energy spin state structure for several different BTC MOF structures. Bond lengths given in Å. Stretching frequencies given in cm^{-1}. CO was adsorbed to the second metal site, i.e., Ni for CuNi. * Denotes the spin state is an antiferromagnetic spin state. CO shows a bond length (R_{CO}) of 1.1278 Å and a stretching frequency (ν_{CO}) of 2209.42 cm^{-1} in the gas phase using the M06L density functional. All stretching modes have been corrected for using a scaling factor of 0.9965.[1]

	Spin Multiplicity		ΔH^{298}	ΔR_{CO}	$\Delta \nu_{CO}$	Metal-Metal Bond Length			Metal-CO bond length
	Before Adsorption	After Adsorption				Before Adsorption	After Adsorption	ΔR	
CuMn	S = 2	S = 2	-12.98	-0.004	16.3	2.663	2.728	0.065	2.225
CuFe	S = 1.5	S = 2.5	-13.99	0.007	-60.6	2.619	2.674	0.055	2.014
CuCo	S = 1	S = 2	-20.40	0.003	-32.5	2.417	2.566	0.149	1.956
CuNi	S = 0.5	S = 1.5	-14.83	-0.002	8.4	2.431	2.497	0.066	1.974
CuCu	S = 0*	S = 0*	-7.79	-0.003	21.1	2.474	2.533	0.059	2.217
CuRu	S = 1.5	S = 0.5	-42.92	0.025	-141.8	2.515	2.556	0.041	1.787
CuRh	S = 1	S = 1	-19.29	-0.002	-96.8	2.436	2.523	0.087	2.052
CuIr	S = 1	S = 0	-41.86	0.017	-83.9	2.459	2.497	0.038	1.830
RhRh	S = 0	S = 0	-26.56	0.011	-89.6	2.366	2.418	0.052	1.960
RuRh	S = 0.5	S = 0.5	-18.23	0.010	-102.0	2.365	2.427	0.062	2.025
RhRu	S = 0.5	S = 0.5	-40.61	0.023	-161.1	2.365	2.463	0.098	1.875

hydroformylation. Of the eleven tested structures, only four resulted in small shifts (< 40 cm^{-1}) of the CO stretching frequency: CuCu, CuNi, CuMn, and CuCo. These four also showed very little change in the R_{CO} bond length as well, with a maximum displacement of -0.004 Å for the CuMn case. This indicates that the interaction with the metal node and CO is weak enough to not perturb the electronic structure of the CO molecule significantly. As these four support metals (Mn, Co, Ni, Cu) are all first-row transition metals, their metal-metal and metal-CO bond lengths should all be similar and thus an appropriate comparison may be made between them to further examine the metal-CO interactions. Upon adsorption, all the first-row transition metal complexes (including Fe) show an increase in their respective metal-metal bond lengths with CuCo showing the largest increase (0.149 Å). Likewise, the CuCo complex results in the shortest metal-CO bond length (1.956 Å) which is commensurate with the largest binding energy of the five, first-row cases. This binding energy may be too high, however, for room temperature studies and thus, the CuMn, CuFe, and CuNi may be better candidates to be suggested to experimentalists to examine how well the synthesized MOF materials may perform with regards to the hydroformylation reaction. These calculations are not able to explain the phenomenon observed experimentally whereby CO addition causes the Rh(II) sites to reduce to pure metallic Rh(0), however. The binding of CO to the CuRhBTC MOF does not result in a large binding energy (19.3 kcal/mol), however, it does perturb the ν_{CO} to a large degree (-96.8 cm^{-1}) which may indicate a significant change in the electronic structure of the complex upon the binding of CO. Further examination of this species must be performed to better understand the deformation upon CO addition especially regarding the interaction with thermally accessible excited states. In addition, certain spin states may be accessible at lower temperatures which may allow for hydroformylation to occur on different spin state surfaces of the MOF complexes.

4.4.7 Conclusions

The Rh(II) transmetalated, HKUST-1 MOF has been shown experimentally to be highly active towards propene hydrogenation but not ethyne hydrogenation. A theoretical examination of the reaction pathway for both hydrogenation reactions as well as that of the hydrogenation of

ethene was performed. These studies identified a reaction intermediate where a C_2H_3 group is bound to the Rh(II) site of the CuRhBTC MOF complex which is very stable and thus results in a large barrier towards the final protonation step to form ethane. Similar calculations were unable to reproduce an experimental trend which showed stronger interactions of the CuRhBTC MOF with ethyne than either propene or ethene and may be a result of the ethyne poisoning the catalyst.

Through a systematic study of the binding of CO towards different BTC MOF complexes, it was found that when the Cu(II) center of the CuBTC MOF is transmetalated with either Mn, Fe, Co, or Ni, the interaction with CO is neither too strong nor too weak; between 15 – 25 kcal/mol. Strong interactions would disallow for hydroformylation as the CO bound complex would most likely not be able to lead to any new reaction intermediates while weak interactions would only allow for passive interactions with the metal site. These four complexes show binding energies between 13 and 20 kcal/mol with relatively small perturbations in the stretching frequency of CO ($\Delta\nu_{CO} < 60$ cm^{-1}). Thus, it may be advantageous to study these four transmetalated MOF materials as possible catalysts towards hydroformylation.

Spectroscopic studies were performed to positively identity the metal-metal stretching frequency of the CuBTC, CuRhBTC, and RhBTC MOF materials and good agreement was made between the experimental and theorical results. Likewise, the identification of peaks that appear in the Raman spectra during *in situ* experiments for the CuRhBTC MOF were possibly identified as methyl rotations on the adsorbed propene molecule.

4.5 Conclusions

Computational methods have been used extensively for the characterization and elucidation of the properties of catalysts. The use of electronic structure theory (EST) for catalytic studies allow for an atom-by-atom examination of the reaction mechanism which is impossible to achieve through experimental methods. The examination of three catalytic mechanisms was performed using computational methods which includes the dehalogenation of CH_2Cl_2 to form CH_4, the hydrogen evolution reaction to produce H_2, and the hydrogenation of either of propene, ethyne, and ethene. The catalysts used for these reactions utilize either d- or p-block metals such

as Cu, Rh, Sb, or Sn. The dehalogenation of CH_2Cl_2 was shown to proceed with very low barriers for both the C-Cl bond breaking and C-H bond formation steps which never exceeds a value of 0.45 eV (10.4 kcal/mol). Likewise, the hydrogen evolution reactions studied using both Sb and Sn molecular catalysts proceeds with low barriers in similar mechanisms: H_2 is produced in an ECEC fashion where E stands for an electric step (reduction) and C stands for a chemical step (protonation). Finally, the hydrogenation of propene, ethyne, and ethene was examined using the CuRhBTC MOF complex and low barriers were observed for propene and ethyne hydrogenation while a single high barrier of 22.5 kcal/mol was observed for ethene hydrogenation which may limit the feasibility the production of ethane. These results help to illustrate the usefulness of computational methods towards a better understanding of different reaction mechanisms. Knowing the steps that a reaction takes may help experimental chemists in some form to design better catalysts which either show lower barriers, such as those seen for the dehalogenation reaction, more selectivity, such as the CuRhBTC MOF catalyst performing well for propene hydrogenation but not for hydroformylation, or high activity, such as the Sb and Sn HER catalysts which showed Faradaic efficiencies upwards of 100%. Computational methods allow for a view into the chemical realm that is simply impossible through experimental efforts and thus is invaluable towards the study of catalytic systems. This work may serve as a knowledge base for future calculations such as the high-throughput screening of MOF catalysts for the electroreduction of CO_2.

Chapter 5

Summary and Outlook

In this work, we performed a large number of calculations on chemical systems ranging from diatomic molecules to larger, molecular complexes which are of a challenging nature. To perform this work, coupled cluster is utilized to perform highly accurate calculations on small systems (between two and five atoms) while density functional theory is used for the high-throughput examination of larger systems. In Chapter 2, a detailed discussion was made regarding the reproduction of the rotational, vibrational, and rovibrational spectra of the CF^+ molecular cation and a H_2O molecule in a parahydrogen HCP lattice. The spectroscopic constants of CF^+ were reproduced with very low error due to the highly accurate potential energy curve that was developed in conjunction with a vibrational Hamiltonian. This method can be easily applied to other diatomic molecules of interest, such as those already identified in the interstellar medium (ISM) or more unique molecules such as lanthanide diatomics. Likewise, the construction of three potential energy curves and two potential energy surfaces were combined with a rotational Hamiltonian to examine what effects the environment of a perturbed HCP lattice of parahydrogen molecules has on a rotating H_2O molecule. The underlying theoretical framework constructed for this work is applicable to other systems such a NH_3 molecule in an HCP lattice.

Following this, calculations which involved lanthanide compounds was performed in Chapter 3. Initial calculations were performed to identify the theoretical methodology which best describe the electronic structure of lanthanide diatomics as well as larger, molecular complexes using density functional theory (DFT). This theoretical methodology was then applied to the calculation of the free binding energies of lanthanide-tris-β-diketone complexes. By using high-throughput screening and machine learning, we were able to formulate a mathematical function which can map the chemical structure of a β-diketone ligand to its free binding energy towards

either lanthanum or lutetium. Good predictability was achieved using the many-body tensor representation with the kernel ridge regression learner.

Finally, Chapter 4 introduced work done utilizing computational methods to study catalytic systems. The dehalogenation of CH_2Cl_2 and the hydrogen evolution reaction were both examined using molecule complexes. Experimental results showed high activity for the former reaction which was explained through the presence of low activation barriers (< 10 kcal/mol). Likewise, the production of hydrogen was examined on two surfaces and the high experimental activity could be attributed to a thermodynamically and kinetically favorable processes. After this, calculations which used metal-organic frameworks (MOF) were examined to understand how hydrogenation occurs on a transmetalated form of the HKUST-1 MOF. Low activation barriers were observed for propylene and acetylene hydrogenation while a slightly higher activation barrier was observed for ethylene hydrogenation. Following this, several forms of the HKUST-1 MOF were examined to test their applicability towards the production of aldehydes via hydroformylation. Three of the tested 3d transition metals resulted in perturbations to the MOF which may be advantageous towards aldehyde production.

Bibliography

[1] Kesharwani, M. K.; Brauer, B.; Martin, J. M. L. Frequency and Zero-Point Vibrational Energy Scale Factors for Double-Hybrid Density Functionals (and Other Selected Methods): Can Anharmonic Force Fields Be Avoided? *J. Phys. Chem. A* **2015**, *119* (9), 1701–1714.

[2] Martin Gruebele, Mark Polak, R. J. S. Velocity Modulation Laser Spectroscopy of Vibrationally Excited CF^+: Determination of the Molecular Potential Function. *Chem. Phys. Lett.* **1986**, *125* (2), 165–169.

[3] Schrödinger, E. Quantisierung Als Eigenwertproblem. *Ann. Phys.* **1926**, *79*, 361.

[4] Dirac, P. A. M. The Quantum Theory of the Electron. *Proc. R. Soc. Lond. A* **1928**, *117*, 610.

[5] He, Z. H.; Ye, X. B.; Pan, B. C. Linear Scaling Algorithm for Tight-Binding Molecular Dynamics Simulations. *J. Chem. Phys.* **2019**, *150* (11).

[6] Hofer, T. S. From Macromolecules to Electrons-Grand Challenges in Theoretical and Computational Chemistry. *Front. Chem.* **2013**, *1* (May), 1–4.

[7] Dill, K. A.; Maccallum, J. L.; Folding, P. The Protein-Folding Problem , 50 Years On. **2012**, No. NOVEMBER, 1042–1047.

[8] Bash, E. Water and Proteins: A Love- Hate Relationship Yaakov. *PhD Propos.* **2015**, *1* (10), 3325–3326.

[9] Zaccai, G.; Daniel, R. M.; Cupane, A.; Hooper, R.; Franks, F. The Effect of Water on Protein Dynamics. *Philos. Trans. R. Soc. B Biol. Sci.* **2004**, *359* (1448), 1269–1275.

[10] De Simone, A.; Dodson, G. G.; Verma, C. S.; Zagari, A.; Fraternali, F. Prion and Water: Tight and Dynamical Hydration Sites Have a Key Role in Structural Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (21), 7535–7540.

[11] Mamontov, E.; Chu, X. Q. Water-Protein Dynamic Coupling and New Opportunities for

Probing It at Low to Physiological Temperatures in Aqueous Solutions. *Phys. Chem. Chem. Phys.* **2012**, *14* (33), 11573–11588.

[12] Xu, X.; Ma, Z.; Wang, X.; Xiao, Z. T.; Li, Y.; Xue, Z. H.; Wang, Y. H. Water's Potential Role: Insights from Studies of the P53 Core Domain. *J. Struct. Biol.* **2012**, *177* (2), 358–366.

[13] Varghese, J. J.; Mushrif, S. H. Origins of Complex Solvent Effects on Chemical Reactivity and Computational Tools to Investigate Them: A Review. *React. Chem. Eng.* **2019**, *4* (2), 165–206.

[14] Monard, G.; Rivail, J.-L. Solvent Effects in Quantum Chemistry. In *Handbook of Computational Chemistry*; 2017; pp 727–739.

[15] Das, M.; Gogoi, A. R.; Sunoj, R. B. Molecular Insights on Solvent Effects in Organic Reactions as Obtained through Computational Chemistry Tools. *J. Org. Chem.* **2022**, *87* (3), 1630–1640.

[16] Gastegger, M.; Schütt, K. T.; Müller, K. R. Machine Learning of Solvent Effects on Molecular Spectra and Reactions. *Chem. Sci.* **2021**, *12* (34), 11473–11483.

[17] Wang, M.; Lu, Z.; Yang, W. Nuclear Quantum Effects on an Enzyme-Catalyzed Reaction with Reaction Path Potential: Proton Transfer in Triosephosphate Isomerase. *J. Chem. Phys.* **2006**, *124* (12).

[18] Agarwal, P. K.; Billeter, S. R.; Hammes-Schiffer, S. Nuclear Quantum Effects and Enzyme Dynamics in Dihydrofolate Reductase Catalysis. *J. Phys. Chem. B* **2002**, *106* (12), 3283–3293.

[19] Krilov, G.; Sim, E.; Berne, B. J. Quantum Time Correlation Functions from Complex Time Monte Carlo Simulations: A Maximum Entropy Approach. *J. Chem. Phys.* **2001**, *114* (3), 1075–1088.

[20] Habershon, S.; Braams, B. J.; Manolopoulos, D. E. Quantum Mechanical Correlation Functions, Maximum Entropy Analytic Continuation, and Ring Polymer Molecular

Dynamics. *J. Chem. Phys.* **2007**, *127* (17).

[21] Paesani, F.; Voth, G. A. Nonlinear Quantum Time Correlation Functions from Centroid Molecular Dynamics and the Maximum Entropy Method. *J. Chem. Phys.* **2008**, *129* (19).

[22] Bonella, S.; Monteferrante, M.; Pierleoni, C.; Ciccotti, G. Path Integral Based Calculations of Symmetrized Time Correlation Functions. II. *J. Chem. Phys.* **2010**, *133* (16), 164105.

[23] Bonella, S.; Monteferrante, M.; Pierleoni, C.; Ciccotti, G. Path Integral Based Calculations of Symmetrized Time Correlation Functions. I. *J. Chem. Phys.* **2010**, *133* (16), 164104.

[24] Krylov, A.; Windus, T. L.; Barnes, T.; Marin-Rimoldi, E.; Nash, J. A.; Pritchard, B.; Smith, D. G. A.; Altarawy, D.; Saxe, P.; Clementi, C.; Crawford, T. D.; Harrison, R. J.; Jha, S.; Pande, V. S.; Head-Gordon, T. Perspective: Computational Chemistry Software and Its Advancement as Illustrated through Three Grand Challenge Cases for Molecular Science. *J. Chem. Phys.* **2018**, *149* (18).

[25] Gary, J. H. Petroleum Refining. In *Encyclopedia of Physical Science and Technology*; 2001; pp 741–761.

[26] Nicolaou, K. C. Catalyst: Synthetic Organic Chemistry as a Force for Good. *Chem* **2016**, *1* (3), 331–334.

[27] Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the Computational Design of Solid Catalysts. *Nat. Chem.* **2009**, *1* (1), 37–46.

[28] Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2019**, *119* (4), 2453–2523.

[29] Scholes, G. D. Introduction: Light Harvesting. *Chem. Rev.* **2017**, *117* (2), 247–248.

[30] Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and Their Functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6* (3), 197–208.

[31] Tompa, P. Intrinsically Disordered Proteins: A 10-Year Recap. *Trends Biochem. Sci.* **2012**, *37* (12), 509–516.

[32] Puzzarini, C. Rotational Spectroscopy Meets Theory. *Phys. Chem. Chem. Phys.* **2013**, *15* (18), 6595–6607.

[33] Ehrenfreund, P.; Charnley, S. B. Organic Molecules in the Interstellar Medium, Comets, and Meteorites: A Voyage from Dark Clouds to the Early Earth. *Annu. Rev. Astron. Astrophys.* **2000**, *38* (1), 427–488.

[34] Fortenberry, R. C. Quantum Astrochemical Spectroscopy. *Int. J. Quantum Chem.* **2017**, *117* (2), 81–91.

[35] Fortenberry, R. C. Interstellar Anions: The Role of Quantum Chemistry. *J. Phys. Chem. A* **2015**, *119* (39), 9941–9953.

[36] Baiano, C.; Lupi, J.; Tasinato, N.; Puzzarini, C.; Barone, V. The Role of State-of-the-Art Quantum-Chemical Calculations in Astrochemistry: Formation Route and Spectroscopy of Ethanimine as a Paradigmatic Case. *Molecules* **2020**, *25* (12).

[37] Fortenberry, R. C.; Huang, X.; Yachmenev, A.; Thiel, W.; Lee, T. J. On the Use of Quartic Force Fields in Variational Calculations. *Chem. Phys. Lett.* **2013**, *574*, 1–12.

[38] Piccardo, M.; Bloino, J.; Barone, V. Generalized Vibrational Perturbation Theory for Rotovibrational Energies of Linear, Symmetric and Asymmetric Tops: Theory, Approximations, and Automated Approaches to Deal with Medium-to-Large Molecular Systems. *Int. J. Quantum Chem.* **2015**, *115* (15), 948–982.

[39] Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-Chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29* (2), 273–367.

[40] Puzzarini, C. Grand Challenges in Astrochemistry. *Front. Astron. Sp. Sci.* **2020**, *7* (May), 8–11.

[41] Chyba, C.; Sagan, C. Endogenous Production, Exogenous Delivery and Impact-Shock Synthesis of Organic Molecules: An Inventory for the Origins of Life. *Nature* **1992**, *355*

(6356), 125–132.

[42] Dolg, M. *Computational Methods in Lanthanide and Actinide Chemistry*; John Wiley & Sons, 2015.

[43] Siegbahn, P. E. M. A New Direct CI Method for Large CI Expansions in a Small Orbital Space. *Chem. Phys. Lett.* **1984**, *109* (5), 417–423.

[44] Handy, N. C. Multi-Root Configuration Interaction Calculations. *Chem. Phys. Lett.* **1980**, *74* (2), 280–283.

[45] Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. Determinant Based Configuration Interaction Algorithms for Complete and Restricted Configuration Interaction Spaces. *J. Chem. Phys.* **1988**, *89* (4), 2185–2192.

[46] Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. Multireference Electron Correlation Methods: Journeys along Potential Energy Surfaces. *Chem. Rev.* **2020**, *120* (13), 1–33.

[47] Kolarik, Z. Complexation and Separation of Lanthanides(III) and Actinides(III) by Heterocyclic N-Donors in Solutions. *Chem. Rev.* **2008**, *108* (10), 4208–4252.

[48] Autschbach, J. Perspective: Relativistic Effects. *J. Chem. Phys.* **2012**, *136* (15).

[49] Pyykkö, P. Relativistic Effects in Chemistry: More Common than You Thought. *Annu. Rev. Phys. Chem.* **2012**, *63*, 45–64.

[50] Van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-Component Hamiltonians. *J. Chem. Phys.* **1993**, *99* (6), 4597–4610.

[51] Heully, J.-L.; Lindgrent, I.; Lindrothi, E.; Lundqvist, S.; Martensson-Pendrill, A.-M. Diagonalisation of the Dirac Hamiltonian as a Basis for a Relativistic Many-Body Procedure. *J. Phys. B At. Mol. Phys.* **1986**, *19*, 2799–2815.

[52] Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys. (N. Y).* **1974**, *82*, 89–155.

[53] Hess, B. A. Applicability of the No-Pair Equation with Free-Particle Projection Operators to Atomic and Molecular Structure Calculations. *Phys. Rev. A* **1985**, *32* (2), 756–763.

[54] Li, Z.; Ji, S.; Liu, Y.; Cao, X.; Tian, S.; Chen, Y.; Niu, Z.; Li, Y. Well-Defined Materials for Heterogeneous Catalysis: From Nanoparticles to Isolated Single-Atom Sites. *Chem. Rev.* **2020**, *120* (2), 623–682.

[55] Wang, A.; Li, J.; Zhang, T. Heterogeneous Single-Atom Catalysis. *Nat. Rev. Chem.* **2018**, *2* (6), 65–81.

[56] Schlexer, P. Computational Modeling in Heterogeneous Catalysis. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; 2017.

[57] T. Bligaard, J. K. N. Chapter 4 - Heterogeneous Catalysis. In *Chemical Bonding at Surfaces and Interfaces*; 2008; pp 255–321.

[58] Wisniak, J. The History of Catalysis. From the Beginning to Nobel Prizes. *Educ. Quim.* **2010**, *21* (1), 60–69.

[59] Erisman, J. W.; Sutton, M. A.; Galloway, J.; Klimont, Z.; Winiwarter, W. How a Century of Ammonia Synthesis Changed the World. *Nat. Geosci.* **2008**, *1* (10), 636–639.

[60] McCarver, G. A.; Rajeshkumar, T.; Vogiatzis, K. D. Computational Catalysis for Metal-Organic Frameworks: An Overview. *Coord. Chem. Rev.* **2021**, *436*, 213777.

[61] Cheung, A. C.; Rank, D. M.; Townes, C. H.; Thornton, D. D.; Welch, W. J. Detection of NH_3 Molecules in the Interstellar Medium by Their Microwave Emission. *Phys. Rev. Lett.* **1968**, *21* (25), 1701–1705.

[62] McGuire, B. A. 2018 Census of Interstellar, Circumstellar, Extragalactic, Protoplanetary Disk, and Exoplanetary Molecules. *Astrophys. J. Suppl. Ser.* **2018**, *239* (2), 17.

[63] Lambert, D. L.; Sheffer, Y.; Federman, S. R. Hubble Space Telescope Observations of C_2 Molecules in Diffuse Interstellar Clouds. *Astrophys. J.* **1995**, *438*, 740–749.

[64] Souza, S. P.; Lutz, B. L. Detection of C_2 In the Interstellar Spectra of Cygnus OB2 Number 12 (VI Cygni Number 12)*. *Astrophys. J.* **1977**, *216*, L49–L51.

[65] Geballe, T. R.; Oka, T. Detection of H_3^+ in Interstellar Space. *Nature* **1996**, *384* (November), 334.

[66] Irvine, W. M.; Avery, L. W.; Friberg, P.; Matthews, H. E.; Ziurys, L. M. Newly Detected Molecules in Dense Interstellar Clouds. *Astron. Lett. Commun.* **1988**, *26*, 167–180.

[67] Cami, J.; Bernard-Salas, J.; Peeters, E.; Malek, S. E. Detection of C_{60} and C_{70} in a Young Planetary Nebula. *Science (80-.).* **2010**, *329* (September), 1180–1182.

[68] Carruthers, G. R. Rocket Observation of Interstellar Molecular Hydrogen. *Astrophys. J.* **1970**, *161*, L81–L85.

[69] Ziurys, L. M.; Frieberg, P.; Irvine, W. M. Interstellar SiO as a Tracer of High-Temperature Chemistry. *Astrophys. J.* **1989**, *343*, 201–207.

[70] Furuya, R. S.; Walmsley, C. M.; Nakanishi, K.; Schilke, P.; Bachiller, R. Interferometric Observations of FeO towards Sagittarius B2. *Astron. Astrophys.* **2003**, *409* (2), 21–24.

[71] Bondybey, V. E.; Räsänen, M.; Lammers, A. Rare-Gas Matrices, Their Photochemistry and Dynamics: Recent Advances in Selected Areas. *Annu. Reports Prog. Chem. - Sect. C* **1999**, *95* (10), 331–372.

[72] Haupa, K. A.; Strom, A. I.; Anderson, D. T.; Lee, Y. P. Hydrogen-Atom Tunneling Reactions with Methyl Formate in Solid Para -Hydrogen: Infrared Spectra of the Methoxy Carbonyl [•$C(O)OCH_3$] and Formyloxy Methyl [$HC(O)OCH_2$•] Radicals. *J. Chem. Phys.* **2019**, *151* (23), 0–14.

[73] Wonderly, W. R.; Anderson, D. T. Reactions of Atomic Hydrogen with Formic Acid and Carbon Monoxide in Solid Parahydrogen II: Deuterated Reaction Studies. *J. Phys. Chem. A* **2014**, *118* (36), 7653–7662.

[74] Wonderly, W. R.; Anderson, D. T. Transient HDO Rovibrational Satellite Peaks in Solid Parahydrogen: Evidence of Hydrogen Atoms or Vacancies? *Low Temp. Phys.* **2012**, *38* (8), 673–678.

[75] Kufeld, K. A.; Wonderly, W. R.; Paulson, L. O.; Kettwich, S. C.; Anderson, D. T.

Transient H_2O Infrared Satellite Peaks Produced in UV Irradiated Formic Acid Doped Solid Parahydrogen. *J. Phys. Chem. Lett.* **2012**, *3* (3), 342–347.

[76] Ruzi, M.; Anderson, D. T. Fourier Transform Infrared Studies of Ammonia Photochemistry in Solid Parahydrogen. *J. Phys. Chem. A* **2013**, *117* (50), 13832–13842.

[77] Ruzi, M.; Anderson, D. T. Quantum Diffusion-Controlled Chemistry: Reactions of Atomic Hydrogen with Nitric Oxide in Solid Parahydrogen. *J. Phys. Chem. A* **2015**, *119* (50), 12270–12283.

[78] Ruzi, M.; Anderson, D. T. Matrix Isolation Spectroscopy and Nuclear Spin Conversion of NH_3 and ND_3 Suspended in Solid Parahydrogen. *J. Phys. Chem. A* **2012**, *117*, 9712–9724.

[79] Pinelo, L. F.; Klotz, E. R.; Wonderly, W. R.; Paulson, L. O.; Kettwich, S. C.; Kubelka, J.; Anderson, D. T. Solid Parahydrogen Infrared Matrix Isolation and Computational Studies of Li_n-$(C_2H_4)_m$ Complexes. *J. Phys. Chem. A* **2018**, *122* (4), 985–991.

[80] Mutunga, F. M.; Anderson, D. T. Infrared Spectroscopy and 193 Nm Photochemistry of Methylamine Isolated in Solid Parahydrogen. *J. Phys. Chem. A* **2015**, *119* (11), 2420–2428.

[81] Raston, P. L.; Anderson, D. T. Infrared-Induced Reaction of Cl Atoms Trapped in Solid Parahydrogen. *Phys. Chem. Chem. Phys.* **2006**, *8* (26), 3124–3129.

[82] Nuth III, J. A.; Charnley, S. B.; Johnson, N. M. Chemical Processes in the Interstellar Medium : Source of the Gas and Dust in the Primitive Solar Nebula. In *Meteorites and the Early Solar System II*; 2006; pp 147–167.

[83] Endres, C. P.; Schlemmer, S.; Schilke, P.; Stutzki, J.; Müller, H. S. P. The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC. *J. Mol. Spectrosc.* **2016**, *327*, 95–104.

[84] Duley, W. W.; Hu, A. Fullerenes and Proto-Fullerenes in Interstellar Carbon Dust. *Astrophys. J. Lett.* **2012**, *745* (1), 2008–2011.

[85] Tenenbaum, E. D.; Ziurys, L. M. Exotic Metal Molecules in Oxygen-Rich Envelopes:

Detection of AlOH ($X^1\Sigma^+$) in VY Canis Majoris. *Astrophys. J. Lett.* **2010**, *712*, 93–97.

[86] Agúndez, M.; Cernicharo, J.; Decin, L.; Encrenaz, P.; Teyssier, D. Confirmation of Circumstellar Phosphine. *Astrophys. J. Lett.* **2014**, *790* (2), 2–5.

[87] Zack, L. N.; Halfen, D. T.; Ziurys, L. M. Detection of FeCN (X4Δi) in IRC+10216: A New Interstellar Molecule. *Astrophys. J. Lett.* **2011**, *733*.

[88] Neufeld, D. A.; Schilke, P.; Menten, K. M.; Wolfire, M. G.; Black, J. H.; Schuller, F.; Müller, H. S. P.; Thorwith, S.; Güsten, R.; Philipp, S. Discovery of Interstellar CF+. *Astron. Astrophys.* **2006**, *454*, L37–L40.

[89] Kumar, V.; Pulpytel, J.; Rauscher, H.; Mannelli, I.; Rossi, F.; Arefi-Khonsari, F. Fluorocarbon Coatings via Plasma Enhanced Chemical Vapor Deposition of 1H,1H,2H,2H-Perfluorodecyl Acrylate - 2, Morphology, Wettability and Antifouling Characterization. *Plasma Process. Polym.* **2010**, *7* (11), 926–938.

[90] Sicard, A. J.; Baker, R. T. Fluorocarbon Refrigerants and Their Syntheses: Past to Present. *Chem. Rev.* **2020**, *120* (17), 9164–9303.

[91] Green, S. W.; Slinn, D. S. L.; Simpson, R. N. F.; Woytek, A. J. Perflurocarbon Fluids. In *Organofluroine Chemistry, Principles and Commerical Applications*; 1994; p 89.

[92] Stoffels, A.; Kluge, L.; Schlemmer, S.; Brünken, S. Laboratory Rotational Ground State Transitions of NH_3D^+ and CF^+. *Astron. Astrophys.* **2016**, *593*, 1–7.

[93] Wu, Y. J.; Chen, H. F.; Chou, S. L.; Lin, M. Y.; Cheng, B. M. Vacuum-Ultraviolet Photolysis of H_3CF in Solid Neon: Infrared Spectra of HCF and CF^+. *Chem. Phys. Lett.* **2010**, *497* (1–3), 12–17.

[94] Cazzoli, G.; Cludi, L.; Puzzarini, C.; Gauss, J. Rotational Spectra of CF^+ and $^{13}CF^+$: Accurate Rest Frequencies and Spectroscopic Parameters. *Astron. Astrophys.* **2010**, *509* (1), 1–5.

[95] Cazzoli, G.; Cludi, L.; Buffa, G.; Puzzarini, C. Precise THz Measurements of HCO^+, N_2H^+, and CF^+ for Astrophysical Observations. *Astrophys. Journal, Suppl. Ser.* **2012**, *203*

(1).

[96] Muller, S.; Kawaguchi, K.; Black, J. H.; Amano, T. Detection of Extragalactic CF^+ toward PKS 1830 - 211: Chemical Differentiation in the Absorbing Gas. *Astron. Astrophys.* **2016**, *589*, 1–5.

[97] Peterson, K. A.; Claude Woods, R.; Rosmus, P.; Werner, H. J. Spectroscopic Properties of the $X^1\Sigma^+$ and a $^3\Pi$ Electronic States of CF^+, SiF^+, and CCl^+ by Multireference Configuration Interaction. *J. Chem. Phys.* **1990**, *93* (3), 1889–1894.

[98] Guzmán, V.; Roueff, E.; Gauss, J.; Pety, J.; Gratier, P.; Goicoechea, J. R.; Gerin, M.; Teyssier, D. The Hyperfine Structure in the Rotational Spectrum of CF^+ (Research Note). *Astron. Astrophys.* **2012**, *548*, 1–2.

[99] Inostroza, N.; Letelier, J. R.; Senent, M. L.; Fuentealba, P. Theoretical Ro-Vibrational Spectrum of CF^+. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2008**, *71* (3), 798–802.

[100] Petsalakis, I. D.; Theodorakopoulos, G. Electronic States of CF^+. *Chem. Phys.* **2000**, *254* (2–3), 181–186.

[101] Abe, M.; Kajita, M.; Hada, M.; Moriwaki, Y. Ab Initio Study on Vibrational Dipole Moments of XH^+ Molecular Ions: X = ^{24}Mg, ^{40}Ca, ^{64}Zn, ^{88}Sr, ^{114}Cd, ^{138}Ba, ^{174}Yb and ^{202}Hg. *J. Phys. B At. Mol. Opt. Phys.* **2010**, *43* (24).

[102] M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei, and Á. G. MRCC, a Quantum Chemical Program Suite.

[103] Kállay, M.; Nagy, P. R.; Mester, D.; Rolik, Z.; Samu, G.; Csontos, J.; Csóka, J.; Szabó, P. B.; Gyevi-Nagy, L.; Hégely, B.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Petrov, K.; Farkas, M.; Mezei, P. D.; Ganyecz, Á. The MRCC Program System: Accurate Quantum Chemistry from Water to Proteins. *J. Chem. Phys.* **2020**, *152* (7).

[104] Kállay, M.; Gauss, J. Approximate Treatment of Higher Excitations in Coupled-Cluster

Theory. II. Extension to General Single-Determinant Reference Functions and Improved Approaches for the Canonical Hartree-Fock Case. *J. Chem. Phys.* **2008**, *129* (14).

[105] Kállay, M.; Gauss, J. Approximate Treatment of Higher Excitations in Coupled-Cluster Theory. *J. Chem. Phys.* **2005**, *123* (21).

[106] Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. Coupled-Cluster Methods Including Noniterative Corrections for Quadruple Excitations. *J. Chem. Phys.* **2005**, *123* (5).

[107] Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron through Neon. *J. Chem. Phys.* **1995**, *103* (11), 4572–4585.

[108] Pansini, F. N. N.; Neto, A. C.; Varandas, A. J. C. Extrapolation of Hartree–Fock and Multiconfiguration Self-Consistent-Field Energies to the Complete Basis Set Limit. *Theor. Chem. Acc.* **2016**, *135* (12), 1–6.

[109] Pansini, F. N. N.; Varandas, A. J. C. Toward a Unified Single-Parameter Extrapolation Scheme for the Correlation Energy: Systems Formed by Atoms of Hydrogen through Neon. *Chem. Phys. Lett.* **2015**, *631–632*, 70–77.

[110] Pansini, F. N. N.; Neto, A. C.; Varandas, A. J. C. On the Performance of Various Hierarchized Bases in Extrapolating the Correlation Energy to the Complete Basis Set Limit. *Chem. Phys. Lett.* **2015**, *641*, 90–96.

[111] Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22).

[112] Dunham, J. L. The Energy Levels of a Rotating Vibrator. *Phys. Rev.* **1932**, *41*, 721–731.

[113] Londoño, B. E.; Mahecha, J. E.; Luc-Koenig, E.; Crubellier, A. Shape Resonances in Ground-State Diatomic Molecules: General Trends and the Example of RbCs. *Phys. Rev. A - At. Mol. Opt. Phys.* **2010**, *82* (1), 1–15.

[114] Larsson, M. Conversion Formulas between Radiative Lifetimes and Other Dynamical Variables for Spin-Allowed Electronic Transitions in Diatomic Molecules. *Astron.*

Astrophys. **1983**, *128* (291–298).

[115] Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A Simple DFT-Based Diagnostic for Nondynamical Correlation. *Theor. Chem. Acc.* **2013**, *132* (1), 1–9.

[116] Cooley, J. W. An Improved Eigenvalue Corrector Formula for Solving the Schrodinger Equation for Central Fields. *Math. Comput.* **1961**, *15* (76), 363.

[117] Kállay, M.; Surján, P. R. Higher Excitations in Coupled-Cluster Theory. *J. Chem. Phys.* **2001**, *115* (7), 2945–2954.

[118] Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2* (1), 73–78.

[119] Karton, A.; Martin, J. M. L. Comment on: "Estimating the Hartree-Fock Limit from Finite Basis Set Calculations" [Jensen F (2005) Theor Chem Acc 113:267]. *Theor. Chem. Acc.* **2006**, *115* (4), 330–333.

[120] Huh, S. B.; Lee, J. S. Basis Set and Correlation Dependent Extrapolation of Correlation Energy. *J. Chem. Phys.* **2003**, *118* (7), 3035–3042.

[121] Shagam, Y.; Klein, A.; Skomorowski, W.; Yun, R.; Averbukh, V.; Koch, C. P.; Narevicius, E. Molecular Hydrogen Interacts More Strongly When Rotationally Excited at Low Temperatures Leading to Faster Reactions. *Nat. Chem.* **2015**, *7* (11), 921–926.

[122] Lebedev, V. I. Quadratures on a Sphere. *USSR Comput. Math. Math. Phys.* **1976**, *16* (2), 10–24.

[123] Partridge, H.; Bauschlicher, C. W.; Stallcop, J. R.; Levin, E. Ab Initio Potential Energy Surface for H-H_2. *J. Chem. Phys.* **1993**, *99* (8), 5951–5960.

[124] Herzberg, G. Forbidden Transitions in Diatomic Molecules. I. The Quadrupole Rotation-Vibration Spectrum of H_2. *Can. J. Res. Sect. A.* **1950**, *28*, 144.

[125] Herzberg, G.; Monfils, A. The Dissociation Energies of H_2, HD, and D_2 Molecules. *J. Mol. Spect.* **1960**, *5*, 482.

[126] Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19* (4), 553–566.

[127] McKellar, A. R. W. Infrared Spectra of Hydrogen Dimers. *J. Chem. Phys.* **1990**, *92* (6), 3261–3277.

[128] McKellar, A. R. W.; Schaefer, J. Far-Infrared Spectra of Hydrogen Dimers: Comparisons of Experiment and Theory for $(H_2)_2$ and $(D_2)_2$ at 20 K. *J. Chem. Phys.* **1991**, *95* (5), 3081–3091.

[129] Norman, M. J.; Watts, R. O.; Buck, U. A Spherical Potential for Hydrogen from Solid State and Scattering Data. *J. Chem. Phys.* **1984**, *81* (8), 3500–3504.

[130] Hinde, R. J. A Six-Dimensional H_2 - H_2 Potential Energy Surface for Bound State Spectroscopy. *J. Chem. Phys.* **2008**, *128* (15).

[131] Lillestolen, T. C.; Hinde, R. J. Vibrationally Averaged Isotropic Dispersion Energy Coefficients of the Parahydrogen Dimer. *J. Chem. Phys.* **2012**, *136* (20).

[132] Green, S. Rotational Excitation in H_2-H_2 Collisions: Close-Coupling Calculations. *J. Chem. Physics1* **1975**, *62*, 2271.

[133] Kufeld, K. A.; Wonderly, W. R.; Paulson, L. O.; Kettwich, S. C.; Anderson, D. T. Transient H_2O Infrared Satellite Peaks Produced in UV Irradiated Formic Acid Doped Solid Parahydrogen. *J. Phys. Chem. Lett.* **2012**, *3* (3), 342–347.

[134] Kühn, O.; Manz, J.; Schild, A. Quantum Effects of Translational Motions in Solid Para-Hydrogen and Ortho-Deuterium: Anharmonic Extension of the Einstein Model. *J. Phys. Condens. Matter* **2010**, *22* (13).

[135] Silvera, I. F. The Solid Molecular Hydrogens in the Condensed Phase: Fundamentals and Static Properties. *Rev. Mod. Phys.* **1980**, *52* (2), 393–452.

[136] Sterling, M.; Li, Z.; Apkarian, V. A.; Voth, G. A.; Introduction, I. Simulations of Quantum Crystals by Classical Dynamics. **1995**, *103* (13), 5679–5683.

[137] McCarver, G. A.; Hinde, R. J. High Accuracy Ab Initio Potential Energy Surface for the H_2O-H van Der Waals Dimer. *J. Chem. Phys.* **2021**, *155* (11).

[138] Barreto, P. R. B.; Albernaz, A. F.; Capobianco, A.; Palazzetti, F.; Lombardi, A.; Grossi, G.; Aquilanti, V. Potential Energy Surfaces for Interactions of H_2O with H_2, N_2 and O_2: A Hyperspherical Harmonics Representation, and a Minimal Model for the H_2O-Rare-Gas-Atom Systems. *Comput. Theor. Chem.* **2012**, *990*, 53–61.

[139] Lombardi, A.; Pirani, F.; Bartolomei, M.; Coletti, C.; Laganà, A. Full Dimensional Potential Energy Function and Calculation of State-Specific Properties of the $CO+N_2$ Inelastic Processes within an Open Molecular Science Cloud Perspective. *Front. Chem.* **2019**, *7*, 1–16.

[140] Hinde, R. J. Constructing Atom-Molecule Potential Surfaces from Ab Initio Data: A Method Combining Quadrature and Interpolation. *Comput. Phys. Commun.* **2000**, *130* (1), 1–11.

[141] Hellström, M.; Behler, J. High-Dimensional Neural Network Potentials for Atomistic Simulations. *ACS Symp. Ser.* **2019**, *1326*, 49–59.

[142] Shao, Y.; Hellström, M.; Mitev, P. D.; Knijff, L.; Zhang, C. PiNN: A Python Library for Building Atomic Neural Networks of Molecules and Materials. *J. Chem. Inf. Model.* **2020**, *60*, 1184–1193.

[143] Neutsch, W. Optimal Spherical Designs and Numerical Integration on the Sphere. *J. Comput. Phys.* **1983**, *51* (2), 313–325.

[144] Freeden, W.; Nashed, M. Z.; Sonar, T. *Handbook of Geomathematics*; 2010; Vol. 2.

[145] Shirin, S. V.; Polyansky, O. L.; Zobov, N. F.; Barletta, P.; Tennyson, J. Spectroscopically Determined Potential Energy Surface of $H_2{}^{16}O$ up to 25000 Cm^{-1}. *J. Chem. Phys.* **2003**, *118* (5), 2124–2129.

[146] Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J.*

Chem. Inf. Model. **2019**, *59*, 4814–4820.

[147] Beentjes, C. H. L. Quadrature on a Spherical Surface. *Int. Conf. Math. Comput. Methods Appl. to Nucl. Sci. Eng.* **2015**, 1–15.

[148] Zeng, T.; Li, H.; Le Roy, R. J.; Roy, P. N. Adiabatic-Hindered-Rotor Treatment of the Parahydrogen-Water Complex. *J. Chem. Phys.* **2011**, *135* (9).

[149] Mizus, I. I.; Kyuberis, A. A.; Zobov, N. F.; Makhnev, V. Y.; Polyansky, O. L.; Tennyson, J. High-Accuracy Water Potential Energy Surface for the Calculation of Infrared Spectra. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2018**, *376* (2115), 1–11.

[150] Moller, C.; Plesset, M. S. Note on an Approximate Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

[151] Rose, M. *Elementary Theory of Angular Momentum*; John Wiley & Sons: New York, 1957.

[152] Tajima, N. Analytical Formula for Numerical Evaluations of the Wigner Rotation Matrices at High Spins. *Phys. Rev. C - Nucl. Phys.* **2015**, *91* (1), 1–5.

[153] Toth, R. A. Linelists of Water Vapor Parameters from 500 to 8000 Cm^{-1}.

[154] Bulthuis, J.; Möller, J.; Loesch, H. J. Brute Force Orientation of Asymmetric Top Molecules. *J. Phys. Chem. A* **1997**, *101* (41), 7684–7690.

[155] Andreev, S. N.; Makarov, V. P.; Tikhonov, V. I.; Volkov, A. A. Ortho and Para Molecules of Water in Electric Field. *arXiv* **2021**.

[156] Kumada, T. Experimental Determination of the Mechanism of the Tunneling Diffusion of H Atoms in Solid Hydrogen: Physical Exchange versus Chemical Reaction. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2003**, *68* (5), 16–18.

[157] Miyazaki, T.; Mori, S.; Nagasaka, T.; Kumagai, J.; Aratono, Y.; Kumada, T. Decay Dynamics of H Atoms in Solid Hydrogen at 4.2 K. Controlling Factor of Tunneling Reaction H + Para-H_2 → Para-H_2 + H. *J. Phys. Chem. A* **2000**, *104* (42), 9403–9407.

[158] Miyazaki, T.; Hiraku, T.; Fueki, K.; Tcuchihashi, Y. Effect of Rotational Quantum States (J=0, 1) on the Tunnelling Reaction $H_2 + H \rightarrow H + H_2$ in Parahydrogen Solid at 4.2 K. *J. Phys. Chem.* **1991**, *95*, 26–29.

[159] Miyazaki, T.; Iwata, N.; Lee, K. P.; Fueki, K. Decay of H (D) Atoms in Solid Hydrogen at 4.2 K. Rate Constant for Tunneling Reaction H_2 (D_2, HD) + H (D). *J. Phys. Chem.* **1989**, *93* (8), 3352–3355.

[160] Mutunga, F. M. Infrared Spectroscopy, UV Photochemistry, and H-Atom Tunneling Reactions of Molecules Trapped in Parahydrogen Crystals, 2016.

[161] IPIECA. *Petroleum Refinery Waste Management and Minimization*; 2014.

[162] Xu, J.; Gulzar, A.; Yang, P.; Bi, H.; Yang, D.; Gai, S.; He, F.; Lin, J.; Xing, B.; Jin, D. Recent Advances in Near-Infrared Emitting Lanthanide-Doped Nanoconstructs: Mechanism, Design and Application for Bioimaging. *Coord. Chem. Rev.* **2019**, *381*, 104–134.

[163] Lu, Q.; Peterson, K. A. Correlation Consistent Basis Sets for Lanthanides: The Atoms La-Lu. *J. Chem. Phys.* **2016**, *145*, 054111.

[164] Albrecht-Schmitt, T. E. Actinide Chemistry at the Extreme. *Inorg. Chem.* **2019**, *58* (3), 1721–1723.

[165] Anastasiadis, N. C.; Granadeiro, C. M.; Mayans, J.; Raptopoulou, C. P.; Bekiari, V.; Cunha-Silva, L.; Psycharis, V.; Escuer, A.; Balula, S. S.; Konidaris, K. F.; Perlepes, S. P. Multifunctionality in Two Families of Dinuclear Lanthanide(III) Complexes with a Tridentate Schiff-Base Ligand. *Inorg. Chem.* **2019**, *58* (15), 9581–9585.

[166] Mikami, K.; Terada, M.; Matsuzawa, H. “Asymmetric” Catalysis by Lanthanide Complexes. *Angew. Chemie - Int. Ed.* **2002**, *41* (19), 3554–3572.

[167] Dicken, R. D.; Motta, A.; Marks, T. J. Homoleptic Lanthanide Amide Catalysts for Organic Synthesis: Experiment and Theory. *ACS Catal.* **2021**, *11* (5), 2715–2734.

[168] Zhang, Y.; Liu, S.; Zhao, Z. S.; Wang, Z.; Zhang, R.; Liu, L.; Han, Z. B. Recent Progress

in Lanthanide Metal-Organic Frameworks and Their Derivatives in Catalytic Applications. *Inorg. Chem. Front.* **2021**, *8* (3), 590–619.

[169] Walsh, K. A. *Physical Properties of Nickel Carbonyl*; 1953.

[170] Vander Hoogerstraete, T.; Souza, E. R.; Onghena, B.; Banerjee, D.; Binnemans, K. Mechanism for Solvent Extraction of Lanthanides from Chloride Media by Basic Extractants. *J. Solution Chem.* **2018**, *47* (8), 1351–1372.

[171] Gupta, C. K.; Krishnamurthy, N. *Extractive Metallurgy of Rare Earths*; CRC Press, 2005.

[172] Bhattacharyya, A.; Gadly, T.; Kanekar, A. S.; Ghosh, S. K.; Kumar, M.; Mohapatra, P. K. First Report on the Separation of Trivalent Lanthanides from Trivalent Actinides Using an Aqueous Soluble Multiple N-Donor Ligand, 2,6-Bis(1 H-Tetrazol-5-Yl)Pyridine: Extraction, Spectroscopic, Structural, and Computational Studies. *Inorg. Chem.* **2018**, *57* (9), 5096–5107.

[173] Perreault, L. L.; Giret, S.; Gagnon, M.; Florek, J.; Larivière, D.; Kleitz, F. Functionalization of Mesoporous Carbon Materials for Selective Separation of Lanthanides under Acidic Conditions. *ACS Appl. Mater. Interfaces* **2017**, *9* (13), 12003–12012.

[174] Wehbie, M.; Arrachart, G.; Karamé, I.; Ghannam, L.; Pellet-Rostaing, S. Triazole Diglycolamide Cavitand for Lanthanide Extraction. *Sep. Purif. Technol.* **2016**, *169*, 17–24.

[175] Ellis, R. J.; Brigham, D. M.; Delmau, L.; Ivanov, A. S.; Williams, N. J.; Vo, M. N.; Reinhart, B.; Moyer, B. A.; Bryantsev, V. S. "Straining" to Separate the Rare Earths: How the Lanthanide Contraction Impacts Chelation by Diglycolamide Ligands. *Inorg. Chem.* **2016**, *56* (3), 1152–1160.

[176] Shimojo, K.; Aoyagi, N.; Saito, T.; Okamura, H.; Kubota, F.; Goto, M.; Naganawa, H. Highly Efficient Extraction Separation of Lanthanides Using a Diglycolamic Acid Extractant. *Anal. Sci.* **2014**, *30* (2), 263–269.

[177] Raut, D. R.; Mohapatra, P. K.; Ansari, S. A.; Godbole, S. V.; Iqbal, M.; Manna, D.; Ghanty, T. K.; Huskens, J.; Verboom, W. Complexation of Trivalent Lanthanides and Actinides with Several Novel Diglycolamide-Functionalized Calix[4]Arenes: Solvent Extraction, Luminescence and Theoretical Studies. *RSC Adv.* **2013**, *3* (24), 9296–9303.

[178] Sengupta, A.; Bhattacharyya, A.; Verboom, W.; Ali, S. M.; Mohapatra, P. K. Insight into the Complexation of Actinides and Lanthanides with Diglycolamide Derivatives: Experimental and Density Functional Theoretical Studies. *J. Phys. Chem. B* **2017**, *121* (12), 2640–2649.

[179] Baldwin, A. G.; Ivanov, A. S.; Williams, N. J.; Ellis, R. J.; Moyer, B. A.; Bryantsev, V. S.; Shafer, J. C. Outer-Sphere Water Clusters Tune the Lanthanide Selectivity of Diglycolamides. *ACS Cent. Sci.* **2018**, *4* (6), 739–747.

[180] Brigham, D. M.; Ivanov, A. S.; Moyer, B. A.; Delmau, L. H.; Bryantsev, V. S.; Ellis, R. J. Trefoil-Shaped Outer-Sphere Ion Clusters Mediate Lanthanide(III) Ion Transport with Diglycolamide Ligands. *J. Am. Chem. Soc.* **2017**, *139* (48), 17350–17358.

[181] Sharova, E. V; Artyushin, O. I.; Turanov, A. N.; Karandashev, V. K.; Meshkova, S. B.; Topilova, Z. M.; Odinets, I. L. N-Tris[(2-Aminoethyl)-2-(Diphenylphosphoryl) Acetamide)] - Novel CMPO Tripodand: Synthesis, Extraction Studies and Luminescent Properties of Lanthanide Complexes. *Cent. Eur. J. Chem.* **2012**, *10* (1), 146–156.

[182] Rosario-Amorin, D.; Ouizem, S.; Dickie, D. A.; Wen, Y.; Paine, R. T.; Gao, J.; Grey, J. K.; De Bettencourt-Dias, A.; Hay, B. P.; Delmau, L. H. Synthesis, Lanthanide Coordination Chemistry, and Liquid-Liquid Extraction Performance of CMPO-Decorated Pyridine and Pyridine N-Oxide Platforms. *Inorg. Chem.* **2013**, *52* (6), 3063–3083.

[183] Patterson, M. G.; Mulville, A. K.; Connor, E. K.; Henry, A. T.; Hudson, M. L.; Tissue, K.; Biros, S. M.; Werner, E. J. Lanthanide Extraction Selectivity of a Tripodal Carbamoylmethylphosphine Oxide Ligand System. *Dalt. Trans.* **2018**, *47* (40), 14318–14326.

[184] Gordon, M. S.; McCann, B. W.; Hay, B. P.; Silva, N. De; Windus, T. L.; Bryantsev, V. S.;

Moyer, B. A. Computer-Aided Molecular Design of Bis-Phosphine Oxide Lanthanide Extractants. *Inorg. Chem.* **2016**, *55* (12), 5787–5803.

[185] Platt, A. W. G. Lanthanide Phosphine Oxide Complexes. *Coord. Chem. Rev.* **2017**, *340* (3), 62–78.

[186] Elistratova, Y. G.; Mustafina, A. R.; Tatarinov, D. A.; Mironov, V. F.; Burilov, V. A.; Tananaev, I. G.; Konovalov, A. I. Extraction of Lanthanide Ions from Acidic and Strongly Acidic Media by Phosphine Oxide Derivatives Using Temperature-Induced Phase Separation. *Russ. Chem. Bull.* **2011**, *60* (5), 790–796.

[187] Turanov, A. N.; Karandashev, V. K.; Artyushin, O. I.; Sharova, E. V. Solvent Extraction of Lanthanides(III) from Nitric Acid Solutions with Novel Functionalized Ionic Liquids Based on the Carbamoyl(Methyl)Phosphine Oxide Pattern in Molecular Diluents. *Solvent Extr. Ion Exch.* **2021**, *00* (00), 1–13.

[188] Safiulina, A. M.; Matveeva, A. G.; Dvoryanchikova, T. K.; Sinegribova, O. A.; Tu, A. M.; Tatarinov, D. A.; Kostin, A. A.; Mironov, V. F.; Tananaev, I. G. Acetyl-Containing Phosphine Oxides as Extractants for Actinides and Lanthanides. *Russ. Chem. Bull.* **2012**, *61* (2), 392–398.

[189] Xu, L.; Pu, N.; Li, Y.; Wei, P.; Sun, T.; Xiao, C.; Chen, J.; Xu, C. Selective Separation and Complexation of Trivalent Actinide and Lanthanide by a Tetradentate Soft-Hard Donor Ligand: Solvent Extraction, Spectroscopy, and DFT Calculations. *Inorg. Chem.* **2019**, *58* (7), 4420–4430.

[190] Hasegawa, Y.; Tamaki, S.; Yajima, H.; Hashimoto, B.; Yaita, T. Selective Separation of Samarium(III) by Synergistic Extraction with β-Diketone and Methylphenylphenanthroline Carboxamide. *Talanta* **2011**, *85* (3), 1543–1548.

[191] Whittaker, D. M.; Griffiths, T. L.; Helliwell, M.; Swinburne, A. N.; Natrajan, L. S.; Lewis, F. W.; Harwood, L. M.; Parry, S. A.; Sharrad, C. A. Lanthanide Speciation in Potential SANEX and GANEX Actinide/Lanthanide Separations Using Tetra-N-Donor Extractants. *Inorg. Chem.* **2013**, *52* (7), 3429–3444.

[192] Jaoul, A.; Nocton, G.; Clavaguéra, C. Assessment of Density Functionals for Computing Thermodynamic Properties of Lanthanide Complexes. *ChemPhysChem* **2017**, *18* (19), 2688–2696.

[193] Xiao, C. L.; Wang, C. Z.; Yuan, L. Y.; Li, B.; He, H.; Wang, S.; Zhao, Y. L.; Chai, Z. F.; Shi, W. Q. Excellent Selectivity for Actinides with a Tetradentate 2,9-Diamide-1,10-Phenanthroline Ligand in Highly Acidic Solution: A Hard-Soft Donor Combined Strategy. *Inorg. Chem.* **2014**, *53* (3), 1712–1720.

[194] Ustynyuk, Y. A.; Borisova, N. E.; Babain, V. A.; Gloriozov, I. P.; Manuilov, A. Y.; Kalmykov, S. N.; Alyapyshev, M. Y.; Tkachenko, L. I.; Kenf, E. V.; Ustynyuk, N. A. N,N′-Dialkyl-N,N′-Diaryl-1,10-Phenanthroline-2,9-Dicarboxamides as Donor Ligands for Separation of Rare Earth Elements with a High and Unusual Selectivity. DFT Computational and Experimental Studies. *Chem. Commun.* **2015**, *51* (35), 7466–7469.

[195] Mccarver, G. A.; Hinde, R. J.; Vogiatzis, K. D. Selecting Quantum-Chemical Methods for Lanthanide-Containing Molecules: A Balance between Accuracy and Efficiency. *Inorg. Chem.* **2020**.

[196] U.S Geological Survey. *Mineral Commodity Summaries 2020*; 2020.

[197] U.S. DOE. Critical Materials Strategy Report. *Cult. Stud.* **2010**, *10* (1).

[198] Cotton, S. *Lanthanide and Actinide Chemistry*; Wiley, 2006.

[199] Hasegawa, Y.; Tamaki, S.; Yajima, H.; Hashimoto, B.; Yaita, T. Selective Separation of Samarium(III) by Synergistic Extraction with β-Diketone and Methylphenylphenanthroline Carboxamide. *Talanta* **2011**, *85* (3), 1543–1548.

[200] Xie, F.; Zhang, T. A.; Dreisinger, D.; Doyle, F. A Critical Review on Solvent Extraction of Rare Earths from Aqueous Solutions. *Miner. Eng.* **2014**, *56*, 10–28.

[201] Hoogerstraete, T. V.; Souza, E. R.; Onghena, B.; Banerjee, D.; Binnemans, K. Mechanism for Solvent Extraction of Lanthanides from Chloride Media by Basic Extractants. *J. Solution Chem.* **2018**, *47*, 1351–1372.

[202] Veliscek-Carolan, J.; Hanley, T. L.; Jolliffe, K. A. The Impact of Structural Variation in Simple Lanthanide Binding Peptides. *RSC Adv.* **2016**, *6*, 75336–75346.

[203] Vassilev, N.; Tosheva, T.; Atanassova, M.; Shenkov, S.; Tashev, E.; Dukov, I.; Lachkova, V.; Varbanov, S. Hexa(1,1,3,3-Tetramethyl-Butyl)-Hexakis(Dimethylphosphinoyl-Methoxy)-Calix[6]Arene: Synthesis, Characterization and Implementation as a Synergistic Agent in the Solvent Extraction of Lanthanoids. *Polyhedron* **2017**, *134*, 135–142.

[204] Le Fur, M.; Molnár, E.; Beyler, M.; Fougère, O.; Esteban-Gómez, D.; Rousseaux, O.; Tripier, R.; Tircsó, G.; Platas-Iglesias, C. Expanding the Family of Pyclen-Based Ligands Bearing Pendant Picolinate Arms for Lanthanide Complexation. *Inorg. Chem.* **2018**, *57*, 6932–6945.

[205] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), 864–871.

[206] Lu, Q.; Peterson, K. A. Correlation Consistent Basis Sets for Lanthanides: The Atoms La-Lu. *J. Chem. Phys.* **2016**, *145* (5).

[207] McCann, B. W.; De Silva, N.; Windus, T. L.; Gordon, M. S.; Moyer, B. A.; Bryantsev, V. S.; Hay, B. P. Computer-Aided Molecular Design of Bis-Phosphine Oxide Lanthanide Extractants. *Inorg. Chem.* **2016**, *55* (12), 5787–5803.

[208] Healy, M. R.; Ivanov, A. S.; Karslyan, Y.; Bryantsev, V. S.; Moyer, B. A.; Jansone-Popova, S. Efficient Separation of Light Lanthanides(III) Using Bis-Lactam Phenthronline Ligands. *Chem. - A Eur. J.* **2019**, *25* (25), 6326–6331.

[209] Ivanov, A. S.; Bryantsev, V. S. A Computational Approach to Predicting Ligand Selectivity for the Size-Based Separation of Trivalent Lanthanides. *Eur. J. Inorg. Chem.* **2016**, *2016* (21), 3474–3479.

[210] Grimmel, S.; Schoendorff, G.; Wilson, A. K. Gauging the Performance of Density Functionals for Lanthanide-Containing Molecules. *J. Chem. Theory Comput.* **2016**, *12* (3), 1259–1266.

[211] Aebersold, L. E.; Yuwono, S. H.; Schoendorff, G.; Wilson, A. K. Efficacy of Density Functionals and Relativistic Effective Core Potentials for Lanthanide-Containing Species: The Ln54 Molecule Set. *J. Chem. Theory Comput.* **2017**, *13* (6), 2831–2839.

[212] Clark, A. E. Density Functional and Basis Set Dependence of Hydrated Ln(III) Properties. *J. Chem. Theory Comput.* **2008**, *4* (5), 708–718.

[213] Vo, M. N.; Bryantsev, V. S.; Johnson, J. K.; Keith, J. A. Quantum Chemistry Benchmarking of Binding and Selectivity for Lanthanide Extractants. *Int. J. Quantum Chem.* **2017**, *118* (7), e25516.

[214] Söderlind, P.; Turchi, P. E. A.; Landa, A.; Lordi, V. Ground-State Properties of Rare-Earth Metals: An Evaluation of Density-Functional Theory. *J. Phys. Condens. Matter* **2014**, *26* (41).

[215] Kuta, J.; Clark, A. E. Trends in Aqueous Hydration across the 4f Period Assessed by Reliable Computational Methods. *Inorg. Chem.* **2010**, *49* (17), 7808–7817.

[216] Ciupka, J.; Cao-Dolg, X.; Wiebke, J.; Dolg, M. Computational Study of Lanthanide(III) Hydration. *Phys. Chem. Chem. Phys.* **2010**, *12* (40), 13215–13223.

[217] Liu, W.; Dolg, M. Benchmark Calculations for Lanthanide Atoms: Calibration of Ab Initio and Density-Functional Methods. *Phys. Rev. A - At. Mol. Opt. Phys.* **1998**, *57* (3), 1721–1728.

[218] Pantazis, D. A.; Neese, F. All-Electron Scalar Relativistic Basis Sets for the Lanthanides. *J. Chem. Theory Comput.* **2009**, *5* (9), 2229–2238.

[219] Aravena, D.; Neese, F.; Pantazis, D. A. Improved Segmented All-Electron Relativistically Contracted Basis Sets for the Lanthanides. *J. Chem. Theory Comput.* **2016**, *12* (3), 1148–1156.

[220] Noro, T.; Sekiya, M.; Koga, T. Sapporo-(DKH3)-NZP (n = D, T, Q) Sets for the Sixth Period s-, d-, and p-Block Atoms. *Theor. Chem. Acc.* **2013**, *132* (5), 1–5.

[221] Gulde, R.; Pollak, P.; Weigend, F. Error-Balanced Segmented Contracted Basis Sets of

Double-ζ to Quadruple-ζ Valence Quality for the Lanthanides. *J. Chem. Theory Comput.* **2012**, *8* (11), 4062–4068.

[222] Cao, X.; Dolg, M. Valence Basis Sets for Relativistic Energy-Consistent Small-Core Actinide Pseudopotentials. *J. Chem. Phys.* **2001**, *115* (16), 7348–7355.

[223] Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P. O. New Relativistic ANO Basis Sets for Transition Metal Atoms. *J. Phys. Chem. A* **2005**, *109* (29), 6575–6579.

[224] Lu, J. B.; Cantu, D. C.; Nguyen, M. T.; Li, J.; Glezakou, V. A.; Rousseau, R. Norm-Conserving Pseudopotentials and Basis Sets to Explore Lanthanide Chemistry in Complex Environments. *J. Chem. Theory Comput.* **2019**, *15*, 5987–5997.

[225] Solomonik, V. G.; Smirnov, A. N. Toward Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy of Lanthanide Compounds: Assessing Core-Valence Correlation, Second-Order Spin-Orbit Coupling, and Higher Order Effects in Lanthanide Diatomics. *J. Chem. Theory Comput.* **2017**, *13* (11), 5240–5254.

[226] Chang, C.; Pelissier, M.; Durand, P. Regular Two-Component Pauli-Like Effective Hamiltonians in Dirac Theory. *Phys. Scr.* **1986**, *34* (5), 394–404.

[227] Cao, X.; Weigand, A. Relativistic Pseudopotentials and Their Applications. In *Computational Methods in Lanthanide and Actinide Chemistry*; 2015; pp 147–179.

[228] Pantazis, D. A.; Neese, F. All-Electron Basis Sets for Heavy Elements. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (4), 363–374.

[229] Hong, G.; Dolg, M.; Li, L. A Comparison of Scalar-Relativistic ZORA and DKH Density Functional Schemes: Monohydrides, Monooxides and Monofluorides of La, Lu, Ac and Lr. *Chem. Phys. Lett.* **2001**, *334*, 396–402.

[230] Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59* (11), 4814–4820.

[231] Slater, J. C.; Johnson, K. H. Self-Consistent-Field Xα Cluster Method for Polyatomic Molecules and Solids. *Phys. Rev. B* **1972**, *5* (3), 844–853.

[232] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211.

[233] Becke, A. D. Density-Functional Exchange Approximation with Correct Asymptotic Behaviour. *Phys. Rev. A* **1988**, *38* (6), 3098–3100.

[234] Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33* (12), 8822–8824.

[235] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

[236] Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.

[237] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

[238] Handy, N. C.; Cohen, A. J. Left-Right Correlation Energy. *Mol. Phys.* **2001**, *99* (5), 403–412.

[239] Hoe, W. M.; Cohen, A. J.; Handy, N. C. Assessment of a New Local Exchange Functional OPTX. *Chem. Phys. Lett.* **2001**, *341* (3–4), 319–328.

[240] Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125* (19), 194101.

[241] Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

[242] Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable

Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170.

[243] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2007**, *120* (1–3), 215–241.

[244] Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119* (23), 12129–12137.

[245] Chai, J. Da; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128* (8), 084106.

[246] Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[247] Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124* (3), 034108.

[248] Kalinowski, J.; Wennmohs, F.; Neese, F. Arbitrary Angular Momentum Electron Repulsion Integrals with Graphical Processing Units: Application to the Resolution of Identity Hartree-Fock Method. *J. Chem. Theory Comput.* **2017**, *13* (7), 3160–3170.

[249] Stoychev, G. L.; Auer, A. A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *J. Chem. Theory Comput.* **2017**, *13* (2), 554–562.

[250] Ames, L. L.; Walsh, P. N.; White, D. Rare Earths. IV. Dissociation Energies of the Gaseous Monoxides of the Rare Earths. *J. Phys. Chem.* **1967**, *71* (8), 2707–2718.

[251] Zmbov, K. F.; Margrave, J. L. Mass Spectrometric Studies of Scandium, Yttrium, Lanthanum, and Rare-Earth Fluorides. **1968**, 267–290.

[252] Fatila, E. M.; Hetherington, E. E.; Jennings, M.; Lough, J.; Preuss, K. E. Syntheses and Crystal Structures of Anhydrous Ln(Hfac)$_3$(Monoglyme). Ln = La, Ce, Pr, Sm, Eu, Gd,

Tb, Dy, Er, Tm. *Dalt. Trans.* **2012**, *41*, 1352–1362.

[253] Behrsing, T.; Deacon, G. B.; Luu, J.; Junk, P. C.; Skelton, B. W.; White, A. H. Structural Diversity of Lanthanoid Salicylate Hydrates. *Polyhedron* **2016**, *120*, 69–81.

[254] Drew, M. G. B.; Foreman, M. R. S.; Hudson, M. J.; Kennedy, K. F. Structural Studies of Lanthanide Complexes with Tetradentate Nitrogen Ligands. *Inorg. Chim. Acta* **2004**, *357* (14), 4102–4112.

[255] Peterson, C. C.; Penchoff, D. A.; Auxier, J. D.; Hall, H. L. Establishing Cost-Effective Computational Models for the Prediction of Lanthanoid Binding in $[Ln(NO_3)]2+$ (with Ln = La to Lu). *ACS Omega* **2019**, *4* (1), 1375–1385.

[256] Penchoff, D. A.; Peterson, C. C.; Quint, M. S.; Auxier, J. D.; Schweitzer, G. K.; Jenkins, D. M.; Harrison, R. J.; Hall, H. L. Structural Characteristics, Population Analysis, and Binding Energies of $[An(NO_3)]+2$ (with An = Ac to Lr). *ACS Omega* **2018**, *3* (10), 14127–14143.

[257] Lavrov, H. V.; Ustynyuk, N. A.; Matveev, P. I.; Gloriozov, I. P.; Zhokhov, S. S.; Alyapyshev, M. Y.; Tkachenko, L. I.; Voronaev, I. G.; Babain, V. A.; Kalmykov, S. N.; Ustynyuk, Y. A. A Novel Highly Selective Ligand for Separation of Actinides and Lanthanides in the Nuclear Fuel Cycle. Experimental Verification of the Theoretical Prediction. *Dalt. Trans.* **2017**, *46* (33), 10926–10934.

[258] Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102* (11), 1995–2001.

[259] Sun, T.; Xu, C.; Xie, X.; Chen, J.; Liu, X. Quantum Chemistry Study on the Extraction of Trivalent Lanthanide Series by Cyanex301: Insights from Formation of Inner- and Outer-Sphere Complexes. *ACS Omega* **2018**, *3* (4), 4070–4080.

[260] Waller, M. P.; Braun, H.; Hojdis, N.; Bühl, M. Geometries of Second-Row Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2007**, *3* (6), 2234–2242.

[261] Cirera, J.; Via-Nadal, M.; Ruiz, E. Benchmarking Density Functional Methods for Calculation of State Energies of First Row Spin-Crossover Molecules. *Inorg. Chem.* **2018**, *57* (22), 14097–14105.

[262] Huang, P. W. Understanding the Stability Trend Along Light Lanthanide Complexes with an Ehtylenediamine-Type Ligand: A Quantum Chemical Study. *ChemistrySelect* **2019**, *4* (42), 12368–12374.

[263] Wilson, A. M.; Bailey, P. J.; Tasker, P. A.; Turkington, J. R.; Grant, R. A.; Love, J. B. Solvent Extraction: The Coordination Chemistry behind Extractive Metallurgy. *Chem. Soc. Rev.* **2014**, *43* (1), 123–134.

[264] Llanos, E. J.; Leal, W.; Luu, D. H.; Jost, J.; Stadler, P. F.; Restrepo, G. Exploration of the Chemical Space and Its Three Historical Regimes. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (29), 14779.

[265] Gromski, P. S.; Henson, A. B.; Granda, J. M.; Cronin, L. How to Explore Chemical Space Using Algorithms and Automation. *Nat. Rev. Chem.* **2019**, *3* (2), 119–128.

[266] Solov'ev, V. P.; Ustynyuk, Y. A.; Zhokhova, N. I.; Karpov, K. V. Predictive Models for HOMO and LUMO Energies of N-Donor Heterocycles as Ligands for Lanthanides Separation. *Mol. Inform.* **2018**, *37* (11), e1800025.

[267] Chaube, S.; Goverapet Srinivasan, S.; Rai, B. Applied Machine Learning for Predicting the Lanthanide-Ligand Binding Affinities. *Sci. Rep.* **2020**, *10* (1), 1–11.

[268] Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K. R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights into Chemical Systems. *Chem. Rev.* **2021**, *121* (16), 9816–9872.

[269] Townsend, J.; Vogiatzis, K. D. Data-Driven Acceleration of the Coupled-Cluster Singles and Doubles Iterative Solver. *J. Phys. Chem. Lett.* **2019**, *10* (14), 4129–4135.

[270] Townsend, J.; Micucci, C. P.; Hymel, J. H.; Maroulas, V.; Vogiatzis, K. D. Representation of Molecular Structures with Persistent Homology for Machine Learning Applications in

Chemistry. *Nat. Commun.* **2020**, *11* (1), 1–9.

[271] Umetani, S.; Kawase, Y.; Le, Q. T. H.; Matsui, M. Acylpyrazolone Derivatives of High Selectivity for Lanthanide Metal Ions: Effect of the Distance between the Two Donating Oxygens. *J. Chem. Soc. Dalt. Trans.* **2000**, *15* (16), 2787–2791.

[272] Le, Q. T. H.; Umetani, S.; Suzuki, M.; Matsui, M. α-Substituted β-Diketones: Effect of the α Substituent on the Complexation and Selectivity for Lanthanides. *J. Chem. Soc. Dalt. Trans.* **2002**, *56* (4), 643–648.

[273] Ioannidis, E. I.; Gani, T. Z. H.; Kulik, H. J. MolSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016**, 2106–2117.

[274] Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020**, *22* (14), 7169–7192.

[275] Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[276] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15).

[277] Becke, A. D.; Johnson, E. R. A Density-Functional Model of the Dispersion Interaction. *J. Chem. Phys.* **2005**, *123* (15).

[278] Murphy, K. P. *Machine Learning A Probabilistic Perspective*; 2012.

[279] Ji Zhu, Hui Zou, Saharon Rosset, T. H. Multi-Class AdaBoost. *Stat. Interface* **2009**, *2*, 349–360.

[280] Freund, Y.; Schapire, R. E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55* (1), 119–139.

[281] Friedman, J. H. Stochastic Gradient Boosting. *Comput. Stat. Data Anal.* **2002**, *38* (4), 367–378.

[282] Varoquaux, G.; Buitinck, L.; Louppe, G.; Grisel, O.; Pedregosa, F.; Mueller, A. Scikit-Learn. *GetMobile Mob. Comput. Commun.* **2015**, *19* (1), 29–33.

[283] Rupp, M.; Tkatchenko, A.; Müller, K. R.; Von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, *108* (5), 1–5.

[284] Hansen, K.; Biegler, F.; Ramakrishnan, R.; Pronobis, W.; Von Lilienfeld, O. A.; Müller, K. R.; Tkatchenko, A. Machine Learning Predictions of Molecular Properties: Accurate Many-Body Potentials and Nonlocality in Chemical Space. *J. Phys. Chem. Lett.* **2015**, *6* (12), 2326–2331.

[285] Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2013**, *87* (18), 1–16.

[286] Huo, H.; Rupp, M. Unified Representation of Molecules and Crystals for Machine Learning. **2017**, No. i.

[287] Himanen, L.; Jäger, M. O. J.; Morooka, E. V.; Federici Canova, F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. DScribe: Library of Descriptors for Machine Learning in Materials Science. *Comput. Phys. Commun.* **2020**, *247*, 106949.

[288] Tralie, C.; Saul, N.; Bar-On, R. Ripser.Py: A Lean Persistent Homology Library for Python. *J. Open Source Softw.* **2018**, *3* (29), 925.

[289] Townsend, J.; Micucci, C. P.; Hymel, J. H.; Maroulas, V.; Vogiatzis, K. D. Persistent Images for Chemistry https://maroulaslab.github.io/PersistentImages_Chemistry/index.html.

[290] Liu, L.; Corma, A. Metal Catalysts for Heterogeneous Catalysis: From Single Atoms to Nanoclusters and Nanoparticles. *Chem. Rev.* **2018**, *118* (10), 4981–5079.

[291] Chepaikin, E. G. Oxidative Functionalization of Alkanes under Dioxygen in the Presence of Homogeneous Noble Metal Catalysts. *J. Mol. Catal. A Chem.* **2014**, *385*, 160–174.

[292] Stock, N.; Biswas, S. Synthesis of Metal-Organic Frameworks (MOFs): Routes to Various

MOF Topologies, Morphologies, and Composites. *Chem. Rev.* **2012**, *112* (2), 933–969.

[293] Zhu, J.; Hu, L.; Zhao, P.; Lee, L. Y. S.; Wong, K. Y. Recent Advances in Electrocatalytic Hydrogen Evolution Using Nanoparticles. *Chem. Rev.* **2020**, *120* (2), 851–918.

[294] Shakya, D. M.; Ejegbavwo, O. A.; Rajeshkumar, T.; Senanayake, S. D.; Brandt, A. J.; Farzandh, S.; Acharya, N.; Ebrahim, A. M.; Frenkel, A. I.; Rui, N.; Tate, G. L.; Monnier, J. R.; Vogiatzis, K. D.; Shustova, N. B.; Chen, D. A. Selective Catalytic Chemistry at Rhodium(II) Nodes in Bimetallic Metal-Organic Frameworks. *Angew. Chemie Int. Ed.* **2019**, *58*, 1–6.

[295] Mattson, B.; Foster, W.; Greimann, J.; Hoette, T.; Le, N.; Mirich, A.; Wankum, S.; Cabri, A.; Reichenbacher, C.; Schwanke, E. Heterogeneous Catalysis: The Horiuti-Polanyi Mechanism and Alkene Hydrogenation. *J. Chem. Educ.* **2013**, *90* (5), 613–619.

[296] Fukuzumi, S.; Yamada, Y.; Suenobu, T.; Ohkubo, K.; Kotani, H. Catalytic Mechanisms of Hydrogen Evolution with Homogeneous and Heterogeneous Catalysts. *Energy Environ. Sci.* **2011**, *4* (8), 2754–2766.

[297] Cui, X.; Li, W.; Ryabchuk, P.; Junge, K.; Beller, M. Bridging Homogeneous and Heterogeneous Catalysis by Heterogeneous Single-Metal-Site Catalysts. *Nat. Catal.* **2018**, *1* (6), 385–397.

[298] Collis, A. E. C.; Horváth, I. T. Heterogenization of Homogeneous Catalytic Systems. *Catal. Sci. Technol.* **2011**, *1* (6), 912–919.

[299] Yang, D.; Momeni, M. R.; Demir, H.; Pahls, D. R.; Rimoldi, M.; Wang, T. C.; Farha, O. K.; Hupp, J. T.; Cramer, C. J.; Gates, B. C.; Gagliardi, L. Tuning the Properties of Metal-Organic Framework Nodes as Supports of Single-Site Iridium Catalysts: Node Modification by Atomic Layer Deposition of Aluminium. *Faraday Discuss.* **2017**, *201*, 195–206.

[300] Desai, S. P.; Ye, J.; Zheng, J.; Ferrandon, M. S.; Webber, T. E.; Platero-Prats, A. E.; Duan, J.; Garcia-Holley, P.; Camaioni, D. M.; Chapman, K. W.; Delferro, M.; Farha, O. K.; Fulton, J. L.; Gagliardi, L.; Lercher, J. A.; Penn, R. L.; Stein, A.; Lu, C. C. Well-

Defined Rhodium-Gallium Catalytic Sites in a Metal-Organic Framework: Promoter-Controlled Selectivity in Alkyne Semihydrogenation to E-Alkenes. *J. Am. Chem. Soc.* **2018**, *140* (45), 15309–15318.

[301] Bernales, V.; Yang, D.; Yu, J.; Gümüşlu, G.; Cramer, C. J.; Gates, B. C.; Gagliardi, L. Molecular Rhodium Complexes Supported on the Metal-Oxide-Like Nodes of Metal Organic Frameworks and on Zeolite HY: Catalysts for Ethylene Hydrogenation and Dimerization. *ACS Appl. Mater. Interfaces* **2017**, *9* (39), 33511–33520.

[302] Jones, D. R.; Iqbal, S.; Kondrat, S. A.; Lari, G. M.; Miedziak, P. J.; Morgan, D. J.; Parker, S. F.; Hutchings, G. J. An Investigation of the Effect of Carbon Support on Ruthenium/Carbon Catalysts for Lactic Acid and Butanone Hydrogenation. *Phys. Chem. Chem. Phys.* **2016**, *18* (26), 17259–17264.

[303] Li, Z.; Peters, A. W.; Bernales, V.; Ortuño, M. A.; Schweitzer, N. M.; Destefano, M. R.; Gallington, L. C.; Platero-Prats, A. E.; Chapman, K. W.; Cramer, C. J.; Gagliardi, L.; Hupp, J. T.; Farha, O. K. Metal-Organic Framework Supported Cobalt Catalysts for the Oxidative Dehydrogenation of Propane at Low Temperature. *ACS Cent. Sci.* **2017**, *3* (1), 31–38.

[304] Ortuño, M. A.; Bernales, V.; Gagliardi, L.; Cramer, C. J. Computational Study of First-Row Transition Metals Supported on MOF NU-1000 for Catalytic Acceptorless Alcohol Dehydrogenation. *J. Phys. Chem. C* **2016**, *120* (43), 24697–24705.

[305] Barona, M.; Ahn, S.; Morris, W.; Hoover, W.; Notestein, J. M.; Farha, O. K.; Snurr, R. Q. Computational Predictions and Experimental Validation of Alkane Oxidative Dehydrogenation by Fe_2M MOF Nodes. *ACS Catal.* **2020**, *10* (2), 1460–1469.

[306] Paluka, V.; Maihom, T.; Probst, M.; Limtrakul, J. Dehydrogenation of Ethanol to Acetaldehyde with Nitrous Oxide over the Metal–Organic Framework NU-1000: A Density Functional Theory Study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13622–13628.

[307] Bernales, V.; League, A. B.; Li, Z.; Schweitzer, N. M.; Peters, A. W.; Carlson, R. K.; Hupp, J. T.; Cramer, C. J.; Farha, O. K.; Gagliardi, L. Computationally Guided Discovery

of a Catalytic Cobalt-Decorated Metal-Organic Framework for Ethylene Dimerization. *J. Phys. Chem. C* **2016**, *120* (41), 23576–23583.

[308] Liu, J.; Ye, J.; Li, Z.; Otake, K. I.; Liao, Y.; Peters, A. W.; Noh, H.; Truhlar, D. G.; Gagliardi, L.; Cramer, C. J.; Farha, O. K.; Hupp, J. T. Beyond the Active Site: Tuning the Activity and Selectivity of a Metal-Organic Framework-Supported Ni Catalyst for Ethylene Dimerization. *J. Am. Chem. Soc.* **2018**, *140* (36), 11174–11178.

[309] Pahls, D. R.; Ortuño, M. A.; Winegar, P. H.; Cramer, C. J.; Gagliardi, L. Computational Screening of Bimetal-Functionalized Zr_6O_8 MOF Nodes for Methane C-H Bond Activation. *Inorg. Chem.* **2017**, *56* (15), 8739–8743.

[310] Maihom, T.; Choomwattana, S.; Khongpracha, P.; Probst, M.; Limtrakul, J. Formaldehyde Encapsulated in Lithium-Decorated Metal-Organic Frameworks: A Density Functional Theory Study. *ChemPhysChem* **2012**, *13* (1), 245–249.

[311] Mondloch, J. E.; Katz, M. J.; Isley, W. C.; Ghosh, P.; Liao, P.; Bury, W.; Wagner, G. W.; Hall, M. G.; Decoste, J. B.; Peterson, G. W.; Snurr, R. Q.; Cramer, C. J.; Hupp, J. T.; Farha, O. K. Destruction of Chemical Warfare Agents Using Metal-Organic Frameworks. *Nat. Mater.* **2015**, *14* (5), 512–516.

[312] Chen, H.; Liao, P.; Mendonca, M. L.; Snurr, R. Q. Insights into Catalytic Hydrolysis of Organophosphate Warfare Agents by Metal-Organic Framework NU-1000. *J. Phys. Chem. C* **2018**, *122* (23), 12362–12368.

[313] Troya, D. Reaction Mechanism of Nerve-Agent Decomposition with Zr-Based Metal Organic Frameworks. *J. Phys. Chem. C* **2016**, *120* (51), 29312–29323.

[314] Wang, G.; Sharp, C.; Plonka, A. M.; Wang, Q.; Frenkel, A. I.; Guo, W.; Hill, C.; Smith, C.; Kollar, J.; Troya, D.; Morris, J. R. Mechanism and Kinetics for Reaction of the Chemical Warfare Agent Simulant, DMMP(g), with Zirconium(IV) MOFs: An Ultrahigh-Vacuum and DFT Study. *J. Phys. Chem. C* **2017**, *121* (21), 11261–11272.

[315] Momeni, M. R.; Cramer, C. J. Dual Role of Water in Heterogeneous Catalytic Hydrolysis of Sarin by Zirconium-Based Metal-Organic Frameworks. *ACS Appl. Mater. Interfaces*

2018, *10* (22), 18435–18439.

[316] Momeni, M. R.; Cramer, C. J. Structural Characterization of Pristine and Defective $[Zr_{12}(\mu_3\text{-}O)_8(\mu_3\text{-}OH)_8(\mu_2\text{-}OH)_6]^{18+}$ Double-Node Metal-Organic Framework and Predicted Applications for Single-Site C. *Chem. Mater.* **2018**, *30* (13), 4432–4439.

[317] Islamoglu, T.; Ortuño, M. A.; Proussaloglou, E.; Howarth, A. J.; Vermeulen, N. A.; Atilgan, A.; Asiri, A. M.; Cramer, C. J.; Farha, O. K. Presence versus Proximity: The Role of Pendant Amines in the Catalytic Hydrolysis of a Nerve Agent Simulant. *Angew. Chemie Int. Ed.* **2018**, *57* (7), 1967–1971.

[318] Yang, D.; Ortuño, M. A.; Bernales, V.; Cramer, C. J.; Gagliardi, L.; Gates, B. C. Structure and Dynamics of Zr_6O_8 Metal-Organic Framework Node Surfaces Probed with Ethanol Dehydration as a Catalytic Test Reaction. *J. Am. Chem. Soc.* **2018**, *140* (10), 3751–3759.

[319] Svane, K. L.; Bristow, J. K.; Gale, J. D.; Walsh, A. Vacancy Defect Configurations in the Metal-Organic Framework UiO-66: Energetics and Electronic Structure. *J. Mater. Chem. A* **2018**, *6* (18), 8507–8513.

[320] Platero-Prats, A. E.; Mavrandonakis, A.; Gallington, L. C.; Liu, Y.; Hupp, J. T.; Farha, O. K.; Cramer, C. J.; Chapman, K. W. Structural Transitions of the Metal-Oxide Nodes within Metal-Organic Frameworks: On the Local Structures of NU-1000 and UiO-66. *J. Am. Chem. Soc.* **2016**, *138* (12), 4178–4185.

[321] Xiao, D. J.; Bloch, E. D.; Mason, J. A.; Queen, W. L.; Hudson, M. R.; Planas, N.; Borycz, J.; Dzubak, A. L.; Verma, P.; Lee, K.; Bonino, F.; Crocellà, V.; Yano, J.; Bordiga, S.; Truhlar, D. G.; Gagliardi, L.; Brown, C. M.; Long, J. R. Oxidation of Ethane to Ethanol by N_2O in a Metal-Organic Framework with Coordinatively Unsaturated Iron(II) Sites. *Nat. Chem.* **2014**, *6* (7), 590–595.

[322] Vitillo, J. G.; Bhan, A.; Cramer, C. J.; Lu, C. C.; Gagliardi, L. Quantum Chemical Characterization of Structural Single Fe(II) Sites in MIL-Type Metal-Organic Frameworks for the Oxidation of Methane to Methanol and Ethane to Ethanol. *ACS Catal.* **2019**, *9* (4), 2870–2879.

[323] Vogiatzis, K. D.; Klopper, W.; Mavrandonakis, A.; Fink, K. Magnetic Properties of Paddlewheels and Trinuclear Clusters with Exposed Metal Sites. *ChemPhysChem* **2011**, *12* (17), 3307–3319.

[324] Simons, M. C.; Vitillo, J. G.; Babucci, M.; Hoffman, A. S.; Boubnov, A.; Beauvais, M. L.; Chen, Z.; Cramer, C. J.; Chapman, K. W.; Bare, S. R.; Gates, B. C.; Lu, C. C.; Gagliardi, L.; Bhan, A. Structure, Dynamics, and Reactivity for Light Alkane Oxidation of Fe(II) Sites Situated in the Nodes of a Metal-Organic Framework. *J. Am. Chem. Soc.* **2019**, *141* (45), 18142–18151.

[325] Simons, M. C.; Ortuño, M. A.; Bernales, V.; Gaggioli, C. A.; Cramer, C. J.; Bhan, A.; Gagliardi, L. C-H Bond Activation on Bimetallic Two-Atom Co-M Oxide Clusters Deposited on Zr-Based MOF Nodes: Effects of Doping at the Molecular Level. *ACS Catal.* **2018**, *8* (4), 2864–2869.

[326] Latimer, A. A.; Kulkarni, A. R.; Aljama, H.; Montoya, J. H.; Yoo, J. S.; Tsai, C.; Abild-Pedersen, F.; Studt, F.; Nørskov, J. K. Understanding Trends in C-H Bond Activation in Heterogeneous Catalysis. *Nat. Mater.* **2017**, *16* (2), 225–229.

[327] Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Structure-Activity Relationships That Identify Metal-Organic Framework Catalysts for Methane Activation. *ACS Catal.* **2019**, *9* (4), 3576–3587.

[328] Impeng, S.; Siwaipram, S.; Bureekaew, S.; Probst, M. Ethane C-H Bond Activation on the Fe(IV)-Oxo Species in a Zn-Based Cluster of Metal-Organic Frameworks: A Density Functional Theory Study. *Phys. Chem. Chem. Phys.* **2017**, *19* (5), 3782–3791.

[329] Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. High-Valent Metal–Oxo Species at the Nodes of Metal–Triazolate Frameworks: The Effects of Ligand Exchange and Two-State Reactivity for C−H Bond Activation. *Angew. Chemie Int. Ed.* **2020**, *59*, 2–11.

[330] Ikuno, T.; Zheng, J.; Vjunov, A.; Sanchez-Sanchez, M.; Ortuño, M. A.; Pahls, D. R.; Fulton, J. L.; Camaioni, D. M.; Li, Z.; Ray, D.; Mehdi, B. L.; Browning, N. D.; Farha, O. K.; Hupp, J. T.; Cramer, C. J.; Gagliardi, L.; Lercher, J. A. Methane Oxidation to

Methanol Catalyzed by Cu-Oxo Clusters Stabilized in NU-1000 Metal-Organic Framework. *J. Am. Chem. Soc.* **2017**, *139* (30), 10294–10301.

[331] Zheng, J.; Ye, J.; Ortuño, M. A.; Fulton, J. L.; Gutiérrez, O. Y.; Camaioni, D. M.; Motkuri, R. K.; Li, Z.; Webber, T. E.; Mehdi, B. L.; Browning, N. D.; Penn, R. L.; Farha, O. K.; Hupp, J. T.; Truhlar, D. G.; Cramer, C. J.; Lercher, J. A. Selective Methane Oxidation to Methanol on Cu-Oxo Dimers Stabilized by Zirconia Nodes of an NU-1000 Metal-Organic Framework. *J. Am. Chem. Soc.* **2019**, *141* (23), 9292–9304.

[332] Verma, P.; Vogiatzis, K. D.; Planas, N.; Borycz, J.; Xiao, D. J.; Long, J. R.; Gagliardi, L.; Truhlar, D. G. Mechanism of Oxidation of Ethane to Ethanol at Iron(IV)−Oxo Sites in Magnesium-Diluted Fe_2(Dobdc). *J. Am. Chem. Soc.* **2015**, *137* (17), 5770–5781.

[333] Liu, Y. Y.; Leus, K.; Bogaerts, T.; Hemelsoet, K.; Bruneel, E.; Van Speybroeck, V.; Van Der Voort, P. Bimetallic–Organic Framework as a Zero-Leaching Catalyst in the Aerobic Oxidation of Cyclohexene. *ChemCatChem* **2013**, *5* (12), 3657–3664.

[334] Baek, J.; Rungtaweevoranit, B.; Pei, X.; Park, M.; Fakra, S. C.; Liu, Y. S.; Matheu, R.; Alshmimri, S. A.; Alshehri, S.; Trickett, C. A.; Somorjai, G. A.; Yaghi, O. M. Bioinspired Metal-Organic Framework Catalysts for Selective Methane Oxidation to Methanol. *J. Am. Chem. Soc.* **2018**, *140* (51), 18208–18216.

[335] Vogiatzis, K. D.; Haldoupis, E.; Xiao, D. J.; Long, J. R.; Siepmann, J. I.; Gagliardi, L. Accelerated Computational Analysis of Metal-Organic Frameworks for Oxidation Catalysis. *J. Phys. Chem. C* **2016**, *120* (33), 18707–18712.

[336] Liao, P.; Getman, R. B.; Snurr, R. Q. Optimizing Open Iron Sites in Metal−Organic Frameworks for Ethane Oxidation: A First-Principles Study. *ACS Appl. Mater. Interfaces* **2017**, *9* (39), 33484–33492.

[337] Kazaryan, A.; Baerends, E. J. Ligand Field Effects and the High Spin−High Reactivity Correlation in the H Abstraction by Non-Heme Iron(IV)−Oxo Complexes: A DFT Frontier Orbital Perspective. *ACS Catal.* **2015**, *5* (3), 1475–1488.

[338] Suh, B. L.; Kim, J. Ligand Insertion in MOF-74 as Effective Design for Oxidation of

Ethane to Ethanol. *J. Phys. Chem. C* **2018**, *122* (40), 23078–23083.

[339] Luo, W.; Chen, G.; Xiao, S.; Wang, Q.; Huang, Z.; Wang, L. The Enzyme-like Catalytic Hydrogen Abstraction Reaction Mechanisms of Cyclic Hydrocarbons with Magnesium-Diluted Fe-MOF-74. *RSC Adv.* **2019**, *9* (41), 23622–23632.

[340] Szécsényi, A.; Li, G.; Gascon, J.; Pidko, E. A. Unraveling Reaction Networks behind the Catalytic Oxidation of Methane with H_2O_2 over a Mixed-Metal MIL-53(Al,Fe) MOF Catalyst. *Chem. Sci.* **2018**, *9* (33), 6765–6773.

[341] Osadchii, D. Y.; Olivos-Suarez, A. I.; Szécsényi, Á.; Li, G.; Nasalevich, M. A.; Dugulan, I. A.; Crespo, P. S.; Hensen, E. J. M.; Veber, S. L.; Fedin, M. V.; Sankar, G.; Pidko, E. A.; Gascon, J. Isolated Fe Sites in Metal Organic Frameworks Catalyze the Direct Conversion of Methane to Methanol. *ACS Catal.* **2018**, *8* (6), 5542–5548.

[342] Noh, H.; Cui, Y.; Peters, A. W.; Pahls, D. R.; Ortuno, M. A.; Vermeulen, N. A.; Cramer, C. J.; Gagliardi, L.; Hupp, J. T.; Farha, O. K. An Exceptionally Stable Metal-Organic Framework Supported Molybdenum(VI) Oxide Catalyst for Cyclohexene Epoxidation. *J. Am. Chem. Soc.* **2016**, *138* (44), 14720–14726.

[343] Maihom, T.; Choomwattana, S.; Wannakao, S.; Probst, M.; Limtrakul, J. Ethylene Epoxidation with Nitrous Oxide over Fe–BTC Metal–Organic Frameworks: A DFT Study. *ChemPhysChem* **2016**, *17* (21), 3416–3422.

[344] Maihom, T.; Sawangphruk, M.; Probst, M.; Limtrakul, J. A Computational Study of the Catalytic Aerobic Epoxidation of Propylene over the Coordinatively Unsaturated Metal-Organic Framework $Fe_3(Btc)_2$: Formation of Propylene Oxide and Competing Reactions. *Phys. Chem. Chem. Phys.* **2018**, *20* (9), 6726–6734.

[345] Ketrat, S.; Maihom, T.; Wannakao, S.; Probst, M.; Nokbin, S.; Limtrakul, J. Coordinatively Unsaturated Metal-Organic Frameworks $M_3(Btc)_2$ (M = Cr, Fe, Co, Ni, Cu, and Zn) Catalyzing the Oxidation of CO by N_2O: Insight from DFT Calculations. *Inorg. Chem.* **2017**, *56* (22), 14005–14012.

[346] Irving, H.; Williams, R. J. P. The Stability of Transition-Metal Complexes. *J. Chem. Soc.*

1953, No. 0, 3192–3210.

[347] Wannakao, S.; Maihom, T.; Probst, M.; Limtrakul, J.; Kongpatpanich, K. Porous Materials as a Platform for Highly Uniform Single-Atom Catalysts: Tuning the Electronic Structure for the Low-Temperature Oxidation of Carbon Monoxide. *J. Phys. Chem. C* **2016**, *120* (35), 19686–19697.

[348] Abdel-Mageed, A. M.; Rungtaweevoranit, B.; Parlinska-Wojtan, M.; Pei, X.; Yaghi, O. M.; Jürgen Behm, R. Highly Active and Stable Single-Atom Cu Catalysts Supported by a Metal-Organic Framework. *J. Am. Chem. Soc.* **2019**, *141* (13), 5201–5210.

[349] Tong, Y.; Xue, G.; Wang, H.; Liu, M.; Wang, J.; Hao, C.; Zhang, X.; Wang, D.; Shi, X.; Liu, W.; Li, G.; Tang, Z. Interfacial Coupling between Noble Metal Nanoparticles and Metal-Organic Frameworks for Enhanced Catalytic Activity. *Nanoscale* **2018**, *10* (35), 16425–16430.

[350] Piscopo, C. G.; Voellinger, L.; Schwarzer, M.; Polyzoidis, A.; Bošković, D.; Loebbecke, S. Continuous Flow Desulfurization of a Model Fuel Catalysed by Titanium Functionalized UiO-66. *ChemistrySelect* **2019**, *4* (9), 2806–2809.

[351] Limvorapitux, R.; Chen, H.; Mendonca, M. L.; Liu, M.; Snurr, R. Q.; Nguyen, S. T. Elucidating the Mechanism of the UiO-66-Catalyzed Sulfide Oxidation: Activity and Selectivity Enhancements through Changes in the Node Coordination Environment and Solvent. *Catal. Sci. Technol.* **2019**, *9* (2), 327–335.

[352] Paille, G.; Gomez-Mingot, M.; Roch-Marchal, C.; Lassalle-Kaiser, B.; Mialane, P.; Fontecave, M.; Mellot-Draznieks, C.; Dolbecq, A. A Fully Noble Metal-Free Photosystem Based on Cobalt-Polyoxometalates Immobilized in a Porphyrinic Metal-Organic Framework for Water Oxidation. *J. Am. Chem. Soc.* **2018**, *140* (10), 3613–3618.

[353] Li, F. L.; Wang, P.; Huang, X.; Young, D. J.; Wang, H. F.; Braunstein, P.; Lang, J. P. Large-Scale, Bottom-Up Synthesis of Binary Metal–Organic Framework Nanosheets for Efficient Water Oxidation. *Angew. Chemie - Int. Ed.* **2019**, *58* (21), 7051–7056.

[354] Vitillo, J. G.; Crocellà, V.; Bonino, F. ZIF-8 as a Catalyst in Ethylene Oxide and

Propylene Oxide Reaction with CO2 to Cyclic Organic Carbonates. *ChemEngineering* **2019**, *3* (3), 60.

[355] Ye, J.; Li, L.; Johnson, J. K. The Effect of Topology in Lewis Pair Functionalized Metal Organic Frameworks on CO_2 Adsorption and Hydrogenation. *Catal. Sci. Technol.* **2018**, *8* (18), 4609–4617.

[356] Ye, J.; Johnson, J. K. Design of Lewis Pair-Functionalized Metal Organic Frameworks for CO_2 Hydrogenation. *ACS Catal.* **2015**, *5* (5), 2921–2928.

[357] Babu, R.; Roshan, R.; Gim, Y.; Jang, Y. H.; Kurisingal, J. F.; Kim, D. W.; Park, D. W. Inverse Relationship of Dimensionality and Catalytic Activity in CO_2 Transformation: A Systematic Investigation by Comparing Multidimensional Metal-Organic Frameworks. *J. Mater. Chem. A* **2017**, *5* (30), 15961–15969.

[358] Dey, S.; Mondal, B.; Chatterjee, S.; Rana, A.; Amanullah, S.; Dey, A. Molecular Electrocatalysts for the Oxygen Reduction Reaction. *Nat. Rev. Chem.* **2017**, *1* (12).

[359] Jiang, J.; Materna, K. L.; Hedström, S.; Yang, K. R.; Crabtree, R. H.; Batista, V. S.; Brudvig, G. W. Antimony Complexes for Electrocatalysis: Activity of a Main-Group Element in Proton Reduction. *Angew. Chemie* **2017**, *129* (31), 9239–9243.

[360] Chen, Z.; Liu, Y.; Wei, W.; Ni, B. J. Recent Advances in Electrocatalysts for Halogenated Organic Pollutant Degradation. *Environ. Sci. Nano* **2019**, *6* (8), 2332–2366.

[361] Fache, E.; Mercier, C.; Pagnier, N.; Despeyroux, B.; Panster, P. Selective Hydrogenation of α,β-Unsaturated Aldehydes Catalyzed by Supported Aqueous-Phase Catalysts and Supported Homogeneous Catalysts. *J. Mol. Catal.* **1993**, *79* (1–3), 117–131.

[362] Riisager, A.; Jørgensen, B.; Wasserscheid, P.; Fehrmann, R. First Application of Supported Ionic Liquid Phase (SILP) Catalysis for Continuous Methanol Carbonylation. *Chem. Commun.* **2006**, No. 9, 994–996.

[363] Williams, C. K.; McCarver, G. A.; Lashgari, A.; Vogiatzis, K. D.; Jiang, J. J. Electrocatalytic Dechlorination of Dichloromethane in Water Using a Heterogenized

Molecular Copper Complex. *Inorg. Chem.* **2021**, *60* (7), 4915–4923.

[364] Schlosser, P. M.; Bale, A. S.; Gibbons, C. F.; Wilkins, A.; Cooper, G. S. Human Health Effects of Dichloromethane: Key Findings and Scientific Issues. *Environ. Health Perspect.* **2015**, *123* (2), 114–119.

[365] Martin, E. T.; McGuire, C. M.; Mubarak, M. S.; Peters, D. G. Electroreductive Remediation of Halogenated Environmental Pollutants. *Chem. Rev.* **2016**, *116* (24), 15198–15234.

[366] Shestakova, M.; Sillanpää, M. Removal of Dichloromethane from Ground and Wastewater: A Review. *Chemosphere* **2013**, *93* (7), 1258–1267.

[367] Sonoyama, N.; Ezaki, K.; Sakata, T. Continuous Electrochemical Decomposition of Dichloromethane in Aqueous Solution Using Various Column Electrodes. *Adv. Environ. Res.* **2001**, *6* (1), 1–8.

[368] Kotsinaris, A.; Kyriacou, G.; Lambrou, C. Electrochemical Reduction of Dichloromethane to Higher Hydrocarbons. *J. Appl. Electrochem.* **1998**, *28* (6), 613–616.

[369] Yang, L.; Chen, Z.; Ma, T.; Zhang, S.; Dai, W.; Xiao, X.; Luo, X.; Zou, J.; Tu, X.; Yang, L.; Luo, S. Efficient Electrochemical Dehalogenation of Florfenicol without Discharging Toxic Intermediates via Direct Electron Transfer over Electrochromic WO_3. *Chem. Eng. J.* **2021**, *412* (September 2020), 127481.

[370] Escobedo, E.; Kim, J.; Oh, D.; Cho, K.; Chang, Y. S. Electrocatalytic Dehalogenation of Aqueous Pollutants by Dealloyed Nanoporous Pd/Ti Cathode. *Catal. Today* **2021**, *361* (September 2019), 63–68.

[371] Tsyganok, A. I.; Yamanaka, I.; Otsuka, K. Electrocatalytic Dehalogenation of Chloroaromatics on Palladium-Loaded Carbon Felt Cathode in Aqueous Medium. *Chem. Lett.* **1998**.

[372] Souza, F. D.; Choi, J.; Kutner, W. Electrocatalytic Dehalogenation of 1,2-Dihaloethanes by the C_{60}, C_{70}, C_{76}, C_{78}, and C_{84} Fullerene Anions: Structure-Reactivity Aspects. *Society*

1999, 2892–2896.

[373] Wirtz, M.; Klucik, J.; Rivera, M. Ferredoxin-Mediated Electrocatalytic Dehalogenation of Haloalkanes by Cytochrome P450(Cam). *J. Am. Chem. Soc.* **2000**, *122* (6), 1047–1056.

[374] Puente Santiago, A. R.; Sanad, M. F.; Moreno-Vicente, A.; Ahsan, M. A.; Cerón, M. R.; Yao, Y. R.; Sreenivasan, S. T.; Rodriguez-Fortea, A.; Poblet, J. M.; Echegoyen, L. A New Class of Molecular Electrocatalysts for Hydrogen Evolution: Catalytic Activity of $M_3N@C_{2n}$ (2n = 68, 78, and 80) Fullerenes. *J. Am. Chem. Soc.* **2021**, *143* (16), 6037–6042.

[375] Bullock, R. M.; Helm, M. L. Molecular Electrocatalysts for Oxidation of Hydrogen Using Earth-Abundant Metals: Shoving Protons Around with Proton Relays. *Acc. Chem. Res.* **2015**, *48* (7), 2017–2026.

[376] Artero, V.; Saveant, J. M. Toward the Rational Benchmarking of Homogeneous H_2-Evolving Catalysts. *Energy Environ. Sci.* **2014**, *7* (11), 3808–3814.

[377] Straistari, T.; Fize, J.; Shova, S.; Réglier, M.; Artero, V.; Orio, M. A Thiosemicarbazone–Nickel(II) Complex as Efficient Electrocatalyst for Hydrogen Evolution. *ChemCatChem* **2017**, *9* (12), 2262–2268.

[378] Ren, S.; Joulié, D.; Salvatore, D.; Torbensen, K.; Wang, M.; Robert, M.; Berlinguette, C. P. Molecular Electrocatalysts Can Mediate Fast, Selective CO_2 Reduction in a Flow Cell. *Science (80-.).* **2019**, *365* (6451), 367–369.

[379] Izsák, R.; Neese, F. An Overlap Fitted Chain of Spheres Exchange Method. *J. Chem. Phys.* **2011**, *135* (14).

[380] Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.

[381] Helm, M. L.; Stewart, M. P.; Bullock, R. M.; DuBois, M. R.; DuBois, D. L. A Synthetic Nickel Electrocatalyst with a Turnover Frequency Above 100,000 s^{-1} for H_2 Production. *Science (80-.).* **2011**, *333* (August), 863–866.

[382] Graham, D. J.; Nocera, D. G. Electrocatalytic H_2 Evolution by Proton-Gated Hangman Iron Porphyrins. *Organometallics* **2014**, *33* (18), 4994–5001.

[383] Bhugun, I.; Lexa, D.; Savéant, J. M. Homogeneous Catalysis of Electrochemical Hydrogen Evolution by Iron(0) Porphyrins. *J. Am. Chem. Soc.* **1996**, *118* (16), 3982–3983.

[384] Li, C. B.; Gong, P.; Yang, Y.; Wang, H. Y. Cobalt(II)–Salen Complexes for Photocatalytic Hydrogen Production in Noble Metal-Free Molecular Systems. *Catal. Letters* **2018**, *148* (10), 3158–3164.

[385] McNamara, W. R.; Han, Z.; Yin, C. J.; Brennessel, W. W.; Holland, P. L.; Eisenberg, R. Cobalt-Dithiolene Complexes for the Photocatalytic and Electrocatalytic Reduction of Protons in Aqueous Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (39), 15594–15599.

[386] Andreiadis, E. S.; Jacques, P. A.; Tran, P. D.; Leyris, A.; Chavarot-Kerlidou, M.; Jousselme, B.; Matheron, M.; Pécaut, J.; Palacin, S.; Fontecave, M.; Artero, V. Molecular Engineering of a Cobalt-Based Electrocatalytic Nanomaterial for H_2 Evolution under Fully Aqueous Conditions. *Nat. Chem.* **2013**, *5* (1), 48–53.

[387] Rodríguez-López, N.; Wu, Y.; Ge, Y.; Villagrán, D. Hydrogen Evolution Catalyzed by a Metal-Free Corrole: An Experimental and Theoretical Mechanistic Study. *J. Phys. Chem. C* **2020**, *124* (19), 10265–10271.

[388] Bates, R. G.; Macaskill, J. B. Standard Potential of the Silver-Silver Chloride Electrode. *Pure Appl. Chem.* **1978**, *50* (11–12), 1701–1706.

[389] Xiao, W. C.; Tao, Y. W.; Luo, G. G. Hydrogen Formation Using a Synthetic Heavier Main-Group Bismuth-Based Electrocatalyst. *Int. J. Hydrogen Energy* **2020**, *45* (15), 8177–8185.

[390] Maher, A. G.; Passard, G.; Dogutan, D. K.; Halbach, R. L.; Anderson, B. L.; Gagliardi, C. J.; Taniguchi, M.; Lindsey, J. S.; Nocera, D. G. Hydrogen Evolution Catalysis by a Sparsely Substituted Cobalt Chlorin. *ACS Catal.* **2017**, *7* (5), 3597–3606.

[391] Wu, Y.; Rodríguez-López, N.; Villagrán, D. Hydrogen Gas Generation Using a Metal-Free Fluorinated Porphyrin. *Chem. Sci.* **2018**, *9* (20), 4689–4695.

[392] Patra, B. C.; Khilari, S.; Manna, R. N.; Mondal, S.; Pradhan, D.; Pradhan, A.; Bhaumik, A. A Metal-Free Covalent Organic Polymer for Electrocatalytic Hydrogen Evolution. *ACS Catal.* **2017**, *7* (9), 6120–6127.

[393] Thompson, E. J.; Berben, L. A. Electrocatalytic Hydrogen Production by an Aluminum(III) Complex: Ligand-Based Proton and Electron Transfer. *Angew. Chemie - Int. Ed.* **2015**, *54* (40), 11642–11646.

[394] Barrozo, A.; Orio, M. Molecular Electrocatalysts for the Hydrogen Evolution Reaction: Input from Quantum Chemistry. *ChemSusChem* **2019**, *12* (22), 4905–4915.

[395] Luo, G. G.; Zhang, H. L.; Tao, Y. W.; Wu, Q. Y.; Tian, D.; Zhang, Q. Recent Progress in Ligand-Centered Homogeneous Electrocatalysts for Hydrogen Evolution Reaction. *Inorg. Chem. Front.* **2019**, *6* (2), 343–354.

[396] Haddad, A. Z.; Cronin, S. P.; Mashuta, M. S.; Buchanan, R. M.; Grapperhaus, C. A. Metal-Assisted Ligand-Centered Electrocatalytic Hydrogen Evolution upon Reduction of a Bis(Thiosemicarbazonato)Cu(II) Complex. *Inorg. Chem.* **2017**, *56* (18), 11254–11265.

[397] Jiang, J.; Materna, K. L.; Hedström, S.; Yang, K. R.; Crabtree, R. H.; Batista, V. S.; Brudvig, G. W. Antimony Complexes for Electrocatalysis: Activity of a Main-Group Element in Proton Reduction. *Angew. Chemie - Int. Ed.* **2017**, *56* (31), 9111–9115.

[398] Wang, N.; Lei, H.; Zhang, Z.; Li, J.; Zhang, W.; Cao, R. Electrocatalytic Hydrogen Evolution with Gallium Hydride and Ligand-Centered Reduction. *Chem. Sci.* **2019**, *10* (8), 2308–2314.

[399] Aroua, S.; Todorova, T. K.; Mougel, V.; Hommes, P.; Reissig, H. U.; Fontecave, M. New Cobalt-Bisterpyridyl Catalysts for Hydrogen Evolution Reaction. *ChemCatChem* **2017**, *9* (12), 2099–2105.

[400] Metavarayuth, K.; Ejegbavwo, O.; McCarver, G.; Myrick, M. L.; Makris, T. M.;

Vogiatzis, K. D.; Senanayake, S. D.; Manley, O. M.; Ebrahim, A. M.; Frenkel, A. I.; Hwang, S.; Rajeshkumar, T.; Jimenez, J. D.; Chen, K.; Shustova, N. B.; Chen, D. A. Direct Identification of Mixed-Metal Centers in Metal-Organic Frameworks: $Cu_3(BTC)_2$Transmetalated with Rh^{2+} Ions. *J. Phys. Chem. Lett.* **2020**, *11* (19), 8138–8144.

[401] Batten, S. R.; Champness, N. R.; Chen, X. M.; Garcia-Martinez, J.; Kitagawa, S.; Öhrström, L.; O'Keeffe, M.; Suh, M. P.; Reedijk, J. Coordination Polymers, Metal-Organic Frameworks and the Need for Terminology Guidelines. *CrystEngComm* **2012**, *14* (9), 3001–3004.

[402] Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q. Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal-Organic Framework Database: CoRE MOF 2019. *J. Chem. Eng. Data* **2019**, *64* (12), 5985–5998.

[403] Li, J.-R.; Sculley, J.; Zhou, H.-C. Metal-Organic Frameworks for Separations. *Chem. Rev.* **2012**, *112* (2), 869–932.

[404] Kreno, L. E.; Leong, K.; Farha, O. K.; Allendorf, M.; Van Duyne, R. P.; Hupp, J. T. Metal-Organic Framework Materials as Chemical Sensors. *Chem. Rev.* **2012**, *112* (2), 1105–1125.

[405] Wang, L.; Han, Y.; Feng, X.; Zhou, J.; Qi, P.; Wang, B. Metal-Organic Frameworks for Energy Storage: Batteries and Supercapacitors. *Coord. Chem. Rev.* **2016**, *307*, 361–381.

[406] Wang, L.; Zheng, M.; Xie, Z. Nanoscale Metal-Organic Frameworks for Drug Delivery: A Conventional Platform with New Promise. *J. Mater. Chem. B* **2018**, *6* (5), 707–717.

[407] Wang, Q.; Astruc, D. State of the Art and Prospects in Metal-Organic Framework (MOF)-Based and MOF-Derived Nanocatalysis. *Chem. Rev.* **2020**, *120* (2), 1438–1511.

[408] Islamoglu, T.; Chen, Z.; Wasson, M. C.; Buru, C. T.; Kirlikovali, K. O.; Afrin, U.; Mian, M. R.; Farha, O. K. Metal-Organic Frameworks against Toxic Chemicals. *Chem. Rev.* **2020**, *120* (16), 8130–8160.

[409] Bavykina, A.; Kolobov, N.; Khan, I. S.; Bau, J. A.; Ramirez, A.; Gascon, J. Metal-Organic Frameworks in Heterogeneous Catalysis: Recent Progress, New Trends, and Future Perspectives. *Chem. Rev.* **2020**, *120* (16), 8468–8535.

[410] Jiang, J.; Yaghi, O. M. Brønsted Acidity in Metal-Organic Frameworks. *Chem. Rev.* **2015**, *115* (14), 6966–6997.

[411] Zhu, L.; Liu, X. Q.; Jiang, H. L.; Sun, L. B. Metal-Organic Frameworks for Heterogeneous Basic Catalysis. *Chem. Rev.* **2017**, *117* (12), 8129–8176.

[412] Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions. *J. Chem. Phys.* **2005**, *123* (2).

[413] Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124* (17).

[414] Sagara, T.; Klassen, J.; Ortony, J.; Ganz, E. Binding Energies of Hydrogen Molecules to Isoreticular Metal-Organic Framework Materials. *J. Chem. Phys.* **2005**, *123* (1), 1–5.

[415] Sagara, T.; Klassen, J.; Ganz, E. Computational Study of Hydrogen Binding by Metal-Organic Framework-5. *J. Chem. Phys.* **2004**, *121* (24), 12543–12547.

[416] Mavrandonakis, A.; Klopper, W. Comment on "Kinetics and Mechanistic Model for Hydrogen Spillover on Bridged Metal-Organic Frameworks." *J. Phys. Chem. C* **2008**, *112* (8), 3152–3154.

[417] Hübner, O.; Glöss, A.; Fichtner, M.; Klopper, W. On the Interaction of Dihydrogen with Aromatic Systems. *J. Phys. Chem. A* **2004**, *108* (15), 3019–3023.

[418] Klontzas, E.; Mavrandonakis, A.; Froudakis, G. E.; Carissan, Y.; Klopper, W. Molecular Hydrogen Interaction with IRMOF-1: A Multiscale Theoretical Study. *J. Phys. Chem. C* **2007**, *111* (36), 13635–13640.

[419] Vogiatzis, K. D.; Klopper, W.; Friedrich, J. Non-Covalent Interactions of CO2 with Functional Groups of Metal-Organic Frameworks from a CCSD(T) Scheme Applicable to Large Systems. *J. Chem. Theory Comput.* **2015**, *11* (4), 1574–1584.

[420] Cotton, S. A. Rhodium and Iridium. In *Chemistry of Precious Metals*; 1997; pp 78–172.

[421] Hendon, C. H.; Walsh, A. Chemical Principles Underpinning the Performance of the Metal-Organic Framework HKUST-1. *Chem. Sci.* **2015**, *6* (7), 3674–3683.

[422] Todaro, M.; Alessi, A.; Sciortino, L.; Agnello, S.; Cannas, M.; Gelardi, F. M.; Buscarino, G. Investigation by Raman Spectroscopy of the Decomposition Process of HKUST-1 upon Exposure to Air. *J. Spectrosc.* **2016**, *2016*.

[423] Heinz, W. R.; Kratky, T.; Drees, M.; Wimmer, A.; Tomanec, O.; Günther, S.; Schuster, M.; Fischer, R. A. Mixed Precious-Group Metal-Organic Frameworks: A Case Study of the HKUST-1 Analogue $[Ru_xRh_{3-x}(BTC)_2]$. *Dalt. Trans.* **2019**, *48* (32), 12031–12039.

[424] Goto, M.; Miyagi, Y.; Minami, M.; Sanda, F. Synthesis and Crosslinking Reaction of Polyacetylenes Substituted with Benzoxazine Rings: Thermally Highly Stable Benzoxazine Resins. *J. Polym. Sci. Part A Polym. Chem.* **2018**, *56* (16), 1884–1893.

[425] Onfroy, T.; Marie, O. IR Operando Study of the Acetylene Effect on the NO Reduction Mechanism on Pd–Rh/CeO_2–ZrO_2 Three-Way Catalyst. *Top. Catal.* **2019**, *62* (1–4), 336–344.

[426] Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125* (19), 194101.

[427] Pirillo, J.; Hijikata, Y. Trans Influence across a Metal-Metal Bond of a Paddle-Wheel Unit on Interaction with Gases in a Metal-Organic Framework. *Inorg. Chem.* **2020**, *59* (2), 1193–1203.

Vita

Gavin Alexander McCarver grew up in West Valley City, UT but spent several years in Wake Forest, NC and Saratoga Springs, UT where he graduated high school alongside his twin brother, Garrett. He attended Utah State University where he studied chemistry and gained an interest in both physical and inorganic chemistry. He performed research under the guidance of Dr. Yujie Sun and Dr. David Farrelly where he gained a new love for computational and theoretical chemistry. After graduating with a BS in Chemistry with a Professional emphasis, he moved across the country to Knoxville, TN to pursue a Ph.D. in chemistry. He joined the groups of Dr. RJ Hinde and Dr. Konstantinos Vogiatzis where he explored numerous areas of chemistry but found a passion in spectroscopic studies of small molecules and catalysis. During his time at UT, Gavin received several awards from the University and from the chemistry department including the Tennessee Fellowship for Graduate Excellence, the Extraordinary Professional Promise Award, the Eugene John Barber Fellowship in Physical Chemistry, the Outstanding Teaching Award, and the Second Year Candidacy Award. He was awarded a two-year fellowship at the National Institute of Standards and Technology (NIST) where he plans to study under Dr. Wei Zhou and Dr. Taner Yildirim.