January 2018

Computational Studies Of Oxides Relevant To Clean Energy, Catalytic Processing Of Renewables, And Biological Systems

Jason M. Hicks

COMPUTATIONAL STUDIES OF OXIDES RELEVANT TO CLEAN ENERGY, CATALYTIC PROCESSING OF RENEWABLES, AND BIOLOGICAL SYSTEMS

by

Jason Michael Hicks

Master of Science, Chemistry, University of North Dakota, 2013
Bachelor of Science, Chemistry, University of North Dakota, 2010
Associate of Arts, Northland Community and Technical College, 2008

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota
August
2018

PERMISSION

Title	Computational Studies of Oxides Relevant to Clean Energy, Catalytic Processing of Renewables, and Biological Systems
Department	Chemistry
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Jason M. Hicks
7/19/2018

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES viii

ACKNOWLEDGEMENTS x

ABSTRACT xiii

1 INTRODUCTION 1
- 1.1 Computational Chemistry . . . 1
- 1.2 Theoretical Investigations of Metalloid Oxides . . . 4
- 1.3 Computational Studies of the Catalytic Degradation of Lignin . . . 5
- 1.4 Theoretical Studies of Isomerization of the $Cu_2O_2^{2+}$ Core . . . 7
- 1.5 Structure of Dissertation . . . 9

2 METHODS OF ELECTRONIC STRUCTURE 10
- 2.1 Theoretical Foundation . . . 10
- 2.2 Hartree–Fock Approximation . . . 15
- 2.3 Second Order Møller–Plesset Perturbation Theory . . . 19
- 2.4 Multiconfigurational Self-Consistent Field Theory, Active Spaces, and Macroconfigurations . . . 24
- 2.5 Second-Order Generalized van Vleck Perturbation Theory . . . 28
- 2.6 Multireference Configuration Interaction with Singles and Doubles . . . 36
- 2.7 Choice of Active Space . . . 38

3 THEORETICAL INVESTIGATIONS OF ARSENIC OXIDES RELEVANT TO COAL COMBUSTION: A GVVPT2 MULTIREFERENCE PERTURBATION THEORY STUDY 44
- 3.1 Introduction . . . 44
- 3.2 Theoretical Methodology . . . 47
  - 3.2.1 Computational Details . . . 48
  - 3.2.2 Active Spaces . . . 49
- 3.3 Results and Discussion . . . 51
  - 3.3.1 Arsenic monoxide (AsO) . . . 52
  - 3.3.2 Arsenic dioxide (AsO_2) . . . 54
  - 3.3.3 Arsenic trioxide (AsO_3) . . . 54
  - 3.3.4 Diarsenic trioxide (As_2O_3) . . . 57
  - 3.3.5 Diarsenic pentoxide (As_2O_5) . . . 63
- 3.4 Summary . . . 72

4 GVVPT2 MULTIREFERENCE PERTURBATION THEORY DESCRIPTION OF ANTIMONY OXIDES RELEVANT TO COAL COMBUSTION 74
4.1 Introduction . . . 74
4.2 Theoretical Methodology . . . 76
4.2.1 Computational Details . . . 77
4.2.2 Active Spaces . . . 78
4.3 Results and Discussion . . . 80
4.3.1 Antimony Monoxide (SbO) . . . 81
4.3.2 Antimony Dioxide (SbO_2) . . . 83
4.3.3 Antimony Trioxide (SbO_3) . . . 83
4.3.4 Diantimony Trioxide (Sb_2O_3) . . . 85
4.3.5 Diantimony Pentoxide (Sb_2O_5) . . . 88
4.4 Summary . . . 93

5 ELECTRONIC STRUCTURE CALCULATIONS OF CATALYZED LIGNIN DECOMPOSITION 94
5.1 Introduction . . . 94
5.2 Methods . . . 97
5.2.1 Model Construction . . . 97
5.2.2 Computational Details . . . 98
5.3 Results and Discussion . . . 100
5.4 Summary . . . 107

6 THEORETICAL INVESTIGATIONS OF THE ISOMERIZATION OF $Cu_2O_2^{2+}$ 108
6.1 Introduction . . . 108
6.2 Computational Details . . . 111
6.3 Active Space Development . . . 112
6.3.1 Generalized Valence Bond-Derived Active Space . . . 112
6.3.2 Numerically Based Active Space Development . . . 115
6.4 Relative Energies of the $Cu_2O_2^{2+}$ Isomerization . . . 124
6.4.1 Active Space Effect . . . 125
6.4.2 Effect of Including Quadruple Excitations . . . 128
6.4.3 Comparison of GVVPT2 Results to Previous Studies . . . 131
6.5 Future Work . . . 135

7 OVERALL CONCLUDING REMARKS 136

REFERENCES 137

LIST OF FIGURES

2.1 Descriptions of various interaction variables in a system consisting of nuclei A and B along with electrons i and j (Image taken from Szabo and Ostlund).[1] 12
2.2 The Hamiltonian matrix represented as a box. (a) The $\mathbf{H}_{MM}$ matrix has a small number of electron configurations and neglects dynamic correlation while the full H matrix has a large number of configurations and includes dynamic correlation. (b) By a final diagonalization, dynamic electron correlation is included into $\mathbf{H}_{MM}^{eff}$ by the contributions from $\mathbf{H}_{QM}$ and $\mathbf{H}_{MQ}$ while only the diagonal elements of $\mathbf{H}_{QQ}$ are kept. 32
2.3 Schematic representation of the FCI matrix. Here it is explicitly showing up through triple excitations, but in general, it continues to the right and down till all possible excitations are accounted for. The location that the FCI matrix is truncated to get MRCISD is shown by the red box. Drawing is not to scale. 37

3.1 Structure of isomer 1 of diarsenic trioxide (As_2O_3). 57
3.2 Relative GVVPT2/cc-pVTZ energies of diarsenic trioxide (As_2O_3) isomers and TSs (in kcal/mol). 61
3.3 Structure of isomer 1 of diarsenic pentoxide (As_2O_5). 63
3.4 Relative GVVPT2/cc-pVTZ energies of diarsenic pentoxide (As_2O_5) isomers and TSs (in kcal/mol). 69

5.1 (a) An example of the structure of a small piece of lignin. (b) The three common monomer units, from left to right: p-hydroxyphenyl, guaiacyl, and syringyl. 95
5.2 SEM of fractured surface of an amorphous silica–alumina sample.[2] 96
5.3 (a) Optimized structure of p-hydroxyphenyl alcohol, (b) Optimized structure of a β–O–4 linked model compound (i.e., a p-hydroxyphenyl linked to a guaiacyl). . . . 101
5.4 (a) The full B3LYP/6-31G* optimized structure of p-hydroxyphenyl adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of p-hydroxyphenyl with the surface showing the hydrogen bond distances. 102
5.5 (a) The full B3LYP/6-31G* optimized structure of the β–O–4 linked p-hydroxyphenyl and guaiacyl (i.e., the dimer) adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of the dimer with the surface showing the hydrogen bond distances. 103
5.6 (a) The full B3LYP/6-31G* optimized structure of the dimer, with the β–O–4 ether bond cleaved and capped with hydrogens, adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of the dimer fragments with the surface showing the hydrogen bond and aluminum to oxygen interaction distances. . . 104

5.7 The principle of how a volcano plot works. The optimal catalyst is when the heat of adsorption (or binding energy) is "just right" so that the rate of reaction can be maximized. 106

6.1 The most common bonding motifs of Cu_2O_2.[3] 109
6.2 B3LYP orbitals for bis(μ-oxo) (left) and side-on μ-η^2:η^2-peroxo (right). This illustrates the nominal $2b_{1g}$ HOMO and $3b_{3g}$ LUMO orbitals at the B3LYP level.[3] . . . 110
6.3 Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing how the energy changed when AS-4 (solid lines) was modified to get AS-5 (dashed lines). Both active spaces included SDT excitations. 125
6.4 Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing how the energy changed when AS-5 (solid lines) was modified to get AS-6 (dashed lines). Both active spaces included SDTQ excitations. 127
6.5 Bar chart showing how the difference in the energies (kcal/mol) change at different points along the isomerization of $Cu_2O_2^{2+}$ for both MCSCF and GVVPT2. The MCSCF and GVVPT2 levels of theory showing the effects of adding quadruple excitations (i.e., SDTQ [dashed lines] vs. SDT [solid lines]) when using AS-5. . . . 130
6.6 Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing the effects of adding quadruple excitations (i.e., SDTQ [dashed lines] vs. SDT [solid lines]) when using AS-5. 131
6.7 Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory compared to previous studies (CR-CCSD(TQ)$_L$ values obtained from Ref. 3 and RASPT2(24e,28o)//4 values obtained from Ref. 4). 134

LIST OF TABLES

2.1 Comparison of the orbital classifications for CASSCF and RASSCF. 27

3.1 Comparison of thermodynamic constants for the ground state ($X^2\Pi$) of AsO. 52
3.2 Leading configurations and weights of MCSCF reference functions of arsenic oxides given as CSFs in the GVVPT2 output. 53
3.3 Comparison of structural parameters (Å and degrees) of AsO and AsO_2, optimized at the B3LYP/6-311G* and GVVPT2/cc-pVTZ levels with experimental data. . . . 54
3.4 The geometrical parameters (Å and degrees) of the AsO_3 doublet and quartet, optimized at the B3LYP/6-311G* level. 55
3.5 Reaction enthalpies (kcal/mol) for the oxidation of monomeric arsenic oxides. . . . 56
3.6 Optimized B3LYP/6-311G* geometrical parameters (Å and degrees) of the diarsenic trioxide (As_2O_3) isomers. 58
3.7 Frequency of diarsenic trioxide (As_2O_3) isomers and transition states in cm^{-1} calculated at the level of B3LYP/6-311G*. 59
3.8 Relative energies (kcal/mol) at the B3LYP, CR-CC(2,3) and GVVPT2 levels of theory of the diarsenic trioxide (As_2O_3) isomers and transition states. 60
3.9 Total energies (a.u.) of AsO, AsO_2 and As_2O_3 (isomer 1), and predicted reaction energies (kcal/mol). 62
3.10 Optimized B3LYP/6-311G* geometrical parameters (Å and degrees) of the diarsenic pentoxide (As_2O_5) isomers and transition states. 64
3.11 Frequencies of diarsenic pentoxide (As_2O_5) isomers obtained at the B3LYP/6-311G* level of theory. 66
3.12 Frequencies of diarsenic pentoxide (As_2O_5) transistion states obtained at the B3LYP/6-311G* level of theory. 67
3.13 Relative energies (kcal/mol) at the B3LYP, CR-CC(2,3) and GVVPT2 levels of theory of the diarsenic pentoxide (As_2O_5) isomers and transition states. 68
3.14 Total energies (a.u.) of O_2, AsO_3 and As_2O_5 (isomer 1), and predicted reaction energies (kcal/mol). 71

4.1 Comparison of thermodynamic constants from geometries optimized at various levels of theory for the ground state of SbO. 81
4.2 Leading configurations and weights of MCSCF reference functions of antimony oxides given as CSFs in the GVVPT2 output. 82
4.3 Comparison of structural parameters (Å and degrees) of SbO and SbO_2, optimized at the B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ levels. 83
4.4 The geometrical parameters (Å and degrees) of the SbO_3 doublet, optimized at the B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ levels. 84

4.5 Reaction enthalpies (kcal/mol) for the oxidation of monomeric antimony oxides using B3LYP and GVVPT2//B3LYP using the SDB-aug-cc-pVTZ basis set. 84
4.6 Optimized B3LYP geometrical parameters (Å and degrees) of the diantimony trioxide (Sb_2O_3) isomers. 85
4.7 Frequency of diantimony trioxide (Sb_2O_3) isomers in cm^{-1} calculated at the B3LYP level of theory. 86
4.8 Relative energies (kcal/mol) at the B3LYP and GVVPT2 levels of theory of the diantimony trioxide (Sb_2O_3) isomers. 87
4.9 Total energies (a.u.) of SbO, SbO_2 and Sb_2O_3 (isomer 1), and predicted reaction energies (kcal/mol). 87
4.10 Optimized B3LYP/SDB-aug-cc-pVTZ geometrical parameters (Å and degrees) of the diantimony pentoxide (Sb_2O_5) isomers. 89
4.11 Frequencies of diantimony pentoxide (Sb_2O_5) isomers obtained at the B3LYP/SDB-aug-cc-pVTZ level of theory. 90
4.12 Relative energies (kcal/mol) at the B3LYP and GVVPT2 levels of theory of the diantimony pentoxide (Sb_2O_5) isomers. 91
4.13 Total energies (a.u.) of O_2, SbO_3 and Sb_2O_5 (isomer 1), and predicted reaction energies (kcal/mol). 92

6.1 MCSCF and GVVTP2 total energies (a.u.) for Cu_2 + 2O along with the GVB-derived active space (22e,14o). 114
6.2 GVB-derived active space applied to the isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo, (20e,14o). 114
6.3 Orbital partitioning in AS-Big, (30e,30o). SD excitations were considered with this active space. 118
6.4 Orbital partitioning in AS-RedBig, (22e,22o). SDTQ excitations were considered with this active space. 119
6.5 Orbital partitioning in AS-2, (20e,13o). SD excitations were considered with this active space. 120
6.6 Orbital partitioning in AS-3, (20e,14o). SDT excitations were considered with this active space. 121
6.7 Orbital partitioning in AS-4, (18e, 16o). SDT excitations were considered with this active space. 122
6.8 Orbital partitioning in AS-5, (16e, 17o). Both SDT and SDTQ excitations were considered with this active space. 123
6.9 Orbital partitioning in AS-6, (12e, 13o). Only SDTQ excitations were considered with this active space. 124
6.10 Absolute energies (Hartree) and the difference at each given point (kcal/mol) along the isomerization of $Cu_2O_2^{2+}$ showing the how the energies changed when quadruple excitations were add to AS-5. Relative energies were calculated at each point (e.g., at F= 100%) as $E_{rel} = E_{SDTQ} - E_{SDT}$. 129
6.11 Comparisons of relative energies (kcal/mol) of various levels of theory along the isomerization path of $Cu_2O_2^{2+}$ compared to best estimates from CAS, RAS and CR-CC results. Energies were related by equating all the side-on μ-η^2:η^2-peroxo energies. 133

Acknowledgements

Firstly, I would like to thank my advisor, Prof. Mark Hoffmann for his continual guidance in my education and research; I would not be where I am now if not for him. His masterful understanding of theoretical and computational chemistry, along with his willingness to help has enabled me to grasp computational chemistry and my research far more than I would have thought I could have. He has also always been so willing to meet with me, to answer my barrage of questions, many times even without an appointment, and he has been very patient with me throughout my stay at UND. He has been a true and great mentor over the years and I consider myself very fortunate to have had the privilege of being advised by him.

I would also like to thank Prof. Evguenii Kozliak, Prof. Harmon Abrahamson, Prof. Jerome Delhommelle, and Prof. Kanishka Marasinghe for all of their help and for agreeing to serve on my committee. I have had the privilege of taking many courses from both Dr. Kozliak and Dr. Abrahamson over the years, and they have proven to be mentors to me in more ways than just academics. They have given me so much of their time and guidance over the years, my time here would have been nowhere near as fruitful without their counsel. My first semester at UND I had Dr. Delhommelle for physical chemistry and I have loved it ever since, he got me hooked. Not

only that, but he was the one who introduced me to research and when he was my co-advisor for my masters, he was always so helpful in getting me deeper into the literature and helping me hone my presentation skills. Dr. Marasinghe has always been so kind, helpful and personable. Though I have never had the pleasure of taking a class from him, he has always been very helpful in my educational experience, even letting me tag along with him to a PICUP meeting.

I would also like to thank all my fellow research group members, both current and alumni for all of their help and support during my time here at UND, you have all proven to valuable colleagues as well as great friends. I would also like to thank all the chemistry graduate students for all the great times. Graduate school would not have been anywhere near as good without all of your friendships.

I would like to thank all of my other friends and family who have so instrumental to my life. My mom and dad deserve an indescribable amount of thanks for all of their continual support and prayers.

I would especially like to thank my wonderful wife Erica Hicks for all of her love and encouragement throughout my stay here. She has so lovingly put up with all the late nights I have put in finishing this dissertation. There have been so many times that her loving encouragement has given me energy and drive to push beyond anything I thought I could do. My world would be so much darker without her lovely smile and selfless support to help me through. I am so incredibly fortunate to be married to my best friend. I will even say that we have great chemistry together! I owe her a lot for everything and now even more so since she delivered our first child, Jase Thomas Hicks, on December 18th, 2017!

Last, but certainly not least, I would like to thank God for all of his guidance, love, presence and wisdom. He has been there from the beginning pushing me to step outside my comfort zone

and giving me the wisdom, strength and motivation I need to make it through. I am deeply grateful for all of He has done and is still doing for me. There were so many times throughout my graduate work, when things were not going well, that He filled me with and indescribable peace through His Holy Spirit or miraculously came through and helped me in ways I did not even think possible. He is always drawing me closer to Him and I really have no way to properly describe how incredibly awesome that is.

Abstract

Computational chemistry has grown into a large field and is continuing to grow every year in both number and variety of applications. This dissertation will give a few such applications relevant to cleaner energy production from coal, catalytic degradation of renewable agricultural and forest waste into valuable chemicals, and extending the reach of electronic structure methods to systems of biological and macromolecular interest. The first two studies presented in this dissertation are concerned with the remediation of trace elements released into the environment through the combustion of coal for power production. In flue gases, arsenic and antimony exists most often as oxides. Despite the prevalence and importance of remediating these oxides, critical information on the thermodynamics of plausible intermediates and transition states in reaction pathways have been missing prior to these studies. Several of the intermediates, and essentially all transition states, were found to be electronically multiconfigurational for the arsenic oxides. In this work, the electronic structures of several oxides of arsenic, As_xO_y, where x = 1, 2 and y = 1–5, were investigated using the second-order generalized van Vleck variant of multireference perturbation theory (GVVPT2), using the cc-pVTZ basis set, with comparison to multi-reference configuration interaction (MRCISD) and the linked completely renormalized coupled cluster through perturbative triple excitations (CR-CCSD(T)$_L$ or CR-CC(2,3)) when relevant. Calculated oxidation reaction energies for the formation of AsO_2 and AsO_3 from AsO were predicted to be energetically favorable and formation energies of the lowest energy compounds containing two metalloid atoms,

called dimers for brevity, from the monomers were also predicted to be energetically favorable. The energetics of the monomers, five isomers of As_2O_3 and eleven isomers of As_2O_5 were characterized using a composite methodology along with the key transition states between the isomers. Geometry optimizations as well as harmonic vibrational frequencies of As_xO_y were obtained at the B3LYP/6-311G* level of theory and gave satisfactory agreement with experimental data when available. It was discovered that several isomers of As_2O_3 and As_2O_5 have comparable energies and relatively low barrier heights. Therefore, we expect these isomers to be chemically relevant.

The antimony oxides were also found to be electronically multiconfigurational. The electronic structures of several antimony oxides, Sb_xO_y, where x = 1, 2 and y = 1–5, were investigated using GVVPT2 and the SBD-aug-cc-pVTZ basis set. The oxidation reaction energies of elemental antimony toward the formation of SbO and SbO_2 was found to be energetically favorable, while the further oxidation of those species to SbO_3 was found to be unfavorable. It was found that the accretion of the monomers into Sb_2O_3 was highly energetically favorable at both the B3LYP/SBD-aug-cc-pVTZ and GVVPT2/SBD-aug-cc-pVTZ//B3LYP/SBD-aug-cc-pVTZ levels of theory. However, while the reaction of SbO_2 and SbO_3 toward Sb_2O_5 was found to be favorable, it was found to be unfavorable for Sb_2O_5 to form from the oxidation of Sb_2O_3.The energetics of the monomers, three isomers of Sb_2O_3 and four isomers of Sb_2O_5 were characterized using the same composite methodology as the arsenic oxides. Geometry optimizations and harmonic vibrational frequencies of all antimony oxides were obtained at the B3LYP/SBD-aug-cc-pVTZ level of theory. Several of the dimeric antimony oxide isomer structures were found to be quite similar to the arsenic dimers and are also expected to be chemically relevant.

The third study is pertinent to the catalytic degradation of lignin, one of the most renewable carbon sources on Earth. Unfortunately, it has only seen limited industrial use due to its chemical

stability and complex structure; thus lignin is typically disposed of as non-commercialized waste product. However, if a viable path could be found for the decomposition of lignin, the by-products could be used to replace high-value petrochemicals. Catalytic decomposition by amorphous silica-alumina (ASA) based catalysts may be a viable path, but the mechanisms and the effects of metal doping are not strictly known. In this work, DFT calculations with the B3LYP hybrid functional, as implemented in the NWChem software package, are used to elucidate this information. A cluster model of an amorphous silica-alumina catalyst has been studied and a monomer and β–O–4 linked dimer have been adsorbed to the surface. They were both found to have a favorable interaction with the surface and, for the dimer, a favorable cleavage of the β–O–4 bond. The desorption energy of the cleaved monomers was also shown to be favorable when compared to the free dimer, suggesting ASA is a viable catalyst.

The fourth study, the $Cu_2O_2^{2+}$ core, has both biological and computational relevance and has received much attention over the years. This is due to its importance in biological systems and to the computational difficulties associated with modeling its relevant isomers. The complexity for computation arises, inter alia, to both a varying degree of biradical character as it isomerizes, as well as a rapidly changing degree of dynamic and static electron correlation effects along the isomerization coordinate. In this work, the two dominant isomers, bis(μ-oxo) and μ-η^2:η^2 peroxo, along with four points along the reaction coordinate, were considered. MCSCF and GVVPT2 were used with a variety of active spaces. The starting active spaces were developed using a recently established approach, in which a valence picture is used rather than the typical Hartree–Fock description. The active spaces were then modified based on numerical considerations in an iterative fashion. Finally, a stable active space of 13 electrons in 12 orbitals (13e,12o) was found and used to obtain results which were then compared to the more expensive methods of

complete active space second-order perturbation theory (CASPT2) using an active space of (16e, 14o), restricted active space second-order perturbation theory (RASPT2) using an active space of (24e,28o) which considered up to quadruple excitations, and linked completely renormalized coupled cluster through perturbative triple and quadruple excitations (CR-CCSD(TQ)$_L$ or CR-CC(2,4)). GVVPT2 was found to agree well with the CASPT2 results. It was also determined that a larger active space than (13e,12o) will likely be required for it to approach the relative energies of RASPT2(24e,28o)//4 or CR-CCSD(TQ)$_L$.

1 INTRODUCTION

1.1 Computational Chemistry

Since the first computational calculations of semi-empirical atomic orbital calculations in the early 1950s,[5] computational chemistry has grown into an increasingly useful tool that has been brought it into the limelight, having been the topic of several Nobel Prizes in Chemistry and Physics. The continued advancement of computational chemistry in the broadness and depth of application, along with its excellent qualitative and quantitative capabilities, shows no sign of slowing down. One of the pillars of modern computational chemistry, Henry F. Schaefer III, has accurately predicted this growth by stating that every year there would be another 1% of peer reviewed chemical publications that have some sort of computational component.

What is the driving force behind this profound proliferation that this field of chemistry has been experiencing? Computational chemistry has moved from merely a qualitative tool, which provided validation, into a highly respectable quantitative tool that has been used to predict reactions, provide invaluable insight in ways experimentation is unable to and has even been shown to overturn established experimental results, as later resolved by further experiment.[6] The driv-

ing force for this shift and maturing of an entire field of chemistry is fourfold. Firstly, the theory driving the computations has been advancing, allowing for greater accuracy and decreased computational cost. Secondly, the computational implementation of the theory has vastly improved over the years with more efficient algorithms and code optimization, allowing for faster calculations, which make better use of the available computational architecture. Thirdly, computational power available has grown exponentially, following Moore's Law.[7] It shows no signs of plateauing anytime soon, with new innovations such as allowing for calculations on GPUs, increasing powerful massively parallel supercomputers, and even the incredible potential of quantum computers for certain applications. Fourthly, computational chemists have pushed the field forward by applying these computational implemented methods to solve chemical problems through the development of novel computational protocols. This fourth vehicle of growth can provide the scientific community with both valuable insight, many times previously unobtainable, into chemical problems as well as with new computational procedures to obtain insights into other similar systems. This is the avenue which the body of work presented in this dissertation has striven to follow.

The field of computational chemistry has many methods at its disposal. These methods can be loosely divided into 3 categories: (i) classical methods such as Monte Carlo and molecular dynamics, (ii) semi-empirical methods, and (iii) *ab initio* quantum mechanical methods. Though there have been many great advances in both development and application in all of these areas, the methods used in this dissertation all fall under the category of *ab initio* methods and will therefore be the focus of further discussion.

Ab initio methods, also known as first-principle methods, are named as such because they do not need any experimental parameters in their pure form. They operate fully from the theoretical methods that drive them, the mathematical description of the atomic orbitals, and the geometrical

arrangement of the atoms. *Ab initio* methods are powerful in that they can provide the electronic structure of the system, from which nearly all relevant chemical information and useful properties can be derived. Just a few examples are geometries of stable species and reaction intermediates, reaction pathways and mechanisms, NMR shifts, harmonic vibrational frequencies, spectroscopic constants, potential energy curves, and with the help of some statistical mechanics, various kinetic and thermodynamic quantities.

However, in order for the electronic structure of a system to be obtained, the well-known Schrödinger equation must be solved, which unfortunately is solvable exactly only for the hydrogen atom and hydrogen-like atoms. Because of this, each of the *ab initio* methods must use various techniques of approximation with varying degrees of accuracy and computational efficiency. Typically, the methods that are used can be categorized as either a variational or a perturbation method. They can also be further categorized as single-reference (SR) or multireference (MR) methods, depending on whether they assume the energy of system can be qualitatively described by a single Slater determinant reference state (which, in simplest cases, can be thought of as a single electron configuration reference state), or whether it allows for more than one reference state.

The most foundational methods used in modern quantum chemistry use a variational approach and are either known as the Hartree–Fock (HF) approximation (or self-consistent field (SCF) method) or Density Functional Theory (DFT) depending on whether correlation energy is neglected or approximated. HF is a highly efficient method that owes its favorable scaling with respect to system size primarily to two approximations, it neglects both dynamic and static electron correlation. This is because HF is a mean field method (neglects dynamic) and a SR method (neglects static). However, both static and dynamic electron correlation are important for accurately describing the electronic structure of a system. Therefore, in practice, HF is typically only used

as a starting point for other, higher-level methods. These methods aim to add electron correlation onto the HF wavefunction, and are therefore aptly named as post-HF methods. These include SR variational methods, such as configuration interaction (CI); MR variational methods, such as multiconfigurational self-consistent field (MCSCF),[8,9] multireference configuration interaction including single and double excitations (MRCISD); SR perturbation methods, such as second order Møller–Plesset perturbation theory (MP2);[10] and MR perturbation methods, such as second-order Generalized Van Vleck Perturbation Theory (GVVPT2).[11–14] There is another exceedingly popular method known as density functional theory (DFT), which is also considered a mean field method. However, practical implementations of DFT includes dynamic electron correction through the use of parameterizations and considers the wavefunctions as functionals of electron density rather than as functions of electron coordinates. Descriptions of each of these above-mentioned methods are given in Chapter 2, including the rationale for when and how to use them.

1.2 Theoretical Investigations of Metalloid Oxides

The combustion of coal for power production is one of the primary ways by which the trace elements such as arsenic and antimony are released into the environment.[15] In flue gases, these trace elements exist most often as oxides.[16,17] Because of high temperatures, low concentrations, and complex environments, experimental results are difficult to obtain. Consequently, little is known of the activity of these oxides in flue gases, especially soon after they are formed. Despite the prevalence and importance of mitigating these trace element oxides, critical information on the thermodynamics of plausible multiconfigurational intermediates and transition states

in reaction pathways have been missing prior to these studies.[18–20] Therefore, in this work, reaction and isomerization thermodynamics, as well as the potential pathways for gas phase accretion were investigated. Studies that have been conducted to date have all used single-reference *ab initio* methods, thus the multireference character of these trace metal oxides needed to be assessed.

In this work, the electronic structures of arsenic oxides, AsO, AsO_2, AsO_3, their dimers, As_2O_3 and As_2O_5, along with the corresponding antimony monomeric and dimeric oxides, were investigated using GVVPT2 with comparison to MRCISD when possible. Geometry optimizations as well as harmonic vibrational frequencies of both the monomers and dimers were obtained using DFT. One of the chief challenges of this work was to develop active spaces (i.e., the molecular orbitals (MOs) and electron configurations that are considered important to bonding and describing the molecule) that could be used to describe all isomers within a given series of the oxide. The difficulty was due, *inter alia*, to the complex variations in bond formation and breaking along the isomerization coordinates while keeping the active spaces small enough to allow for timely calculation convergence. The rationale behind developing the active spaces used in these studies is given in Chapter 2. The results along with the computational details are presented in Chapter 3. This study was recently extended to a new series of oxides also relevant to coal combustion, SbO, SbO_2, SbO_3, their dimers, Sb_2O_3 and Sb_2O_5 presented in Chapter 4.

1.3 Computational Studies of the Catalytic Degradation of Lignin

Though lignin is one of the most renewable carbon sources on earth,[21] it has only seen limited industrial use due to its chemical inertness and complex structure; thus lignin is typically disposed of as non-commercialized waste product. However, if a viable path could be found for the decomposition of lignin, the by-products could be used to replace high-value petrochemicals. Because of this, much research has gone into the search for a catalyst that provides a low energy degradation pathway. Catalytic decomposition by amorphous silica-alumina (ASA) based catalysts may be a viable path. However, there have been no computational studies on the mechanisms of catalytic bond cleavage of lignin. Moreover, the effects of metal doping of ASA on its catalytic activity toward lignin have not been investigated computationally.

In this work, DFT calculations with the B3LYP hybrid functional, as implemented in the NWChem software package, were used to elucidate this information. First, a cluster model of an ASA catalyst was developed and used, being that the band gap of ASA is sufficiently large that periodic boundary conditions are not required. Secondly, in an effort to successfully model a cluster of an amorphous catalyst, a new computational protocol needed to be developed. This protocol was used to study the adsorption of a monomer and a β–O–4 linked dimer to the surface. The catalytic activity was then adjusted by doping the ASA with various metals.

One of the key goals of this research was to aid experimentalists in the design of novel catalysts for lignin decomposition by developing a protocol, which could be used to screen various catalysts and elucidate a deeper mechanistic understanding of the catalytic process. For this, the above DFT protocol was developed. This work was set up as part of the DakotaBioCon collaboration, in which four institutions across North and South Dakota participated.

1.4 Theoretical Studies of Isomerization of the $Cu_2O_2^{2+}$ Core

The various bonding motifs of $Cu_2O_2^{2+}$ have drawn considerable attention over the years for several reasons. It is known to play an important role in various metalloenzymes such as oxy-tyrosinase, which oxidizes tyrosine residues to their corresponding *o*-quinones. In oxy-tyrosinase the $Cu_2O_2^{2+}$ core is thought to rapidly interconvert between the bis(μ-oxo) and side-on μ-η^2:η^2-peroxo isomers.[22] Understanding this rapid equilibrium then becomes crucial in the study of the oxidation mechanism of tyrosine. Therefore, several computational studies have been devoted to attempting to better understand this interconversion. This then brings us to the other key reason for the considerable attention this system has received.

During the isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo there is a rapid change in both dynamic and static electron correlation while possessing a variable amount of biradical character. These properties are very difficult for one single method to adequately model. Thus, this system has become a type of benchmark "torture track" for various theoretical models.[3] Though some methods, e.g. CR-CC(2,3),[23,24] have shown that they can do fairly well on this system, multireference perturbation theories, e.g., CASPT2,[25] have had limited success. This comes, in part, due to the difficulty of selecting a physically reasonable, balanced active space which captures all the important chemistry that happens along the isomerization coordinate. Along with that difficulty, the proper perturbative treatment of the electron correlation effects for transition metal complexes is needed. The GVVPT2 method has shown itself to provide a high level of accuracy for difficult systems comparable to the $Cu_2O_2^{2+}$ core while still enjoying favorable scaling (i.e. N^5, where N is the number of basis functions). This task is accomplished while not suffering from

the well-known intruder state problem that vexes most of the other multireference perturbation theories which possess similar scaling.

Despite the advantages of GVVPT2, it has been limited in the case of $Cu_2O_2^{2+}$ due to the limitations of MCSCF. It was previously thought that a very large active space was required in order to correctly describe transition metal complexes, for which MCSCF had a difficult time converging. However, work recently conducted in our group by Dr. Patrick Tamukong showed that a small valence picture active space can capture the important chemistry of transition metals.[26,27] In these active spaces, it is assumed that the most important contributors to bonding are the MOs derived from the valence shells of the participating atoms. Although atomic sub-shells just above or below the valence may contribute, such contributions are generally minimal. These MOs are often further partitioned into subspaces in which bonding and antibonding orbitals of the same kind are in the same subspace (e.g., grouping pi orbitals together). An active space like this had not been previously attempted in a calculation on $Cu_2O_2^{2+}$. Using this valence picture, an active space was developed as a starting guess and converged MCSCF results were obtained at a few geometries. These results were then used to develop a numerically-based technique for active space refinement that improves the active space in an iterative fashion (more detail of this technique is given in Chapter 2). After several iterations, a stable active space that had excellent MCSCF convergence behavior for the entire isomerization coordinate was obtained using both triple and quadruple excitations. GVVPT2 was then run using those orbitals and compared to other, more expensive methods.

1.5 Structure of Dissertation

This dissertation continues with Chapter 2 laying the theoretical foundations followed by an introductory section for the various methods used in the body of this work. The descriptions of the methods are given in a way that highlights the advantages and disadvantages of each and shows how they can be used together to gain chemical insight. Important methodological development or manipulation that was useful for completing the work presented in this dissertation (e.g., active space development) will also be discussed. In Chapter 3, the application of these methods in gaining insight into the chemistry of various monomeric and dimeric arsenic oxides is presented. The study was then extended in Chapter 4 to various antimony oxides in which relativistic effects were required. The computational prediction of the catalytic degradation of lignin is presented in Chapter 5. Lastly, the results obtained from the calculations of the isomerization of the biologically and theoretically relevant system of $Cu_2O_2^{2+}$ as well as a description are presented in Chapter 6. Finally, the overall conclusion and future directions are given in Chapter 7.

2 METHODS OF ELECTRONIC STRUCTURE

2.1 Theoretical Foundation

In order to understand and fully appreciate the results presented in this dissertation, the development of computational protocols, and the manipulation of existing techniques presented in later chapters, an introductory knowledge of the theoretical methodology must be provided. This chapter is written in an effort to accomplish that goal. It is with that spirit that we shall begin with one of the most important equations in all of quantum chemistry, the time-independent Schrödinger equation. When written using bra–ket notation, it is given as

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{2.1}$$

It is with this equation, if correctly solved, all of the time-independent information about the system in question can be obtained. Here, $\hat{H}$ is the Hamiltonian operator for electrons and

nuclei interacting via Coulomb interactions, $|\Psi\rangle$ is the wave function of the system and E is the corresponding energy of that system. $|\Psi\rangle$ can be thought of as an oracle of sorts, where if one could figure out the 'right question' to ask, one would be able to find out any information one could ever need to know about that system. Typically, we are most interested in finding the energy of the system, but many other useful quantities can be found. The nonrelativistic Hamiltonian operator for one particle, which corresponds to the total energy of the system, consists of a kinetic energy operator, $\hat{T}$, and a potential energy operator, $\hat{V}$, and is given by

$$\hat{H} = \hat{T} + \hat{V} = -\frac{\hbar}{2m}\nabla^2 + V(\mathbf{r}) \tag{2.2}$$

Here $\hbar = h/2\pi$, m is the mass of the particle and ∇^2 is the Laplacian operator. When the Hamiltonian operator is expanded, applied to a system of M atoms and N electrons, and atomic units (a.u.) are used (i.e., $e^2 = \hbar = m_e = 1$, energies are in Hartrees and distances are in Bohr radii), Eq. 2.2 then becomes[1]

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.3}$$

The first and second terms represent the contributions from the kinetic energy operators for the electrons and nuclei, respectively. The last three terms make up the potential energy operator of the system in the form of the Coulombic attraction between electrons and nuclei, the repulsion potential between electrons, and the repulsion potential between the nuclei, respectively. Here, ∇_A^2, M_A, and Z_A denotes the spatial gradient, the ratio of the mass of the nucleus to the mass of an electron, and the charge of nucleus A, respectively; and ∇_i^2 is the second spatial derivative of electron i. Lastly, the three variables, r_{iA}, r_{ij}, and R_{AB} represent the distances between electron

i and nucleus A, electrons i and j, and nucleus A and B. These interactions are each described pictorially below in Figure 2.1 and are each described formally as

$$r_{iA} = \left|\vec{r}_{iA}\right| = \left|\vec{r}_i - \vec{R}_A\right|, \tag{2.4}$$

$$r_{ij} = \left|\vec{r}_{ij}\right| = \left|\vec{r}_i - \vec{r}_j\right|, \tag{2.5}$$

$$R_{AB} = \left|\vec{R}_{AB}\right| = \left|\vec{R}_A - \vec{R}_B\right|. \tag{2.6}$$

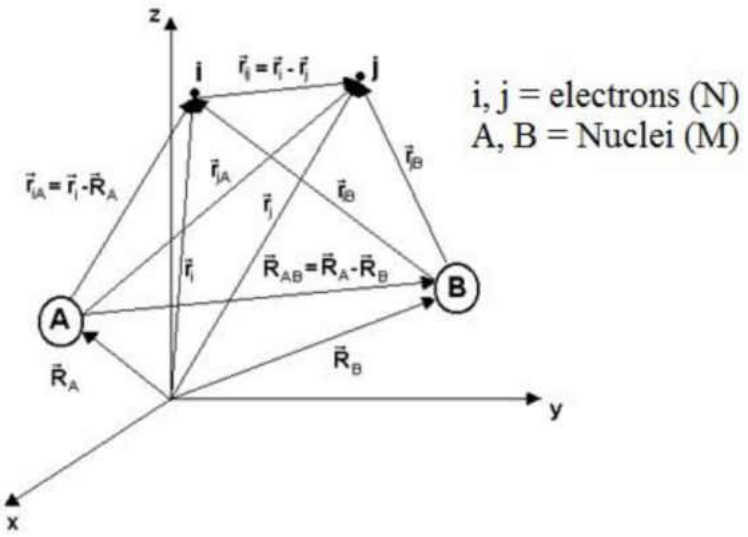


Figure 2.1. Descriptions of various interaction variables in a system consisting of nuclei A and B along with electrons i and j (Image taken from Szabo and Ostlund).[1]

As mentioned in Chapter 1, approximations must be taken in order to solve the Schrödinger equation (Eq. 2.1) for anything larger than 1-electron hydrogen-like atoms. The most well-known of these is the so-called Born–Oppenheimer approximation.[28] The fundamental assumption behind the Born–Oppenheimer approximation is that the motions of the electrons and of the nuclei can be

separated. This is because the much slower speeds of the relatively massive nuclei can be approximated as stationary when compared to the relatively light and fast moving electrons. Therefore, the electrons are thought of as moving in a stationary potential generated by the nuclei fixed at a given geometry. This greatly simplifies the evaluation of the Schrödinger equation. An example of this can be seen by considering the computation of the energy and wave function of ethanol with its 9 nuclei and 26 electrons. Without the Born–Oppenheimer approximation, the Schrödinger equation becomes a partial differential eigenvalue equation with 105 variables to solve for, due to the three spatial coordinates of each particle. However, with the Born–Oppenheimer approximation, the problem can be divided into two separate, easier to solve partial differential eigenvalue equations and combined later. This can be accomplished because the wave function of the electrons and the nuclei can be treated separately from each other as

$$|\Psi\rangle \cong |\Psi^{el}\rangle|\Psi^{nuc}\rangle \tag{2.7}$$

and the Hamiltonian, Eq. 2.3, can be modified by neglecting the second term, i.e., the kinetic energy term of the nuclei, and the fifth term, i.e., the repulsion potential between the nuclei, is considered constant. This is known as the electronic Hamiltonian and is given by

$$\hat{H_{el}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{2.8}$$

Then, the electronic wave function and the corresponding electronic energy are determined by the electronic Schrödinger equation, which can be written as

$$\hat{H}_{el}|\Psi^{el}\rangle = \left[\sum_{i=1}^{N}\hat{T}_i^{el} + \sum_{i=1}^{N}\sum_{A=1}^{M}\hat{V}_{i,A}^{el,nuc} + \sum_{i=1}^{N}\sum_{j>i}^{N}\hat{V}_{i,j}^{el,el}\right]|\Psi^{el}\rangle = E_{el}|\Psi^{el}\rangle \tag{2.9}$$

where the electronic kinetic energy, the electron-nuclear attraction, and the electron-electron repulsion have been denoted as $\hat{T}_i^{el}$, $\hat{V}_{i,A}^{el,nuc}$, and $\hat{V}_{i,j}^{el,el}$, respectively. Once Eq. 2.8 is solved, the total energy of the system, E_{tot} can be obtained by adding on the constant potential resulting from the fixed nuclei to the electronic energy, E_{el}, giving

$$E_{tot} = E_{el} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.10}$$

Here, E_{tot} can be thought of as the potential energy of the system at a given molecular geometry. If Eq. 2.8 is solved at a sufficient number of geometrical points, the potential energy surface that influences the motion of the nuclei will be obtained. This is useful in many ways with respect to chemistry, e.g., when attempting to find equilibrium and transition state molecular geometries or vibrational frequencies.

Unfortunately, Eq. 2.8 can only be solved exactly for a few very simple systems due to the problematic term, $\hat{V}_{i,j}^{el,el}$ (i.e., the electron-electron repulsion term). Solving and approximating this term in various ways in order to solve Eq. 2.8 has been much of the focus of quantum chemists since the formulation of the Schrödinger Equation. These various approximations of Eq. 2.8, which intrinsically hold various degrees of accuracy and computational efficiency, will be the focus of the rest of the methods presented in this chapter.

2.2 Hartree–Fock Approximation

It will serve us well to begin our discussion of the various ways to solve Eq. 2.7 by starting at the same place most modern day computational calculations in quantum chemistry begin, the well-known Hartree–Fock (HF) approximation.

One simple way to approximate $\hat{V}_{i,j}^{el,el}$ is to replace it with an average or mean electric field generated by all of the other electrons in the system, $\hat{V}_{i,j}^{el,el} \approx V_i(\phi_1, \phi_2, ..., \phi_N)$. Therefore, instantaneous interactions of the electrons would be neglected and the electronic Hamiltonian for each electron along with the corresponding energy could then be separated. Similarly, the electronic wave function, $|\Psi^{el}(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)\rangle$, would become merely the product of N one-electron wave functions, $\prod_i^N |\phi^{el}(\mathbf{r}_i)\rangle$. This mean-field approximation is known as the Hartree approximation. Through the use of these 'Hartree products,' the electronic Schrödinger equation becomes

$$\hat{H}_{el}|\Psi^{el}\rangle = \sum_{i=1}^{N}\left[\hat{T}_i^{el} + \sum_{A=1}^{M}\hat{V}_{i,A}^{el,nuc} + \hat{V}_i(\phi_1, \phi_2, ..., \phi_N)\right]\prod_i^N |\phi_i^{el}(\mathbf{r}_i)\rangle = \sum_i^N E_i \prod_i^N |\phi_i^{el}(\mathbf{r}_i)\rangle = E_{el}|\Psi^{el}\rangle \tag{2.11}$$

However, there a serious problem in the Hartree approximation that is unaddressed at this point. It does not account for the Pauli exclusion principle or, more generally, the antisymmetry principle. Electrons are fermions, meaning a wave function of electrons must be antisymmetric with respect to interchange of any set of spin and space coordinates, denoted as $\mathbf{x} = \{\mathbf{r}, \omega\}$. In this notation, $\mathbf{r}$ corresponds to the three spatial degrees of freedom and ω represents the spin coordinate (i.e., spin α or spin β).

When including the spin, the Hartree product of orbitals changes from the exclusively spatial orbital form, $\phi^{el}(\mathbf{r}_i)$, into spin orbitals, $\chi^{el}(\mathbf{x}_i)$, then we have

$$\Psi^{el}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \chi_i^{el}(\mathbf{x}_1), \chi_j^{el}(\mathbf{x}_2), ..., \chi_k^{el}(\mathbf{x}_N) \tag{2.12}$$

In order to satisfy the antisymmetry principle the relation,

$$\Psi^{el}(\mathbf{x}_1, \mathbf{x}_2) = -\Psi^{el}(\mathbf{x}_2, \mathbf{x}_1) \tag{2.13}$$

must be obeyed. This is accomplished through the use of the so-called Slater determinant. It is the inclusion of Slater determinants that moved the Hartree approximation into what is now known as the Hartree–Fock approximation. A Slater determinant is a determinant of spin orbitals, and, for a system of N electrons, it can be written as

$$\Psi^{el}(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x_1}) & \chi_2(\mathbf{x_1}) & ... & \chi_N(\mathbf{x_1}) \\ \chi_1(\mathbf{x_2}) & \chi_2(\mathbf{x_2}) & ... & \chi_N(\mathbf{x_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x_N}) & \chi_2(\mathbf{x_N}) & ... & \chi_N(\mathbf{x_N}) \end{vmatrix} \tag{2.14}$$

where the factor $\frac{1}{\sqrt{N!}}$ is the normalization factor. The rows of the determinant are labeled by electron coordinates and the columns are labeled by the spin orbitals. Therefore, if two rows are interchanged, it would be as if the coordinates of two electrons were interchanged. Due to the mathematical nature of the determinant, this would change the sign of the wave function, resulting in satisfaction of the requirements of the antisymmetry principle. Furthermore, if two electrons with the same coordinates occupy the same spin orbital, there would be two equal columns in the

determinant, making the determinant equal to zero. This representation fully satisfies the Pauli exclusion principle.[1] By this representation, as an example, the Hartree–Fock wave function for a simple two electron system is given as

$$\Psi^{el}(\mathbf{x}_1, \mathbf{x}_2) \cong \Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2!}} \left[\chi_1(\mathbf{x_1})\chi_2(\mathbf{x_2}) - \chi_1(\mathbf{x_2})\chi_2(\mathbf{x_1})\right] \tag{2.15}$$

Typically, in the HF approximation, the electronic Hamiltonian is written as

$$\hat{H}_{el} = \sum_{i=1}^{N} \left[h(i) + \hat{V}_i(\phi_1, \phi_2, ..., \phi_N)\right] = \sum_{i=1}^{N} \left[h(i) + v^{HF}(i)\right] = \sum_{i=1}^{N} \hat{f}_i \tag{2.16}$$

where $\hat{f}_i$ is the Fock operator, $v^{HF}(i)$ is the average potential experienced by the ith electron due to the electric field from the rest of the electrons, which includes both the classical Hartree term and the so-called exchange term that takes into account antisymmetry, and $h(i)$ is a Hamiltonian containing only one-electron terms, given by

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \tag{2.17}$$

In the work presented in this dissertation, the HF method was used to obtain a starting guess of the MOs from the AOs. The MOs would then be used in some higher levels of theory to generate a more accurate description of the system. In practice, the MOs are obtained as a linear combination of basis functions to describe the AOs, which are typically Gaussian-type functions centered around the atoms.[1] Thus, the optimized MOs, denoted by ψ_i, are given by

$$\psi_i = \sum_{\mu=1}^{n} C_{\mu i}\phi_\mu \tag{2.18}$$

where ϕ_μ is the μth basis function and n is the total number of the basis functions in the basis set. The way that the HF method obtains the MOs is by iteratively minimizing the energy, E_{HF}, where each iteration results in a new, more correct value of the coefficients, $C_{\mu i}$. They are then used to update the HF wave function for the next iteration. The iterations continuously improve the energy until the change in energy with respect to the change in the coefficient, $C_{\mu i}$, is negligible. This method, in which the HF equation is solved, is called the self-consistent field (SCF) method. When fully converged, the resultant MOs can be a reasonable starting guess for further post-HF calculations. This minimization process is an example of the variational method. The variational theorem states that the energy of an approximate wave function is always larger than the true energy. The HF energy, when the wave functions are normalized, is calculated by

$$E_{HF} = \langle \Psi_{HF} | \hat{H}_{el} | \Psi_{HF} \rangle \tag{2.19}$$

There are two ways the HF approximation can handle spin: unrestricted Hartree–Fock (UHF) and restricted Hartree–Fock (RHF). The type used in this dissertation is RHF, where the same set of spatial orbitals are used to describe two electrons that possess opposite spins. Though the HF method is computationally inexpensive and can give adequate starting estimates for MOs, neglecting electron correlation can lead to large deviations from experimental results. There have been a number of approaches developed in an attempt to overcome this limitation, collectively these methods are known as post-HF methods. These methods have been formulated in a way to include electron correlation into the wave function. One of these approaches, 2nd order Møller-Plesset perturbation theory (MP2),[10] treats correlation as a perturbation of the HF wave function, and a summary of that method is given below.

2.3 Second Order Møller–Plesset Perturbation Theory

Møller–Plesset (MP) perturbation theory improves on the Hartree–Fock method by adding electron correlation effects by means of Rayleigh-Schrödinger (RS) perturbation theory using an unperturbed Hamiltonian of Slater determinants and one-electron energies. In practice, this perturbational correction is usually to second order, i.e., MP2. MP2 is considered the simplest and computationally least expensive *ab initio* method for including electron correlation effects.[29,30] In the work presented in this dissertation, MP2 was used to take the HF solution, including the basis set and molecular geometry information, and create a rough estimate of the unoccupied MO's energy and give the relative occupancies of the occupied MOs (i.e., it gives natural orbitals). This information was then used to develop an active space, i.e., the MOs which are considered as the most chemically relevant.

In MP2, rather than approaching the system variationally, like the HF treatment of orbitals, the total Hamiltonian of the system is expanded into two parts. The first part, $\hat{H}_0$, is called the zero-order term, while the second part, $\lambda\hat{V}$, is called the perturbation term. The form of the MP Hamiltonian is then

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{2.20}$$

where λ is an arbitrary real parameter between 0 and 1 and $\hat{H}_0$ is given by

$$\hat{H}_0 = \sum_{i=1}^{N} \left[h(i) + v^{HF}(i) \right] \tag{2.21}$$

and the perturbation, V, is given by

$$V = \sum_{i<j}^{N} \frac{1}{r_{ij}} - \sum_{i=1}^{N} v^{HF}(i) \tag{2.22}$$

MP perturbation theory is a special case of RS perturbation theory. In RS perturbation theory the perturbed wave function and perturbed energy are expressed as a power series in λ. The perturbed wave function of the system is

$$\Psi = \lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i \Psi^{(i)} \tag{2.23}$$

and the perturbed energy of the system is given as

$$E = \lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i E^{(i)} \tag{2.24}$$

When Eqs. 2.21, 2.23 and 2.24 are substituted into the time-independent Schrödinger equation (Eq. 2.1), a power series representation of the equation is obtained,

$$\left(\hat{H}_0 + \lambda\hat{V}\right)\left(\lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i \Psi^{(i)}\right) = \left(\lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i E^{(i)}\right)\left(\lim_{n\to\infty} \sum_{i=0}^{n} \lambda^i \Psi^{(i)}\right) \tag{2.25}$$

where $n \to \infty$ and λ^i corresponds to a ith-order perturbation. In the case of MP2, $i = 2$ for the energy term (second order energy), but the first order perturbed wave function.

Since the Hartree–Fock Slater determinant, $|\Psi_0^{(0)}\rangle$, is an eigenfunction of $\hat{H}_0$, we have

$$\hat{H}_0|\Psi_0^{(0)}\rangle = E_0^{(0)}|\Psi_0^{(0)}\rangle \tag{2.26}$$

where the corresponding zeroth-order energy eigenvalue is given by

$$E_0^{(0)} = \sum_i \varepsilon_i \tag{2.27}$$

in which ε_i is the orbital energy of the ith electron. The first-order correction to the energy of the system is given by

$$E_0^{(1)} = \langle \Psi_0 | \hat{V} | \Psi_0 \rangle \tag{2.28}$$

and after the expression for $\hat{V}$, given in Eq. 2.22, is substituted in we have

$$E_0^{(1)} = \langle \Psi_0 | \sum_{i<j} \frac{1}{r_{ij}} | \Psi_0 \rangle - \langle \Psi_0 | \sum_i v^{HF}(i) | \Psi_0 \rangle \tag{2.29}$$

It is now useful to introduce a shorthand notation for two types of integrals over spin orbitals in which square brackets are used.

$$[i|h|j] \equiv \int d\mathbf{x}_1 \chi_i^*(1) h(\mathbf{r_1}) \chi_j(1)) \tag{2.30}$$

where the integration variable $\mathbf{x}_i$ denotes both the spatial and spin coordinates of electron i, as mention above, and $\mathbf{r}_i$ denotes only the spatial part, which is used here because we are assuming the Hamiltonian does not depend on spin. For a two-electron operator we have

$$[ij|kl] \equiv \int d\mathbf{x}_1 \int d\mathbf{x}_2 \chi_i^*(1) \chi_j(1) \frac{1}{r_{12}} \chi_k^*(2) \chi_l(2) \tag{2.31}$$

Using this notation and the relation

$$\langle \Psi_0 | \sum_{i<j} \frac{1}{r_{ij}} | \Psi_0 \rangle = \frac{1}{2} \sum_i \sum_j \left([ii|jj] - [ij|ji] \right) \tag{2.32}$$

Eq. 2.29 becomes

$$E_0^{(1)} = \frac{1}{2} \sum_i \sum_j \left([ii|jj] - [ij|ji] \right) - \sum_i \left[i | v^{HF}(i) | j \right] \tag{2.33}$$

and since the expectation value of $v^{HF}(i)$ is given by

$$\sum_i \left[i | v^{HF}(i) | j \right] = \sum_i \sum_j \left([ii|jj] - [ij|ji] \right) \tag{2.34}$$

it simplifies to give the first-order energy

$$E_0^{(1)} = -\frac{1}{2} \sum_i \sum_j \left([ii|jj] - [ij|ji] \right) \tag{2.35}$$

Note that when $E_0^{(0)}$ and $E_0^{(1)}$ are summed together, the HF energy, E_0^{HF} is simply obtained. Thus, in order to obtain the needed correlation energy, second-order or higher is needed.

The second-order correction to the energy of the system is given by

$$E_0^{(2)} = \sum_n \frac{\left| \langle \Psi_0 | \hat{V} | \Psi_n \rangle \right|^2}{E_0^{(0)} - E_n^{(0)}} \tag{2.36}$$

where $|\Psi_n\rangle$ is a Slater determinant corresponding to an excited electron configuration relative to the HF reference state. Only a double excitation determinant, $|\Psi_{ij}^{ab}\rangle$, yields a non-zero contribution to the energy with

$$\langle \Psi_0 | \sum_{i<j} \frac{1}{r_{ij}} | \Psi_{ij}^{ab} \rangle = [ia|jb] - [ib|ja] \tag{2.37}$$

The zero-order energy eigenvalue of the doubly excited determinant is $E_0^{(0)} - \varepsilon_i - \varepsilon_j + \varepsilon_a + \varepsilon_b$, i.e.,

$$\hat{H}_0 | \Psi_{ij}^{ab} \rangle = (E_0^{(0)} - \varepsilon_i - \varepsilon_j + \varepsilon_a + \varepsilon_b) | \Psi_{ij}^{ab} \rangle \tag{2.38}$$

Therefore, the second-order energy correction is given by

$$E_0^{(2)} = \sum_{i<j} \sum_{a<b} \frac{\left| \langle \Psi_0 | \sum_{i<j} \frac{1}{r_{ij}} | \Psi_{ij}^{ab} \rangle \right|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} = \sum_{i<j} \sum_{a<b} \frac{\left| [ia|jb] - [ib|ja] \right|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \tag{2.39}$$

The MP2 energy can then be written as

$$E_0^{MP2} = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} = E_0^{HF} + E_0^{(2)} \tag{2.40}$$

The type of MP2 used in this dissertation is known as Restricted Second Order Møller–Plesset Perturbation Theory (RMP2) where the same set of spatial orbitals are used to describe two electrons that possess opposite spins. RMP2 provides an initial approximation of the MOs in the form of what is known as natural orbitals (NOs) in its output. These NOs give values ranging from 2.00-0.00, where a NO with a value of 2.00 would correspond to an MO that is always doubly occupied and a NO with a value of 0.00 would correspond to an MO that is never occupied. The NOs generated from the RMP2 calculations are helpful in devising the active spaces used in later, higher level calculations.

Since RMP2 is a single-reference method, only a single ground state electron configuration can be well described, thus not accounting for static electron correlation. However, this limitation

can be addressed through the use of other post-HF methods, such as the multiconfigurational self-consistent field (MCSCF) method. MCSCF includes static electron correlation by expanding the wave function in terms of a linear combination of Slater determinants rather than a single Slater determinant. MCSCF can also serve as a reference for multireference methods and require the use of what is known as an active space. MCSCF, active spaces, and macroconfigurations are described in the next section.

2.4 Multiconfigurational Self-Consistent Field Theory, Active Spaces, and Macroconfigurations

MCSCF is a multiconfigurational/multireference method since several electron configurations are considered important rather than just one. This is accomplished through the inclusion of many Slater determinants to describe the ground and possibly low-lying excited state of a system rather than just a single Slater determinant, which was the approach that the previous methods discussed have used. In the presented work of this dissertation, MCSCF was used to refine the MOs and energies obtained from RMP2.

Though there are some systems which can be reasonably described through the use of single-reference methods (i.e., where Hartree–Fock is a decent zero-order approximation from which correlation energies can be calculated), there are many cases where a single-reference description is inadequate to capture the correct chemistry. Processes such as bond breaking and forming as well as many excitations of the electronic state of an atom or molecule require multireference

methods. This is because the HF Slater determinant may no longer dominate the system's wave function due to other important electronic configurations. Several well-known examples where a single-reference description fails are: diradicals, unsaturated transition metals, excited states and often transition states. In these cases, the use of a HF wave function as a reference fails to produce sufficiently accurate results, thereby causing subsequent higher-level, but still approximate, methods to produce inaccurate results. Therefore, a multiconfigurational method is needed to generate a reference wave function which can be used by a multireference perturbation theory, such as GVVPT2. This is particularly true for the work presented in this dissertation, since many of the systems involve the formation and breaking of one or more bonds and many of the molecules considered are transition state structures. The MCSCF method has shown itself to be an excellent way to generate a multireference wave function, the general form of which is given by

$$|\Psi_{MCSCF}\rangle = \sum_{I} A_I |\Phi_I\rangle \tag{2.41}$$

where A_I are the corresponding configuration mixing coefficients of the Ith configuration state function (CSF), Φ_I, which is given by

$$\Phi_I = \mathcal{A}_I \left\{ \prod_{i \in I} \psi_i \right\} \tag{2.42}$$

CSFs are linear combinations of Slater determinants that transform as an eigenfunction of S^2.

The difference in the arrangement of electrons in the MOs, ψ_i, and the spin-coupling differentiate one electronic configurations from another, where ψ_i was given earlier in Eq. 2.18 as a linear combination of basis functions. It is worth noting that in MCSCF the configuration mixing coefficients, A_I, as well as the MO expansion coefficients, $C_{\mu i}$, are variationally optimized. As

will be seen later on, this is a strategy different from that of other methods (e.g., configuration interaction (CI)), where only the configuration mixing coefficients are optimized.

Unlike HF and MP2, a user defined active space is required when a MCSCF calculation is conducted. The active space (or model space) dictates which MOs and electron configurations are most important to the chemical activity of the system. There are two widely used approaches to include multiple electron configurations in the chosen active space MOs, the complete active space SCF (CASSCF)[8] approach and the restricted active space SCF (RASSCF).[9] In CASSCF the valence electrons are distributed in all possible ways among MOs in the active space. However, this can be very computationally expensive, since the number of configuration state functions (CSFs) rapidly increases with the number of active MOs. In RASSCF, a smaller set of CSFs can be obtained by restricting the number of electrons in certain orbital groups. More specifically, the orbitals in a RAS model space are broken up into three categories: orbitals with a limited number of vacancies (i.e., holes), fully active orbitals and orbitals with a limited number of electrons in the CSFs used to build up the multiconfigurational wave function. CASSCF and RASSCF are compared below in Table 2.1.

Rather than being limited to a CAS-type or a RAS-type active space, the work in this dissertation used what is known as the macroconfiguration approach developed in our group.[31] In the macroconfiguration approach, the construction of chemically relevant, physically reasonable and relatively compact model spaces can be achieved. This method provides far more efficient treatment of chemically relevant configurations than CAS-type active spaces while providing even more flexibility than RAS-type active spaces. By using the macroconfiguration approach, when the active space and excitations are set up by the user in a chemically reasonable way, methods like MCSCF, MRCISD and GVVPT2 are sped up considerably while still enjoying the high-level

Table 2.1. Comparison of the orbital classifications for CASSCF and RASSCF.

DESCRIPTION	CASSCF	RASSCF	DESCRIPTION
All orbitals have occupations of 0	EXTERNAL	EXTERNAL	All orbitals have occupations of 0
Active Space: Orbitals may have occupation number from 0 to 2	CAS	RAS3	**Active Space:** Orbitals may have occupation number from 0 to 2, but there cannot be more than n electrons
		RAS2	**Active Space:** Orbitals may have occupation number from 0 to 2, just like CAS
		RAS1	**Active Space:** Orbitals may have occupation number from 0 to 2, but there cannot be more than n holes
Frozen Core: All orbitals have occupations of 2	INACTIVE	INACTIVE	**Frozen Core:** All orbitals have occupations of 2

post-HF accuracies that these methods are capable of. With this approach, the user is enabled to construct truncated configuration spaces, which describe the static electron correlation while still providing good references for obtaining a proper description of the dynamic electron correlation. This can be accomplished due to the macroconfiguration approach's excellent systematic flexibility.

Generally, when considering a fixed partitioning, $\boldsymbol{N} = (N_1, N_2, ..., N_g)$ of N electrons in g disjoint orbital groups $\{G_1, G_2, ..., G_g\}$, the corresponding macroconfiguration is given by

$$\kappa(\boldsymbol{N}) : \{G_1^{N_1}, G_2^{N_2}, ..., G_g^{N_g}\} \tag{2.43}$$

Rather than giving the occupancies of separate orbitals as in a conventional electron configuration, a macroconfiguration specifies occupation numbers of fixed groups. Here the user can choose both the number of groups and which orbitals are in each group. The user is also able to

specify which occupancies they would like to consider in each group. In MCSCF-type schemes, when based on macroconfiguration descriptions of configuration spaces, the adequate treatment of static electron correlation effects for all states considered can be accomplished without suffering from the steep scaling of CASSCF.

Whether the macroconfiguration approach is used or not, MCSCF still has a few drawbacks. It requires a computationally intensive iterative procedure which is highly dependent upon the specified model space. This procedure will fail for sufficiently complex systems using certain model spaces. Many times, MCSCF has even taken longer to finish than higher-level methods that use MCSCF results as a starting point. This has led our group to experiment with other alternative methods to generate starting guesses for the MOs; one such method that was recently developed was to use DFT orbitals as a starting point. This will be discussed later in this chapter in the DFT section. The other important drawback of MCSCF is that it is generally only able to account for static electron correlation, thus missing a crucial amount of dynamic electron correlation energy. In order to accurately and effectively characterize the species studied in the work presented in this dissertation, dynamic electron correlation must be accounted for. This can be done very effectively and efficiently through the use of a method developed in our research group, GVVPT2, which is described below.

2.5 Second-Order Generalized van Vleck Perturbation Theory

Second-Order Generalized Van Vleck Perturbation Theory (GVVPT2) is a multireference perturbative method developed by Dr. Hoffmann's group.[13,14,32] In GVVPT2 the MCSCF wave function, which is generally only able to account for static electron correlation, can be used as the zeroth-order wave function. Then, through a perturbative treatment, dynamic electron correlation is added to the MCSCF wave function while still allowing the static electron correlation to be accounted for. It is worth noting that recently GVVPT2 was also shown to produce good results when starting with DFT orbitals.[?] Hoffmann2015)

Besides GVVPT2, dynamic electron correlation can be added onto the MCSCF wave function by other methods as well. The multireference configuration interaction including single and double excitations (MRCISD) method is one such method and is known to be highly accurate and it makes for a great benchmarking tool when the systems are small enough. However, MRCISD scales as approximately N^6, where N is the number of basis functions (i.e., a calculation twice as large takes 64 times longer to complete).[1] Thus, it is considered computationally expensive and is typically only used on relatively small systems, e.g., composed of a few atoms when a potential energy surface is needed. In this dissertation, MRCISD was used to confirm the validity of GVVPT2 results, and will be described in further detail later in this chapter.

There are techniques like multireference perturbation theory (MRPT) or quasidegenerate perturbation theory (QDPT), which are significantly more computationally efficient, since they scale as approximately N^5 (i.e., a calculation twice as large takes 32 times longer to complete). However, these methods have drawbacks, most notably by what is known as the "intruder state" problem. This problem occurs when the energy of an excited configuration is comparable to that of the zeroth-order wave function. This results in a denominator of the perturbative correction having a value close to zero, thus causing a highly divergent behavior in the potential energy

curve. Because of this "intruder state" problem and the complexity of their removal, the general applicability of MRPT and QDPT methods have been limited.

GVVPT2, however, is able to generate smooth potential energy surfaces, similar to the MRCISD method, while still retaining the scalability of MRPT and QDPT (i.e., as approximately N^5).[13] GVVPT2 is considered to be of the "diagonalize-then-perturb-then-diagonalize-again" variant of multireference perturbation theory and it is also an intermediate Hamiltonian variant of QDPT, making it a subspace selective method. It is worth noting that intermediate Hamiltonians have been shown to be a useful framework in a variety of both multireference (or quasidegenerate) perturbation and coupled cluster theories,[12,33–40] and they are seeing continued use and development.

In GVVPT2, a configuration space, L, that is used to describe the states of interest contains an expansion of the set of target wave functions,

$$|\vec{\Psi}_P\rangle = \{|\Psi_1\rangle, |\Psi_2\rangle, ..., |\Psi_{N_P}\rangle\} \tag{2.44}$$

where N_P is the number of the low-lying electronic states of interest. The reference space, L, is partitioned into two subspaces: (i) the model space, L_M, and (ii) the external space, L_Q. The model space, L_M (where the dimension of $L_M > N_P$), is usually of the MCSCF-type (i.e., it generally coincides with the MCSCF space). It describes, semi-quantitatively, the most important configurations needed to describe the state(s) of interest and is further divided into two subspaces: (i) the primary subspace, L_P, and (ii) the secondary subspace, L_S.

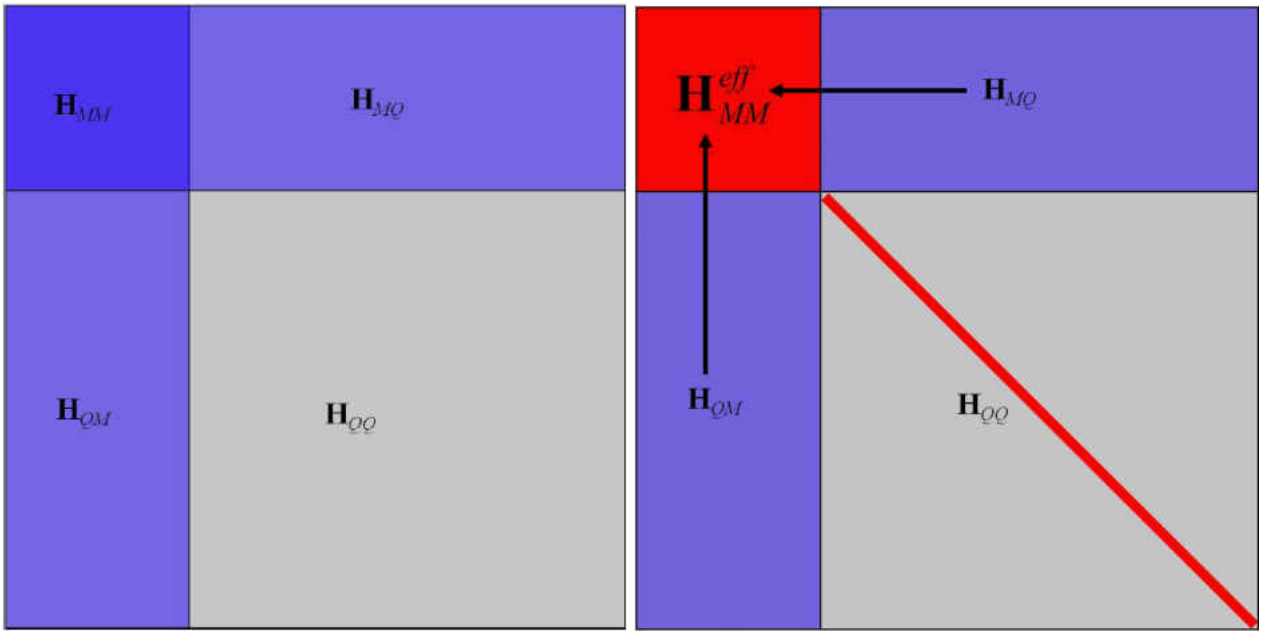
The primary subspace is composed of a small set of reference functions (e.g., one or more MCSCF CI vectors or MCSCF wave functions), $\{\Phi_P\}_{P=1}^{N_P}$, which involve internal (doubly occupied

core plus active) orbitals and are used to generate the L_M space. The secondary subspace, L_S, is the orthogonal complement to L_P in the model space, so

$$L_M = L_P \oplus L_S \tag{2.45}$$

The external space, L_Q, is composed of electron configurations that are connected to and obtained from the model space, L_M, through single and double electron excitations. In order to generate external configurations that can be expected to have only a perturbative effect on the lowest MCSCF states of interest, they need to be energetically well separated, which requires a physically reasonable MCSCF model space. It is also worth noting that the possible problem of quasidegeneracy between the secondary subspace, L_S, and the external subspace, L_Q, is entirely avoided in the GVVPT2 method.[12,33,41] The way in which the GVVPT2 method takes into account the interactions of the perturbed primary states in L_P with the unperturbed secondary states in L_S is by a final diagonalization of the effective Hamiltonian, $\mathbf{H}^{eff}$, (which is the same size as the MCSCF matrix) over the entire model space, L_M. This allows for the effects of dynamic electron correlation on the static electron correlation part. A graphical representation of this process, in which the Hamiltonian matrix is shown as a box divided up into its various parts, is given in Figures 2.2a and2.2b.

To gain a deeper understanding of the workings of the GVVPT2 method, consider a more general construction of an effective Hamiltonian in which the P-Q interactions are taken into account to the first-order in the system's wave function using a non-diagonal variant of QDPT[41] where the block-diagonal part of the Hamiltonian matrix in the L_M space is considered as the unperturbed Hamiltonian, given by



(a) (b)

Figure 2.2. The Hamiltonian matrix represented as a box. (a) The $\mathbf{H}_{MM}$ matrix has a small number of electron configurations and neglects dynamic correlation while the full H matrix has a large number of configurations and includes dynamic correlation. (b) By a final diagonalization, dynamic electron correlation is included into $\mathbf{H}^{eff}_{MM}$ by the contributions from $\mathbf{H}_{QM}$ and $\mathbf{H}_{MQ}$ while only the diagonal elements of $\mathbf{H}_{QQ}$ are kept.

$$H_0 = PHP + QHQ \tag{2.46}$$

and the off-diagonal block is considered as the perturbation, given by

$$V = PHQ + QHP \tag{2.47}$$

Then the expressions for the second-order perturbatively corrected Hamiltonian matrix, i.e., the intermediate effective Hamiltonian matrix, $\mathbf{H}^{eff}_{MM}$, is given in the model space L_M as

$$\mathbf{H}^{eff}_{PP} = \mathbf{H}_{PP} + \frac{1}{2}\left(\mathbf{H}_{PQ}\mathbf{X}_{QP} + \mathbf{X}^{\dagger}_{QP}\mathbf{H}_{QP}\right), \tag{2.48}$$

$$\mathbf{H}_{SP}^{eff} = \mathbf{H}_{SQ}\mathbf{X}_{QP}, \tag{2.49}$$

$$\mathbf{H}_{SS}^{eff} = \mathbf{H}_{SS}. \tag{2.50}$$

In the above equations, the elements of the anti-Hermitian matrix, $\mathbf{X}_{QP}$, (i.e., the P-Q rotational parameters) determine the contributions from the external CSFs to the final wave function. Without further approximation the P-Q rotation parameters satisfy the system of linear equations

$$\left(\mathbf{H}_{QQ} - E_0^P\right)\mathbf{X}_{QP} = -\mathbf{H}_{QP} \tag{2.51}$$

where $P \in [1, N_P]$ (i.e., P a member of the primary space), and $\{E_0^P = \langle\Phi_P|H|\Phi_P\rangle\}_{P=1}^{N_P}$ are the corresponding energies of the reference states, i.e., the MCSCF energies within the L_M space. However, in GVVPT2, the above second-order QDPT basic equations are approximated further, since they are block-diagonal and scale approximately as N^6. However, in accord with true QDPT methods, the effects of the secondary states in L_S on the perturbed primary states in L_P is allowed by $\mathbf{H}_{SP}^{eff} = \mathbf{H}_{SQ}\mathbf{X}_{QP}$. This is not the case in several other widely used multireference second-order perturbation methods, e.g., MRPT2,[42] CASPT2[43] and MCQDPT.[44] It has been shown that these effects can be important, though they are usually small.[13]

In the GVVPT2 method, as mentioned above and shown in Figure 2.2b, the off-diagonal matrix elements of the $\mathbf{H}_{QQ}$ matrix are discarded. The required diagonal elements, $H_{qq} - E_0^P$ of the matrices $\left(\mathbf{H}_{QQ} - E_0^P\right)$ are also approximated.[11,45] This could be done with conventional, Møller-Plesset-type, one-electron Hamiltonians,

$$\left(\mathbf{H}_{qq} - E_0^p\right) \approx \varepsilon_q^p - \varepsilon_0^p. \tag{2.52}$$

In many cases, this approximation works well. However, in some cases the above mentioned "intruder states" problem would occur, where the difference $\varepsilon_q^P - \varepsilon_0^P$ would be small and in some cases, even negative. This issue was overcome in GVVPT2;[13] "intruder state" are no longer a problem and smooth potential energy surfaces (PES) can be obtained even when two surfaces are quite close to one another. This was accomplished through the use of what could be called stabilized perturbation theory where the rotational parameter, $\mathbf{X}_{qp}$, can be calculated by

$$X_{qp} = -\frac{\tanh\left(\varepsilon_{\mathbf{m_e}}^p - \tilde{E}_{0,\mathbf{m_e}}^p\right)}{\varepsilon_{\mathbf{m_e}}^p - \tilde{E}_{0,\mathbf{m_e}}^p} H_{qp} \tag{2.53}$$

where $\tilde{E}_{0,\mathbf{m_e}}^p$ is the degeneracy-corrected zero-order energy of the pth primary state and $\mathbf{m_e}$ represents a given external configuration. Thus, all external CSFs that are created by a given external configuration, $\mathbf{m_e}$ (which possesses degenerate energy $\varepsilon_{\mathbf{m_e}}^p$), are considered simultaneously. More specifically, $\tilde{E}_{0,\mathbf{m_e}}^p$ is given by

$$\tilde{E}_{0,\mathbf{m_e}}^p = \frac{1}{2}\left(\varepsilon_{\mathbf{m_e}}^p - \varepsilon_0^p\right) + \frac{1}{2}\sqrt{\left(\varepsilon_{\mathbf{m_e}}^p - \varepsilon_0^p\right)^2 + 4\sum_{q\in\mathbf{m_e}} H_{qp}^2}. \tag{2.54}$$

These new X_{qp} parameters are continuous for all regions of the PES regardless of how small the difference $\varepsilon_q^p - \varepsilon_0^p$ becomes. It is worth noting that where $\varepsilon_{\mathbf{m_e}}^p \gg \varepsilon_0^p$, Eqs. 2.53 and 2.54 reduce to the same result as the original theory (Eq. 2.52), yet where $\varepsilon_{\mathbf{m_e}}^p \leq \varepsilon_0^p$, Eqs. 2.53 and 2.54 give small finite rotation parameters and small negative contributions of the "intruder states" to the correction energy of a given primary state. Thus, the "intruder state" problem is not an issue with GVVPT2,

while it still possesses a lower computational cost than MRCISD with a scaling of approximately N^5. This gives GVVPT2 excellent prospects and it is particularly useful for complicated chemical systems. GVVPT2 can give information on the occupancies of the MOs and the multireference character of the systems studied while costing significantly less than other methods of comparable accuracy. Thus GVVPT2 was a natural choice for the flagship method used in the work presented in much of this dissertation.

It is important to note that GVVPT2, in a similar way to MCSCF and MRCISD, is sensitive to the user-specified active space, so a reasonably accurate, chemically intuitive description is required. GVVPT2 does not have any restrictions on the model space that can be used and both complete and incomplete model spaces are supported. The construction of a well-balanced model space is known to be the main challenge when using multireference methods like GVVPT2. Fortunately, the macroconfiguration approach, described earlier,[31] is available and it can provide significant aid in this challenge.

In GVVPT2, as well as in other similar multireference methods, the use of the macroconfiguration approach enables: (i) the use of chemical intuition in constructing compact model spaces and in generating external spaces; (ii) the elimination of the need to store large lists of external configurations with subsets of such configurations efficiently regenerated during the calculation; and (iii) a drastic decrease of the number of configuration pairs that must be explicitly calculated, because of the screening of the effective Hamiltonian matrix blocks corresponding to non-interacting macroconfiguration pairs. Therefore, the macroconfiguration approach was used in the construction of the active spaces used in GVVPT2.

2.6 Multireference Configuration Interaction with Singles and Doubles

As mentioned earlier MRCISD is a highly accurate multireference method that can be used to add dynamic electron correlation to a MCSCF wave function. MRCISD adds dynamic electron correlation by substituting occupied orbitals with virtual orbitals in the individual CSFs. The SD part of MRCISD is a specification of the truncation of the expansion space to single and double substitutions (SD) which is usually required due to the steep increase in the number of CSFs as you approach what is known as full configuration interaction (FCI), and consequently the computational cost. In fact, FCI scales as $N!$, so it can only ever be used on the simplest of systems. The SD of MRCISD corresponds to single and double excitations from the occupied reference CSFs to the unoccupied virtual CSFs. For a $|\Phi_0\rangle$ corresponding to a MCSCF wave function, the FCI wave function can be expanded out as

$$|\Phi_0\rangle = c_0|\Psi_0\rangle + \sum_{i}^{occ.}\sum_{r}^{vir.} c_i^r|\Psi_i^r\rangle + \sum_{i<j}^{occ.}\sum_{r<s}^{vir.} c_{ij}^{rs}|\Psi_{ij}^{rs}\rangle + \sum_{i<j<k}^{occ.}\sum_{r<s<t}^{vir.} c_{ijk}^{rst}|\Psi_{ijk}^{rst}\rangle + \cdots \tag{2.55}$$

where the i, j, k,... indices represent occupied orbitals that electrons are being excited from and r, s, t,... indices represent unoccupied virtual orbitals that electrons are being excited to. When this FCI wave function is truncated at the second term, it becomes the MRCISD wave function. This can be seen by looking at the FCI matrix given below in Figure 2.3.

Though MRCISD scales relatively steeply at N^6, it is highly accurate and gives excellent results for systems small enough to be computationally feasible. Therefore, it works well as a

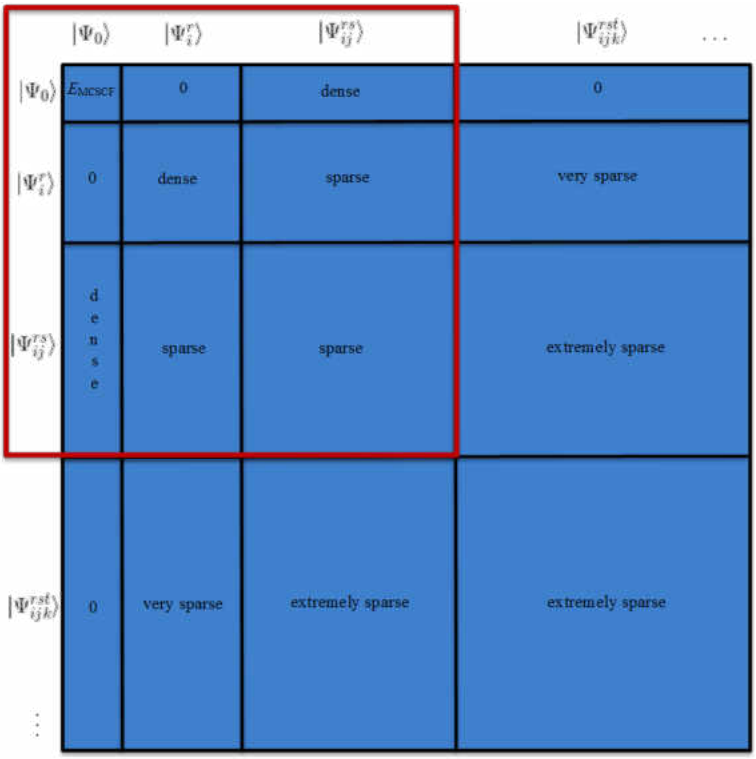


Figure 2.3. Schematic representation of the FCI matrix. Here it is explicitly showing up through triple excitations, but in general, it continues to the right and down till all possible excitations are accounted for. The location that the FCI matrix is truncated to get MRCISD is shown by the red box. Drawing is not to scale.

method to compare and validate results with other methods on smaller test case calculations, which then allows the extrapolation from the test calculations onto larger systems with the method that was validated. This is a particularly useful strategy when the larger systems are too large for MRCISD to handle, yet the validated method can handle them. This is the way in which MRCISD is used in the work presented in this dissertation. It will be seen in Chapters 3 and 4 when it is used to validate GVVPT2 results as needed.

2.7 Choice of Active Space

One area that is particularly important to the work presented in this dissertation is the development of active spaces. The general challenge when developing an active space for a given system is that the balance must be found in the overall size. It must be large enough to capture all the necessary chemistry, yet it must be small enough to retain computational feasibility. This can be quite challenging for any given case, e.g., a single point energy calculation, but when dealing with isomerization or reaction coordinates, it gets considerably more complex. In the work presented in this dissertation this was a huge challenge and it was only accomplished through multiple iterations.

In the arsenic oxide project presented in Chapter 3, the active spaces had the extra requirement to fit the above described criteria while also needing to work equally well for all the isomers considered along relatively complex reaction pathways. Many different active spaces were developed using various strategies which will be described below. There were also significant challenges in the infant stages of the $Cu_2O_2^{2+}$ project, presented in Chapter 6, which was so problematic that we ended up having to bench the study for several years until recently. This was due to a varying degree of biradical character and a rapidly changing degree of dynamical and static electron correlation effects along the isomerization coordinate. Therefore, the chosen active space needs to adequately describe very different chemical environments while still remaining small enough to be computationally practical.

Finally, another type of challenge arose in the work presented in this dissertation. When dealing with a large molecule that needs a high-level, multireference method, the difficult choice

of an active space becomes considerably more challenging. This is due to the fact that only a small subset of orbitals that appears necessary to adequately describe the chemistry can be realized because of severe system size constraints. This happened in the arsenic oxide project and may need to be addressed in the extensions of the antimony project as well (Chapter 4). Originally a sophisticated high-level single-reference method, CR-CC(2,3) was thought to provide calibration for our GVVPT2 results. However, the arsenic oxides proved to have a high degree of multi-center bonding character, necessitating the use of a more flexible high-level multireference method like MRCISD to validate the GVVPT2 results against. The difficulty arose in trying to reduce the active space from a single and double (SD) excited (28e,20o) active space (where the common notation of "e" standing for electrons and "o" for orbitals is used) down to a much smaller SD excited (6e,6o) active space, while still accounting for all the important chemistry.

Though a lot of active space development is through trial and error, there are several techniques that can aid in the development of good active spaces. One such technique is just the straight forward use of HF+RMP2, in which the natural orbitals printed out in the RMP2 output file will give the computed approximate occupancies of the orbitals. When using this technique, it is useful to make a balanced active space, choosing a similar number of mostly occupied and mostly unoccupied orbitals. Many times there will be natural divides that can be seen in orbital occupation rather than a smooth increase or decrease in occupancy of the orbitals. These places have those sharp changes to occupancies for chemical reasons, so they many times will give a good active space. This point leads naturally into another technique that takes more time to develop a sense for.

One of the most important things to consider when constructing an active space is the chemistry that is attempting to be described. It is here that years of well-developed chemical

intuition can be highly advantageous. Fortunately, such intuition can be developed, it just takes time and practice of working with many difference chemical species. This technique is what guided us to our current successful active space for the arsenic oxide series.

Another important protocol that is available to users of UNDMOL is a highly customizable orbital partitioning scheme through the use of macroconfigurations. Sometimes there will be several groups of orbitals that should be considered rather than just previously mentioned "mostly occupied" and "mostly unoccupied" groups. Sometimes it is advantageous to break up orbitals further into groups based on symmetry or spatial arrangement. An example of this would be what we used at one time to break up the active space for $Cu_2O_2^{2+}$ where the orbitals were partitioned based on whether they were in-plane or out-of-plane from the plane that intersected each of the atoms in the molecule. Macroconfigurations can be used to specify how many electrons each orbital should have excited to and from, vide supra. This gives a lot of power in the hands of the user to make an active space that is too large for a CAS, or even SD, into a something far more manageable by only considering the excitations and occupancies that are most important.

Another technique that can be used is flexibility in the use of the active core orbitals for GVVPT2 and MRCISD. This allows the user to have a smaller active space at the MCSCF level by putting low-lying orbitals, which may partially contribute to the chemical activity of a species, into the frozen orbitals while still giving them a correlated treatment at the GVVPT2 or MRCISD levels. This can speed up the MCSCF calculation and even improve overall results in the right situation.

In the case that was mentioned earlier about significantly truncating an active space, e.g., in order to run MRCISD on a large molecule, the technique that was used in this dissertation involved a few steps. First, a MCSCF+GVVPT2 calculation should be run with the larger active

space. Then, the printed out configurations for the GVVPT2 calculation will give weights of each configuration. Typically if there is a weight that is more than ± 0.04 then it constitutes an important enough configuration that it should be in the active space. Once those are all identified, the variable occupancy orbitals should be located within the configurations. In order to develop a minimal active space, the orbitals that have the most hole/excited electron activity should be used. Again, as in the first technique described, it is important to strive for a balanced active space.

Another recently developed technique that has proven useful for transition metals was discovered by Patrick Tamukong, who showed that a small valence picture active space can still capture the important chemistry.[26,27] The traditional active spaces types (for MO based calculations) rely on a Hartree–Fock MO picture, which the group has demonstrated success in describing organic molecules as well as other main group systems. However, the active spaces that were found to work well for transition metals were based on valence bond theory (similar to a generalized valence bond (GVB) picture). Here, the chemical bonding results from overlap of atomic orbitals, which are many times hybridized prior to bonding in order to satisfy geometry constraints. These types of active spaces also assume that the most important contributors to bonding are the MOs derived from the valence shells of the participating atoms. Certainly, atomic sub-shells just above or below the valence may contribute, but such contributions are generally minimal. However, these MOs are then further partitioned into subspaces in which bonding and anti-bonding orbitals of the same kind are in the same subspace. The last step would then be to include those atomic sub-shell orbitals just below the valence MOs into the active core subgroup for the GVVPT2 and MRCISD calculations. This type of active space was what allowed us to continue work on the benched $Cu_2O_2^{2+}$ project, presented in Chapter 6. This GVB-derived active space was used to generate an initial guess for the important $Cu_2O_2^{2+}$ orbitals. Though this gave us a much better start than the

above mentioned HF-type of active space, the active space had to go under several refinement iterations through our most recent protocol for active space development.

This final technique is primarily a numerical approach rather than a type of chemical approach (which both the GVB- and HF-type techniques are) and was developed as a way to get good orbitals that allow the MCSCF calculations to converge at all points of the $Cu_2O_2^{2+}$ isomerization. As mentioned above, neither the HF-type or the initial GVB-derived of active space was adequate for describing the chemistry of the $Cu_2O_2^{2+}$ isomerization. We found this was because it did not fit with the main block-type of systems (for which the HF-type of active space works well) or with the transition metal-type of systems (for which the GVB-derived active space works well). Therefore, it proved quite challenging to get the MCSCF calculations to converge and whenever they did converge, they would not form a smooth energy curve for the entire isomerization. It should be noted that the GVB-derived active space worked significantly better than the HF-type. However, the MCSCF still had a lot of trouble converging due to the active space. In effort to determine what was going wrong with the active space, we analyzed the MCSCF quasi-canonical orbital energies, the one-particle density matrix, and both the electron configurations that contribute to the total wave function and their respective amplitudes. From this analysis we were able to refine the active space iteratively. In the end the active space was broken up into two subspaces, similar to the HF-type of active spaces mentioned above: (i) mostly occupied and (ii) mostly unoccupied. The general rules we enacted, if it was true across the whole isomerization, were as follows: (i) if the quasi-canonical orbital energy was below -1.0 Ha and it remained doubly occupied for all CSF amplitudes greater than 0.03 then the orbital would be moved to the frozen core, (ii) if the orbital energy was between -1.0 Ha and 0.0 Ha the orbital would be moved into the “mostly occupied” subspace, (iii) if the orbital energy was between 0.0 Ha and about 1.0 Ha and there was at least one configuration that

showed occupation with an amplitude of greater than 0.03 then the orbital would be moved into "mostly unoccupied" subspace, and finally (iv) if the orbital energy was above 1.0 Ha and there was no configuration that showed occupation with an amplitude of greater than 0.03 then the orbital would be moved out of the active space and it would be treated as a virtual orbital. There were a few minor exceptions to these rules that are system dependent that won't be discussed here, but one important exception for our study was to rule (iii) and rule (iv). This exception was because $Cu_2O_2^{2+}$ is a cation, so there will always be a few negative energies for some of the lowest lying virtual orbitals; however when they are put into the "mostly unoccupied" subspace, their energies grow past 1.0 Ha and show no significant occupied configurations as expected for virtual orbitals. Therefore, several iterations were used to test the lowest lying virtual orbitals by placing them into the "mostly unoccupied" subspace and then taking them out again if they acted as virtual orbitals. One last step that is only used for the GVVPT2 or MRCISD levels was to take any frozen orbitals that had a quasi-canonical orbital energy between -2.0 Ha and -1.0 Ha and put them into the active core subgroup. It was through the above described numerical active space development approach, along with allowing for triple and quadruple excitations between the active space groups, that we were able to get MCSCF to converge while also getting a smooth energy curve along the isomerization coordinate. This technique applied to the $Cu_2O_2^{2+}$ system is described in more detail below in Chapter 6.

3 THEORETICAL INVESTIGATIONS OF ARSENIC OXIDES RELEVANT TO COAL COMBUSTION: A GVVPT2 MULTIREFERENCE PERTURBATION THEORY STUDY

3.1 Introduction

The production of electricity through the combustion of coal has been widespread historically and is still being used more extensively than any other fuel to generate electricity.[46] This is due to coal's availability, relatively high energy density and low costs. However, the emitted flue gases are known to contain potentially hazardous trace elements.[15,47] Since these trace elements

can be harmful to the environment and human health, as well as have adverse effects on the production of power, there has been an increasing interest in their remediation. This requires a more detailed understanding of their kinetic reactivities and thermodynamic stabilities. This is especially true for their respective oxides present in the flue gases. A deeper understanding of their chemical forms, oxidation states and structures present in the flue gases may aid in the development of novel strategies to combat these trace elements.[18–20] One of the particularly hazardous trace elements found in flue gases is arsenic.[15,16,47–49]

Arsenic's level of toxicity depends on its form, concentration and oxidation state. The oxidation states in which arsenic can be found are -3, +3 and +5, where arsenic in the +3 oxidation state in an inorganic compound is its most toxic form. Arsenic is a chronic poison and poses a significant risk to humans, animals and plants alike while it also negatively impacts the production of power. Specifically, it poisons the catalytic converter used in the selective catalytic reduction of NO_x, thereby increasing the cost of NO_x control.[50] It is clear that controlling the emission of arsenic in flue gases is of utmost importance.

In order to control the emission of arsenic, it is necessary to understand the mechanisms that lead to the formation of the various arsenic containing compounds as well as the relative energies of their corresponding isomers. Arsenic is likely to occur in flue gases during coal combustion in either its elemental form or in the form of oxides.[16,17] Though there have been a few theoretical studies on the monomeric oxides, AsO, AsO_2 and AsO_3,[51–54] several isomers of As_2O_3,[54,55] and on As_4O_6,[55–57] to the best of our knowledge a potential gas phase accretion mechanism from monoarsenic oxides has yet to be investigated.

Previously, Sun et al. used B3LYP, B3P86, B3PW91 and MP2 to predict the structural and bonding properties and the vibrational frequencies of AsO_2.[51] They also calculated the relevant

state–state energy separations using CCSD(T) and QCISD(T). They found that the geometrical parameters determined at several theoretical levels for AsO_2 did not deviate significantly from each other and that the lowest state of cyclic AsO_2 was $^2B_2$ with a $^2B_1$ state about 28 kcal/mol higher. More related to this work, several arsenic and selenium monomeric oxides were studied by Urban and Wilcox with a variety of methods.[52] They showed that the QCISD, QCISD(T), CCSD, and CCSD(T) methods with small core relativistic effective core potentials (RECP) for arsenic outperformed B3LYP/RECP. In a paper with a similar motivation to our work, Monahan-Pendergast et al. used several *ab initio* methods to study gaseous arsenic and selenium monomeric oxides released into the atmosphere due to coal-fired power plant emissions, and they found five thermodynamically favored reactions which were more likely to go to completion under tropospheric conditions than stratospheric conditions.[53] More recently, Rosli et al. used LDA to calculate the structures and vibrational frequencies of various monoarsenic and diarsenic oxide clusters relevant to arsenic oxide glass.[54] They found that AsO_2, AsO_3, As_2O_2 and AsO_4 may all be present in the glassy state due to their calculated vibrational frequencies comparing reasonably well with those found in the experimental Raman spectra. Even more relevant to our work, da Hora et al. used the chemometric multivariate technique, principal component analysis (PCA), to select B3PW91/6-311+G(3df) in the study of nine structures of As_2O_3.[55] They found three of them be stable isomers, and that the least stable isomer found and its enantiomer were able to form As_4O_6 spontaneously in the gas phase.

Though each of these studies gave valuable insights on various arsenic oxide species, each of them used single reference-based approaches exclusively. Therefore, the multireference character of the arsenic oxides has not yet been but should be assessed. In this work, a relatively recently developed multireference method, i.e., second-order Generalized Van Vleck Perturbation Theory

(GVVPT2),[13,14] will be used. The purpose of this work was to study theoretically both monoarsenic oxides (i.e., AsO, AsO_2, AsO_3) and diarsenic oxides (i.e., various isomers of As_2O_3 and As_2O_5); in particular, the equilibrium structures, heats of oxidation reactions, and relative stabilities of the possible isomers of As_2O_3 and As_2O_5 were investigated. The overall goal was to determine the degree of multireference character these arsenic oxides possess and whether they will accrete in the gas phase through a similar mechanism of the previously studied selenium oxides or not.[58]

3.2 Theoretical Methodology

For this study, the GVVPT2 method was used, which is described in detail in Chapter 2. GVVPT2 has proven itself successful in many applications where difficult multireference structures were studied. Some examples include disulfur monoxides (S_2O),[59] azabenzenes,[60] the modeling of mercury oxidation and binding on activated carbons in flue gas,[61] and, of particular relevance to this study, the study of dioxiranes.[62] GVVPT2 has also been shown to produce results in good agreement with CCSD[63] and CCSD(T)[64] when those methods are applicable. In our previous study of both monomeric and dimeric selenium oxides,[58] GVVPT2 was shown to produce results in good agreement with the linked completely renormalized coupled cluster method,[23,65–68] including single and double excitations and perturbative triples, i.e., CR-CCSD(T)$_L$, or CR-CC(2,3), as is recommended by its developer. It is worth noting that the CR-CC(2,3) method has been shown to be more accurate than the original CCSD(T) method.[24]

3.2.1 Computational Details

The computational protocol employed in this study was validated in our previous work with selenium oxides and can be broken up into the following six steps. First, the geometries were generated through the use of density functional theory (DFT). More specifically, the B3LYP hybrid generalized gradient approximation variant[69–72] with the 6-311G(d,p) (i.e., 6-311G*) basis set.[73,74] Second, those structures were confirmed to be either a stationary point or a saddle point through a vibrational analysis using the same level of theory. It should be pointed out that B3LYP/6-311G* has been shown to give geometries and vibrational frequencies with a reasonable compromise of accuracy and computational expense and was recently shown in our selenium oxide study to perform well with systems similar to this present work. Third, a restricted Hartree–Fock (RHF) and a restricted Møller–Plesset perturbation (RMP2) calculation using the cc-pVTZ basis set[75,76] was performed to obtain an initial approximation of the molecular orbitals (MOs) and determine an active space. Fourth, a multiconfigurational self-consistent field (MCSCF) calculation using the cc-pVTZ basis set was performed to further optimize the MOs. Fifth, a GVVPT2/cc-pVTZ single point energy calculation using the MCSCF orbitals was conducted. Sixth, the GVVPT2 results were then assessed against experimental data, when available as well as MRCISD/cc-pVTZ and CR-CC(2,3)/cc-pVTZ single point energy calculations. It should also be noted that both the structural and thermodynamic data were compared with the results obtained from the B3LYP/6-311G* calculations since they were used in the geometry optimizations of the monomers and were used to locate the various isomers and transition states of the dimers.

The B3LYP/6-311G* geometry optimizations and frequency calculations, as well as the

CR-CC(2,3) single point energy calculations, were performed using the GAMESS-US program.[77,78] All of the B3LYP calculations were of the RPA-derived Vosko-Wilke-Nusair correlation functional,[79] with 144 radial and 302 angular points for the Lebedev grid. The RHF+RMP2/cc-pVTZ, MCSCF/cc-pVTZ, MRCISD/cc-pVTZ, and GVVPT2/cc-pVTZ single point energy calculations were performed using our local electronic structure program, known as UNDMOL.[14] Version 1.2 of UNDMOL was particularly helpful in this study, because it has enhanced capabilities of resolving the angular momentum components which transform as the same irreducible representation of the Abelian point groups (e.g., σ and δ orbitals in AsO) used in most of the algorithms. Spectroscopic constants of AsO were obtained by 5–point numerical second- and third-differentiation of energies with step sizes of 0.003 Bohr using a program developed by Dr. Hoffmann called difreq.

3.2.2 Active Spaces

In this study, model spaces of the MCSCF incomplete variety were used for all GVVPT2 calculations. These active spaces were defined as all configuration state functions (CSFs) obtained from distributing all electrons in one orbital group (denoted here as G_1), all but 1 electron in G_1 (and one electron in a second orbital group, G_2), and all but 2 electrons in G_1 (and 2 electrons in G_2). The effectiveness of this type of active space for reactive oxygen-rich systems was previously demonstrated[80] and confirmed by comparing results obtained using these active spaces for GVVPT2 by comparing them to CASSCF.[58] Specifically for the GVVPT2 calculations, there were single and double excitations from all macroconfigurations used to describe the model space, and excitations were allowed from the higher-lying doubly occupied orbitals as well.

3.2.2.1 Monoarsenic Oxides

The active spaces used for the monoarsenic oxides consisted of the highest one-electron energy occupied molecular orbitals, G_1 (i.e., G_1 = {HOMO-(n – 1), ..., HOMO-1, HOMO}, where n is the number of occupied MOs in the active space) and the low-lying unoccupied molecular orbitals, G_2 (i.e., G_2 = {LUMO, LUMO+1, ..., LUMO+(n – 1)}, where n is the number of unccupied MOs in the active space). As with the monoselenium oxides, the active spaces were derived from a valence picture after preliminary studies.[58]

For AsO, the set of 8 valence molecular orbitals (i.e., 2s and 2p orbitals of O, and the 4s and 4p orbitals of As) having 11 electrons (11e, 8o), were divided into two subgroups: G_1 = {$9a_1$, $10a_1$, $4b_1$, $11a_1$, $4b_2$, $5b_1$}, and G2 = {$5b_2$, $12a_1$}. The AsO_2 molecule had 12 valence orbitals in its active space (i.e., 2s and 2p orbitals of O, and the 4s and 4p orbitals of As) with 17 electrons (17e, 12o) divided into two subgroups: G_1 = {$9a_1$, $5b_2$, $10a_1$, $11a_1$, $6b_2$, $4b_1$, $7b_2$, $2a_2$, $12a_1$} and G_2 = {$5b_1$, $13a_1$, $8b_2$}. And in the same way for AsO_3: G_1 = {$10a_1$, $11a_1$, $5b_2$, $12a_1$, $13a_1$, $4b_1$, $6b_2$, $5b_1$, $14a_1$, $7b_2$, $2a_2$, $8b_2$} and G_2 = {$6b_1$, $15a_1$, $16a_1$, $9b_2$}, with an active space of (23e, 16o).

3.2.2.2 Diarsenic Oxides

The active orbitals of the diarsenic oxides were divided into two subgroups as well. The energy calculations of the isomers and transition states of As_2O_3 used 20 active orbitals (i.e., the 2s and 2p of O, and the 4s and 4p of As) which contained 28 electrons (28e, 20o) that were distributed as G_1 = {13–18a_1 5a_2 6–9b_1 11–13b_2, in C_{2v} symmetry for isomer 1, 23–31a' 10–14a",

in C_s symmetry for isomer 2, 24–34a' 9–11a", in C_s symmetry for isomer 4, or simply 32–45a in C_1 symmetry} and G_2 = {19–20a_1 6a_2 10b_1 14–15b_2, in C_{2v} symmetry, 32–35a' 15–16a", in C_s symmetry, 35–39a' 12–13a", in C_s symmetry for isomer 4, or simply 46–51a in C_1 symmetry}. The active space used for MRCISD, which was used on isomer 1, was chosen based on the dominant CSFs of the GVVPT2 calculation. Single and double excitations were considered from G_1 to G_2 with G_1 = {5a_2 12–13b_1, in C_{2v} symmetry}, G_2 = {6a_2 14–15b_1, in C_{2v} symmetry}. The rest of the orbitals that were considered active for the MCSCF + GVVPT2 calculations presented in this work were put into the frozen core for the MCSCF and moved into the active core for MRCISD.

The single point energy calculations of the isomers and transition states of As_2O_5 used 28 active orbitals (i.e., the 2s and 2p of O, and the 4s and 4p of As) containing 40 electrons (40e, 28o) that were distributed as G_1 = {14–21a_1 5–6a_2 6–10b_1 12–16b_2, in C_{2v} symmetry for isomer 1, 18–25a_1 3–4a_2 8–12b_1 8–12b_2, in C_{2v} symmetry for isomer 3, 25–37a' 10–16a", in C_s symmetry, or simply 34–53a in C_1 symmetry} and G_2 = {22–24a_1 7a_2 11b_1 17–19b_2, in C_{2v} symmetry for isomer 1, 26–29a_1 13–14b_1 13–14b_2, in C_{2v} symmetry for isomer 3, 38–43a' 17–18a", in C_s symmetry or simply 54–61a in C_1 symmetry}.

3.3 Results and Discussion

The goals of this work were to analyze the arsenic oxides' multireference character through GVVPT2 and to determine the likelihood of a gas phase accretion mechanism from the monoarsenic oxides into diarsenic oxides through careful structural and thermodynamic analysis. These goals will be addressed, starting with the monomers and moving to the dimers.

3.3.1 Arsenic monoxide (AsO)

Since we already assessed the effect of basis set and model space choice in our previous study,[58] we need only assess the effect of inclusion of dynamic correlation here. This was accomplished through comparing spectroscopic constants obtained via B3LYP/6-311G*, GVVPT2/cc-pVTZ and CR-CC(2,3)/cc-pVTZ calculations on the ground state of AsO with available experimental data and is given below in Table 3.1.

Table 3.1. Comparison of thermodynamic constants for the ground state ($X^2\Pi$) of AsO.

Method	r_e (Å)	ω_e (cm^{-1})	$\omega_e x_e$ (cm^{-1})	α_e (cm^{-1})	B_e (cm^{-1})	D_e (kcal/mol)
B3LYP/6-311G*	1.631	979.27	-95.15	-0.0060	0.48025	48.7
GVVPT2(11e,8o)-B_2	1.666	979.85	-298.26	-0.0110	0.46092	50.0
GVVPT2(11e,8o)-B_1	1.642	–	–	–	–	50.7
CR-CC(2,3)	1.637	972.50	-52.40	-0.0058	0.47709	47.7
Exp.[a,b]	1.6236	967.08 (965.90)	–	–	0.48482 (0.48552)	–

[a]Ref. 81; [b]The values for the $^2\Pi_{1/2}$ and $^2\Pi_{3/2}$ states with the latter in parenthesis.

Although AsO possesses $C_{\infty v}$ symmetry, it was studied using the C_{2v} and C_1 point groups when using UNDMOL and with C_{4v} and C_1 symmetry in GAMESS calculations due to software constraints. The geometry of AsO was optimized with three methods to determine the variation of geometrical parameters in arsenic oxides with varying levels of theory. The r_e bond length of AsO was found to be 1.631 Å at the B3LYP level of theory, showing great agreement with experimental value of 1.6236 Å.[81] At the GVVPT2 level r_e for the B_2 state was found to be 1.666 Å and 1.642 for the B_1 state, showing good agreement with experiment for the B_1 state. At the CR-CC(2,3) level it was 1.637 Å, also showing good agreement with experiment. The value of

Table 3.2. Leading configurations and weights of MCSCF reference functions of arsenic oxides given as CSFs in the GVVPT2 output.

	Abelian (sub)group	amplitude	configuration
AsO	C_{2v}	0.875	$9a_1^2 10a_1^2 11a_1^2 4b_1^2 5b_1^1 4b_2^2 12a_1^0 5b_2^0$
		-0.284	$4b_2 \rightarrow 5b_2$
		-0.108	$4b_2^2 \rightarrow 5b_2^2$
		0.093	$11a_1 4b_2 \rightarrow 12a_1 5b_2$
		-0.091	$4b_1 4b_2 \rightarrow 5b_1 5b_2$
		-0.090	$4b_1 \rightarrow 5b_1$
AsO_2	C_{2v}	0.867	$9a_1^2 10a_1^2 11a_1^2 12a_1^1 2a_2^2 4b_1^2 5b_2^2 6b_2^2 7b_2^2 13a_1^0 5b_1^0 8b_2^0$
		-0.142	$4b_1 \rightarrow 5b_1$
		0.104	$2a_2 7b_2 \rightarrow 12a_1 5b_1$
		-0.099	$11a_1 \rightarrow 12a_1$
		0.092	$2a_2 6b_2 \rightarrow 12a_1 5b_1$
AsO_3	C_{2v}	0.877	$10a_1^2 11a_1^2 12a_1^2 13a_1^2 14a_1^2 2a_2^2 4b_1^2 5b_1^2 5b_2^2 6b_2^2 7b_2^2$ $8b_2^1 15a_1^0 16a_1^0 6b_1^0 9b_2^0$
		–	–

ω_e was found to be 979.27 cm^{-1}, 979.85 cm^{-1} and 972.50 cm^{-1} at the B3LYP, GVVPT2 and CR-CC(2,3) levels of theory respectively with the experimental value being 967.08 cm^{-1} for the $^2\Pi_{1/2}$ state and 965.90 cm^{-1} for the $^2\Pi_{3/2}$ state.[81] From this it can be seen that the geometrical parameters do not show a large dependence on the method used. This suggests that B3LYP/6-311G*, which can generate geometries much faster than GVVPT2 and CR-CC(2,3), can be used to generate reasonable geometries of the other arsenic oxide species.

Table 3.2 displays the dominant amplitudes of the monomeric arsenic oxides and the corresponding configurations of the MCSCF wave function near the equilibrium geometry for the purpose of displaying the multireference character of the monomeric species. As can be seen, while the wave function for AsO has a dominant configuration with an amplitude of 0.875, there are five additional significant (≥ 0.09) configurations, showing significant multireference behavior.

3.3.2 Arsenic dioxide (AsO_2)

AsO_2 was predicted to possess C_{2v} symmetry with the structural parameters given below in Table 3.3. The AsO_2 r_e bond length calculated at the B3LYP level of theory is longer than the r_e of AsO by 0.013 Å, the GVVPT2 r_e bond lengths were predicted to be identical, and the CR-CC(2,3) bond length for AsO_2 r_e was found to be significantly shorter by 0.056 Å.

Table 3.3. Comparison of structural parameters (Å and degrees) of AsO and AsO_2, optimized at the B3LYP/6-311G* and GVVPT2/cc-pVTZ levels with experimental data.

	Symmetry	Parameter	B3LYP	GVVPT2	CR-CC(2,3)	Exp.
AsO	$C_{\infty v}$	r_e(As–O)	1.631	1.642	1.637	1.6236[a]
AsO_2	C_{2v}	r_e(As–O)	1.644	1.642	1.581	–
		∠ O–As–O	126.7	129.8	129.8	–

[a]Ref. 81.

The formation of AsO_2 from the reaction of AsO and $\frac{1}{2}O_2$ was predicted to be energetically favorable at all levels of theory. The energies of reactions are -28.5 kcal/mol at the B3LYP level, -36.5 kcal/mol at the GVVPT2 level and -26.9 kcal/mol at the CR-CC(2,3) level of theory (see Table 3.5 below).

3.3.3 Arsenic trioxide (AsO_3)

AsO_3 was predicted to also possess C_{2v} symmetry with one As–O bond being longer than the other two. The energetically lowest doublet AsO_3 r_{e1} bond length calculated at the B3LYP level of theory was found to be longer than the r_e of AsO by 0.113 Å and longer than the r_e of AsO_2 by 0.100 Å. However, the doublet AsO_3 r_{e2} bond length calculated at the B3LYP level of theory

was found to be shorter than the r_e of AsO by 0.013 Å and shorter than the r_e of AsO_2 by 0.026 Å. The energetically lowest quartet AsO_3 r_e bond length calculated at the B3LYP level of theory was found to be longer than the r_e of AsO by 0.156 Å and longer than the r_e of AsO_2 by 0.143 Å.

Table 3.4. The geometrical parameters (Å and degrees) of the AsO_3 doublet and quartet, optimized at the B3LYP/6-311G* level.

	Symmetry	Multiplicity	Parameter	B3LYP
AsO_3	C_{2v}	doublet	r_e(As–O1)	1.744
			r_e(As–O2/O3)	1.618
			∠ O1–As–O2/O3	111.5
			∠ O2–As–O3	137.0
AsO_3	C_{3v}	quartet	r_e(As–O)	1.787
			∠ O–As–O	79.0

Reaction enthalpies are presented below in Table 3.5. The formation of the doublet state of AsO_3 from the reaction of AsO and O_2 was predicted to be energetically favorable at all levels of theory. The energies of reactions are -25.6 kcal/mol at the B3LYP level, -34.6 kcal/mol at the GVVPT2 level and -29.3 kcal/mol at the CR-CC(2,3) level of theory. In contrast, the formation of doublet state of AsO_3 from the reaction of AsO_2 and $\frac{1}{2}O_2$ was predicted to be slightly energetically unfavorable at both the B3LYP and GVVPT2 levels of theory and slightly energetically favorable at the CR-CC(2,3) level of theory.

Table 3.5. Reaction enthalpies (kcal/mol) for the oxidation of monomeric arsenic oxides.

Reaction	Product Multiplicity	B3LYP 6-311G*	GVVPT2 cc-pVTZ	CR-CC(2,3) cc-pVTZ
$As + \frac{1}{2}O_2 \rightarrow AsO$	doublet	-48.7	-50.0	-47.7
$AsO + \frac{1}{2}O_2 \rightarrow AsO_2$	doublet	-28.5	-36.5	-26.9
$AsO + O_2 \rightarrow AsO_3$	doublet	-25.6	-34.6	-29.3
$AsO_2 + \frac{1}{2}O_2 \rightarrow AsO_3$	doublet	2.9	1.9	-2.4
$AsO + O_2 \rightarrow AsO_3$	quartet	24.0	–	–
$AsO_2 + \frac{1}{2}O_2 \rightarrow AsO_3$	quartet	52.4	–	–

The formation of the quartet state of AsO_3 from the reaction of AsO and O_2 and the formation of the quartet state of AsO_3 from the reaction of AsO_2 and $\frac{1}{2}O_2$ were both also predicted to be considerably more energetically unfavorable than the reactions forming the doublet state. GVVPT2 shows that the order of stability for the mono-arsenic oxides is as follows

(least stable) AsO_3 (quartet) < AsO < AsO_3 (doublet) AsO_2 (most stable)

3.3.4 Diarsenic trioxide (As_2O_3)

A total of four minima (i.e., isomers 1, 2, 3 and 4) and three transition states (i.e., TS12, TS23 and TS34) were found for the diarsenic trioxide (As_2O_3) molecule. The most stable isomer is given in Figure 3.1.

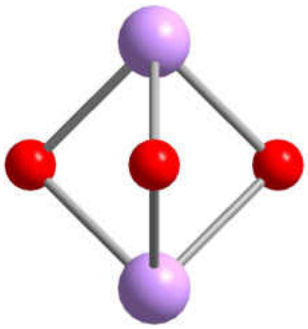


Figure 3.1. Structure of isomer 1 of diarsenic trioxide (As_2O_3).

Isomer 1 was predicted to be a triple oxygen bridging structure that possesses D_{3h} symmetry. Isomer 2 was predicted to be a ring structure with an oxygen jutting out of the plane of the ring and it possesses C_s symmetry. Isomer 4 was found to be a planar structure has C_s symmetry. Isomer 3 was found to possess C_2 symmetry. Each of the transition states were predicted to be of the C_1 point group. B3LYP optimized geometrical parameters are given below in Table 3.6.

The optimized bond lengths fall into two general categories: short bond length and long bond length. The only exception to this was isomer 2, which had two As–O bonds in-between the short and long bond length observed in the As_2O_3 isomers at 1.700 Å. When comparing the geometrical parameters for As_2O_3 to that of the monomers, it was found that the long As–O bonds in the equilibrium structures of As_2O_3 (i.e., 1.819–2.146 Å) were significantly longer than the r_e(As–O) bond lengths in AsO (i.e., 1.631 Å), AsO_2 (i.e., 1.644 Å) and AsO_3 (i.e., 1.744 Å and

Table 3.6. Optimized B3LYP/6-311G* geometrical parameters (Å and degrees) of the diarsenic trioxide (As_2O_3) isomers.

Parameters	Isomer 1 (D_{3h})	Isomer 2 (C_s)	Isomer 3 (C_2)	Isomer 4 (C_s)	TS12 (C_1)	TS23 (C_1)	TS34 (C_1)
$R_{As1\text{-}O3}$	1.865	1.700	1.819	1.845	1.738	1.765	1.840
$R_{As1\text{-}O4}$	1.865	1.700	1.620	1.617	1.738	1.645	1.617
$R_{As2\text{-}O3}$	1.865	2.146	1.819	1.803	2.035	1.903	1.798
$R_{As2\text{-}O4}$	1.865	2.146	–	–	2.034	2.791	–
$R_{As2\text{-}O5}$	1.865	1.612	1.620	1.625	1.671	1.614	1.625
∠ O3–As2–O5	82.0	103.2	106.9	106.3	89.3	103.9	106.0
∠ As1–O3–As2	81.5	96.4	134.7	129.6	86.0	112.3	131.4
∠ O3–As1–O4	82.0	95.4	106.9	103.4	92.1	100.5	104.8

1.618 Å). Notably, each of the short As–O bonds of the As_2O_3 isomers (i.e., 1.612–1.625 Å) are shorter than the r_e(As–O) bond lengths in AsO, AsO_2 and AsO_3.

The frequencies of the As_2O_3 isomers and their transition states were obtained at the B3LYP/6-311G* level and show similar characteristics to those of the monomeric arsenic oxides. They are given below in Table 3.7. Each of the TS's were confirmed as such by each exhibiting an imaginary frequency. Similarly, each of the isomers were confirmed as such by having only real frequencies, except for the high lying isomer 3, which may be due to the geometry optimization mistakenly converging onto a saddle point.

Table 3.7. Frequency of diarsenic trioxide (As_2O_3) isomers and transition states in cm^{-1} calculated at the level of B3LYP/6-311G*.

Normal Mode	Isomer 1 (D_{3h})	Isomer 2 (C_s)	Isomer 3 (C_2)	Isomer 4 (C_s)	TS12 (C_1)	TS23 (C_1)	TS34 (C_1)
1	315.84 (E")	130.70 (A')	36.13*i* (B)	38.97 (A")	229.36*i*	101.81*i*	43.86*i*
2	322.15 (E")	150.28 (A")	32.66 (A)	94.42 (A')	173.12	100.11	52.52
3	406.20 (E')	204.80 (A")	57.14 (A)	133.30 (A")	328.90	134.56	108.21
4	410.36 (E')	240.01 (A')	265.88 (B)	225.52 (A')	333.84	267.88	278.35
5	421.62 (A_1')	330.92 (A')	324.59 (A)	314.33 (A')	346.25	324.78	302.02
6	588.03 (A_2")	471.67 (A')	430.38 (A)	497.77 (A')	447.54	507.22	464.14
7	643.96 (E')	780.26 (A")	694.10 (B)	695.68 (A')	693.68	675.67	700.83
8	647.44 (E')	842.76 (A')	989.00 (B)	982.48 (A')	762.14	927.97	980.39
9	739.86 (A_1')	1013.25 (A')	997.60 (A)	1001.83 (A')	859.93	1009.99	1002.50

The relative energies found for each level of theory and with each basis set used is given in Table 3.8. Also, for a visual aid, the relative energies are plotted for the cc-pVTZ basis set and are given in Figure 3.2. In the comparison between the single point energy calculations using GVVPT2/cc-pVTZ and CR-CC(2,3)/cc-pVTZ, both methods found the same relative ordering. B3LYP/6-311G* mistakenly predicted the relative energies of the each of the isomers to be nearly isoenergetic. It predicted isomer 2 to be the most stable isomer with an energy of -0.68 kcal/mol relative to isomer 1 (0.00 kcal/mol). It then predicted isomer 3 (0.06 kcal/mol) to be the next most stable followed by isomer 4 (0.61 kcal/mol) as the least stable.

Table 3.8. Relative energies (kcal/mol) at the B3LYP, CR-CC(2,3) and GVVPT2 levels of theory of the diarsenic trioxide (As_2O_3) isomers and transition states.

	Symmetry	B3LYP/6-311G*	CR-CC(2,3)/cc-pVTZ	GVVPT2/cc-pVTZ
Isomer 1	C_{2v}	0.00	0.00	0.00
TS12	C_1	12.50	17.39	16.14
Isomer 2	C_s	-0.70	8.10	4.39
TS23	C_1	1.07	8.57	20.53
Isomer 3	C_2	0.05	8.41	20.25
TS34	C_1	0.63	8.85	20.92
Isomer 4	C_s	0.60	9.04	21.23

Although the relative ordering of GVVPT2/cc-pVTZ agrees with CR-CC(2,3)/cc-pVTZ, a significant deviation was found between the two theories for both TS23 and TS34. TS12 has the closest agreement between B3LYP/6-311G* and GVVPT2/cc-pVTZ, with relatively small differences in energy. However, the other transition states show significant variation between B3LYP/6-311G* and GVVPT2/cc-pVTZ.

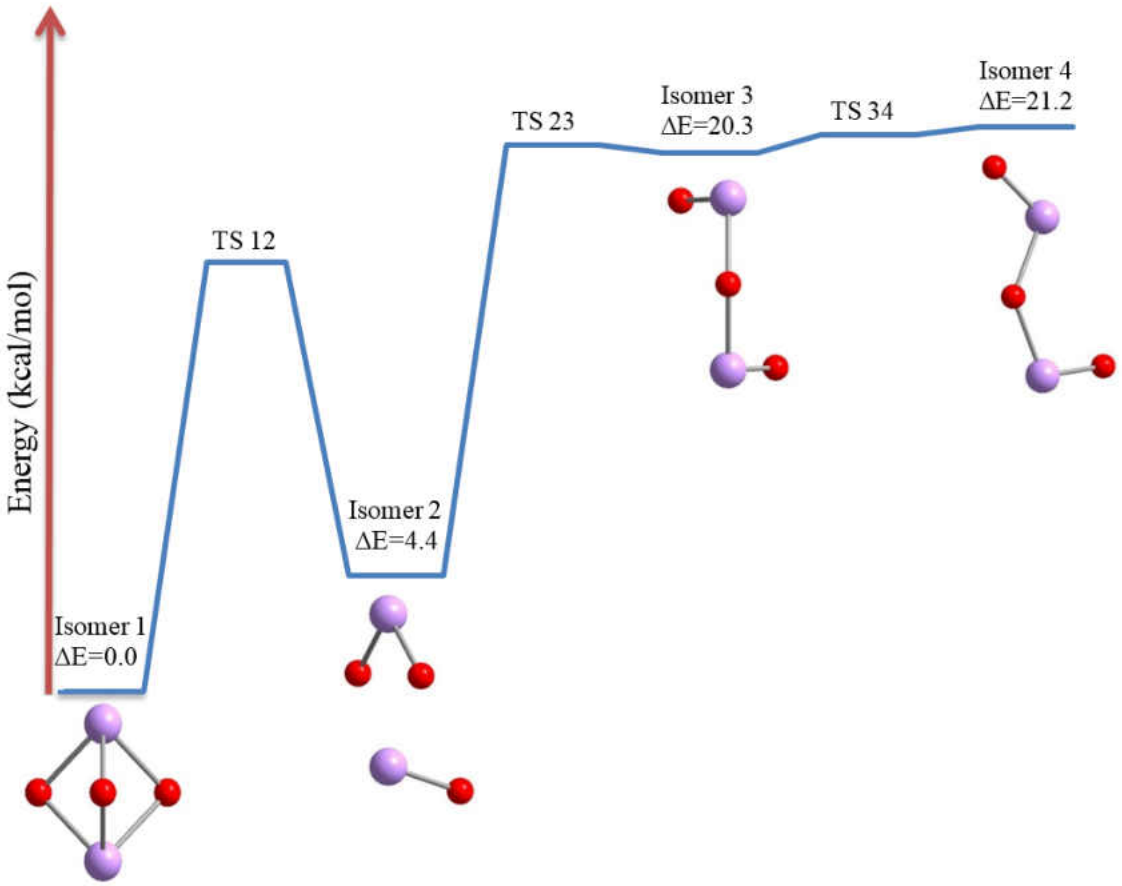


Figure 3.2. Relative GVVPT2/cc-pVTZ energies of diarsenic trioxide (As_2O_3) isomers and TSs (in kcal/mol).

In analogy with the reaction energy predictions made for the monomers (cf. Table 3.5), the stability of the monomeric oxides to formation of As_2O_3 was studied; i.e.,

$$AsO + AsO_2 \rightarrow As_2O_3 \tag{3.1}$$

The reaction energies as calculated by B3LYP, GVVPT2, CR-CC(2,3), and MRCISD are given below in Table 3.9. All predicted reaction energies were found to be highly energetically favorable. Both B3LYP and CR-CC(2,3) were found to agree reasonably well, but GVVPT2 predicts that the reaction is slightly less favorable (15.7% less than CR-CC(2,3)). This is likely due to the high degree of multireference character that was found in these species, and GVVPT2 does an excellent job in such cases. MRCISD calculations were not performed because of computational expense. These highly favorable reaction energies support the possibility of a gas phase accretion mechanism.

Table 3.9. Total energies (a.u.) of AsO, AsO_2 and As_2O_3 (isomer 1), and predicted reaction energies (kcal/mol).

	AsO	AsO_2	As_2O_3 (**1**)	Δ E (kcal/mol)
B3LYP/6-311G*	-2311.061 689	-2386.287 494	-4697.486 593	-86.2
GVVPT2/cc pVTZ//B3LYP	-2309.447 848	-2384.565 117	-4694.131 540	-74.4
CR-CC(2,3)/cc pVTZ//B3LYP	-2309.466 793	-2384.574 673	-4694.182 209	-88.3
MRCISD/cc pVTZ//B3LYP	–	–	-4693.996 429	–

3.3.5 Diarsenic pentoxide (As_2O_5)

Calculations on the diarsenic pentoxide (As_2O_5) molecules produced a total of seven minima (i.e., isomers 1, 2, 3, 4, 5, 6 and 7) and five transition states (i.e., TS13, TS24, TS36, TS67 and TS53). The most stable isomer is given in Figure 3.3.

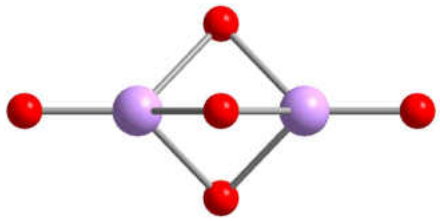


Figure 3.3. Structure of isomer 1 of diarsenic pentoxide (As_2O_5).

Isomer 1 was predicted to possess D_{3h} symmetry. Isomer 2 was found to possess C_{2v} symmetry. Isomers 3, 5 and 6 all were predicted to have C_s symmetry. Isomers 4 and 8 were found to be of the C_1 point group. Each of the transition states were predicted to have C_s symmetry except for TS 34, which was found to belong to the C_1 point group. Isomer 1 and isomer 6 were predicted to be cage-like structures while isomer 3 and TS23 were found to be ring structures. Isomer 1 has a cage-like structure, similar to the As_2O_3 isomer 1 cage, only with two additional oxygens jutting out from the two arsenics away from the cage. Isomer 6 was predicted to be a cage-like structure similar to isomer 1 of As_2O_3, yet with the two additional oxygens jutting out from one of the two arsenics away from the cage. B3LYP optimized geometrical parameters for the isomers and transition states are given below in Table 3.10.

Table 3.10. Optimized B3LYP/6-311G* geometrical parameters (Å and degrees) of the diarsenic pentoxide (As_2O_5) isomers and transition states.

Parameters	Isomer 1 (D_{3h})	Isomer 2 (C_{2v})	Isomer 3 (C_s)	Isomer 4 (C_1)	Isomer 5 (C_s)	Isomer 6 (C_s)	Isomer 7 (C_2)	Isomer 8 (C_1)	TS12 (C_s)	TS23 (C_s)	TS34 (C_1)	TS56 (C_s)
$R_{As1\text{-}O3}$	1.604	2.071	1.875	1.629	1.765	1.759	1.611	1.871	1.605	1.685	1.659	1.757
$R_{As1\text{-}O4}$	–	2.071	2.726	1.795	–	1.915	–	–	1.669	–	1.763	2.383
$R_{As1\text{-}O5}$	1.852	–	1.844	1.612	1.677	1.806	1.790	–	1.990	1.922	1.611	1.724
$R_{As1\text{-}O6}$	1.852	1.611	1.844	2.675	1.677	1.806	–	1.767	1.990	1.922	2.694	1.724
$R_{As1\text{-}O7}$	1.852	1.611	–	–	1.764	1.775	1.614	1.830	–	1.613	–	1.766
$R_{As2\text{-}O4}$	1.604	1.693	1.835	2.807	1.604	1.800	1.611	1.615	2.477	1.668	3.045	1.667
$R_{As2\text{-}O5}$	1.852	1.604	1.816	–	2.211	1.913	1.790	1.612	1.735	1.773	–	2.057
$R_{As2\text{-}O6}$	–	–	–	1.906	–	–	1.614	1.841	1.735	1.773	2.034	2.057
$R_{As2\text{-}O7}$	1.852	–	1.605	1.611	–	–	–	–	1.602	–	1.610	–
$\angle_{(O3)(As1)(O4)}$	–	75.3	30.5	109.7	–	102.7	–	–	140.6	–	106.9	98.4
$\angle_{(O5)(As1)(O7)}$	83.6	–	–	–	123.9	113.7	111.6	–	–	117.3	–	120.2
$\angle_{(O6)(As1)(O4)}$	129.7	105.8	68.5	30.4	–	82.4	–	–	90.9	–	29.1	76.2
$\angle_{(O5)(As2)(O7)}$	83.6	–	126.7	–	–	–	–	–	132.6	–	–	–
$\angle_{(O6)(As2)(O4)}$	129.7	–	94.2	28.4	103.2	82.5	138.4	109.4	74.5	106.5	23.7	87.4
$\angle_{(O6)(As2)(O7)}$	83.6	–	126.7	99.4	–	–	–	–	132.6	–	100.6	–
$\angle_{(As1)(O5)(As2)}$	79.4	–	89.2	–	93.1	80.4	125.8	–	84.0	89.0	–	83.4
$\angle_{(As1)(O4)(As2)}$	–	94.0	–	93.4	–	–	–	–	–	–	80.5	74.9

The frequencies of the As_2O_5 isomers and their transition states were calculated at the B3LYP/6-311G* level. The frequencies of the isomers are given below in Table 3.11 and the frequencies of the transition states are given in Table 3.12.

Table 3.11. Frequencies of diarsenic pentoxide (As_2O_5) isomers obtained at the B3LYP/6-311G* level of theory.

Normal Mode	1 (D_{3h})	2 (C_{2v})	3 (C_s)	4 (C_1)	5 (C_s)	6 (C_s)	7 (C_s)	8 (C_1)
1	140.10 (E")	64.62 (B_2)	198.11 (A")	69.94	87.23 (A')	59.64 (A")	56.47 (A)	78.42
2	174.12 (E')	125.50 (B_1)	198.22 (A')	91.29	107.72 (A")	174.77 (A')	70.68 (B)	97.56
3	176.11 (E')	191.42 (A_2)	227.18 (A")	144.27	158.86 (A")	182.16 (A")	93.21 (A)	124.26
4	179.25 (E')	192.08 (B_2)	318.48 (A")	169.04	178.69 (A')	296.89 (A')	214.64 (B)	210.70
5	350.90 (E")	223.59 (A_1)	335.37 (A')	204.57	204.12 (A")	362.11 (A")	222.66 (A)	244.30
6	351.19 (E")	245.21 (B_1)	383.37 (A')	240.06	259.30 (A')	385.28 (A')	231.87 (A)	269.23
7	394.76 (A_1')	321.74 (B_2)	403.94 (A')	246.90	316.03 (A')	446.83 (A")	291.74 (B)	331.11
8	444.71 (E')	325.96 (A_1)	450.27 (A")	315.44	335.17 (A")	453.16 (A')	294.83 (B)	342.44
9	446.07 (E')	365.52 (B_1)	549.12 (A')	399.68	419.08 (A')	520.63 (A')	335.35 (A)	530.08
10	538.01 (A_2")	501.49 (A_1)	568.14 (A')	564.22	578.74 (A')	590.29 (A')	517.86 (A)	543.65
11	690.72 (E')	819.60 (B_1)	644.27 (A')	633.04	625.98 (A')	627.27 (A')	753.15 (B)	578.15
12	694.25 (E')	822.39 (A_1)	647.68 (A")	861.15	811.54 (A')	693.94 (A")	919.93 (B)	699.95
13	748.93 (A_1')	921.03 (A_1)	728.95 (A')	906.15	844.87 (A")	699.28 (A')	931.28 (A)	862.34
14	1005.24 (A_2')	1026.85 (A_1)	843.45 (A')	1018.09	901.12 (A')	760.15 (A')	1046.28 (B)	923.72
15	1025.51 (A_1')	1045.51 (B_2)	1014.46 (A')	1025.86	1036.59 (A')	862.85 (A')	1049.73 (A)	1044.00

Table 3.12. Frequencies of diarsenic pentoxide (As_2O_5) transistion states obtained at the B3LYP/6-311G* level of theory.

Normal Mode	TS12 (Cs)	TS23 (Cs)	TS34 (C1)	TS56 (Cs)
1	161.73*i* (A')	315.71*i* (A')	108.74*i*	182.74*i* (A')
2	116.79 (A")	162.09 (A")	101.21	62.79 (A")
3	162.35 (A')	201.08 (A")	122.54	130.92 (A')
4	196.63 (A")	202.28 (A')	165.43	139.80 (A")
5	244.95 (A')	234.02 (A')	186.46	302.45 (A")
6	320.96 (A")	285.11 (A")	204.09	328.04 (A')
7	386.17 (A")	295.18 (A')	261.47	375.36 (A")
8	400.64 (A')	345.18 (A')	322.60	392.13 (A')
9	414.26 (A')	368.77 (A")	377.81	432.26 (A')
10	491.90 (A')	457.52 (A')	427.40	580.03 (A')
11	746.28 (A")	656.20 (A")	666.15	623.34 (A')
12	753.96 (A')	710.39 (A')	812.87	738.91 (A")
13	847.80 (A')	774.65 (A')	868.30	759.64 (A')
14	1018.06 (A')	815.07 (A')	1004.79	877.57 (A')
15	1031.42 (A')	976.88 (A')	1018.13	886.95 (A')

The similarities in the structural data of the monomer and of several isomers of As_2O_3 to that of several isomers of As_2O_5 suggest that several characteristics of their bonding motifs are retained. As observed for As_2O_3, the optimized bond lengths again fall into the two general categories of short and long bond lengths. Though with As_2O_5, the line between the two was somewhat blurred. As for As_2O_3, the similarity of bond lengths and angles in the found isomers of As_2O_5 raises the likelihood of isomerization.

The relative energies found for each level of theory and with each basis set used are given below in Table 3.13. Also, for a visual aid, the relative energies are plotted for the cc-pVTZ basis set and are given in Figure 3.4. It can be seen that isomer 1 is the minimum for GVVPT2 and for CR-CC(2,3), but isomer 3 is the minimum for B3LYP/6-311G*.

Table 3.13. Relative energies (kcal/mol) at the B3LYP, CR-CC(2,3) and GVVPT2 levels of theory of the diarsenic pentoxide (As_2O_5) isomers and transition states.

	Symmetry	B3LYP/6-311G*	CR-CC(2,3)/cc-pVTZ	GVVPT2/cc-pVTZ
Isomer 1	D_{3h}	0.00	0.00	0.00
TS12	(C_s)	7.70	14.38	15.58
Isomer 2	(C_{2v})	-3.34	6.51	1.88
TS23	(C_s)	39.73	50.92	47.73
Isomer 3	(C_s)	-7.23	4.80	13.09
TS34	(C_1)	25.60	52.00	72.00
Isomer 4	(C_1)	23.45	47.31	60.13
Isomer 5	(C_s)	24.61	37.65	44.17
TS56	(C_s)	36.23	44.39	52.10
Isomer 6	(C_s)	31.02	34.53	42.80
Isomer 7	(C_2)	-1.51	10.11	7.94
Isomer 8	(C_1)	17.65	33.78	49.87

GVVPT2 predicted that isomer 1, with a cage-like structure similar to the most stable isomer of As_2O_3, (i.e., isomer 1) was found to be the most stable isomer. Isomer 2 was found to be the next most stable with an energy of 1.88 kcal/mol above isomer 1, followed by isomer 7 at 7.94 kcal/mol, isomer 3 at 13.09 kcal/mol, isomer 6 at 42.80 kcal/mol, isomer 5 at 44.17 kcal/mol, isomer 8 at 49.87 kcal/mol, and isomer 4 at 60.13 kcal/mol. For CR-CC(2,3), isomer 1 was predicted to be lowest in energy and the same relative energy ordering for the rest of the isomers. In the B3LYP/6-311G* relative energy calculations, isomer 3 was mistakenly predicted as the most stable isomer at an energy of -7.23 kcal/mol with respect to isomer 1.

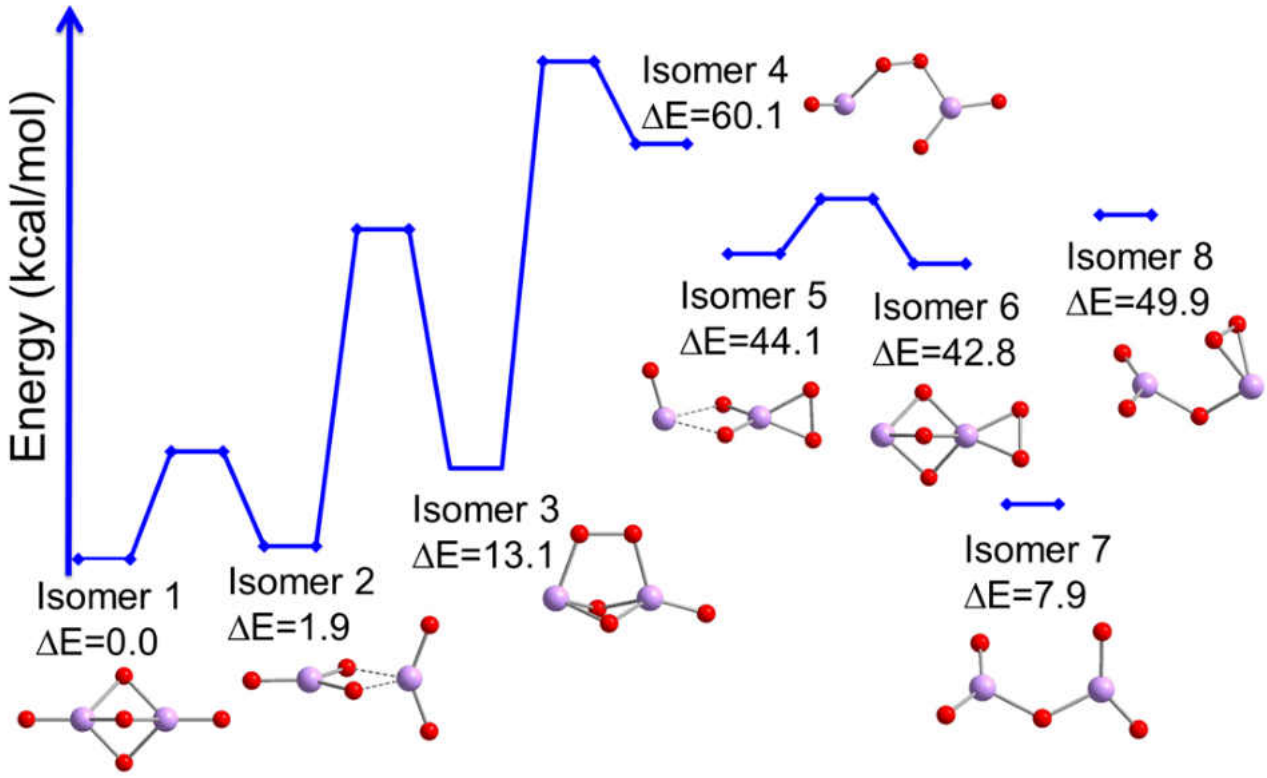


Figure 3.4. Relative GVVPT2/cc-pVTZ energies of diarsenic pentoxide (As_2O_5) isomers and TSs (in kcal/mol).

The energics of two reactions were studied for diarsenic pentoxide, given as

$$AsO_2 + AsO_3 \rightarrow As_2O_5 \tag{3.2}$$

and

$$As_2O_3 + O_2 \rightarrow As_2O_5 \tag{3.3}$$

The reaction energies as calculated by B3LYP, GVVPT2 and CR-CC(2,3) are given in Table 3.14. At each level of theory, reaction 3.2 was found to be highly energetically favorable with an excellent agreement between GVVPT2 and CR-CC(2,3). However, B3LYP predicted the reaction to be about 20 kcal/mol less favorable. Reaction 3.3 was predicted to be far less favorable. B3LYP again predicted the reaction to be a little more than 20 kcal/mol less favorable than CR-CC(2,3). GVVPT2 and CR-CC(2,3) did not agree well in this reaction, with GVVPT2 predicting reaction 3 to be 18 kcal/mol more favorable. This is likely due to the high degree of multireference character of the arsenic oxides in this reaction, and CR-CC(2,3) is not designed to handle molecules that are exceedingly multireference. We are in the process of testing all of the species included in our thermodynamics calculations to be sure of this. MRCISD calculations will either corroborate or supplant CR-CC(2,3). Again, it should be noted that these favorable reaction energies support the possibility of a gas phase accretion mechanism.

Table 3.14. Total energies (a.u.) of O_2, AsO_3 and As_2O_5 (isomer 1), and predicted reaction energies (kcal/mol).

	O_2	AsO_3	As_2O_5 (**1**)	Δ E (AsO_2 + AsO_3 $\rightarrow$$As_2O_5$ (**1**))	Δ E (As_2O_5 (**1**)+O_2 $\rightarrow$$As_2O_5$ (**1**))
B3LYP/6-311G*	-150.360 823	-2461.463 296	-4847.848 039	-61.0	-0.4
GVVPT2/cc pVTZ //B3LYP	-150.118 225	-2459.621 141	-4844.314 106	-80.2	-40.4
CR-CC(2,3)/cc pVTZ //B3LYP	-150.130 032	-2459.643 514	-4844.347 898	-81.4	-22.4
MRCISD/cc-pVTZ //B3LYP	–	–	–	–	–

3.4 Summary

On the basis of good agreement between B3LYP/6-311G*, CR-CC(2,3)/cc-pVTZ, and GVVPT2/cc-pVTZ optimized geometries for similar systems[58] as well as for AsO, B3LYP/6-311G* geometry optimizations were performed for AsO, AsO_3, As_2O_3 and As_2O_5. Four isomers were obtained for As_2O_3 and eight isomers were obtained for As_2O_5, along with a number of relevant transition states for each species. Single point energy calculations of the obtained geometries were then conducted at the GVVPT2/cc-pVTZ, CR-CC(2,3)/cc-pVTZ, and MRCISD/cc-pVTZ levels. Without exception, GVVPT2, CR-CC(2,3), and when used, MRCISD predicted identical energy orderings, with comparable relative energies were the data was available, especially so for the low-lying isomers. In contrast, B3LYP failed to produce semi-quantitatively useful energy results for As_2O_3 or As_2O_5.

From the reaction thermodynamics calculations, it was found that a gas phase accretion would be energetically favorable. Whether starting from As and O_2 or from the monoarsenic oxides, the accretion into As_2O_3 and As_2O_5 was found to be favored at all levels of theory (i.e., B3LYP/6-311G*, GVVPT2/cc-pVTZ, and CR-CC(2,3)/cc-pVTZ). All of the predicted reaction energies for the formation for the dimers were found to be highly favorable. The formation of As_2O_3 from AsO and AsO_2 is qualitatively energy neutral, with both B3LYP and CR-CC(2,3) showing slight exothermicity. GVVPT2 does, however, predict that the reaction is slightly less energetically favorable. In the absence of experimental numbers or MRCISD calculations, one would be tempted to rely on the GVVPT2 results. The formation of As_2O_5, from AsO_2 and AsO_3, was found to be energetically favorable with an excellent agreement between GVVPT2

and CR-CC(2,3) while B3LYP predicted the reaction to be about 20 kcal/mol less favorable. For the formation of As_2O_5 form As_2O_3 and O_2, it was predicted to be far less favorable. B3LYP again predicted the reaction to be a little more than 20 kcal/mol less favorable than CR-CC(2,3). However, GVVPT2 did not agree well with CR-CC(2,3) in this reaction, predicting the reaction to be 18 kcal/mol more favorable, which is likely due to the high degree of multireference character. Based on both the structural data and thermodynamic information obtained, it appears that a gas phase accretion mechanism is possible. MRCISD calculations were used as a method replace with for smaller systems with high multireference character.

This study also confirms earlier studies that GVVPT2 can be used for accurate energy differences similarly to CCSD(T), in this case extended to greater range through linking and complete renormalization, i.e. CR-CC(2,3), and we expect good agreement with MRCISD calculations. Additionally, this study further validates the conclusions of earlier studies with GVVPT2 using incomplete model spaces.[58,80] Therefore, with GVVPT2's favorable scaling (i.e., N^5) along with its rigorous spin-adaptation, it should be given serious consideration for studies of complex potential energy surfaces.

4 GVVPT2 MULTIREFERENCE PERTURBATION THEORY DESCRIPTION OF ANTIMONY OXIDES RELEVANT TO COAL COMBUSTION

4.1 Introduction

As mentioned in the previous chapter, the combustion of coal for the production of electricity has seen considerable use historically and it still ranks the highest globally when compared to other methods of power production.[46] This is due to coal's low costs, global availability, and relatively high energy density. However, the emitted flue gases are known to contain potentially hazardous trace elements like arsenic, selenium and antimony.[15,47] Due to the harmful effects they pose to the environment and human health all the while adversely affecting power production

costs,[50] there has been an increasing interest in their remediation. In order to move toward accomplishing this, a more detailed understanding of their reactivities and thermodynamic stabilities must be obtained. This is particularly true for their various molecular oxide forms present in the flue gases. Insights into their chemical forms, oxidation states and structures present in the flue gases may aid in the development of novel strategies to combat these trace elements.[18–20] As mentioned above, one of the hazardous trace elements found in flue gases is antimony.[19,47,82] Like arsenic, antimony's level of toxicity depends on its form, concentration and oxidation state. The oxidation states in which antimony can be found are -3, +3, and +5. Inhalation of Sb_2O_3 is considered harmful and suspected of causing cancer.[82] It is clear that controlling the emission of antimony in flue gases is of great significance.

As is the case for other trace elements present in coal during combustion, it is necessary to understand the mechanisms that lead to the formation of the various antimony containing compounds as well as the relative energies of their corresponding isomers if total remediation of antimony escaping into the environment is to be achieved. The chemical forms of antimony in the flue gases are most likely to be either in its elemental form or as oxides. Though there have been a few theoretical studies on the monomeric oxides, SbO, SbO_2 and SbO_3,[83,84] the Sb_4O_6 cluster,[85] other Sb_xO_y clusters,[86] and the solid state of antimony oxides,[87] to the best of our knowledge a mechanism for the potential gas phase accretion from monoantimony oxides has yet to be investigated. Moreover, as was the case for the arsenic oxides, all of the previous studies used single reference-based approaches, thus the multireference character of the antimony oxides should be assessed. This is particularly true because the arsenic oxides proved to be posses a high degree of multireference character, which had significant effects on the reaction energies.

In this work, a relatively recently developed multireference method, i.e., second-order gen-

eralized van Vleck variant of multireference perturbation theory (GVVPT2),[13,14] will be used. The goal of this work was to study theoretically both monoantimony oxides (i.e., SbO, SbO_2, SbO_3) and diantimony oxides (i.e., various isomers of Sb_2O_3 and Sb_2O_5). Therefore, the equilibrium structures, heats of oxidation reactions, and relative stabilities of the possible isomers of Sb_2O_3 and Sb_2O_5 were investigated. There were two motivating questions to answer in this study, and that was to determine the degree of multireference character these species possess and whether or not they will accrete in the gas phase through a similar mechanism of the previously studied selenium and arsenic oxides.[58]

4.2 Theoretical Methodology

Similarly for the study on arsenic oxides presented in Chapter 3, the GVVPT2 method was used, which is described in detail in Chapter 2. GVVPT2 has proven itself successful in many applications where difficult multireference structures were studied.[58–62] When they are applicable, GVVPT2 has been shown to produce results in good agreement with CCSD[63] and CCSD(T).[64]

In our previous study of both monomeric and dimeric selenium oxides,[58] GVVPT2 was shown to produce results in good agreement with CR-CC(2,3).[23,65–68] In our previous study of arsenic oxides, GVVPT2 was shown to produce good agreement with MRCISD rather than with CR-CC(2,3). This because the arsenic oxides had such a high degree of simultaneous multiple bond breaking and forming that CR-CC(2,3) could no longer be used as a validating tool of comparison.

4.2.1 Computational Details

The general idea behind the computational protocol employed in this study was validated in our previous work with selenium and arsenic oxides and can be broken up into the following five steps. First, the geometries were generated through the use of density functional theory (DFT). More specifically, the B3LYP hybrid generalized gradient approximation variant[69–72] with the SDB-aug-cc-pVTZ basis set[88] was used for the antimony atoms, in which relativistic effects are included in the effective core potential, and the aug-cc-pVTZ basis set[75,89] for oxygen atoms. Second, those structures were confirmed to be either a stationary point or a saddle point through a vibrational analysis using the same level of theory. It should be pointed out that B3LYP has been shown to give geometries and vibrational frequencies with a reasonable compromise of accuracy and computational expense and was recently shown in our selenium and arsenic oxide studies to perform well with systems similar to this present work. Third, a Hartree-Fock (HF) and a restricted Møller–Plesset perturbation (RMP2) calculation using the above described SDB-aug-cc-pVTZ/aug-cc-pVTZ basis set was performed to obtain an initial approximation of the MOs and to determine an active space. Fourth, a multiconfigurational self-consistent field (MCSCF) calculation using the same basis set was performed to further optimize the MOs. Fifth, a GVVPT2 single point energy calculation using the MCSCF orbitals was conducted. It should also be noted that both the structural and thermodynamic data obtained from the B3LYP calculations was compared with the results from GVVPT2 because B3LYP was used in the geometry optimizations of the monomers and were used to locate the various isomers of the dimers and needed to be shown to agree well with a higher level of theory.

The details of the B3LYP calculations were intentionally identical to both the selenium and arsenic studies. The B3LYP geometry optimizations were performed using the GAMESS-US program.[77,78] All of the B3LYP calculations were of the RPA-derived Vosko-Wilke-Nusair correlation functional,[79] with 144 radial and 302 angular points for the Lebedev grid. The RHF + RMP2, MCSCF, and GVVPT2 single point energy calculations were performed using our local electronic structure program, known as UNDMOL.[14] Version 1.2 of UNDMOL was particularly helpful in this study, because it has enhanced capabilities of resolving the angular momentum components which transform as the same irreducible representation of the Abelian point groups (e.g., σ and δ orbitals in SbO) used in most of the algorithms. A 5–point numerical second- and third-differentiation of the SbO energy was used to obtain spectroscopic constants. The program used was called difreq, a small program developed by Dr. Hoffmann, and the step sizes used were 0.003 Bohr.

4.2.2 Active Spaces

In this study, model spaces of the MCSCF incomplete variety were used for all GVVPT2 calculations. These active spaces were defined as all configuration state functions (CSFs) obtained from distributing all electrons in one orbital group (denoted here as G_1), all but 1 electron in G_1 (and one electron in a second orbital group, G_2), and all but 2 electrons in G_1 (and 2 electrons in G_2). The effectiveness of this type of active space for reactive oxygen-rich systems was previously demonstrated[80] and confirmed by comparing results obtained using these active spaces for GVVPT2 by comparing them to CASSCF.[58] Furthermore, since antimony is directly below arsenic

on the periodic table, the number of active electrons in its valence shells are the same as arsenic. Therefore, the active spaces that were shown to work well for the arsenic oxides were used for the antimony oxides. Specifically for the GVVPT2 calculations, there were single and double excitations from all macroconfigurations used to describe the model space, and excitations were allowed from the higher-lying doubly occupied orbitals as well.

4.2.2.1 Monoantimony Oxides

The active spaces used for the monoantimony oxides consisted of the highest one-electron energy occupied molecular orbitals, G_1 (i.e., G_1 = {HOMO-(n − 1), ..., HOMO-1, HOMO}, where n is the number of occupied MOs is the active space) and the low-lying unoccupied molecular orbitals, G_2 (i.e., G_2 = {LUMO, LUMO+1, ..., LUMO+(n − 1)}, where n is the number of unoccupied MOs in the active space). As with the monoselenium oxides and monoarsenic oxides, the active spaces were derived from a valence picture, after preliminary studies.[58]

For SbO, the active space consisted of a set of 8 valence molecular orbitals (i.e., 2s and 2p orbitals of O, and the 5s and 5p orbitals of Sb) having 11 electrons (11e, 8o) and was divided into two subgroups: G_1 = {$2a_1$, $3a_1$, $1b_1$, $4a_1$, $1b_2$, $2b_2$, $2b_1$}, and G_2 = {$5a_1$}. Because SbO is linear, both the $2b_1$ and $2b_2$ (i.e., physically there is no reason to favor the the p_x vs. p_y orbitals) orbitals were included in G_1 rather than splitting them up among the G_1 and G_1 subgroups. The SbO_2 molecule had 12 valence orbitals in its active space (i.e., 2s and 2p orbitals of O, and the 5s and 5p orbitals of Sb) with 17 electrons (17e, 12o) divided into two subgroups: G_1 = {$2a_1$, $2b_2$, $3a_1$, $4a_1$, $3b_2$, $1b_1$, $4b_2$, $1a_2$, $5a_1$} and G_2 = {$2b_1$, $6a_1$, $5b_2$}. Similarly, the valence molecular orbitals for SbO_3 were partitioned as: G_1 = {3a'–10a', 2a"–5a"} and G_2 = {11a'–12a', 6a"–7a"}, giving an

active space of (23e, 16o).

4.2.2.2 Diantimony Oxides

The active orbitals of the diantimony oxides were also divided into two subgroups, G_1 and G_2. The single point energy calculations of the Sb_2O_3 structures used 20 active orbitals (i.e., the 2s and 2p of O, and the 5s and 5p of Sb) which contained 28 electrons (28e, 20o) with 14 orbitals in G_1 and 6 orbitals in G_2. The orbitals were distributed as G_1 = {$3a_1$, $4a_1$, $5a_1$, $6a_1$, $6a_1$, $7a_1$, $8a_1$, $1a_2$, $1b_1$, $2b_1$, $3b_1$, $2b_2$, $3b_2$, $4b_2$, $5b_2$} and G_2 = {$9a_1$, $10a_1$, $2a_2$, $4b_1$, $5b_1$, $6b_2$} for isomer 1, G_1 = {3a'–11a', 2a"–6a"} and G_2 = {11a'–15a', 7a"–8a"} for isomer 2, and G_1 = {4a'–14a', 1a"–3a"} and G_2 = {15a'–18a', 4a"–5a"} for isomer 3.

The single point energy calculations of the isomers of Sb_2O_5 used 28 active orbitals (i.e., the 2s and 2p of O, and the 5s and 5p of Sb) containing 40 electrons (40e, 28o) with 20 orbitals in G_1 and 8 orbitals in G_2. The orbitals were distributed as G_1 = {$4a_1$–$11a_1$, $1a_2$–$2a_2$, $2b_1$–$6b_1$, $2b_2$–$6b_2$} and G_2 = {$12a_1$–$14a_1$, $3a_2$, $7b_1$, $7b_2$–$9b_2$} for isomer 1, G_1 = {$4a_1$–$11a_1$, $1a_2$–$2a_2$, $2b_1$–$6b_1$, $2b_2$–$6b_2$} and G_2 = {$12a_1$–$15a_1$, $7b_1$–$8b_1$, $7b_2$–$8b_2$} for isomer 2, G_1 = {5a'–17a', 2a"–8a"} and G_2 = {18a'–23a', 9a"–10a"} for isomer 3, and trivially G_1 = {6a–25a} and G_2 = {26a–33a} for isomer 4.

4.3 Results and Discussion

The goals of this work, as was the case for the arsenic oxides presented in Chapter 3, were

to analyze the antimony oxides' multireference character through GVVPT2 and to determine the likelihood of a gas phase accretion mechanism from the monoantimony oxides into diantimony oxides through careful structural and thermodynamic analysis. These goals will be addressed as the results are presented, starting with the monomeric antimony oxides followed by the dimeric isomers.

4.3.1 Antimony Monoxide (SbO)

Since the effects of basis set and model space variations was already assessed in our previous study,[58] the effect of inclusion of dynamic correlation is all we need to consider here. This effect was assessed through comparing spectroscopic constants obtained via B3LYP and GVVPT2 calculations on the ground state of SbO and is given below in Table 4.1.

Table 4.1. Comparison of thermodynamic constants from geometries optimized at various levels of theory for the ground state of SbO.

Method	r_e (Å)	ω_e (cm^{-1})	$\omega_e x_e$ (cm^{-1})	α_e (cm^{-1})	B_e (cm^{-1})	D_e (kcal/mol)
B3LYP	1.821	849.96	3.19	-0.0038	0.35989	-33.5
GVVPT2(11e,8o)	1.842	791.35	–	-0.0036	0.35189	-33.3

Though SbO possesses $C_{\infty v}$ symmetry, it was studied using the C_{2v} point group when using UNDMOL and with C_{4v} in GAMESS calculations due to software constraints. The ground electronic state of SbO was predicted to be $^2B_1$. The geometry of SbO was optimized with both B3LYP and GVVPT2 to study the effect that various levels of theory would have on the geometrical parameters in the antimony oxides. The r_e bond length of SbO was found to be 1.821 Å with B3LYP and slightly longer at 1.842 Å with GVVPT2 giving a difference of only 0.021 Å.

Table 4.2. Leading configurations and weights of MCSCF reference functions of antimony oxides given as CSFs in the GVVPT2 output.

	Abelian (sub)group	amplitude	configuration
SbO	C_{2v}	0.905	$2a_1^2 3a_1^2 4a_1^2 1b_1^2 2b_1^1 1b_2^2 2b_2^0 5a_1^0$
		-0.167	$1b_2 \rightarrow 2b_2$
		-0.117	$1b_1 \rightarrow 2b_1$
		-0.115	$1b_1 \rightarrow 2b_1$, $1b_2 \rightarrow 2b_2$
SbO_2	C_{2v}	0.843	$2a_1^2 3a_1^2 4a_1^2 5a_1^1 1a_2^2 1b_1^2 2b_2^2 3b_2^2 4b_2^2 6a_1^0 2b_1^0 5b_2^0$
		-0.159	$1b_1 \rightarrow 2b_1$
		-0.142	$1a_2 4b_2 \rightarrow 5a_1 2b_1$
		0.139	$4a_1 \rightarrow 5a_1$
SbO_3	C_s	0.883	$3a'^2$–$9a'^2$, $10a'^1$, $2a''^2$–$5a''^2$, $11a'^0 12a'^0 6a''^0 7a''^0$

Table 4.2 displays the dominant amplitudes of the monomeric antimony oxides and the corresponding configurations of the MCSCF wave function near the equilibrium geometry for the purpose of displaying the multireference character of the monomeric species. As can be seen, while the wave function for SbO has a dominant configuration with an amplitude of 0.818, there are six additional significant (≥ 0.09) configurations, showing substantial multireference behavior. Likewise, SbO_2 exhibits considerable multireference character. In contrast, SbO_3 wasn't predicted to have any secondary configurations of major importance.

4.3.2 Antimony Dioxide (SbO_2)

SbO_2 was predicted to possess C_{2v} symmetry with the structural parameters given below in Table 4.3. The ground electronic state of SbO_2 was predicted to be $^2A_1$. The SbO_2 r_e bond length calculated at the B3LYP level of theory is longer than the r_e of SbO by 0.013 Å, which was interestingly the same difference observed for the r_e of AsO compared to AsO_2, shown in Chapter 3.

Table 4.3. Comparison of structural parameters (Å and degrees) of SbO and SbO_2, optimized at the B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ levels.

	Symmetry	Parameter	B3LYP	GVVPT2
SbO	$C_{\infty v}$	r_e(Sb–O)	1.821	1.853
SbO_2	C_{2v}	r_e(Sb–O)	1.834	1.851
		∠ O–Sb–O	121.8	125.3

The formation of SbO_2 from the reaction of SbO and $\frac{1}{2}O_2$ was predicted to be energetically favorable at both the B3LYP and GVVPT2 level of theory. The reaction energies were found to be -23.4 kcal/mol at the B3LYP level, -29.3 kcal/mol at the GVVPT2 level (see Table 4.5 below).

4.3.3 Antimony Trioxide (SbO_3)

SbO_3 was predicted to possess C_s symmetry with two Sb–O bonds being longer than the other one while being in a trigonal pyramidal-type structure. The ground electronic state of SbO_3 was predicted to be 2A'.The predicted geometrical parameters are given below in Table 4.4 The SbO_3 r_{e1} bond length calculated at the B3LYP level of theory was found to be longer than the r_e of SbO by 0.146 Å and longer than the r_e of SbO_2 by 0.133 Å. Similarly, the SbO_3 r_{e2} bond length

calculated at the B3LYP level of theory was found to be longer than the r_e of SbO by 0.142 Å and shorter than the r_e of SbO_2 by 0.129 Å.

Table 4.4. The geometrical parameters (Å and degrees) of the SbO_3 doublet, optimized at the B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ levels.

	Symmetry	Parameter	B3LYP	GVVPT2
SbO_3	C_s	r_{e1}(Sb–O1)	1.966	1.961
		r_{e2}(Sb–O2/O3)	1.963	1.973
		∠ O1–Sb–O2/O3	101.3	101.1
		∠ O2–Sb–O3	46.1	47.6

Reaction enthalpies are presented below in Table 4.5. The formation of SbO_3 from the reaction of SbO and O_2 was predicted to be favorable with B3LYP and to be nearly isoenergetic with GVVPT2. However, the formation of SbO_3 from the reaction of SbO_2 and $\frac{1}{2}O_2$ was predicted to be energetically unfavorable at both the B3LYP and GVVPT2 levels of theory.

Table 4.5. Reaction enthalpies (kcal/mol) for the oxidation of monomeric antimony oxides using B3LYP and GVVPT2//B3LYP using the SDB-aug-cc-pVTZ basis set.

Reaction	B3LYP	GVVPT2
$Sb + \frac{1}{2}O_2 \rightarrow SbO$	-33.5	-33.1
$SbO + \frac{1}{2}O_2 \rightarrow SbO_2$	-23.4	-29.3
$SbO + O_2 \rightarrow SbO_3$	-12.6	1.7
$SbO_2 + \frac{1}{2}O_2 \rightarrow SbO_3$	10.9	31.0

Therefore, based on the oxidation reaction energies, GVVPT2 shows that the order of stability for the antimony monoxides is as follows

$$\text{(least stable)} \quad SbO_3 \quad < \quad SbO \quad < \quad SbO_2 \quad \text{(most stable)}$$

4.3.4 Diantimony Trioxide (Sb_2O_3)

Three minima (i.e., isomers 1, 2, and 3) of diantimony trioxide (Sb_2O_3) molecule were found. Similarly to the arsenic dimers, the Sb_2O_3 isomer 1 was predicted to be a triple oxygen bridging structure that possesses D_{3h} symmetry. Due to software limitations, it was rotated 90° and treated as C_{2v} with a ground electronic state of $^1A_1$. Isomer 2 was predicted to be a ring structure with an oxygen jutting out of the plane of the ring that possesses C_s symmetry and a ground electronic state of ^{1}A'. Isomer 3 was predicted to have a single oxygen bridge between the antimony atoms, which are then each bonded to an additional oxygen. It has a symmetry of C_s and a ground electronic state of ^{1}A'. B3LYP optimized geometrical parameters are given below in Table 4.6. The numbering scheme lists atoms as 1–5 and has the antimony atoms always being atom 1 and 2.

Table 4.6. Optimized B3LYP geometrical parameters (Å and degrees) of the diantimony trioxide (Sb_2O_3) isomers.

Parameters	Isomer 1 (D_{3h})	Isomer 2 (C_s)	Isomer 3 (C_s)
$R_{Sb1\text{-}O3}$	2.008	1.880	1.988
$R_{Sb1\text{-}O4}$	2.008	1.880	1.805
$R_{Sb2\text{-}O3}$	2.008	2.199	1.955
$R_{Sb2\text{-}O4}$	2.008	2.199	–
$R_{Sb2\text{-}O5}$	2.008	1.809	1.813
$R_{Sb1\text{-}Sb2}$	2.691	3.097	3.606
∠ O3–Sb2–O5	79.9	99.2	102.2
∠ Sb1–O3–Sb2	84.3	98.5	132.3
∠ O3–Sb1–O4	79.9	89.1	101.0

As was the case for the arsenic oxides study presented in Chapter 3, the optimized bond

lengths fall into two general categories: short bond length and long bond length. When comparing the geometrical parameters for Sb_2O_3 to that of the monomers, it was found that many of the long Sb–O bonds in the equilibrium structures of Sb_2O_3 (i.e., 1.988–2.199 Å) were significantly longer than the r_e(Sb–O) bond lengths in SbO (i.e., 1.821 Å), SbO_2 (i.e., 1.834 Å) and SbO_3 (i.e., 1.963 Å and 1.966 Å). Notably, some of the short Sb–O bonds of the Sb_2O_3 isomers (i.e., 1.805–1.955 Å) were found to be shorter than some of the r_e(Sb–O) bond lengths exhibited by the monomers.

The frequencies of the Sb_2O_3 isomers were obtained at the B3LYP/SDB-aug-cc-pVTZ level. They are given below in Table 4.7. Each of the isomers, for which the frequencies have been calculated, were confirmed as such by having no imaginary frequencies.

Table 4.7. Frequency of diantimony trioxide (Sb_2O_3) isomers in cm^{-1} calculated at the B3LYP level of theory.

Normal Mode	Isomer 1 (D_{3h})	Isomer 2 (C_s)	Isomer 3 (C_s)
1	322.95 (A_1')	91.13 (A')	48.59 (A")
2	351.23 (E')	194.72 (A")	66.42 (A')
3	353.12 (E')	209.64 (A')	116.37 (A")
4	387.36 (E")	248.64 (A")	182.88 (A')
5	387.56 (E")	280.28 (A')	246.29 (A')
6	580.04 (A_2")	437.32 (A')	438.40 (A')
7	595.05 (E')	703.34 (A")	716.57 (A')
8	595.25 (E')	768.76 (A')	862.67 (A')
9	683.26 (A_1')	863.70 (A')	874.47 (A')

The relative energies found for each level of theory used are given in Table 4.8. When the comparison between the single point energy calculations using B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ, both methods found the same relative ordering for where GVVPT2 results are available. As was the case for As_2O_3, the lowest lying isomer predicted by GVVPT2 was the triple oxygen bridging structure. The predicted stability of the rest of the isomers goes in

order of their number.

Table 4.8. Relative energies (kcal/mol) at the B3LYP and GVVPT2 levels of theory of the diantimony trioxide (Sb_2O_3) isomers.

	Symmetry	B3LYP	GVVPT2
Isomer 1	D_{3h}/C_{2v}	0.00	0.00
Isomer 2	C_s	27.1	28.4
Isomer 3	C_s	36.4	51.15

In analogy with the reaction energy predictions made for the monomers (cf. Table 4.5), the stability of the monomeric oxides to formation of Sb_2O_3 were studied; i.e.,

$$SbO + SbO_2 \rightarrow Sb_2O_3 \tag{4.1}$$

The reaction energies as calculated by B3LYP/SDB-aug-cc-pVTZ and GVVPT2/SDB-aug-cc-pVTZ are given below in Table 4.9. All predicted reaction energies were found to be highly energetically favorable. Favorable reaction energies of this magnitude support the possibility of a gas phase accretion mechanism.

Table 4.9. Total energies (a.u.) of SbO, SbO_2 and Sb_2O_3 (isomer 1), and predicted reaction energies (kcal/mol).

	SbO	SbO_2	Sb_2O_3 (**1**)	Δ E (kcal/mol)
B3LYP	-80.640 485	-155.867 723	-236.713 331	-128.7
GVVPT2//B3LYP	-80.467 518	-155.580 066	-236.232 482	-116.0

4.3.5 Diantimony Pentoxide (Sb_2O_5)

Four minima have been located on the diantimony pentoxide (Sb_2O_5) surface (i.e., isomers 1, 2, 3, and 4). Isomer 1 was predicted to have a triple oxygen bridging structure that possesses D_{3h} symmetry, similar to the structural motif of the Sb_2O_3 isomer 1, with one of each of the additional oxygens jutting out from the two antimony atoms away from the triple oxygen bridging structure. Again, due to software limitations, it was rotated 90° and treated as C_{2v} with a ground electronic state of $1A_1$. Isomer 2 was found to be a ring structure with an oxygen bonded above and below the ring on one antimony and a third oxygen in plane jutting out away from the ring on the other antimony. This structure has C_{2v} symmetry and a ground electronic state of $1A_1$. Isomer 3, found to be C_s symmetry and have a ground electronic state of 1A', has the two antimony atoms triply bridging by oxygens, similar to isomer 1, but one of the bridges has two oxygen atoms, with a O–O distance of 1.466 Å, while the other two bridges have one. The fifth oxygen is jutting out on one of the antimony away from the center of mass. Isomer 4 was predicted to have C_2 symmetry, but due to software limitations, it was run at C_1. It has a zigzag structure with an extra oxygen bonded to each antimony and half of the zigzag is rotated out of plane along the bond to the bridging oxygen. B3LYP/SDB-aug-cc-pVTZ optimized geometrical parameters for the isomers are given below in Table 4.10. The numbering scheme lists atoms as 1–7 and always has the antimony atoms listed as 1 and 2.

The frequencies of the Sb_2O_5 isomers were calculated at the B3LYP/SDB-aug-cc-pVTZ level and are given below in Table 4.11.

The similarities in the structural data of the monomer and of several isomers of Sb_2O_3 to

Table 4.10. Optimized B3LYP/SDB-aug-cc-pVTZ geometrical parameters (Å and degrees) of the diantimony pentoxide (Sb_2O_5) isomers.

Parameters	Isomer 1 (D_{3h})	Isomer 2 (C_{2v})	Isomer 3 (C_s)	Isomer 4 (C_2)
$R_{Sb1\text{-}O3}$	1.798	1.804	1.990	1.800
$R_{Sb1\text{-}O4}$	1.999	1.804	1.990	1.804
$R_{Sb1\text{-}O5}$	1.999	2.166	2.029	1.945
$R_{Sb1\text{-}O6}$	1.999	2.166	–	–
$R_{Sb2\text{-}O4}$	1.999	–	1.970	–
$R_{Sb2\text{-}O5}$	1.999	1.873	–	1.945
$R_{Sb2\text{-}O6}$	1.999	1.873	1.997	1.804
$R_{Sb2\text{-}O7}$	1.798	1.796	1.801	1.800
$\angle_{(O3)(Sb1)(O4)}$	131.5	142.6	80.3	142.4
$\angle_{(O6)(Sb1)(O4)}$	80.9	104.6	–	–
$\angle_{(O5)(Sb2)(O7)}$	131.5	134.6	–	110.6
$\angle_{(O6)(Sb2)(O4)}$	80.9	–	91.4	–
$\angle_{(O6)(Sb2)(O7)}$	131.5	134.6	125.4	142.4
$\angle_{(Sb1)(O5)(Sb2)}$	83.0	96.6	–	128.3
$\angle_{(Sb1)(O4)(Sb2)}$	83.0	–	91.4	–

that of the isomers of Sb_2O_5 suggest that several characteristics of their bonding motifs are retained. This is especially true for isomers 1 and 2 of Sb_2O_5 when compared to isomers 1 and 2 of Sb_2O_3. As observed for Sb_2O_3, the optimized bond lengths again fall into the two general categories of short and long bond lengths. The similarity of bond lengths and angles in the found isomers of Sb_2O_5 raises the likelihood of isomerization. The relative energies found for each level of theory are given below in Table 4.12.

GVVP2 predicted that isomer 1, with a cage-like structure similar to the most stable isomer of Sb_2O_3, (i.e., isomer 1) was the most stable isomer of the Sb_2O_5. Isomer 3 was found to be the next most stable with an energy of 4.4 kcal/mol above isomer 1, which B3LYP had mistaken for the lowest energy isomer at 11.9 kcal/mol below isomer 1.

Table 4.11. Frequencies of diantimony pentoxide (Sb_2O_5) isomers obtained at the B3LYP/SDB-aug-cc-pVTZ level of theory.

Normal Mode	Isomer 1 (D_{3h})	Isomer 2 (C_{2v})	Isomer 3 (C_s)	Isomer 4 (C_2)
1	130.19 (E')	50.21 (B_2)	138.65 (A')	38.36 (B)
2	133.02 (E')	118.15 (B_1)	154.47 (A")	40.45 (A)
3	133.38 (E')	123.87 (B_2)	196.33 (A")	56.77 (A)
4	142.78 (E")	168.99 (A_2)	270.74 (A")	138.99 (B)
5	311.21 (A_1')	170.16 (A_1)	298.80 (A')	150.12 (A)
6	374.58 (E')	181.72 (B_1)	305.24 (A')	170.74 (A)
7	376.56 (E')	248.68 (A_1)	344.33 (A')	209.25 (B)
8	399.43 (E")	270.71 (B_2)	486.61 (A")	242.48 (A)
9	399.75 (E")	345.82 (B_1)	513.62 (A')	246.68 (B)
10	543.36 (A_2")	458.66 (A_1)	573.27 (A')	456.95 (A)
11	640.72 (E')	726.86 (B_1)	615.34 (A')	749.08 (B)
12	642.05 (E')	755.88 (A_1)	634.17 (A')	814.75 (B)
13	704.61 (A_1')	810.86 (A_1)	685.53 (A')	820.07 (A)
14	855.12 (A_2")	872.83 (A_1)	856.15 (A')	887.07 (B)
15	863.89 (A_1')	880.71 (B_2)	864.68 (A')	888.07 (A)

The energics of two reactions were studied for diantimony pentoxide, given as

$$SbO_2 + SbO_3 \rightarrow Sb_2O_5 \tag{4.2}$$

and

$$Sb_2O_3 + O_2 \rightarrow Sb_2O_5 \tag{4.3}$$

The reaction energies as calculated by B3LYP and GVVPT2 are given in Table 4.13. Reaction 4.2 was found to be highly energetically favorable. However, interestingly, reaction 4.3 was predicted to be energetically unfavorable at the B3LYP level. This shows a favorable gas phase accretion to Sb_2O_5 from the monomers, but not from Sb_2O_3. Therefore, a gas phase accretion

Table 4.12. Relative energies (kcal/mol) at the B3LYP and GVVPT2 levels of theory of the diantimony pentoxide (Sb_2O_5) isomers.

	Symmetry	B3LYP	GVVPT2
Isomer 1	D_{3h}	0.0	0.0
Isomer 2	C_{2v}	20.2	15.9
Isomer 3	C_s	-11.9	4.4
Isomer 4	C_2	30.4	36.2

mechanism is supported, but the most stable form of Sb_xO_y that was considered in this study was found to be Sb_2O_3.

Table 4.13. Total energies (a.u.) of O_2, SbO_3 and Sb_2O_5 (isomer 1), and predicted reaction energies (kcal/mol).

	O_2	SbO_3	Sb_2O_5 (**1**)	Δ E (SbO_2 + SbO_3 $\rightarrow$$Sb_2O_5$ (**1**))	Δ E (Sb_2O_3 (**1**)+O_2 $\rightarrow$$Sb_2O_5$ (**1**))
B3LYP	-150.379 839	-231.040 314	-387.059 456	-95.0	21.2
GVVPT2 //B3LYP	-150.131 720	-230.596 594	-386.361 977	-116.3	1.4

4.4 Summary

On the basis of good agreement between B3LYP, GVVPT2 optimized geometries for similar systems[58] as well as for the study on the arsenic oxides presented in Chapter 3, B3LYP/SBD-aug-cc-pVTZ geometry optimizations were performed for SbO, SbO_3, Sb_2O_3 and Sb_2O_5. Five isomers were obtained for Sb_2O_3 and two isomers were obtained for Sb_2O_5. Single point energy calculations of the obtained geometries were then conducted at the GVVPT2/SBD-aug-cc-pVTZ level after finding that the antimony oxides possess a high degree of multireference character. GVVPT2 has shown excellent applicability for studying systems with a large amount of multireference character.[58,59]

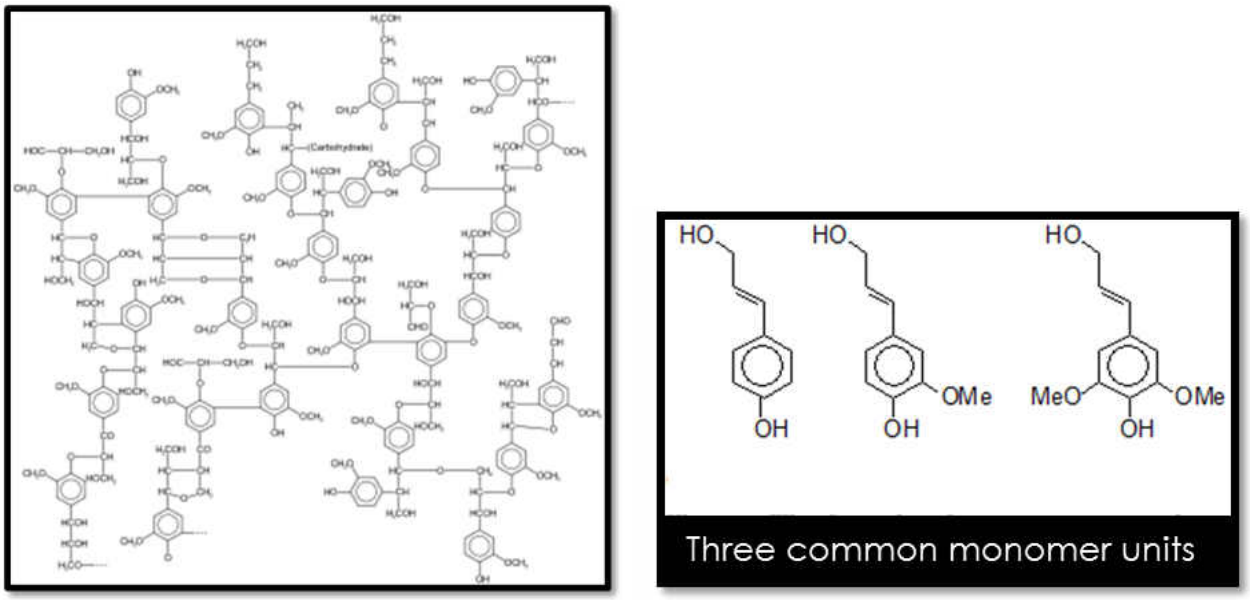
From the reaction thermodynamics calculations, it was found that a gas phase accretion would be energetically favorable. Whether starting from Sb and O_2 or from the monoantimony oxides, the accretion into Sb_2O_3 was found to be favored at the B3LYP/SBD-aug-cc-pVTZ and GVVPT2/SBD-aug-cc-pVTZ//B3LYP/SBD-aug-cc-pVTZ levels of theory. In contrast, the accretion into Sb_2O_5 was only favorable with respect to the monomers, and not when compared to Sb_2O_3.

5 ELECTRONIC STRUCTURE CALCULATIONS OF CATALYZED LIGNIN DECOMPOSITION

5.1 Introduction

Lignin is a highly branched and cross-linked naturally occurring network polymer which exists in the secondary cell walls of wood and various plants where it binds the cells, cellulose fibers and vessels. Second only to cellulose, lignin is one of the most abundant natural organic polymers on Earth. Despite its abundance, lignin has only seen limited use due to its chemical stability and complex structure. There has been significant efforts to characterize the structure of the various types of lignin and several important features of its structures and the various respective functions have become more apparent from these studies. There have been many experimental studies, involving spectroscopy[90] coupled with other techniques,[91–95] and some computational studies[96,97] devoted to this effort and have advanced our understanding of the important structural

features, constituents, and various linkages of lignin. An example of a potential structure of lignin is given below, in Figure 5.1).



(a) (b)

Figure 5.1. (a) An example of the structure of a small piece of lignin. (b) The three common monomer units, from left to right: p-hydroxyphenyl, guaiacyl, and syringyl.

Lignin can be obtained from many sources which are typically treated as waste (e.g., forest waste, agricultural waste, switchgrass, etc.). If lignin could be broken down in an economically viable way into renewable chemicals and polymeric alternatives to petrochemicals, the national dependence on imported fossil fuels would be reduced. These high-value lignin-derived chemicals, particularly aromatic compounds, would be coming from one of the most renewable carbon sources on the planet.[21] Moreover, rather than competing with food sources (e.g., ethanol from corn), lignin would augment the economics of food production, because it would be coming from the non-edible stalks of plants.

Although a renewable source like lignin has great potential for a wide variety of uses, it remains largely untapped because its complicated structure makes it too expensive to break down

into useful products on a large scale. Therefore, there has been great recent interest in finding and designing catalysts for the purpose of controlled and inexpensive degradation of lignin.[21] One potential catalyst that shows great potential is known as amorphous silica–alumina (ASA). ASA is a microporous synthetic catalyst typically consisting of silicon, aluminum and oxygen. A scanning electron micrograph (SEM) of ASA is given below in Figure 5.2.

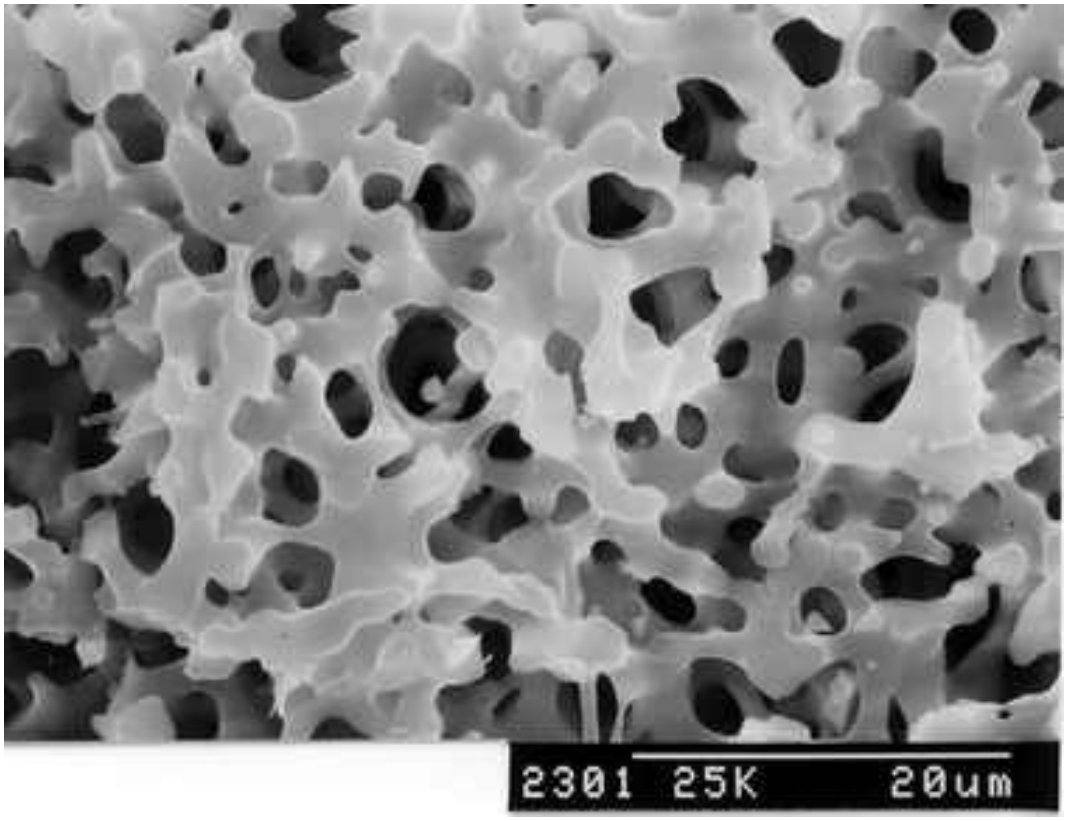


Figure 5.2. SEM of fractured surface of an amorphous silica–alumina sample.[2]

Despite its promise as a catalyst for lignin degradation, the underlying mechanisms of catalysis and the effects of metal doping are not fully known. A better understanding of the mechanisms of the relevant reactions and gaining a better understanding of the effects of metal doping would be highly beneficial to aid in the design of novel catalysts derived from ASA. Computational chemistry, in particular density functional theory (DFT), can provide valuable insights into the inner

workings of this catalytic lignin degradation. DFT can be used to efficiently screen potential catalysts and determine what makes certain catalysts better than others, thereby aiding in the rational development the next generation of catalysts. In this work, DFT is used with the aim of aiding experimentalists, in the ways mentioned above, as part of a multi-institutional collaboration known as DakotaBioCon and the rest of this chapter of my dissertation will give the highlights of those efforts.

5.2 Methods

5.2.1 Model Construction

When constructing a model for the catalyst and lignin system a balance of size and speed needed to be found. The larger the size of the catalyst and the larger the fragment of lignin, the closer the numerical calculations should be expected to correspond to the actual system. However, the larger the system the longer the calculations take. Thus, several test calculations needed to be conducted to find the optimum system size. When considering a model for the catalyst and its size, several things needed to be taken into account (i.e., most notably the electronic structure of the catalyst). Whether we could use a cluster model or would be forced to do periodic boundary conditions would depend on the nature of the catalyst, in particular, did it have a large or small band gap. ASA, like zeolites, has a comparatively wide band gap, so we were able to use a cluster model. Several sizes were analyzed, but a good balance of accuracy and speed was found in a system of 183 silicon, oxygen, aluminum and hydrogen (to cap dangling bonds) with 2 aluminum atoms for

every 23 silicon atoms.[98] Similar considerations needed to be taken into account when modeling lignin. Lignin is known to be a difficult system to model effectively, both experimentally and theoretically. Our approach for modeling lignin has been to start with the simplest fragments possible and gradually increase size and complexity. We chose to start with a p-hydroxyphenyl monomer and a dimer consisting of p-hydroxyphenyl and guaiacyl β–O–4 linked monomers, which are two well-known fragments from the literature.[21] The cleavage of the β–O–4 ether bond is known to be the primary way lignin is depolymerized and typically results in compounds containing phenolic hydroxyl groups.[90]

5.2.2 Computational Details

Though several DFT functionals were tested, the hybrid functional B3LYP[69,70,72] was chosen due to its comparatively high level of accuracy while still providing relatively fast convergence. To account for long range interactions, Grimme's D3 dispersion correction[99] has been used. After various basis sets were tested, a good balance of speed vs. accuracy for our system was found in 6-31G*.[100–102] The total number of atoms for the p-hydroxyphenyl monomer with the catalyst model is 204 atoms and the total number for the β–O–4 linked dimer with the catalyst model is 227 atoms. Partial geometry optimizations were conducted with the lower two levels of the catalyst surface (i.e., the lower 86 out of 183 atoms) frozen in place to better simulate its connection to the bulk. When modeling the ASA cluster, the amount of aluminum substitution in the silicon T sites was tested, but the optimal value that was used was obtained from the literature, where they found it to work best for aromatic organics. This variation is defined in the literature as Si/(Al x

0.5)=23,[98] which essentially means that for every 23 Si atoms, there should be two Al atoms.

For the construction and modification of the bare catalyst as well as capping the dangling bonds with hydrogens, the freely available program Avogadro was used.[103] After the model was constructed, Avogadro's molecular mechanics driven geometry optimization was used to prepare the system for the B3LYP/6-31G* optimizations on the University of North Dakota's local Linux cluster, Hodor, located in the UND's Computational Research Center (CRC). This step was also used to allow the system to "relax" out of a more perfect crystallization state and into a more amorphous state. All the DFT calculations were conducted with the use of the freely available NWChem software package.[104]

Various tests were also conducted to determine the optimal number of nodes on the Hodor, and it was found that using 12 nodes with 8 processor cores each gave the best speed up. There was significant reworking in both the job queuing scripts as well the input files in order to set up a program like NWChem, which appeared to be optimized, by default, for large workstations rather than supercomputer architecture. The massive amount of I/O that NWChem generated by the way it calculated and stored the integrals and other scratch files needed to be treated carefully by storing them locally on the node running the calculations. After determining this and reconfiguring, the calculations sped up 20- to 30-fold when compared to its initial configuration.

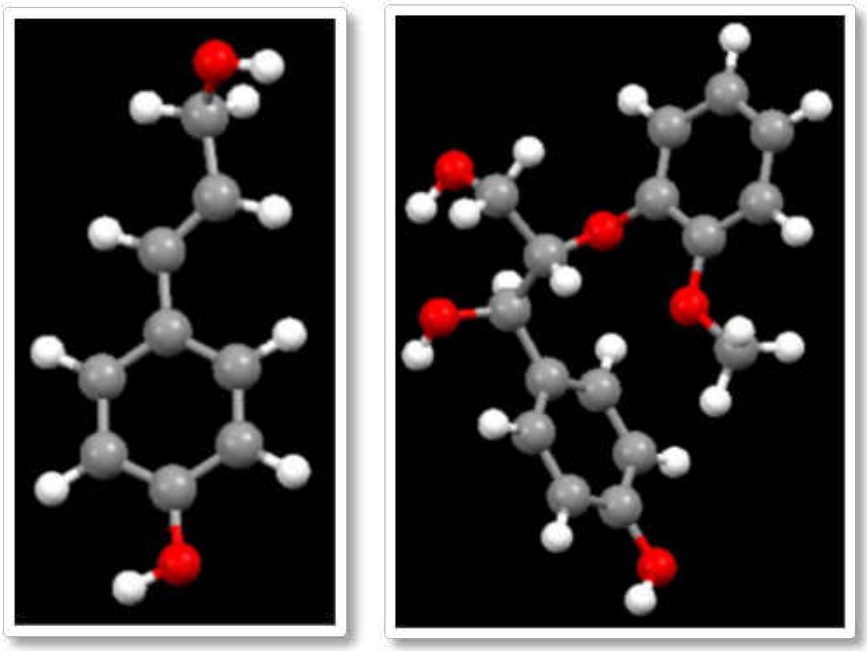
In order to obtain the binding energies of the various lignin fragments on the ASA model surface, several calculations needed to be completed. First, a B3LYP/6-31G* geometry optimization of an isolated lignin fragment was conducted. Second, a constrained B3LYP/6-31G* optimization of the isolated ASA cluster was performed. Finally, the systems were combined and re-optimized, again using the B3LYP/6-31G* level of the entire system. The binding energies could then be calculated as

$$E_B = E_{\text{Sub+Surf}} - E_{\text{Surf}} - E_{\text{Sub}} \tag{5.1}$$

where $E_{\text{Sub+Surf}}$ is the B3LYP energy of the combined system, E_{Surf} is the B3LYP energy of the ASA model, and E_{Sub} is the B3LYP energy of the substrate (i.e., the isolated monomer or dimer).

5.3 Results and Discussion

Both qualitative and quantitative results have been obtained. The above summarized computational protocol has been successfully developed. The optimized structures of many systems have been obtained. The optimized structures of the monomer and dimer are given below in Figure 5.3.



(a) (b)

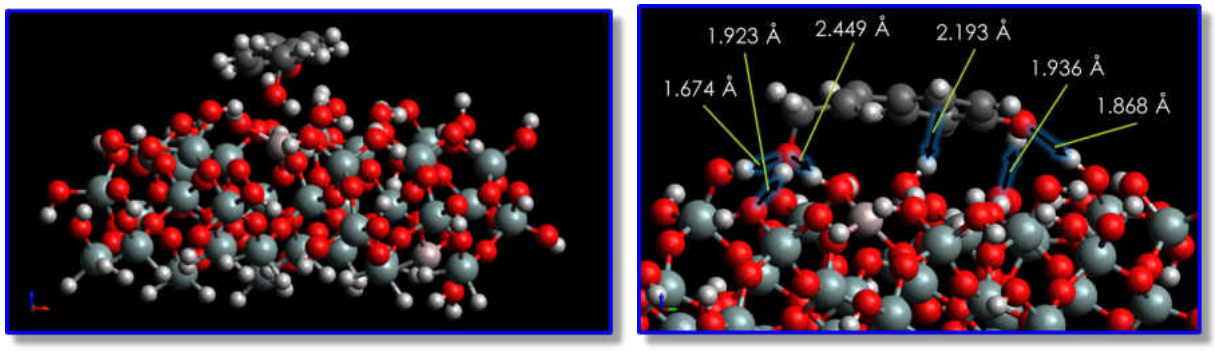
Figure 5.3. (a) Optimized structure of p-hydroxyphenyl alcohol, (b) Optimized structure of a β–O–4 linked model compound (i.e., a p-hydroxyphenyl linked to a guaiacyl).

The binding energies of several systems have been obtained and the relative energies of the products and reactants of the β–O–4 bond cleavage has been calculated. Three of the systems are given below.

The binding energy, E_B, was obtained for the monomer, p-hydroxyphenyl, adsorbed onto the ASA model surface given by the reaction

$$\text{monomer}_{\text{free}} + \text{ASA}_{\text{clean surface}} \rightarrow \text{monomer}_{\text{bound}} + \text{ASA}_{\text{adsorbate on surface}} \tag{5.2}$$

The binding energy was found to be exothermic with a value of -45.8 kcal/mol. Below, in Figure 5.4, the optimized system is given. The large magnitude of the binding energy is likely due to the six strong hydrogen bonds formed between the monomer and the surface. For comparison to the hydrogen bond strengths of the catalyst surface, the hydrogen bond distance in liquid water at room temperature is 1.97 Å. It can be seen that several of the surface hydrogen bonds are stronger. The large magnitude of the binding energies should also decrease when solvent effects are taken into account.



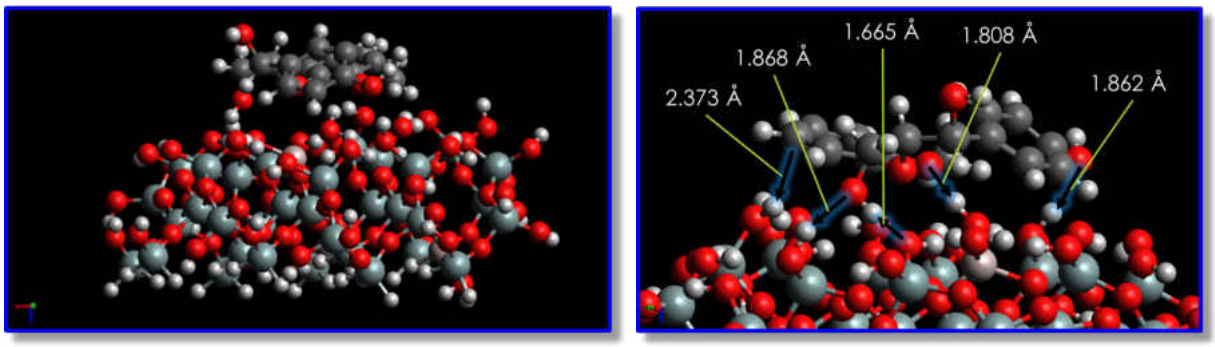
(a) (b)

Figure 5.4. (a) The full B3LYP/6-31G* optimized structure of p-hydroxyphenyl adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of p-hydroxyphenyl with the surface showing the hydrogen bond distances.

The binding energy, E_B, for the β–O–4 linked dimer adsorbed onto the silica–alumina model surface was calculated for the reaction

$$\text{dimer}_{\text{free}} + \text{ASA}_{\text{clean surface}} \rightarrow \text{dimer}_{\text{bound}} + \text{ASA}_{\text{adsorbate on surface}} \tag{5.3}$$

The binding energy was found to be exothermic by -54.7 kcal/mol. Again the magnitude is unrealistically high for the same reasons mentioned above. However, it is important to note that this study is more interested in relative energies as the catalyst surface is modified, than the absolute values. Therefore, the numbers are considered useful. The optimized geometry is given below in Figure 5.5.



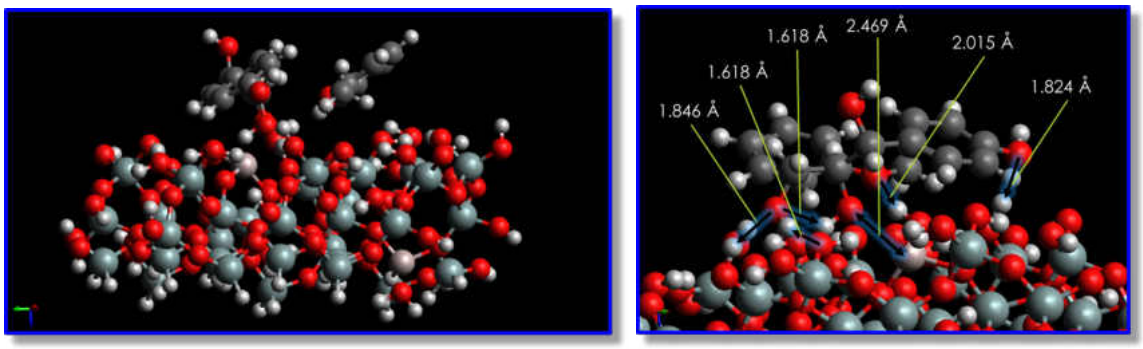
(a) (b)

Figure 5.5. (a) The full B3LYP/6-31G* optimized structure of the β–O–4 linked p-hydroxyphenyl and guaiacyl (i.e., the dimer) adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of the dimer with the surface showing the hydrogen bond distances.

The reaction energy, ΔE, for the cleavage of the β–O–4 bond of the adsorbed dimer with the dangling bonds capped with hydrogens was obtained for the reaction

$$\text{dimer}_{\text{bound}} + \text{ASA}_{\text{adsorbate on surface}} + \text{H}_2 \rightarrow \frac{1}{2}\text{dimer}_{\text{p-hydroxyphenyl+H}} + \frac{1}{2}\text{dimer}_{\text{guaiacyl+H}} + \text{ASA}_{\text{capped ads frags on surf}} \tag{5.4}$$

This reaction was found to be energetically favorable on the studied surface with a ΔE of -29.3 kcal/mol. The optimized geometry is given below in Figure 5.6. Though not shown here, the same reaction without the hydrogens to cap the dangling bonds on the fragments was also calculated and found to be energetically unfavorable. However, in the reaction conditions of the catalytic degradation, H_2 is many times present in excess.



(a) (b)

Figure 5.6. (a) The full B3LYP/6-31G* optimized structure of the dimer, with the β–O–4 ether bond cleaved and capped with hydrogens, adsorbed onto the simulated ASA surface. (b) A close-up view of the interactions of the dimer fragments with the surface showing the hydrogen bond and aluminum to oxygen interaction distances.

The desorption energies of both fragments were also calculated and were found to be energetically favorable when compared to the energies of the free dimer, free catalyst and free H_2, by -61.79 kcal/mol, suggesting ASA is a good catalyst for the process of β–O–4 bond cleavage in lignin degradation. When compared to the adsorbed dimers, the desorption products were found to be unfavorable by 23.06 kcal/mol for one fragment desorbed and by 22.24 kcal/mol for both

fragments removed. However, this is likely due to the stabilization that their hydrogen bonds provide on the surface. Yet, when the solvation effects are taken into account, this desorption process should become energetically favorable. One way to test this would be to run these same calculations using the so-called "conductor-like screening model" or COSMO to describe the dielectric screening effects in solvents.[105]

As mentioned earlier, the end goal of this study is to provide useful information to the experimentalist to aid in the rational design and modification of catalysts for lignin degradation. One way in which this can be done is to construct a volcano plot, which uses the Sabatier principle, from the calculated binding energies. A pedagogical illustration of a volcano plot is given below in Figure 5.7. This plot can be constructed by substituting various metals into the surface of our ASA model and find how that affects the binding energy. The volcano plot could then be used by experimentalists who could then dope the surface of ASA with the specific metals that allow for optimal binding energies. Another way to generate a volcano plot that could be helpful for experimentalists is to use the faster semiempirical calculations to screen through other types of catalysts and plot each of those as well.

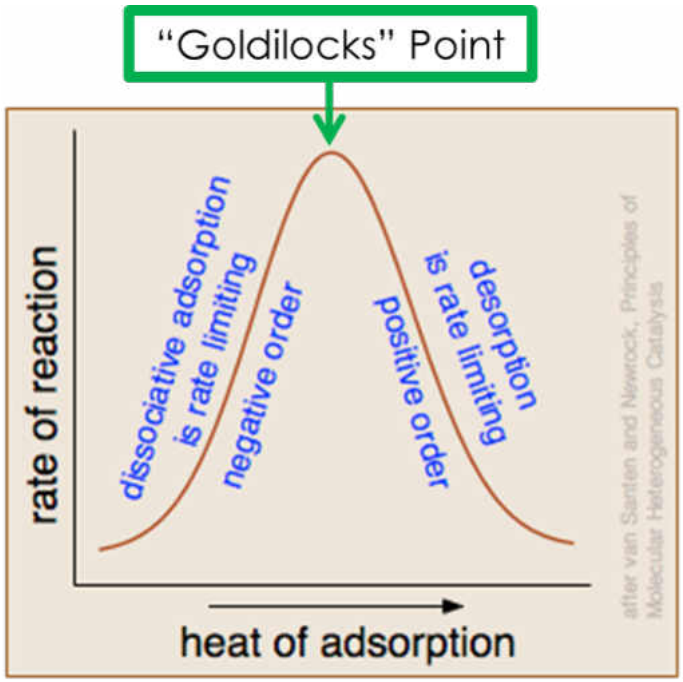


Figure 5.7. The principle of how a volcano plot works. The optimal catalyst is when the heat of adsorption (or binding energy) is "just right" so that the rate of reaction can be maximized.

5.4 Summary

Currently lignin is considered waste, but it has great econonimc potential if it could be broken down in an cost-effective way. It can be broken down more efficiently through catalysis. Amorphous silica–alumina (ASA) has been showing promise for this application, though the development of ASA as a catalyst for lignin degradation is limited by a lack of information on the mechanism of degradation as well as the effects of metal substitution. DFT has shown itself in the past to provide relevant information in similar cases. In this study the hybrid functional B3LYP with the 6-31G* basis set as implemented in NWChem has been used to study several systems involving the ASA catalyst as well as several fragments of lignin. Results have been obtained for the binding energies of a monomer and a dimer adsorbed onto the surface of ASA, for the reaction energy for breaking apart the dimer, and for the desorption of the resultant fragments of the dimer form the ASA surface. The binding energies for both the monomer and the dimer were found to be energetically favorable. The cleavage of the β–O–4 ether bond that linked the dimer together was found to be favorable when the dangling bonds were capped with hydrogen. Although it was found to be an energetically unfavorable process when compared to the bound cleaved and capped monomers, overall the desorption of the resultant monomers was found to be energetically favorable, suggesting ASA is a good catalyst for the degradation of lignin.

6 THEORETICAL INVESTIGATIONS OF THE ISOMERIZATION OF $Cu_2O_2^{2+}$

6.1 Introduction

The various bonding motifs of $Cu_2O_2^{2+}$ have drawn considerable attention over the years for several reasons. The moiety is known to play an important role in various metalloenzymes, e.g., oxy-tyrosinase, which oxidizes tyrosine residues to their corresponding *o*-quinones. In oxy-tyrosinase, which has imidazole rings from six histidine residues serving as ligands, the $Cu_2O_2^{2+}$ core is thought to rapidly interconvert between the bis(μ-oxo) and side-on μ-η^2:η^2-peroxo isomers,[106–109] both of which are shown in Figure 6.1. Understanding this rapid equilibrium then becomes crucial in the study of the oxidation mechanism of tyrosine. Therefore, many computational studies have been devoted to attempting to better understand this interconversion.[3,4,22,110–114] This then brings us to the other key reason for the considerable attention this system has received.

During the isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo there is a rapid change in both dynamic and nondynamic electron correlation while possessing a variable amount of birad-

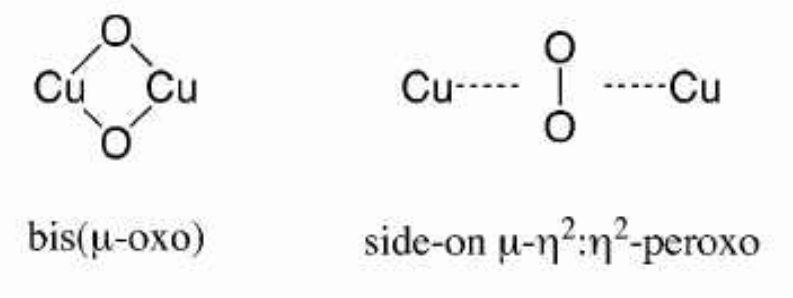


Figure 6.1. The most common bonding motifs of Cu_2O_2.[3]

ical character. These properties are very difficult for any method to adequately model. Thus, this system has become a sort of benchmark "torture track" for various theoretical models.[3] Though some methods, e.g. CR-CC(2,3),[23,24] have shown that they can do a good job describing this system, multireference perturbation theories, e.g., CASPT2,[25] have met with serious challenges.

The difficulties that perturbation theories have had with this system in the past comes, in part, due to the difficulty of forming a physically reasonable active space that captures all the important chemistry that happens along the isomerization coordinate. Along with that difficulty, a proper perturbative treatment of the electron correlation effects for transition metal complexes, in which there are many low-lying electronic states, is needed. GVVPT2 has shown itself to provide a high level of accuracy for difficult systems, and can be expected to similarly be successful with the $Cu_2O_2^{2+}$ core, while still enjoying favorable scaling (i.e., N^5). This could be hoped for because the method does not suffer from the well-known intruder state problem that vexes most of the other MRPTs which possess similar scaling. Therefore, armed with these three motives, it is clear that $Cu_2O_2^{2+}$ is an important system to study. In order to give a better understanding of what happens to several important orbitals as $Cu_2O_2^{2+}$ isomerizes, the B3LYP orbitals are given below, in Figure 6.2.

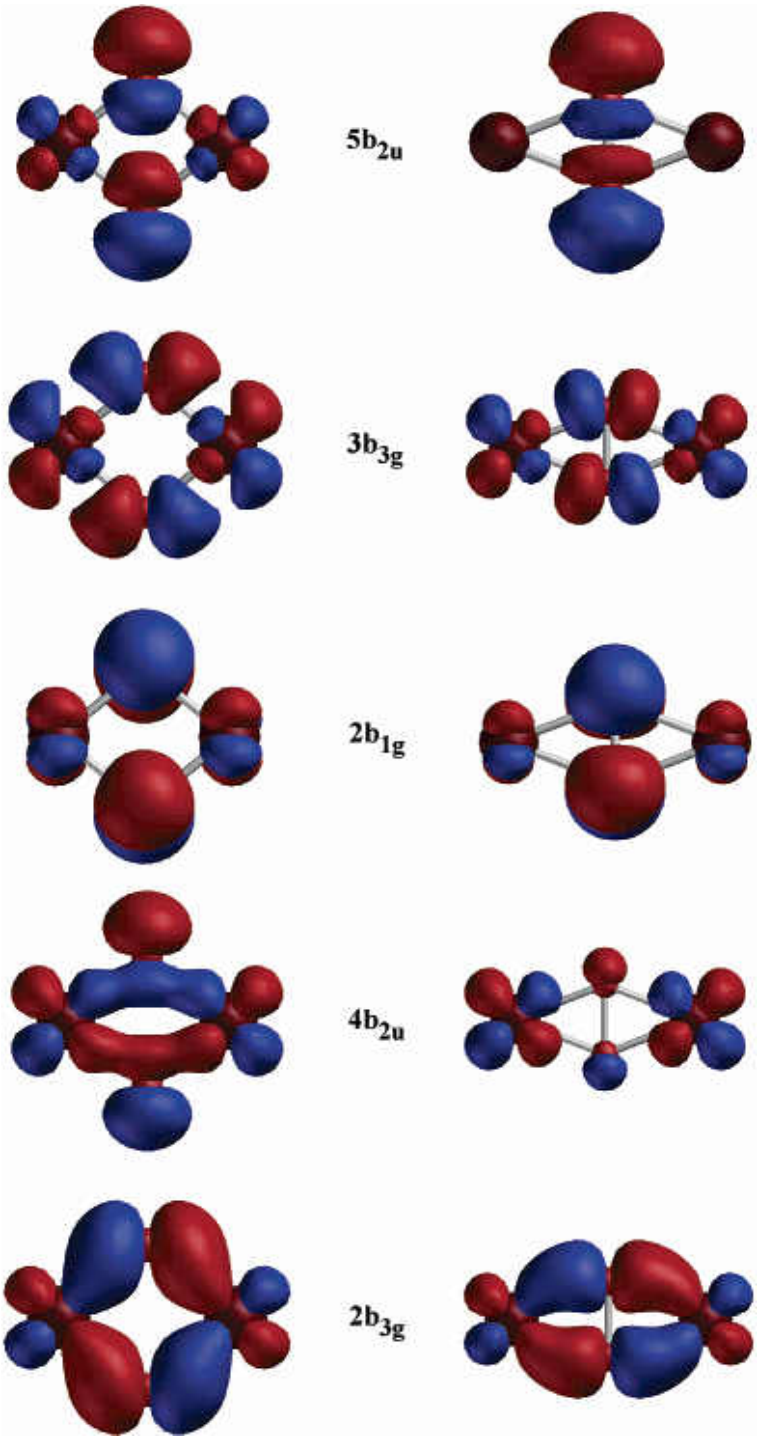


Figure 6.2. B3LYP orbitals for bis(μ-oxo) (left) and side-on μ-η^2:η^2-peroxo (right). This illustrates the nominal $2b_{1g}$ HOMO and $3b_{3g}$ LUMO orbitals at the B3LYP level.[3]

6.2 Computational Details

The geometries in this study have been obtained from the literature to match the various benchmark studies. For the bis(μ-oxo) geometry, the Cu–Cu and O–O distances in the D_{2h} core were taken as 2.8 and 2.3 Å, respectively. For the side-on μ-η^2:η^2-peroxo geometries the Cu–Cu and O–O distances in the D_{2h} core were taken as 3.6 and 1.4 Å, respectively. The isomerization between bis(μ-oxo) and side-on μ-η^2:η^2-peroxo will be tracked using six different points along the isomerization coordinate. The geometries can be smoothly changed with the use of

$$q_i(F) = q_i(0) + \frac{F}{100}[q_i(100) - q_i(0)] \tag{6.1}$$

where q_i is a given atomic Cartesian coordinate and F is the fraction of progress along the isomerization coordinate. Here F=0% corresponds to the bis(μ-oxo) isomer and F=100% corresponds to the side-on μ-η^2:η^2-peroxo isomer. Thus, the six points will be at F=0%, 20%, 40%, 60%, 80%, and 100%.

The correlation-consistent polarized valence triple-zeta (cc-pVTZ) basis set was used in all calculations.[75,115] The HF, B3LYP, MCSCF, and GVVPT2 single point energy calculations were performed using our local electronic structure program, known as UNDMOL (version 1.2).[14] For comparison, each of the methods used to compare relative energies also used the Stuttgart pseudopotential and associated basis functions for Cu[116] and the atomic natural orbital (ANO) basis set from by Pierloot et al.[117] for oxygen, contracted as [10s6p3d|4s3p2d]. This was simply denoted at basis set 1 (BS1).

6.3 Active Space Development

In this section the most important active spaces used in this study will be presented, the logic behind their development will be presented, and the rationale for going from one active space to the next will be given. When this study was in its infancy, a Hartree–Fock (HF) MO picture was used to develop several active spaces for the isomerization of $Cu_2O_2^{2+}$. These HF-type active spaces were originally chosen because of the previous successes our group had demonstrated in describing organic compounds, other systems made up of main-group elements,[58,80,118] as well as for my own studies presented earlier in Chapters 3 and 4. These types of active spaces typically include single and double excitations between two or more orbital subgroups. However, we found that the standard logic for constructing these active spaces could not be applied to the $Cu_2O_2^{2+}$ system. These active spaces either (i) exhibited serious convergence issues in the MCSCF or (ii) gave nonsensical results when the calculations did converge. Therefore this project had to be benched until we could find a new approach to generate orbitals and active spaces for GVVPT2 to use, thinking MCSCF could not be used for this purpose.

6.3.1 Generalized Valence Bond-Derived Active Space

Several years later, there was a new development in our group by Dr. Patrick Tamukong, who showed that a small valence picture active space can capture the important chemistry of transition metals.[26,27] The development of these generalized valence bond (GVB)-derived active spaces is based on valence bond theory in which chemical bonding results from overlap of atomic orbitals

which may be hybridized prior to bonding in order to satisfy geometric constraints. It is assumed that the most important contributors to bonding are the MOs derived from the valence shells of the participating atoms. Although atomic sub-shells just above or below the valence may contribute, such contributions are generally minimal. These MOs were often further partitioned into subspaces in which bonding and antibonding orbitals of the same kind are in the same subspace. The last step was to correlate the near frontier MOs at the GVVPT2 level, when their energies were close enough to the frontier MO energies. He was able to show that GVB-derived active space can successfully overcome both of the above mentioned issues with MCSCF for transition metals for both ground- and low-lying-excited electronic states.

Therefore, we decided to revisit the $Cu_2O_2^{2+}$ isomerization system using Patrick's GVB-derived active space in an attempt to adequately describe the bonding for this system. This active space was developed by Dr. Patrick Tamukong following the above mentioned logic applied to a system comprised of a Cu_2 moiety with oxygen atoms added to either side, Cu_2 + 2O. The structure had Cu–Cu and O–O distances of 1.78 and 3.08 Å, respectively with the same D_{2h} symmetry as the $Cu_2O_2^{2+}$ core. Several preliminary calculations were conducted to test this active space as well as energy difference of various electronic states. The MCSCF orbitals were obtained followed by GVVPT2 energies; the absolute energies along with the relative energies are given below, in Table 6.1. Based on the energies presented here, it was decided to run the rest of the calculations presented in this chapter with the irreducible representation of singlet A_g.

Following the success of these preliminary calculations, I applied this active space to the various geometries of $Cu_2O_2^{2+}$ along its isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo. The details of this active space, denoted as active space-1 (AS-1), is given below, in Table 6.2. Single and double excitations from valence group 1 (G_1) to valence group 2 (G_2) were considered.

Table 6.1. MCSCF and GVVTP2 total energies (a.u.) for Cu_2 + 2O along with the GVB-derived active space (22e,14o).

State	MCSCF/ cc-pVTZ	MCSCF Relative E	GVVPT2/ cc-pVTZ	GVVPT2 Relative E
Singlet A_g	-3427.554 114	0.000	-3428.965 341	0.000
Singlet B_{1u}	-3427.499 892	0.054	-3428.793 463	0.172
Triplet A_g	-3359.172 917	68.381	-3374.174 633	54.791
Triplet A_u	-3355.118 318	72.436	-3373.557 394	55.408
Triplet B_{1u}	-3358.837 335	68.717	-3376.981 104	51.984
Triplet B_{2u}	-3358.456 981	69.097	-3374.661 673	54.304
Triplet B_{3u}	-3360.155 218	67.399	-3374.475 816	54.900

To accomplish this, three macroconfigurations were used, denoted as (number of electrons in G_1 | number of electrons in G_2): (12 | 8), (11 | 9), and (10 | 10). For the subsequent GVVPT2 calculations based on orbitals generated from MCSCF, six of the high-lying frozen core orbitals were correlated by being placed in an active core group. This group consisted of two a_g orbitals, one b_{1g} orbital, two b_{2u} orbitals, and one b_{3u} orbital.

Table 6.2. GVB-derived active space applied to the isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo, (20e,14o).

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	8	1	2	2	0	5	5	3	26 total orbitals
Valence Group 1	2	0	0	1	1	1	0	1	6 total orbitals
Valence Group 2	1	1	1	1	0	3	1	0	8 total orbitals

Using this approach we were able to get converged MCSCF optimized $Cu_2O_2^{2+}$ MOs for GVVPT2 to use at a few geometries moving from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo. However, as I approached side-on μ-η^2:η^2-peroxo (i.e., F $\geq$ 70, which is the biradical region of the $Cu_2O_2^{2+}$ isomerization), the MCSCF began to show serious convergence issues once again. Therefore, though the GVB-derived active space gave a much better start then the HF-type active space our

group had used before, AS-1 had to go through significant modification before getting an active space that worked well for all geometries. It is worth noting that it was because of the partial success of the GVB-derived active space that we were able to develop the approach which proved successful for the entire $Cu_2O_2^{2+}$ isomerization. This was accomplished through numerical considerations from the MCSCF output which now provided meaningful values at several geometries due to the successful convergence. The rationale for this numerically based active space development technique was first introduced in a more generalized way in Chapter 2 and is given below in more detail as applied to $Cu_2O_2^{2+}$.

6.3.2 Numerically Based Active Space Development

As mentioned above, AS-1 had to go through significant modification. This was accomplished through numerical considerations in an iterative way based on the output of converged MCSCF calculations rather than based fully on chemical intuition. This way of active space development had not been used in our group prior and it yielded an active space that was somewhere between the originally attempted HF-type active space (which was shown to work well for main-block elements)[58,80,118] and the GVB-derived active space (which was shown to work well for transition metals).[26,27] This makes sense being that the chemistry behind the $Cu_2O_2^{2+}$ isomerization is driven by both transition metal and main-block elements and needed to be described in such a way that adequately accounted for both.

The iterative process of active space improvement started from AS-1. However, it did prove to be a challenge to get converged MCSCF results for even a few geometries with AS-1.

Thankfully there were enough points to give us an idea of what was going wrong with the MCSCF convergence. This analysis involved studying the MCSCF quasi-canonical orbital energies, the one-particle density matrix, and both the electron configurations that contribute to the total wave function and their respective amplitudes. After a few iterations of this, we were able to develop and execute several general rules to guide our modifications to the active space. These rules were given above in Chapter 2, but are given again here for the readers convenience. As our active spaces improved and MCSCF calculations were able to converge across the entire isomerization, any new modification to the active space needed to meet these criterion for all geometries. These rules are as follows: (i) if the quasi-canonical orbital energy was below -1.0 Ha and it remained doubly occupied for all CSF amplitudes greater than 0.03 the orbital would be moved to the frozen core, (ii) if the orbital energy was between -1.0 Ha and 0.0 Ha the orbital would be moved into the "mostly occupied" valence group 1 (G_1), (iii) if the orbital energy was between 0.0 Ha and about 1.0 Ha and there was at least one configuration that showed occupation with an amplitude of greater than 0.03 the orbital would be moved into "mostly unoccupied" valence group 2 (G_2), and finally (iv) if the orbital energy was above 1.0 Ha and there was no configuration that showed occupation with an amplitude of greater than 0.03 the orbital would be moved out of the active space and it would be treated as a virtual orbital. One last step that only applies to the GVVPT2 level was to take any frozen core orbitals that had a quasi-canonical orbital energy between -2.0 Ha and -1.0 Ha and put them into the active core group.

There are a few minor exceptions to these rules that are system dependent that won't be discussed here, but one important exception for our study was to rule (iii) and rule (iv). This exception was because $Cu_2O_2^{2+}$ is a cation, so there were always a few negative energies for some of the lowest lying virtual orbitals, however when they were put into G_1, their energies grew past 1.0

Ha and showed no significant occupied configurations as expected for virtual orbitals. Therefore, several iterations were used to test the lowest lying virtual orbitals by placing them into the "mostly unoccupied" subspace and then taking them out again if they acted as a virtual orbitals.

It is worth emphasizing that though having a good active space is important for capturing the chemistry of the system in question, the allowed excitations between valence groups is also a crucial component. Since our code has the capability to use macroconfigurations, we are able to easily adjust the electron occupations in each valence group to allow for whatever excitation levels were of interest. Throughout these active space modifications, we used single and double (SD) excitations (like both the HF-type and GVB-derived of active spaces), SD along with triple (SDT) excitations, and SDT along with quadruple (SDTQ) excitations. We had no need to go further toward a complete active space (CAS) level of excitation, where all possible excitations are considered, since all the chemistry along the isomerization should be describable through SDTQs. This saves a lot of computational effort and is one of the advantages that the flexibility of macroconfigurations provides us.

6.3.2.1 Active Space Iterations

After we determined that AS-1 was not going to work across the entire isomerization from bis(μ-oxo) to side-on μ-η^2:η^2-peroxo, we needed to determine how to modify the active space in a way that worked across all geometries. As mentioned above, this was approached numerically and there were several iterations and tests of modifications to existing active spaces. Originally this was taken in two directions. One way was to trim down AS-1 as far as reasonable and slowly add in orbitals until a stable active space was reached. The other approach was to include all possible

reasonable orbitals, making a huge active space, and slowly chopping away until a stable active space was obtained. The latter method will be discussed first because only two active spaces were developed using this approach before the smaller active space approach showed more promise.

The first active space, denoted here as (for lack of a better name) big active space (AS-Big), was developed by looking at B3LYP output and taking the orbitals that had a noticeable energy gap from themselves to the next lowest occupied orbital for G_1 or next highest virtual orbital for G_2. In AS-Big both G_1 and G_2 were kept the same size. This active space was also used to test many of the virtual orbitals which had a negative quasi-canonical orbital energy in previous MCSCF outputs. The orbital partitioning is given below, in Table 6.3. AS-Big allowed SD excitations between valence groups by using three macroconfigurations, (30 | 0), (29 | 1), and (28 | 2). Though the active space was relatively large compared to AS-1, it converged relatively painlessly. GVVPT2 was not run using these orbitals, however, since there were many obvious changes to be made to the active space from the MCSCF output to make it smaller. These trims brought about the next active space.

Table 6.3. Orbital partitioning in AS-Big, (30e,30o). SD excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	7	0	2	2	0	5	3	2	21 total orbitals
Valence G_1	3	2	1	1	1	3	2	2	15 total orbitals
Valence G_2	3	0	2	2	0	3	3	2	15 total orbitals

The next active space, denoted here as reduced big active space (AS-RedBig), was made from AS-Big by removing the four lowest-lying orbitals from G_1 and the four highest-lying orbitals from G_2 and is presented in Table 6.4. Four orbitals were chosen on each side because there was another natural break in the B3LYP energies and we were still trying to keep the valence

groups balanced. While running several iterations of the smaller active spaces, presented below, we had realized the importance of including SDTQ excitations for capturing the chemistry of the isomerization. Since these larger active spaces were being developed and tested in parallel with the smaller active spaces, SDTQ excitations were attempted for this active space as well. This required five macroconfigurations, (22 | 0), (21 | 1), (20 | 2), (19 | 3), and (18 | 4). Though the MCSCF converged in only 13 iterations, it took over 100 hours, while the smaller SDTQ active spaces would typically finish in a few hours. Therefore, this reason along the fact that the quasi-canonical orbital energies and MCSCF configurations were stabilizing quickly for the smaller active spaces, the larger active spaces were abandoned for the smaller which are discussed next.

Table 6.4. Orbital partitioning in AS-RedBig, (22e,22o). SDTQ excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	8	0	2	2	0	6	4	3	25 total orbitals
Valence G_1	2	2	1	1	1	2	1	1	11 total orbitals
Valence G_2	3	0	1	2	0	2	2	1	11 total orbitals

As mentioned above, there were two directions of active space development that took place after AS-1 failed to converge at all geometries. The large active spaces have been presented above, now the smaller active spaces will be discussed. The first of which, denoted as active space-2 (AS-2), was obtained from the MCSCF output from AS-1 at three points along the isomerization coordinate, bis(μ-oxo) (F=0%), F=70%, and side-on μ-η^2:η^2-peroxo (F=100%). F=70% was used because that was where there appeared to be some sort of potential curve crossing while using AS-1 and we could only get "mostly converged" orbitals, though they were good enough to serve in the analysis of the MCSCF output for the necessary modifications. The orbital partitioning in AS-2 is given below, in Table 6.5. Two orbitals, $6b_{1u}$ and $3b_{2g}$, were consistently doubly occupied in the

MCSCF configurations and had values of -1.0 or lower in their energies, so they were moved from G_1 and G_2, respectively, to the frozen core. Two orbitals, $1b_{1g}$ and $5b_{2u}$, had energies between -1.0 and 0.0 at one or more geometries, so they were moved from the frozen core to G_1. One orbital, $11a_g$ showed no occupation in the MCSCF configuration in this active space, so it was moved from G_2 to virtual space. AS-2 allowed for SD excitations between valence groups by using three macroconfigurations, (14 | 6), (13 | 7), and (12 | 8). AS-2 was unable to give convergent MCSCF behavior, therefore, more modification was necessary.

Table 6.5. Orbital partitioning in AS-2, (20e,13o). SD excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	8	0	3	2	0	6	4	3	26 total orbitals
Valence G_1	2	1	0	1	1	0	0	1	7 total orbitals
Valence G_2	0	1	0	1	0	3	1	0	6 total orbitals

To arrive at the next active space, active space-3 (AS-3), there were just two modifications taken. The $11a_g$ orbital was brought back from the virtual space into G_2 because of a re-evaluation of the quasi-canonical energies at various geometries, and it was found to be important for (F=100%). The other change was bringing the $7b_{1u}$ down from G_2 to G_1 because of its negative energy for all geometries analyzed. The orbital partitioning in AS-3 is given below, in Table 6.6. It was with this active space that we decided that at least triple excitations should be allowed when generating orbitals for GVVPT2, and it is possible that a converged MCSCF may not be obtainable with a small active space without considering higher than SD excitations. The SDTs were obtained by using four macroconfigurations, (16 | 4), (15 | 5), (14 | 6), and (13 | 7). As was the case for AS-2, AS-3 was unable to give convergent MCSCF behavior for all geometries, therefore, more modifications were necessary before GVVPT2 calculations could be conducted.

Table 6.6. Orbital partitioning in AS-3, (20e,14o). SDT excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	8	0	3	2	0	6	4	3	26 total orbitals
Valence G_1	2	1	0	1	1	1	0	1	8 total orbitals
Valence G_2	1	1	0	1	0	2	1	0	6 total orbitals

This next active space, active space-4 (AS-4), was the first moderately successful active space. We changed the structure of our active spaces here and set up the two valence groups in a similar way to the HF-type of active space, where G_1 contained the high-lying mostly occupied orbitals and G_2 contained the low-lying mostly unoccupied active space. The rationale for this was taken from the larger active spaces presented earlier, since they were run in parallel with these smaller active spaces and were exhibiting better MCSCF convergence behavior. As can be seen above, in AS-2 there were 34 orbitals in the combined frozen core and G_1 and a total of 72 electrons in the $Cu_2O_2^{2+}$ system. Therefore, two more orbitals needed to be moved into the frozen core or G_1 from G_2 or the virtual space in order to have room for all 72 electrons and were decided on based on the lowest energy orbitals. The two orbitals, $8b_{1u}$ and $2b_{1g}$, were moved from G_2 to G_1. Besides this main change, there were also a other few modifications. The $7b_{1u}$ orbital was moved from G_1 to the frozen core because its energy had dropped below -2.0. There were also three other relatively low-lying virtual orbitals, $12a_g$, $5b_{3u}$, and $4b_{2g}$, that were added to help balance out G_1 and G_2. The orbital partitioning in AS-4 is given below, in Table 6.7. Due to convergence issues for both AS-2 and AS-3, the modifications mentioned here were based off of unconverged MCSCF output after 20 iterations. AS-4 also considered SDT excitations by using four macroconfigurations, (18 | 0), (17 | 1), (16 | 2), and (15 | 3). This was the first active space that gave converged MCSCF results at all geometries while maintaining a smooth

relative energy throughout the isomerization. Therefore, GVVPT2 calculations were conducted using these orbitals, and the energies are presented in section 6.4. As was the case when using AS-1, the subsequent GVVPT2 calculations based on orbitals generated from MCSCF had seven of the high-lying frozen core orbitals correlated by being placed in an active core group. This group consisted of two a_g orbitals, one b_{2g} orbital, two b_{1u} orbitals, one b_{2u} orbital, and one b_{3u} orbital.

Table 6.7. Orbital partitioning in AS-4, (18e, 16o). SDT excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	8	0	3	2	0	7	4	3	27 total orbitals
Valence G_1	2	2	0	1	1	1	1	1	9 total orbitals
Valence G_2	2	0	1	1	0	1	1	1	7 total orbitals

Though AS-4 was able to give converged results for all points along the $Cu_2O_2^{2+}$ isomerization, the shape of the GVVPT2 relative energy curve possessed some strange features when compared with higher level methods. These features are discussed more in section 6.4. Along with this, the MCSCF output still gave more numeric clues as to how to improve the active space. The $9a_g$ orbital energy was around -1.8, so it was moved from G_1 to the frozen core. Two more low-lying virtual orbitals, $13a_g$ and $10b_{1u}$, were moved into G_2. The orbital partitioning in AS-5 is given below, in Table 6.8. AS-5 considered both SDT and SDTQ excitations by using four macroconfigurations for the SDTs, (16 | 0), (15 | 1), (14 | 2), and (13 | 3) and five macroconfigurations for the SDTQs, (16 | 0), (15 | 1), (14 | 2), (13 | 3), and (12 | 4). It was during the development of AS-5 that we decided to test the effect of SDT vs. SDTQ excitations on the numerics in the MCSCF output and on the relative energy curve of the $Cu_2O_2^{2+}$ isomerization. The latter effect was particularly interesting to us, because of the strange features mentioned earlier. It should be mentioned that we had planned on doing SDTQ excitations all along for the final calculations once

a reasonable active space was found, since two bonds are broken and formed along the isomerization coordinate, thus needing four electrons to be allowed to change occupation between valence groups. However, SDTQ excitations are relatively expensive compared to SD and even SDT excitations. So we were attempting to find a good active space using the faster SDT excitations and only conduction SDTQ excitations once it was found. AS-5 was the first active space that SDTQ excitations were attempted on all geometries at both the MCSCF and GVVPT2 levels. It was found that including quadruple excitations drastically changed the relative GVVPT2 energy, and this is discussed in section 6.4. The GVVPT2 calculations again correlated eight of the high-lying frozen core orbitals. This active core group consisted of three a_g orbitals, one b_{2g} orbital, two b_{1u} orbitals, one b_{2u} orbital, and one b_{3u} orbital. Though AS-5 was stable enough to test quadruple excitations, there were still some modifications that could be made to trim it down.

Table 6.8. Orbital partitioning in AS-5, (16e, 17o). Both SDT and SDTQ excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	9	0	3	2	0	7	4	3	28 total orbitals
Valence G_1	1	2	0	1	1	1	1	1	8 total orbitals
Valence G_2	3	0	1	1	0	2	1	1	9 total orbitals

This most recent modification to the active space, active space-6 (AS-6), was obtained by considering the SDTQ MCSCF output. It was found that both $13a_g$ and $4b_{2g}$ could be safely moved into the virtual space from G_2 due to large positive energies. The orbital $4b_{3g}$ was moved from G_2 to G_1 because of a consistently negative energy across all geometries. And finally, both $1b_{1g}$ and $1a_u$ were moved from G_1 to the frozen core because of energies of around -1.0 for all geometries. This trimming significantly decreased the size of the active space, yet the relative energies are very comparable to the larger AS-5 as presented in section 6.4. The orbital partitioning in AS-6 is given

below, in Table 6.9. AS-6 considered only SDTQ excitations, since it was smaller than AS-5 and quadruple were shown to greatly effect the energies. The SDTQ excitations were obtained by using five macroconfigurations, (12 | 0), (11 | 1), (10 | 2), (9 | 3), and (8 | 4). It is worth noting that not all of G_1 is doubly occupied in AS-6, as was the case for AS-4 and AS-5 (i.e., there are two holes present in G_1 in AS-6). For GVVPT2, 10 high-lying frozen core orbitals were correlated. This active core group consisted of three a_g orbitals, one b_{1g} orbital, one b_{2g} orbital, one a_u orbital, two b_{1u} orbitals, one b_{2u} orbital, and one b_{3u} orbital. AS-6 is the active space presented below as the "best-so-far" active space. We have several ideas on how to improve the active space to get even better agreement between GVVPT2 and the higher level methods given in Reference 3, but that is outside the scope of this dissertation.

Table 6.9. Orbital partitioning in AS-6, (12e, 13o). Only SDTQ excitations were considered with this active space.

Symmetry	a_g	b_{1g}	b_{2g}	b_{3g}	a_u	b_{1u}	b_{2u}	b_{3u}	
Frozen Core	9	1	3	2	1	7	4	3	30 total orbitals
Valence G_1	1	1	0	2	0	1	1	1	7 total orbitals
Valence G_2	2	0	0	0	0	2	1	1	6 total orbitals

6.4 Relative Energies of the $Cu_2O_2^{2+}$ Isomerization

As mentioned above, adjusting the active space and accounting for quadruple excitations had a significant effect on the relative energies. This is particularly true for the inclusion of quadruples, which can be seen particularly well in the data presented below. However it should be noted that without a suitable active space, the energies cannot even be obtained for the entire isomer-

ization coordinate with MCSCF. Therefore, only the most refined active spaces energies will be presented here, AS-4, AS-5, and AS-6.

6.4.1 Active Space Effect

The first comparison is showing the effect on the relative energies that occurred as we adjusted AS-4 to get AS-5 and is presented graphically below, in Figure 6.3. Both active spaces here only considered SDT excitations.

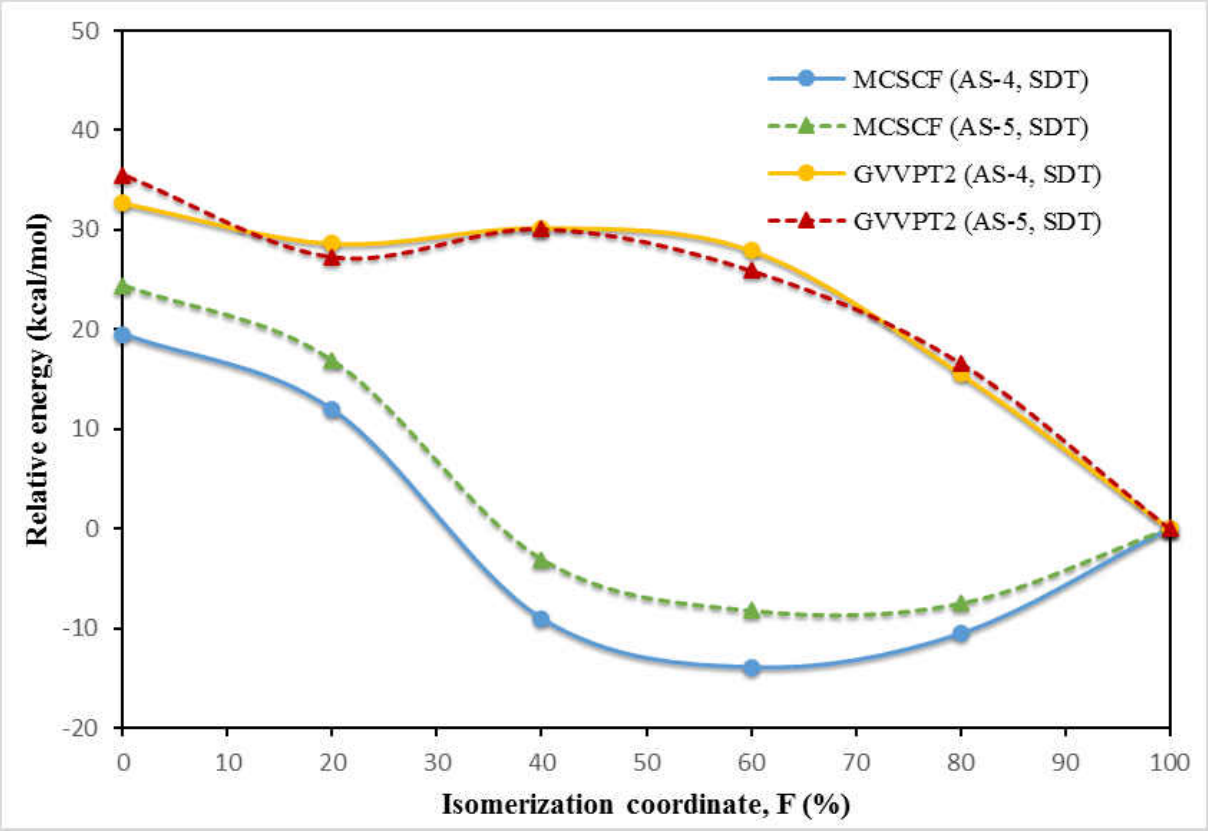


Figure 6.3. Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing how the energy changed when AS-4 (solid lines) was modified to get AS-5 (dashed lines). Both active spaces included SDT excitations.

It can be seen that although MCSCF had a significant change, GVVPT2 remained fairly

stable to the change. One interesting feature of these calculations was the large dip in the MCSCF energies starting from F=20% to F=40% for both active spaces. This was then accompanied by a large bump in the corresponding GVVPT2 energies also starting from F=20% to F=40% for both active spaces. This is likely due to GVVPT2 attempting to correct for MCSCF's odd behavior. This behavior was later found out to be from the lack of inclusion of quadruple excitations.

The comparison of AS-5 and AS-6 is given below, in Figure 6.4, where both included quadruple excitations. Again GVVPT2 was shown to be fairly stable to the active space modification, while MCSCF exhibited more change. Here again MCSCF incorrectly predicts the minimum, only now it is closer to the true minimum (i.e., F=100%),[3] at around F=70%–F=80%. This is likely because this is the biradical region of the isomerization of $Cu_2O_2^{2+}$, and MCSCF is known to struggle in such regions with over stabilization.[119]

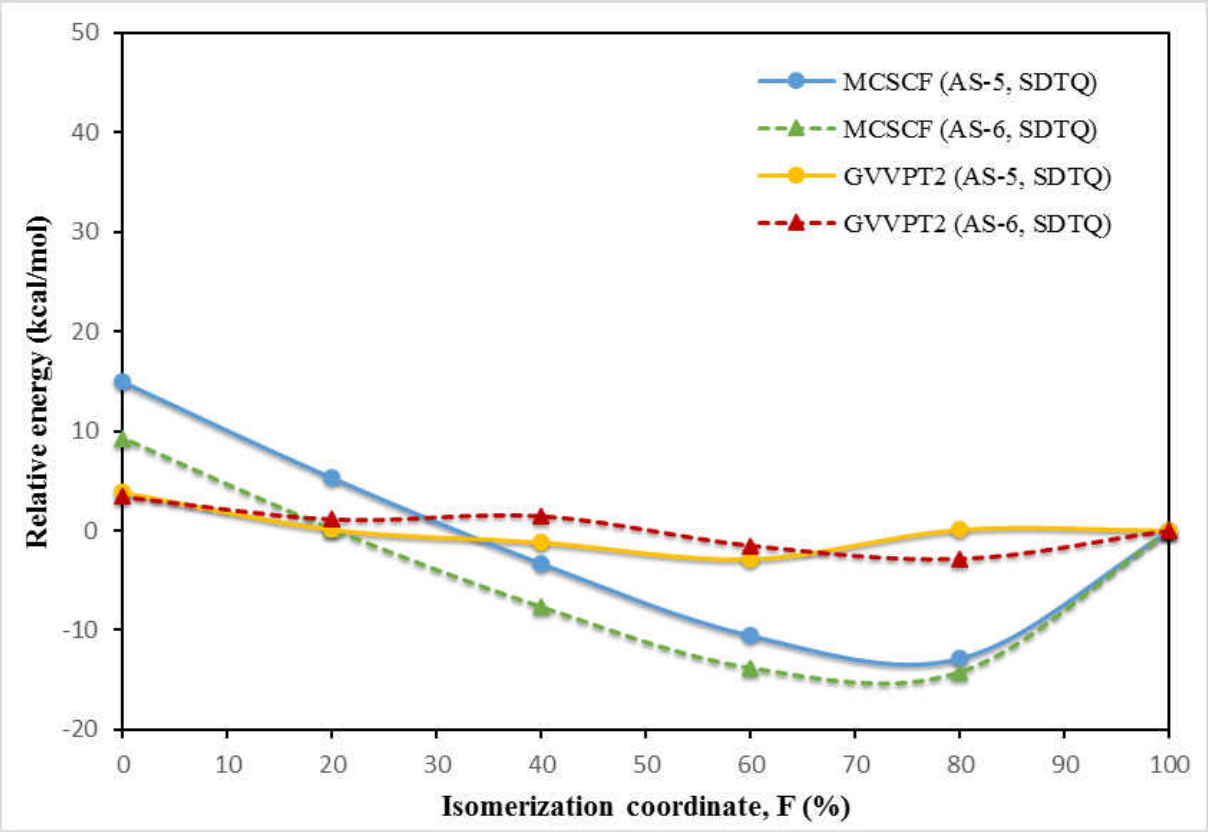


Figure 6.4. Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing how the energy changed when AS-5 (solid lines) was modified to get AS-6 (dashed lines). Both active spaces included SDTQ excitations.

6.4.2 Effect of Including Quadruple Excitations

When quadruple excitations are considered, the relative energy landscape of the isomerization of $Cu_2O_2^{2+}$ drastically changes. This is due to the nature of the isomerization where two bonds are simultaneously formed/broken, thus involving four electrons. However, calculations involving quadruple excitations can take considerably longer than when only up to triple excitations are accounted for. Therefore, we tested the convergence and stability of various active spaces using SD and SDT excitations before we decided to do SDTQ calculations. As can be seen above, AS-4 and AS-5 both converged for the MCSCF and were relatively stable to modification, so we decided to consider quadruple excitations starting with AS-5. The difference in the absolute energies at various geometries (i.e., for given value of F) were taken at both the MCSCF and GVVPT2 levels of theory and are presented below, in Table 6.10, and in Figure 6.5. It is worth clarifying that these energies are not set equal at side-on μ-η^2:η^2-peroxo (F=100%), but are rather a comparison of the absolute energy (given in Hartree) for each value of F individually (e.g., $E_{MCSCF(AS-5,SDT)}$ at F=0% compared to $E_{MCSCF(AS-5,SDTQ)}$ also at F=0%).

Table 6.10. Absolute energies (Hartree) and the difference at each given point (kcal/mol) along the isomerization of $Cu_2O_2^{2+}$ showing the how the energies changed when quadruple excitations were add to AS-5. Relative energies were calculated at each point (e.g., at F= 100%) as $E_{rel} = E_{SDTQ} - E_{SDT}$.

	Isomerization Coordinate, F					
Level of Theory	0%	20%	40%	60%	80%	100%
MCSCF (AS-5, SDT), Hartree	-3427.016 453	-3427.028 445	-3427.060 230	-3427.068 433	-3427.067 323	-3427.055 318
MCSCF (AS-5, SDTQ), Hartree	-3427.064 346	-3427.079 715	-3427.093 424	-3427.104 963	-3427.108 591	-3427.088 185
MCSCF Relative Energy, kcal/mol	-30.1	-32.2	-20.8	-22.9	-25.9	-20.6
GVVPT2 (AS-5, SDT), Hartree	-3428.281 999	-3428.295 148	-3428.290674	-3428.297 315	-3428.312 097	-3428.338 618
GVVPT2 (AS-5, SDTQ), Hartree	-3428.269 313	-3428.275 295	-3428.277 342	-3428.280 026	-3428.275 295	-3428.275 446
GVVPT2 Relative Energy, kcal/mol	8.0	12.5	8.4	10.8	23.1	39.6

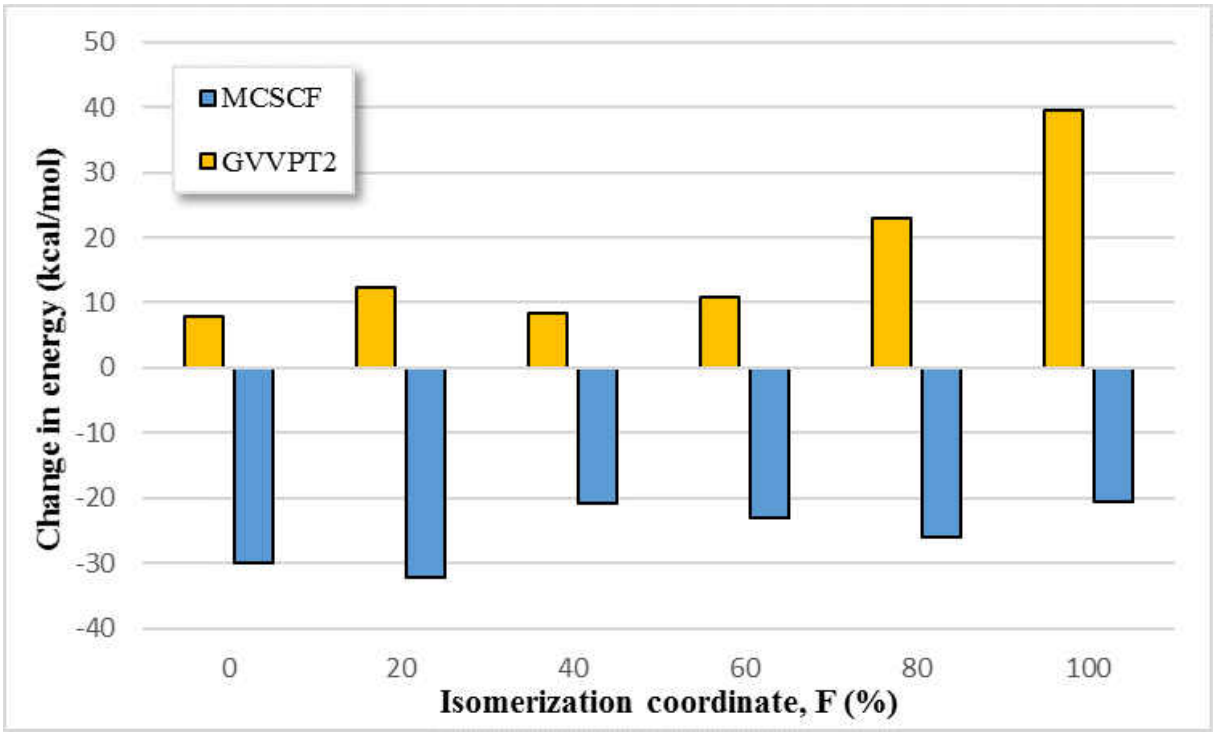


Figure 6.5. Bar chart showing how the difference in the energies (kcal/mol) change at different points along the isomerization of $Cu_2O_2^{2+}$ for both MCSCF and GVVPT2. The MCSCF and GVVPT2 levels of theory showing the effects of adding quadruple excitations (i.e., SDTQ [dashed lines] vs. SDT [solid lines]) when using AS-5.

Considering quadruple excitations at the MCSCF level was found to stabilize all points along the isomerization. However, for geometries near bis(μ-oxo) (i.e., F=0% and 20%), they were stabilized more than the rest of the geometries by about 10 kcal/mol. Considering quadruple excitations at the GVVPT2 level was found to destabilize all points along the isomerization. However, for geometries near side-on μ-η^2:η^2-peroxo (i.e., F=100% and 80%), they were destabilized more than the rest of the geometries by about 10 kcal/mol for F=80% and nearly 20 kcal/mol for F=100%. This massive destabilization near side-on μ-η^2:η^2-peroxo is what is primarily responsible for the fairly isoenergetic GVVPT2 isomerization curve.

While the above Table 6.10 and Figure 6.5 compare each value of F, Figure 6.6, given be-

low, compares their relative energies by setting both energies of side-on μ-η^2:η^2-peroxo (F=100%) equal to zero, as was the convention before and will be for the rest of this chapter.

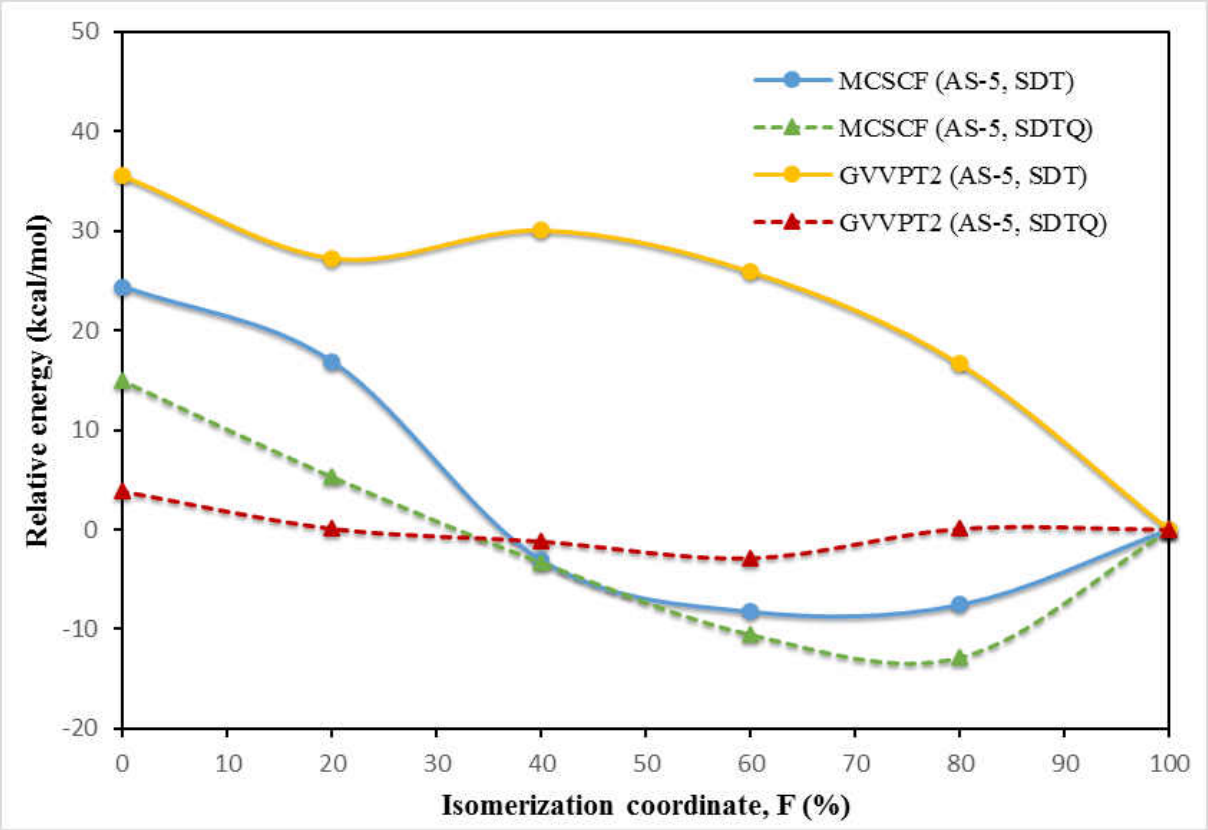


Figure 6.6. Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory showing the effects of adding quadruple excitations (i.e., SDTQ [dashed lines] vs. SDT [solid lines]) when using AS-5.

6.4.3 Comparison of GVVPT2 Results to Previous Studies

Now that we have a converged and reasonably stable active space (AS-6), we will compare our results with previous theoretical studies of this system. As mention in the introduction, this $Cu_2O_2^{2+}$ isomerization is a kind of benchmark system. Thus far the best estimates of the actual relative energy along the isomerization from bis(μ-oxo) (i.e., F=0%) to side-on μ-η^2:η^2-peroxo (i.e.,

F=100%) has been predicted by a massive calculation of CR-CCSD(TQ)$_L$, which was computed as CR-CCSD(T)$_L$ + CR-CCSD(TQ) – CR-CCSD(T).[3] Another great estimate was made a few years later using second order perturbation theory using a exceedingly large restricted active space (RASPT2) of (24e,28o). They also that allowed for SDTQ excitations between the valence RAS groups (denoted by the authors by "//4", i.e., RASPT2(24e,28o)//4.[4] Both of these methods used BS1, mentioned above in section 6.2.

Two other relatively expensive methods that were used on this system was the complete active space self consistent field (CASSCF) method and the complete active space second order perturbation (CASPT2) method, which is applied to the CASSCF reference function in a similar way that GVVPT2 is applied to the MCSCF reference function. Before CASSCF and CASPT2 were applied to this $Cu_2O_2^{2+}$ system, it was widely used for quantitative modeling of transition metal chemistry.[120,121,121] Similarly to RASSCF and RASPT2 (or MCSCF and GVVPT2 in our study), the quality of the results of CASSCF and CASPT2 depends greatly on the choice of active space. The primary difference of CAS from RAS is that it must consider all possible excitations in its active orbitals, making it relatively expensive and limits the size of active space that can be used. When applied to $Cu_2O_2^{2+}$, the authors had attempted several active spaces varying considerably in size, the smallest and largest of which are shown here for comparison, (8e,8o) and (16e,14o).[3] All of the calculations presented here for CASSCF and CASPT2 used BS1, mentioned above in section 6.2. The comparison of our MCSCF and GVVPT2 results using various active spaces and degree of excitations are given below, in Table 6.11. A select few isomerization relative energy curves are also presented graphically below, in Figure 6.7.

Though the results of both CR-CCSD(TQ)$_L$ and RASPT2(24e,28o)//4 are included, they are here for reference purposes only. The main point we are making at the current state in this study

is that GVVPT2 can reproduce large active space CASSCF/CASPT2 results using a smaller space with AS-6 (12e, 13o) and only considering up to quadruple excitations rather than considering all possible excitations. Future studies will determine how large of an active space will be necessary to have close agreement with the CR-CCSD(TQ)$_L$ results (in a similar way as RASPT2(24e,28o)//4 did).

Table 6.11. Comparisons of relative energies (kcal/mol) of various levels of theory along the isomerization path of $Cu_2O_2^{2+}$ compared to best estimates from CAS, RAS and CR-CC results. Energies were related by equating all the side-on μ-η^2:η^2-peroxo energies.

	Isomerization Coordinate, F					
Level of Theory	0%	20%	40%	60%	80%	100%
MCSCF (AS-4, SDT)	19.5	11.9	-9.0	-13.9	-10.6	0.0
MCSCF (AS-5, SDT)	24.4	16.9	-3.1	-8.2	-7.5	0.0
MCSCF (AS-5, SDTQ)	15.0	5.3	-3.3	-10.5	-12.8	0.0
MCSCF (AS-6, SDTQ)	9.3	0.1	-7.7	-13.8	-14.2	0.0
[a]CASSCF (8e, 8o)	-3.6	-11.1	-16.5	-19.5	-16.7	0.0
[a]CASSCF (16e, 14o)	0.2	-7.2	-12.7	-16.3	-14.0	0.0
GVVPT2 (AS-4, SDT)	32.7	28.6	30.2	28.0	15.6	0.0
GVVPT2 (AS-5, SDT)	35.5	27.3	30.1	25.9	16.6	0.0
GVVPT2 (AS-5, SDTQ)	3.9	0.1	-1.2	-2.9	0.1	0.0
GVVPT2 (AS-6, SDTQ)	3.4	1.1	1.5	-1.6	-2.9	0.0
[a]CASPT2 (8e, 8o)	6.5	0.4	-3.1	-5.8	-6.1	0.0
[a]CASPT2 (16e, 14o)	1.4	-2.7	-5.4	-7.6	-8.7	0.0
[b]RASPT2(24e,28o)//4	28.6	21.2	15.1	7.1	0.8	0.0
[c]CR-CCSD(TQ)$_L$	33.8	24.2	15.9	8.2	1.8	0.0

[a]All CAS results obtained from Ref. 3. [b]RASPT2(24e,28o)//4 (a two part RAS with a maximum of four excitations allowed between the two valence groups) results obtained from Ref. 4. [c]CR-CCSD(TQ)$_L$ (computed as CR-CCSD(T)$_L$ + CR-CCSD(TQ) – CR-CCSD(T)) results obtained from Ref. 3.

It can be seen that MCSCF curve agrees exceedingly well with CASSCF (16e,14o) results in the region near side-on μ-η^2:η^2-peroxo (i.e., F=100%) and reasonably well as the isomerization

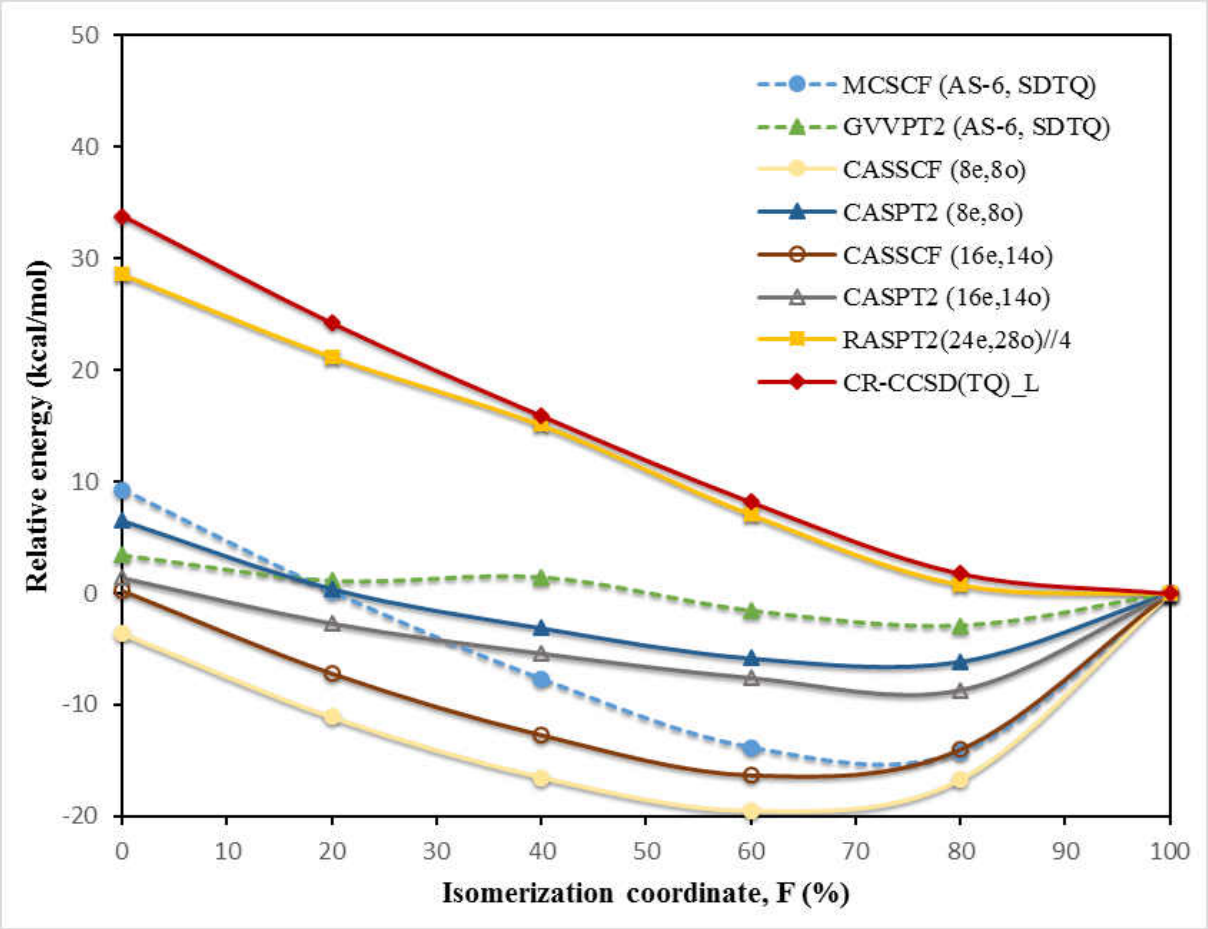


Figure 6.7. Relative energies (kcal/mol) of $Cu_2O_2^{2+}$ at the MCSCF and GVVPT2 levels of theory compared to previous studies (CR-CCSD(TQ)$_L$ values obtained from Ref. 3 and RASPT2(24e,28o)//4 values obtained from Ref. 4).

moves closer to bis(μ-oxo) (i.e., F=0%). GVVPT2 is also shown to agree reasonably well with both of the CASPT2 results. These agreements highlight the versatility of the macroconfiguration approach, in that we were able to specify that we only consider the important excitations, thus getting results comparable to much larger, time intensive calculations. This also shows the validity of the numerical approach of active space development, as described above in section 6.3.2. There are several ways in which GVVPT2 calculations can be expanded to increase agreement with CR-CCSD(TQ)$_L$ results, and they are described briefly below in section 6.5.

6.5 Future Work

There are several directions in which we plan on taking this study to improve our results. The most obvious of which is to attempt to imitate the large RAS active space of (24e,28o) with quadruple excitations from Ref. 4. One way in which we intend on doing that is to approach it starting from our converged MCSCF orbitals for the larger active spaces we have attempted (i.e., AS-Big and AS-RedBig). Another factor that needs investigating is the effect of using the smaller basis set that was used in the studies mentioned,[3,4] which had the Stuttgart pseudopotential and associated basis functions for Cu[116] and the atomic natural orbital (ANO) basis set from by Pierloot et al.[117] for oxygen, contracted as [10s6p3d|4s3p2d].

Another direction to take is based on numerical considerations of the MCSCF output using AS-6 with SDTQ excitations. There are still some valid modifications to be tested. Though this should not have a large effect on the relative energies, it is worth testing.

The fourth direction investigates the use of DFT orbitals, rather than MCSCF orbitals, as a reference for GVVPT2. This was shown to work well for several test systems[122] and may prove to work well with the isomerization of $Cu_2O_2^{2+}$. Once the ability for GVVPT2 to run using DFT generated orbitals is fully implemented into UNDMOL, we will also be able to study far larger systems than previously possible. In the case of the $Cu_2O_2^{2+}$ system, we would be able to pursue the more realistic system containing ammonia ligands, which has been shown to give a good approximation for the biological compounds that use this copper core.[22]

7 OVERALL CONCLUDING REMARKS

There are several key insights that were obtained from the research presented in this dissertation. The protocol presented for main group oxides seems to be fairly general. The single and double excitation HF-type active space works well for these systems and other main group oxides could be attempted with this type of active space. GVVPT2 was shown to work well in describing these multireference systems. For both As_xO_y (Chapter 3) and Sb_xO_y (Chapter 4) species were shown to have energetically favorable gas phase accretion.

The protocol for the catalytic degradation of lignin (Chapter 5) was able to describe the binding energy of both a monomer and a dimer (representing lignin) onto a model of an amorphous silica–alumina catalyst. This protocol could be attempted with other lignin fragments. The described cluster model could be extended to study other catalysts, provided it has a large enough band gap.

GVVPT2 was shown to be capable of describing the isomerization curve between the bis(μ-oxo) and side-on μ-η^2:η^2-peroxo isomers of the $Cu_2O_2^{2+}$ moiety (Chapter 6). We found that the combination of a GVB-derived and HF-type active space was needed in order to properly describe $Cu_2O_2^{2+}$. This type of active space should be extended to other metal oxo-bridged compounds.

REFERENCES

1. Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, revised 1st ed.; McGraw-Hill Publishing Co.: Singapore, 1986.

2. Yabuki, M.; Takahashi, R.; Sato, S.; Sodesawa, T.; Ogura, K. Silica–alumina catalysts prepared in sol–gel process of TEOS with organic additives. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4830–4837.

3. Cramer, C. J.; Włoch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical models on the Cu_2O_2 torture track: mechanistic implications for oxytyrosinase and small-molecule analogues. *J. Phys. Chem. A* **2006**, *110*, 1991–2004.

4. Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* **2008**, *128*, 204109.

5. Smith, S. J.; Sutcliffe, B. T. *The Development of Computational Chemistry in the United Kingdom*; pp 271–316.

6. Schaefer, H. F. Methylene: A Paradigm for Computational Quantum Chemistry. *Science (80).* **1986**, *231*, 1100–1107.

7. Denning, P. J.; Lewis, T. G. Exponential laws of computing growth. *Commun. ACM* **2016**, *60*, 54–65.

8. Roos, B. O. In *Ab Initio Methods Quantum Chem.*; Lawley, K. P., Ed.; Wiley: New York, 1987; pp 399–445.

9. 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*, 2185–2192.

10. Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

11. Hoffmann, M. R. Canonical Van Vleck quasidegenerate perturbation theory with trigonometric variables. *J. Phys. Chem.* **1996**, *100*, 6125–6130.

12. Khait, Y. G.; Hoffmann, M. R. A self-consistent version of quasidegenerate perturbation theory. *J. Chem. Phys.* **1998**, *108*, 8317.

13. Khait, Y. G.; Song, J.; Hoffmann, M. R. Explication and revision of generalized Van Vleck perturbation theory for molecular electronic structure. *J. Chem. Phys.* **2002**, *117*, 4133–4145.

14. Jiang, W.; Khait, Y. G.; Hoffmann, M. R. Configuration-Driven Unitary Group Approach for Generalized Van Vleck Variant Multireference Perturbation Theory. *J. Phys. Chem. A* **2009**, *113*, 4374–4380.

15. Shah, P.; Strezov, V.; Stevanov, C.; Nelson, P. F. Speciation of arsenic and selenium in coal combustion products. *Energy & Fuels* **2007**, *21*, 506–512.

16. Jadhav, R. A.; Fan, L. S. Capture of gas-phase arsenic oxide by lime: Kinetic and mechanistic studies. *Environ. Sci. Technol.* **2001**, *35*, 794–799.

17. Zhao, Y.; Zhang, J.; Huang, W.; Wang, Z.; Li, Y.; Song, D.; Zhao, F.; Zheng, C. Arsenic emission during combustion of high arsenic coals from Southwestern Guizhou, China. *Energy Convers. Manag.* **2008**, *49*, 615–624.

18. Swaine, D. J. Why trace elements are important. *Fuel Process. Technol.* **2000**, *65*, 21–33.

19. Ebdon, L., Pitts, L., Cornelis, R., Crews, H., Donard, O. F. X., Quevauviller, P., Eds. *Trace Element Speciation for Environment, Food and Health*; Royal Society of Chemistry: Cambridge, 2001.

20. Fitzpatrick, S.; Ebdon, L.; Foulkes, M. E. Separation and detection of arsenic and selenium species in environmental samples by HPLC-ICP-MS. Int. J. Environ. Anal. Chem. 2002; pp 835–841.

21. Zakzeski, J.; Bruijnincx, P. C. A.; Jongerius, A. L.; Weckhuysen, B. M. The Catalytic Valorization of Lignin for the Production of Renewable Chemicals. *Chem. Rev.* **2010**, *110*, 3552–3599.

22. Gherman, B. F.; Cramer, C. J. Quantum chemical studies of molecules incorporating a $Cu_2O_2^{2+}$ core. *Coord. Chem. Rev.* **2009**, *253*, 723–753.

23. Piecuch, P.; Włoch, M. Renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed Hamiltonian. *J. Chem. Phys.* **2005**, *123*, 224105.

24. Piecuch, P.; Włoch, M.; Gour, J. R.; Kinal, A. Single-reference, size-extensive, non-iterative coupled-cluster approaches to bond breaking and biradicals. *Chem. Phys. Lett.* **2006**, *418*, 467–474.

25. Andersson, K.; Malmqvist, P.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

26. Tamukong, P. K.; Theis, D.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 multireference perturbation theory description of diatomic scandium, chromium, and manganese. *J. Phys. Chem. A* **2012**, *116*, 4590–4601.

27. Tamukong, P. K.; Hoffmann, M. R.; Li, Z.; Liu, W. Relativistic GVVPT2 multireference perturbation theory description of the electronic states of Y_2 and Tc_2. *J. Phys. Chem. A* **2014**, *118*, 1489–1501.

28. Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457–484.

29. Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

30. Grimme, S. Accurate Calculation of the Heats of Formation for Large Main Group Compounds with Spin-Component Scaled MP2 Methods. *J. Phys. Chem. A* **2005**, *109*, 3067–3077.

31. Khait, Y. G.; Song, J.; Hoffmann, M. R. Macroconfigurations in molecular electronic structure theory. *Int. J. Quantum Chem.* **2004**, *99*, 210–220.

32. Song, J.; Khait, Y. G.; Hoffmann, M. R. Abstracts of Papers, 222nd ACS National Meeting, Chicago, IL, United States, August 26-30, 2001, PHYS-292.

33. Kirtman, B. Simultaneous calculation of several interacting electronic states by generalized Van Vleck perturbation theory. *J. Chem. Phys.* **1981**, *75*, 798.

34. Malrieu, J.; Durand, P.; Daudey, J. Intermediate Hamiltonians as a new class of effective Hamiltonians. *J. Phys. A. Math. Gen.* **1985**, *18*, 809–826.

35. Lindgren, I.; Morrison, J. *Atomic Many-Body Theory*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1986.

36. Mukhopadhyay, D.; Datta (nee Kundu), B.; Mukherjee, D. The construction of a size-extensive intermediate Hamiltonian in a coupled-cluster framework. *Chem. Phys. Lett.* **1992**, *197*, 236–242.

37. Malrieu, J.-P.; Heully, J.-L.; Zaitsevskii, A. Multiconfigurational second-order perturbative methods: Overview and comparison of basic properties. *Theor. Chim. Acta* **1995**, *90*, 167–187.

38. Meissner, L. Fock-space coupled-cluster method in the intermediate Hamiltonian formulation: Model with singles and doubles. *J. Chem. Phys.* **1998**, *108*, 9227.

39. Landau, A.; Eliav, E.; Kaldor, U. Intermediate Hamiltonian Fock-space coupled-cluster method. *Chem. Phys. Lett.* **1999**, *313*, 399–403.

40. Eliav, E.; Borschevsky, A.; Shamasundar, K. R.; Pal, S.; Kaldor, U. Intermediate Hamiltonian Hilbert space coupled cluster method: Theory and pilot application. *Int. J. Quantum Chem.* **2009**, *109*, 2909–2915.

41. Kuhler, K.; Hoffmann, M. R. A nondiagonal quasidegenerate fourth-order perturbation theory. *J. Math. Chem.* **1996**, *20*, 351–364.

42. Murphy, R. B.; Messmer, R. P. Generalized Møller–Plesset perturbation theory applied to general MCSCF reference wave functions. *Chem. Phys. Lett.* **1991**, *183*, 443–448.

43. Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chem. Phys. Lett.* **1998**, *288*, 299–306.

44. Nakano, H.; Nakatani, J.; Hirao, K. Second-order quasi-degenerate perturbation theory with quasi-complete active space self-consistent field reference functions. *J. Chem. Phys.* **2001**, *114*, 1133–1141.

45. Hoffmann, M. R. Third-order complete active space self-consistent field based generalized Van Vleck perturbation theory. *Chem. Phys. Lett.* **1993**, *210*, 193–200.

46. U.S. Energy Information Administration, Annual Energy Outlook 2015. *Off. Integr. Int. Energy Anal.* **2015**, *1*, 1–244.

47. Nelson, P. F. Trace metal emissions in fine particles from coal combustion. Energy and Fuels. 2007; pp 477–484.

48. Van Der Hoek, E. E.; Comans, R. N. Modeling arsenic and selenium leaching from acidic fly ash by sorption on iron (hydr)oxide in the fly ash matrix. *Environ. Sci. Technol.* **1996**, *30*, 517–523.

49. Huggins, F. E.; Senior, C. L.; Chu, P.; Ladwig, K.; Huffman, G. P. Selenium and arsenic speciation in fly ash from full-scale coal-burning utility plants. *Environ. Sci. Technol.* **2007**, *41*, 3284–3289.

50. Staudt, J. E.; Engelmeyer, T.; Weston, W. H.; Sigling, R. The Impact Of Arsenic On Coal Fired Power Plants Equipped With SCR. ICAC Forum. Houston, TX, 2002.

51. Sun, L.; Bu, Y.; Yan, S. Structures and bonding character of cyclic AsO_2 and AsS_2 systems in the doublet state. *Chem. Phys. Lett.* **2004**, *387*, 160–164.

52. Urban, D. R.; Wilcox, J. A theoretical study of properties and reactions involving arsenic and selenium compounds present in coal combustion flue gases. *J. Phys. Chem. A* **2006**, *110*, 5847–52.

53. Monahan-Pendergast, M.; Przybylek, M.; Lindblad, M.; Wilcox, J. Theoretical predictions of arsenic and selenium species under atmospheric conditions. *Atmos. Environ.* **2008**, *42*, 2349–2357.

54. Rosli, A. N.; Zabidi, N. A.; Kassim, H. A.; Shrivastava, K. N. Ab initio calculation of vibrational frequencies of AsO glass. *J. Non. Cryst. Solids* **2010**, *356*, 428–433.

55. Da Hora, G. C. A.; Longo, R. L.; Da Silva, J. B. P. Calculations of structures and reaction energy profiles of As_2O_3 and As_4O_6 species by quantum chemical methods. *Int. J. Quantum Chem.* **2012**, *112*, 3320–3324.

56. Tossell, J. Theoretical studies on arsenic oxide and hydroxide species in minerals and in aqueous solution. *Geochim. Cosmochim. Acta* **1997**, *61*, 1613–1623.

57. Jensen, J. O.; Gilliam, S. J.; Banerjee, A.; Zeroka, D.; Kirkby, S. J.; Merrow, C. N. A theoretical study of As4O6: Vibrational analysis, infrared and Raman spectra. *J. Mol. Struct. THEOCHEM* **2003**, *664-665*, 145–156.

58. Mokambe, R. M.; Hicks, J. M.; Kerker, D.; Jiang, W.; Theis, D.; Chen, Z.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 multireference perturbation theory study of selenium oxides. *Mol. Phys.* **2013**, *111*, 1078–1091.

59. Dudley, T. J.; Hoffmann, M. R. Theoretical study of the ground and first excited singlet state potential energy surfaces of disulphur monoxide (S_2O). *Mol. Phys.* **2003**, *101*, 1303–1310.

60. Devarajan, A.; Gaenko, A. V.; Khait, Y. G.; Hoffmann, M. R. Generalized van Vleck perturbation theory (GVVPT2) study of the excited states of benzene and the azabenzenes. *J Phys Chem A* **2008**, *112*, 2677–2682.

61. Azenkeng, A.; Laumb, J.; Jensen, R.; Olson, E.; Benson, S.; Hoffmann, M. Abstracts of Papers, 231st ACS National Meeting, Atlanta, GA, March 26–30, 2006, PHYS–448.

62. Song, J.; Khait, Y. G.; Wang, H.; Hoffmann, M. R. Low-lying electronic states of difluorodioxirane. *J. Chem. Phys.* **2003**, *118*, 10065–10072.

63. Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

64. Pople, J. A.; Head-Gordon, M.; Raghavachari, K. Quadratic configuration interaction. A general technique for determining electron correlation energies. *J. Chem. Phys.* **1987**, *87*, 5968.

65. Kowalski, K.; Piecuch, P. The method of moments of coupled-cluster equations and the renormalized CCSD[T], CCSD(T), CCSD(TQ), and CCSDT(Q) approaches. *J. Chem. Phys.* **2000**, *113*, 18.

66. Kowalski, K.; Piecuch, P. Renormalized CCSD(T) and CCSD(TQ) approaches: Dissociation of the N_2 triple bond. *J. Chem. Phys.* **2000**, *113*, 5644–5652.

67. Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musiał, M. Efficient computer implementation of the renormalized coupled-cluster methods: The R-CCSD[T], R-CCSD(T), CR-CCSD[T], and CR-CCSD(T) approaches. *Comput. Phys. Commun.* **2002**, *149*, 71–96.

68. Włoch, M.; Gour, J. R.; Piecuch, P. Extension of the Renormalized Coupled-Cluster Methods Exploiting Left Eigenstates of the Similarity-Transformed Hamiltonian to Open-Shell Systems: A Benchmark Study. *J. Phys. Chem. A* **2007**, *111*, 11359–11382.

69. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

70. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

71. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157*, 200–206.

72. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

73. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

74. Curtiss, L. A.; McGrath, M. P.; Blaudeau, J.-P.; Davis, N. E.; Binning, R. C.; Radom, L. Extension of Gaussian-2 theory to molecules containing third-row atoms Ga–Kr. *J. Chem. Phys.* **1995**, *103*, 6104–6113.

75. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

76. Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. *J. Chem. Phys.* **1999**, *110*, 7667–7676.

77. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

78. Gordon, M. S.; Schmidt, M. W. In *Theory Appl. Comput. Chem.*; Dykstra, C., Frenking, G., Kim, K., Scuseria, G., Eds.; Elsevier, 2005; Chapter 41, pp 1167–1189.

79. 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*, 1200–1211.

80. Mbote, Y. E. B.; Khait, Y. G.; Hardel, C.; Hoffmann, M. R. Multireference Generalized Van Vleck Perturbation Theory (GVVPT2) Study of the NCO + HCNO Reaction: Insight into Intermediates. *J. Phys. Chem. A* **2010**, *114*, 8831–8836.

81. Knight, L. B.; Jones, G. C.; King, G. M.; Babb, R. M.; McKinley, A. J. Electron spin resonance and theoretical studies of the PO_2 and AsO_2 radicals in neon matrices at 4 K: Laser vaporization and x-irradiation radical generation techniques. *J. Chem. Phys.* **1995**, *103*, 497–505.

82. Grund, S. C.; Hanusch, K.; Breunig, H. J.; Wolf, H. U. *Ullmann's Encycl. Ind. Chem.*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006.

83. Lee, E. P. F.; Dyke, J. M.; Chau, F.-T.; Chow, W.-K.; Mok, D. K. W. An ab initio study on the ground and low-lying doublet electronic states of SbO_2. *J. Chem. Phys.* **2006**, *125*, 064307.

84. Lee, E. P. F.; Dyke, J. M.; Mok, D. K. W.; Chau, F.-t.; Chow, W.-k. Ab initio calculations on low-lying electronic states of SbO_2^- and Franck-Condon simulation of its photodetachment spectrum. *J. Chem. Phys.* **2007**, *127*, 094306.

85. Gilliam, S. J.; Jensen, J. O.; Banerjee, A.; Zeroka, D.; Kirkby, S. J.; Merrow, C. N. A theoretical and experimental study of Sb_4O_6: vibrational analysis, infrared, and Raman spectra. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2004**, *60*, 425–434.

86. Reddy, B.; Jena, P. Signature of crystalline order in ultra-small metal-oxide clusters. *Chem. Phys. Lett.* **1998**, *288*, 253–260.

87. Allen, J. P.; Carey, J. J.; Walsh, A.; Scanlon, D. O.; Watson, G. W. Electronic Structures of Antimony Oxides. *J. Phys. Chem. C* **2013**, *117*, 14759–14769.

88. Martin, J. M. L.; Sundermann, A. Correlation consistent valence basis sets for use with the Stuttgart-Dresden-Bonn relativistic effective core potentials: the atoms Ga-Kr and In-Xe. *J. Chem. Phys.* **2001**, *114*, 3408–3420.

89. Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

90. Chakar, F. S.; Ragauskas, A. J. Review of current and future softwood kraft lignin process chemistry. *Ind. Crops Prod.* **2004**, *20*, 131–141.

91. Crestini, C.; D'Auria, M. Singlet oxygen in the photodegradation of lignin models. *Tetrahedron* **1997**, *53*, 7877–7888.

92. Wang, R.; Chen, C. L.; Gratzl, J. S. Ozonation of pine kraft lignin in alkaline solution. Part 1: Ozonation, characterization of kraft lignin and its ozonated preparations. *Holzforschung* **2004**, *58*, 622–630.

93. Akiyama, T.; Goto, H.; Nawawi, D. S.; Syafii, W.; Matsumoto, Y.; Meshitsuka, G. Erythro/threo ratio of β-O-4-structures as an important structural characteristic of lignin. Part 4: Variation in the erythro/threo ratio in softwood and hardwood lignins and its relation to syringyl/guaiacyl ratio. *Holzforschung* **2005**, *59*, 276–281.

94. Holmgren, A.; Brunow, G.; Henriksson, G.; Zhang, L.; Ralph, J. Non-enzymatic reduction of quinone methides during oxidative coupling of monolignols: Implications for the origin of benzyl structures in lignins. *Org. Biomol. Chem.* **2006**, *4*, 3456–3461.

95. Xie, X.; Goodell, B.; Zhang, D.; Nagle, D. C.; Qian, Y.; Peterson, M. L.; Jellison, J. Characterization of carbons derived from cellulose and lignin and their oxidative behavior. *Bioresour. Technol.* **2009**, *100*, 1797–1802.

96. Beste, A.; Buchanan, A. C. Computational study of bond dissociation enthalpies for lignin model compounds. Substituent effects in phenethyl phenyl ethers. *J. Org. Chem.* **2009**, *74*, 2837–2841.

97. Younker, J. M.; Beste, A.; Buchanan, A. C. Computational study of bond dissociation enthalpies for substituted β-O-4 lignin model compounds. *ChemPhysChem* **2011**, *12*, 3556–3565.

98. Mo, N.; Tandar, W.; Savage, P. E. Aromatics from saturated and unsaturated fatty acids via zeolite catalysis in supercritical water. *J. Supercrit. Fluids* **2015**, *102*, 73–79.

99. 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*, 154104.

100. Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

101. Francl, M. M.; Petro, W.; Hehre, W.; Binkley, J.; Gordon, M.; DeFrees, D.; Pople, J. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654.

102. Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* basis set for atoms K through Zn. *J. Chem. Phys.* **1998**, *109*, 1223.

103. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **2012**, *4*, 17.

104. Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Van Dam, H.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.

105. Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

106. Tolman, W. B. Making and Breaking the Dioxygen O–O Bond: New Insights from Studies of Synthetic Copper Complexes. *Acc. Chem. Res.* **1997**, *30*, 227–237.

107. Holland, P. L.; Tolman, W. B. Dioxygen activation by copper sites: relative stability and reactivity of (μ-η^2:η^2-peroxo)- and bis(μ-oxo)dicopper cores. *Coord. Chem. Rev.* **1999**, *190-192*, 855–869.

108. Que, Jr., L.; Tolman, W. B. Bis(μ-oxo)dimetal "Diamond" Cores in Copper and Iron Complexes Relevant to Biocatalysis. *Angew. Chemie Int. Ed.* **2002**, *41*, 1114–1137.

109. Lewis, E. A.; Tolman, W. B. Reactivity of Dioxygen–Copper Systems. *Chem. Rev.* **2004**, *104*, 1047–1076.

110. Piquemal, J. P.; Pilmé, J. Comments on the nature of the bonding in oxygenated dinuclear copper enzyme models. *J. Mol. Struct. THEOCHEM* **2006**, *764*, 77–86.

111. Cramer, C. J.; Kinal, A.; Włoch, M.; Piecuch, P.; Gagliardi, L. Theoretical Characterization of End-On and Side-On Peroxide Coordination in Ligated Cu_2O_2 Models. *J. Phys. Chem. A* **2006**, *110*, 11557–11568.

112. Saito, T.; Kataoka, Y.; Nakanishi, Y.; Matsui, T.; Kitagawa, Y.; Kawakami, T.; Okumura, M.; Yamaguchi, K. Which hybrid GGA DFT is suitable for Cu_2O_2 systems if the spin contamination error is removed? *Chem. Phys.* **2010**, *368*, 1–6.

113. Yanai, T.; Kurashige, Y.; Neuscamman, E.; Chan, G. K. L. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J. Chem. Phys.* **2010**, *132*.

114. Hoffmann, A.; Herres-Pawlis, S. Donor-driven conformational flexibility in a real-life catalytic dicopper(ii) peroxo complex. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6430–6440.

115. Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.

116. Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted ab initio pseudopotentials for the first row transition elements. *J. Chem. Phys.* **1987**, *86*, 866–872.

117. Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor. Chim. Acta* **1995**, *90*, 87–114.

118. Mokambe, R. M.; Khait, Y. G.; Hoffmann, M. R. Ground and Low-Lying Excited Electronic States of [3,3'] Bidiazirinylidene (C_2N_4). *J. Phys. Chem. A* **2010**, *114*, 8119–8125.

119. Kozlowski, P. M.; Dupuis, M.; Davidson, E. R. The Cope Rearrangement Revisited with Multireference Perturbation Theory. *J. Am. Chem. Soc.* **1995**, *117*, 774–778.

120. Persson, B. J.; Roos, B. O.; Pierloot, K. A theoretical study of the chemical bonding in $M(CO)_x$ (M=Cr, Fe, and Ni). *J. Chem. Phys.* **1994**, *101*, 6810–6821.

121. Ghosh, A.; Taylor, P. R. High-level ab initio calculations on the energetics of low-lying spin states of biologically relevant transition metal complexes: A first progress report. 2003.

122. Hoffmann, M. R.; Helgaker, T. Use of density functional theory orbitals in the GVVPT2 variant of second-order multistate multireference perturbation theory. *J. Phys. Chem. A* **2015**, *119*, 1548–1553.