COMPUTATIONAL MODEL OF THE CATALYTIC CYCLE OF ORGANOSELENIUM ANTIOXIDANTS

by

Gavin S. Heverly-Coulson

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
July 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF CHEMISTRY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "COMPUTATIONAL MODEL OF THE CATALYTIC CYCLE OF ORGANOSELENIUM ANTIOXIDANTS" by Gavin S. Heverly-Coulson in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: July 11, 2012

External Examiner: J. W. Gauld

Research Supervisor: R. J. Boyd

Examining Committee: A. D. Becke

D. J. Burnell

D. F. Weaver

Departmental Representative:

DALHOUSIE UNIVERSITY

DATE: July 11, 2012

AUTHOR: Gavin S. Heverly-Coulson

TITLE: COMPUTATIONAL MODEL OF THE CATALYTIC CYCLE OF ORGANOSELENIUM ANTIOXIDANTS

DEPARTMENT OR SCHOOL: Department of Chemistry

DEGREE: Ph.D. CONVOCATION: October YEAR: 2012

Signature of Author

For my teachers,

both professional and otherwise.

Table of Contents

List of Tables . . . . . **ix**

List of Figures . . . . . **x**

Abstract . . . . . **xiii**

List of Abbreviations and Symbols Used . . . . . **xiv**

Acknowledgements . . . . . **xviii**

Chapter 1 Introduction **1**

Chapter 2 Chemical Background **4**

2.1 Selenium Chemistry 4

2.2 Biochemical Selenium 5
2.2.1 Selenium uptake and transport in plants and animals 6
2.2.2 Selenocysteine 7
2.2.3 Glutathione peroxidase 9

2.3 Glutathione Peroxidase Mimics 11
2.3.1 *N,N*-dimethyl-benzylamine-2-selenol 12

Chapter 3 Theoretical Background **18**

3.1 The Schrödinger Equation 19
3.1.1 The Born-Oppenheimer approximation 20
3.1.2 The orbital approximation 21
3.1.3 Linear combination of atomic orbitals 23

3.2 Basis Sets 23
3.2.1 Slater-type orbitals 24
3.2.2 Gaussian-type orbitals 24
3.2.3 Basis set construction 25
3.2.4 Commonly used basis sets 27

3.3 Hartree-Fock Theory 28
3.3.1 The variational principle 29

3.3.2 Roothaan-Hall equations 30
3.3.3 Self-consistent field procedure 32

3.4 Multi-Configuration Methods 33
3.4.1 Configuration interaction 35
3.4.2 Møller-Plesset perturbation theory 36

3.5 Density-Functional Theory 38
3.5.1 Hohenberg-Kohn Theorem 39
3.5.2 Kohn-Sham Theory 40
3.5.3 The exchange-correlation hole 41
3.5.4 Density-Functional Approximation 43

3.6 Solvent Effects 44
3.6.1 Self-consistent reaction field 45

3.7 Quantum Theory of Atoms in Molecules 47

Chapter 4 Benchmark of Density Functional Methods for Use with Organoselenium Compounds 53

4.1 Motivation 53

4.2 Model Systems and Computational Methods 54

4.3 Results and Discussion 56
4.3.1 Geometry prediction 56
4.3.2 Energy prediction 61

4.4 Conclusions 68

Chapter 5 Reduction of Hydrogen Peroxide by Selenol 69

5.1 Computational Methods 70

5.2 Results and Discussion 70
5.2.1 Diselenide Reaction 70
5.2.2 Selenol Reaction 73
5.2.3 Zwitterion Reaction 74

5.3 Conclusion 76

Chapter 6 Reaction of Selenenic Acid with Thiol 77

6.1 Introduction 77

6.2 Computational Methods 78

6.3 Results and Discussion 79
6.3.1 Thiolate attack 79
6.3.2 Thiol attack 81

6.4 Conclusions 86

Chapter 7 Reduction of Selenylsulfide by Thiol 87

7.1 Introduction 87

7.2 Computational Methods 88

7.3 Results and Discussion 89
7.3.1 Small Model Systems 89
7.3.2 Attempts to Model DMBS 93

7.4 Conclusions 93

Chapter 8 Effects of Monosubstitution on the Peroxide Reduction Mechanism 95

8.1 Motivation 95

8.2 Computational Methods 96

8.3 Results and Discussion 97
8.3.1 Zwitterion formation 97
8.3.2 Peroxide reduction 99

8.4 Conclusions 107

Chapter 9 Conclusions and Future Work 108

9.1 Benchmarking DFT Methods 108

9.2 Modelling the Catalytic Cycle of an Organoselenium Antioxidant . . 109
9.2.1 Peroxide reduction 109
9.2.2 Selenenic acid substitution 110
9.2.3 Selenylsulfide reduction 110
9.2.4 Peroxide reduction using monosubstituted DMBS 111

9.3 Future Directions 111
9.3.1 Testing robustness of DMBS 111

9.3.2 Competing side reactions 112
9.3.3 Improve GPx-like activity 114

Bibliography **116**

Appendix A Copyright Permission Letters **125**

List of Tables

4.1 Arithmetic mean C-Se bond lengths from **1**, **2**, **3**, **6**, **7** (in angstroms). 59

4.2 BDEs calculated at the B3PW91/6-311+G(2df,p) level using geometries optimized at either the B3PW91/6-31+G(d,p) or the B3PW91/6-311+G(2df,p) level (in kJ/mol). 68

5.1 Calculated Gibbs energies of reaction in kJ/mol. 73

6.1 Reaction energies for nucleophilic attack of various thiols on the selenenic acid of DMBS, in kJ/mol. Energy barriers are the difference between the TS and RC and total energy of reaction is the difference between the PC and RC. 81

8.1 Gibbs energies for formation of the zwitterion from the selenol of DMBS (energies relative to the selenol form in kJ/mol). . . . 98

8.2 QTAIM charges for the selenium atom in DMBS (au). 99

8.3 Gibbs energies for reaction of zwitterionic DMBS with hydrogen peroxide (in kJ/mol). Δ G‡ is difference between the TS and RC energies, while Δ G$_{rxn}$ is the difference between the PC and RC energies. 101

8.4 The volume of the selenium atom in the zwitterionic form of each molecule, calculated at the 0.001 au density cutoff, in au^3. 103

List of Figures

2.1 The catalytic cycle of GPx. 10

2.2 Various GPx mimics proposed in the literature. 12

2.3 Catalytic cycle proposed by Reich and Jasperse and modified by Wilson *et al.* 13

2.4 Catalytic cycle proposed by Iwaoka and Tomoda. 14

3.1 Flowchart showing the SCF procedure. 32

3.2 General forms of the parallel (a) and opposite (b) spin exchange correlation holes. 42

3.3 (a) Electron density contour plot of formaldehyde in the plane of the molecule. (b) Gradient vector field of the density in (a). Wavefunction calculated at B3LYP/6-31G level. 48

3.4 Molecular graph of malonic acid. Carbon, oxygen, and hydrogen atoms are large grey, red, and white spheres, respectively. BCPs are green spheres, the RCP is small red sphere, bond paths are black lines. Wavefunction calculated at the B3LYP/6-31G level. 50

4.1 The series of organoselenium and organosulfur compounds included in this benchmark. 55

4.2 The series of homolytic bond cleavage reactions included in this benchmark. 55

4.3 Optimized C−Se bond lengths in selected molecules. 58

4.4 Optimized Se−O bond length in methyl selenenic acid (**3**). . . 60

4.5 Optimized Se−Se bond length in dimethyl diselenide (**6**). . . . 61

4.6 RMSD for the bond length of all selenium-containing bonds. . 62

4.7 Bond dissociation energies calculated for methyl selenol. . . . 63

4.8 Bond dissociation energies for the Se−O bond in methyl selenenic acid (**III**). 64

4.9 Bond dissociation energies for the Se−Se bond in dimethyl diselenide (**V**). 65

4.10 Bond dissociation energies for the Se−S bond in dimethyl selenylsulfide (**VI**). 66

4.11 RMSD for the BDE for the seven bonds shown in figure 4.2. . 67

5.1 The reaction of N,N-dimethyl-benzylamine diselenide with hydrogen peroxide. 71

5.2 Gibbs energy profile of the reaction of N,N-dimethyl-benzylamine diselenide with hydrogen peroxide. 72

5.3 The expected reaction of neutral DMBS with hydrogen peroxide. 73

5.4 Two-step mechanism for the reaction of neutral DMBS with a peroxide. 74

5.5 The Gibbs energy profile of the two-step reaction of neutral DMBS with hydrogen peroxide. 74

5.6 Gibbs energy profile of the reaction of charged analogues of DMBS with hydrogen peroxide. 75

6.1 Schematic outline of the reaction of a thiolate with the selenenic acid of DMBS. 79

6.2 Reaction profile for the reaction of a thiolate with the selenenic acid of DMBS. Energies calculated in aqueous solvent are shown with a solid line and methanolic solvent with a dashed line. . . 80

6.3 Solvent-assisted nucleophilic attack of a thiol on the selenenic acid of DMBS. 82

6.4 Reaction profile for the reaction of a thiol with the selenenic acid of DMBS assisted by two water molecules. 83

6.5 Structures found for the water-assisted reaction of methanethiol with the selenenic acid of DMBS. 84

6.6 Molecular graphs of the water-assisted reaction of methanethiol with the selenenic acid of DMBS. Bond and ring critical points are denoted by green and red spheres, respectively. Bond critical point densities given in atomic units. 85

7.1 Addition-elimination reaction of a selenylsulfide and thiolate, as proposed by Bachrach *et al.* 90

7.2 Reaction of methanethiol with dimethyl selenylsulfide using water as a proton shuttle. 91

7.3 Electrostatic potential maps of two selenylsulfides. Blue isosurface corresponds to an electrostatic potential of 0.4 au and the red isosurface corresponds to -0.03 au. 92

8.1 Selenol to zwitterion formation reaction. The substituents listed are placed in the ortho, meta, or para positions relative to the selenol. 97

8.2 Peroxide reduction reaction being modelled. The substituents listed are placed in the ortho, meta, or para positions relative to the selenol. 100

8.3 Gibbs energy barriers for peroxide reduction reaction with substituted DMBS. Heavy black line indicates barrier of the reaction with the unsubstituted compound. 100

8.4 Charge on selenium atom in the zwitterionic form of DMBS vs Gibbs energy barrier for peroxide reduction reaction 102

8.5 Electrostatic potential maps of various systems. Blue isosurface corresponds to an electrostatic potential of 0.2 au and the red isosurface corresponds to -0.04 au. Negative isosurfaces have been faded to show positive regions underneath. 104

8.6 Volume of the selenium atom in the zwitterionic form of DMBS vs Gibbs energy barrier for peroxide reduction reaction 105

8.7 Atomic basin of Se at the 0.001 au isodensity envelope. 106

Abstract

The chemistry of the enzyme glutathione peroxidase and synthetic organoselenium enzyme mimics has been a significant research interest for more than three decades. In this work, the results of a computational study employing modern electronic structure methods to model the reactions of a synthetic glutathione peroxidase mimic are presented.

The ability of nine density-functional theory methods and thirteen basis sets to predict both molecular geometries and bond dissociation energies in organoselenium compounds is examined. This is used to determine the best methodology to employ for the study of glutathione peroxidase mimics.

The key reactions in the catalytic mechanism of the organoselenium antioxidant *N,N*-dimethyl-benzylamine-2-selenol are the focus of the remainder of the document. This is a three-step mechanism which includes many of the organic forms adopted by selenium compounds, including selenol, oxoacids, and selenylsulfides. In the first step of the cycle, the well-studied reduction of hydrogen peroxide by a selenol and a diselenide is modelled. The second step modelled is a substitution reaction at the selenium centre of a selenenic acid with a thiol. The final step discussed is the reduction of the selenium centre in a selenylsulfide, regenerating the selenol and forming a disulfide species. Each mechanism is evaluated by discussing both molecular geometries and reaction energetics.

To close the document, the peroxide reduction reaction is revisited to determine the effects of substitution on the phenyl ring of the synthetic antioxidant. This serves as a preliminary attempt to improve the antioxidant efficiency of this compound. In addition to a discussion of the changes in reaction energetics predicted, the topology of the electron density is studied using the quantum theory of atoms in molecules to better understand how the distribution of electron density is affected by substituents.

List of Abbreviations and Symbols Used

Abbreviations

AO	Atomic Orbital
BDE	Bond Dissociation Energy
cGPx	Cytosolic Glutathione Peroxidase
CGTO	Contracted Gaussian-type Orbital
CI	Configuration Interaction Method
CIS	Configuration Interaction with Single Excitations
CISD	Configuration Interaction with Single and Double Excitations
CPCM	Conductor-like Polarizable Continuum Model
Cys	Cysteine
DFT	Density-Functional Theory
DMBS	*N,N*-dimethyl-benzylamine-2-selenol
DNA	Deoxyribonucleic Acid
GGA	Generalized Gradient Approximation
giGPx	Gastrointestinal Glutathione Peroxidase
GPx	Glutathione Peroxidase
GSH	Glutathione
GTO	Gaussian-type Orbital
HF	Hartree-Fock theory
IAS	Interatomic Surface
IPCM	Isodensity Polarizable Continuum Model
IRC	Intrinsic Reaction Coordinate
LCAO	Linear Combination of Atomic Orbitals
LDA	Local-Density Approximation
LSDA	Local-Spin-Density Approximation
MO	Molecular Orbital

MPPT	Møller-Plesset Perturbation Theory
NMR	Nuclear Magnetic Resonance Spectroscopy
NPA	Natural Population Analysis
PC	Product Complex
PCM	Polarizable Continuum Model
pGPx	Plasma Glutathione Peroxidase
PHGPx	Phospholipid Hydroperoxide Glutathione Peroxidase
QCISD	Quadratic Configuration Interaction with Single and Double Excitations
QTAIM	Quantum Theory of Atoms in Molecules
RC	Reactant Complex
RMSD	Root Mean Square Deviation
ROS	Reactive Oxygen Species
SCF	Self-consistent Field
SCRF	Self-consistent Reaction Field
Sec	Selenocysteine
Ser	Serine
STO	Slater-type Orbital
STQN	Synchronous Transit-guided Quasi-Newton method
tRNA	Transfer Ribonucleic Acid
TS	Transition State

Symbols

$\mathrm{p}K_a$	$-\log_{10} K_a$, Negative logarithm of the acid dissociation constant
pH	$-\log_{10}$ [H^+], Negative logarithm of the concentration of hydrogen cations
Ψ	Molecular wavefunction
$\mathbf{r}$	Vector describing the positions of all electrons
$\mathbf{R}$	Vector describing the positions of all nuclei
$\hat{H}$	Hamiltonion operator
E	Eigenvalues of Schrödinger equation, Total atomic or molecular energy
$\hat{T}$	Kinetic energy operator, Coupled cluster perturbation operator
$\hat{V}$	Potential energy operator, MPPT perturbation operator
$\hbar$	Reduced Planck's constant
m_e	Mass of an electron
e	Charge of an electron
ϵ_o	Permittivity of free space
m_i	Mass of particle i
∇^2	$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$
Z_A	Nuclear charge of nucleus A
r_{ij}	Distance between electron i and j
r_{iA}	Distance between electron i and nucleus A
R_{AB}	Distance between nucleus A and B
χ	One-electron spin orbital
$\phi(\mathbf{r})$	Atomic orbital
c_{ki}	Atomic orbital expansion coefficient
ζ	Orbital exponent
θ, ϕ	Polar angles
g_{ijk}	Primitive Gaussian function
$\mathbf{F}$	Fock matrix
$\mathbf{C}$	Orbital coefficient matrix

$\mathbf{S}$	Overlap matrix
ε	Orbital energy matrix or dielectric constant
$\mathbf{P}$	Charge density matrix
λ	Perturbation parameter or diagonal term of Hessian matrix
$v_{nuc}(\mathbf{r})$	Nuclear potential of a molecular system
$\rho(\mathbf{r})$	Electron density
$T(\rho)$	Kinetic energy functional
$V_{ee}(\rho)$	Electron-electron potential energy functional
$T_0(\rho)$	Non-interacting kinetic energy functional
$J(\rho)$	Classical Coulomb repulsion energy
$E_{XC}[\rho(\mathbf{r})]$	Exchange-correlation energy functional
h_{XC}	Exchange-correlation hole
q_i	Surface element charge
$\mathbf{E}_i$	Electric field
μ	Dipole moment
Λ	Diagonalized Hessian matrix
$N(\Omega)$	Electron population of an atom
$q(\Omega)$	Atomic charge on atom Ω
$\Delta G^{\ddagger}$	Gibbs energy barrier for reaction
ΔG_{rxn}	Gibbs energy of reaction

Acknowledgements

First and foremost, I would like to thank Prof Russ Boyd for welcoming me into his research group. His guidance has been inspiring and the things learned under his supervision will remain with me.

The members of my supervisory committee, Profs Axel Becke, Jean Burnell, and Don Weaver, have provided much appreciated support. They are responsive and quick to provide assistance when requested.

It has been a pleasure to work with all the members of the Boyd research group: Laura Albrecht, Sarah Belong, Hugo Bohórquez, Sean Collins, David Hally, Shenna LaPointe, Michelle Lu, Corey MacDonald, Elias Machaalani, Alexis Taylor, Vikki Walker, Sarah Whittleton, and Jian Wu. They have always provided a friendly environment and innumerable fruitful discussions and helpful suggestions.

My family and friends have been supportive and understanding throughout this whole process. Their kind words and encouragement have been appreciated more than they know. In particular, I would like to acknowledge my parents, who were my first teachers and have been the inspiration to achieve all I have done.

Finally, the Killam Trusts, Nova Scotia Health Research Foundation, Natural Sciences and Engineering Research Council of Canada, and Dalhousie University are thanked for their generous financial support during my time at Dalhousie University. Additionally, thanks are given to ACEnet for the use of their computing facilities, without which I would have been unable to complete the work presented herein.

Chapter 1

Introduction

The ability to predict chemical phenomena and reaction outcomes using mathematical rules was a goal of many early chemists. Some of the foremost chemists of the late 18th and early 19th centuries, like Antoine Lavoisier, John Dalton, Joseph Gay-Lussac, and Amedeo Avogadro, worked to study chemistry in a more quantitative fashion than their predecessors, and in fact many of their contemporaries. Gay-Lussac was optimistic that their work would come to fruition when he wrote, in a contribution to Mémoires de Physique et de Chimie de la Société d'Arcueil, in 1809, "...we are perhaps not far removed from the time when we shall be able to submit the bulk of chemical phenomena to calculation."[1] Through the 19th century, work on thermodynamics and chemical kinetics, much of which was built on the work of Gay-Lussac and his colleagues, led to a new branch of chemistry known as physical chemistry, which applied the concepts of physics to the study of chemistry.

Although physical chemistry introduced new levels of quantitative study to the world of chemistry, it was primarily focused on analysing phenomena in the bulk. With the advent of both atomic and quantum theory, it became desirable, and indeed possible, to study chemistry on the molecular and atomic scales. The work of Erwin Schrödinger introduced a mathematical model for the interactions of electrons and nuclei in atomic systems and marked the beginning of modern electronic structure theory. Unfortunately, the solutions for Schrödinger's equation become exceptionally complex for large systems, so it took some time before his work could be realistically applied to the vast majority of chemical problems. Through the mid to late twentieth century, the development of theoretical methods that allowed chemists to solve accurate approximations of the Schrödinger equation for large molecular systems, in conjunction with rapid advances in computing power, led to the introduction

of computational chemistry. This work was acknowledged with the 1998 Nobel prize in chemistry awarded to Walter Kohn and John Pople for their contributions to the development of density-functional theory and computational methods in quantum chemistry, respectively.

In the early days of computational chemistry, electronic structure calculations on a single diatomic molecule could constitute an entire doctoral thesis; however, with today's advances in the theoretical methods and computing power available, these same calculations can be performed on a desktop computer in seconds and have been incorporated into undergraduate chemistry curricula. Highly accurate calculations on systems containing a few dozen atoms have become routine in chemistry research and the maximum size of systems that can be studied is constantly increasing.

Computational chemistry provides a new way to study chemistry; in the view of some, it achieves the lofty vision of Gay-Lussac. With computational chemistry, one can model discrete molecules or infinitely repeating materials and learn about specific interactions in chemical systems that can be very difficult, or even impossible, to study using traditional experimental chemistry techniques. It is particularly useful for the study of chemicals that are dangerous or simply malodorous, as is the case for the compounds in the work presented here, as one can learn a great deal without needing to come in physical contact with them. Computational studies are increasingly used in chemical research to probe a chemical process of interest quickly before expending resources to study potentially unfavourable reactions, synthesising expensive compounds, or investigating unexpected experimental results.

This document will present the results of the application of modern electronic structure methods to the catalytic redox reaction of hydrogen peroxide and a thiol to produce water and a disulfide under the influence of an organoselenium catalyst.

Chapter two highlights the relevant chemistry and biochemistry of selenium, with special emphasis on the use of this element as an antioxidant in both natural selenoenzymes and synthetic molecules, which mimic the behaviour of these enzymes. This section is intended to introduce readers who are unfamiliar with the current

research surrounding biochemical selenium to key concepts required to understand the motivations for this study.

In chapter three, the theoretical basis for the computational methods used in this study will be shown. A brief discussion of the historical development of electronic structure methods will be given followed by a more in depth study of the methods used herein, including Hartree-Fock (HF) and density-functional theories (DFT), implicit solvation methods, and the quantum theory of atoms in molecules (QTAIM).

Chapter four presents the results of a benchmarking study to determine the optimal DFT method and basis set pair for the study of organoselenium compounds. This benchmark attempts to find a balance between accuracy of the geometries and bond energies in the compounds of interest.

The following three chapters are devoted to modelling the three-step, catalytic reaction mechanism for the reduction of hydrogen peroxide with thiol using *N,N*-dimethyl-benzylamine-2-selenol (DMBS) as the organoselenium catalyst. Chapter five analyses the oxidation of the selenol to a selenenic acid by hydrogen peroxide. Chapter six determines the mechanism for conversion of the selenenic acid to a selenylsulfide, liberating water. Chapter seven studies the recovery of catalytically active selenol through reduction of the selenylsulfide with an additional thiol molecule.

The penultimate chapter shows some attempts to optimize the performance of *N,N*-dimethyl-benzylamine-2-selenol in the peroxide reduction step of the mechanism through various substitutions on the phenyl ring of the catalyst. The electronic and steric effects of these substituents are analysed through atomic charges, electrostatic potential maps, and atomic volumes.

The document closes with some overall conclusions of the findings of this study and the author's view of the future outlook of this area of research.

Chapter 2

Chemical Background

The relevant chemical and biochemical background for the research presented herein will be introduced in this chapter. The intention is to introduce readers who are unfamiliar with the current research surrounding biochemical selenium to the key concepts required to understand the motivations of the work presented in the following chapters. The discussion will include an overview of the biochemical role of selenium, with special focus on selenium-containing enzymes (selenoenzymes), leading into a review of nearly 30 years of work producing small oragnoselenium compounds that mimic the behaviour of these selenoenzymes.

2.1 Selenium Chemistry

Selenium (atomic number 34) is the third element in the chalcogen group, found below oxygen and sulfur in the periodic table. It was originally discovered as a byproduct of sulfuric acid production and copper ore roasting by the Swedish chemist Berzelius in 1818.[2] Existing in group 16 of the periodic table between sulfur, a non-metal, and tellurium, a metal, selenium is a metalloid. It has a variety of allotropic forms resembling non-metals (Se_8 rings, red selenium) and metals (helical chains, grey selenium), as well as an amorphous form (black selenium).[3–5]

Much of the historical use of selenium is dominated by its unique electrical properties. The grey and black allotropes are semiconductors with bandgaps of 2.6 and 1.8 eV, respectively,[3] corresponding to light with wavelengths of 477 and 689 nm, respectively. The grey form exhibits low conductivity in the dark, but on exposure to light, the conductivity increases more than one hundred times and a small current

is generated, making it a particularly interesting material for incorporation in photovoltaic cells and photoreceptors.[4] It is also capable of converting AC power to DC, giving it early use in rectifiers,[5] although selenium-based rectifiers have mostly been replaced by cheaper and more compact silicon-based ones.

The compounds of selenium, both organic and inorganic, are often analogues of sulfur compounds. Selenium derivatives, however, often have properties distinct from their sulfur analogues, despite the two elements having very similar properties. These differences will be discussed in more detail below.

Selenium chemistry is a varied field and an in-depth review of much of it is beyond the scope of this document. The remainder of the discussion will therefore focus on the biochemistry of selenium and specifically its role as an antioxidant.

2.2 Biochemical Selenium

After its discovery, and for over a century thereafter, selenium was considered toxic and assumed to have no beneficial biochemical role. This view began to change in 1939 with the discovery that selenium was essential for growth of some plants in the vetch family.[6] Nearly two decades later, it was found that selenium is essential for animals as well. While studying the prevention of necrotic liver degeneration in rats and chickens by feeding the animals various combinations of α-tocopherol (vitamin E), cysteine, and Factor 3, a dietary supplement derived from brewer's yeast, it was found that Factor 3 contained selenium.[7] By varying the selenium content in Factor 3, Schwarz and Foltz found that dietary selenium levels were correlated with protection against necrotic degeneration. They further found that injections of sodium selenite ($NaSeO_3$) or potassium selenate (K_2SeO_4) were also capable of protecting against necrotic tissue degeneration, and were many orders of magnitude more effective than vitamin E or cysteine alone, prompting them to suggest for the first time that selenium was an essential micronutrient. Their results were quickly confirmed by others.[8,9] Selenium has since been discovered to play a key role in the metabolism of various bird and mammal species,[10,11] some bacteria,[12,13] and plankton.[14]

2.2.1 Selenium uptake and transport in plants and animals

As an essential micronutrient, the question of how selenium makes its way into the human body is an obvious one. Similar to many other dietary minerals, it begins with uptake from the soil by a plant's roots. The selenate anion (SeO_4^{2-}) is taken up by an active transport system in the roots competitively with sulfate. Selenite (SeO_3^{2-}), another common inorganic form of selenium in soil, appears to enter roots through passive diffusion.[15] Some organoselenium molecules, such as selenomethionine, are also absorbed through active transport mechanisms, again competitively with their sulfur analogues.[16]

Selenate taken up by plants is easily transported from the roots and accumulates in the leaves. Conversely, in seedlings treated with selenite or selenomethionine, more selenium is observed in the root system than the leaves after one week.[17] Inorganic selenium in the leaves is reduced to the selenide anion (Se^{2-}) in the chloroplasts and subsequently synthesised into the amino acid selenocysteine, using the same enzymes that produce cysteine from sulfide. It is then non-specifically incorporated into proteins.[18] Selenocysteine can be further metabolized into selenomethionine and also incorporated non-specifically into proteins in a methionine site.[19] The predominant form of selenium found in plants is selenomethionine.

Once ingested by animals, selenium is absorbed in the small intestine, as with most nutrients. Overall, about eighty percent of the selenium ingested is absorbed, with a higher proportion of selenomethionine being absorbed than selenite.[20] As with in plants, selenate and selenomethionine are absorbed via active transport mechanisms, competitively with their sulfur analogues. Selenite absorbs through passive diffusion. There is evidence that selenocysteine is absorbed using the cysteine active transporters, although this has not been fully documented.[4] Selenate and selenomethionine are transported through the bloodstream intact, however selenite is absorbed by erythrocytes and reduced through reaction with glutathione (GSH) to form an analogue of glutathione trisulfide (GSSSG) where the central sulfur is replaced by selenium (GSSeSG).[21]

2.2.2 Selenocysteine

After the discovery that selenium was an essential trace element, the next challenge was determining its role in metabolism. In 1973, it was discovered that selenium played a role in the function of glutathione peroxidase (GPx),[22] an antioxidant enzyme discovered sixteen years prior.[23] At the time, it was known that GPx catalytically reduced peroxides in erythrocytes, while oxidatively coupling two equivalents of the tripeptide glutathione through formation of a disulfide bond.[23] In the same year, it was shown that GPx was a selenoenzyme with four atoms of selenium per enzyme unit.[24] In 1978, the selenium moiety in GPx was identified as selenocysteine (Sec), the selenium analogue of serine (Ser) and cysteine (Cys), and it was confirmed that this amino acid was responsible for the catalytic action of the enzyme.[25] Previously, Sec had been identified in the glycine reductase enzyme of the bacterium *Clostridium sticklandii*,[13] but GPx was the first mammalian protein discovered to contain the amino acid.

After its initial discovery, it was assumed that Sec was incorporated into proteins via post-translational modification of serine or cysteine. However, in 1986, it was discovered that the genetic sequence that coded GPx in mice[26] and formate dehydrogenase (another Sec containing enzyme) in *E. coli*[27] contained the UGA stop codon, which is used to signal the end of protein translation, in the middle of the protein sequence. This same codon has been found in the genetic sequence for other selenoproteins, leading to Sec being referred to as the twenty-first proteinogenic amino acid.[28] It has been determined that selenocysteine is synthesized *in situ* during protein synthesis. Directed by the UGA codon, serine is esterified to a specialized transfer RNA (tRNA) and converted to selenocysteine through a two-step process using the enzyme selenocysteine synthase and monoselenophosphate as a selenium donor.[29]

By 1994, thirty different proteins had been identified which included selenium, and Sec was confirmed to exist in half of those.[30] The advent of large genomic databases and improved bioinformatics methods has allowed scientists to scan the genome and identify twenty-five proteins coded to include Sec.[31–33] Of these, most mammalian

selenoproteins fall into three families of redox proteins: glutathione peroxidases, thyroid hormone deiodinases, and thioredoxin reductases, although other functions have also been identified.[34,35]

Why use selenium?

Incorporating selenocysteine into a protein introduces added synthetic complexity, which raises the question: what makes selenium important enough to warrant this extra work? In many ways, selenium is quite similar to its neighbour in the periodic table, sulfur. Selenium and sulfur have nearly identical Pauling electronegativities (2.55 and 2.58, respectively) and their atomic and ionic radii are similar, with selenium no more than 15% larger, depending on the atomic radius scale used.[36] One major difference between the two elements is observed in their bonding. Selenium-carbon and selenium-hydrogen bonds show longer bond lengths and lower bond dissociation energies than their sulfur analogues. This, in part, leads to a significantly lower pK_a for the selenol group compared to the thiol. For instance, the selenol in free selenocysteine has a pK_a of 5.2, while the thiol in cysteine is 8.3.[37] Although the surrounding protein environment has an effect on the deprotonation of an amino acid side chain, at physiological pH (6.5-7.5) most Sec residues will be in their deprotonated selenolate form, while most Cys residues will be in the protonated thiol form.

The key differentiating characteristic of seleno-compounds is their nucleophilicity, which has a direct effect on their reactivity. It has been shown that the reaction rate of diphenyl selenide with $(MeO)_2SO_2$ is twelve times higher than the same reaction using diphenyl sulfide.[38] Computational models employing high-level electronic structure theory showed that both selenol and selenolate oxidation by hydrogen peroxide has lower energy barriers than oxidation of their sulfur analogues.[39,40] This high nucleophilicity makes Sec particularly effective for performing redox reactions at non-equilibrium states, whereas Cys is better suited to maintaining a redox equilibrium.[41] Many of the known selenoenzymes, such as glutathione peroxidase and thioredoxin reductase, which both work to counteract oxidative stress, perform their

roles at non-equilibrium conditions where there is an imbalance of reactive oxygen species.

2.2.3 Glutathione peroxidase

As previously mentioned, glutathione peroxidase was the first selenium-containing protein identified in mammals. This enzyme functions as part of the body's natural antioxidant system protecting various tissues from oxidative damage by reducing hydroperoxides. Four distinct families of selenium-containing glutathione peroxidase enzymes have been identified in mammals. The first to be discovered was classical cytosolic GPx (cGPx). It is found inside cells in the cytosolic and mitochondrial compartments.[22–24] The phospholipid hydroperoxide GPx (PHGPx) family contains intracellular, membrane-bound enzymes, which protect cell membranes from oxidative damage.[42–44] The plasma GPx (pGPX) family is found in blood plasma[45] and the gastrointestinal GPx (giGPx) family is found in the gastrointestinal tract.[46] All four families of GPx enzymes fulfill similar functions in the body, but have different specificities for various organic peroxides and glutathione. Each family of GPx enzymes exists as homotetramers except PHGPx, which is a monomeric protein smaller than the subunits of the other families.[47] Because it is the most widely studied, the cGPx family will be discussed for the remainder of this document, unless otherwise specified, and will be simply referred to as GPx.

Glutathione peroxidase is an 84 kilodalton protein, containing a Sec residue at position 35 in each subunit. The active site, which contains Sec, lies in a flat depression on the surface of the protein allowing for easy access by substrates. This exposure of the catalytic site is consistent with the high rates of reaction observed for GPx.[48] Replacing the Sec residue with Cys or Ser, the other naturally occurring group sixteen analogues, greatly reduces the activity of GPx,[49] illustrating the importance of the Sec residue in this selenoprotein.

All aerobic organisms form small amounts of superoxide (O_2^-), hydroxyl radical (HO•), and hydroperoxides (ROOH), collectively known as reactive oxygen species

(ROS), due to incomplete reduction of O_2 during normal metabolism. There is evidence that in small quantities these species play a vital physiological role in intercellular signalling and redox regulation, however when too many accumulate, a condition known as oxidative stress, they can have negative effects on proteins, lipids, DNA, and cell membranes. Oxidative stress has been linked to a number of diseases, including atherosclerosis and some cancers, and the aging process, through damage to mitochondrial DNA.[50]

The human body contains an elaborate defense mechanism to protect against oxidative stress, which reduces ROS to alcohol or water through a multi-step process. The key steps involve the conversion of superoxide to hydrogen peroxide and molecular oxygen by the enzyme superoxide dimutase[51] and reduction of hydrogen peroxide or organic peroxides to water or the corresponding alcohol by GPx, oxidising two equivalents of GSH in the process.[23]

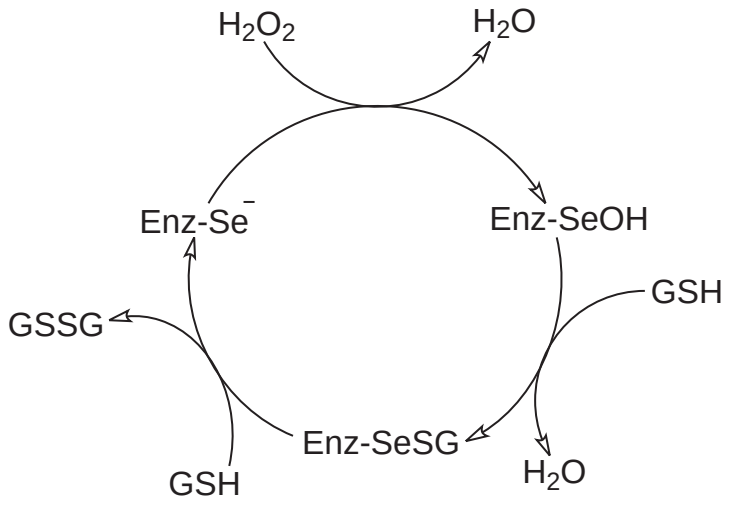


Figure 2.1: The catalytic cycle of GPx.

The catalytic cycle of GPx (figure 2.1) was proposed alongside the first crystal structure of the enzyme.[48] It has since been confirmed by applying both molecular mechanics[52] and hybrid QM/MM procedures[53,54] on the full enzyme as well as quantum mechanics on only the active site residues.[55] Starting with the reduced form of

the selenium moiety in the active site (Enz-Se^-), the selenolate is oxidised to a selenenic acid (Enz-SeOH) through reduction of the peroxide to water. A computational study has suggested that this follows a two-step mechanism, with the selenol first deprotonated by a nearby glutamine residue, followed by reaction of the selenolate anion with the peroxide.[55] However, it was argued in the initial mechanism proposal that the selenol will be rapidly deprotonated by solvent due to its low pK_a.[48] Glutathione (GSH) then attacks the selenenic acid, liberating water and forming a selenyl-sulfide (Enz-SeSG) intermediate. A second molecule of glutathione attacks this intermediate, regenerating the selenol and liberating oxidised glutathione (GSSG). The overall reaction reduces one peroxide molecule to two water molecules and oxidatively couples two glutathione molecules.

2.3 Glutathione Peroxidase Mimics

It is expected that an increase in glutathione peroxidase (GPx) concentration would allow the body to better cope with oxidative stress. Biological studies have shown that adding inorganic selenium salts, such as sodium selenite, to the diet can lead to increased concentrations of selenoproteins and greater GPx activity.[7,10,56] But, these salts can be toxic at high doses, so an alternative means of supplementing the body's antioxidant systems, which does not introduce inorganic selenium to the body, is desired. One method that has been proposed is to create organoselenium compounds that mimic the reductive behaviour of GPx *in vivo* and can act catalytically in the presence of thiols.

A recent computational study showed that the calculated reaction rate for the reduction of hydrogen peroxide by free Sec is several orders of magnitude smaller than the rate associated with GPx,[40] showing that the enzyme environment is very important to this reaction. Therefore, it is necessary to include groups that stabilize all steps of the reaction in order to produce effective GPx mimics.

The first organoselenium compound reported to exhibit GPx-like activity was ebselen (2-phenyl-1,2-benzisoselenazol-3(2H)-one),[57,58] shown in figure 2.2a. It has

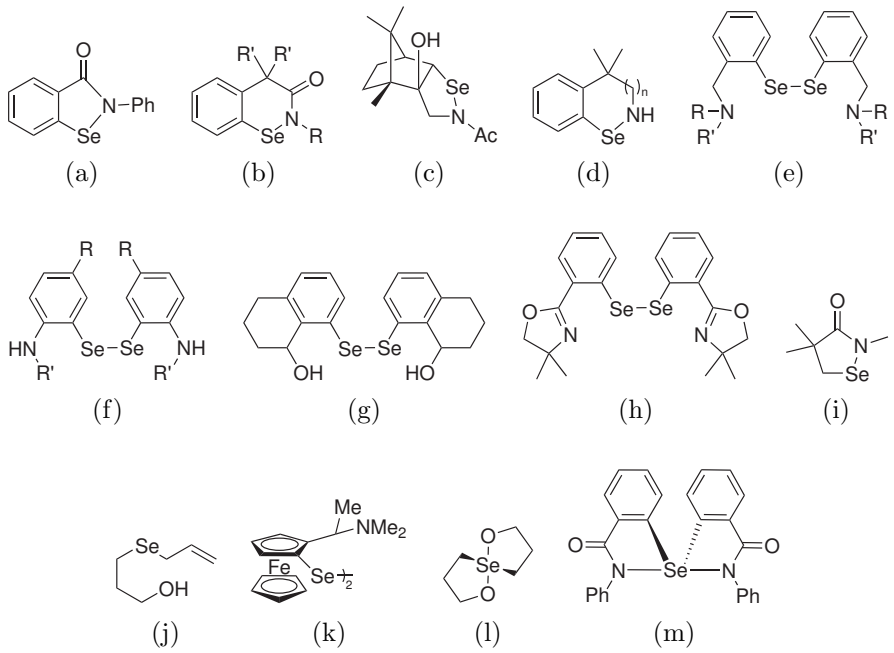


Figure 2.2: Various GPx mimics proposed in the literature.

since been studied extensively *in vitro*[59–63] and computationally.[64–67] On top of its antioxidant effect, ebselen has also been identified as a substrate for the selenoenzyme thioredoxin reductase, increasing the reductase activity of the protein.[68]

Ebselen is a cyclic selenamide, a structure common to many proposed GPx mimics (figures 2.2b,[69] 2.2c,[70] 2.2d[71]), but many other structures have been proposed, including diselenides (figures 2.2e,[72,73] 2.2f,[74] 2.2g,[75] 2.2h[76]), alkyl and aryl selenides (figures 2.2i[77] and 2.2j[78]), and some more exotic compounds like diferrocenyl diselenides (figure 2.2k[79]) and spirocompounds (figures 2.2l[80] and 2.2m[81]). Most GPx mimics contain a basic nitrogen near the selenium, forming covalent or non-covalent selenium-nitrogen interactions. The importance of these interactions to the activity of GPx mimics has been demonstrated,[73,77,82] although some compounds have been proposed with oxygen, which is significantly less basic, in place of the nitrogen.[75]

2.3.1 *N,N*-dimethyl-benzylamine-2-selenol

Although ebselen is one of the most extensively studied, it is not the most promising organoselenium compound that has been proposed as a GPx mimic, in terms of

in vitro GPx-like activity. In 1989, five years after ebselen was proposed as a GPx mimic, Wilson *et al.* synthesised two diselenides which had activities an order of magnitude greater than ebselen.[72] These compounds, *N,N*-dimethyl-benzylamine-2-diselenide (figure 2.2e) and pyrrol-benzylamine-2-diselenide, contain a selenium moiety ortho to an amine functionality, like found in ebselen.

The group constructed these molecules based on two observations made by Reich and Jasperse two years earlier during their study of isoselenazolidin-3-one (figure 2.2i), another GPx mimic.[77] The first was that diselenides function as effectively as cyclic species, like ebselen, while being easier to synthesize. The second stated that a strongly basic group near selenium catalyses the reaction of thiols with diselenide and selenylsulfide intermediates. However, the strength of the interaction between selenium and the basic group needs to be moderated because there is evidence that if these interactions are too strong, the catalyst cannot be regenerated easily due to a preference for thiol exchange reactions rather than disulfide formation.[64]

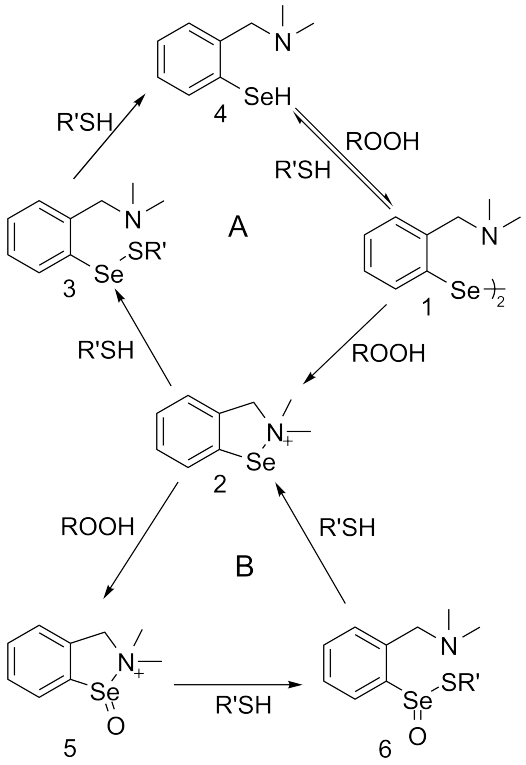


Figure 2.3: Catalytic cycle proposed by Reich and Jasperse[77] and modified by Wilson *et al.*[72]

Wilson *et al.* proposed a catalytic mechanism for the reaction of their tertiary benzylamine diselenides with organic peroxides, shown in figure 2.3. This mechanism is based on one proposed by Reich and Jasperse for the cyclic selenamide GPx mimics they had studied[77] and is significantly different than the accepted cycle for GPx outlined in figure 2.1. It assumes that the role of the nearby nitrogen is to form covalent bonds with the selenium upon oxidation, as in other GPx mimics like ebselen and the selenamide studied by Reich and Jasperse. However, in these other organoselenium compounds the nitrogen is a secondary amide with a hydrogen which is lost upon cyclization to the selenamide producing a neutral tertiary nitrogen. In the compound proposed by Wilson *et al.*, the nitrogen is a tertiary amine, which forms a less stable cationic quaternary nitrogen after cyclization to the selenamine.

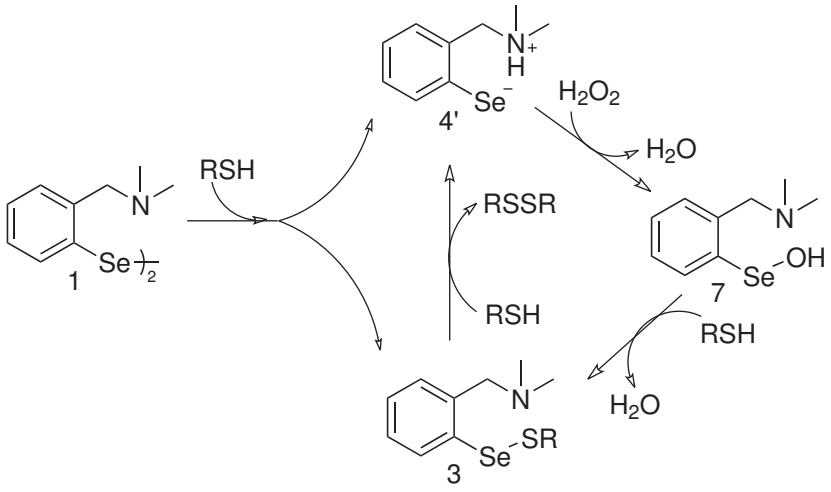


Figure 2.4: Catalytic cycle proposed by Iwaoka and Tomoda.[73]

Later, Iwaoka and Tomoda performed a kinetic study of hydrogen peroxide reduction by tertiary-benzylamine diselenides.[73] They made a number of key observations, which effectively invalidated the mechanism proposed by Wilson *et al.* for this class of GPx mimics. First, their rate equations were extremely similar to those found for the catalytic reaction of GPx with peroxides, indicating that the mechanism for these compounds is likely very similar to that of GPx. Second, the reaction of the diselenide (**1**) with H_2O_2 was very slow, however reaction with thiophenol rapidly

produced a mixture of the selenylsulfide (**3**) and the selenol (**4**).

$$\mathrm{R{-}SeSe{-}R} + \mathrm{R'SH} \longrightarrow \mathrm{R{-}SeSR'} + \mathrm{R{-}SeH}$$

After this initial reaction with thiol, the diselenide form is not observed in the reaction mixture, indicating that it functions as a precatalyst rather than a part of the catalytic cycle. Third, ^{77}Se NMR of the selenol (**4**) had a large upfield shift suggesting a significant negative charge on the selenium, indicating that the catalytically active form of the selenium is its selenolate form, like in the GPx mechanism. *Ab initio* modelling showed that the fully optimized geometry of the selenol was the zwitterionic form (**4'**) with a deprotonated selenium and a protonated amine. Fourth, when the selenol was treated with H_2O_2 in the absence of a thiol, a mixture of the selenenic (**7**) and seleninic acids was observed.

$$\mathrm{R{-}SeH} + \mathrm{H_2O_2} \longrightarrow \mathrm{R{-}SeOH} + \mathrm{R{-}SeO_2H}$$

Upon addition of an excess of peroxide, a large proportion of seleninic acid was observed along with the introduction of selenonic acid.

$$\mathrm{R{-}SeH} + \underset{\text{excess}}{\mathrm{H_2O_2}} \longrightarrow \mathrm{R{-}SeOH} + \mathrm{R{-}SeO_2H} + \mathrm{R{-}SeO_3H}$$

In the presence of a thiol, none of these oxoacids are observed indicating that the conversion of these acids to the selenylsulfide is rapid. All these observations led to the proposal of a new catalytic mechanism for these compounds (figure 2.4).

The role of the intramolecular N···Se non-bonded interaction in *N,N*-dimethylbenzylamine-2-selenol (DMBS) and other diaryl diselenides has been studied extensively. Iwaoka and Tomoda proposed a number of roles for the basic amino group of DMBS in the catalytic cycle.[73] Mugesh and coworkers have performed a series of studies to further elucidate the role of this interaction in this class of compounds.[76,79] They concluded that the basic amino group plays the following roles in the reaction:

to *a)* activate the Se−Se bond for cleavage by a thiol, *b)* abstract the proton on the selenol to produce the more reactive selenolate anion, *c)* stabilize the selenenic acid against further oxidation, *d)* enhance nucleophilic attack at the sulfur in the selenyl-sulfide intermediate, and *e)* deprotonate the thiol to give a high local concentration of more nucleophilic thiolates. To determine the effect of the strength of the N···Se interaction on GPx-like activity, Mugesh *et al.* produced a compound with the nitrogen in an imine bond (figure 2.2h), which demonstrates a stronger and shorter interaction than in DMBS, and found that it was inactive under the same conditions that DMBS is used. Additionally, maintaining the tertiary amine structure, but allowing the nitrogen to participate in π-conjugation, as in 1-selenol-9-dimethylamino-naphthalene, also eliminates much of the GPx-like activity.[76]

Recently, Bhabak and Mugesh[83] have produced a series of secondary and tertiary amine and amide organoselenium compounds using the DMBS structure as a template. They found that the amine compounds perform much better than their amide analogues for catalytic reduction of H_2O_2 or bulky organic peroxides like tertiary-butyl hydroperoxide. Additionally, the secondary amines were about twice as effective as tertiary amines in the same reaction, even though both classes of GPx mimics follow the same mechanism. For the secondary amine compounds, they searched for a cyclic selenamine intermediate, like that in the original mechanism proposed for these catalysts (figure 2.3), however they were unable to find any evidence for cyclization.

Others have attempted to improve the GPx activity of the DMBS system by changing the nature of the aromatic ring. Kumar and Singh[84] replaced the phenyl group in DMBS with a naphthyl group, keeping the selenol and amino groups in the 1 and 2 positions, respectively, and found a ten-fold increase in GPx-like activity. Unfortunately, they make no suggestions for the origin of this greatly enhanced activity. Bhabak and Mugesh[85] have added a methoxy substituent ortho to the selenium moiety and also observed an increase in activity of an order of magnitude compared to DMBS. Since the second ring of the naphthyl-substituted compound is also ortho to the selenium, this may be an indication that the steric effects of substituents in

this position can have a large effect on the GPx-like activity of DMBS.

As outlined above, there has been a great deal of work studying the GPx-like activity of tertiary-benzylamine selenols *in vitro*, however there has been relatively little *in silico* analysis of this class of GPx mimics. By modelling the catalytic reduction of hydrogen peroxide by thiols using DMBS as the catalyst, the importance of each functional group to the high GPx-like activity can be determined. With a complete picture of the full catalytic cycle, it could be possible to design improved GPx mimics through a rational approach by altering functional groups most important to the desired effect. Through the application of computational methods, the relative effectiveness of these substitutions can be determined without synthesising these compounds by studying their reaction energies in the catalytic cycle.

Chapter 3

Theoretical Background

It is possible to study chemistry from a purely mathematical standpoint by applying the laws of quantum mechanics to molecules, which are systems of electrons, protons, and neutrons. Within the field of quantum chemistry, there are many different methods that can be used to solve the electronic structure of a chemical system, each of which has its own strengths and weaknesses. Therefore it is important to apply the right method to the system of interest. To do that, an understanding of the underlying chemical, physical, and mathematical principles is required. A complete discussion of every method is beyond the scope of this document, for that the reader is referred to a standard quantum chemistry textbook.[86–88]

In the following chapters, results obtained from the application of quantum mechanics to systems of organoselenium compounds are presented. This chapter will outline the theories forming the foundation of the methods used. First, a discussion of the basic concepts required for an understanding of quantum chemistry are presented, followed by an overview of various electronic structure theories. Special attention is paid to density-functional theory (DFT), since the work presented in this document was performed using DFT methods. This chapter concludes with an introduction to implicit solvation methods and the quantum theory of atoms in molecules (QTAIM).

3.1 The Schrödinger Equation

The modern study of quantum chemistry was initiated in 1926 when Erwin Schrödinger developed his famous eigenvalue/eigenfunction equation,[89] shown below in its time-independent short hand form.

$$\hat{H}\Psi(\mathbf{r},\mathbf{R}) = E\Psi(\mathbf{r},\mathbf{R}) \tag{3.1}$$

The eigenfunctions of this equation, Ψ, are functions of the spatial coordinates of all nuclei, $\mathbf{R}$, and electrons, $\mathbf{r}$. The Hamiltonian operator, $\hat{H}$, is a differential operator representing the total energy of the system. The eigenvalues of the above equation are the total atomic or molecular energies, E, for the state associated with a particular eigenfunction. There are many solutions to the Schrödinger equation, each corresponding to a different state of the molecule. The state with the lowest energy, E_0, is the ground state and the eigenfunction associated with it, Ψ_0, is the ground state molecular wavefunction.

The Hamiltonian operator represents the total energy of a molecule containing N electrons and M nuclei. It is composed of two parts: the kinetic energy operator, $\hat{T}$, and the potential energy operator, $\hat{V}$, shown below in atomic units,[i]

$$\hat{T} = \sum_{i}^{N+M} -\frac{1}{2m_i}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = \sum_{i}^{N+M} -\frac{1}{2m_i}\nabla_i^2 \tag{3.2}$$

$$\hat{V} = -\sum_{i}^{N}\sum_{A}^{M}\left(\frac{Z_A}{r_{iA}}\right) + \sum_{i}^{N}\sum_{j>i}^{N}\left(\frac{1}{r_{ij}}\right) + \sum_{A}^{M}\sum_{B>A}^{M}\left(\frac{Z_A Z_B}{R_{AB}}\right) \tag{3.3}$$

where m_i is the mass of particle i, Z_A and Z_B are the nuclear charges, equivalent to the atomic numbers, of nuclei A and B, r_{ij}, r_{iA}, and R_{AB} are the distances between electrons i and j, electron i and nucleus A, and nuclei A and B, respectively. The

[i]Atomic units are a set of units used in theoretical chemistry to ensure that the computed properties are independent of the values of the fundamental constants, which can change as they are refined. In atomics units, the reduced Planck's constant, $\hbar$, the mass of an electron, m_e, the charge of an electron, e, and a constant multiplied by the permittivity of free space, $4\pi\epsilon_o$, are all set equal to 1. Atomic units will be used for the remainder of this discussion, unless otherwise specified.

kinetic energy operator accounts for the energy arising from the momentum of all particles, while the potential energy operator accounts for the energy arising from the classical Coulombic interactions in a molecular system. These interactions are separated into three parts: electron-nucleus attraction, electron-electron repulsion, and nucleus-nucleus repulsion. To calculate the total energy from the Hamiltonian operator, the expectation value must be taken,

$$E = \left\langle \hat{H} \right\rangle \tag{3.4}$$

Solving the Schrödinger equation provides the molecular wavefunction and associated energy of a molecular system. A great deal of information can be extracted from these two pieces of information. Unfortunately, an analytical solution of the Schrödinger equation is only possible for systems containing a single electron, such as the hydrogen atom (a two-body problem) or cationic dihydrogen (in the Born-Oppenheimer approximation discussed below). It is computationally possible to find approximate solutions for larger systems, but it is costly, even with the most powerful computers currently available. In order to apply quantum mechanics to larger systems, it is necessary to make practical approximations to the Schrödinger equation, which provide solutions in a reasonable amount of time.

3.1.1 The Born-Oppenheimer approximation

The first simplification made to the Schrödinger equation was the Born-Oppenheimer approximation,[90] proposed a year after the publication of the Schrödinger equation. In this approximation, the kinetic energy of the nuclei is assumed to be zero. The mass of a single proton is roughly three orders of magnitude greater than that of an electron, therefore the velocity of the nucleus will be significantly less than that of an electron, making this a reasonable approximation. Thus, it can be assumed that the nuclei hold fixed positions relative to the electrons and therefore have no kinetic energy. This means that an electronic Schrödinger equation can be constructed omitting all purely nuclear terms, which considers only the kinetic energy of the electrons, the

Coulombic repulsion between electrons, and electron-nucleus attraction.

$$\hat{H}^{elec}\Psi(\mathbf{r},\mathbf{R}) = E^{elec}\Psi(\mathbf{r},\mathbf{R}) \tag{3.5}$$

Note that in this wavefunction calculation only the positions of the electrons, $\mathbf{r}$, can change while the nuclear coordinates, $\mathbf{R}$, are fixed. This means that for any set of fixed nuclear coordinates, an electronic energy can be calculated, using the electronic Hamiltonian operator,

$$\hat{H}^{elec} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{i}^{N}\sum_{A}^{M}\left(\frac{Z_A}{r_{iA}}\right) + \sum_{i}^{N}\sum_{j>i}^{N}\left(\frac{1}{r_{ij}}\right). \tag{3.6}$$

In order to obtain an accurate total energy for the system, the classical Coulombic repulsion of the nuclei is added to the quantum electronic energy.

$$E^{total} = E^{elec} + \sum_{A}^{M}\sum_{B>A}^{M}\left(\frac{Z_A Z_B}{R_{AB}}\right) \tag{3.7}$$

3.1.2 The orbital approximation

The electronic wavefunction, $\Psi(\mathbf{r})$, is a $3N$-dimensional function of the positions of all electrons in the system, where N is the number of electrons. To simplify the calculation of $\Psi(\mathbf{r})$, it is often approximated by products of one-electron functions called molecular orbitals (MO), Ψ_i. Molecular orbitals are functions of the Cartesian coordinates of a single electron, i.e. $\Psi_i(x_i, y_i, z_i)$. The square modulus, $|\Psi_i(x_i, y_i, z_i)|^2$, of a MO is the probability distribution of its electron in space. The Cartesian coordinates alone, however, do not provide a complete description of an electron. For that, the electron spin must also be included. An electron's spin is denoted α or β, often called spin up or spin down, respectively. A complete description of an electron can be given using spin orbitals, $\chi_i(\mathbf{r}, s)$, which are the product of the spatial molecular orbital, Ψ_i, and a spin description, $\alpha(s)$ or $\beta(s)$.

$$\chi_i(\mathbf{r},s) = \Psi_i(x_i,y_i,z_i)\alpha(s) \text{ or } \chi_i(\mathbf{r},s) = \Psi_i(x_i,y_i,z_i)\beta(s) \tag{3.8}$$

It is important to combine the spin orbitals, which form the total electronic wavefunction, in the proper way to maintain antisymmetry. For a wavefunction to be antisymmetric, interchanging the coordinates of any two electrons must result in a change in the sign of the overall wavefunction.

$$\Psi(1,2,\ldots,i,\ldots,j,\ldots,N-1,N) = -\Psi(1,2,\ldots,j,\ldots,i,\ldots,N-1,N) \tag{3.9}$$

Since electrons are indistinguishable, from the uncertainty principle, the total electron distribution, $|\Psi(\mathbf{r})|^2$, typically represented by $\rho(\mathbf{r})$, should not change after interchanging any two electrons.

$$|\Psi(1,2,\ldots,i,\ldots,j,\ldots,N-1,N)|^2 = |-\Psi(1,2,\ldots,j,\ldots,i,\ldots,N-1,N)|^2 \tag{3.10}$$

As a result of wavefunction antisymmetry, electrons of the same spin are forced away from each other and therefore the Pauli exclusion principle, which states that no two electrons can occupy the same spin orbital, is maintained.

In 1929, Slater[91,92] found that a determinant of the form shown below, known as a Slater determinant, satisfies antisymmetry requirements and can be used to approximate a total electronic wavefunction composed of N spin orbitals representing N electrons.

$$\Psi(1,2,\ldots,N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{vmatrix} \tag{3.11}$$

In this determinant, all the elements in a given column represent one spin orbital and all the elements in a given row represent one electron. Since interchanging any two rows, corresponding to the interchange of the coordinates of any two electrons, will change the sign of the determinant, the antisymmetry requirements are met. Additionally, if any two rows are equal, corresponding to two electrons occupying the

same spin orbital, the determinant will vanish, satisfying the Pauli exclusion principle.

3.1.3 Linear combination of atomic orbitals

As discussed above, the total molecular wavefunction is approximated as an antisymmetrized product of molecular orbitals, $\Psi_i(\mathbf{r})$. These MOs are further approximated as linear combinations of atom-centered functions, $\phi(\mathbf{r})$, called atomic orbitals (AO).

$$\Psi_i(\mathbf{r}) = \sum_k^K c_{ki}\phi_k(\mathbf{r}) \tag{3.12}$$

In the above expression, the MO is a linear combination of K AOs, also referred to as basis functions, and their associated expansion coefficients, c_{ki}. The set of K basis functions used to produce a molecular wavefunction is known as a basis set. This approach of using AOs to produce the MOs is called linear combination of atomic orbitals (LCAO). The quality of the MOs produced using this procedure is directly related to the quality of the basis set used. By using a basis set with the right combination of orbitals and appropriate orbital flexibility for the system, the best MOs can be produced.

3.2 Basis Sets

To simplify quantum chemical calculations, basis sets of atomic orbitals are used as a starting point for producing the molecular orbitals and the molecular wavefunction. All basis functions are composed of three parts: an angular component, a radial component, and a normalization constant. The angular component describes the shape of the orbital and is used to define the type of orbital (e.g. s, p, d, etc.) that is being modelled. The radial component describes the extent of the orbital in space (e.g. $1s$, $2s$, etc.). The normalization constant is used to ensure that the orbital described, $\phi(\mathbf{r})$, integrates to the proper value. Over time, basis sets have evolved from minimal basis sets composed only of one basis function for each active free atomic orbital to the large basis sets used today that include many more basis

functions on each atom to produce more accurate wavefunctions by allowing more flexibility of the active atomic orbitals.

3.2.1 Slater-type orbitals

The first mathematical functions used to model atomic orbitals were proposed by Slater in 1930.[93] These orbitals, known as Slater-type orbitals (STO), are modelled using exponential functions for the radial component, which produces a cusp at the nucleus. The general form of a STO centred on atom a is shown below,

$$\phi(\mathbf{r}) = N r_a^{(n-1)} e^{-\zeta r_a} Y_l^m(\theta_a, \phi_a) \tag{3.13}$$

where N is the normalization constant, $r_a^{(n-1)} e^{-\zeta r_a}$ is the radial component, and $Y_l^m(\theta_a, \phi_a)$ is the angular component. The parameter ζ is the orbital exponent and is used to fine tune the shape and behaviour of the orbital. The principal, angular, and magnetic quantum numbers are n, l, and m, respectively.

A small number of well-chosen STOs can be used to accurately model any atomic orbital, and linear combinations of these can produce very accurate wavefunctions. Unfortunately, performing the multi-centre integrations required for any calculation (to be discussed later) using STOs becomes very time consuming and reaches the point of being unfeasible for systems larger than diatomic molecules due to the cusp at the nucleus.

3.2.2 Gaussian-type orbitals

In 1950, Boys[94] proposed using Gaussian functions, rather than exponential functions, to describe the radial component of atomic orbitals to speed up the evaluation of integrals. A Cartesian Gaussian-type orbital (GTO) centred on atom a is defined as

$$g_{ijk} = N x_a^i y_a^j z_a^k e^{-\zeta r_a^2} \tag{3.14}$$

where x_a, y_a, and z_a are Cartesian coordinates and $i+j+k=l$, the angular quantum number.

Gaussian functions are smooth and continuous at the nucleus, which greatly simplifies the calculations required for their integration, but this does not accurately model the behaviour of real atomic orbitals at the nucleus. In order to overcome this, each atomic orbital, $\phi(\mathbf{r})$, is represented as the sum of several GTOs, known as primitive Gaussians, each with a different orbital exponent, ζ.

$$\phi(\mathbf{r}) = \sum_p d_p g_p \tag{3.15}$$

The factor d_p is a contraction coefficient. The basis functions produced in this way are known as contracted GTOs (CGTO) and closely resemble STOs without actually having a cusp at the nucleus. Ideally, the ζ_p and d_p values would be varied in every calculation to produce the best functions for the system, but this proves to be too computationally expensive to be practical, so they are chosen to best fit a given STO during basis set construction and fixed thereafter.

3.2.3 Basis set construction

The simplest, and smallest, basis set that can be constructed is a minimal basis set. In a minimal basis set, one STO or CGTO is used to represent each atomic orbital present in the system and only the occupied atomic orbitals are included in the calculation of the molecular orbitals. This means that hydrogen and helium are described using a single basis function representing a $1s$ orbital, while row two elements have a $1s$, a $2s$, and three $2p$ ($2p_x, 2p_y, 2p_z$) orbitals for a total of five basis functions and so on.

Using a single basis function to describe each atomic orbital produces very poor results, due to the inflexibility of their radial components. Therefore, it is common to double the number of basis functions used, called a double-zeta basis set. In this case, each occupied atomic orbital is described by two STOs or CGTOs, with different orbital exponents. Larger basis sets can be made by increasing the number of basis

functions used to describe each atomic orbital, producing triple-zeta, quadruple-zeta, etc. basis sets.

It has been found that it is primarily the valence orbitals that require more than one basis function, while a minimal basis set describes the core orbitals well, since they do not require as much radial flexibility. It is therefore possible to reduce the number of basis functions required in a calculation by designing split-valence basis sets that use one basis function to describe each core orbital and multiple basis functions to describe each valence orbital. Like the larger n-zeta basis sets described above, split-valence basis sets can use two, three or more basis functions to describe valence orbitals, producing valence double-zeta, valence triple-zeta, and larger basis sets.

When placed in a molecular environment, atomic orbitals are typically distorted to reflect the greater electron density in internuclear regions compared to other parts of the molecule. To better model this, and to allow a greater degree of angular flexibility, basis functions one angular quantum number higher than the highest occupied orbitals in a given atom (e.g. d-type orbitals on a carbon atom), called polarization functions, are added to the basis set. As with the other basis sets discussed, it is often desirable to add even more basis functions to a system, in which case multiple sets of polarization functions (e.g. two d-type orbitals with different orbital exponents) or functions two angular quantum numbers higher than the occupied orbitals (e.g. both d and f-type orbitals on a carbon atom) are added to the basis set. Since heavier atoms typically experience more polarization than hydrogen, it is common to have a basis set that includes polarization functions only for all non-hydrogen atoms to moderate the basis set size. Basis sets including polarization functions for hydrogen can be used in cases where hydrogen polarization is important, such as hydrogen bonded complexes.

The basis functions described above do a reasonable of job representing normal covalently bound molecules, in which the electron density far from the nucleus is very small, but for some systems, like anions, atoms with lone pairs, and in non-covalent interactions, this is not the case. In order to better model the increased electron density far from the nucleus, another type of basis function, called a diffuse function,

is added to the basis set. Diffuse functions are of the same type as the valence orbitals, but have very small orbital exponents, giving very large radial distributions allowing them to better model regions of space far from the nucleus. As with polarization functions, sometimes it is desirable to only include diffuse functions on non-hydrogen atoms, so there are basis sets that only include diffuse functions on heavy atoms.

3.2.4 Commonly used basis sets

While many different basis sets have been developed, the two most widely used are the Pople[95] and Dunning[96,97] basis sets. Both are split-valence basis sets that use contracted Gaussian functions to represent the atomic orbitals, however each has its own advantages.

The Pople basis sets, developed by Pople and coworkers over many years, include both minimal and split valence basis sets. The nomenclature used to describe the minimal basis sets composed of CGTOs is STO-nG,[98] where n is the number of primitive GTOs used to produce each atomic orbital. The split valence basis sets follow the general formula k-lmnG for their naming, where k represents the number of primitive GTOs used for the core orbitals, l the number of primitive GTOs for the first set of valence orbitals, m for the second set, and so on. There have been valence double-zeta and valence triple-zeta basis sets of this style developed. Following this naming scheme, the 3-21G basis set uses three primitive GTOs to describe each core orbital and two primitive GTOs for one set of valence orbitals and one primitive GTO for the other set. Many Pople basis sets also include polarization functions, which are denoted by naming the orbital types used after the rest of the basis set definition (e.g. 6-311G(2df,p) is a valence triple-zeta basis set that includes two sets of d and one set of f orbitals on all non-hydrogen atoms and one set of p orbitals on all hyrodgens). Diffuse functions are added by including one 'plus sign' after the valence orbital description to only include diffuse functions on non-hydrogen atoms and two 'plus signs' for diffuse functions on all atoms (e.g. 3-21+G is a valence double-zeta basis set that includes diffuse functions on all non-hydrogen atoms).

The Dunning basis sets were developed for use in high level *ab initio* calculations because they smoothly converge to the basis set limit (the point where introducing additional basis functions does not improve the quality of a calculation). These basis sets use the naming convention cc-pVnZ, where n indicates the number of basis functions used to represent each valence atomic orbital. The 'p' indicates that polarization functions are included and 'cc' stands for correlation consistent. Diffuse functions can be included by adding the keyword 'aug' to the beginning of the basis set name (e.g. aug-cc-pVTZ is a valence triple-zeta basis set with polarization and diffuse functions). While the Pople basis sets give the user more control over some basis set features, such as the number of polarization functions, the Dunning basis sets are favoured for high level calculations because much larger basis sets, up to valence sextuple-zeta, have been developed and these were specifically designed to introduce similar changes to the correlation energy (to be discussed later) with each increase in basis set size.

3.3 Hartree-Fock Theory

The simplest, and one of the earliest, *ab initio* electronic structure methods is the Hartree-Fock (HF) method. Many of the more sophisticated *ab initio* methods use HF as their starting point. Although this method is no longer widely used, a good understanding of it will aid in the understanding of other methods discussed later. An account of the development of modern HF theory with extensive references has been published by Slater[99] and the reader is directed there for complete details.

Hartree proposed a method for solving the many-body Schrödinger equation for atoms in 1927,[100] one year after the publication of Schrödinger's equation, which uses the mean field approximation to solve this many-body problem. Specifically, he assumed that each electron moves in an electric field generated from the nuclei and the spherically averaged charge distribution of all other electrons. Shortly thereafter, Gaunt[101] and Slater[102] showed that Hartree's proposal could be solved by using the product of one electron functions associated with each electron, leading to the use of

a Slater determinant, (3.11), to determine the many electron wavefunction. It was soon realized by Slater[103] and Fock[104] that to accurately solve the wavefunction using the product of one electron functions, proposed by Gaunt and Slater, the variational principle, discussed below, should be applied. Fock's proposal to apply the variational principle to Hartree's equations did not use a determinant method, but rather more complex permutation methods. The modern determinant form of the HF method was formulated by Hartree in 1935.[105] The final step in the development of the modern HF method was the use of atomic basis functions to approximate the one electron functions of the determinant[93,106] as opposed to the original method of determining the one electron functions through numerical integration of a radial differential equation.

3.3.1 The variational principle

Slater[103] and Fock[104] proposed the use of the variational principle to solve the wavefunction using the framework established by Hartree. The variational principle states that any approximate wavefunction, Φ, will have an energy greater than or equal to the energy of the exact wavefunction Ψ.

$$\left\langle \Phi|\hat{H}|\Phi \right\rangle \geq \left\langle \Psi|\hat{H}|\Psi \right\rangle \tag{3.16}$$

This provides both a measure of the quality of an approximate wavefunction (if the exact wavefunction is known) and a method for finding the best approximate wavefunction. If the exact wavefunction is not known, the variational principle can still be applied by testing if Φ satisfies the virial theorem, which states that kinetic energy in a quantum mechanical system is twice the total potential energy.

$$2\left\langle \Psi|\hat{T}^{elec}|\Psi \right\rangle = \left\langle \Psi|\hat{V}^{elec}|\Psi \right\rangle \tag{3.17}$$

An exact wavefunction will satisfy this relationship, so the closer an approximate wavefunction is to satisfying the virial theorem the closer it is to the exact wavefunction.

Since the wavefunction with the lowest energy possible will be the best approximation to the exact wavefunction, the components of the wavefunction can be varied to minimize the energy. The wavefunction is found through the Slater determinant (3.11), which is composed of one electron spin orbitals (3.8). These spin orbitals are in turn defined through the linear combination of atomic orbitals, or basis functions (3.12). The expansion coefficients, c_{ki}, are the one part of this formulation that is not held constant; they are varied to produce the lowest energy wavefunction in order to give the best description of the MOs that is possible with the basis set used.

3.3.2 Roothaan-Hall equations

To make the solution of the Hartree-Fock method more computationally amenable when using atomic basis functions, Roothaan[107] and Hall[108] separately developed what are now known as the Roothaan-Hall equations in 1951. Versions of these equations exist for both open and closed-shell systems, but for the sake of simplicity only the closed-shell form will be discussed here. They generalize the Hartree-Fock method to a linear algebra problem, written below in matrix form.

$$\mathbf{FC} = \mathbf{SC}\varepsilon \tag{3.18}$$

Where $\mathbf{F}$ is the Fock matrix, whose elements are derived from the Hamiltonian operator (discussed in greater detail below), $\mathbf{C}$ is the orbital coefficient matrix constructed from the coefficients of the molecular orbitals,

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1i} \\ c_{21} & c_{22} & \cdots & c_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{ii} \end{pmatrix} \tag{3.19}$$

S is the overlap matrix, whose elements are the overlap between pairs of basis functions.

$$S_{ij} = \langle \phi_i | \phi_j \rangle \tag{3.20}$$

Since the basis functions are normalized, the diagonal elements of the overlap matrix, $S_{i=j}$, will equal 1. However, the basis functions in a standard basis set are not orthogonal, so some degree of overlap will exist between any two different basis functions. Therefore, the remaining matrix elements will be nonzero. Lastly, ε is a diagonal matrix of the one-electron orbital energies.

The Fock matrix elements are composed of a slightly altered Hamiltonian operator.

$$F_{ij} = H_{ij}^{core} + \sum_{k=1}^{K} \sum_{l=1}^{K} P_{kl} \left[(ij|kl) - \frac{1}{2} (ik|jl) \right] \tag{3.21}$$

The Hamiltonian used in construction of the Fock matrix, H_{ij}^{core}, is derived from the electronic Hamiltonian (3.6), but neglects the electron-electron repulsion term. Therefore it is called the core Hamiltonian, since it only accounts for kinetic energy and the electron-nuclear attraction. The core Hamiltonian, in terms of the coordinates of electron one, is:

$$\hat{H}^{core}(1) = -\frac{1}{2}\nabla_1^2 - \sum_{A}^{M} \frac{Z_A}{r_{1A}} \tag{3.22}$$

The matrix elements, H_{ij}^{core}, are then obtained as an expectation value of the one-electron core Hamiltonian,

$$H_{ij}^{core} = \int \phi_i(1) \hat{H}^{core}(1) \phi_j(1) \mathrm{d}\mathbf{r}_1 \tag{3.23}$$

where $\phi_i(1)$ and $\phi_j(1)$ are basis functions depending on the coordinates of electron 1.

The electron-electron repulsion term neglected in the core Hamiltonian is reintroduced as the second term of the Fock operator. This term gives the Coulombic repulsion between electrons and accounts for electron exchange. The summations are

over all the basis functions, K, and the following shorthand is used,

$$(ij|kl) = \iint \frac{\phi_i(1)\phi_j(1)\phi_k(2)\phi_l(2)}{r_{12}} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{3.24}$$

P_{kl} is an element of the charge density matrix, $\mathbf{P}$. This matrix element is a function of the orbital coefficients of the occupied MOs.

$$P_{kl} = 2\sum_{x}^{N/2} c_{xk}c_{xl} \tag{3.25}$$

3.3.3 Self-consistent field procedure

At this point, a problem with solving the Roothaan-Hall equations becomes obvious. The construction of the charge density matrix relies on the orbital coefficients, c_{xy}, which are solved by the Roothaan-Hall equations. In order to solve the orbital coefficients, these equations are solved iteratively using the self-consistent field (SCF) procedure to obtain the best values of the coefficients in order to produce the lowest energy wavefunction.

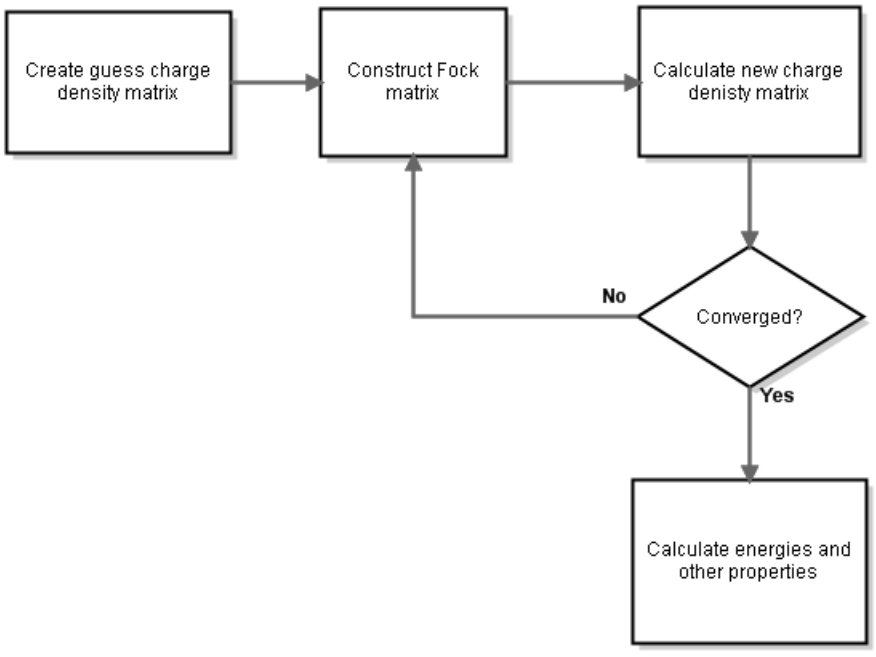


Figure 3.1: Flowchart showing the SCF procedure.

The SCF procedure, outlined in figure 3.1, start with a reasonable initial guess for the charge density and overlap matrices. From these, the Fock matrix is constructed and used to calculate new coefficient, **C**, and charge density, **P**, matrices. This process continues until the new charge density matrix equals the previous one, within the limits of the convergence criteria defined by the user, at which point the matrix has converged, or is self-consistent. From the self-consistent charge density matrix, the orbital energies, ε, and optimized molecular orbital coefficients, **C**, are obtained.

Hartree-Fock theory is a valuable tool for computational chemists, but as mentioned previously, it is not widely used anymore due to the limitations of the method. The HF method uses the mean field approximation, which assumes each electron is moving in the average electric field generated by the remaining electrons and the nuclei, to allow for easy calculation of many-electron systems. However, the movements of the electrons are correlated, meaning that the position of any one electron affects the positions of the other electrons. HF theory only accounts for the correlation between electrons of parallel spin, but neglects the correlation between electrons of opposite spin, so other methods have been developed which build upon this method to account for the opposite spin electron correlation.

3.4 Multi-Configuration Methods

The failure of Hartree-Fock theory to account for electron correlation is its largest shortcoming. While the energies associated with correlation between opposite-spin electrons are small, about 0.5–1% of the electronic energy,[87,88] they are important for describing many chemical phenomena. The correlation energy is defined as the difference between the exact energy and the HF energy.

$$E_{corr} = E_{exact} - E_{HF} \tag{3.26}$$

The mean field approximation assumes each electron is moving through the average electric field generated by the nuclei and other electrons without individual

electrons affecting the motion of the other electrons. This simplifies the calculations considerably and allows the entire wavefunction to be described using a single determinant; however, in order to account for the correlated motions of electrons, more than one determinant is required. Normally a HF wavefunction will be calculated first, since it provides a good starting point. Other wavefunctions, representing various excited states of the system, will be added to the HF result,

$$\Psi = a_0\Phi_{SCF} + \sum_{i=1} a_i\Phi_i \tag{3.27}$$

where the first term uses the ground state Hartree-Fock wavefunction, Φ_{SCF}, and the a_0 coefficient is normally close to one. The other wavefunctions, Φ_i, are excited states of the Hartree-Fock wavefunction and their coefficients, a_i, can be determined through various means, which is the major difference between multi-configuration methods.

A solution of the ground state HF wavefunction with N electrons and M basis functions will produce $\frac{N}{2}$ occupied MOs and $M-\frac{N}{2}$ unoccupied, or virtual, MOs. The excited states can be obtained by swapping one or more of the occupied MOs in the ground state determinant with one or more virtual MOs to produce a determinant representing an excited state wavefunction. If only one virtual orbital is used, this is referred to as a single excitation, if two are used, it is a double excitation and so on, up to a potential N excitations. The number of excitations used to produce the final wavefunction will affect the quality of the wavefunction, with more excitations producing a better wavefunction. However, since each excitation to the computational cost, the number of excitations is normally limited. For consistency, all excitations of a particular type (e.g. single, double, etc.) are normally included.

Many multi-configurational methods have been proposed, which can be loosely classified into two categories: variational and peturbative methods. Since most of these methods are outside the scope of this project, only the configuration interaction and Møller-Plesset perturbation methods will be described.

3.4.1 Configuration interaction

The configuration interaction (CI) method is the simplest of the multi-configuration methods and is based on the variational principle, like HF theory. This method uses a linear combination of determinants with their expansion coefficients varied to produce the lowest energy wavefunction.

$$\Psi_{CI} = a_0\Phi_{SCF} + \sum_S a_S\Phi_S + \sum_D a_D\Phi_D + \sum_T a_T\Phi_T + \cdots + \sum_N a_N\Phi_N \tag{3.28}$$

where S, D, and T indicate single, double and triple excitations, respectively, and N indicates the promotion of every electron to a virtual orbital. The MOs used for the excited states are generated from the HF wavefunction and held fixed to simplify the calculation. The above equation represents the full CI method, where all possible excitations are considered. Performing a full CI calculation on a given system will produce the exact wavefunction and energy, within the constraints of the basis set used and other approximations employed.

Since the number of configurations grows rapidly with the number of electrons and basis set size, it is common to use truncated CI methods, which only consider some of the possible excitations. The simplest possible truncated method would use only single excitations (CIS), considering every possible electron configuration with one electron promoted to a virtual orbital.

$$\Psi_{CIS} = a_0\Phi_{SCF} + \sum_i^{occ}\sum_r^{virt} a_i^r\Phi_i^r \tag{3.29}$$

The CIS wavefunction, Ψ_{CIS}, is not an improvement upon the ground state HF wavefunction, however, since singly excited determinants do not interact with the ground state determinant, from Brillouin's theorem.[109] The lowest level CI method that improves on the HF wavefunction considers all single and double excitations (CISD).

$$\Psi_{CISD} = a_0\Phi_{SCF} + \sum_i^{occ}\sum_r^{virt} a_i^r\Phi_i^r + \sum_i^{occ}\sum_{j>i}^{occ}\sum_r^{virt}\sum_{s>r}^{virt} a_{ij}^{rs}\Phi_{ij}^{rs} \tag{3.30}$$

The CI wavefunction can be improved by including triple excitations (CISDT), quadruple excitations (CISDTQ), and so on up to the full CI wavefunction. Due to computational limitations, it is not currently common to include anything above double excitations, although it is possible to introduce triple excitations for very small systems.

Constraining the CI wavefunction to a limited number of configurations is often not enough to reduce the size of the calculations, so the frozen-core approximation[95] is often used as well. In this approach, the electrons in core orbitals (e.g. the $1s$ orbital for second-row atoms and the $1s$, $2s$, and $2p$ orbitals for third-row atoms) are excluded from the excitations. This does have an effect on the total energy of the system, but it has been found that it is essentially constant for similar systems and therefore relative energies will not be affected.

3.4.2 Møller-Plesset perturbation theory

Perturbative methods can be used as an alternative to the variational method employed by CI to add electron correlation to a Hartree-Fock wavefunction. The motivation behind such treatments is an assumption that the desired solution is only slightly different than the solution that has already been obtained. It was proposed by Møller and Plesset in 1934[110] that electron correlation could be viewed as a perturbation to the HF Hamiltonian operator. In this formulation (hereafter referred to as MPPT), the total Hamiltonian operator, $\hat{H}$, is decomposed into two parts: the Hamiltonian operator from an SCF calculation, called the zeroth-order Hamiltonian, $\hat{H}_0$, which has known eigenfunctions and eigenvalues, and a perturbation that includes the electron correlation, $\hat{V}$.

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{3.31}$$

In order to improve the eigenfunctions ($\Psi_i^{(0)}$) and eigenvalues ($E_i^{(0)}$) of the SCF Hamiltonian, a perturbation parameter, λ, which has values between 0 and 1, is added to the perturbation term.

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{3.32}$$

The eigenfunctions and eigenvalues of the total Hamiltonian operator, $\hat{H}$, can now be represented as Taylor expansions containing the perturbation parameter.

$$\Psi_i = \Psi_i^{(0)} + \lambda\Psi_i^{(1)} + \lambda^2\Psi_i^{(2)} + \cdots \tag{3.33}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots \tag{3.34}$$

The higher order terms in these expansions are treated as corrections to the original HF system. The energy corrections are evaluated using the following expressions

$$E_i^{(0)} = \langle\Psi_i^{(0)}|\hat{H}_0|\Psi_i^{(0)}\rangle \tag{3.35}$$

$$E_i^{(1)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(0)}\rangle \tag{3.36}$$

$$E_i^{(2)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(1)}\rangle \tag{3.37}$$

$$E_i^{(3)} = \langle\Psi_i^{(0)}|\hat{V}|\Psi_i^{(2)}\rangle \tag{3.38}$$

The Hartree-Fock ground state energy is equal to the sum of the two lowest-order energies.

$$E_0 = E_0^{(0)} + E_0^{(1)} \tag{3.39}$$

Therefore, in order for MPPT to improve upon the HF energy, at least the second order perturbation must be applied. The level of perturbation applied is denoted in the general name as MPn, where n is the highest order perturbation applied in a calculation.

In order to obtain a perturbation energy of a given order, it can be seen above that it is necessary to calculate the wavefunction of one order lower, e.g. for a second-order perturbation, the first-order wavefunction, $\Psi_0^{(1)}$, is required, which can be obtained as a linear combination of solutions to the zeroth-order Hamiltonian.

$$\Psi_0^{(1)} = \sum_j c_j^{(1)}\Psi_j^{(0)} \tag{3.40}$$

These wavefunctions include excitations, performed in a manner similar to CI described earlier, up to a level defined by the highest order of perturbation, e.g. MP4 includes single, double, triple, and quadruple excitations. For the MP2 method, the second-order energy, $E_0^{(2)}$, is defined by

$$E_0^{(2)} = \sum_{a<b,r<s} \frac{\left| \left\langle \Psi_0 \left| \sum_{i<j} \frac{1}{r_{ij}} \right| \Psi_{ab}^{rs} \right\rangle \right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{3.41}$$

The first summation is over the set of occupied orbitals a and b and virtual orbitals r and s and the second is over the electrons i and j. Ψ_0 refers to the ground state HF wavefunction and Ψ_{ab}^{rs} is a series of doubly excited determinants. The perturbation operator, $\hat{V}$, is given by the electron-electron repulsion, corrected to eliminate the mean-field introduced by HF. The ε in the denominator corresponds to the energy of the orbitals, a, b, r, and s. The higher order energy expressions are significantly more complex and will not be discussed here.

3.5 Density-Functional Theory

Density-functional theory (DFT) is an alternative electronic structure theory, which, like the multi-configurational methods discussed in the previous section, explicitly accounts for electron correlation. It has one major advantage over multi-configurational methods: it is based on a single determinant like HF, so it is about as computationally expensive as HF. Due to this, in the two decades since accurate DFT methods were introduced, it has become the most widely used electronic structure method of computational chemists. A complete description of the method will not be given in this document; for that the reader is directed to a textbook reference[111] or any of the many review articles published on the subject.

3.5.1 Hohenberg-Kohn Theorem

The modern study of DFT began with Hohenberg and Kohn[112] when they proposed that the properties of a multi-electron system can be uniquely determined by its electron density, $\rho(\mathbf{r})$. The conceptual framework for this originated with the Thomas-Fermi model[113,114] of an electron gas. This concept can be explained by starting with an external potential, *e.g.* the nuclear potential of a molecular system, $v_{nuc}(\mathbf{r})$. Given N electrons, this potential generates a ground state wavefunction, Ψ, through a Hamiltonian the same as (3.6),

$$\hat{H}_v = \sum_i^N (-\frac{1}{2}\nabla_i^2 + v_{nuc}) + \frac{1}{2}\sum_i^N \sum_{j\neq i}^N \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \tag{3.42}$$

Recall that $\rho(\mathbf{r})$ is generated from Ψ, so the external potential can be connected to the electron density through the wavefunction.

$$v \Rightarrow \Psi \Rightarrow \rho(\mathbf{r}) \tag{3.43}$$

Hohenberg and Kohn proved that this relationship is reversible, so it can be said that,

$$\rho(\mathbf{r}) \Rightarrow v_\rho \Rightarrow \Psi_\rho. \tag{3.44}$$

Because of this mapping, in principle, an energy expression can be written that depends exclusively on the external potential and the electron density,

$$E_v(\rho) = \langle \Psi_\rho \left| \hat{H}_v \right| \Psi_\rho \rangle = T(\rho) + V_{ee}(\rho) + \int v\rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \tag{3.45}$$

which can be solved variationally using the electron density as the trial function. Unfortunately, although Hohenberg and Kohn proved that the energy functionals $T(\rho)$ and $V_{ee}(\rho)$ exist, they provided no simple expressions for them.

3.5.2 Kohn-Sham Theory

One year after the proposal of the Hohenberg-Kohn theorem, Kohn and Sham[115] introduced a framework to usefully approximate the kinetic energy functional, $T(\rho)$. It uses the concept of a system of N non-interacting electrons, meaning there are no electron-electron repulsions, which has the same density as the real N-electron density, $\rho(\mathbf{r})$. For a closed-shell system, this non-interacting density is

$$\rho(\mathbf{r}) = 2\sum_{i}^{N/2} \psi_i^2(\mathbf{r}) \tag{3.46}$$

with orbitals satisfying a simplified Schrödinger equation, whose Hamiltonian does not include an electron-electron repulsion term,

$$-\frac{1}{2}\nabla^2\psi_i(\mathbf{r}) + v_0\psi_i = \varepsilon_i\psi_i(\mathbf{r}) \tag{3.47}$$

where v_0 is the external potential that produced the non-interacting density.

Within Kohn-Sham theory, the kinetic energy of the non-interacting system, $T_0(\rho)$, is assumed to be approximately equal to the kinetic energy of the interacting system, $T(\rho)$. The non-interacting kinetic energy can be given by,

$$T_0(\rho) = -\frac{1}{2}\sum_{i} 2\int \psi_i(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\mathrm{d}\mathbf{r}. \tag{3.48}$$

Additionally, the electron-electron repulsion energy can be approximated as the classical Coulomb repulsion energy of the density, $J(\rho)$,

$$J(\rho) = \frac{1}{2}\iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2. \tag{3.49}$$

These two approximations produce the relationship,

$$T(\rho) + V_{ee}(\rho) = T_0(\rho) + J(\rho) + E_{XC}(\rho) \tag{3.50}$$

where $E_{XC}(\rho)$ is a term that corrects for the approximations in $T_0(\rho)$ and $J(\rho)$ and is called the exchange-correlation energy. These energies can be substituted into (3.45),

$$E_v(\rho) = T_0(\rho) + J(\rho) + E_{XC}(\rho) + \int v\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{3.51}$$

giving an energy equation with only one term for which an exact expression is not known, $E_{XC}(\rho)$. Therefore, in order to use this method, approximations that model this term must be found.

3.5.3 The exchange-correlation hole

In order to better understand $E_{XC}(\rho)$, it is useful to connect the non-interacting density of Kohn-Sham theory with the real density, known as the adiabatic connection.[116–119] These densities are connected by visualizing a continuum of electronic systems with the degree of electron-electron repulsion varied from the non-interacting system to the fully interacting system. The Coulomb repulsion energy for a partially interacting system is,

$$\lambda G = \frac{\lambda}{2}\sum_i \sum_{j\neq i} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{3.52}$$

where λ represents the coupling strength, with $\lambda = 0$ as the non-interacting system and $\lambda = 1$ the fully interacting system. Through this, the exchange-correlation energy can be rigorously expressed as,

$$E_{XC}(\rho) = \frac{1}{2}\iint \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}[\Pi_{avg}(\mathbf{r}_1, \mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)]\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \tag{3.53}$$

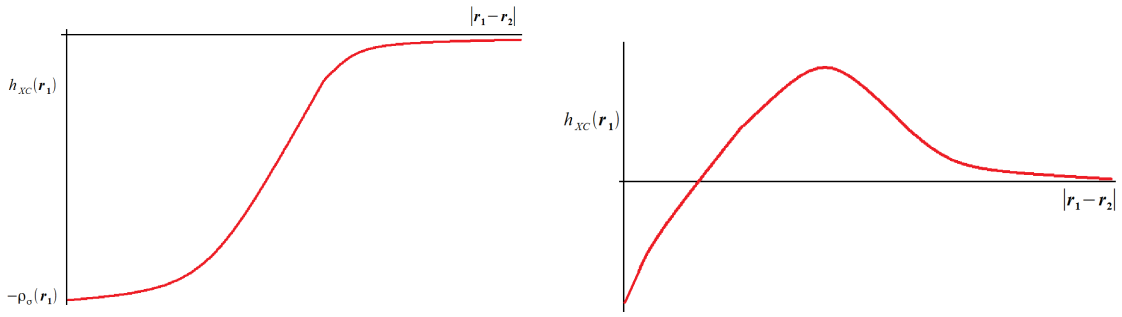
where $\Pi_{avg}(\mathbf{r}_1, \mathbf{r}_2)$ is the pair density averaged from $\lambda = 0$–1. From this definition of $E_{XC}(\rho)$, a conceptual "hole" can be defined as

$$h_{XC}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\Pi_{avg}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - \rho(\mathbf{r}_2). \tag{3.54}$$

The first term in this expression is the λ-averaged conditional pair density, representing the probability of finding an electron at $\mathbf{r}_2$ if one is known to exist at $\mathbf{r}_1$.

The hole is therefore the difference in conditional probabilities between the coupling strength averaged correlated system and the fully uncorrelated density. More conceptually, the hole represents the effects of electron-electron interactions: if an electron is found at $\mathbf{r}_1$, the probability of finding one at $\mathbf{r}_2$ is reduced. Substituted back into the expression for $E_{XC}(\rho)$, the energy is an integration over h_{XC},

$$E_{XC}(\rho) = \frac{1}{2}\iint \frac{\rho(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} h_{XC}(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2. \tag{3.55}$$



(a) (b)

Figure 3.2: General forms of the parallel (a) and opposite (b) spin exchange correlation holes.

Due to the Pauli exclusion principle, electrons with parallel spin interact differently than ones with opposite spins. Therefore, it is useful to decompose the exchange-correlation energy and hole into parallel and opposite spin components. The general forms of the two spin decomposed holes is shown in figure 3.2. Although there are many different models for these holes, they share the same constraints. The parallel spin hole (figure 3.2a) has a known depth at $\mathbf{r}_2 = \mathbf{r}_1$ and normalization,

$$h_{XC}^{\sigma\sigma}(\mathbf{r}_1, \mathbf{r}_2 = \mathbf{r}_1) = -\rho_\sigma(\mathbf{r}_1) \tag{3.56}$$

$$\int h_{XC}^{\sigma\sigma}(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r}_2 = -1 \tag{3.57}$$

while the opposite spin hole (figure 3.2b) has a known normalization and its value at

$\mathbf{r}_2 = \mathbf{r}_1$ is known to have a cusp, but the exact value cannot be easily determined.

$$\int h_{XC}^{\sigma\sigma'}(\mathbf{r}_1, \mathbf{r}_2)\, \mathrm{d}\mathbf{r}_2 = 0 \tag{3.58}$$

Where σ and σ' represent the spin indices of the electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$. The spin-dependant exchange correlation holes can be further decomposed into their exchange and correlation components – how this is done will not be shown here.

3.5.4 Density-Functional Approximation

Since the exact form of the exchange-correlation hole in DFT is not known, approximations to the hole must be found in order to make DFT useful in chemistry. The earliest such approximation to provide reasonably accurate results is known as the local-density approximation (LDA).[115] This method approximates h_{XC} in a chemical system as the hole from a uniform electron gas, h_{XC}^{UEG}. The uniform electron gas hole depends only on the local density and interelectron separation, $|\mathbf{r}_1 - \mathbf{r}_2|$. The uniform electron gas can be parameterized from Monte-Carlo simulations or similar calculations. The LDA hole can be substituted into the exchange-correlation energy expression (3.55),

$$E_{XC}^{LDA}(\rho) = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} h_{XC}^{UEG}(\rho(\mathbf{r}_1), |\mathbf{r}_1 - \mathbf{r}_2|)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2. \tag{3.59}$$

The LDA method can be extended to the spin-dependant holes introduced in figure 3.2 by relying on spin-dependant uniform electron gases instead, a method known as the local-spin-density approximation (LSDA).

Atomization energies calculated using LSDA overestimate bond strengths by about 160 kJ/mol.[111] This is about half the magnitude of the errors found using HF, although HF tends to underestimate bond energies. In order to correct this, functionals that depend not only on the electron density, but also the gradient of the density have been proposed, collectively known as the generalized gradient approximation (GGA).[120,121] The E_{XC} depending on the gradient of the density is normally

taken as a correction factor to the LSDA E_{XC}, giving an exchange correlation energy expression of the form,

$$E_{XC} = E_{XC}^{LSDA}(\rho_\sigma) + bE_{XC}^{GGA}(\rho_\sigma, \nabla\rho_\sigma) \tag{3.60}$$

where the parameter b is normally fit to experimental data to provide the best results. GGA functionals improve on the LSDA error by about an order of magnitude. The GGA exchange correlation energy can be further improved by adding additional terms in the gradient expansion, so it relies on the density, the gradient of the density, and the Laplacian of the density, and the kinetic energy density, τ_σ $E_{XC}(\rho_\sigma, \nabla\rho_\sigma, \nabla^2\rho_\sigma, \tau_\sigma)$. Methods of this type are called meta-GGA functionals.

Recall from (3.54) that h_{XC} is defined from the coupling strength averaged pair density. At $\lambda = 0$, in GGA methods h_X (the exchange component of h_{XC}) is too localized, whereas the delocalized exchange hole from Hartree-Fock theory is accurate. In order to correct for this, it is reasonable to form a hybrid DFT method that incorporates a fraction of the HF exchange energy into E_{XC} to improve on the GGA methods.[122] In this formalism, the exchange correlation energy expression has the form,

$$E_{XC}^{hybrid} = E_{XC}^{GGA} + a_0(E_X^{HF} - E_X^{GGA}) \tag{3.61}$$

where the a_0 parameter is fit to experimental data, similar to the b parameter in the GGA energy expression.

3.6 Solvent Effects

All the methods described above assume the system of interest is isolated, generally viewed as either *in vacuo* or as a low pressure gas. While this will produce adequate results for some systems, the vast majority of chemistry takes place in the solution phase, which contains solvent effects that are entirely neglected by traditional computational models.

Treating the solvent explicitly (*i.e.* adding discrete solvent molecules to the model)

would be desirable, but there are many problems introduced by doing this. For instance – how does one determine the number of solvent molecule required for complete solvation of a system? One way to solve this is to add solvent molecules until the properties of interest converge, but this can be time consuming and practical limits of computational power may prevent complete solvation from ever being reached. If attempting to model an aqueous environment, the addition of each water molecule adds ten electrons and many extra degrees of freedom to the system. This method still does not fully describe the solvent, since it is inherently static and the solvent molecules are always in motion around the solute. From this, it is easy to see that even a relatively small number of solvent molecules can increase the computational cost beyond what is currently practical. Therefore, a method that introduces solvent effects in a simpler way is required.

3.6.1 Self-consistent reaction field

To model the effect of a solvent on the solute, Onsager[123] proposed placing it inside a spherical cavity in a medium with a dielectric constant, ε, equal to the desired solvent. This provides a stabilisation energy dependant on the dielectric constant of the solvent and the dipole moment, μ, of the solute,

$$\Delta G_{solv} = -\frac{2(\varepsilon - 1)}{a^3(2\varepsilon + 1)}\mu \tag{3.62}$$

where a is the radius of the cavity.

The concept proposed by Onsager is appealing, but simply using a static spherical cavity does not effectively model most solute-solvent interactions. Most molecules of interest are not spherical and therefore much of the solute will be far from the walls of the cavity. Additionally, introducing an external electric field will affect the electron distribution in the solute, changing its dipole moment, an effect which is neglected in the Onsager model. To better model these interactions a new formalism, known as the self-consistent reaction field (SCRF) was proposed.

To capture the polarization effects of the solvent medium on the solute and *vice*

versa the surface of the cavity is divided into small rectangular elements and each is assigned a charge, whose magnitudes are defined as

$$q(\mathbf{r}_s) = -\frac{\varepsilon - 1}{4\pi\varepsilon}\mathbf{E}(\mathbf{r}_s) \tag{3.63}$$

where $\mathbf{E}(\mathbf{r}_s)$ is the electric field perpendicular to the surface. The potential energy of the interaction between the solute dipole and the cavity surface charges is added to the Hamiltonian operator as an additional term.

$$\hat{H} = \hat{H}_0 + \int \frac{q(\mathbf{r}_s)}{|\mathbf{r} - \mathbf{r}_s|}\,\mathrm{d}\mathbf{r}_s \tag{3.64}$$

This potential now affects the molecular wavefunction and can be solved iteratively, like the SCF procedure described previously.

Better methods for defining the cavity have also been proposed. One of the most popular today is the polarizable continuum model (PCM),[124] which builds a cavity from overlapping spheres, normally centred on the nuclei, to keep the solute a constant distance from the walls of the cavity. To produce an even more accurate cavity, the surface can be defined at a particular electron density isosurface (IPCM).[125] This, however, can become very computationally expensive.

In the work presented in the following chapters, the conductor-like PCM (CPCM) method[126,127] is used. In this model, the cavity is defined in the same way as the PCM model, however the charges on the cavity surface are treated differently. Rather than using polarizable dielectric charges, as in other PCM methods, this method treats the solvent medium as an electrical conductor. For non-polar solvents, this formalism is less accurate than a polarizable dielectric, however the two methods become equivalent as the dielectric of a solvent tends toward infinity. Using this conductor-like picture of the solvent is attractive, despite its unphysical representation of the solvent, because it is computationally much simpler and therefore becomes more useful for large systems.

Often, to find an acceptable compromise between accuracy of a computational

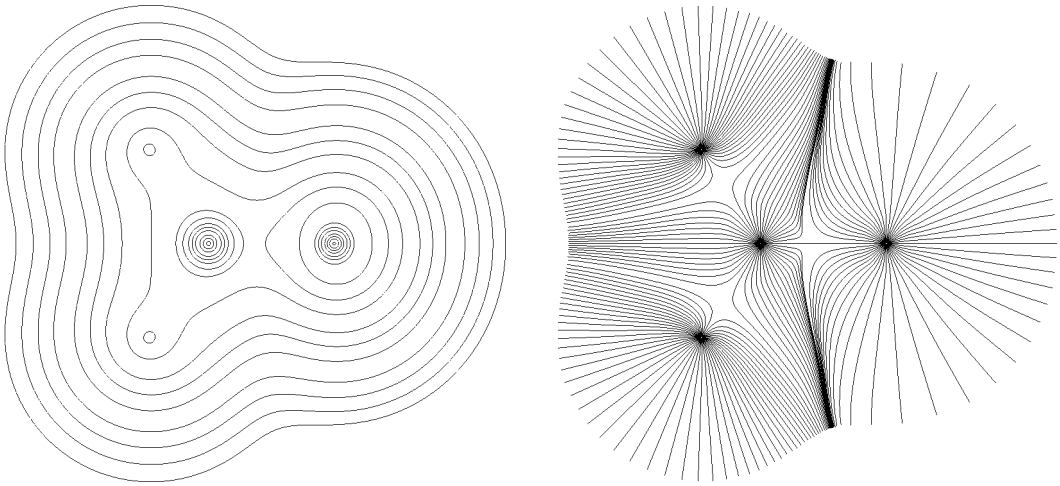
model and computational cost, a combination of some explicit solvation and an implicit solvent model will be used. This allows important solute-solvent interactions, such as solvent stabilisation of transition states or solvent mediated proton transfers, to be included in the model system, while keeping the overall computational cost at a reasonable level.

3.7 Quantum Theory of Atoms in Molecules

Using the methods described above, one can obtain information on the geometry of a chemical system and its total energy. The molecular wavefunction, $\Psi(\mathbf{r})$, describing the distribution of the electrons in the system, is also obtained from these quantum chemical calculations. This wavefunction, and its derived electron density, $\rho(\mathbf{r})$, contain a large amount of information about the system. Unfortunately, it is very difficult to analyse to obtain meaningful results, due to its complexity. The quantum theory of atoms in molecules (QTAIM) developed by Bader and coworkers[128] is a method whose goal is to partition a molecule into regions which define an atom in the molecule through a topological analysis of the electron density. Due to the number of properties which can be calculated using QTAIM, only the required background and those properties used in the following chapters will be discussed here. For a complete discussion of all that can be done with QTAIM analysis, the reader is referred to more in-depth reviews of the subject.[128–130]

The positively charged nuclei exert a strong attractive force on the negatively charged electrons, so a local maximum (a cusp) in $\rho(\mathbf{r})$ is found at each nucleus, as seen in figure 3.3a. In the figure, the carbon and oxygen nuclei of formaldehyde can clearly be identified as the large peaks on the centre and right-hand side of the plot, respectively. The hydrogen nuclei are characterised by the much smaller peaks toward the left-hand side of the plot.

Although the maxima of the density are useful for identification of nuclei, there are other critical points, such as minima and saddle points, of interest in the electron density. A critical point is defined as the point where the first derivatives of the



(a) (b)

Figure 3.3: (a) Electron density contour plot of formaldehyde in the plane of the molecule. (b) Gradient vector field of the density in (a). Wavefunction calculated at B3LYP/6-31G level.

density are zero,

$$\nabla\rho = \hat{i}\frac{\partial\rho}{\partial x} + \hat{j}\frac{\partial\rho}{\partial y} + \hat{k}\frac{\partial\rho}{\partial z} = \vec{0} \tag{3.65}$$

where the zero vector, $\vec{0}$, is used to indicate that all partial derivatives are zero and not just their sum. The gradient alone is not enough to discriminate between the various types of critical points, for that the second derivative of $\rho(\mathbf{r})$ at each critical point is also required. The relevant second derivatives of the density are normally represented by a diagonalized Hessian matrix, $\mathbf{\Lambda}$,

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{\partial^2\rho}{\partial x^2} & 0 & 0 \\ 0 & \frac{\partial^2\rho}{\partial y^2} & 0 \\ 0 & 0 & \frac{\partial^2\rho}{\partial z^2} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{3.66}$$

The three diagonal elements of this matrix, λ_1, λ_2, and λ_3, are the curvatures of the density with respect to the axis system and are used to classify the critical points.

The critical points are classified by finding their rank, ω, and signature, σ, and are symbolized as (ω, σ). The rank is the number of non-zero curvatures at a critical point. Since the density is a function in three-dimensional space, the maximum rank

possible is three and it is found that critical points with $\omega < 3$ are mathematically unstable and vanish with small perturbations of density caused by nuclear motion. The signature is the sum of the signs of the curvatures, *i.e.* each curvature contributes ± 1. There are four types of critical points that can be described in this way:

- $(3, -3)$ $\rho(\mathbf{r})$ is a maximum. These are found at the nuclei and are referred to as nuclear critical points (NCP).
- $(3, -1)$ $\rho(\mathbf{r})$ is a saddle point with a maximum in the plane defined by the negative curvatures and a minimum along the axis perpendicular to the plane. These are found in the space between two neighbouring nuclei and are referred to as bond critical points (BCP).
- $(3, +1)$ $\rho(\mathbf{r})$ is a saddle point with a minimum in the plane defined by the positive curvatures and a maximum along the axis perpendicular to the plane. These points are found in the centre of cyclic rings of nuclei and are referred to as ring critical points (RCP).
- $(3, +3)$ $\rho(\mathbf{r})$ is a local minimum. These points are found in the space enclosed by several rings and are referred to as cage critical points (CCP).

The number and types of critical points that exist in any given molecular system follow a strict relationship, known as the Poincaré-Hopf relationship,

$$n_{NCP} - n_{BCP} + n_{RCP} - n_{CCP} = 1. \tag{3.67}$$

If this relationship is not met, it is likely that some critical points have been missed or misidentified. Satisfaction of this relationship, however, does not prove that all critical points in a system have been found, but it is unlikely in practise that the correct combination of critical points have been missed to give an error.

If one examines the density between two nuclei connected by a BCP, a line of locally maximum density can be found connecting the nuclei with its lowest point at the BCP. This line is termed a bond path and is used in QTAIM as an indicator of

chemical interactions of all types, weak and strong, open and closed shell interactions. Similarly, a line of maximum density can be found connecting RCPs and the BCPs in the ring termed ring paths and RCPs and CCPs termed cage paths, although in practise these provide less useful information than the bond path. The collection of critical points and bond paths are known as a molecular graph and can be used to visualize molecular structure. The molecular graph of malonic acid is shown in figure 3.4. Note that all interactions, including the intramolecular hydrogen bond, are treated identically in this representation.

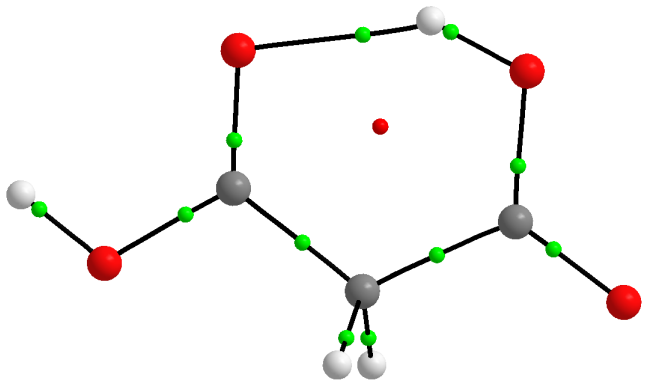


Figure 3.4: Molecular graph of malonic acid. Carbon, oxygen, and hydrogen atoms are large grey, red, and white spheres, respectively. BCPs are green spheres, the RCP is small red sphere, bond paths are black lines. Wavefunction calculated at the B3LYP/6-31G level.

Since there is a maximum in the electron density at each nucleus, it is natural to attempt to partition the density into individual atoms in the molecule, denoted Ω. One finds that there is a surface in the gradient vector field of the density (see figure 3.3b) that has zero flux, meaning it satisfies the condition

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \text{for all } \mathbf{r} \text{ belonging to the surface } S(\Omega) \tag{3.68}$$

where $\mathbf{r}$ is the position vector and $\mathbf{n}(\mathbf{r})$ is the unit vector normal to surface $S(\Omega)$. A zero-flux surface is found separating each pair of interacting atoms in a molecule and is therefore referred to as the interatomic surface (IAS). Note that the interatomic surface corresponding to a pair of atoms will pass through the BCP connecting those

atoms.

There are a number of properties that can be calculated from a QTAIM representation of a molecular system to give insight into its bonding and atomic properties. The density at a given BCP, ρ_{BCP}, can be used as a measure of bond strength – a larger ρ_{BCP} is indicative of a stronger bond. This measure can only be used to compare similar bonds, so in the malonic acid example from figure 3.4, one could compare ρ_{BCP} at the two carbonyl groups to determine the hydrogen bond's effect on the C=O bond, while it would be meaningless to compare a carbonyl BCP to a C−C BCP.

Another useful property of a BCP is the trace of $\mathbf{\Lambda}$, known as the Laplacian,

$$\nabla^2\rho(\mathbf{r}) = \lambda_1 + \lambda_2 + \lambda_3 \tag{3.69}$$

which is a measure of whether the two negative curvatures or the positive curvature dominate the nature of that BCP, or in chemical terms, the extent to which the density is concentrated in the region of the bond path or IAS. It is found that for covalent bonds, $\nabla^2\rho_{\mathrm{BCP}} < 0$, meaning that the density is concentrated along the bond path. When $\nabla^2\rho_{\mathrm{BCP}} > 0$, the density is depleted along the IAS and concentrated in their respective basins, commonly found in closed-shell interactions like hydrogen bonds, ionic bonds, or van der Waals interactions.

Various atomic properties can be found by integrating over a particular atomic basin, Ω. For instance, the electron population in an atom can be found by integrating $\rho(\mathbf{r})$,

$$N(\Omega) = \int_{\Omega} \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{3.70}$$

to give the number of electrons found within that basin. The atomic charge, more commonly used in chemistry, can then be obtained by subtracting the electron population from the nuclear charge, Z_Ω,

$$q(\Omega) = Z_\Omega - N(\Omega). \tag{3.71}$$

An atomic volume, $Vol(\Omega)$, can also be calculated by simply determining the space that is bounded by all IASs associated with that atom and a threshold density (typically 0.001 au, since this is close to experimental van der Waals surfaces in the gas phase). By comparing atomic volumes for an atom in different bonding patterns (for instance the oxygen in water and an ether) one can determine how the atom's size is affected by neighbouring groups.

Chapter 4

Benchmark of Density Functional Methods for Use with Organoselenium Compounds

Reproduced in part with permission from Heverly-Coulson, G. S.; Boyd, R. J. *J. Phys. Chem. A* **2011**, *115*, 4827-4831.

In this chapter, the results of a benchmarking study to determine the optimal DFT functional/basis set pair to employ for modelling organoselenium compounds used as GPx mimics are summarized. To do this, a series of small organoselenium and organosulfur compounds were studied as models for relevant bonding during the catalytic cycle of antioxidant GPx mimics. They are modelled using functionals commonly used in the literature (*e.g.* B3LYP, B3PW91, PBE, B971) and compared to other functionals which have not yet seen widespread use for these types of compounds. A variety of double and triple-zeta Pople basis sets, with varying numbers of polarization and diffuse functions, are examined to determine the appropriate basis set for predicting accurate geometries and energies. High-level conventional electronic structure calculations are used as a reference for judging the accuracy of the DFT methods.

4.1 Motivation

Organoselenium chemistry is an active research area for both synthetic chemistry and computational modelling. In particular, the biochemical applications of these compounds as antioxidants[57,58,72,77,80,131] has received a great deal of attention, as discussed in chapter 2. Additionally, seleno-substituted nucleic acids and nucleobases[132–134] and the use of organoselenium antioxidants as substrates for some enzymes[68,135] are especially topical. In order to better understand the phenomena

observed experimentally, many theoretical studies have been carried out, modelling organoselenium compounds with high-level quantum chemical methods.[55,64–66,81,136–140]

Despite the many computational studies on biologically relevant organoselenium compounds, relatively little has been done to evaluate appropriate methods for studying them. Bacharach and Jiang noted this over a decade ago,[141] at which time many theoretical studies had already been published. Six years ago, Pearson *et al.*[142] evaluated the B3LYP and B3PW91 density-functional theory methods for use with organoselenium compounds. However, there have been no recent systematic studies that examine more recently proposed functionals for this purpose.

When designing a benchmarking study, it is important to take a number of factors into account. The first is the nature of the chemical systems under examination and the properties that are of most interest. For instance, a study of molecules at their ground-state, equilibrium geometries may use different methods than one studying excited state molecules or molecules at non-equilibrium geometries. It is also important to balance the competing factors of accuracy and computational cost, taking into account not only the size of the model systems used in the benchmark, but also the compounds of interest for the larger project. It is not useful to use exclusively multi-configurational methods with large basis sets in a benchmark if the purpose is to determine methods for studying systems with hundreds of electrons.

4.2 Model Systems and Computational Methods

The series of small organoselenium and organosulfur compounds shown in figure 4.1 was studied as models representing the relevant bonding patterns in the catalytic cycle of antioxidant GPx mimics. These include a selenol (**1**) and its corresponding selenolate (**2**), a selenenic acid (**3**), a thiol (**4**) and its corresponding thiolate (**5**), a diselenide (**6**), a selenylsulfide (**7**), and a disulfide (**8**). The model compounds listed are suitable for evaluating the ability of various methods to predict molecular geometries. They, however, cannot be used for energy predictions; to do that effectively, a reaction must be studied. Figure 4.2 shows a series of seven homolytic bond cleavage

reactions that are used to evaluate bond dissociation energy (BDE) predictions. The dissociation energies of Se−H (**I**), C−Se (**II**), Se−O (**III**), S−H (**IV**), Se−Se (**V**), Se−S (**VI**), and S−S (**VII**) bonds are included. The BDE is calculated as the difference between the total energy of the full molecule and the sum of the energies of the two radicals formed from bond cleavage.

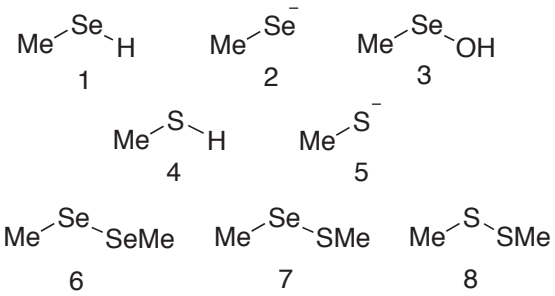


Figure 4.1: The series of organoselenium and organosulfur compounds included in this benchmark.

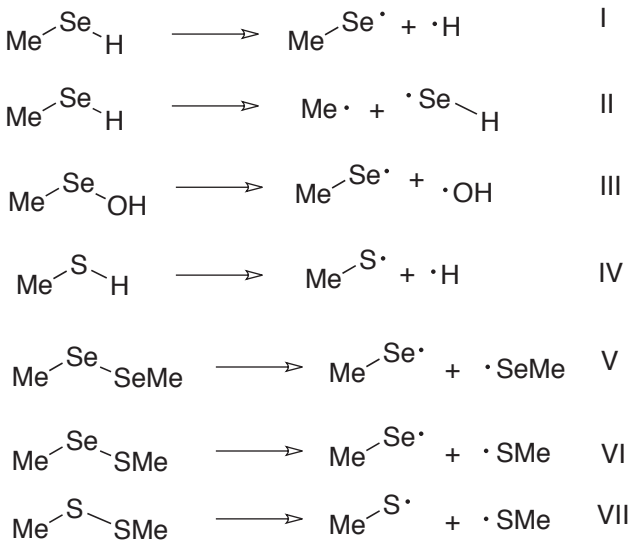


Figure 4.2: The series of homolytic bond cleavage reactions included in this benchmark.

All calculations were performed using the Gaussian 09[143] suite of programs. Geometry optimizations and frequency calculations were performed using a variety of Pople basis sets[95] paired with several pure and hybrid DFT methods. The pure

functionals studied include PBE,[144,145] HCTH,[146] and B97D.[147] Hybrid functionals studied include B3LYP,[122,148] B3P86,[122,149] B3PW91,[122,150] O3LYP,[151] B971,[146] and τHCTHhybrid.[152] High-level reference calculations were performed using the quadratic configuration interaction (QCISD) method with Dunning's correlation consistent triple-zeta basis set (cc-pVTZ).[153]

All energies include zero-point vibrational corrections, calculated from frequency analysis using the same method and basis set as the geometry optimization, with the exception of QCISD reference calculations for **3**, **6**, **7**, and **8**, which use frequencies calculated at the QCISD/cc-pVDZ level due to the overwhelming computational cost required to calculate the zero-point energy correction using the triple-zeta basis set.

4.3 Results and Discussion

In order to assess the accuracy of the computational results, it would be ideal to compare to accurate experimental geometries and BDEs. However, there is a lack of these for the series of molecules studied. Therefore, results obtained at the QCISD/cc-pVTZ level will be used for comparison, which Pearson *et al.*[142] previously showed agree well with experimentally obtained geometries for hydrogen selenide and dimethyl selenide.

4.3.1 Geometry prediction

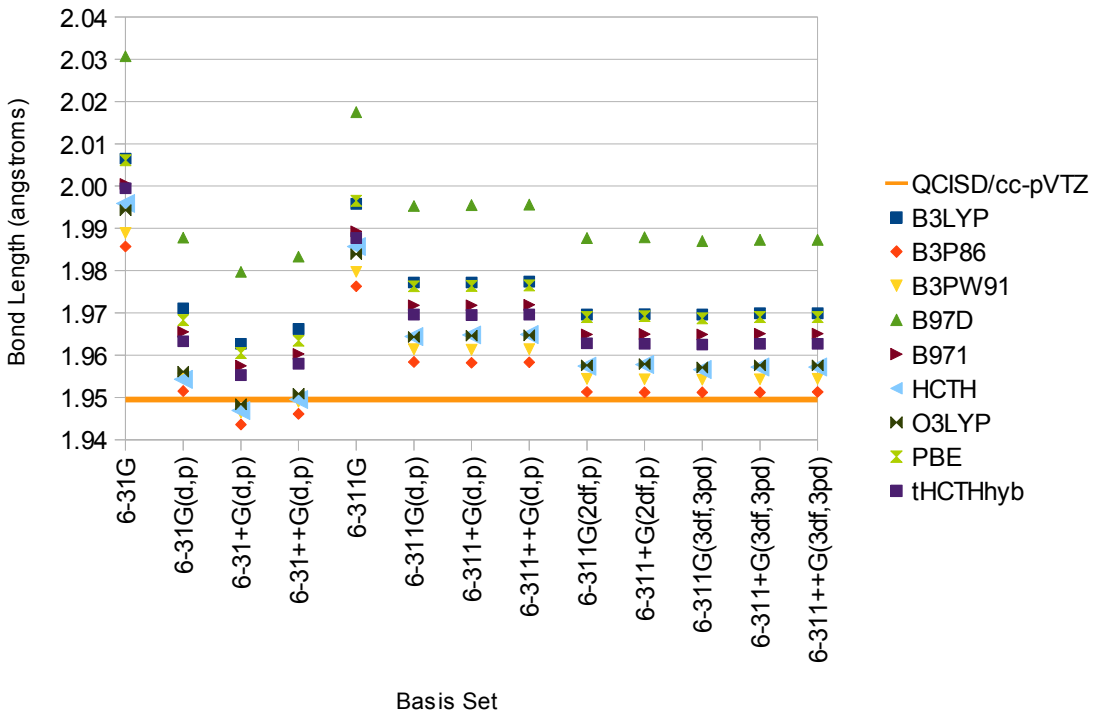
The first step was to optimize the geometries of the eight structures shown in figure 4.1 with the nine DFT functionals in conjunction with a series of basis sets. Figure 4.3 shows the optimized carbon–selenium bond in methyl selenol and dimethyl diselenide. The 6-31G and 6-311G basis sets, which contain no polarization or diffuse functions, predict remarkably different results than any of the other basis sets and predict significantly longer bonds than the reference calculations. Across all molecules studied, these two basis sets consistently predicted significantly longer bonds than their more

flexible counterparts containing polarization and diffuse functions. At its best performance, with the B3P86 functional, the root mean square deviation (RMSD)[i] for the 6-31G basis set for all selenium containing bonds is 0.061 Å, while the 6-311G performs marginally better with an RMSD of 0.058 Å. It is well known that basis sets without polarization functions predict bond lengths that are too long, a result reproduced here with organoselenium compounds. Since these basis sets produce unacceptable results with respect to the rest of the set, for clarity, they will be omitted from the remainder of the discussion.

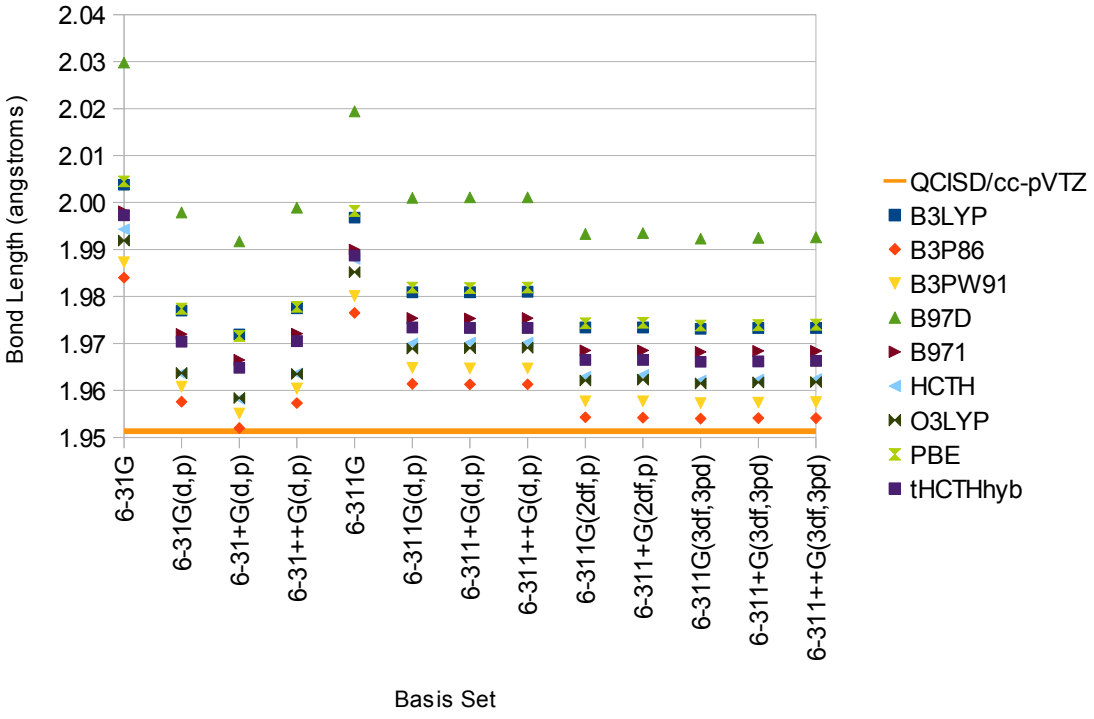
Once polarization functions are included, there is very little change in the C−Se bond lengths when diffuse functions are added as well. It is also interesting to note that the relative ordering of the DFT functionals remains consistent regardless of basis set used. The B97D functional predicts the longest bond lengths for all molecules studied, giving a C−Se bond about 0.02 Å longer than the functionals predicting the next longest bonds (PBE and B3LYP) in methyl selenol, while the remaining eight functionals predicts results that are clustered in a range of about 0.02 Å. The C−Se bond lengths predicted by each DFT/basis set pair are quite consistent across the various molecules studied, so the bond lengths can be studied in aggregate, as the arithmetic mean, without a significant loss of accuracy, listed in table 4.1. Every combination of DFT functional and basis set predicted longer C−Se bonds, on average, than the QCISD reference calculations, except B3P86/6-31+G(d,p). Across the entire series, B3P86 and B3PW91 predicted C−Se bonds that were closest to the QCISD reference values.

Methyl selenenic acid (**3**) contains both a C−Se and a Se−O single bond. Figure 4.4 shows the optimized Se−O bond length in this molecule. A similar trend is seen for this bond as was observed for the C−Se bonds. There is, however, a greater variability in these bond lengths across basis sets. Carbon–selenium bonds increased by about 0.01 Å upon going from a double-zeta basis set to a triple-zeta basis set with only the (*d*,*p*) set of polarization functions, while this increase doubled to 0.02

[i] RMSD $= \sqrt{\frac{\sum(x-x_{ref})^2}{n}}$, where x is the bond length of interest, x_{ref} is the reference bond length from QCISD/cc-pVTZ calculations, and n is the number of bonds in the set.



(a) C−Se bond lengths in methyl selenol (**1**).



(b) C−Se bond lengths in dimethyl diselenide (**6**).

Figure 4.3: Optimized C−Se bond lengths in selected molecules.

Table 4.1: Arithmetic mean C-Se bond lengths from **1**, **2**, **3**, **6**, **7** (in angstroms).

Functional	6-31G(d,p)	6-31+G(d,p)	6-31++G(d,p)	6-311G(d,p)	6-311+G(d,p)	6-311++G(d,p)
B3LYP	1.9741	1.9699	1.9731	1.9800	1.9806	1.9806
B3P86	1.9546	1.9499	1.9528	1.9605	1.9608	1.9608
B3PW91	1.9578	1.9530	1.9559	1.9638	1.9641	1.9641
B97D	1.9933	1.9890	1.9927	1.9994	2.0000	1.9999
B971	1.9686	1.9641	1.9672	1.9741	1.9746	1.9746
HCTH	1.9590	1.9551	1.9580	1.9677	1.9689	1.9687
O3LYP	1.9601	1.9556	1.9583	1.9670	1.9679	1.9677
PBE	1.9727	1.9685	1.9717	1.9797	1.9803	1.9802
τHCTHhyb	1.9666	1.9621	1.9652	1.9719	1.9724	1.9723

	6-311G(2df,p)	6-311+G(2df,p)	6-311G(3df,3pd)	6-311+G(3df,3pd)	6-311++G(3df,3pd)
B3LYP	1.9729	1.9736	1.9729	1.9734	1.9733
B3P86	1.9539	1.9541	1.9537	1.9540	1.9539
B3PW91	1.9572	1.9574	1.9570	1.9572	1.9572
B97D	1.9922	1.9929	1.9914	1.9919	1.9918
B971	1.9677	1.9682	1.9676	1.9680	1.9679
HCTH	1.9612	1.9624	1.9604	1.9613	1.9611
O3LYP	1.9607	1.9615	1.9602	1.9609	1.9607
PBE	1.9727	1.9734	1.9723	1.9729	1.9727
τHCTHhyb	1.9656	1.9659	1.9653	1.9656	1.9655
QCISD/cc-pVTZ				1.9520	

Å for the selenium–oxygen bond. As with the C−Se bonds, the B3P86 and B3PW91 functionals perform the best, while B97D and PBE produce bond lengths significantly longer than the reference values for the Se−O bond.

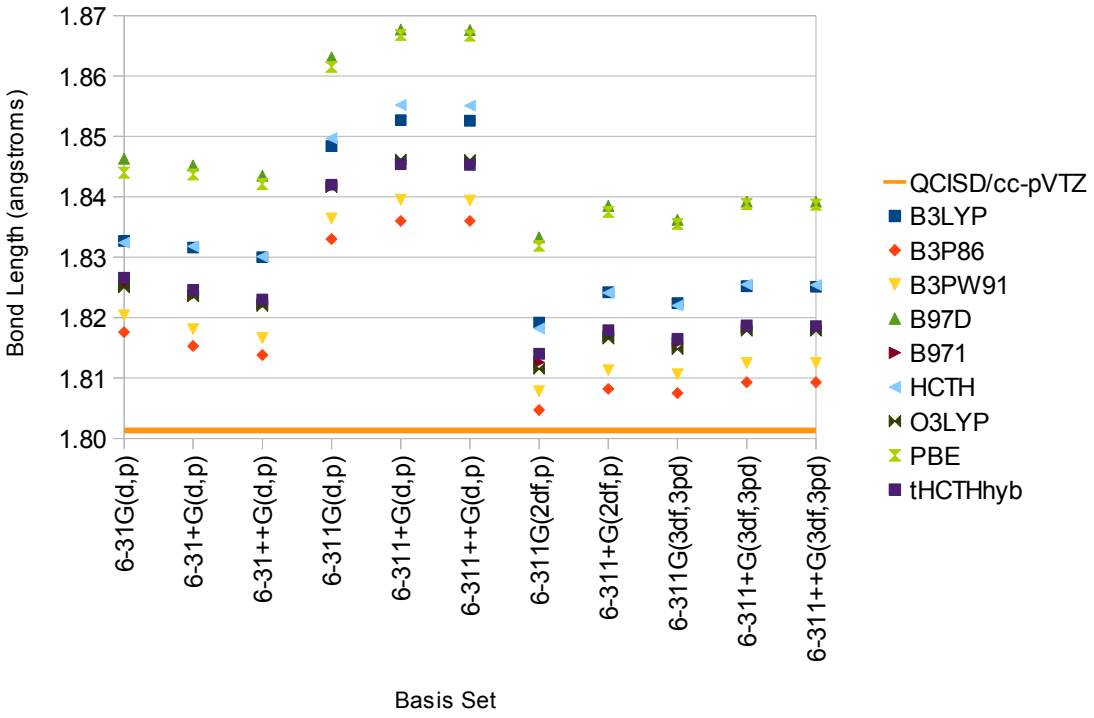


Figure 4.4: Optimized Se−O bond length in methyl selenenic acid (**3**).

The dimethyl diselenide (**6**) molecule contains a selenium–selenium single bond. Although similar trends to those discussed for C−Se and Se−O bonds are observed, the Se−Se bond is a rare case where some of the DFT functionals predict a bond that is shorter than the QCISD reference (figure 4.5). Among the double-zeta basis sets, B3P86, B3PW91, and O3LYP all predict Se−Se bonds about 0.01 Å shorter than the reference bond length. As observed for other bond types, the large triple-zeta basis sets with at least *d*- and *f*-type polarization functions on heavy atoms paired with the B3PW91 and O3LYP functionals predict the most accurate bond lengths, while the B3P86 and HCTH functionals also perform well.

Overall, there were ten different selenium-containing bonds in this study. To assess the ability of each DFT/basis set pair to accurately predict geometries involving selenium, the RMSD of each is summarized in figure 4.6. There are three functionals

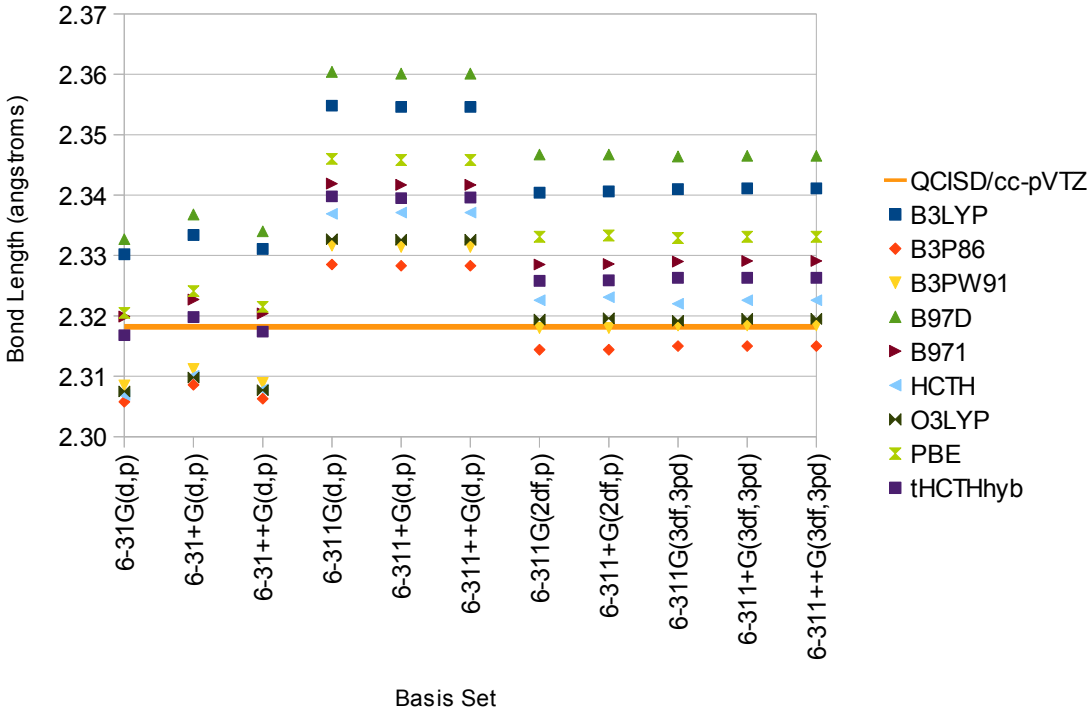


Figure 4.5: Optimized Se−Se bond length in dimethyl diselenide (**6**).

which consistently predict geometries to within 0.01 Å deviation when paired with both double-zeta and triple-zeta basis sets: B3P86, B3PW91, and O3LYP. There are also three functionals (B97D, B3LYP, and PBE), which predict selenium containing bond lengths with a deviation of 0.02 Å or more.

4.3.2 Energy prediction

Another important component to test the accuracy of a computational method is its ability to predict energies, especially if interested in studying reactions. To do this, the bond dissociation energies (BDE) for the homolytic bond cleavage reactions shown in figure 4.2 were calculated.

There is much less consistency across systems studied for the BDEs than the geometries discussed above. In general, the double-zeta basis sets predict higher BDEs than the triple-zeta basis sets for the selenium-containing bonds, while they predict smaller energies for the two bonds involving no selenium atoms, S−H (**IV**) and S−S (**VII**). The magnitude of the difference between the highest and lowest predicted

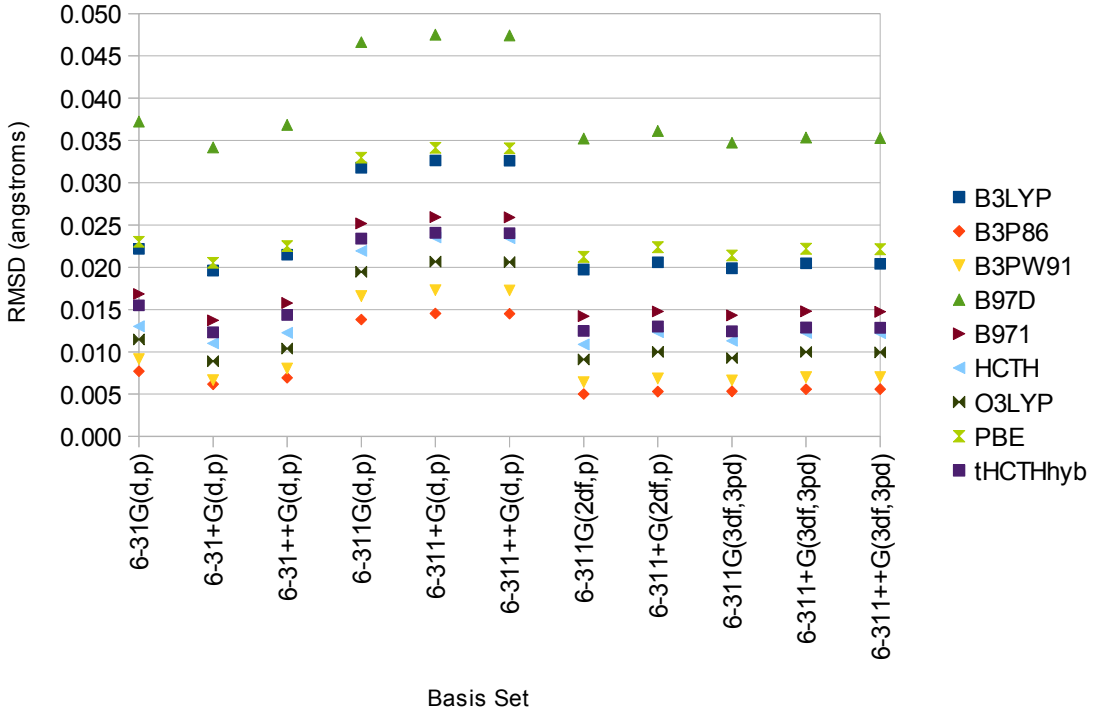
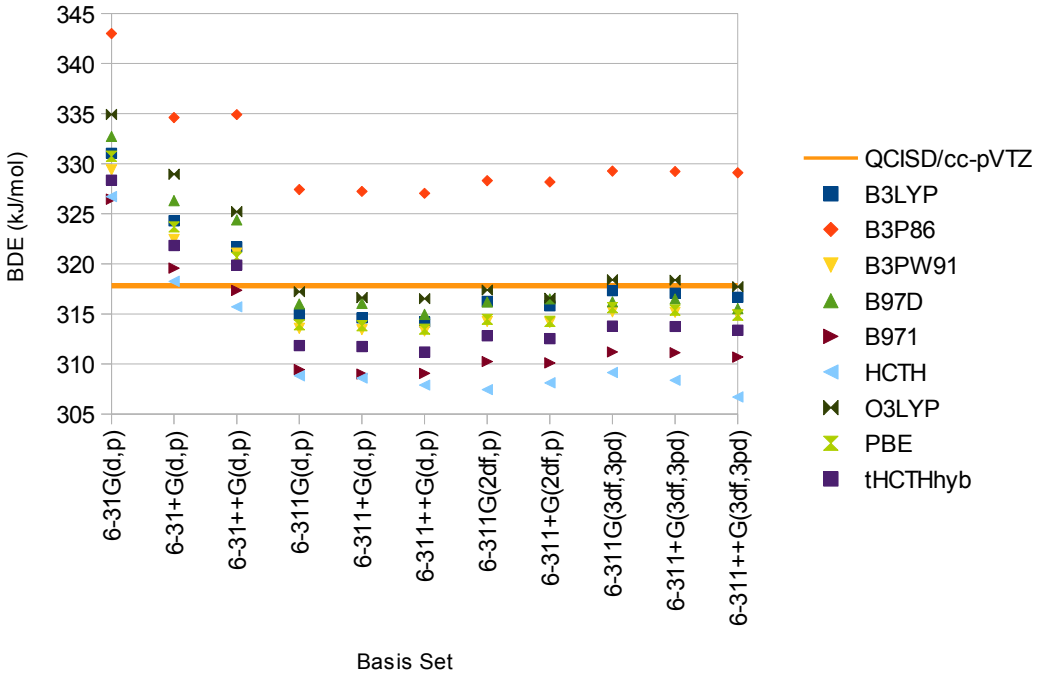


Figure 4.6: RMSD for the bond length of all selenium-containing bonds.

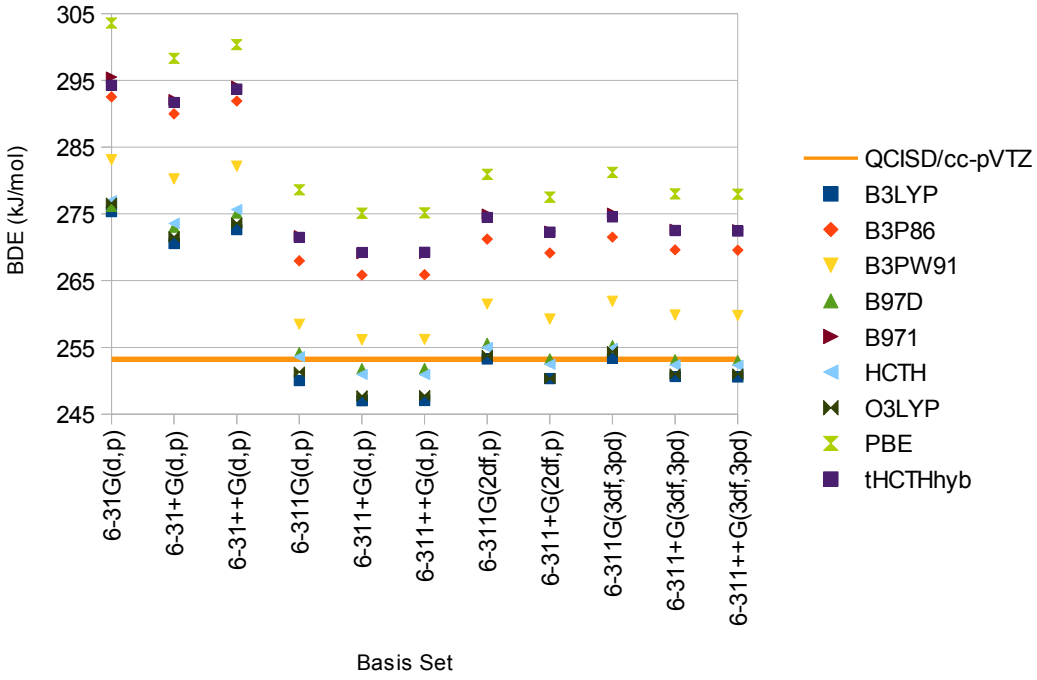
energies within the same DFT method range from about 15 kJ/mol for Se−H cleavage in methyl selenol (**I**) to 50 kJ/mol for Se−Se cleavage in dimethyl diselenide (**V**).

The energies associated with the Se−H bond cleavage (**I**) essentially converge by the first triple-zeta basis set, changing by no more than 2 kJ/mol upon increasing the size beyond 6-311G(d,p), as shown in figure 4.7a. Interestingly, the B3P86 functional, which predicts some of the best geometries, converges to a BDE 11 kJ/mol higher than the QCISD reference value. Conversely, B3LYP and B97D, which predict some of the poorest geometries overall, converge to BDEs within 2 kJ/mol of the reference value. Overall, every functional studied, except B3P86, B971, and HCTH, predicts BDEs within 5 kJ/mol of the reference value for this bond. Most of these, with the exception of B3P86, predict BDEs that are lower than the reference value.

The C−Se bond cleavage (**II**) in the methyl selenol molecule shows BDE trends similar to Se−H cleavage in the same molecule. For the best functionals, the double-zeta basis sets predict energies 20 kJ/mol higher than the QCISD reference. At the triple-zeta level, five DFT functionals (B3LYP, B3PW91, B97D, HCTH, and



(a) Se−H bond (**I**)



(b) C−Se bond (**II**)

Figure 4.7: Bond dissociation energies calculated for methyl selenol.

O3LYP) predict energies which converge to within 5 kJ/mol of the reference. While for the Se−H bond, most of the DFT methods predict BDEs lower than the QCISD reference, for the C−Se bond the majority of the functionals predict BDEs higher than the reference.

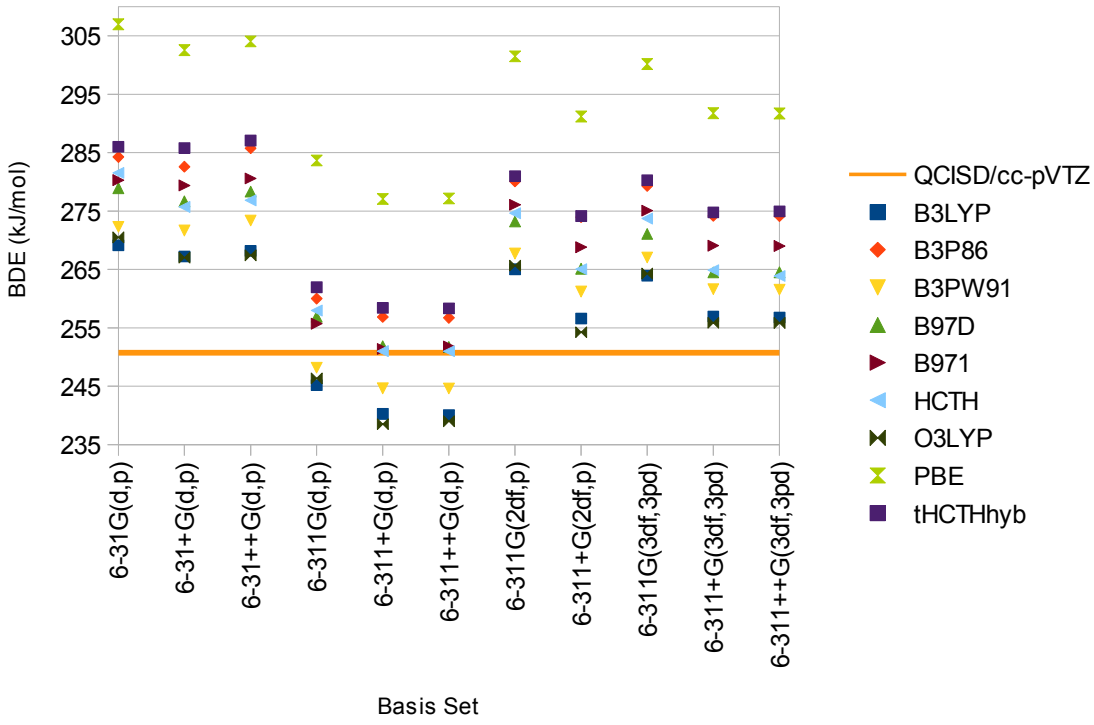


Figure 4.8: Bond dissociation energies for the Se−O bond in methyl selenenic acid (**III**).

The Se−O bond in methyl selenenic acid (**III**) is an unusual case where, unlike the two BDEs discussed above for methyl selenol, the largest basis sets converge to an energy more than 5 kJ/mol higher than the QCISD reference for most functionals (see figure 4.8). Additionally, the small double-zeta basis sets predict BDEs very similar to larger triple-zeta basis sets with many polarization functions but no diffuse functions. The addition of diffuse functions to these large basis sets decreases the predicted energy by about 10 kJ/mol. This trend is observed in the C−Se dissociation as well (figure 4.7b), however in that case the energy change is only 2-3 kJ/mol. The B3LYP

and O3LYP functionals converge to a BDE about 5 kJ/mol higher than the QCISD reference when paired with a triple-zeta basis set with at least two sets of d- and one set of f-type polarization functions on heavy atoms and p-type functions on hydrogen paired along with diffuse functions. B3PW91 converges to an energy 10 kJ/mol above the reference with the same basis sets.

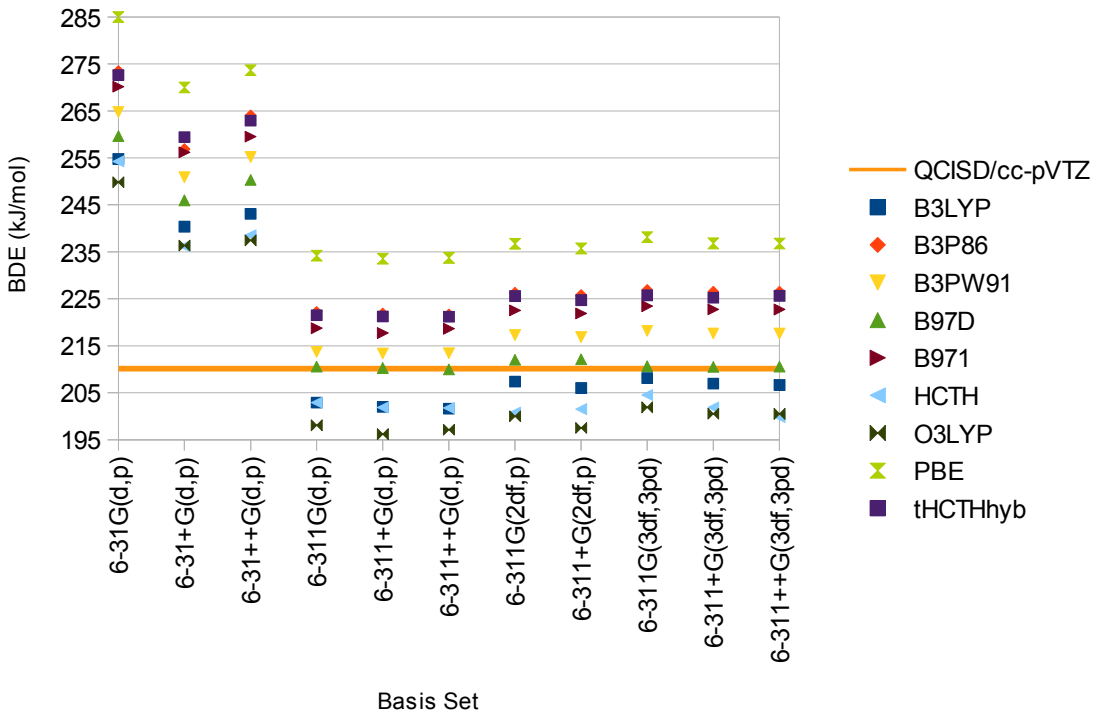


Figure 4.9: Bond dissociation energies for the Se−Se bond in dimethyl diselenide (**V**).

The BDEs predicted for the Se−Se bond in dimethyl diselenide (**V**) follow a similar trend to the Se−H bond in methyl selenol discussed above, in that there is a large energy difference between double and triple-zeta basis sets and the energies converge at the smallest triple-zeta basis set containing polarization functions. This bond is unique in having the largest change in energies between double and triple-zeta basis sets, a decrease of about 50 kJ/mol between 6-31G(d,p) and 6-311G(d,p). The B97D functional converges to within 1 kJ/mol of the reference BDE and B3LYP, B3PW91, HCTH, and O3LYP all converge to within 10 kJ/mol of the reference.

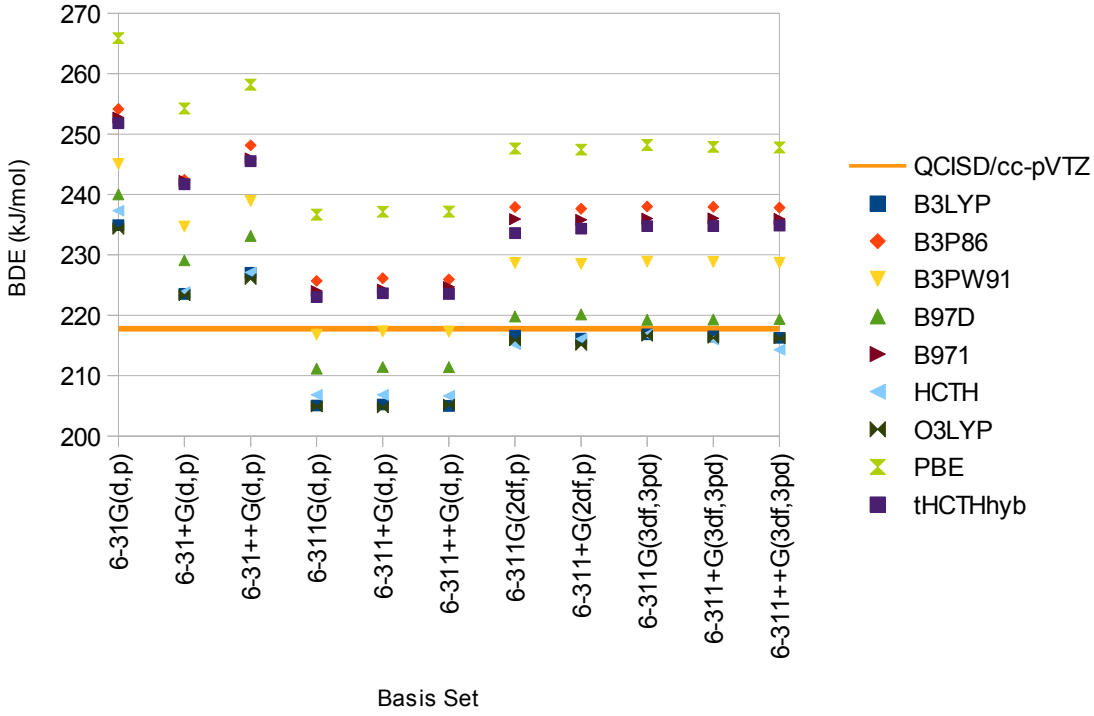


Figure 4.10: Bond dissociation energies for the Se−S bond in dimethyl selenylsulfide (**VI**).

The predicted BDEs for the Se−S (**VI**) bond in dimethyl selenylsulfide (figure 4.10) converge at the triple-zeta basis set with many polarization functions. B3LYP, B97D, HCTH, and O3LYP all converge to within 2 kJ/mol of the reference and B3PW91 converges to 10 kJ/mol above the reference.

As in the geometries, with the diversity of bonds studied, it is difficult to draw conclusions from examining each type of bond individually, since each one is an important component of bio-organic selenium chemistry. Therefore, studying the RMSD over all BDEs (figure 4.11) can provide valuable information. As expected, the RMSD for every DFT/double-zeta basis set pair is large, over 15 kJ/mol in the best cases and with deviations as high as 45 kJ/mol using the PBE functional. Interestingly, for the set of triple-zeta basis sets with only *d*- and *p*-type polarization functions, every functional except PBE shows an RMSD of 10 kJ/mol or less. For these three

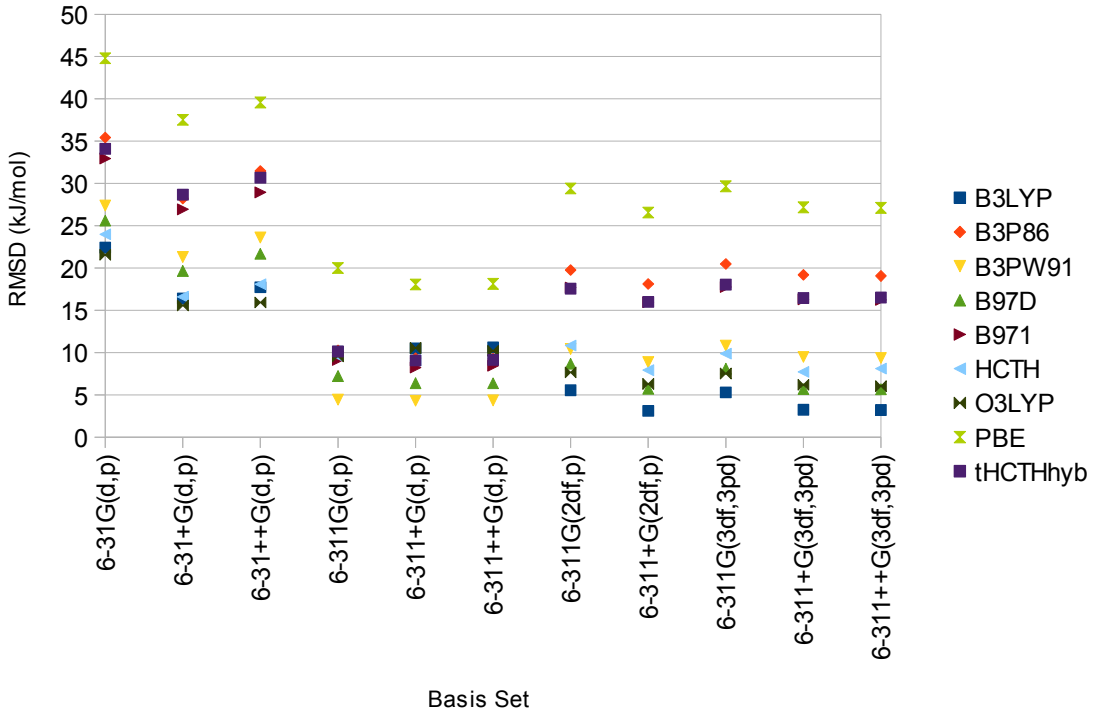


Figure 4.11: RMSD for the BDE for the seven bonds shown in figure 4.2.

basis sets, the B3PW91 functional predicts BDEs with an RMSD less than 5 kJ/mol. When more polarization functions are added, the RMSD for B3PW91 increases to about 10 kJ/mol, while the B3LYP functional shows a decrease in error and predicts BDEs with less than 5 kJ/mol RMSD.

As seen in figure 4.6, the geometries predicted with a small double-zeta basis set, such as 6-31+G(d,p), are very close to those predicted with a triple-zeta basis set, such as 6-311+G(2df,p). As outlined in figure 4.11, the BDEs computed using these two basis sets are quite different. This shows that the common procedure of predicting geometries with a small basis set and then computing single point energies at that geometry with a larger basis set may work well with these systems. The set of bond dissociation reactions shown in figure 4.2 were computed using geometries and energies calculated at the B3PW91/6-311+G(2df,p) level and using geometries calculated at B3PW91/6-31+G(d,p) with energies at B3PW91/6-311+G(2df,p). These results, summarized in table 4.2, show that the BDEs computed using the two different methodologies have less than 1 kJ/mol discrepancy over the test set.

Table 4.2: BDEs calculated at the B3PW91/6-311+G(2df,p) level using geometries optimized at either the B3PW91/6-31+G(d,p) or the B3PW91/6-311+G(2df,p) level (in kJ/mol).

Method	Bond Dissociation Energy (kJ/mol)	
	B3PW91/6-311+G(2df,p)// B3PW91/6-31+G(d,p)	B3PW91/6-311+G(2df,p)// B3PW91/6-311+G(2df,p)
I	313.9	314.2
II	258.6	259.2
III	261.5	261.2
IV	342.9	343.2
V	216.6	216.8
VI	228.4	228.5
VII	244.0	244.5

4.4 Conclusions

The B3PW91 functional, which provides a good compromise between accuracy in both geometries and energies, is recommended for calculations on organoselenium systems. The B3P86 functional predicts geometries marginally closer to the reference than B3PW91, but is the second worst for predicting BDEs modelling the systems included in the following chapters. Conversely, B3LYP performs slightly better than B3PW91 for predicting BDEs, but is consistently one of the least accurate for predicting geometries. Therefore, B3PW91, while not the best performing in any of the criteria evaluated, provides a good compromise between accuracy in both predicted geometries and energies.

Since the geometries predicted using small double-zeta basis sets are nearly identical to those predicted with larger triple-zeta basis sets, it is possible to increase computational efficiency by calculating geometries with a small basis set, such as 6-31+G(d,p), and performing a single point energy calculation on the optimized geometry using a larger basis set, such as 6-311+G(2df,p). Although additional polarization functions could be included, there does not appear to be a significant benefit from using more than the (2df,p) set of polarization functions.

Chapter 5

Reduction of Hydrogen Peroxide by Selenol

As outlined in section 2.3.1, the tertiary-benzylamine selenol family of GPx mimics, first proposed by Wilson *et al.*, appears to be an especially promising class of organoselenium antioxidants. It is widely accepted that these compounds catalytically reduce peroxides using thiols following the catalytic cycle shown in figure 2.4. There have been many *in vitro* studies performed that provide evidence for this mechanism,[73,76,79,83] however to date, there has been no complete computational study of this reaction. By modelling the reaction as predicted by experiment and looking at potential alternatives or side reactions, further insight into not only why the reaction proceeds as observed, but also why it does not follow other potential mechanisms can be gained.

When *N,N*-dimethyl-benzylamine-2-selenol (DMBS) was first proposed as a GPx mimic,[72] it was synthesised as the diselenide. The diselenide form was mixed with peroxide and thiol directly and the reaction proceeded as expected with a depletion of peroxide and the emergence of a disulfide. It was therefore assumed that the diselenide was reacting with peroxide directly, reducing it and producing an oxidised form of the diselenide. Later kinetic studies[73] that examined the reaction step-by-step determined that the diselenide does not react directly with the peroxide, but acts as a precatalyst, cleaving to form the selenol and never reappearing in the reaction mixture.

In this chapter, the reaction of hydrogen peroxide with both the selenol and diselenide form of DMBS is modelled to better understand why one is observed to react

and the other not. Additionally, to better understand the role of the various groups involved in the reaction, both the neutral and charged variants of them will be modelled to determine their effect on the energies for peroxide reduction.

5.1 Computational Methods

All calculations were performed using the Gaussian 03 suite of programs.[154] Geometry optimizations were performed with the B3LYP hybrid DFT functional, composed of Becke's three-parameter exchange correlation functional[122] and the correlation functional proposed by Lee, Yang, and Parr,[148] and the 6-31G(d,p) Pople basis set for all neutral and cationic atoms and the 6-31+G(d) basis set for anionic selenium atoms, as suggested by Pearson *et al.*[142] for organoselenium compounds. Transition states were found with Schlegel's synchronous transit-guided quasi-Newton (STQN) method[155,156] and were linked to the reactants and products by intrinsic reaction coordinate calculations.[157,158] Frequency calculations were performed on all optimized structures at the same level of theory to confirm whether the structure was a local minimum or first-order saddle point. A scaling factor of 0.9614[159] was used to obtain more accurate thermochemical data. Accurate energies were obtained for all systems at the B3LYP/6-311++G(3df,3pd) level. Solvent effects for an aqueous environment (dielectric constant of 78.39) were included in single point calculations using the conductor-like polarizable continuum model (CPCM).

All energy barriers are calculated from the transition state relative to the complex immediately preceding it. Gibbs energies of reaction are calculated as the difference between the product complex and reactant complex energies.

5.2 Results and Discussion

5.2.1 Diselenide Reaction

To determine why the diselenide form of DMBS is not the active form, its reaction with hydrogen peroxide was studied. A schematic of the expected reaction is shown

in figure 5.1. Since the reaction kinetics reported by Iwaoka and Tomoda imply a one-step mechanism,[73] it will be assumed that an oxygen forms a bond to selenium simultaneously with a proton transfer from one oxygen to the other. Figure 5.2 shows the relative Gibbs energy profile calculated for this reaction, with all energies relative to the infinitely separated reactant molecules.

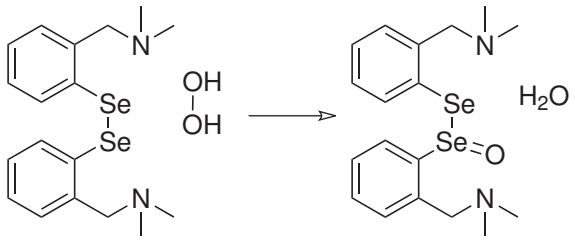


Figure 5.1: The reaction of N,N-dimethyl-benzylamine diselenide with hydrogen peroxide.

The reaction barrier for reduction of hydrogen peroxide by the diselenide was found to be 114.5 kJ/mol. This is equivalent to the barrier found by Pearson and Boyd[65] for ebselen reduction of hydrogen peroxide, but 11.6 kJ/mol higher than the barrier using ebselen selenolate, their most energetically favourable form of ebselen (see table 5.1). Studying the geometrical changes during the reaction shows that oxidation destabilizes the molecule. The selenium-selenium bond in the diselenide is 2.387 Å, which is slightly longer than the average Se−Se bond length of 2.33 Å.[160] At the transition state, this bond is lengthened to 2.448 Å and in the product it is 2.510 Å. With a separation of over 2.5 Å, it is difficult to classify this as a covalent interaction.

These reactions are normally carried out at neutral to slightly basic pH. At these pHs, it is expected that some of the amines would be protonated. With a protonated amine near the active site of the molecule, it is possible that it could be used as a proton shuttle to ease the strain of the intramolecular proton transfer in the peroxide and lower the barrier height.

Initially, only one of the two amines was protonated. The optimized structure of the gas phase protonated molecule shows a Se−Se bond length of 2.451 Å. Attempts to form a reactant complex with hydrogen peroxide lengthened this bond even more,

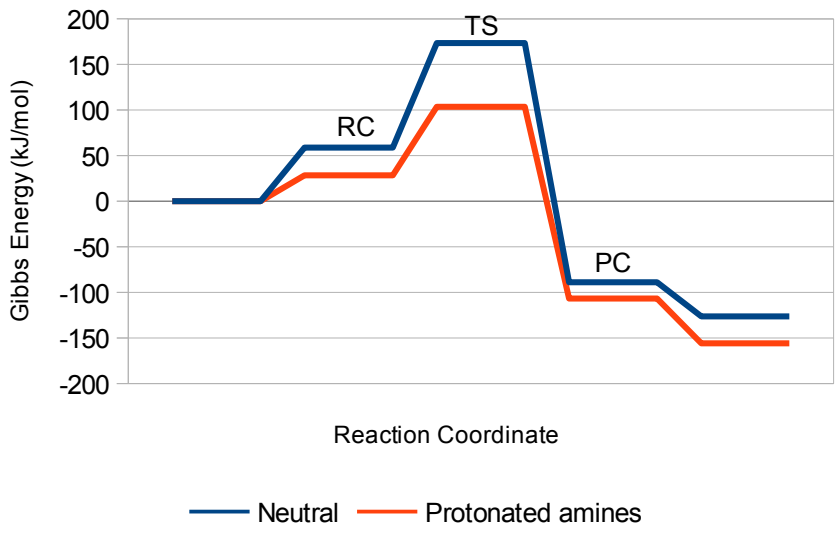


Figure 5.2: Gibbs energy profile of the reaction of N,N-dimethyl-benzylamine diselenide with hydrogen peroxide.

breaking the Se-Se bond. Therefore, attempts to model the reaction of this system with hydrogen peroxide were abandoned.

If both amines in the diselenide are protonated, the Se−Se bond length is 2.373 Å, which is slightly shorter than the neutral form. Following the reaction of this compound with hydrogen peroxide while using the adjacent amine as a proton shuttle, a barrier height of 75.0 kJ/mol is found, 39.5 kJ/mol lower than the neutral compound. Attempts were made to find a transition state that resembled the neutral compound, that is direct oxygen to oxygen proton transfer, for comparison, but all attempted structures optimized either to a local minimum or a transition structure incorporating the amine proton in the reaction, showing that the accessible reaction paths on the potential energy surface use the amine as a proton shuttle. The Se−Se bond lengthening observed during this reaction is smaller than in the neutral molecule, but is still considerable. In this case, the bond starts at 2.372 Å in the reactant molecule, remains the same in the transition state, and lengthens to 2.447 Å in the oxidised product.

Table 5.1: Calculated Gibbs energies of reaction in kJ/mol.

System	$\Delta G_1^{\ddagger}$	$\Delta G_2^{\ddagger}$	ΔG_{rxn}
N,N-dimethyl-benzylamine diselenide	114.5		-147.9
protonated amines	75.0		-135.2
N,N-dimethyl-benzylamine selenol	90.7	141.9	-214.0
zwitterion	41.2		-196.0
selenolate	34.3		-172.5
protonated amine	72.8		-252.5
ebselen[a]	114.2		-141.8
ebselen diselenide[a]	118.4		-136.0
ebselen selenololate[a]	102.9		-176.1

[a] Results from Pearson and Boyd[65] at the B3LYP/6-311++G(3df,3pd)/CPCM//B3LYP/6-31G(d,p) level.

5.2.2 Selenol Reaction

In the catalytic cycle outlined in figure 2.4, the catalytically active form of the GPx mimic is not the diselenide, but the zwitterionic form of the selenol. First, the reaction using the neutral selenol will be studied. It is assumed that in this reaction, the peroxide bond will cleave forming a Se−O bond and the selenol proton will be transferred to the other oxygen forming water, as shown in figure 5.3. As with the diselenide species, this reaction is expected to proceed via a one-step mechanism with the Se−O bond forming simultaneously with the proton transfer.[73]

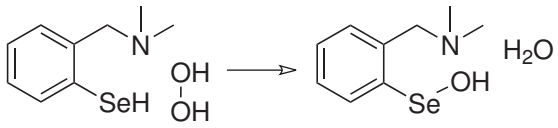


Figure 5.3: The expected reaction of neutral DMBS with hydrogen peroxide.

The mechanism found for the reaction of the selenol with hydrogen peroxide is not one step, as expected, but instead two step, as outlined in figure 5.4. In the first step of this reaction, a peroxide oxygen forms a bond with selenium with a simultaneous proton transfer from this oxygen to the other, producing water. The second step is a proton transfer from selenium to the oxygen bonded to it. The barrier height for the first step is 90.7 kJ/mol, which leads to an intermediate in an energy well 72.6 kJ/mol lower than the infinitely separated reactants. The barrier to go from this intermediate

to the product is 141.9 kJ/mol. The Gibbs energy profile of this two-step process is shown in figure 5.5. The barriers for both steps of this mechanism are quite large and are not consistent with the fast reactions observed by Wilson *et al.*[72] in their studies of these molecules. The kinetic studies performed by Iwaoka and Tomoda[73] suggest a one-step mechanism for the peroxide reduction step, which is also not consistent with this finding.

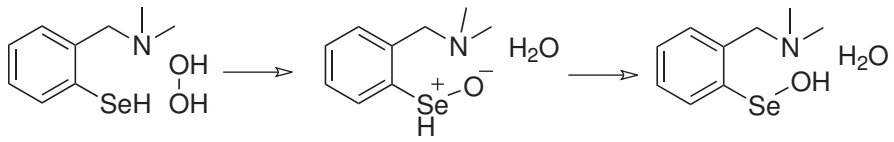


Figure 5.4: Two-step mechanism for the reaction of neutral DMBS with a peroxide.

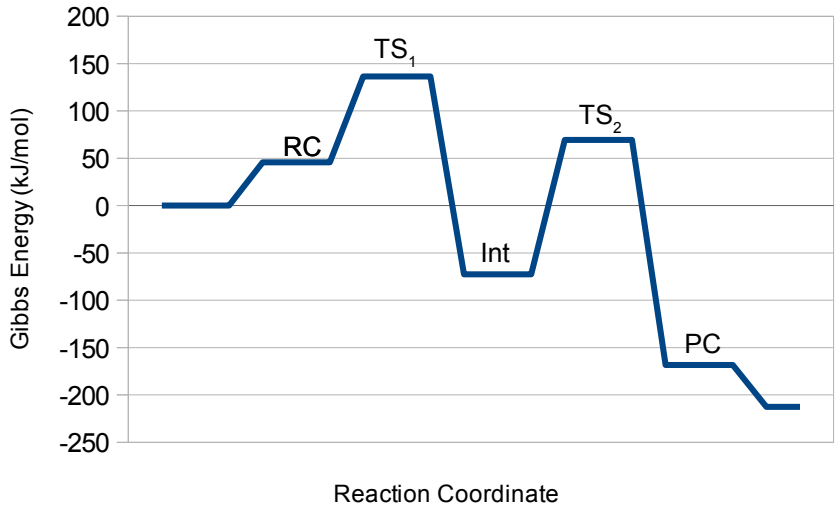


Figure 5.5: The Gibbs energy profile of the two-step reaction of neutral DMBS with hydrogen peroxide.

5.2.3 Zwitterion Reaction

The proposed catalytic cycle suggests that the neutral selenol is not the active form of the compound, but instead the zwitterion is used. The mechanism found using the zwitterion is a one-step process with an energy barrier of 41.2 kJ/mol. This is more

than 50 kJ/mol lower than either step of the reaction with the neutral selenol and less than half the barrier height predicted for ebselen by Pearson and Boyd.[65] In the zwitterion, there is a formal negative charge on selenium, which should increase its activity as a GPx mimic. Also, in this reaction, the proton on the amine is transferred to oxygen to form water, which reduces strain on the system, as with other proton transfer mechanisms, and therefore lowers the activation energy.

The zwitterionic form of the selenol is made up of two charged groups, an anionic selenolate and a cationic protonated amine. Studying the reactions of hydrogen peroxide and selenol analogues with these groups individually will provide some insight into which portion is most important to the reactivity of the zwitterion.

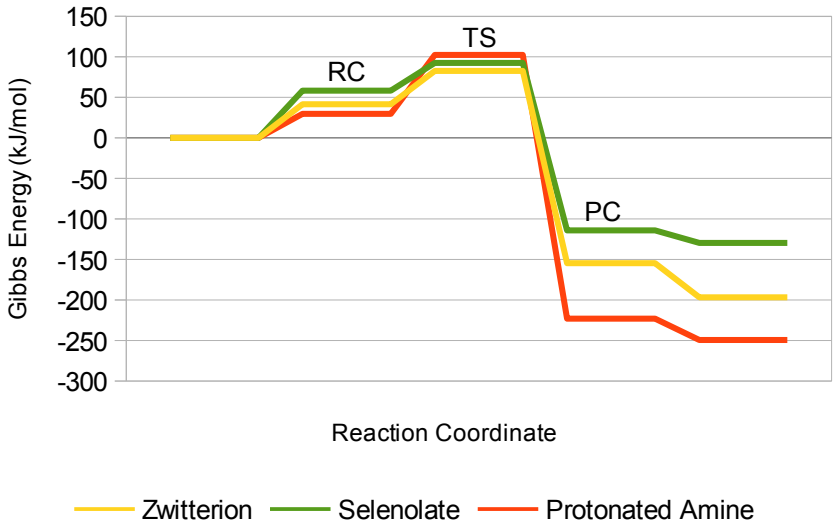


Figure 5.6: Gibbs energy profile of the reaction of charged analogues of DMBS with hydrogen peroxide.

In the reaction of the selenolate with hydrogen peroxide, a barrier of 34.3 kJ/mol is predicted. The transition state for this reaction shows the formation of a selenium-oxygen bond simultaneous with a proton transfer from the reacting oxygen to the other, which is similar to the first step observed for the neutral selenol, leading to a final product with a deprotonated oxygen bonded to selenium. The barrier height of this reaction is 7 kJ/mol lower than for the zwitterion. Natural population analysis

(NPA) charges on the selenium atom, calculated at the same level as the geometry optimization, are higher in the selenolate than the zwitterion, -0.54 and -0.33 au, respectively. This, in particular, demonstrates the importance of a nucleophilic selenium centre for this reaction.

The reaction between the cationic selenol, arising from the protonated amine, and hydrogen peroxide shows a barrier of 72.8 kJ/mol, which is about the same as for the diselenide with protonated amines. As with the zwitterion, the amine proton is used to form water, but in this case the selenol proton is simultaneously transferred to the amine. The additional selenium-to-nitrogen proton transfer may be the source of the larger barrier height of this mechanism.

5.3 Conclusion

It was found that the selenolate form of DMBS showed the most favourable energy barrier for peroxide reduction at 34.3 kJ/mol. Using the zwitterionic form as the reducing agent a barrier 7 kJ/mol higher is predicted, which reinforces the catalytic cycle for these compounds proposed by Iwaoka and Tomoda.[73] A previously unreported two-step mechanism was found for the neutral selenol, which has large barriers of 90.7 and 141.9 kJ/mol. The reaction of the diselenide with hydrogen peroxide has an energy barrier of 114.5 kJ/mol and shows considerable lengthening of the Se−Se bond to over 2.5 Å. By protonating the amines and using them as a proton shuttle, the energy barrier is lowered to 75.0 kJ/mol, but the Se−Se bond in this system is still lengthened to 2.447 Å. A similar barrrier is found by protonating the amine of the neutral selenol.

Although the zwitterionic form of the selenol does not have the lowest predicted energy barrier, its barrier is still quite low, reinforcing the proposed catalytic mechanism. It is unlikely that the more active selenolate form would be found experimentally as the selenol will be deprotonated and the amime will be protonated at the pHs of interest.

Chapter 6

Reaction of Selenenic Acid with Thiol

In this chapter, the mechanism for the substitution reaction between a selenenic acid and a thiol is examined. This is the first step in the regeneration of organoselenium antioxidants after reduction of a peroxide. The chemistry of selenenic acids is not well-characterised, although they are widely accepted as intermediates in the catalytic cycle of GPx and its mimics, they are very difficult to isolate and even observe *in situ*.

6.1 Introduction

In experimental studies using DMBS as a catalyst for the reaction of hydrogen peroxide with a thiol, the first proposed catalytic intermediate, a selenenic acid, was not observed,[73] instead a rapid conversion of the selenol to a selenylsulfide is seen. This implies that the intermediate between these two species is short-lived and is quickly converted to the selenylsulfide. Iwaoka and Tomoda[73] assigned this intermediate a selenenic acid structure through spectroscopic observations after reacting DMBS with hydrogen peroxide without any thiol present to continue the reaction. They also observed the higher oxoacids (seleninic and selenonic acids) in their reaction mixtures. Regardless of the relative ratios of the three oxoacids, after adding a thiol all the acids are rapidly converted to the selenylsulfide. Similar behaviour is found for GPx, where the seleninic acid form has been observed through crystallographic methods.[48]

The instability and high reactivity of selenenic acids has long been known. In solution they often react to form seleninic acids or diselenides. Some of the first stable selenenic acids reported were later proven to be selenenic acid anhydrides.[161,162] Throughout the 1980s, various stable phenyl-selenenic acids were reported, mostly stabilised

by coordination with ortho electron withdrawing groups, such as esters[163] and nitro groups.[164] Others have used bulky alkyl groups, as in 2,4,6-tri-*tert*-butylphenyl selenenic acid, to stabilise the selenenic acid group.[165] It was not until 1997 that a stable selenenic acid was characterised as a crystal rather than in solution. This was accomplished by burying the selenenic acid moiety in the conical cavity of *p-tert*-butylcalix[6]arene.[166] Although the surrounding calixarene prevented the formation of diselenides or anhydrides of this species, other common reactions of selenenic acids, such as with peroxide to form the seleninic acid or thiol to produce the selenylsulfide, were still observed.

In spite of the reaction of a selenenic acid with thiol being well known in the literature, this reaction has received little attention by computational chemists. In a DFT study of the complete enzymatic cycle of GPx, the Morokuma group[55] found that the oxidised selenocysteine reacts with ethanethiol via a one-step reaction with synchronous transfer of the S−H hydrogen to the leaving hydroxyl group and formation of the Se−S bond. A DFT study of the catalytic cycle of GPx mimics,[167] using benzeneselenenic acid and hydrogen sulfide, found a mechanism very similar to the one found by the Morokuma group. In this model, the barrier for this reaction is predicted to be 68 kJ/mol.

6.2 Computational Methods

All calculations were performed with the Gaussian 09[143] suite of programs. Geometry optimizations were performed using the B3PW91 hybrid DFT functional, composed of Becke's three-parameter exchange correlation functional (B3)[122] and Perdew and Wang's correlation functional (PW91)[150] with the 6-31+G(d,p) Pople basis set, as suggested by the benchmark in chapter 4. Transition states were found with Schlegel's synchronous transit-guided quasi-Newton (STQN) method[155,156] and the reaction path was followed using intrinsic reaction coordinate (IRC)[157,158] calculations. Frequency calculations were performed on all optimized structures at the same

level of theory to confirm whether the structure is a local minimum or first-order saddle point and to obtain thermochemical corrections. Accurate energies were obtained for all systems at the B3PW91/6-311+G(2df,p) level. Solvent effects for an aqueous (dielectric constant of 78.3553) or methanolic (dielectric constant of 32.613) environment were included in all calculations using the conductor-like polarizable continuum model (CPCM).

The topology of the electron density was studied using quantum theory of atoms in molecules (QTAIM)[128,129] calculations performed using AIMAll (Version 12.06.03).[168]

6.3 Results and Discussion

As with the peroxide reduction reaction in chapter 5, different mechanisms for this reaction will be examined to determine which is most favourable. The mechanisms studied primarily differ in the nature of the attacking sulfur group, using either a thiol or an anionic thiolate. These reactions produce the same selenylsulfide product, but differ in having either water or hydroxide as the leaving group.

6.3.1 Thiolate attack

If, in solution, the thiol is deprotonated, a highly nucleophilic thiolate can attack the selenium centre of the selenenic acid. Using this nucleophile, the expected reaction would be nucleophilic attack of the thiolate at the selenium centre with a hydroxide ion as the leaving group, outlined schematically in figure 6.1.

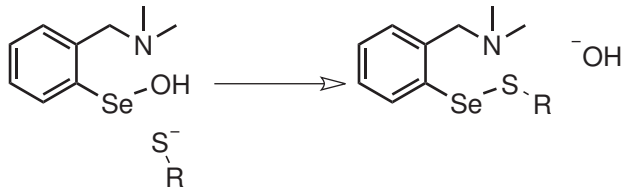


Figure 6.1: Schematic outline of the reaction of a thiolate with the selenenic acid of DMBS.

The reaction described was modelled using the thiolates of methanethiol and thiophenol. Modelling their reaction with the selenenic acid in water predicts energy

barriers of 175.4 and 177.0 kJ/mol, respectively. An aqueous environment would be found *in vivo*, however many of the *in vitro* studies of GPx mimics are conducted in methanolic environments. Modelling this reaction with methanol as the solvent increases the energy barriers slightly to 178.5 and 181.2 kJ/mol for the thiolates of methanethiol and thiophenol, respectively.

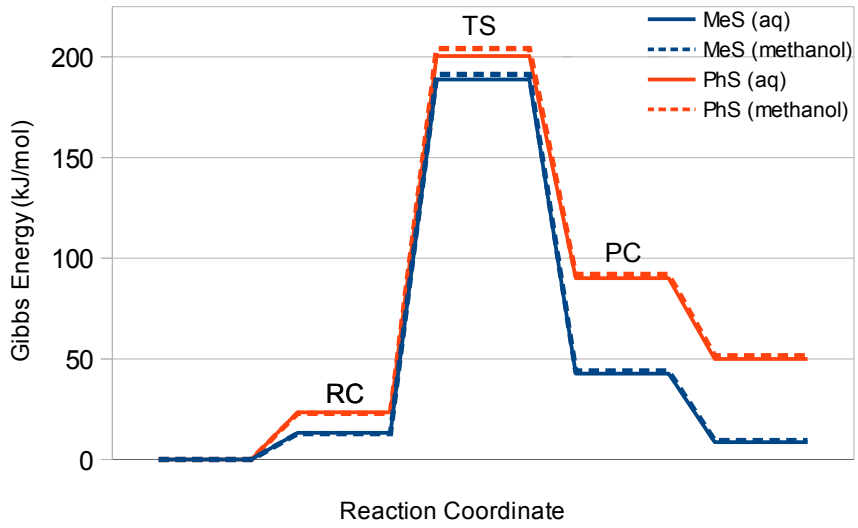


Figure 6.2: Reaction profile for the reaction of a thiolate with the selenenic acid of DMBS. Energies calculated in aqueous solvent are shown with a solid line and methanolic solvent with a dashed line.

Although the energy barriers do not change significantly with the use of methanethiol or thiophenol, the total energy of reaction is quite different between the two. With both solvents and both thiolates the reaction is endothermic, by about 30 kJ/mol for the alkyl-thiol and double that for the aryl-thiol. The large difference in stability of the products may be due to differences in the selenylsulfides produced. The Se−S bond is 0.04 Å longer with thiophenol than methanethiol, indicating a weaker, less stable bond. There will be significant steric clash between the two phenyl rings, which sit perpendicular to each other, whereas the methyl group is significantly smaller and will have little interaction with the phenyl ring in DMBS.

Experimental studies have found that this step of the DMBS catalytic cycle is very fast; in some studies, the researchers were unable to observe selenenic acid at

Table 6.1: Reaction energies for nucleophilic attack of various thiols on the selenenic acid of DMBS, in kJ/mol. Energy barriers are the difference between the TS and RC and total energy of reaction is the difference between the PC and RC.

Nucleophile	Solvent	$\Delta G^{\ddagger}$	ΔG_{rxn}
MeS^-	Water	175.4	29.3
MeS^-	Methanol	178.5	31.2
PhS^-	Water	177.0	66.7
PhS^-	Methanol	181.2	67.0
MeSH	Water	49.1	-76.3
PhSH	Water	43.0	-66.1

all. Knowing this, it is unlikely that the reaction proceeds via nucleophilic attack of a thiolate. This mechanism is predicted not only to have very large energy barriers, but also to be endothermic by up to 67 kJ/mol. Much of this is likely due to the hydroxide ion, which is a very poor leaving group. If a neutral thiol is used instead, it could transfer its hydrogen to the hydroxyl group, forming water as the leaving group, which should be more favourable.

6.3.2 Thiol attack

As mentioned above, in their modelling of the GPx mechanism, the Morokuma group[55] found a concerted transfer of the thiol hydrogen to the leaving hydroxyl group with formation of the Se−S bond. This results in a four-centred transition state that is likely to be under a lot of strain. Additionally, in the reactant complex, it is more favourable for the hydroxyl group to hydrogen bond to the sulfur than for the thiol to bond to the oxygen. In the GPx active site, this is overcome through a nearby glutamine residue that accepts a hydrogen bond from the hydroxyl to the carbonyl in its side chain, allowing the incoming thiol to donate a hydrogen bond to the oxygen. However, there are no such groups in DMBS to facilitate this type of interaction, so it is unlikely that such a complex would be formed. Instead, if the reaction is performed in a protic solvent, solvent molecules could be used as a proton shuttle, transferring the thiol hydrogen to the hydroxyl group through a chain of solvent molecules. A reaction of this type is outlined schematically in figure 6.3.

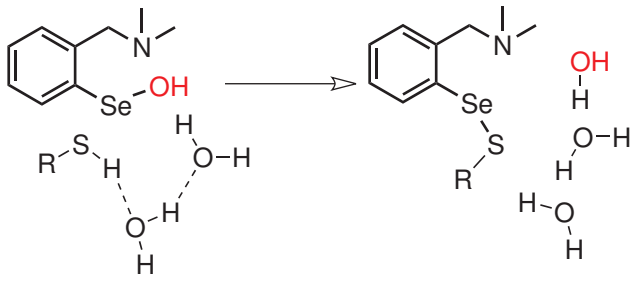


Figure 6.3: Solvent-assisted nucleophilic attack of a thiol on the selenenic acid of DMBS.

Attempts were made to follow the reaction of the selenenic acid of DMBS with methanethiol both with and without explicit solvent molecules. As expected, without solvent molecules a reactant complex forms with the hydroxyl group donating a hydrogen bond to the sulfur in the thiol. With this arrangement of the atoms, there is no way to transfer the thiol hydrogen to the hydroxyl group. To facilitate this, solvent molecules were added to the reactant complex to act as a proton shuttle. A single water molecule does not have enough flexibility to accept both a hydrogen bond from the thiol and also donate a hydrogen bond to the hydroxyl group. Therefore, a single water molecule will not be effective as a proton shuttle. The addition of a second water molecule forms a chain in the reactant complex that accepts a hydrogen bond from the thiol and donates to the hydroxyl (see figure 6.5a). The transition state shows a concerted transfer of three protons (from thiol to water, water to water, and water to hydroxyl) with a lengthening of the Se−O bond (see figure 6.5b). Following the reaction path around the transition state, using IRC, shows that the hydrogens are transferred in a "cascade" mechanism, with the thiol hydrogen transferred first, followed by the water-to-water transfer, and finally the water-to-hydroxyl transfer.

The concerted reaction, using two water molecules as proton shuttles, has a reasonably low energy barrier shown in the reaction profile in figure 6.4, varying slightly depending on the thiol used. If methanethiol is used, the energy barrier is 49.1 kJ/mol, while the barrier for thiophenol is 43.0 kJ/mol. The lower barrier with thiophenol is likely due to its lower pK_a.[169] These barriers are roughly a quarter of those found for the same reaction using the thiolate. Additionally, they are about the same as

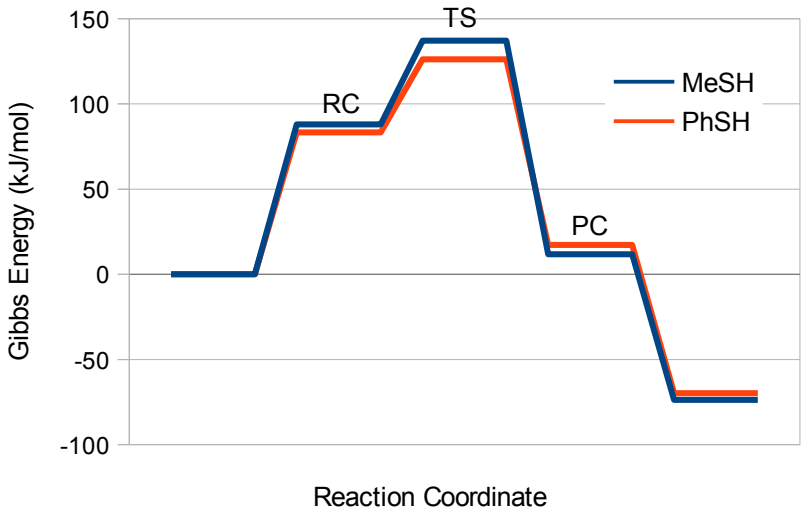
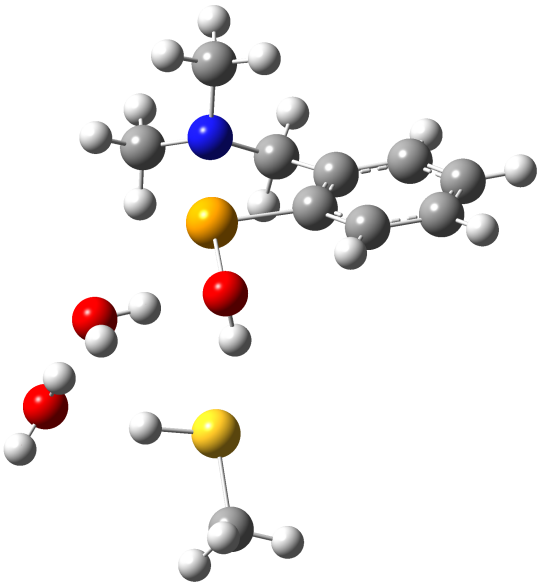


Figure 6.4: Reaction profile for the reaction of a thiol with the selenenic acid of DMBS assisted by two water molecules.

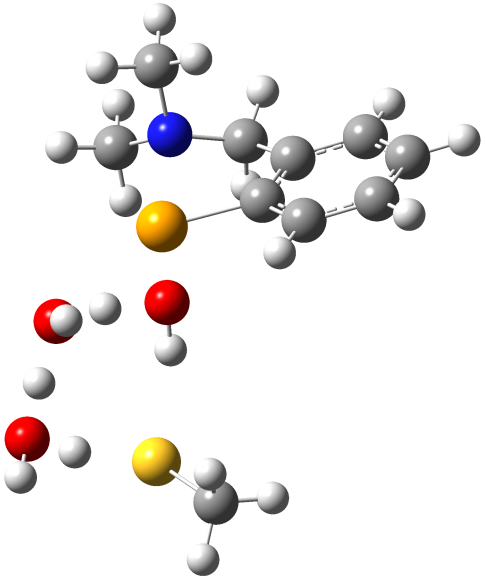
found for the preceding peroxide reduction step of this reaction, discussed in chapter 5, which is in agreement with experimental kinetic studies that indicate that these two steps proceed quickly. Unlike when using thiolates, the reaction of a thiol with the selenenic acid of DMBS is exothermic by about 65-75 kJ/mol, demonstrating the increased stability of water as a leaving group compared to a hydroxide ion.

The reactant complex (RC) energy for these systems is 80 to 90 kJ/mol higher than the infinitely separated reactants, however the magnitude of this is not physically meaningful. When modelling systems in the gas phase, RC formation is stabilizing, however with implicit solvation methods, it is common for the RC energy to be higher than the infinitely separated reactants.

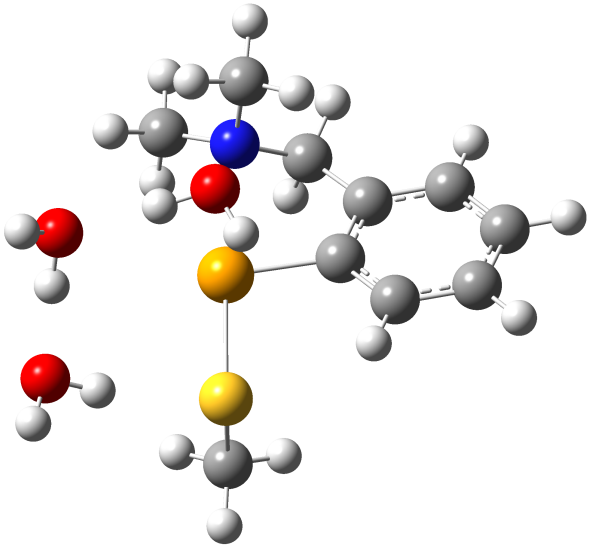
The QTAIM molecular graphs for the reactant complex (RC) and transition state (TS) of the solvent-assisted reaction mechanism are shown in figure 6.6. These can be used to learn more about the proton transfers in this mechanism by comparing the electron densities at the bond critical points in the RC and TS. The QTAIM results, agree well with mechanism observed through IRC. In the transition state, the first proton transfer (from the thiol to a water) is nearly complete, with ρ_{O-H} of 0.236 au



(a) Reactant Complex

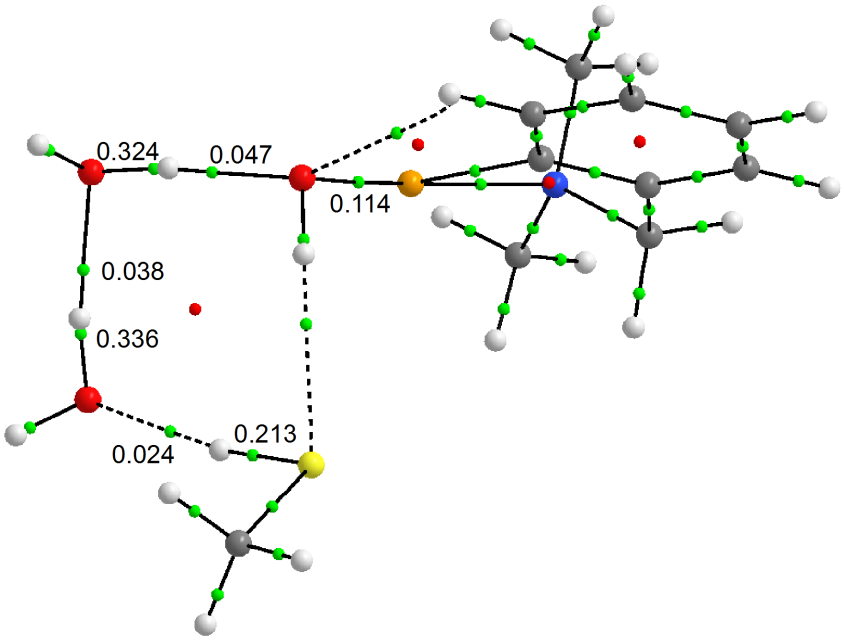


(b) Transition State

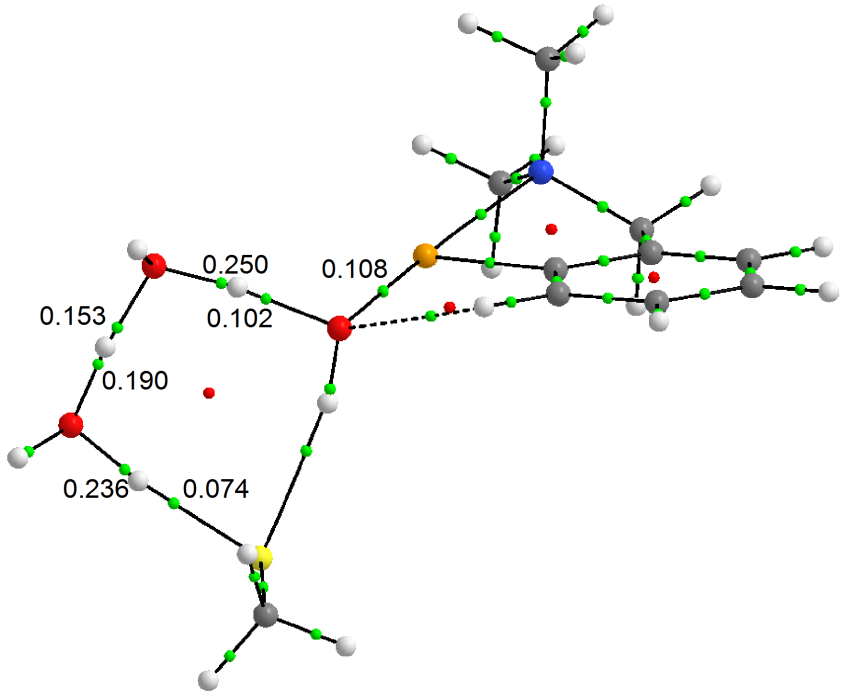


(c) Product Complex

Figure 6.5: Structures found for the water-assisted reaction of methanethiol with the selenenic acid of DMBS.



(a) Reactant Complex



(b) Transition State

Figure 6.6: Molecular graphs of the water-assisted reaction of methanethiol with the selenenic acid of DMBS. Bond and ring critical points are denoted by green and red spheres, respectively. Bond critical point densities given in atomic units.

(note that ρ_{O-H} for the O−H bonds not participating in the proton transfer are about 0.36 au). The second proton transfer (from water to water) is about half complete with ρ_{O-H} of 0.190 and 0.153 au for the bonds breaking and forming, respectively. The final proton transfer (from water to the acid hydroxyl) has only just begun with ρ_{O-H} of 0.250 and 0.102 au for the water and hydroxyl oxygens, respectively.

6.4 Conclusions

Two mechanisms for the reaction of the selenenic acid of DMBS with a thiol have been modelled, using either a highly nucleophilic thiolate or a more moderate thiol. It is found that the use of a thiol is more favourable than the thiolate.

With a thiolate as the nucleophile, the energy barrier for this reaction is around 175 kJ/mol. This barrier is much higher than expected, since this step of the catalytic cycle proceeds very quickly. Additionally, the reaction is predicted to be endothermic by 31 or 67 kJ/mol for the thiolates of methanethiol and thiophenol, respectively. Both the high barrier and endothermic products are related to the leaving group in this mechanism, a hydroxide ion. Hydroxide is a poor leaving group, so there will be a high barrier associated with breaking the Se−O bond and it will be less stable than the reactants.

An alternative mechanism uses a thiol as the nucleophile. In this mechanism, the hydrogen on the sulfur is transferred to the hydroxyl of the selenenic acid, making water the leaving group. This proton transfer does not occur directly, but instead makes use of two molecules of water from the surrounding solvent as a proton shuttle. This mechanism has a low energy barrier of 49 and 43 kJ/mol for methanethiol and thiophenol, respectively. It is also exothermic overall, by 76 and 66 kJ/mol for methanethiol and thiophenol, respectively, easily making this mechanism more favourable than the one using thiolates.

Chapter 7

Reduction of Selenylsulfide by Thiol

In this chapter, the final step in the catalytic reduction of peroxides by thiols is studied. This involves the regeneration of the active form of the organoselenium catalyst. Specifically, a selenylsulfide reacts with a thiol to produce a selenol and a disulfide. As will be discussed below, the chemistry of this reaction is poorly understood compared to the thiol–disulfide analogue.

7.1 Introduction

A common refrain in publications studying the reactions of selenylsulfides and the mechanisms of GPx mimics is "...in contrast to thiol–disulfide exchange reactions, only a few studies of analogous reactions involving selenium have been published." [170] Indeed, there are very few in-depth examinations of the thiol–selenylsulfide exchange reaction.[64,167,170,171] Of these, very few study the reaction using compounds commonly used experimentally with GPx mimics, instead opting to use smaller model systems like benzeneselenol, methaneselenol, methanethiol, or hydrogen sulfide.

A computational model of the catalytic cycle of GPx mimics has been published by Benkova *et al.*[167] using benzeneselenol, hydrogen peroxide, and hydrogen sulfide. Their results predict that the reaction of the selenylsulfide with hydrogen sulfide follows a one-step mechanism.

$$\mathrm{PhSe{-}SH} + \mathrm{H_2S} \longrightarrow \mathrm{PhSeH} + \mathrm{HS{-}SH}$$

This reaction proceeds via an energy barrier twice as high as the other two steps in the cycle, easily identifying this as the rate-determining step. Their isolated reaction

mechanism involves a four-centred intermediate in the transition state with simultaneous formation of the selenium-hydrogen and sulfur-sulfur bonds. It is possible that introducing a water molecule to act as a proton shuttle, forming a six-centred intermediate, could lower that barrier.

Bachrach *et al.*[171] have performed a computational study of nucleophilic attack of thiolates and selenolates on diselenides and selenylsulfides. Their work examines the reactions using hydrogen and methyl-substituted chalcogens in the gas phase. For the nucleophilic attack of a thiolate on a selenylsulfide, they followed the reaction at both ends of the Se−S bond. In both cases, a two-step addition-elimination reaction with a stable intermediate is predicted. In the first step, the thiolate attacks one end of the selenylsulfide bond, forming an intermediate structure with a hypervalent central atom. In the second step, the intermedate breaks apart forming a selenolate and a disulfide, for attack at sulfur, or a thiolate and a selenylsulfide, for attack at selenium.

$$R_1Se{-}SR_2 + R_3S^- \longrightarrow [R_1Se{-}SR_2{-}SR_3]^- \longrightarrow R_1Se^- + R_2S{-}SR_3$$

$$R_1Se{-}SR_2 + R_3S^- \longrightarrow [R_3S{-}SeR_1{-}SR_2]^- \longrightarrow R_3S{-}SeR_1 + R_2S^-$$

Their work predicts that attack at selenium in this system is more favourable than at sulfur. Sarma and Mugesh[64] have found that in the selenylsulfide intermediate of ebselen, thiol attack at selenium is competitive with attack at sulfur, leading to an unproductive thiol exchange reaction which diminishes its efficiency as a GPx mimic.

7.2 Computational Methods

All calculations were performed with the Gaussian 09[143] suite of programs. Geometry optimizations were performed using the B3PW91 hybrid DFT functional, composed of Becke's three-parameter exchange correlation functional (B3)[122] and Perdew and Wang's correlation functional (PW91)[150] with the 6-31+G(d,p) Pople basis set, as suggested by the benchmark in chapter 4. Transition states were found

with Schlegel's synchronous transit-guided quasi-Newton (STQN) method[155,156] and the reaction path was followed using intrinsic reaction coordinate (IRC)[157,158] calculations. Frequency calculations were performed on all optimized structures at the same level of theory to confirm whether the structure is a local minimum or first-order saddle point and to obtain thermochemical corrections. Accurate energies were obtained for all systems at the B3PW91/6-311+G(2df,p) level. Solvent effects for an aqueous environment (dielectric constant of 78.3553) were included in all calculations using the conductor-like polarizable continuum model (CPCM).

The topology of the electron density was studied using quantum theory of atoms in molecules (QTAIM) calculations performed using AIMAll (Version 11.10.16).[168]

7.3 Results and Discussion

In the following section, the results of a mechanism search using methyl-substituted model systems are presented. These have been used as an initial way to probe the reactivity of selenylsulfide systems. Using what was learned from the model systems, preliminary results using the selenylsulfide of DMBS are discussed.

7.3.1 Small Model Systems

The addition-elimination mechanism proposed by Bachrach *et al.*[171] was examined first. In the gas phase, this mechanism was successfully modelled using methyl substituted model systems, shown in figure 7.1. However, when attempting to re-optimize the geometries in an implicit aqueous environment, TS_1, the transition state for the addition reaction, cannot be isolated. In those attempts that converged to a first-order saddle point, the two sulfur atoms were more than 3.5 Å apart, and the movement of the atoms in the vibrational mode associated with the imaginary frequency did not indicate the formation of a bond between the two sulfur atoms, but rather a rotation in the thiolate methyl group.

Since Bachrach's addition-elimination mechanism is unstable in solution, a mechanism that can work in an aqueous environment was pursued. As in section 6.3.2, it

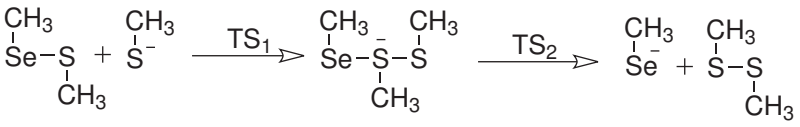
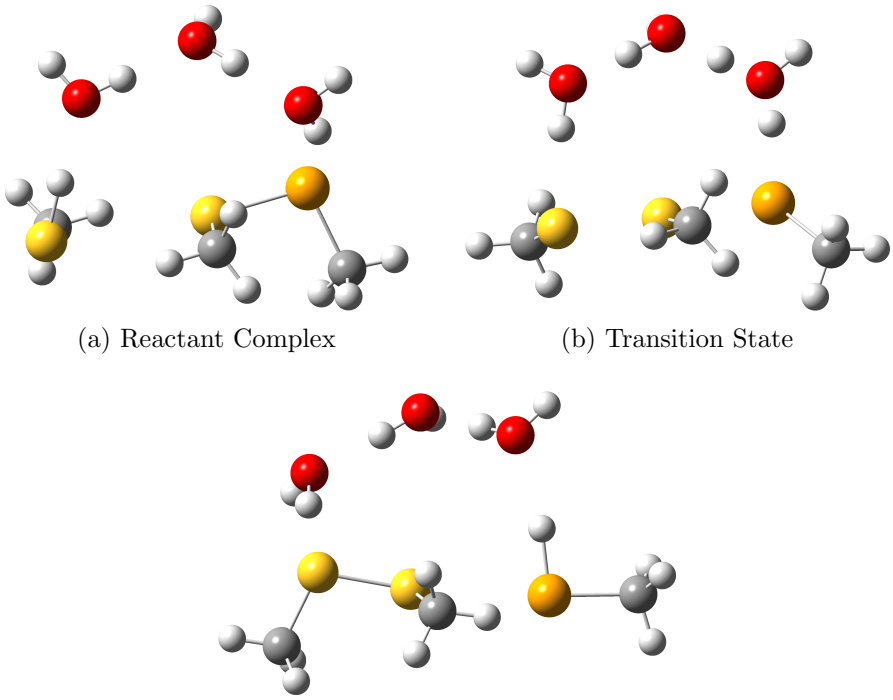


Figure 7.1: Addition-elimination reaction of a selenylsulfide and thiolate, as proposed by Bachrach *et al.*[171]

is desirable to use a thiol as the nucleophile and include a chain of solvent molecules to transfer the thiol hydrogen to the selenol leaving group. A water-assisted mechanism was isolated for methyl substituted model systems. It can be seen in figure 7.2 that there is a nearly linear arrangement of the three chalcogen atoms throughout the reaction with three water molecules parallel to them. Like in the water-assisted reaction of thiol with the selenenic acid presented in section 6.3.2, the water molecules act as a shuttle for the thiol hydrogen, transferring a proton to the selenium centre to produce a selenol leaving group. With the methyl-substituted model systems, this mechanism has a Gibbs energy barrier of 120 kJ/mol. However, the reactant complex geometry of the water-assisted mechanism is unstable when substituting the methyl group on selenium for the benzylamine found in DMBS. Geometry optimizations of the RC converge to a structure with the two sulfur atoms more than 6 Å apart, while maintaining the chain of water molecules connecting the thiol to the selenium centre.

In a selenylsulfide bond, the charge is expected to be polarized toward the sulfur, since it is more electronegative than selenium. This is reproduced in dimethylselenylsulfide, which has QTAIM charges on the sulfur and selenium atoms of -0.169 and 0.157 au, respectively. The sulfur atom in methanethiol has a small negative charge of -0.041 au, so forming a complex where the two sulfur atoms interact will be somewhat unfavourable. In the selenylsulfide of DMBS, the Se−S bond is even more polarized, with charges of -0.264 and 0.280 au on sulfur and selenium, respectively. This increased charge is due to a further polarization of the selenium density through its interaction with the highly anionic nitrogen atom (q_N = -1.050).

The differences in the electronic environment around the Se−S bond between the methyl-substituted model system and selenylsulfide of DMBS can also be observed



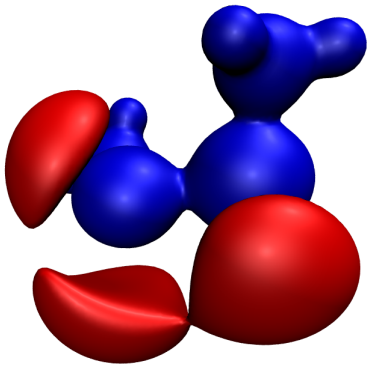
(a) Reactant Complex (b) Transition State

(c) Product Complex

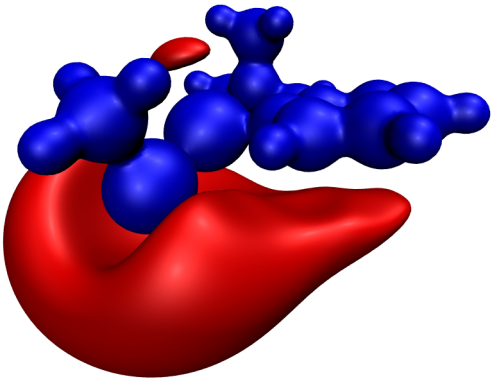
Figure 7.2: Reaction of methanethiol with dimethyl selenylsulfide using water as a proton shuttle.

through their electrostatic potentials (figure 7.3). In the methyl-model system, there are small, compact regions of negative potential around both the sulfur and selenium atoms, but neither centre has complete coverage. Conversely, in the selenylsulfide of DMBS, there is a large, diffuse region of negative potential surrounding the exposed surface of the sulfur centre. This region extends to partially cover the selenium centre and over one face of the phenyl ring. In the attempts to isolate a RC using the selenylsulfide of DMBS with methanethiol, the thiol molecule always optimised to a position outside of the negative potential region shown in figure 7.3.

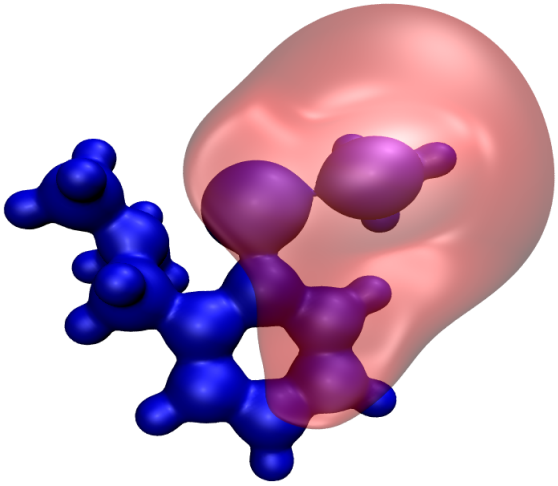
Small model systems are often used in computational chemistry to probe reactions quickly, but care must be exercised in attributing too much significance to results obtained using these models. As demonstrated above, a reaction mechanism can be modelled using small systems that is not possible with the system of interest. In this case, it is due to the very different electronic environments caused by methyl and benzylamine substituents.



(a) Dimethyl selenylsulfide, note that the sulfur atom is on the left and the selenium on the right.



(b) Selenylsulfide of DMBS



(c) Selenylsulfide of DMBS from a different angle, note that the region of negative potential has been made partially transparent to show positive regions underneath.

Figure 7.3: Electrostatic potential maps of two selenylsulfides. Blue isosurface corresponds to an electrostatic potential of 0.4 au and the red isosurface corresponds to -0.03 au.

7.3.2 Attempts to Model DMBS

Mugesh and coworkers have suggested that it is actually an anionic thiolate that acts as the nucleophile in this step of the reaction.[76] Normally, the pH of these reactions is not high enough to deprotonate a thiol, but they propose that the amine in DMBS abstracts the hydrogen to give a local concentration of thiolate. In addition to the changes within DMBS, this would also provide a source of thiolate to act as the nucleophile.

The cationic form of the selenylsulfide of DMBS, with the amine protonated, has quite different characteristics than the neutral form discussed previously. In the neutral form, the N−Se−S bond angle is 177°, with the Se−S bond 20° away from co-planarity with the phenyl ring. With the amine protonated, the N−Se−S angle is 95°, with the Se−S bond roughly perpendicular to the plane of the phenyl ring. There are also significant changes in electron population of the sulfur and selenium centres, with charges of -0.128 and 0.225 au, respectively, the first time the charges of these atoms are not complementary. Additionally, the electrostatic potential map of this species has no negative component.

Other mechanisms for the reaction of methanethiolate with the protonated form of the selenylsulfide have been modelled, but no transition states were isolated.

7.4 Conclusions

The two-step addition-elimination reaction proposed by Bachrach *et al.*[171] for the reaction of a thiolate with selenylsulfide was found to be unstable in an implicit solvent environment. An alternative one-step mechanism has been identified using methanethiol and dimethyl selenylsulfide that incorporates three water molecules as a proton shuttle. However, the reactant complex that leads to this reaction cannot be isolated for the selenylsulfide of DMBS. In all attempts it converges to a complex with no S···S interaction.

Studying the changes in the electronic environment around the Se−S bonds in

DMBS and dimethyl selenylsulfide provides some insight into why the reactant complex found using the model system cannot be isolated using DMBS. The charge difference between the sulfur and selenium centres is much larger in DMBS and additionally it has a large area of negative electrostatic potential surrounding the sulfur centre and extending over the phenyl ring. This makes it unfavourable for the thiol to form a close association with the selenylsulfide bond, instead forming a complex with the thiol outside the negative potential region surrounding DMBS.

Chapter 8

Effects of Monosubstitution on the Peroxide Reduction Mechanism

In this chapter, seven functional groups, ranging from strongly electron-withdrawing to strongly electron-donating, are introduced in the ortho, meta, and para positions relative to the selenol in *N,N*-dimethylbenzylamine-2-selenol (DMBS). By modelling these compounds using the peroxide reduction mechanism reported in chapter 5, the purely electronic effects of these substituents in the meta and para positions can be studied and compared to substitutions in the ortho position, where both electronic and steric effects can influence the reaction.

8.1 Motivation

Since the proposal of ebselen as a glutathione peroxidase mimic nearly thirty years ago,[57,58] much work has been done to determine the most important factors for activity of GPx mimics. One of the key factors is an amine functionality adjacent to the selenium moiety, which allows for a Se···N interaction that appears to be important for GPx-like activity.[73,77,82] This closely resembles the GPx active site, which has an arginine residue near the selenocysteine.[172]

Another important factor for the activity of organoselenium compounds as GPx mimics is the selenol group, or rather the selenolate. Many studies have either observed an anionic selenium experimentally[48,73,173] or predicted it computationally to have greater activity than the neutral selenol as a reducing agent[55,65] in both GPx

and mimic compounds. In chapter 5, it was shown that the mechanism for reduction of hydrogen peroxide by DMBS in its neutral selenol form is a two-step process with large energy barriers for each step. Conversely, the zwitterionic form shows a one-step mechanism with a significantly lower barrier. The selenolate form, which has a more anionic selenium centre than the zwitterion, has an even lower barrier. Therefore, it seems likely that altering the electron density on the selenium atom in the zwitterionic form, by introducing electron withdrawing or donating groups to the phenyl ring, would affect the barrier for peroxide reduction.

In 1989,[174] it was observed that introducing a nitro group ortho to the selenium in ebselen increased its GPx-like activity by a factor of nine. More recently, Bhabak and Mugesh[85] have shown that introducing a methoxy group ortho to the selenol in DMBS increases the GPx-like activity by an order of magnitude, an effect they attribute to a combination of steric and electronic effects. On the other hand, Pearson and Boyd[66] have reported that introducing a methoxy group ortho to the selenium in ebselen increases the energy barrier for hydrogen peroxide reduction by about 21 kJ/mol. In the same study, a teritary-butyl group in the same position also increased the energy barrier by the same amount, indicating a primarily steric effect.

8.2 Computational Methods

Calculations were performed with the Gaussian 09[143] suite of programs. Geometry optimizations were performed with the B3PW91 hybrid DFT functional, composed of Becke's three-parameter exchange correlation functional (B3)[122] and the correlation functional proposed by Perdew and Wang (PW91)[150] with the 6-31+G(d,p) Pople basis set, as suggested by Pearson *et al.*[142] Transition states were found with Schlegel's synchronous transit-guided quasi-Newton (STQN) method[155,156] and the reaction path was followed using intrinsic reaction coordinate (IRC)[157,158] calculations. Frequency calculations were performed on all optimized structures at the same level of theory to confirm whether the structure is a local minimum or first-order

saddle point and to obtain thermochemical corrections. Accurate energies were obtained for all systems at the B3PW91/6-311++G(3df,3pd) level. Solvent effects for an aqueous environment (dielectric constant of 78.3553) were included in single-point calculations using the conductor-like polarizable continuum model (CPCM).

The topology of the electron density was studied using quantum theory of atoms in molecules (QTAIM) calculations[128,129] performed using AIMAll (Version 10.03.25).[168]

8.3 Results and Discussion

8.3.1 Zwitterion formation

Experimental evidence suggests that the zwitterionic form of DMBS is the predominant form in water at room temperature[73] and therefore is likely the form that reacts with peroxides. To determine the relative stability of the zwitterion, both the selenol and zwitterionic forms have been modelled. A transition state search was also performed to connect the selenol to the zwitterion, which found a transition state lower in energy than the selenol for all systems, indicating a barrierless conversion from the selenol to the zwitterion.

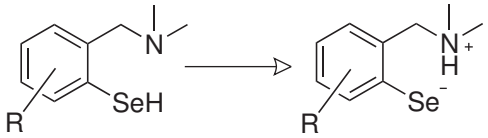


Figure 8.1: Selenol to zwitterion formation reaction. The substituents listed are placed in the ortho, meta, or para positions relative to the selenol.

In all systems studied, the zwitterion is more stable than the selenol, and the conversion from the selenol to the zwitterion is, from the transition states identified, energetically downhill. These results are summarized in table 8.1. The zwitterion of unsubstituted DMBS is 23.1 kJ/mol more stable than the selenol. In the substituted systems, the zwitterions are 20-30 kJ/mol more stable than the selenols. Previous work by Bhabak and Mugesh[85] determined that the zwitterionic form of DMBS is

about 21 kJ/mol more stable than the selenol, which is in agreement with our results. They also found that ortho-methoxy substituted DMBS has a zwitterionic form about 38 kJ/mol more stable than the selenol, which is significantly more stable than reported here.

Table 8.1: Gibbs energies for formation of the zwitterion from the selenol of DMBS (energies relative to the selenol form in kJ/mol).

	Δ G_{rxn}		
Substitution	ortho	meta	para
NH_2	-30.5	-24.1	-19.5
OCH_3	-19.1	-24.3	-20.8
CH_3	-22.0	-28.5	-22.4
F	-25.9	-26.2	-22.5
CN	-30.3	-26.6	-27.3
CF_3	-25.6	-34.9	-26.6
NO_2	-20.5	-27.1	-28.7
Unsubstituted		-23.1	

Using QTAIM analysis, the electron density changes during the proton transfer can the followed. It is observed that in the selenol the Se−H bond critical point (BCP) density (ρ_{BCP}^{Se-H}) has a value of 0.168 au and $\rho_{BCP}^{N\cdots H}$ = 0.042 au. In the zwitterion, after the proton transfer has been completed, ρ_{BCP} is 0.072 and 0.236 au for Se···H and N−H, respectively. This demonstrates that the N−H bond is stronger in the zwitterion than the Se−H bond in the selenol, with ρ_{BCP} 0.068 au higher, and the Se···H interaction in the zwitterion is also stronger than the N···H interaction in the selenol, with ρ_{BCP} 0.030 au higher. This combination of two stronger interactions provides an explanation for the increased stability of the zwitterion relative to the selenol. Similar trends are observed for the substituted DMBSs.

Examining the atomic charges of the selenol and the zwitterionic forms (table 8.2) provides an explanation for the differences in reactivity of these two species with hydrogen peroxide that were reported in chapter 5. The number of electrons in the selenium atomic basin ($N(\Omega_{Se})$), for the unsubstituted species, is 33.871 and 34.311 au for the selenol and zwitterion, respectively, giving atomic charges (q_{Se}) of 0.129 and -0.311 au, respectively. The zwitterion's charge is nearly half an au higher than

Table 8.2: QTAIM charges for the selenium atom in DMBS (au).

	ortho		meta		para	
Substitution	$q_{selenol}$	$q_{zwitterion}$	$q_{selenol}$	$q_{zwitterion}$	$q_{selenol}$	$q_{zwitterion}$
NH_2	0.082	-0.351	0.118	-0.325	0.145	-0.326
OCH_3	0.188	-0.262	0.131	-0.312	0.143	-0.320
CH_3	0.102	-0.317	0.127	-0.316	0.132	-0.316
F	0.186	-0.254	0.131	-0.296	0.138	-0.308
CN	0.163	-0.232	0.136	-0.279	0.111	-0.265
CF_3	0.181	-0.229	0.135	-0.286	0.124	-0.280
NO_2	0.298	-0.162	0.144	-0.271	0.074	-0.244
Unsubstituted			0.129	-0.311		

the selenol, which makes it a stronger nucleophile toward hydrogen peroxide, thus allowing it to react more efficiently. The positive charge in the selenol, combined with its valency of two, makes a poor nucleophile in comparison with the zwitterion which has an anionic selenium with a valency of one. These two factors lead to the two-step reaction found for the selenol with hydrogen peroxide, but allow the zwitterion to react via a one-step process.

By studying the electron density on the atoms nearby the selenium, it is possible to identify where the additional charge comes from during the selenol to zwitterion conversion. In the selenol, $N(\Omega_H) = 0.961$ au for the hydrogen bonded to the selenium atom. After it is transferred to the nitrogen, $N(\Omega_H) = 0.582$ au, meaning that the hydrogen has lost 0.379 au of charge. Over the course of the proton transfer, the nitrogen gains an average of 0.026 au of charge, while the selenium gains 0.426 au implying that the majority of the charge lost by the hydrogen is transferred to the selenium atom.

8.3.2 Peroxide reduction

To determine the effects of changing the electron density on the selenium atom in DMBS, the reaction of substituted systems with hydrogen peroxide was modelled using the same mechanism found in chapter 5. The reaction energies are outlined in table 8.3. The energy barriers for the substituted DMBSs range from 56.9 to 62.2 kJ/mol, with the unsubstituted DMBS showing a barrier of 59.3 kJ/mol. Substitution

with either electron donating (NH_2, OCH_3, and CH_3) or electron withdrawing groups (F, CN, CF_3, and NO_2) had very minor effects on the energy barrier, only varying $\Delta G^{\ddagger}$ by up to 3.5 kJ/mol.

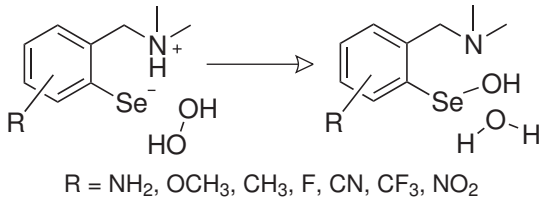


Figure 8.2: Peroxide reduction reaction being modelled. The substituents listed are placed in the ortho, meta, or para positions relative to the selenol.

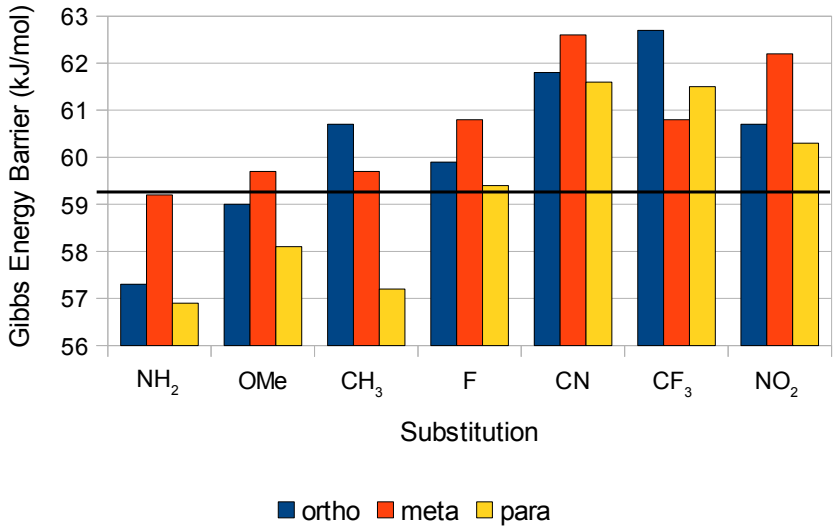


Figure 8.3: Gibbs energy barriers for peroxide reduction reaction with substituted DMBS. Heavy black line indicates barrier of the reaction with the unsubstituted compound.

Within the meta and para substitution patterns, following the series from strongly electron donating to strongly electron withdrawing substituents shows that, in general, the energy barrier increases and stability of the products relative to reactants decreases. In the ortho substitutions, there is less of a clear trend in energies going from electron donating to electron withdrawing groups, however substitution with electron donating groups generally results in lower barriers than substitution with

Table 8.3: Gibbs energies for reaction of zwitterionic DMBS with hydrogen peroxide (in kJ/mol). Δ G‡ is difference between the TS and RC energies, while Δ G$_{rxn}$ is the difference between the PC and RC energies.

	ortho		meta		para	
Substitution	Δ G‡	Δ G$_{rxn}$	Δ G‡	Δ G$_{rxn}$	Δ G‡	Δ G$_{rxn}$
NH_2	57.3	-193.4	59.2	-188.9	56.9	-205.0
OCH_3	59.0	-204.5	59.7	-188.3	58.1	-199.0
CH_3	60.7	-198.2	59.7	-189.4	57.2	-193.7
F	59.9	-181.8	60.8	-185.2	59.4	-190.3
CN	61.8	-173.4	62.6	-179.8	61.6	-175.3
CF_3	62.7	-174.9	60.8	-182.3	61.5	-179.2
NO_2	60.7	-178.2	62.2	-180.2	60.3	-170.2
Unsubstituted			59.3	-188.0		

electron withdrawing groups.

Electronic effects

By comparing the energy barrier for the peroxide reduction reaction to the atomic charge on the selenium atom in the zwitterionic form shown in figure 8.4, the effect of the charge on the energy barrier can be determined. It is seen that the meta and para substituted systems, as well as the unsubstituted system, follow a linear trend with more anionic selenium centres showing lower energy barriers. However, the ortho substitutions are more scattered, with no clear trend. Previous work,[85] has shown that ortho substituents can greatly increase the GPx-like activity of benzylamine selenols and attributed this change in activity to a combination of electronic and steric effects.

The lowest energy barrier observed was with the p-NH_2 substituent, the strongest electron donating group included in this study. The highest barrier is observed for the m-CN substituent, a moderate electron withdrawing group, although m-NO_2, a strongly electron withdrawing group, has a similar barrier. These same substituents had the highest and lowest charge on the selenium atom in the ortho substituted systems, with o-NH_2 showing an atomic charge on selenium more than twice that of o-NO_2. However, the difference in energy barrier for these two ortho-substituted

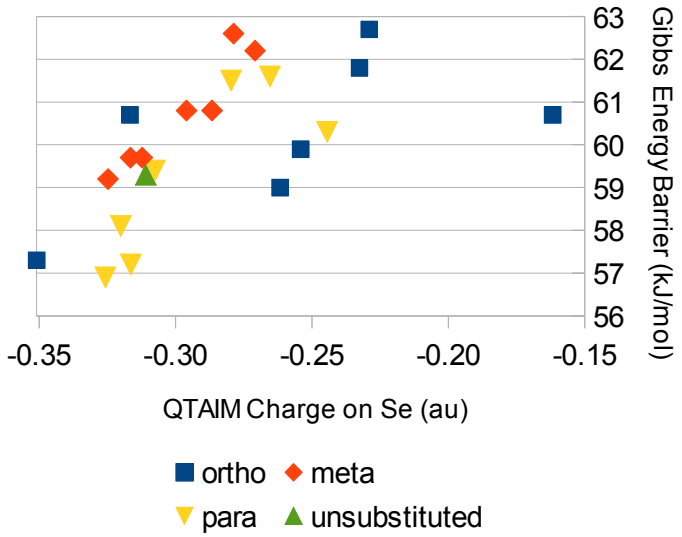


Figure 8.4: Charge on selenium atom in the zwitterionic form of DMBS vs Gibbs energy barrier for peroxide reduction reaction

systems is much lower. This demonstrates that a strongly electron withdrawing or donating substituent adjacent to the selenium moiety has a large effect on the electron density of the selenium atom. However, it is mediated by other effects, such as sterics or polarization of charge induced by a group in close proximity to the reacting centre.

Some of these effects can be seen in the QTAIM analysis as bond paths between the selenium atom and the substituent, which are found for the o-NH_2, o-NO_2, and o-CF_3 systems. At the Se–substituent BCP, ρ_{BCP} in the amino and nitro substituted systems is 0.017 au. The trifluoro-methyl system shows two Se−F bond paths with ρ_{BCP} of 0.010 and 0.009 au. The presence of these bond paths indicate an interaction between the selenium and the substituent, which will contribute to both electronic and steric effects.

The electrostatic potential can provide additional insight into the changes in the electronic environment around the selenium center introduced by ortho substitution. Electrostatic potential diagrams for the zwitterionic form of unsubstituted DMBS and four of the ortho substituted compounds are shown in figure 8.5. In the unsubstituted system, a large region of negative potential is observed around the selenium atom, in the same area where the hydrogen peroxide interacts with the selenium atom to form

the reactant complex. In the substituted systems, this region is extended onto the adjacent functional group. Pearson and Boyd,[66] have also found an extended region of negative electrostatic potential for o-NO_2 substituted ebselen. While this would not prevent the peroxide from associating with the selenium centre, it does present a much larger surface to the peroxide, which makes it less likely that the peroxide would interact with the selenium preferentially.

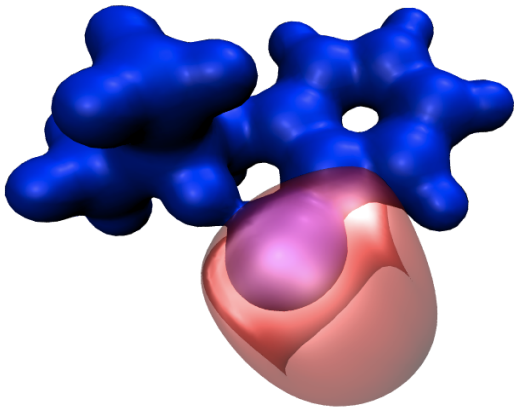
Steric effects

The steric effects arising from ortho substituents can be studied by calculating the atomic volume of the selenium atom via QTAIM. Through this method, the steric effects can be examined both quantitatively by studying changes in the calculated atomic volume and qualitatively by visualising the shape of the selenium atomic basin.

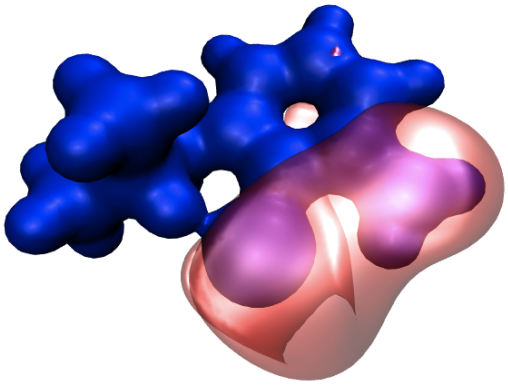
Table 8.4: The volume of the selenium atom in the zwitterionic form of each molecule, calculated at the 0.001 au density cutoff, in au^3.

Substitution	ortho	meta	para
NH_2	282.2	288.4	288.2
OCH_3	278.1	287.1	287.7
CH_3	280.4	287.6	287.6
F	280.0	286.0	286.6
CN	276.6	284.4	283.8
CF_3	267.6	284.9	284.8
NO_2	263.3	283.8	282.5
Unsubstituted		286.9	

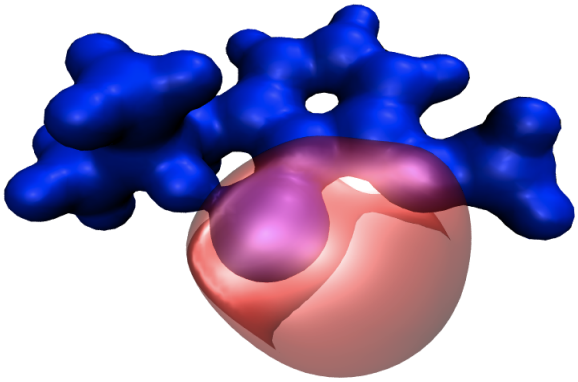
The meta and para substitutions show only minor changes in atomic volume compared to the unsubstituted DMBS, summarized in table 8.4, which arise from the expected contraction or expansion of the atom due to decreased or increased charge. In the ortho series however, the changes in atomic volume are much more pronounced, and are all lower than the unsubstituted system, showing that functional group proximity has an effect on the volume beyond the small electronic effect observed for the other substitutions. The largest decrease in volume is observed for o-NO_2 substitution, with an 8% change. The contraction of the atom due decreased charge for the



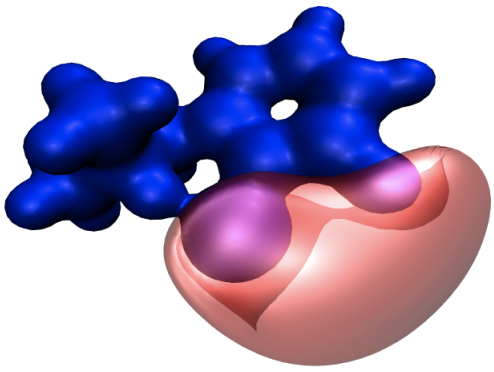
(a) unsubstituted zwitterionic form of DMBS



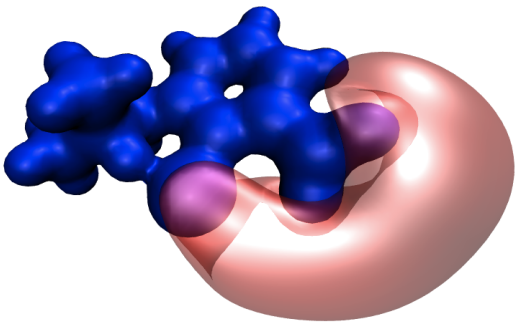
(b) o-NH_2 zwitterionic form of DMBS



(c) o-OCH_3 zwitterionic form of DMBS



(d) o-F zwitterionic form of DMBS



(e) o-NO_2 zwitterionic form of DMBS

Figure 8.5: Electrostatic potential maps of various systems. Blue isosurface corresponds to an electrostatic potential of 0.2 au and the red isosurface corresponds to -0.04 au. Negative isosurfaces have been faded to show positive regions underneath.

other nitro substituted compounds is about 1.5%, indicating that the change due to steric effects is more than 4 times greater than electronic effects.

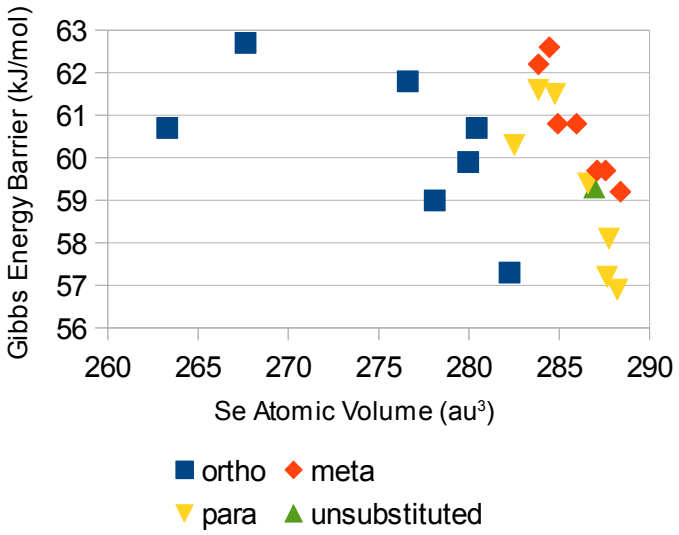
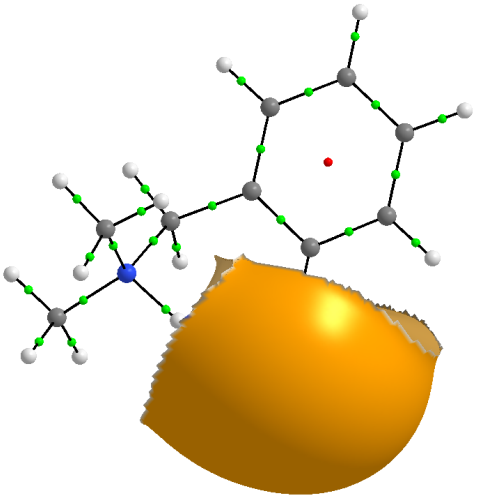


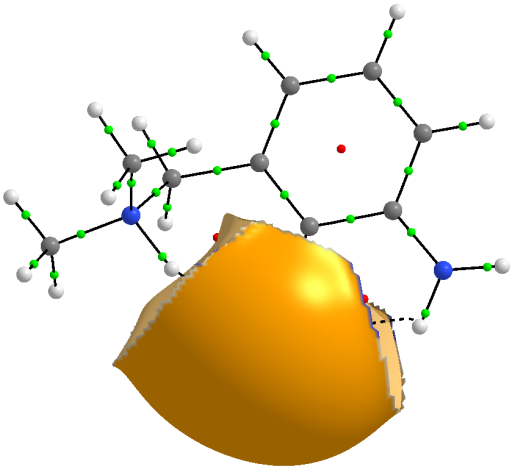
Figure 8.6: Volume of the selenium atom in the zwitterionic form of DMBS vs Gibbs energy barrier for peroxide reduction reaction

The effect of changes in atomic volume on the energy barrier for the peroxide reduction reaction can be seen in figure 8.6. The trend observed for the meta and para substitutions mirrors that seen in figure 8.4 for the changes in atomic charge, confirming that this minor contraction in the atom is correlated with decreases in charge. There is a trend among the ortho substitutions that the systems with smaller volumes for the selenium centre generally have higher energy barriers, indicating that the steric effect does play a role in this reaction.

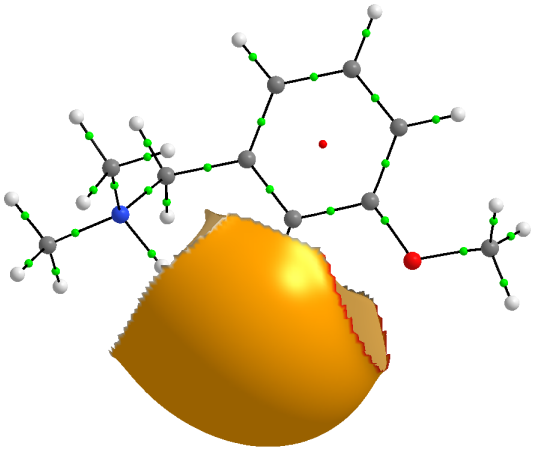
The atomic volumes listed in table 8.4 show that the ortho substituents decrease the size of the selenium atom in DMBS, but they cannot indicate where in the atomic basin that volume is lost. By visualizing the atomic basin of the selenium atom, shown in figure 8.7, it is apparent that the selenium's basin is "cut off" on the edge facing the ortho substituent. The representations in the figure are viewed from the same face of the molecule the peroxide interacts with, so one can see that there is a smaller surface presented to an incoming peroxide. In particular, the o$-CF_3$ substituted system, which has the highest predicted energy barrier in this study has a very small



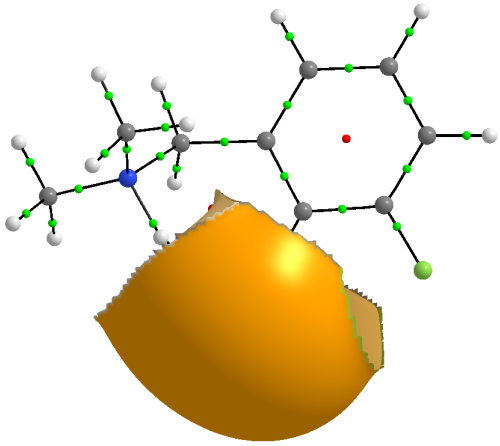
(a) unsubstituted zwitterionic form of DMBS



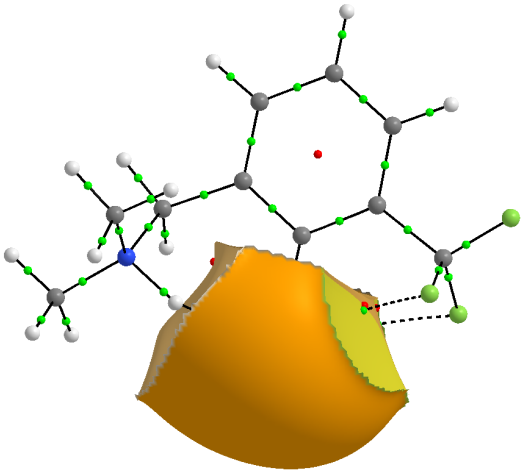
(b) o-NH_2 zwitterionic form of DMBS



(c) o-OCH_3 zwitterionic form of DMBS



(d) o-F zwitterionic form of DMBS



(e) o-CF_3 zwitterionic form of DMBS

Figure 8.7: Atomic basin of Se at the 0.001 au isodensity envelope.

selenium surface on this face of the molecule.

8.4 Conclusions

Experimental evidence suggests that introducing substituents in the ortho position of tertiary-benzylamine-2-selenol compounds increases their GPx activity, in some cases by as much as an order of magnitude. The introduction of substituents in the ortho, meta, or para positions had little effect on the stability of the zwitterion relative to the selenol form of these compounds. Figure 8.3 shows that none of the substituents included in this study significantly changed the energy barrier for the peroxide reduction reaction, although the changes do follow the expected trends. Atoms in molecules analysis and electrostatic potential maps show that some ortho substituents change the environment around selenium such that it becomes less favourable for the peroxide molecule to form a close association with the selenium centre in the molecule.

The increased reactivity observed experimentally for ortho substituted systems likely comes from an increased rate of reaction during regeneration of the oxidised organoselenium compound through hindrance of thiol exchange reactions in the selenylsulfide intermediate.[64] In this case, it is expected that ortho substituents provide needed steric bulk, which has only a minor effect on the peroxide reduction reaction, but blocks the thiol exchange reaction by making the sulfur more accessible.

Chapter 9

Conclusions and Future Work

9.1 Benchmarking DFT Methods

In chapter 4 the results of a benchmarking study of nine DFT methods and thirteen Pople basis sets were presented. The lengths of the selenium-containing bonds and a variety of bond dissociation energies were evaluated, using results from QCISD/cc-pVTZ calculations as the reference. As expected, the 6-31G and 6-311G basis sets predict very poor geometries and energies regardless of the DFT method used.

Across each model system studied, the relative ordering of the bond lengths predicted by each DFT method remained constant. The B3P86 and B3PW91 methods consistently predicted the shortest selenium-containing bonds, while the B3LYP, PBE, and B97D methods predicted the longest bonds, with B97D generally much longer than the other methods. In most systems, the geometries appear to have converged by the 6-311(2df,p) basis set with little change upon adding additional polarization functions. The geometries predicted by the double-zeta basis sets, such as 6-31G(d,p), are generally quite similar to the large triple-zeta basis sets, while the smaller triple-zeta basis sets with fewer polarization functions predict longer bonds.

Similar to the geometries, the relative ordering of the bond dissociation energies predicted by each DFT method remains consistent. The BDEs predicted by double-zeta basis sets were much larger than the triple-zeta basis sets, while the QCISD reference tended to be similar to the triple-zeta energies. Overall, the B3LYP and B97D methods consistently predicted very accurate BDEs, while the B3PW91 is slightly less accurate, but still remains within 10 kJ/mol of the reference.

The method recommended for use with organoselenium compounds is B3PW91/6-311+G(2df,p). It was also shown that, to save on computational cost, geometries can

be predicted with the more modest B3PW91/6-31+G(d,p) followed by a single point energy calculation with the 6-311+G(2df,p) basis set without a significant loss of accuracy.

9.2 Modelling the Catalytic Cycle of an Organoselenium Antioxidant

In chapters 5, 6, and 7 the results of a computational model of the catalytic cycle of *N,N*-dimethyl-benzylamine-2-selenol (DMBS), a particularly promising organoselenium antioxidant were presented. Chapters 5 and 6 give mechanisms previously unreported for this compound in these steps of the catalytic cycle, and while although chapter 7 does not include a likely mechanism for the reaction it models, it does provide new insight into previously reported mechanisms for similar compounds.

9.2.1 Peroxide reduction

Experimental kinetic studies predict that the initial peroxide reduction step of the catalytic mechanism is a rapid, one-step process. The results from this study confirm that finding, predicting a one-step reaction of hydrogen peroxide with the zwitterionic form of DMBS with a low energy barrier of 41.2 kJ/mol. Although it is unlikely to be formed in the real system, the selenolate form of DMBS has a barrier 7 kJ/mol lower, demonstrating the importance of the negative charge on the selenium centre in this reaction. The cationic form of DMBS, with the amine protonated, has a barrier 30 kJ/mol higher than the zwitterionic form, indicating that the incorporation of the amine proton is not as important as the charge on the selenium atom.

It was also shown that it is unlikely that the diselenide form of DMBS would have a significant contribution to the peroxide reduction reaction, as the energy barrier for its reaction with hydrogen peroxide is quite high. If the amines of this form are protonated, forming a dication, the energy barrier is decreased, but it is still much higher than predicted for the zwitterionic form of DMBS.

9.2.2 Selenenic acid substitution

For the reaction of a selenenic acid with a thiol, two potential mechanisms were examined, using either the thiol or a thiolate as the nucleophile. The more favourable mechanism was found to be the one using the thiol. The transfer of the thiol hydrogen to the selenenic acid hydroxyl group is facilitated through inclusion of two explicit solvent molecules, which act a proton shuttle. This mechanism has a low energy barrier of 43-49 kJ/mol, depending on the thiol used.

A mechanism was also isolated that uses a thiolate as the nucleophile. While it was computationally much simpler than the thiol mechanism, not requiring the inclusion of explicit solvent molecules, it has an energy barrier nearly four times larger than the thiol mechanism. Additionally, the reaction is predicted to be endothermic following this mechanism.

9.2.3 Selenylsulfide reduction

Chapter 7 examines some previously reported mechanisms for the reaction of a selenylsulfide and a thiol. The Bachrach group[171] found that the reaction of dimethyl selenylsulfide with methanethiolate follows a two-step addition-elimination mechanism. This mechanism could be reproduced with gas-phase calculations, but upon the addition of an implicit solvent model the transition state associated with the addition reaction could not be isolated. An alternative mechanism was isolated for a one-step solvent-assisted reaction of dimethyl selenylsulfide with methanethiol that uses three solvent molecules to facilitate the transfer of the thiol hydrogen to the leaving selenol.

The solvent-assisted mechanism could not be isolated using the selenylsulfide of DMBS. In all attempts to isolate a reacting complex, the thiol is 6 Å or more away from the other sulfur atom, much too far to react. An analysis of the atomic charges in both dimethyl selenylsulfide and the selenylsulfide of DMBS revealed that DMBS has much larger charges on the sulfur and selenium centres. Examining the electrostatic

potential in both molecules showed that DMBS has a large negative region surrounding the sulfur centre and most of the selenium centre. In the methyl-substituted model system, the negative region is significantly smaller and leaves much of the sulfur centre exposed.

9.2.4 Peroxide reduction using monosubstituted DMBS

In the final results chapter, the peroxide reduction reaction studied in chapter 5 is revisited to determine the effects of introducing substituents to the phenyl ring of DMBS. A range of electron withdrawing and electron donating substituents had only a minor effect on the barrier of the peroxide reduction reaction, changing it by less than 5 kJ/mol. A QTAIM and electrostatic potential analysis on the substituted compound showed that subtituents in the ortho position had large effects on both the electronic and steric environment of the selenium centre.

9.3 Future Directions

An understanding of the reaction mechanisms involved in the catalytic cycle for the reaction of hydrogen peroxide and methanethiol using DMBS provides a starting point for the complete understanding of this system. However, there are numerous other areas that can be examined to better understand the entire system. To close this document, some areas of the chemistry of DMBS that warrant exploration will be introduced and the motivations for their study will be explained.

9.3.1 Testing robustness of DMBS

As highlighted throughout this work and chapter 7 especially, a mechanism that is found for a model system will not always work for the full system. Many of the issues found with model systems related to the use of small models for DMBS, such as dimethyl selenylsulfide, however if a compound like DMBS is to be used in biological systems it will be catalysing the reactions of more complex peroxides and thiols. Therefore, it would be useful to examine the robustness of DMBS as a catalysing

antioxidant by modelling its reaction with other peroxides and thiols.

Throughout the literature on GPx mimics, most studies use hydrogen peroxide as the peroxide. It is important to understand their reaction with this peroxide since it is produced in the body by superoxide dimustase,[51] however there are other organic hydroperoxides that are also formed under oxidative stress. There are some experimental studies that use very sterically hindered hydroperoxides, such as tertiary-butyl hydroperoxide[63,70] or cumene hydroperoxide,[63] and it would be informative to also model the reduction of such compounds with a GPx mimic. Even if one did not want the computational cost of modelling a large organic hydroperoxide, studying a small organic molecule, such as methane hydroperoxide, could give valuable information. Hydrogen peroxide is a symmetric molecule, so it doesn't matter which end of the O−O bond interacts with the selenium centre, but in an symmetrical hydroperoxide depending on which end of the peroxide reacts with selenium either an alcohol and a selenenic acid or water and a selenenic ester will be produced.

In the other two steps of the catalytic cycle, the choice of thiol can also be important. In the body, GPx uses glutathione to regenerate its active site, however the majority of *in vitro* studies of GPx mimics use thiophenol, which does not resemble glutathione and may not even fit into the active site of GPx. A computational model of the GPx active site used ethanethiol as a model for glutathione to reduce computational cost.[55] Due to the existing steric bulk around the selenium centre of DMBS, the use of sterically hindered thiols would likely increase the barriers for these reactions, especially for the initial cleavage of the diselenide bond that was not explored in this work. Modelling these reactions with various thiols could provide insight into the limits of their size.

9.3.2 Competing side reactions

In most chemical reactions, there are a number of pathways that the mixture of compounds can follow to produce a variety of products. Often, one of these reactions will dominate, due to being energetically favoured, but this is not always the case. In

order to fully understand the reactions of a GPx mimic with peroxides and thiols, one must consider not only the reactions in the desired catalytic cycle, but also a number of potentially competing side reactions. Some of these have been suggested for each step of the GPx-like reaction of DMBS.

There have been no reactions proposed that would directly compete with the peroxide reduction reaction, but it has been suggested that once the selenenic acid is formed, it could go on to react with additional peroxides to form higher oxoacids. In the absence of thiols, Iwaoka and Tomoda observed the formation of the seleninic and selenonic acids of DMBS, but could not observe even the selenenic acid in the presence of thiols.[73] This suggests that the reaction of a thiol with the selenenic acid is rapid under ideal conditions and there will not be a high enough concentration of selenenic acid compared to selenol for overoxidation to be a concern. But, this may still become important at non-ideal conditions, such as under a deficiency of thiols. Working under the assumption that these acids are formed, it would be important to know if thiols can still react with them to return them to the catalytic cycle or if the formation of more oxidised selenium centres destroys their catalytic activity.

Many selenenic acids are unstable and will spontaneously form a selenenic acid anhydride in solution.[161,162] More stable selenenic acids have been reported employing electron withdrawing groups on phenyl selenenic acids.[163,164] By modelling the formation of the selenenic acid anhydride, it is possible to determine if the coordination of the amine in DMBS will similarly stabilise its selenenic acid. Kice, *et al.* showed that a thiol will react with a selenenic anhydride to produce the selenylsulfide,[162] so anydride formation would not destroy a GPx mimic's efficacy, but it could reduce its efficiency.

Many reports have indicated that reduction of the selenylsulfide of DMBS to the selenol is the slowest step of its GPx-like catalytic cycle.[64,73,167] This may be due, in part, to a competition between the productive reduction reaction and a non-productive thiol exchange reaction. The work of Bachrach, *et al.* predicts that in small selenylsulfides it is more favourable for a thiolate to attack the selenium centre,

leading to a thiol exchange reaction, than at the sulfur centre, leading to the reduction of selenium.[171] A year later, Sarma and Mugesh reported that thiol exchange reactions are responsible for the poor GPx-like activity of ebselen.[64] Since DMBS has a much higher activity than ebselen, this reaction is likely mediated in this species, but examination of the thiol exchange may give further insight into why it is competitive in some GPx mimics, but not others.

9.3.3 Improve GPx-like activity

With an understanding of the catalytic cycle of DMBS and relevant side reactions, one can use that knowledge to alter the structure of DMBS to improve its antioxidant performance.

One of the simplest ways to change the reactivity of organic molecules is to add functional groups to the system changing the electronic or steric properties of the reacting centre. As was demonstrated in chapter 8, adding electron withdrawing or donating groups to any position on the phenyl ring of DMBS had a very minor effect on the barrier for peroxide reduction. As this step of the reaction proceeds quickly, it is not necessary to attempt to speed it up, but it is similarly important to know that these functional groups do not significantly inhibit this reaction either. The same series of substituents should be studied in the other steps of the catalytic cycle to determine if they have any effect, either adverse or favourable, on the overall rate of the reaction.

It has been previously reported that introduction of functional groups ortho to the selenium in ebselen or DMBS increases the overall GPx-like activity by about an order of magnitude.[85,174] It is likely that these groups have an effect on either reaction of the selenenic acid with thiol or the selenylsulfide reduction. In these systems, the N−Se−O or N−Se−S angle is about 180°, therefore the introduction of a functional group ortho to selenium will necessitate a change in the molecular geometry, which will have an effect on the energy barriers of these reactions.

Rather than indirectly modifying the environment of the selenium moiety through

introduction of functional groups, the selenium itself could be changed. It is well known that replacing the selenium in GPx or a GPx mimic with sulfur has a detrimental effect on its activity.[49,65,175–177] As early as 1992, it was reported that tellurium analogues of known organoselenium GPx mimics had greater activity than the selenium compounds.[59] Like ebselen, its tellurium analogue, ebtellur, has received much attention,[178–180] although many other organotellurium compounds have also been examined.[81,177,181–187] In light of this, studying the tellurium analogue of DMBS could prove interesting. It is unlikely that an organotellurium compound would be used in biological systems, but there are many chemical reactions that produce peroxides as a side product, which require the use of potent antioxidants to protect the desired products from oxidation.

Bibliography

[1] Gay-Lussac, J. L. *Mémoires de Physique et de Chimie de la Société d'Arcueil* **1809**, *2*, 207–234.

[2] Berzelius, J. J. *Annales de chimie et de physique* **1818**, *7*, 199–206.

[3] Shriver, D. F.; Atkins, P. W. *Inorganic Chemistry*, 3rd ed.; W. H. Freeman and Company: New York, 1999.

[4] Reilly, C. *Selenium in Food and Health*, 2nd ed.; Springer: New York, 2006.

[5] Boyd, R. *Nature Chemistry* **2011**, *3*, 570.

[6] Trelease, S. F.; Trelease, H. M. *Am. J. Bot.* **1939**, *26*, 530–535.

[7] Schwarz, K.; Foltz, C. M. *J. Am. Chem. Soc.* **1957**, *79*, 3292–3293.

[8] Patterson, E. L.; Milstrey, R.; Stokstad, E. L. R. *Exp. Biol. Med.* **1957**, *95*, 617–620.

[9] Christensen, F.; Dam, H.; Prange, I.; Søndergaard, E. *Acta Pharmacol. Toxicol.* **1958**, *15*, 181–188.

[10] Thompson, J. N.; Scott, M. L. *J. Nutr.* **1969**, *97*, 335–342.

[11] McCoy, K. E. M.; Weswig, P. H. *J. Nutr.* **1969**, *98*, 383–389.

[12] Turner, D. C.; Stadtman, T. C. *Arch. Biochem. Biophys.* **1973**, *154*, 366–381.

[13] Cone, J. E.; del R$\tilde{\text{i}}$o, R. M.; Davis, J. N.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 2659–2663.

[14] Lindström, K. *Hydrobiologia* **1983**, *101*, 35–47.

[15] Arvy, M. P. *J. Exp. Bot.* **1993**, *44*, 1083–1087.

[16] Sandholm, M.; Oksanen, H. E.; Pesonen, L. *Limnol. Oceanogr.* **1973**, *18*, 496–499.

[17] Zayed, A.; Lytle, C. M.; Terry, N. *Planta* **1998**, *206*, 284–292.

[18] Ng, B. H.; Anderson, J. W. *Phytochemistry* **1978**, *17*, 2069–2074.

[19] Eustice, D. C.; Kull, F. J.; Shrift, A. *Plant Physiol.* **1981**, *67*, 1054–1058.

[20] Stewart, R. D. H.; Griffiths, N. M.; Thomson, C. D.; Robinson, M. F. *Br. J. Nutr.* **1978**, *40*, 45–54.

[21] Jenkins, K. J.; Hidiroglou, M. *Can. J. Physiol. Pharmacol.* **1972**, *50*, 927–935.

[22] Rotruck, J. T.; Pope, A. L.; Ganther, H. E.; Swanson, A. B.; Hafeman, D. G.; Hoekstra, W. G. *Science* **1973**, *179*, 588–590.

[23] Mills, G. C. *J. Biol. Chem.* **1957**, *229*, 189–197.

[24] Flohe, L.; Günzler, W. A.; Schock, H. H. *FEBS Lett.* **1973**, *32*, 132–134.

[25] Forstrom, J. W.; Zakowski, J. J.; Tappel, A. L. *Biochemistry* **1978**, *17*, 2639–2644.

[26] Chambers, I.; Frampton, J.; Goldfarb, P.; Affara, N.; McBain, W.; Harrison, P. R. *EMBO J.* **1986**, *5*, 1221–1227.

[27] Zinoni, F.; Birkmann, A.; Stadtman, T. C.; Böck, A. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 4650–4654.

[28] Böck, A.; Forchhammer, K.; Heider, J.; Leinfelder, W.; Sawers, G.; Veprek, B.; Zinoni, F. *Mol. Microbiol.* **1991**, *5*, 515–520.

[29] Stadtman, T. C. *Annu. Rev. Biochem.* **1996**, *65*, 83–100.

[30] Arthur, J. R.; Beckett, G. J. *Proc. Nutr. Soc.* **1994**, *53*, 615–624.

[31] Hatfield, D. L.; Gladyshev, V. N. *Mol. Cell. Biol.* **2002**, *22*, 3565–3576.

[32] Kryukov, G. V.; Castellano, S.; Novoselov, S. V.; Lobanov, A. V.; Zehtab, O.; Guigó, R.; Gladyshev, V. N. *Science* **2003**, *300*, 1439–1443.

[33] Gladyshev, V. N.; Kryukov, G. V.; Fomenko, D. E.; Hatfield, D. L. *Annu. Rev. Nutr.* **2004**, *24*, 579–596.

[34] Rayman, M. P. *The Lancet* **2000**, *356*, 233–241.

[35] McKenzie, R. C.; Arthur, J. R.; Beckett, G. J. *Antioxid. Redox Signaling* **2002**, *4*, 339–351.

[36] Wessjohann, L. A.; Schneider, A.; Abbas, M.; Brandt, W. *Biol. Chem.* **2007**, *388*, 997–1006.

[37] Huber, R. E.; Criddle, R. S. *Arch. Biochem. Biophys.* **1967**, *122*, 164–173.

[38] Wada, M.; Nobuki, S.; Tenkyuu, Y.; Natsume, S.; Asahara, M.; Erabi, T. *J. Organomet. Chem.* **1999**, *580*, 282–289.

[39] Cardey, B.; Enescu, M. *ChemPhysChem* **2005**, *6*, 1175–1180.

[40] Cardey, B.; Enescu, M. *J. Phys. Chem. A* **2007**, *111*, 673–378.

[41] Arnér, E. S. J. *Exp. Cell Res.* **2010**, *316*, 1296–1303.

[42] Urisini, F.; Maiorina, M.; Valente, M.; Ferri, L.; Gregolin, C. *Biochim. Biophys. Acta* **1982**, *710*, 197–211.

[43] Schuckelt, R.; Brigelius-Flohé, R.; Maiorino, M.; Roveri, A.; Reumkens, J.; Strassburger, W.; Urisini, F.; Wolf, B.; Flohe, L. *Free Rad. Res. Commun.* **1991**, *14*, 343–361.

[44] Brigelius-Flohe, R.; Aumann, K. D.; Blocker, H.; Gross, G.; Kiess, M.; Kloppel, K. D.; Maiorino, M.; Roveri, A.; Schuckelt, R.; Usani, F. *J. Biol. Chem.* **1994**, *269*, 7342–7348.

[45] Takahashi, K.; Avissar, N.; Whitin, J.; Cohen, H. *Arch. Biochem. Biophys.* **1987**, *256*, 677–686.

[46] Chu, F. F.; Doroshow, J. H.; Esworthy, R. S. *J. Biol. Chem.* **1993**, *268*, 2571–2576.

[47] Brigelius-Flohé, R. *Free Radical Biol. Med.* **1999**, *27*, 951–965.

[48] Epp, O.; Ladenstein, R.; Wendel, A. *Eur. J. Biochem.* **1983**, *133*, 51–69.

[49] Rocher, C.; Lalanne, J.; Chaudiére, J. *Eur. J. Biochem.* **1992**, *205*, 955–960.

[50] Nordberg, J.; Arnér, E. S. J. *Free Radical Biol. Med.* **2001**, *31*, 1281–1312.

[51] McCord, J. M.; Fridovich, I. *J. Biol. Chem.* **1969**, *244*, 6056–6063.

[52] Aumann, K. D.; Bedorf, N.; Brigelius-Flohe, R.; Schomburg, D.; Flohe, L. *Biomed. Environ. Sci.* **1997**, *10*, 136–155.

[53] Prabhakar, R.; Musaev, D. G.; Khavrutskii, I. V.; Morokuma, K. *J. Phys. Chem. B* **2004**, *108*, 12643–12645.

[54] Prabhakar, R.; Vreven, T.; Frisch, M. J.; Morokuma, K.; Musaev, D. G. *J. Phys. Chem. B* **2006**, *110*, 13608–13613.

[55] Prabhakar, R.; Vreven, T.; Morokuma, K.; Musaev, D. G. *Biochemistry* **2005**, *35*, 11864–11871.

[56] Rotruck, J. T.; Pope, A. L.; Ganther, H. E.; Hoekstra, W. G. *J. Nutr.* **1972**, *102*, 689–696.

[57] Müller, A.; Cadenas, A.; Graf, P.; Sies, H. *Biochem. Pharmacol.* **1984**, *33*, 3235–3239.

[58] Wendel, A.; Fausel, M.; Safayhi, H.; Tiegs, G. *Biochem. Pharmacol.* **1984**, *33*, 3241–3245.

[59] Engman, L.; Stern, D.; Cotgreave, I. A.; Andersson, C. M. *J. Am. Chem. Soc.* **1992**, *114*, 9737–9743.

[60] Schewe, T. *Gen. Pharmacol. Vasc. Syst.* **1995**, *26*, 1153–1169.

[61] Sies, H.; Masumoto, H. *Adv. Pharmacol.* **1996**, *38*, 229–246.

[62] Mareque, A. M.-M.; Faez, J. M.; Chistiaens, L.; Kohnen, S.; Deby, C.; Hoebeke, M.; Lamy, M.; Deby-Dupont, G. *Redox Report* **2004**, *9*, 81–87.

[63] Bhabak, K. P.; Mugesh, G. *Chem. Eur. J.* **2007**, *13*, 4594–4601.

[64] Sarma, B. K.; Mugesh, G. *J. Am. Chem. Soc.* **2005**, *127*, 11477–11485.

[65] Pearson, J. K.; Boyd, R. J. *J. Phys. Chem. A* **2007**, *111*, 3152–3160.

[66] Pearson, J. K.; Boyd, R. J. *J. Phys. Chem. A* **2008**, *112*, 1013–1017.

[67] Kaura, D.; Sharmaa, P.; Bharatam, P. V. *J. Mol. Struct.* **2010**, *939*, 118–123.

[68] Zhao, R.; Masayasu, H.; Holmgren, A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 8579–8584.

[69] Jacquemin, P. V.; Christiaens, L. E.; Renson, M. J. *Tetrahedron Lett.* **1992**, *33*, 3863–3866.

[70] Back, T. G.; Dyck, B. P. *J. Am. Chem. Soc.* **1997**, *119*, 2079–2083.

[71] Erdelmeier, I.; Tailhan-Lomont, C.; Yadan, J.-C. *J. Org. Chem.* **2000**, *65*, 8152–8157.

[72] Wilson, S. R.; Zucker, P. A.; Huang, R.-R. C.; Spector, A. *J. Am. Chem. Soc.* **1989**, *111*, 5936–5939.

[73] Iwaoka, M.; Tomoda, S. *J. Am. Chem. Soc.* **1994**, *116*, 2557–2561.

[74] Galet, V.; Bernier, J.-L.; Henichart, J.-P.; Lesieur, D.; Abadie, C.; Rochette, L.; Lindenbaum, A.; Chalas, J.; de la Faverie, J.-F. R. *J. Med. Chem.* **1994**, *37*, 2903–2911.

[75] Wirth, T. *Molecules* **1998**, *3*, 164–166.

[76] Mugesh, G.; Panda, A.; Singh, H. B.; Punekar, N. S.; Butcher, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 839–850.

[77] Reich, H. J.; Jasperse, C. P. *J. Am. Chem. Soc.* **1987**, *109*, 5549–5551.

[78] Back, T. G.; Moussa, Z. *J. Am. Chem. Soc.* **2003**, *125*, 13455–13460.

[79] Mugesh, G.; Panda, A.; Singh, H. B.; Punekar, N. S.; Butcher, R. J. *Chem. Commun.* **1998**, 2227–2228.

[80] Back, T. G.; Moussa, Z.; Parvez, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 1268–1270.

[81] Sarma, B. K.; Manna, D.; Minoura, M.; Mugesh, G. *J. Am. Chem. Soc.* **2010**, *132*, 5364–5374.

[82] Mugesh, G.; du Mont, W.-W. *Chem. Eur. J.* **2001**, *7*, 1365–1370.

[83] Bhabak, K. P.; Mugesh, G. *Chem. Eur. J.* **2009**, *15*, 9846–9854.

[84] Kumar, S.; Singh, H. *J. Chem. Sci.* **2005**, *117*, 621–628.

[85] Bhabak, K. P.; Mugesh, G. *Chem. Eur. J.* **2008**, *14*, 8640–8651.

[86] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; McGraw-Hill: New York, 1989.

[87] Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, 1999.

[88] Levine, I. N. *Quantum Chemistry*, 6th ed.; Pearson Prentice Hall: New Jersey, 2009.

[89] Schrödinger, E. *Ann. Physik* **1926**, *79*, 361.

[90] Born, M.; Oppenheimer, R. *Ann. Physik* **1927**, *84*, 457.

[91] Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293–1322.

[92] Slater, J. C. *Phys. Rev.* **1930**, *35*, 509–529.

[93] Slater, J. C. *Phys. Rev.* **1930**, *36*, 57–64.

[94] Boys, S. F. *Proc. R. Soc. Lond. A* **1950**, *200*, 542–554.

[95] Hehre, W. J.; Radom, L.; Pople, J.; Schleyer, P. v. R. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.

[96] Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823–2834.

[97] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[98] Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.

[99] Slater, J. C. *Rev. Mod. Phys.* **1963**, *35*, 484–487.

[100] Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1927**, *24*, 89–110.

[101] Gaunt, J. A. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 328–342.

[102] Slater, J. C. *Phys. Rev.* **1928**, *32*, 339–348.

[103] Slater, J. C. *Phys. Rev.* **1930**, *35*, 210–211.

[104] Fock, V. *Z. Physik* **1930**, *61*, 126.

[105] Hartree, D. R.; Hartree, W. *Proc. R. Soc. Lond. A* **1935**, *150*, 9–33.

[106] Eckart, C. *Phys. Rev.* **1930**, *36*, 878–892.

[107] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

[108] Hall, G. G. *Proc. R. Soc. Lond. A* **1951**, *205*, 541–552.

[109] Brillouin, L. *Actualities Sci. Ind.* **1933**, *71*, 159.

[110] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

[111] Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed.; Wiley-VCH: Weinheim, 2001.

[112] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

[113] Thomas, L. H. *Proc. Cambridge Phil. Soc.* **1927**, *23*, 542–548.

[114] Fermi, E. *Rend. Accad. Naz. Lincei* **1927**, *6*, 602–607.

[115] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[116] Harris, J.; Jones, R. O. *J. Phys. F: Met. Phys.* **1974**, *4*, 1170–1186.

[117] Gunnarsson, O.; Lundqvist, B. I. *Phys. Rev. B* **1976**, *13*, 4274–4298.

[118] Langreth, D. C.; Perdew, J. P. *Phys. Rev. B* **1977**, *15*, 2884–2901.

[119] Harris, J. *Phys. Rev. A* **1984**, *29*, 1648–1659.

[120] Perdew, J. P.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800–8802.

[121] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[122] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[123] Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486–1493.

[124] Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117–129.

[125] Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098–16104.

[126] Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

[127] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.

[128] Bader, R. F. W. *Atoms in Molecules - A Quantum Theory*; Oxford University Press: Oxford, 1990.

[129] Matta, C. F., Boyd, R. J., Eds. *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*; WILEY-VCH: Weinheim, 2007.

[130] Bushmarinov, I. S.; Lyssenko, K. A.; Antipin, M. Y. *Russ. Chem. Rev.* **2009**, *78*, 283–302.

[131] Mugesh, G.; du Mont, W.-W.; Sies, H. *Chem. Rev.* **2001**, *101*, 2125–2179.

[132] Brandt, G.; Carrasco, N.; Huang, Z. *Biochemistry* **2006**, *45*, 8972–8977.

[133] Moroder, H.; Kreutz, C.; Lang, K.; Serganov, A.; Micura, R. *J. Am. Chem. Soc.* **2006**, *128*, 9909–9918.

[134] Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z. *J. Am. Chem. Soc.* **2007**, *129*, 4862?4863.

[135] Fang, J.; Zhong, L.; Zhao, R.; Holmgren, A. *Toxicol. Appl. Pharmacol.* **2005**, *207*, S103–S109.

[136] Pearson, J. K.; Boyd, R. J. *J. Phys. Chem. A* **2006**, *110*, 8979–8985.

[137] Trujillo, C.; Mó, O.; Yáñez, M. *Org. Biomol. Chem.* **2007**, *5*, 3092–3099.

[138] Heverly-Coulson, G. S.; Boyd, R. J. *J. Phys. Chem. A* **2010**, *114*, 1996–2000.

[139] Heverly-Coulson, G. S.; Boyd, R. J. *J. Phys. Chem. A* **2010**, *114*, 10706–10711.

[140] Eizaguirre, A.; Mó, O.; Yáñez, M.; Boyd, R. J. *Org. Biomol. Chem.* **2011**, *9*, 423–431.

[141] Bachrach, S. M.; Jiang, S. *J. Org. Chem.* **1999**, *64*, 8248–8255.

[142] Pearson, J. K.; Ban, F.; Boyd, R. J. *J. Phys. Chem. A* **2005**, *109*, 10373–10379.

[143] Frisch, M. J. et al. Gaussian 09 Revision A.02. Gaussian, Inc., Wallingford CT, 2009.

[144] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[145] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

[146] Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264–6271.

[147] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[148] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

[149] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.

[150] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–13249.

[151] Cohen, A. J.; Handy, N. C. *Mol. Phys.* **2001**, *99*, 607–615.

[152] Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559–9569.

[153] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[154] Frisch, M. J. et al. Gaussian 03 Revision E.01. Gaussian Inc. Wallingford CT 2004.

[155] Peng, C.; Schlegel, H. B. *Isr. J. Chem.* **1993**, *33*, 449–454.

[156] Peng, C.; Ayala, P.; Schlegel, H. B.; Frisch, M. *J. Comput. Chem.* **1996**, *17*, 49–56.

[157] Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154–2161.

[158] Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523–5527.

[159] Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502–16513.

[160] Haynes, W. M., Ed. *CRC Handbook of Chemistry and Physics*, 92nd ed.; CRC Press: Cleveland, 2011.

[161] Reich, H. J.; Willis, W. W.; Wollowitz, S. *Tetrahedron Lett.* **1982**, *23*, 3319–3322.

[162] Kice, J. L.; McAfee, F.; Slebocka-Tilk, H. *Tetrahedron Lett.* **1982**, *23*, 3323–3326.

[163] Reich, H. J.; Hoeger, C. A.; Willis, W. W. *J. Am. Chem. Soc.* **1982**, *104*, 2936–2937.

[164] Kang, S. I.; Kice, J. L. *J. Org. Chem.* **1986**, *51*, 295–301.

[165] Reich, H. J.; Jasperse, C. P. *J. Org. Chem.* **1988**, *53*, 2389–2390.

[166] Saiki, T.; Goto, K.; Okazaki, R. *Angew. Chem., Int. Ed.* **1997**, *36*, 2223–2224.

[167] Benkova, Z.; Kóňa, J.; Gann, G.; Fabian, W. M. F. *Int. J. Quantum Chem.* **2002**, *90*, 555–565.

[168] AIMAll. Todd A. Keith, 2011 (aim.tkgristmill.com).

[169] Danehy, J. P.; Noel, C. J. *J. Am. Chem. Soc.* **1960**, *82*, 2511–2515.

[170] Steinmann, D.; Nauser, T.; Koppenol, W. H. *J. Org. Chem.* **2010**, *75*, 6696–6699.

[171] Bachrach, S. M.; Demoin, D. W.; Luk, M.; Miller, J. V. J. *J. Phys. Chem. A* **2004**, *108*, 4040–4046.

[172] Ren, B.; Huang, W.; Åkesson, B.; Ladenstein, R. *J. Mol. Biol.* **1997**, *268*, 869–885.

[173] Reich, H. J.; Cohen, M. L. *J. Org. Chem.* **1979**, *44*, 3148?3151.

[174] Parnham, M. J.; Biedermann, J.; Bittner, C.; Dereu, N.; Leyck, S.; Wetzig, H. *Agents Actions* **1989**, *27*, 306–308.

[175] Scurlock, R.; Rougee, M.; Bensasson, R.; Evers, M.; Dereu, N. *Photochem. Photobiol.* **1991**, *54*, 733–736.

[176] Moutet, M.; d'Alessio, P.; Malette, P.; Devaux, V.; Chaudière, J. *Free Radical Biol. Med.* **1998**, *25*, 270–281.

[177] Kumar, S.; Engman, L.; Valgimigli, L.; Amorati, R.; Fumo, M. G.; Pedulli, G. F. *J. Org. Chem.* **2007**, *72*, 6046–6055.

[178] Musaev, D. G.; Hirao, K. *J. Phys. Chem. A* **2003**, *107*, 9984–9990.

[179] Sakimoto, Y.; Hirao, K.; Musaev, D. G. *J. Phys. Chem. A* **2003**, *107*, 5631–5639.

[180] Sakimoto, Y.; Hirao, K.; Musaev, D. G. *Int. J. Quantum Chem.* **2009**, *109*, 2297 – 2307.

[181] Andersson, C.-M.; Hallberg, A.; Brattsand, R.; Cotgreave, I. A.; Engman, L.; Persson, J. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2553–2558.

[182] Engman, L.; Stern, D.; Pelcman, M.; Andersson, C. M. *J. Org. Chem.* **1994**, *59*, 1973–1979.

[183] Engman, L.; Persson, J.; Vessman, K.; Ekström, M.; Berglund, M.; Andersson, C.-M. *Free Radical Biol. Med.* **1995**, *19*, 441–452.

[184] Vessman, K.; Ekstroem, M.; Berglund, M.; Andersson, C.-M.; Engman, L. *J. Org. Chem.* **1995**, *60*, 4461–4467.

[185] Kanda, T.; Engman, L.; Cotgreave, I. A.; Powis, G. *J. Org. Chem.* **1999**, *64*, 8161–8169.

[186] Mugesh, G.; Panda, A.; Kumar, S.; Apte, S. D.; Singh, H. B.; Butcher, R. J. *Organometallics* **2002**, *21*, 884–892.

[187] Collins, S. P.; Heverly-Coulson, G. S.; Boyd, R. J. *Comp. Theor. Chem.* **2012**, *981*, 68–72.

Appendix A

Copyright Permission Letters

The following pages include the copyright permission letters from the publishers of the papers included in this thesis. The letters correspond to publications used in the following chapters:

- **Chapter 4**: G. S. Heverly-Coulson and R. J. Boyd, "Systematic Study of the Performance of Density Functional Theory Methods for Prediction of Energies and Geometries of Organoselenium Compounds" *J. Phys. Chem. A* **2011**, *115*, 4827-4831.

- **Chapter 5**: G. S. Heverly-Coulson and R. J. Boyd, "Reduction of Hydrogen Peroxide by Glutathione Peroxidase Mimics: Reaction Mechanism and Energetics" *J. Phys. Chem. A* **2010**, *114*, 1996-2000.

- **Chapter 8**: G. S. Heverly-Coulson and R. J. Boyd, "Theoretical Investigations on the Reaction of Monosubstituted Tertiary-Benzylamine Selenols with Hydrogen Peroxide" *J. Phys. Chem. A* **2010**, *114*, 10706-10711.