Density Functional Theory for the Investigation of Transition Metal Complexes: Structure, Binding, and Spectroscopy of Metal-Siderophores and the Catalase-Peroxidase Enzyme

Bonnie I. Kruft

Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Physical Chemistry Commons

Recommended Citation


https://academicworks.cuny.edu/gc_etds/1013

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
DENSITY FUNCTIONAL THEORY FOR THE INVESTIGATION OF
TRANSITION METAL COMPLEXES: STRUCTURE, BINDING AND
SPECTROSCOPY OF METAL-SIDEROPHORES AND THE
CATALASE-PEROXIDASE ENZYME

by

BONNIE I. KRUFT

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the
requirements for the degree of Doctor of Philosophy, The City University of New York.

2015
This manuscript has been read and accepted by the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Andrzej A. Jarzecki

Date Chair of Examining Committee

Brian R. Gibney

Date Executive Officer

Brian R. Gibney
Richard S. Magliozzo
Louis J. Massa
Supervisory Committee
Abstract

Density Functional Theory for the Investigation of Transition Metal Complexes: Structure, Binding and Spectroscopy of Metal-Siderophores and the Catalase-Peroxidase Enzyme

by

Bonnie I. Kruft

Adviser: Andrzej A. Jarzecki

Since the development of quantum mechanics in the 1920’s and with the introduction of the Schrödinger equation in 1926, various methods to solve the Schrödinger equation have evolved. With advances in these computational methods, we are now able to solve the Schrödinger equation for systems that did not seem possible less than a century ago. Density functional theory (DFT) is a valuable tool for the exploration of the molecular properties of biological systems, and is based upon the theory that the exact energy could be determined from the knowledge of the electron density. The purpose of this dissertation is to explore the use of modern density functional theory methods to compute the structural, spectroscopic and mechanistic properties of biological molecules. In particular, DFT will be used to explore insights into metal-ligand interactions of siderophore-transition metal complexes and to explore the properties of a unique tyrosyl-like radical of the Catalase Peroxidase (KatG) enzyme.
Acknowledgments

I would like to acknowledge the many people that have influenced, supported, and helped me throughout the preparation of this dissertation. First and foremost, I would like to thank my adviser, Andrzej Jarzecki, for sharing his knowledge and love of quantum chemistry. He taught me to be patient, to be careful, and to remain positive. Next, I would like to thank Richard Magliozzo for his support, advice and guidance over the past five years. I would also like to extend my gratitude to the members of my committee, Brian Gibney and Louis Massa, for their advice and support at our annual meetings.

I would also like to thank my family. My father for always pushing me to be the best that I can be. My mother, my very first chemistry teacher, for being my best friend and biggest fan. And my brother, Jon, who has always been and probably always will be the smartest guy I know. Finally, last but certainly not least, I would like to thank my husband and best friend, William Austin Peeler, who has stood by me and always believed in me. I could not have done this without your love and support.
# Contents

1 Introduction 1

2 Electronic Structure Methods 3

2.1 Atoms and Molecules: N-electron systems 3

2.1.1 Basis Sets 8

2.2 Hartree Fock Theory 11

2.2.1 Coulomb and Exchange Integrals 12

2.3 Möller-Plesset Many-Body Perturbation Theory 13

2.4 Multireference Approaches 15

2.4.1 Configuration Interactions 15

2.4.2 Coupled Cluster Methods 16

2.4.3 Complete Active Space (CASSCF) 18

2.5 Density Functional Theory 20

2.5.1 Approximations to the Exchange-Correlation Functional 24

2.5.2 Integration Grid 26

2.6 Time-Dependent Density Functional Theory 27

2.7 Two Approaches to Biological Molecules: QM/MM versus Standard DFT 29

3 Spectroscopy 31

3.1 Vibrational Spectroscopy 32

3.1.1 IR and Raman Intensities 33
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.2 Calculation of Vibrational Frequencies</td>
<td>35</td>
</tr>
<tr>
<td>3.2 EPR Spectroscopy</td>
<td>37</td>
</tr>
<tr>
<td>3.3 UV-Vis Spectroscopy</td>
<td>38</td>
</tr>
<tr>
<td>3.4 X-ray Absorption Spectroscopy</td>
<td>41</td>
</tr>
<tr>
<td>4 Siderophore-Metal Complexes</td>
<td>44</td>
</tr>
<tr>
<td>4.1 Desferrioxamine B Metal Complexes</td>
<td>45</td>
</tr>
<tr>
<td>4.2 Structure</td>
<td>47</td>
</tr>
<tr>
<td>4.3 Binding Energy</td>
<td>56</td>
</tr>
<tr>
<td>4.4 Spectroscopy</td>
<td>64</td>
</tr>
<tr>
<td>4.5 Conclusions</td>
<td>77</td>
</tr>
<tr>
<td>5 Catalase Peroxidase (KatG)</td>
<td>79</td>
</tr>
<tr>
<td>5.1 The Catalase Reaction - Dioxygen Release from the Oxyheme Intermediate</td>
<td>83</td>
</tr>
<tr>
<td>5.1.1 R(MYW)</td>
<td>83</td>
</tr>
<tr>
<td>5.1.2 R(MYW) Oxyheme</td>
<td>91</td>
</tr>
<tr>
<td>5.1.3 Catalase Reaction - Mechanism of Dioxygen Release</td>
<td>126</td>
</tr>
<tr>
<td>5.2 The g-tensor as a function of the MYW opening coordinate</td>
<td>129</td>
</tr>
<tr>
<td>6 A Unique Tyrosyl Radical: (MYW)$^{•+}$</td>
<td>134</td>
</tr>
<tr>
<td>6.1 Modeling of the g-tensor in MYW</td>
<td>135</td>
</tr>
<tr>
<td>6.1.1 The $g_{xx}$ Value as a Function of PCM Environment</td>
<td>139</td>
</tr>
<tr>
<td>6.1.2 Summary</td>
<td>154</td>
</tr>
<tr>
<td>7 Concluding Remarks</td>
<td>155</td>
</tr>
<tr>
<td>Appendices</td>
<td>158</td>
</tr>
<tr>
<td>A Calculation of the J and K Integrals</td>
<td>159</td>
</tr>
</tbody>
</table>
B Theoretical Basis for DFT

B.0.3 The Hohenberg-Kohn Theorem ................................. 166

B.0.4 The Hohenberg-Kohn Variational Theorem ................. 169

C Calculation of EPR properties

Bibliography
**List of Tables**

4.1 Calculated and experimental average interatomic distances, standard deviations and Debye-Waller factors .................................................. 48
4.2 Calculated and experimental average interatomic distances, standard deviations and Debye-Waller factors .................................................. 49
4.3 Calculated and experimental average interatomic distances, standard deviations and Debye-Waller factors .................................................. 50
4.4 Calculated and experimental metal stability constants \( \log K \) ............ 59
4.5 Calculated and experimental frequencies, infrared intensities, and vibrational mode assignments for amide I and amide II of the third coordination shell. 69
4.6 Calculated and experimental frequencies, infrared intensities, and vibrational mode assignments for \( \text{NH}_3 \) bending of the third coordination shell. ........ 70
4.7 Calculated and experimental frequencies, infrared intensities, and vibrational mode assignments for the second coordination shell. ................. 71
5.1 Calculated total spin per residue in the R(MYW)-radical models. ........... 84
5.2 Calculated total charge per residue in the R(MYW) models. ................. 85
5.3 Comparison of the X-ray crystal structure with selected computed geometry parameters, computed relative energies, and \( S^2 \) values of the five spin-states for the closed R(MYW)···oxyheme models. 96
5.4 Comparison of the X-ray crystal structure with selected computed geometry parameters, computed relative energies, and $S^2$ values of the five spin-states for the open R(MYW)$\cdots$oxyheme models. ........................................... 102
5.5 Calculated total spin per residue in the R(MYW)$\cdots$oxyheme models. .... 106
5.6 Calculated total charge per residue in the R(MYW)$\cdots$oxyheme models. ... 107
5.7 Comparison of torsional angles in X-ray crystal structures of KatGs ........ 128
5.8 Relative energies, $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for the scan of the MYW opening coordinate of Arg-remote and Arg-vicinal MYW models. The numbers in bold are the energetic minima (0 kcal/mol) for each structure. .................. 131
5.9 Relative energies, $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for the scan of the YW opening coordinate of Arg-remote and Arg-vicinal YW models. The numbers in bold are the energetic minima (0 kcal/mol) for each structure. .................. 132
6.1 Experimentally and computationally determined g-values for tyrosyl radicals. 136
6.2 Comparison of g-values from multiple methodologies. .......................... 140
6.3 Percentage of spin distributed on the Tyr-ring vs Tyr-O for combinations of M, Y, and W. ................................................................. 144
6.4 Calculated total spin per residue for combinations of M, Y, and W. ......... 144
6.5 Calculated total charge per residue for combinations of M, Y, and W. ...... 145
6.6 Percentage of spin distributed on the Tyr-ring vs Tyr-O for variations of Met and Trp (indole) positions on the MYW adduct. ......................... 148
6.7 Calculated total spin per residue for variations of Met and Trp positions on the MYW adduct. ................................................................. 149
6.8 Calculated total charge per residue for variations of Met and Trp positions on the MYW adduct. ................................................................. 149
6.9 Percentage of spin distributed on the Tyr-ring vs Tyr-O for variations of the MYW adduct that replaced Met with either Cys or His. ............... 152
6.10 Calculated total spin per residue for variations of the MYW adduct that replaced Met with either Cys or His. ................................. 153

6.11 Calculated total charge per residue for variations of the MYW adduct that replaced Met with either Cys or His. ................................. 154
List of Figures

2.1 Comparison of Gaussian-type functions (blue) to Slater-type functions (red). 9

3.1 Absorption and scattering for Infrared and Raman spectroscopy. ...... 35

3.2 The difference between non-resonance Raman and resonance Raman scattering. 36

4.1 Calculated and experimental (EXAFS) average distances of the first, second and third coordination shells for metal-HDFOB complexes versus metal ionic radii. ............................................. 51

4.2 Calculated structures of vanadium-(II), -(III), -(IV) and -(V)-HDFOB. .... 52

4.3 Calculated first-coordination-shell distances vs. experimental distances. .... 55

4.4 Calculated structure of the Pb(II)-HDFOB complex. ............................ 56

4.5 A linear fit of DFT computed (log $K_{Calc}$) and experimental (log $K_{Exp}$) stability constants based on available data for metal(II)-HDFOB complexes. .... 60

4.6 Comparison of experimental (log $K_{Exp}$) and refined (log $K_A$) stability constants. 62

4.7 Computed and experimental binding constants for the first-row transition M(II) and M(III) metal ions plotted against occupation of d-orbitals. ....... 63

4.8 Calculated and experimentally determined metal-HDFOB complex stability constants versus charged normalized ionic radii ............................... 65

4.9 Calculated and experimental IR spectra. (1300 – 1650 cm$^{-1}$ frequency region). 67

4.10 Calculated and experimental IR spectra. (900–1650 cm$^{-1}$ frequency region). 68
4.11 Correlation plots of computed values for selected bond distances vs. the CO bond length ........................................... 74
4.12 Correlation plots of computed values for selected bond distances vs. the amide I-like frequency. ........................................... 76
5.1 Comparison of X-ray crystal structures of KatGs showing the conserved MYW adduct nearby the heme group. ................................. 81
5.2 Calculated spin for the (MYW)•⁺ and net neutral (resting state) MYW-adduct in the R(MYW) models. ................................. 85
5.3 Bar representation of calculated charge distribution for the (MYW)•⁺ and net neutral (resting state) MYW-adduct in the R(MYW) models. ................................. 86
5.4 Molecular orbital diagram of Y-W π-bonding and anti-bonding orbitals in the MYW-radical. ........................................... 87
5.5 The energy profiles for R(MYW) cofactor models as a function of the Y-W torsion angle. ........................................... 90
5.6 Comparison of X-ray crystal structures of oxyferrous ascorbate peroxidase, oxy-cytochrome c peroxidase and M. tuberculosis KatG showing H-bonding of conserved residues. ........................................... 99
5.7 Relative energies for stable oxygen bound and oxygen dissociated structures in four R(MYW)···oxyheme spin states. ........................................... 101
5.8 Potential energy profiles as a function of the Y-W torsional angle for Arg-vicinal R(MYW)···oxyheme states. ........................................... 104
5.9 Potential energy profiles as a function of the Y-W torsional angle for Arg-remote R(MYW)···oxyheme states. ........................................... 108
5.10 Potential energy profiles as a function of Fe-O₂ distance (oxygen-dissociation coordinate) for Arg-vicinal R(MYW)···oxyheme spin states. ........................................... 111
5.11 Potential energy profiles as a function of Fe-O₂ distance (oxygen-dissociation coordinate) for Arg-remote R(MYW)···oxyheme spin states. ........................................... 115
5.12 Stabilization of the \((d^*_{2z} - \pi^*)\) bonding orbital of the Fe-O\(_2\) complex modulated by the MYW-radical opening process in KatG. ........................................ 116

5.13 Bar representation of calculated atomic charge values for the MYW-radical in \(R(\text{MYW}) \cdots \text{oxyheme}\) models for both S=3/2 spin states. ......................... 117

5.14 Representation of calculated atomic charge values for the MYW-radical in \(R(\text{MYW}) \cdots \text{oxyheme}\) models for the two high spin states. ......................... 118

5.15 Computed orbital energies, electron occupations and spin density drawings for four adduct conformations in the \((S=3/2, Ms=1/2)\) spin state. ............... 120

5.16 Computed orbital energies, electron occupations and spin density drawings for four adduct conformations in the \((S=3/2, Ms=3/2)\) spin state. ............... 121

5.17 Computed orbital energies, electron occupations and spin density drawings for four adduct conformations in the \((S=5/2, Ms=5/2)\) spin state. ............... 124

5.18 Computed orbital energies, electron occupations and spin density drawings for four adduct conformations in the \((S=7/2, Ms=7/2)\) spin state. ............... 125

5.19 The D-band EPR spectrum of the MYW radical in KatG.\(^{132}\) ......................... 129

5.20 The \(g_{xx}\) value as a function of the MYW opening coordinate for Arg-remote MYW (purple) and Arg-vicinal MYW (orange) models. .............................. 130

6.1 Comparison between calculations in a vacuum and calculations with the Polarizable Continuum Model. ........................................... 138

6.2 Effect of varying solvents on five computed tyrosyl radicals. .............................. 141

6.3 Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for MYW fragments. .............................. 145

6.4 Effect of varying solvents on models that varied the Met and Trp positions on the MYW adduct. ........................................... 147
6.5 Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for variations of Met and Trp positions on the MYW adduct (pink dots). ................................................................. 148
6.6 Effect of varying solvents on models that replaced Met with either Cys or His (N connectivity). ................................................................. 151
6.7 Effect of varying solvents on models that replaced Met with either His (C connectivity). ................................................................. 151
6.8 Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for variations of the MYW adduct that replaced Met with either Cys or His (black dots). ................................................................. 153
List of Schemes

2.1 Schematic representation of the Hartree-Fock ground state and slater determinants of excited states. 17

4.1 Structure and labeling scheme for desferrioxamine B metal complexes. 46

4.2 Calculated and experimental (EXAFS) average M-O distances for the First Coordination Shell, Second Coordination Shell, and Third Coordination Shell of metal-HDFOB complexes. 47

4.3 The asymmetric (amide I-like) skeletal vibration of metal coordinated oxyamide for various ion coordination depicted by the Lewis structures. 74

5.1 Representation of the four conformers of the isolated R(MYW) adduct, computed Y-W dihedral angles and their relative energies. 88

5.2 R(MYW)···oxyheme model showing a labeling scheme for selected carbon atoms. 92

5.3 Molecular orbital diagram of three models for oxyheme. 94

5.4 Representation of the four conformers of the R(MYW)···oxyheme models. 100

5.5 Proposed stages of the MYW-radical opening process illustrating the effects of reorientation of the SOMO orbital and a hydrogen 1s orbital with respect to the sp² hybrid of the inner oxygen lone pair of oxyheme. 109

5.6 Molecular orbital diagram of R(MYW)···oxyheme models in low (S=3/2) spin states. 119
5.7 Molecular orbital diagram of R(MYW)...oxyheme models in high \((S=5/2)\) spin states. ................................................................. 123

5.8 Molecular orbital diagram of R(MYW)...oxyheme models in the \(S=7/2\) spin state. ................................................................. 123

5.9 Pathway for oxygen release. ................................................................. 127

6.1 The \(S^+\cdots O^*\) interaction and the polarization of electron density in three models: \(Y\), \(MY\), and \(MYW\). ................................................................. 143

6.2 Representation of the four structures calculated to test the effect of the Met and Trp positions on the MYW adduct. ................................................................. 146

6.3 Representation of the structures calculated to test the effect of the identity of the substituent in the Met position on the MYW adduct. Note: structures that omit the Trp residue are not shown here. ................................................................. 150
Chapter 1

Introduction

Nobody ever figures out what life is all about, and it doesn’t matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough.

– R. P. Feynman

Quantum-mechanical calculations have advanced considerably and gained enormous popularity over the past several decades. It is a valuable tool to support or to understand phenomena behind experimental studies. In particular, with recent modern advances in the field, density functional theory is a popular method to study the structural, spectroscopic, and mechanistic properties of molecular systems. A primary goal of this dissertation is to explore the use of density functional theory to investigate the properties of transition metal complexes that are geochemically and biologically relevant as a function of the occupation of $d$-orbital electrons.

This dissertation consists of six chapters. The first three chapters serve as an introduction to the different methodologies available to study biological molecules at the quantum mechanical level. Chapter 1, the current chapter, serves as an overview and introduction to the dissertation. Chapter 2 provides a brief theoretical background of electronic structure methods, and introduces different quantum-mechanical calculations that can be used to probe mechanisms of enzymes and biological molecules. Chapter 3 provides an introduc-
tion to the spectroscopic methods that are relevant to these studies including vibrational spectroscopy, EPR spectroscopy, UV-Vis Spectroscopy, and X-ray absorption spectroscopy.

The remaining chapters discuss applications of density functional theory. Chapter 4 will focus on the use of density functional theory to explore the properties of siderophore-transition-metal complexes. This includes structure, binding, and infrared spectroscopy. Chapter 5 will explore the use of density functional theory to provide insights into a catalytic step of the catalase cycle of the Catalase-Peroxidase (KatG) enzyme. And finally, chapter 6 will focus on the use of density functional theory to explore EPR spectroscopy and the spin distributions of a unique covalently linked tyrosyl-like radical of the KatG enzyme to deepen our understanding of the radical.

Citations to Previously Published Work

The following works have been incorporated into this dissertation.


Chapter 2

Electronic Structure Methods

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws lead to equations much too complicated to be soluble.


2.1 Atoms and Molecules: \( N \)-electron systems

The primary goal of an electronic structure calculation is to solve the non-relativistic time-independent Schrödinger equation,

\[
\hat{H} \Psi_i = E \Psi_i \tag{2.1}
\]

where \( \hat{H} \) is the Hamiltonian operator and \( \Psi \) is the wave function that describes the system in its \( i \)-th state, which is typically the ground-state \( (i = 0) \). The non-relativistic Hamiltonian has the general form:

\[
H = T(\vec{R}) + T(\vec{r}) + V_{N,N} + V_{e,e} + V_{N,e} \tag{2.2}
\]

The kinetic energy operators are given by \( T(\vec{R}) \) and \( T(\vec{r}) \) for nuclei and electrons, respectively. The potential energy operators are given by \( V_{N,N}, V_{e,e}, \) and \( V_{N,e} \) for nuclei, electrons,
and nuclei-electron interactions, respectively. The solution to the hydrogen atom, a one-electron system, is obtained by a separation of variables \((r, \theta, \phi)\) that leads to three equations for the three spatial variables, and their solutions give rise to three quantum numbers that are each associated with a particular variable. The solution gives both the allowed values of the angular momentum and the allowed energies \(E_n\), which agree with the simpler Bohr model.

For many electron systems, the Hamiltonian must be expanded. The full Hamiltonian of a molecular system is then given as:

\[
H_{\text{full}} = -\sum_{i=1}^{\infty} \frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AA}}. \tag{2.3}
\]

In the above equation, \(M_A\) is the ratio of the mass of nucleus \(A\) to the mass of an electron, \(Z_A\) is the atomic number of nucleus \(A\), and the laplacian operators \(\nabla_i^2\) and \(\nabla_A^2\) involve second differentiation with respect to the coordinates of the \(i\)-th electron and the \(A\)-th nucleus. The first and second terms of the equation represent the kinetic energy of the electrons and the nuclei, respectively. The third term represents the coulomb attraction between the electrons and nuclei while the fourth and fifth terms represent the repulsion between electrons and between nuclei, respectively.

Based on the fact that the mass of a nucleus is typically several thousand times larger than that of an electron, it is reasonable to assume that the nuclei in a molecule are stationary with respect to the electrons. Born and Oppenheimer validate this assumption by separation of the molecular Hamiltonian into an electronic and nuclear part.\(^1\) Within this approximation, the electronic Hamiltonian describing the motion of \(N\) electrons in a field of \(M\) point charges, is given by:

\[
H_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} = T_e + V_{en} + V_{ee}. \tag{2.4}
\]

The solution to the time-independent Schrödinger equation involving the electronic Hamil-
tonian \( (H_{\text{elec}} \phi_{\text{elec}} = E_{\text{elec}} \phi_{\text{elec}}) \) is the electronic wave function.

\[
\phi_{\text{elec}} = \phi_{\text{elec}}(r_i; R_A) \tag{2.5}
\]

The electronic wave function and the electronic energy depend on the coordinates of the electron \( (r_i) \) and are parameterized by the nuclear coordinates \( (R_A) \). The total energy must also include the constant nuclear repulsion:

\[
E_{\text{tot}} = E_{\text{elec}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}. \tag{2.6}
\]

If the electronic energy and the total energy have been solved, it is possible to also solve for the motion of the nuclei under the same assumptions. As the electrons move much faster than the nuclei, the electronic coordinates can be replaced by their average over the electronic wave function. The nuclear Hamiltonian for the motion of the nuclei in an average field of electrons is then:

\[
H_{\text{nucl}} = -\sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 + E_{\text{tot}}(R_A). \tag{2.7}
\]

However, for systems with more than one electron, the electronic Schrödinger equation cannot be solved exactly because of the electron-electron interaction, and must be solved approximately. This is a primary reason why many different quantum-mechanical methods are developed. Before these different methods are discussed, the wave function of an \( N \)-electron system will be described.

For an \( N \)-electron system, the wave function is a product of \( N \) one-electron orbitals. To completely describe an electron it is necessary to also include its spin. Two spin functions \( \alpha(\omega) \) and \( \beta(\omega) \) correspond to spin up and spin down, respectively. These two spin functions are complete and orthonormal,

\[
\int d\omega \alpha^*(\omega) \alpha(\omega) = \int d\omega \beta^*(\omega) \beta(\omega) = 1 \tag{2.8}
\]
and
\[ \int d\omega \alpha^*(\omega) \beta(\omega) = \int d\omega \beta^*(\omega) \alpha(\omega) = 0. \] (2.9)

Here, the electron is described not only by three spatial coordinates \( \mathbf{r} \), but also by one spin coordinate \( \omega \). These four coordinates will be denoted collectively by \( \mathbf{x} \),

\[ \mathbf{x} = \{ \mathbf{r}, \omega \}. \]

For \( N \)-electron molecules and atoms, molecular orbitals will describe the electrons. A spatial orbital \( \psi_i(\mathbf{r}) \), is a function of the position vector \( \mathbf{r} \) and describes the probability of finding the electron in a volume element \( d\mathbf{r} \) by \( |\psi_i(\mathbf{r})|^2 d\mathbf{r} \). For each spatial orbital, two different spin orbitals \( \chi(\mathbf{x}) \) may be formed:

\[ \chi(\mathbf{x}) = \psi(\mathbf{r}) \alpha(\omega) \quad \text{or} \quad \psi(\mathbf{r}) \beta(\omega). \] (2.10)

The antisymmetry principle requires that a many electron wave function must be anti-symmetric with respect to the interchange of the coordinate \( \mathbf{x} \) (both space and spin) of any two electrons. Slater determinants satisfy the antisymmetry condition through an appropriate linear combination of Hartree products, which are the non-interacting electron wave functions. The antisymmetric wave function for an \( N \)-electron system can be written as a Slater determinant, where the \( N \) electrons occupy \( N \) orbitals:

\[ \Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = (N!)^{-\frac{1}{2}} \left| \begin{array}{cccc}
\chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\
\chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\
\vdots & \vdots & \ddots & \vdots \\
\chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) 
\end{array} \right|. \]

Interchanging the coordinates of two electrons corresponds to interchanging two rows of the
Slater determinant, which will change the sign of the determinant. Thus, the Slater determinant meets the requirement of the antisymmetry principle. Also, since the determinant will become zero if two electrons occupy the same spin orbital, no more than one electron can occupy a spin orbital (Pauli exclusion principle). A convenient short-hand notation for a normalized Slater determinant that shows only the diagonal elements of the determinant is shown here:

$$\Psi(x_1, x_2, ..., x_N) = |\chi_i(x_1)\chi_j(x_2)\ldots\chi_k(x_N)\rangle.$$  

The $N$-electron Hamiltonian can be separated into two terms: 1) the Hamiltonian acting on one electron of the system $H^{(1)} = T_e + V_{ne}$ and 2) the Hamiltonian acting on two electrons $H^{(2)} = V_{ee}$. The kinetic energy of an electron depends only on that particular electron and the electron-nuclear potential also depends only on a single electron since the nuclear configuration has been fixed. This is in contrast to the electron-electron potential which couples electrons $i$ and $j$ through the $r_{ij}$ distance.

To evaluate the matrix elements of the Hamiltonian, the one- and two-electron integrals must be defined. The one-electron integral is defined as:

$$\langle i|h|j \rangle = \langle \chi_i | h | \chi_j \rangle = \int dx_1 \chi_i^*(x_1) h(r_1) \chi_j(x_1).$$ \hspace{1cm} (2.11)

And the two-electron integral is defined as:

$$\langle ij|kl \rangle = \langle \chi_i \chi_j | \chi_k \chi_l \rangle = \int dx_1 dx_2 \chi_i^*(x_1) \chi_j^*(x_2) r_{12}^{-1} \chi_k(x_1) \chi_l(x_2).$$ \hspace{1cm} (2.12)

The set of rules given for evaluating the matrix elements are known as the Slater-Condon rules. The rules depend on whether the operator is a sum of one-electron operators or two-electron operators, and also the degree to which the two determinants differ.
2.1.1 Basis Sets

An important part of a quantum-mechanical calculation is the basis set. A basis set is a mathematical description of the orbitals within a system, which combine to approximate the total electronic wave function. Basis sets assign a group of basis functions to each atom in a molecule. Two types of basis functions are commonly used: Slater-type functions and Gaussian-type functions. The normalized 1s Slater-type function, centered at $R_A$, has the form

$$\phi_{1s}^{SF}(\zeta, r - R_A) = \left(\frac{\zeta^3}{\pi}\right)^{1/2} e^{-\zeta|r - R_A|} \quad (2.13)$$

where $\zeta$ is the Slater orbital exponent. The normalized 1s Gaussian-type function, centered at $R_A$, has the form

$$\phi_{1s}^{GF}(\alpha, r - R_A) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha|r - R_A|^2} \quad (2.14)$$

where $\alpha$ is the Gaussian orbital exponent. The orbital exponents determine the diffuseness or “size” of the basis functions. A large exponent implies a small dense function, and a small exponent implies a large diffuse function. The major differences between the two functions $e^{-\zeta r}$ and $e^{-\alpha r^2}$ occur at $r = 0$ and large $r$ as shown in Figure 2.1. At $r = 0$ the Slater function has a finite slope and the Gaussian function has zero slope. At larger values of $r$ the Gaussian function decays much more rapidly than the Slater function. For electronic wave function calculations, the Slater functions more correctly describe the qualitative features of the molecular orbitals than Gaussian functions, in particular the 1s orbital of the hydrogen atom. However, in an SCF calculation, the integral involves more than one center, and the Gaussian function is more practical because it allows for a much faster evaluation of the integrals. The incorrect cusp and long range decay of the Gaussian function can be amended by using a linear combination of Gaussian functions, each with a unique orbital exponent, to form a contracted Gaussian basis function,

$$\phi_{\mu}^{CGF}(r - R_A) = \sum_{p=1}^{L} d_{\mu p} \phi_{p}^{GF}(\alpha_{p\mu}, r - R_A) \quad (2.15)$$
where $L$ is the length of the contraction and $d_{\mu \nu}$ is a contraction coefficient. The Gaussian-type orbitals in the contraction are called primitives.

Minimal basis sets use fixed-size atomic-type orbitals and contain the minimum number of basis functions needed for each atom. An example of a minimal basis set is STO-3G, which approximates Slater-type orbitals (STO) with three gaussian primitives per basis function (3G). A major shortcoming of the STO-3G basis set is that these basis functions describe a sphere. Thus, only molecules with spherical environments will be accurately described. Additionally, these basis functions are centered on the atoms. This restricts the description of electron distributions between nuclei, which are an important component of chemical bonds.

A basis set can be increased by increasing the number of basis functions per atom. A split valence basis set, such as 3-21G and 6-31G, represents core atomic orbitals by one set of functions and valence atomic orbitals by two sets of functions. The valence functions are
also split into inner and outer components. Double zeta basis sets, such as the Dunning
Huzinaga basis set (D95), form molecular orbitals from linear combinations of two sizes of
functions for each atomic orbital. Triple split valence basis sets, such as 6-311G, use three
sizes of contracted functions for each orbital.

The basis set can be further improved by providing \( d \)-type functions on main group
elements, and \( p \)-type functions on hydrogen. This allows the electron distribution to extend
between the nuclei. These basis sets are known as polarized basis sets. The 6-31(d) basis
set (also known as 6-31G\(^*\)), adds \( d \) functions to heavy atoms, whereas the 6-31(d,p) (also
known as 6-31G\(^**\)) adds \( p \) functions to hydrogen atoms in addition to the \( d \) functions on
heavy atoms.

For systems in which the electrons are relatively far from the nucleus (for example anions),
diffuse functions may be added. As mentioned above, a small orbital exponent implies a
large diffuse function, which increases the size of the orbitals, allowing for a more accurate
description of long range interactions. The 6-31+G(d) basis set adds diffuse functions to the
heavy atoms, whereas the 6-31++G(d) basis set adds diffuse functions to both the heavy
atoms and the hydrogens.

Additionally, there are correlation-consistent basis sets which were designed to converge
systematically to the complete-basis-set (CBS) limit using empirical extrapolation tech-
niques. For first and second row atoms, the correlation-consistent basis sets are designated
by cc-pV\( N \)Z where \( N=D,T,Q,5,6 \) (D=double, T=triples, etc.). They include successively
larger shells of polarization (correlating) functions (d, f, g, etc.).

For heavier atoms, in the fourth row of the periodic table and below, the core electrons
are often approximated by an effective core potential (ECP), while the valence electrons are
described by traditional Gaussians. The effective potential also corrects for relativistic effects
that are increased for these atoms. The LANL2DZ basis set is the most well-known of these,
and others may be obtained from the basis set exchange library (https://bse.pnl.gov/bse/portal).


2.2 Hartree Fock Theory

The Hartree-Fock (HF) method approximates the difficult electron-electron repulsion term by a single slater-determinant that experiences a mean-electric-field of all other electrons in the atom or molecule. Thus, the essence of the Hartree Fock approximation is to replace the many-electron problem with a one-electron problem in which the electron-electron repulsion is treated in average way. The single-electron problem is solved iteratively to determine the most favorable arrangement of the electrons. The mean-field is then recomputed, and the single-electron problem is solved again using the new mean-field potential. This process continues until the average potential and the solutions to the single-electron problem converge from one iteration to the next. For this reason, mean-field approximations are known as self-consistent field (SCF) methods.

The antisymmetric wave function required to describe the ground state of an \(N\)-electron system has the form of a single Slater determinant,

\[
|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_N\rangle. \tag{2.16}
\]

The lowest possible energy is given according to the variation principle,

\[
E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \tag{2.17}
\]

where \(H\) is the full electronic Hamiltonian. By systematically varying the orbitals \(\chi_a\) while constraining that they are orthonormal \((\langle \chi_a | \chi_b \rangle = \delta_{ab})\) until the electronic minimum is reached, the optimum orbitals can be found and leads to the Hartree Fock equation:

\[
f(i)\chi(x_i) = \epsilon\chi(x_i) \tag{2.18}
\]
where $f(i)$ is an effective one-electron operator called the Fock operator,

$$f(i) = -\frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + v^{HF}(i).$$  \hfill (2.19)

The Fock operator is the mean energy of an electron that occupies an orbital in the presence of all other electrons. It is the sum of the kinetic energy, the nuclear attraction, and the effective potential ($v^{HF}$) experienced by the $i$-th electron that includes a coulomb and exchange interaction.

### 2.2.1 Coulomb and Exchange Integrals

In the Hartree-Fock approximation, electron 1 in $\chi_a$ experiences a one-electron coulomb potential

$$v_{a\text{coul}}^1(1) = \sum_{b \neq a} \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1}. \hfill (2.20)$$

By summing over all $b \neq a$, one obtains the average potential acting on the electron $\chi_a$, arising from the $N-1$ electrons in the other spin orbitals. The coulomb operator is defined as,

$$J_b(1) = \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1}$$  \hfill (2.21)

which represents the average local potential at $\mathbf{x}_1$ arising from an electron in $\chi_b$. The coulomb operator acting on a spin orbital $\chi_a(1)$ results in

$$J_b(1)\chi_a(1) = \left[ \int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_b(2) \right] \chi_a(1). \hfill (2.22)$$

The exchange term arises from the antisymmetric nature of the single determinant. The exchange operator $K_b(1)$ is defined by its effect when operating on a spin orbital $\chi_a(1)$,

$$K_b(1)\chi_a(1) = \left[ \int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_a(2) \right] \chi_b(1). \hfill (2.23)$$
Operating with $K_a(1)$ on $\chi_a(1)$ involves an “exchange” of electron 1 and electron 2.

A major deficiency of Hartree-Fock theory is that it cannot provide a complete description of the correlation between motions of electrons. The single-determinant approximation does not take into account Coulomb correlation, leading to a total electronic energy different from the exact solution of the non-relativistic Schrödinger equation within the Born-Oppenheimer approximation. Therefore, the Hartree-Fock limit is always above this exact energy. This difference is known as the correlation energy. To correct for such a deficiency it is necessary to use wave functions that go beyond a single-determinant. Great effort has been used to develop methods that can recover the correlation energy and thus a variety of different quantum chemical \textit{ab initio} methods have emerged.

### 2.3 Møller-Plesset Many-Body Perturbation Theory

Perturbation theory can be used to recover the correlation energy and improve the accuracy of the Hartree-Fock wave function. As first implemented by Møller and Plesset in 1934\(^2\), these methods are named according to the order of perturbation theory (MP1, MP2, ..., MPn). The MPn method attempts to recover the correlation energy by including a correction term (a perturbation adjustment) to the HF energy.

For electron $m$ in an $N$-electron molecule, the MP unperturbed Hamiltonian is given as follows:

$$H = H_0 + V$$  \hspace{1cm} (2.24) \hspace{1cm}

where $H_0$ is the Hartree-Fock Hamiltonian and $V$ is given as

$$V = \sum_{i>j} \frac{1}{r_{i,j}} - \sum v^{\text{HF}}(i).$$  \hspace{1cm} (2.25) \hspace{1cm}

The Hartree-Fock wave function is the ground state of $H^{\text{HF}}$ with eigenvalue $E^{(0)} = \sum_i \epsilon_i$. 


The first order correction to the energy is then

\[ E^{(1)} = \langle \Psi_{HF} | H^{(1)} | \Psi_{HF} \rangle, \]  

(2.26)

where

\[ H^{(1)} = H - H^{(0)} = \sum_{i} \left( \sum_{j:i>j} \frac{1}{r_{ij}} - V^{HF}(x_i) \right). \]  

(2.27)

The molecular energy is the sum of the first order correction and the zeroth order energy, which yields the Hartree-Fock energy.

\[ E^{(0)} + E^{(1)} = E_{(HF)} \]  

(2.28)

The second order correction (MP2) is the first MP level to go beyond the HF treatment. The MP2 energy is the HF energy plus a correction term (a perturbation adjustment) that represents a lowering of energy brought about by accounting for the interelectron repulsion, and can be written as

\[ E^{(0)} + E^{(1)} + E^{(2)} = E_{(HF)} + E^{(2)}. \]  

(2.29)

The HF term includes the internuclear repulsions, whereas the perturbation correction \( E^{(2)} \) is a purely electronic term. \( E^{(2)} \) is a sum of terms that model the promotion of pairs of electrons, which are also known as double excitations.

\[ E^{(2)}_0 = \sum_{n'} \frac{\langle \Psi_0 | V | \Psi_{n'} \rangle}{E^{(0)}_0 - E^{(0)}_{n'}}, \]  

(2.30)

where \( \Psi_0 \) is the ground state and \( \Psi_{n'} \) represents all possible Slater determinants formed from \( n' \) different spin-orbitals.

Perturbation theory is useful for improving upon the Hartree-Fock solutions because it includes contributions from many other determinants to reproduce dynamical correlations. That said, MPn does not always converge to the exact solution nor does higher order pertur-
bation always improve the wave function. Instead, in some cases, higher order perturbation corrections will have errors greater than the Hartree-Fock solution. In most cases, MP2 is a relatively inexpensive way to recover around 85 to 95% of the correlation energy. Higher orders of Møller-Plesset theory are mostly used for methodology testing rather than practical applications. The remaining correlation energy calls for a multiconfigurational wave function.

2.4 Multireference Approaches

To capture the multireference nature of the electronic problem, various techniques have been developed to address the shortcomings of the single-determinant Hartree-Fock and Møller-Plesset theories by including multi-determinant wave functions. Each determinant represents an idealized configuration, known as a configuration state function (CSF). A CSF is a linear combination of Slater determinants and is an eigenfunction of the spin operators $S^2$ and $S_z$ and satisfies the spatial symmetry requirements of the molecule. Multireference approaches consider contributions from other electronic configurations and thus must include more than one determinant. To do so, the multireference wave function is a linear combination of CSFs.

2.4.1 Configuration Interactions

The goal of the Full Configuration Interaction (FCI) algorithm is to solve the secular equation and obtain the eigenstates for the $N$-electron Hamiltonian by considering all possible determinants. The configuration functions in a CI calculation are classified according the number of electrons that are excited from occupied to virtual orbitals. In the CI expansion $\psi = \sum_i b_i \phi_i$, configuration functions with the same symmetry eigenvalues as the reference state can be included. The number of possible configuration functions with the proper symmetry increases as the number of electrons and the number of basis functions increase. For $N$ electrons and $b$ basis functions, the number of configuration functions is roughly proportional
to $b^N$. Because of the very large number of configuration functions, full CI calculations can only be used for small molecules (small $N$) and small basis sets (small $b$); however, even for small molecules full CI calculations are expensive.

Instead of considering all excitations from the reference state, approximate configuration interaction schemes are often employed. For example, optimization may only include configurations that differ by single excitations from the reference state (CIS), by singles and double excitations (CISD), or by singles, double, and triple excitations (CISDT), see Scheme 2.1. This greatly reduces the computational cost; however, for CI, calculations on simple model systems indicate that the percentage of basis-set correlation energy decreases as the size of the molecule increases, thus producing a basis-set truncation error. Also, it has been shown that these calculations are neither size extensive nor size consistent. A method is said to be size extensive if the energy calculated thereby scales linearly with the number of particles $N$. The word extensive is used in the same sense as in thermodynamics, when referring to an extensive, rather than an intensive property. A method is called size consistent if it gives an energy $E_A + E_B$ for two well separated subsystems $A$ and $B$. While the definition of size extensivity applies at any geometry, the concept of size consistency applies only in the limiting case of infinite separation.

### 2.4.2 Coupled Cluster Methods

Another attempt to simplify the expensive Full CI calculation is by coupled cluster methods. These methods are a class of perturbative techniques that focus on the physical nature of the excitation rather than the configurations. The basic idea is to express the correlated wave function as a sum of determinants by allowing a series of operators ($\hat{T}$) to act on the HF wave function. The fundamental equation in coupled-cluster theory is

$$\psi = e^{\hat{T}} \phi_0.$$  \hspace{1cm} (2.31)
Scheme 2.1: Schematic representation of the Hartree-Fock ground state and Slater determinants of excited states.

Here, $\psi$ is the exact non-relativistic ground-state molecular electronic wave function, $\phi_0$ is the normalized ground-state Hartree-Fock wave function, and the operator $e^{\hat{T}}$ is defined by the Taylor-series expansion

$$e^{\hat{T}} \equiv 1 + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \ldots = \sum_{k=0}^{\infty} \frac{\hat{T}^k}{k!}. \quad (2.32)$$

The cluster operator $\hat{T}$ is given by

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 + \ldots + \hat{T}_n \quad (2.33)$$

where the operators $\hat{T}_1, \hat{T}_2 \ldots \hat{T}_n$ are excitation operators that have the effect of promoting $1, 2, \ldots, n$ electrons into virtual orbitals. Specifically, $\hat{T}_1$ is the one-particle excitation operator,

$$\hat{T}_1 \phi_0 \equiv \sum_{a=n+1}^{\infty} \sum_{i=1}^{n} t^a_i \phi^a_i \quad (2.34)$$
and $\hat{T}_2$ is the two-particle excitation operator,

$$
\hat{T}_2 \phi_0 \equiv \sum_{b=a+1}^\infty \sum_{a=n+1}^\infty \sum_{j=i+1}^{n-1} \sum_{i=1}^{n-1} t_{ij}^{ab} \phi_{ij}^{ab}.
$$

(2.35)

Here, $\phi_i^a$ is a singly excited Slater determinant with the occupied spin-orbital, $u_i$, replaced by the virtual spin-orbital, $u_a$, and $t_{ij}^{n}$ is a numerical coefficient whose value depends on $i$ and $a$. The effect of the $e\hat{T}$ operator is to express the wave function as a linear combination of Slater determinants over all excitations. The aim of a CC calculation is to find the coefficients $t_{ij}^n$, $t_{ij}^{ab}$, $t_{ijk}^{abc}$, etc for all excitations. The coupled-cluster energy is obtained by a projection onto the Hartree-Fock state

$$
E = \left\langle \phi_0 | \hat{H} | e^{\hat{T}} \phi_0 \right\rangle.
$$

(2.36)

And for the excited configuration becomes

$$
\left\langle \phi_{ij}^{ab} | \hat{H} | e^{\hat{T}} \phi_0 \right\rangle = E \left\langle \phi_{ij}^{ab} | e^{\hat{T}} \phi_0 \right\rangle.
$$

(2.37)

The operator $\hat{T}$ is usually truncated at some level. Theory has shown that the most important contribution of $\hat{T}$ is made by $\hat{T}_2$. Inclusion of only double excitations gives an approximate CC approach called the coupled-cluster doubles (CCD) method. Similarly, one can keep singles and doubles (CCSD) or singles, doubles, and triples (CCSDT).

### 2.4.3 Complete Active Space (CASSCF)

Another commonly used multireference method is known as the multiconfigurational SCF (MCSCF) approach. The most commonly used type of MCSCF method is known as the complete active space SCF (CASSCF) method. In this method, a set of orbitals known as the active space, is carefully chosen by the user based on their relevance to the given problem. All excitations using a fixed number of electrons in the chosen orbitals are then permitted.
The molecular wave function is first written as a linear combination of CSFs \((\phi_i)\) and the expansion coefficients \(b_i\) (from \(\psi = \sum_i b_i \phi_i\)) as well as the molecular orbitals are varied. The chosen orbitals are divided into inactive and active orbitals, in which the inactive orbitals are kept doubly occupied. The wave function is then written as a linear combination of all CSFs \((\phi_i)\) that can be formed by distributing the active electrons among the active orbitals in all possible ways. The CASSCF calculation is then performed to find the optimum coefficients \((c_{ri} \text{ and } b_i)\).

The general form of a CASSCF wave function can be represented as

\[
\Psi_{\text{CASSCF}} = \sum_I P_I \Phi_I, \tag{2.38}
\]

where

\[
\Phi_I = Q \Pi \varphi_k, \tag{2.39}
\]

and

\[
\varphi_k = \sum_{\mu} C_{\mu k} \chi_{\mu}. \tag{2.40}
\]

\(\Phi_I\) is the configuration state functions and \(I\) runs over all contributing configurations, \(\varphi_k\) is the molecular orbitals expanded in the basis of atomic basis functions \(\chi_{\mu}\), and \(Q\) is a normalization constant. The CASSCF wave function is written as a linear combination of many-electron functions and can be obtained by simultaneously optimizing the configuration mixing coefficients \(P_I\) and molecular orbital coefficient \(C_{\mu k}\) using a variational method. This approach allows the method to both include the relevant near-degenerate configurations in the electronic states under study and to provide a foundation for more sophisticated theories that would bring the computed properties closer to the exact solution of the Schrödinger equation. Additional electronic correlation effects can then be accounted for by adding a perturbation correction (CASMP2, CASMP3, or CASMP4).

A major shortcoming of the CASSCF methods is the possibility of intruder states. In-
truder states are defined as those of the intermediate states that have energies that are comparable in magnitude to the energy associated to the zero order wave function. In this case, divergent behavior occurs due to the nearly zero denominator in the expression of the perturbative correction.

2.5 Density Functional Theory

The electronic structure methods presented thus far were concerned with obtaining improved wave functions and are therefore called wave function methods. Density functional theory (DFT) introduces a new approach that considers the electron density as a major determinant of the energy and other properties of an atom or molecule. Whereas the wave function is not an observable feature of a molecule or atom, the electron density ($\rho$) is measurable by experiment. Additionally, the electron density is a function of position only, that is, of just three variables ($x, y, z$) while the wave function of an $N$-electron wave function is a function of $3N$ spatial coordinates plus $N$ spin coordinates of each electron. Thus, the electron density function method has benefits over the wave function method because the electron density ($\rho$) is measurable and mathematically more tangible.

The theorems of Hohenberg and Kohn proved that for systems with a non-degenerate ground state, the ground-state electron probability density $\rho_0(r)$ determines the external potential and the number of electrons. Thus, in principle, if the ground-state electron density $\rho_0(r)$ is known, it is possible to calculate all ground-state molecular properties from $\rho_0$, without having to solve the molecular wave function. The energy as a function of density can be written as follows:

$$
\langle \Psi | \hat{H} | \Psi \rangle = E[\rho(r)] = T[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] = \int \rho(r)v(r)dr + F_{HK}[\rho(r)]. \quad (2.41)
$$

Here, $T[\rho(r)]$ describes the kinetic energy, $V_{ne}[\rho(r)] = \int \rho(r)v(r)dr$ describes the nuclei-electron interaction, $V_{ee}[\rho(r)]$ describes the electron-electron interaction, and $F_{HK}[\rho(r)]$ is
the sum of the electron-electron interaction energy and the kinetic energy functionals.

In 1965, Kohn and Sham designed a method for practical implementation for finding \( \rho_0 \) and for finding \( E_0 \) from \( \rho_0 \). The two basic ideas behind the Kohn-Sham (KS) approach are to: (1) express the molecular energy as a sum of terms, only one of which involves an “unknown” functional; and (2) to use an initial guess of the electron density \( \rho_0 \) in the KS equations (analogous to the HF equations) to iteratively refine the KS orbitals and energy levels, in a manner similar to the HF SCF method.

A system of \( n \) non-interacting electrons each experience the same external potential-energy function \( v_s(\mathbf{r}_i) \). The potential-energy function is chosen such that the ground-state electron probability density \( \rho_s(\mathbf{r}) \) of the system is equal to the exact ground-state electron density \( \rho_0(\mathbf{r}) \). Thus, the electron probability density of an \( n \)-particle system whose wave function is a Slater determinant of the spin-orbitals \( u_{KS}^i = \theta_{KS}^i \sigma_i \) is given by \( \sum_{i=1}^{n} |\theta_{KS}^i|^2 \); therefore,

\[
\rho = \rho_s = \sum_{i=1}^{n} |\theta_{KS}^i|^2. \tag{2.42}
\]

The Hamiltonian of the non-interacting reference system is

\[
\hat{H}_s = \sum_{i=1}^{n} \left[ -\frac{1}{2} \nabla_i^2 + v_s(\mathbf{r}_i) \right] \equiv \sum_{i=1}^{n} \hat{h}_{KS}^i,
\]

where \( \hat{h}_{KS}^i \) is the one-electron Kohn-Sham Hamiltonian given by

\[
\hat{h}_{KS}^i \equiv -\frac{1}{2} \nabla_i^2 + v_s(\mathbf{r}_i).
\]

The ground-state wave function \( \psi_{s,0} \) of the reference system is the antisymmetrized product (Slater determinant) of the lowest-energy Kohn-Sham spin-orbitals \( u_{KS}^i \), where the spatial part \( \theta_{KS}^i(\mathbf{r}_i) \) of each spin-orbital is an eigenfunction of the one-electron operator \( \hat{h}_{KS}^i \):
\[ u_i^{KS} = \theta_i^{KS} = \theta_i^{KS}(r_1)\sigma_i \]
\[ h_i^{KS}\theta_i^{KS} = \epsilon_i^{KS}\theta_i^{KS} \]

where \( \sigma_i \) is a spin function (either \( \alpha \) or \( \beta \)), and \( \epsilon_i^{KS} \) is the Kohn-Sham orbital energy.

Kohn and Sham rewrote the Hohenberg-Kohn equation as

\[
E_v[\rho] = \int \rho(r)v(r)dr + T_s[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{r_{12}} + \Delta T[\rho] + \Delta V_{ee}[\rho],
\]

where \( \Delta T \) is the difference in the average ground-state electronic kinetic energy between the molecule and the reference system,

\[ \Delta T[\rho] \equiv T[\rho] - T_s[\rho], \]

and

\[ \Delta V_{ee}[\rho] \equiv V_{ee}[\rho] - \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1dr_2. \]

The exchange-correlation energy functional is defined as

\[ E_{xc}[\rho] \equiv \Delta T[\rho]\Delta V_{ee}[\rho]. \] 

Thus,

\[
E_0 = E_v[\rho] = \int \rho(r)v(r)dr + T_s[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1dr_2 + E_{xc}[\rho].
\]

The first three terms on the right side of Eq. (2.45) are easily evaluated from \( \rho \) and include the main contributions to the ground-state energy. The first term represents the potential energy of nuclear attraction; the second term represents the kinetic energy of the non-interacting electron reference system; and the third term is the classical electrostatic repulsion energy term. The fourth term, \( E_{xc} \), is not known explicitly from \( \rho \). This term
accounts for the non-classical electron-electron interactions, additional kinetic energy terms and the self-interaction of electrons and must be approximated to obtain an accurate KS DFT calculation.

Expanding the first three terms of Eq. 2.45, the expression for the molecular electronic energy becomes

$$E_0 = - \sum_\alpha Z_\alpha \int \frac{\rho(r_1)}{r_{1\alpha}} dr_1 - \frac{1}{2} \sum_{i=1}^n \langle \theta_i^{KS}(1) | \nabla^2 | \theta_i^{KS}(1) \rangle$$

$$+ \frac{1}{2} \int \int \frac{\rho(r_1) \rho(r_2)}{r_{12}} dr_1 dr_2 + E_{xc}[\rho]. \quad (2.46)$$

By determining the KS orbitals ($\theta_i^{KS}$) and approximating $E_{xc}$, $E_0$ can be found from $\rho$. The KS orbitals are found as follows.

First, the Hohenberg-Kohn variational theorem states that the ground-state energy can be found by varying $\rho$ so as to minimize the functional $E_v[\rho]$. However, instead of varying $\rho$, it is equivalent to vary the KS orbitals $\theta_i^{KS}$ which determine $\rho$ by Eq. (2.42). The Kohn-Sham orbitals minimize the expression in Eq. (2.46) for the molecular ground-state energy and satisfy

$$\hat{h}^{KS}(1) \theta_i^{KS}(1) = \epsilon_i^{KS} \theta_i^{KS}(1) \quad (2.47)$$

where

$$\hat{h}^{KS}(1) = \left[ -\frac{1}{2} \nabla^2 - \sum_\alpha \frac{Z_\alpha}{r_{1\alpha}} + \int \frac{\rho(r_2)}{r_{12}} dr_2 + v_{xc}(1) \right] \theta_i^{KS}(1) = \epsilon_i^{KS} \theta_i^{KS}(1). \quad (2.48)$$

The term $v_{xc}$ is the exchange-correlation potential and is related to the exchange-correlation functional through a functional derivative of the exchange-correlation energy $E_{xc}$:

$$v_{xc}(r) \equiv \frac{\delta E_{xc}[\rho(r)]}{\delta \rho(r)}. \quad (2.49)$$
2.5.1 Approximations to the Exchange-Correlation Functional

The earliest and simplest approximation to the exchange-correlation functional is the Local Density Approximation (LDA). In this approximation, the exchange-correlation functional from a uniform electron gas is applied to the system of interest. This produces a local functional whose value at a point depends only on the density at that particular point. If $\rho$ varies with position, then $E_{xc}[\rho]$ is given by

$$E_{xc}^{LDA}[\rho] = \int \rho(r) \epsilon_{xc}(\rho) \, dr$$  \hspace{1cm} (2.50)

where $\epsilon_{xc}(\rho)$ is the exchange plus correlation energy per electron in a uniform electron gas. The LDA method works well in solid state physics applications, where metals are reasonably approximated by a uniform gas. However, most molecules do not have nearly uniform electron densities. This prompted the development of functionals that depend on more than just the local value of the density.

The generalized-gradient approximations (GGA) depend on the density and the gradients of $\rho^\alpha$ and $\rho^\beta$,

$$E_{xc}^{GGA}[\rho^\alpha, \rho^\beta] = \int f(\rho^\alpha(r), \rho^\beta(r), \nabla \rho^\alpha(r), \nabla \rho^\beta(r)) \, dr,$$  \hspace{1cm} (2.51)

where $f$ is a function of the spin densities and their first gradients and $E_{xc}^{GGA}$ is split into exchange and correlation parts,

$$E_{xc}^{GGA} = E_x^{GGA} + E_c^{GGA}.$$  \hspace{1cm} (2.52)

The generalized-gradient exchange and correlation energy functionals are approximated. Commonly used GGA exchange functionals, $E_x$, include Perdew and Wang’s 1986 functional (PW86 or PWx86), Becke’s 1988 functional (B88, Bx88 or Becke88), and Perdew and Wang’s 1991 exchange functional (PWx91). Commonly used GGA correlation functionals, $E_c$, include the Lee-Yang-Parr functional (LYP), the Perdew 1986 correlation functional
(P86 or Pc86), and the Perdew-Wang 1991 parameter-free correlation functional (PW91 or PWc91).

A way to improve the GGA density functionals of Eq. (2.51) is to include the second derivatives of $\rho$ and the non-interacting kinetic energy functional. These functionals are known as meta-GGA (mGGA) functionals and have the form

$$E_{xc}^{MGGA}[\rho^\alpha, \rho^\beta] = \int f(\rho^\alpha, \rho^\beta, \nabla \rho^\alpha, \nabla \rho^\beta, \nabla^2 \rho^\alpha, \nabla^2 \rho^\beta, \tau_\alpha, \tau_\beta) d\mathbf{r}$$  \hspace{1cm} (2.53)

where the Kohn-Sham kinetic-energy density for the spin-$\alpha$ electrons is defined by

$$\tau_\alpha \equiv \frac{1}{2} \sum_i |\nabla \theta_{i\alpha}^{KS}|^2.$$

The most widely used approximation to the exchange-correlation functional are the hybrid functionals. These functionals use linear combinations of GGA’s and the exact exchange contribution derived from Hartree-Fock. The primary functional is B3LYP, which is a linear combination of the LDA, the exact exchange functional, the B88 functional (Becke 1988), and the LYP functional (Lee, Yang, and Parr). The number 3 indicates a three-parameter functional whose values were chosen to fit experimental values. By mixing the exact exchange with approximate exchange, often errors in the correlation can be canceled by errors in the exchange approximation. The success of B3LYP, and the agreement with experimental results, can largely be attributed to a cancellation of errors in the exchange approximation with errors in the correlation approximation.

**Dispersion-corrected DFT**

The success and challenges of DFT can both be attributed to the exchange-correlation functional, $E_{xc}$. The current functionals perform well for the calculation of a range of properties including energetics, geometries, and reaction barriers. However, they fail to correctly describe the long-range dispersion interaction between molecules which greatly limits their
applicability to systems where dispersion interactions are important. Strategies known as
dispersion-corrected DFT have been developed to circumvent this problem. Dispersion-
corrected DFT functionals combine a conventional functional with an add-on energy term.
The dispersion correction energy term is a relatively simple function of interatomic distances
and contains adjustable parameters that are fitted to conformational and interaction ener-
gies computed using CCSD(T)/CBS for a given functional. CBS refers to the complete basis
set limit, which is an extrapolated estimate of a result obtained using an infinitely large
(complete) basis set. DFT-D and DFT-D2 energy corrections consider all pairs of atoms
while DFT-D3 also considers triplets of atoms to account for three-body effects. Because
the dispersion correction is an add-on term it does not directly alter the wave function or
any other molecular property. However, geometry optimizations with dispersion correction
will lead to a different geometry than without because the dispersion correction contributes
to the forces acting on the atoms.

2.5.2 Integration Grid

Another important component of a DFT calculation is the integration grid to be used for
numerical integrations. The accuracy of DFT calculations depends on the number of points
used in the numerical integration. Grids are specified by the number of radial shells around
each atom, each of which contains a set number of integration points.

Uniform grids are those which contain the same number of angular points at each radial
distance. Pruned grids have been optimized to use the minimal number of points required
to achieve a given level of accuracy. These grids are less dense on the shells near the core
and far from the nucleus, and more dense in the region of the atom where properties are
changing most rapidly.

The “fine” integration grid, a pruned (75,302) grid, having 75 radial shells and 302 angular
points per shell, resulting in about 7,000 points per atom, greatly enhances calculation
accuracy at a minimal additional cost. An ultra-fine grid is a pruned (99,590) grid, and
is recommended for molecules containing multiple tetrahedral centers and for computing very low frequency modes of systems. This grid is also useful for optimizations of larger molecules. The superfine grid is about 3x larger than the ultra-fine and is useful when very high accuracy is desired. The grid specification is (150,974) for the first two rows of the periodic table and (225,974) for later elements.

2.6 Time-Dependent Density Functional Theory

Time-dependent density functional theory (TD-DFT) is a formally exact approach to the time-dependent electronic many-bodied problem, and can be used to calculate electronic excitation energies. In contrast, DFT is a ground-state theory, and is used to find the ground state of a quantum system. For most systems, the ground state of the system must be determined before its excitations can be calculated.

The time-dependent Schrödinger equation is given by

\[ i \frac{\partial}{\partial t} \Psi_j(r_1, \ldots, r_N, t), \]  

where the time-dependent Hamiltonian is defined as

\[ \hat{H}(t) = \hat{T} + \hat{V}(t) + \hat{W}. \]

The time-dependent Hamiltonian features an external potential operator that is explicitly time-dependent:

\[ \hat{V}(t) = \sum_{j=1}^{N} v(r_j, t). \]

Starting with \( t_0 = 0 \) and \( \Psi(t_0) \equiv \Psi_0 \) (not necessarily the ground state), the state is propagated in time, and describes how the system evolves under the influence of the time-
dependent potential \( v(r, t) \), which can be written as

\[
v(r, t) = v_0(r + \theta(t - t_0)v_1(r, t).
\] (2.57)

The time-dependent xc potential is formally a functional of the time-dependent density as well as the initial states, \( v_{xc}[n, \Psi_0, \Phi_0](r, t) \). If the system of interest is initially in the ground state, the Hohenberg-Kohn theorem states that the initial states become functionals of the initial (ground-state) density, and the xc functional can be written as a density functional only, \( v_{xc}[n](r, t) \). However, the density dependence of the xc potential is complicated and nonlocal: the xc potential at space-time point \((r, t)\) depends on densities at all other points in space and at all previous times, \( n(r', t') \), where \( t' \leq t \). The most widely used approximation for the xc potential is the adiabatic approximation:

\[
v_{xc}^A(r, t) = v_{xc}^g[n_0](r)|_{n_0(r, t)},
\] (2.58)

where \( v_{xc}^g \), the ground-state xc potential, is evaluated at the instantaneous time-dependent density. This equation is exact for an infinitely slowly varying system; however, in practice, this is generally not the case. Several studies have shown that the adiabatic approximation may break down if the system undergoes very rapid changes; however, it still manages to work quite well and is used often due to its simplicity.

The adiabatic local density approximation (ALDA) can also be used, which uses the exchange-correlation potential of the local density approximation (LDA) as the ground-state functional. As discussed previously, the LDA approximates the exchange-correlation potential as a uniform gas evaluated on the local density at each point in space. The ALDA exchange-correlation potential is given by:

\[
v_{xc}^{ALDA}(r, t) = v_{xc}^{LDA}(n(r, t)).
\] (2.59)
For greater accuracy, the adiabatic approximation can also be based on hybrid functionals such as B3LYP.

### 2.7 Two Approaches to Biological Molecules: QM/MM versus Standard DFT

The use of computational chemistry methods to probe an enzyme mechanism has become a powerful complement and alternative to experimental methods. Currently, two approaches are commonly used: the cluster approach that uses quantum mechanics only (QM)\(^3\)\(^-\)\(^6\) or the hybrid quantum-mechanics/molecular-mechanics method (QM/MM)\(^7\)\(^-\)\(^9\). Both methods have been applied successfully to the study of a variety of enzymes, and in many cases similar results and conclusions have been obtained.

The cluster approach\(^3\)\(^-\)\(^6\) for modeling enzymes has been used successfully for nearly three decades. In this approach, a relatively small but carefully selected part of the enzyme is cut out and treated using quantum mechanical methods. The size of active site models is generally on the order of 100-150 atoms, and is typically modeled using the hybrid DFT method with the B3LYP exchange correlation functional. The remaining part of the protein that is not explicitly included can affect the cluster model approach through both steric and electrostatic effects. First, the protein matrix can impose steric constraints on the active site, and if not properly accounted for, can lead to large artificial movements of various groups. This can lead to an improper description of the active site's reactivity. By applying structural constraints that lock key coordinates of the model, this steric error can be successfully avoided.\(^10\) The constraints prevent large artificial movements of the active site and ensure that the model remains close to the experimentally determined structure. Second, the protein may impart long-range polarization that can affect the computed energies. To account for these electrostatic effects, polarizable continuum techniques are commonly used. Typically, a dielectric constant of \(\epsilon = 4\) is considered to be an accurate representation of the
protein surrounding;\textsuperscript{3–5,11–13} however a larger dielectric constant ($\epsilon = 78$) may be appropriate in cases where a hydrophilic surrounding or a large hydrogen bond network is present.

The second approach, QM/MM, applies quantum mechanics (QM) to the reactive site of the protein and molecular mechanics (MM) to the remaining parts of the system.\textsuperscript{7–9} This method attempts to describe a much larger portion of the active site, and in some cases the entire protein. Similar to the cluster approach, a relatively small but carefully selected part of the enzyme is treated using quantum mechanical methods. However, instead of polarizable continuum methods to account for electrostatic effects, the remaining parts of the protein are described by classical force fields. The interaction between the QM part and MM part is usually treated by a mechanical or electronic embedding scheme. A mechanical embedding (ME) computes the QM part of the system in the absence of the MM part, and then treats the interactions at the boundary between the QM and MM parts using molecular mechanics. These interactions usually include bonding (stretching, bending, and torsional) and non-bonding (electrostatic and van der Waals) interactions. An electrostatic embedding (EE) scheme computes the QM part in the presence of the MM part by including MM point charges in the QM Hamiltonian that describe the electrostatic interaction at the boundary.

A major drawback of the QM/MM approach is that a large number of local minima can be obtained. A recent study on the reaction mechanism of tyrosinase\textsuperscript{14}, showed that the original QM/MM study suffered from convergence to an incorrect reactant state and to different local minima along the reaction pathway\textsuperscript{15}. But once this problem was corrected by an iterative optimization procedure, the QM/MM method gave similar results to a cluster model study.\textsuperscript{14} Additionally, the QM/MM approach is more complicated than the cluster model approach and its success is largely dependent on the choice of the QM region. In principle, if the QM region of QM/MM is sufficiently large, the results should agree with the cluster model approach. However, the typical difference between these two approaches is that the QM part is much larger in the cluster approach while the MM part is very large in the QM/MM approach.
Chapter 3

Spectroscopy

The photons which constitute a ray of light behave like intelligent human beings: out of all possible curves they always select the one which will take them most quickly to their goal.

–Max Planck

Spectroscopy is the study of the interaction between light and matter. Various techniques measure this interaction, and are valuable and powerful tools to probe atoms and molecules. These techniques may be classified by the wavelength of radiation applied in the experiment, and produce a spectrum consisting of a plot of light intensity as a function of frequency or wavelength. The absorption or emission spectrum is associated with a transition between two allowed discrete energy levels of an atom or molecule. The frequency at which energy is absorbed or emitted is related to the energy levels involved in the transitions by

\[ h\nu = |E_2 - E_1|. \] (3.1)

The fact that atoms and molecules possess a set of discrete energy levels is an essential component of spectroscopy. The energy levels are characteristic and structurally unique to a particular atom or molecule, and the spectra provides important structural and electronic
information about the system.

### 3.1 Vibrational Spectroscopy

According to Hooke’s law, an object will oscillate with simple harmonic motion:

$$ F = -kx = m \frac{d^2 x}{dt^2} $$

(3.2)

where $m$ is the mass of the object being displaced from its equilibrium position, $x$ is the magnitude of the displacement, $F$ is the restoring force, and $k$ is the proportionality constant (also known as the force constant).

The simplest molecular case is that of a diatomic molecule. If the masses of the two atoms are given as $m_1$ and $m_2$, Hooke’s Law becomes

$$ \frac{m_1 m_2}{m_1 + m_2} \frac{d^2 (r - r_e)}{dt^2} = k(r - r_e), $$

(3.3)

where the harmonic oscillation around the equilibrium distance is

$$ (r - r_e) = A \sin 2\pi \nu t, $$

(3.4)

and

$$ \nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}}. $$

(3.5)

Here, $r_e$ is the equilibrium internuclear distance, $r$ is the instantaneous internuclear distance, $\mu$ is the reduced mass of the system, and $k$ (force constant) relates to the strength of the bond in the molecule. A characteristic frequency ($\nu$) for a diatomic molecule is equivalent to the frequency of a mass ($\mu$) attached to a spring with the force constant $k$, and thus the two-particle vibrational problem is simplified to a one-particle problem.

In order to find fundamental frequencies for polyatomic molecules ($N$-atoms), one has to
construct a set of equations (\(3N - 6\) or \(3N - 5\) for linear molecules) which can be summarized in matrix notation. All characteristic molecular frequencies along with their vibrational modes can be obtained by solving the matrix equation (eigenvalue equation):

\[
\mathbf{G}\mathbf{F}\mathbf{L} = \lambda_i \mathbf{L},
\]

where the eigenvalue solution is

\[
\lambda_i = 4\pi^2 c^2 \tilde{\nu}_i^2.
\]

Here, \(\tilde{\nu}_i\) is a vibrational frequency measured in wave numbers, \(L\) is the column eigenvector of internal displacement in the corresponding normal mode, and \(\mathbf{G}\) and \(\mathbf{F}\) are the inverse-mass and force constant matrices, respectively. The \(\mathbf{G}\) matrix relates to the kinetic energy by

\[
T = \frac{1}{2} \dot{\mathbf{L}}^T \mathbf{G}^{-1} \mathbf{L},
\]

and the \(\mathbf{F}\) matrix relates to the potential energy by

\[
V = \frac{1}{2} \mathbf{L}^T \mathbf{F} \mathbf{L}.
\]

### 3.1.1 IR and Raman Intensities

The internal oscillations of the atoms in a molecule (vibrations) induces a periodic change in either the molecular dipole moment, \(\vec{\mu}\), or polarizability tensor, \(\alpha\). The intensity of a vibrational mode in IR or Raman spectroscopy is proportional to a change in the dipole moment or polarizability tensor, respectively. For molecules which possess elements of symmetry, vibrations may change the charge distribution of the molecule but not the dipole moment and are thus IR inactive and Raman active. Thus, these two variations of vibrational spectroscopy complement each other because they follow different selection rules. Both spectroscopic methods can be used to study the vibration of molecules, and to determine
the identities of molecules by their characteristic group frequencies.

The most important relation in IR spectroscopy is the dependency of the IR intensity on the change of dipole moment ($\vec{\mu}$) during a vibration along a normal mode ($Q_k$) according to

$$I_{IR}^k \propto \left( \frac{\partial \vec{\mu}}{\partial Q_k} \right)^2_0. \quad (3.10)$$

On the other hand, Raman spectroscopy is based upon the Raman effect, whereby a molecule experiences a nonresonant oscillating electric field, and is caused to oscillate at the same frequency, radiating light in all directions. The electric field distorts the molecule and induces a time-dependent dipole moment ($\mu_{\text{induced}}(t)$). The induced dipole moment is proportional to the magnitude of the electric field:

$$\mu_{\text{induced}}(t) = \alpha E_0 \cos 2\pi \nu t, \quad (3.11)$$

where $\alpha$ is the polarizability. For vibrational modes to be Raman active, the polarizability of the molecule must change as it vibrates. Thus, Raman intensities depend on the change of the polarizability ($\alpha$) during a vibration along a normal mode ($Q_k$) according to

$$I_{\text{Raman}}^k \propto \left( \frac{\partial \alpha}{\partial Q_k} \right)^2_0. \quad (3.12)$$

In contrast to infrared spectroscopy, Raman is a scattering technique. If the light is scattered without a change in frequency, this is known as Rayleigh scattering. A small fraction of the scattered light will have a frequency different than that of the exciting light, $\nu_0$, such that energy is absorbed by the sample by $\Delta E = h|\nu_0 - \nu_k|$. If $\nu_k < \nu_0$ this is known as Stokes scattering whereas if $\nu_k > \nu_0$ this is known as anti-Stokes scattering (see Figure 3.1).

Another important variation of Raman spectroscopy with a wide applicability to biological systems is resonance Raman spectroscopy. Resonance Raman occurs when the exciting
light is resonant with an electronic absorption band of the sample, and the intermediate state is a real rather than virtual state. The intensities of the Raman-active normal modes as a function of excitation wavelength provide insight into the excited electronic state (Figure 3.2). Most vibrational modes which are strong in the resonance Raman spectrum derive their intensities from a geometry change of the molecule in the excited electronic state. This displacement in the ground and excited state potential energies is also responsible for the contribution of certain vibrational modes (said to be “Franck-Condon active”) to the electronic absorption and emission spectra. The intensity of a vibrational transition in resonance Raman is proportional to the square of the excited state gradient; vibrational modes which are parallel to the gradient are expected to be more intense whereas vibrational modes that are perpendicular to the gradient are expected to be less intense. Thus, resonance Raman can reveal the structure of the molecule in its excited electronic state.

![Figure 3.1: Absorption and scattering for Infrared and Raman spectroscopy.](image)

### 3.1.2 Calculation of Vibrational Frequencies

Vibrational spectra of molecules can be calculated in their ground and excited states. In addition to predicting frequencies and intensities of spectral lines, the calculations can also describe the deformations that the system undergoes in its normal modes. In other words, one can calculate the direction and magnitude of the nuclear displacement that occurs when
Figure 3.2: The difference between non-resonance Raman and resonance Raman scattering. Non-resonance Raman scattering occurs when the radiation excites the electrons to an unstable virtual state, and the radiation is emitted (scattered) at a slightly different energy than the incident radiation. For resonance Raman, the incident radiation is at a frequency near the frequency of an electronic transition of the molecule of interest. This provides enough energy to excite the electrons to a higher electronic state.

a system absorbs a quantum of energy.

Vibrational frequencies are taken directly from the second derivatives of the energy with the two perturbations taken as nuclear movements. Due to a combination of electron correlation effects and basis set deficiencies, calculated harmonic force constants and frequencies are usually higher than the experimental values. These discrepancies are usually scaled by an empirical factor that varies by basis set. For a more accurate prediction and assignment of vibrational spectral modes, frequencies can be refined by the scaled quantum-mechanical (SQM) formula.\textsuperscript{16,17} In this procedure, the force-constant matrix for the system is defined in a non-redundant set of Pulay’s internal coordinates\textsuperscript{16,17} and scaled by a set of factors. The total force field is scaled according to

\[
F_{ij}^{scaled} = \sqrt{c_i c_j} F_{ij},
\]

where \(F_{ij}\) is an element of a force constant matrix, and \(c_i\) and \(c_j\) are scale factors associated
with internal coordinate $i$ and $j$, respectively. The use of the SQM formula has been shown to produce results that compare to experimental data within $12 \text{ cm}^{-1}$ or less.$^{16-23}$

### 3.2 EPR Spectroscopy

Electron paramagnetic resonance (EPR) is used to study open shell systems, and thus applies to compounds of the various transition metals and to free radicals. Unpaired electrons occur in the $d$-orbitals of transition-metal ions, or in the uppermost molecular orbitals of radicals. In either case, there is a series of excited electronic states with energies spread over large ranges, $5000 - 40,000 \text{ cm}^{-1}$. Transitions between these levels are the principle concern of UV or visible spectroscopy. However, in EPR spectroscopy, one static magnetic field and one oscillating magnetic field, perpendicular to each other, are applied to the molecules of interest.

Due to the angular momentum, both electrons and magnetic nuclei possess a magnetic moment, $\mu$, which will interact with a magnetic field. Depending on the orientation of the magnetic moment relative to the magnetic field, the energy of the system is increased or decreased. A free electron can be found in one of two spin states $|\alpha\rangle$ or $|\beta\rangle$. In the absence of a magnetic field, the two spin states of the electron are degenerate. Introducing a static magnetic field removes the degeneracy of the system and splits the two spin states by

$$\Delta E = g \mu B$$  \hspace{1cm} (3.14)

where the value of $g$ depends on the identity of the particle, and $\mu$ is its magnetic moment. The oscillating magnetic field induces transitions between these two energy levels. A transition between the two levels, also referred to as Zeeman levels, involves a change in the orientation of the electron spin.

For a free electron, $g$ is $2.002319$. Measurement of EPR spectra depends on determining the $g$-values for the unpaired electrons in a molecule, which differ from those of free electrons
and are sensitive to the chemical environment of the electrons. A marker with a known \( g \)-value is often used as a reference. The \( g \)-value is calculated from the recorded value of the magnetic field at the resonance position:

\[
g_{\text{sample}} = g_{\text{std}} \frac{B_{\text{std}}}{B_{\text{sample}}}.\]

In chemical systems, the unpaired electron occupies an orbital which may be more or less localized on a single atom or may be heavily delocalized across a molecule. The \( g \)-value reflects the nature of this orbital. In a simple free-radical, the \( g \)-value remains close to that of the free-electron value. In other cases, considerable contribution to the magnetic moment of the electron from its orbital motion may alter the \( g \)-value. This interaction is known as spin-orbit coupling, and is most common in systems containing transition metal ions.

When a substance containing unpaired electrons is placed in a magnetic field, the motion of the electrons is affected so that the magnetic moment opposes the applied field. The simplest way to describe this effect is in terms of the system gaining some of the characteristics of excited states. Usually, the excited states have an orbital momentum which gets mixed in to the ground state. The extent of this mixing is inversely related to the energy of the excited state.

\[
g = 2.002319(1 - \frac{f\lambda}{\Delta E})
\]

where \( \lambda \) is the spin-orbit coupling constant, \( f \) is the composite numerical factor, and \( \Delta E \) is the energy of the state which is being mixed in. Note that \( \lambda \) can be positive or negative, so that \( g \) can be greater or less than 2.002319.

### 3.3 UV-Vis Spectroscopy

Energy radiation in the ultraviolet (200 - 400nm) or visible region (400 - 700nm) can cause molecules to undergo electronic transitions by promoting an electron from a lower energy to
a higher energy molecular orbital. Generally, the electron promotion will be from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), resulting in an excited state. When the molecule is exposed to light having an energy that matches a possible electronic transition within the molecule, some of the light energy will be absorbed as the electron is promoted to a higher energy orbital. Common transitions that occur are: $\sigma \rightarrow \sigma^*$, $n \rightarrow \sigma^*$, $n \rightarrow \pi^*$, $\pi \rightarrow \pi^*$, and $d \rightarrow d$. The $\sigma \rightarrow \sigma^*$ transition occurs when a bonding $\sigma$ orbital is excited to the corresponding antibonding orbital. The energy required is large, and is usually not detectable in the UV-vis spectrum. The $n \rightarrow \sigma^*$ transition occurs between a non-bonding electron and the antibonding $\sigma$ orbital. These transitions usually need less energy than $\sigma \rightarrow \sigma^*$ transition. They can be initiated by light whose wavelength is in the range 150 - 250 nm. The majority of absorption spectroscopy of organic compounds and amino acids is based on transitions of $n$ or $\pi$ electrons to the $\pi^*$ excited state. This is because the absorption peaks for these transitions fall in an experimentally convenient region of the spectrum (200 - 700 nm).

The solvent in which the absorbing species is dissolved also has an effect on the spectrum of the species. Peaks resulting from $n \rightarrow \pi^*$ transitions are shifted to shorter wavelengths (blue shift) with increasing solvent polarity. This arises from increased solvation of the lone pair, which lowers the energy of the $n$ orbital. Often the reverse (i.e. red shift) is seen for $\pi \rightarrow \pi^*$ transitions. This is caused by attractive polarization forces between the solvent and the absorber, which lowers the energy levels of both the excited and unexcited states. This effect is greater for the excited state, and so the energy difference between the excited and unexcited states is slightly reduced - resulting in a small red shift. This effect also influences $n \rightarrow \pi^*$ transitions but is overshadowed by the blue shift resulting from solvation of lone pairs.

For inorganic molecules, three types of transitions are important to consider: metal to ligand charge transfer (MLCT), ligand to metal charge transfer (LMCT), and $d \rightarrow d$ transitions. MLCT transitions arise from $\pi$ acceptor ligands and metals that are willing to
donate electrons into the orbitals of ligand character. This type of transition is not common, as metals tend to accept electrons rather than donate them. LMCT transitions involve $\pi$ donor ligands around the metal, and can be considered as a transition from orbitals that are ligand in character to orbitals that are metal in character; for example, an electron travels from a bonding $\pi$ or non-bonding $\pi$ orbital into a $\sigma^*$ orbital. These transitions are very strong and appear very intensely in the absorbance spectrum. The $d \rightarrow d$ transitions correspond to a transition from a metal $d$-orbital to another metal $d$-orbital. According to the Laporte rules, electronic transitions that conserve parity, either symmetry or antisymmetry with respect to an inversion center (i.e., $g \rightarrow g$ or $u \rightarrow u$), are forbidden. Thus, the $d \rightarrow d$ transition ($g \rightarrow g$) is “symmetry” forbidden. However, due to vibronic coupling, the $d \rightarrow d$ transitions can be weakly allowed. Because of their relatively low energy of transition, they can emit visible light upon relaxation which is why many transition metal complexes are brightly colored.

In a UV-vis measurement, an optical spectrometer records the wavelengths at which absorption occurs, and provides information about the electronic transitions occurring in the material. The Beer-Lambert law states that the fraction of light measured as transmittance or reflectance ($I$) versus the incident intensity ($I_0$) is dependent on the path length of light through the sample ($l$), the absorption cross section of the transition ($\sigma$), and the difference in the population of the initial state ($N_1$) and final state ($N_2$) of the initial ($E_1$) and final ($E_2$) electronic energy levels,

$$\frac{I}{I_0} = e^{-\sigma(N_1-N_2)}. \quad (3.15)$$

This equation is often written in a form known as Beer’s law:

$$A = \varepsilon cl = -\log_{10}\left(\frac{I}{I_0}\right), \quad (3.16)$$

where $A$ is the absorbance, $\varepsilon$ is the molar absorptivity coefficient of the material, $c$ is the concentration, and $l$ is the path length of light through the sample.
The calculation of electronic transitions can be achieved by the use of a configuration interaction method (CIS, CISD, CISDT) or TD-DFT. A CI calculation uses the optimized molecular orbitals for the ground state and calculates the energy to move one electron to one of the virtual orbitals. The TD-DFT obtains the wave functions of molecular orbitals that oscillate between the ground state and the first excited states.

3.4 X-ray Absorption Spectroscopy

X-ray absorption spectroscopy is a widely used technique for determining the structure of matter. This form of spectroscopy allows the determination of the chemical environment of a single element in terms of the number and type of its neighbors, interatomic distances and structural disorders. It is a powerful structural local probe, and is particularly useful for the study of biological molecules. The probability that x-ray radiation will be absorbed is given by Beer’s law:

\[ I = I_0 e^{-\mu t}, \]  

where \( I_0 \) is the x-ray intensity incident on a sample, \( t \) is the sample thickness, and \( I \) is the intensity transmitted through the sample. The absorption coefficient (\( \mu \)) is dependent on the sample density (\( \rho \)), the atomic number (\( Z \)), atomic mass (\( A \)), and x-ray energy (\( E \)):

\[ \mu \approx \frac{\rho Z^4}{AE^3}. \]  

When the incident x-ray has an energy equal to the binding energy of a core electron, there is a sharp rise in absorption, which is characteristic of a core electron being ejected as a photoelectron. It shows several discontinuities known as absorption edges which are classified with capital letters (K, L, M...) according to the principal quantum number of the electron in the ground state (\( n = 1, 2, 3, \ldots \)).

The x-ray absorption spectrum consists of two parts: x-ray absorption near-edge spec-
troscopy (XANES) and extended x-ray absorption fine-structure spectroscopy (EXAFS). The XANES spectrum is sensitive to the formal oxidation states and coordination chemistry of the absorbing atom, whereas the EXAFS spectroscopy measures the distances, coordination number, and species of neighboring atoms.

The EXAFS spectrum is concerned with oscillations above the absorption edge, and the EXAFS fine structure function \( \chi(E) \) is defined as:

\[
\chi(E) = \frac{\mu(E)\mu_0(E)}{\Delta\mu_0(E)},
\]  

(3.19)

where \( \mu(E) \) is the measured absorption coefficient, \( \mu_0(E) \) represents the absorption of an isolated atom, and \( \Delta\mu_0(E) \) is the rise in absorption at the threshold energy \( E_0 \).

The ejected photoelectron has the characteristics of both a particle and a wave. It is common practice to convert the x-ray energy to the wave number of the photoelectron \( k \):

\[
k = \sqrt{\frac{2m(E - E_0)}{\hbar^2}}
\]  

(3.20)

where \( E_0 \) is the absorption edge energy and \( m \) is the electron mass. Thus, the oscillations are measured as a function of the photoelectron wave number, \( \chi(k) \).

The different frequencies that occur in the oscillations of \( \chi(k) \) correspond to different near-neighbor coordination shells which can be described according to the EXAFS equation:

\[
\chi(k) = \sum_j N_j f_j(k)e^{2k^2\sigma_j^2} \frac{kR_j^2}{kR_j^2} \sin[2kR_j + \delta_j(k)],
\]  

(3.21)

where \( f(k) \) and \( \delta(k) \) are scattering properties of the atoms neighboring the excited atom, \( N \) is the number of neighboring atoms, \( R \) is the distance to the neighboring atom, and \( \sigma^2 \) is the disorder in the neighbor distance. The EXAFS equation allows the determination each of these quantities, as well as the atomic species of the neighboring atom.

The combination of X-ray absorption spectroscopy and DFT structure calculations is
highly useful for investigating the structural and chemical properties of biological molecules. The structural information derived from the EXAFS can be used to inform and guide DFT structure calculations. In turn, DFT calculations can be used to provide 3D structural information of EXAFS curve-fitting as well as thermodynamic data of the system.
Chapter 4

Siderophore-Metal Complexes

Look deep into nature, and then you will understand everything better.

– Albert Einstein

Density functional theory is a powerful tool that can approximate the structural, binding and spectroscopic properties of molecules. Recent advancements in these techniques have made them suitable for calculations of ground state potential surfaces of transition metal complexes. Computational support of experimental data has become a key element to reveal and understand the functional role of metal sites in proteins and small molecules.

As a preliminary investigation of the ability of density functional theory to accurately depict complex metal binding sites, a comprehensive computational study of aqueous desferrioxamine B (DFOB) metal complexes supported by extended X-ray absorption fine structure (EXAFS) and infrared spectroscopic data was conducted. Structural parameters derived from EXAFS analyses and trends in metal binding constants were well reproduced and validated by the applied computational model, and revealed how structure and vibrational spectra of aqueous desferrioxamine B (DFOB) metal complexes vary with the metal ion identity.
CHAPTER 4. SIDEROPHORE-METAL COMPLEXES

4.1 Desferrioxamine B Metal Complexes

Transition metals are required by many organisms to maintain normal metabolic function. Iron is the most abundant transition metal in the Earth’s crust, and typically has both the highest concentration in soils and sediments as well as the highest intracellular quota of the transition metals. However, in most aerobic aqueous environments, the bioavailability of iron is limited by the low solubility of iron (hydr)oxide minerals. Similarly, in most mammalian biological fluids, the concentration of free iron is tightly regulated to avoid unwanted chemical reactions and opportunistic infections. To facilitate acquisition of iron, organisms including bacteria, fungi, and plants, have developed specific strategies that increase its availability to cells. The exudation of siderophores, generally tetra- or hexadentate chelating agents that form exceptionally strong complexes (log $K > 20$) with Fe(III) via hard Lewis base-bearing functional groups, is a biological response to low iron availability that is widely distributed throughout the microbial tree of life. The most widely studied microbial siderophore is desferrioxamine B (DFOB), which is often used as a convenient model to represent the properties of a common class of siderophores, the trishydroxamates.

Recent work has suggested that siderophores, once solely thought of as iron transport agents, may be involved in the solubilization, binding, and biological uptake of metals other than iron, and that the presence of high concentrations of these metals (e.g., Mn) may interfere with siderophore-mediated iron transport. Although biologically relevant metal-hydroxamate complexes have been extensively studied, there are fewer detailed structural characterizations of siderophore complexes with metals other than iron. Investigating the structure and stability of the resulting metal-siderophore complexes is critical to understanding how metals interact with siderophores. Additionally, structural factors play key roles in regulating the recognition and uptake of the metal at cell membranes.

Modern spectroscopic approaches allow for the study of the molecular and electronic structure of aqueous complexes in solution. In particular, extended X-ray absorption fine structure (EXAFS) spectroscopy is well suited to understanding the molecular structure of
complexes because it probes the local coordination environment of a specific element. X-ray spectroscopic approaches have been used previously to study the speciation and structure of siderophore-metal complexes. Similarly, attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectroscopy provides information about bond energies that are especially complementary to EXAFS observations. These approaches, when used in conjunction with computational techniques, have provided specific insights into the structural factors that control metal-siderophore complex speciation and stability.

A systematic quantum-mechanical study was applied to assess the structural, binding, and vibrational properties that can be reproduced, revealed, and established for aqueous DFOB complexes using modern density functional methods. The calculations predict and complement experimental data for aqueous DFOB complexes with environmentally and biologically important metals (Scheme 4.1), and investigate spectral changes in the core and outer coordination shells associated with the size, spin, and oxidation state of the bound metal ion.

4.2 Structure

Structures were computed at the level of B3LYP density functional in combination with moderate 6-31G* basis sets as implemented in the Gaussian 09 program package. The 6-31G* basis sets are not available for structures with the heavier metal ions Cd(II), Hg(II) and Pb(II). In these cases, the cc-pVDZ-PP relativistic effective core potentials were employed in combination with 6-31G* for the remaining atoms. Relativistic core potentials were obtained from the Basis Set Exchange Library. An aqueous environment was modeled by the polarizable continuum model (PCM) that hosts a computed structure in a polarizable fixed van der Waals cavity to represent the solute-solvent boundary.

Computed structures of aqueous metal-HDFOB complexes are compared with interatomic distances derived from EXAFS spectroscopy in Table 4.1, Table 4.2, Table 4.3, and Figure 4.1. The structural data show the average size of the metal’s three coordination shells (e.g., O atoms bound directly to the metal, N or C atoms proximal to the metal center in the hydroxamate groups (A, B and C), and the two sequential carbon atoms (alpha and beta) in the molecule’s backbone) (Scheme 4.2).

Scheme 4.2: Calculated and experimental (EXAFS) average M-O distances [Å] for A) First Coordination Shell, B) Second Coordination Shell, and C) Third Coordination Shell of metal-HDFOB complexes. Standard deviations (calc.) and Debye-Waller factors (expt.) are indicated as horizontal error bars. Coordination shell atoms are shown in red.
Table 4.1: Calculated and experimental (EXAFS) average interatomic distances [Å], standard deviations (STDV: calc.) and Debye-Waller factors (DWF: expt.) for the first coordination shell.

<table>
<thead>
<tr>
<th>Complex</th>
<th>1st Shell</th>
<th>R_0 (STDV/DWF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal</td>
<td>d^n</td>
<td>2S + 1</td>
</tr>
<tr>
<td>V^{2+}</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Cr^{2+}</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Mn^{2+}</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Fe^{2+}</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Co^{2+}</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>^1Ni^{3+}</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>^3Ni^{3+}</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Cu^{2+}</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Zn^{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Cd^{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Hg^{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Pb^{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Sc^{3+}</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ti^{3+}</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>V^{3+}</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>^4V^{4+}</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>^5V^{5+}</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cr^{3+}</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Mn^{3+}</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Fe^{3+}</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Co^{3+}</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Ga^{3+}</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

^a An axial distance.
^b An equatorial distance.
Table 4.2: Calculated and experimental (EXAFS) average\textsuperscript{43,66,69,72} interatomic distances [Å], standard deviations (STDV: calc.) and Debye-Waller factors (DWF: expt.) for the second coordination shell.

<table>
<thead>
<tr>
<th>Complex</th>
<th>R\textsubscript{C/N} (STDV/DWF)</th>
<th>2nd Shell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal</td>
<td>d\textsuperscript{a}</td>
<td>2S + 1</td>
</tr>
<tr>
<td>V\textsuperscript{2+}</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Cr\textsuperscript{2+}</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Mn\textsuperscript{2+}</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Fe\textsuperscript{2+}</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Co\textsuperscript{2+}</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>\textsuperscript{1}Ni\textsuperscript{2+}</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>\textsuperscript{3}Ni\textsuperscript{2+}</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Cu\textsuperscript{2+}</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Zn\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Cd\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Hg\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Pb\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Sc\textsuperscript{3+}</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ti\textsuperscript{3+}</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>V\textsuperscript{3+}</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>V\textsuperscript{4+}</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>V\textsuperscript{5+}</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cr\textsuperscript{3+}</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Mn\textsuperscript{3+}</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Fe\textsuperscript{3+}</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Co\textsuperscript{3+}</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Ga\textsuperscript{3+}</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

\textsuperscript{a} An axial distance.
\textsuperscript{b} An equatorial distance.
Table 4.3: Calculated and experimental (EXAFS) average\textsuperscript{43,66,69,72} interatomic distances [Å], standard deviations (STDV: calc.) and Debye-Waller factors (DWF: expt.) for the third coordination shell.

<table>
<thead>
<tr>
<th>Complex</th>
<th>Metal</th>
<th>$d^n$</th>
<th>$2S + 1$</th>
<th>$R_{\alpha\beta}$ (STDV/DWF)</th>
<th>Calc. Å</th>
<th>Exp. Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>V\textsuperscript{2+}</td>
<td>3</td>
<td>4</td>
<td>4.604 (0.357)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr\textsuperscript{2+}</td>
<td>4</td>
<td>5</td>
<td>4.677 (0.434)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn\textsuperscript{2+}</td>
<td>5</td>
<td>6</td>
<td>4.663 (0.352)</td>
<td>4.69 (0.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fe\textsuperscript{2+}</td>
<td>6</td>
<td>5</td>
<td>4.608 (0.354)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co\textsuperscript{2+}</td>
<td>7</td>
<td>4</td>
<td>4.569 (0.359)</td>
<td>4.57 (0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{1}\text{Ni}\textsuperscript{2+}$</td>
<td>8</td>
<td>1</td>
<td>4.578 (0.511)</td>
<td>4.42 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{3}\text{Ni}\textsuperscript{2+}$</td>
<td>8</td>
<td>3</td>
<td>4.529 (0.364)</td>
<td>4.42 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu\textsuperscript{2+}</td>
<td>9</td>
<td>2</td>
<td>4.578 (0.411)</td>
<td>4.28 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zn\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
<td>4.571 (0.359)</td>
<td>4.48 (0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cd\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
<td>4.787 (0.343)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hg\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
<td>4.890 (0.352)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pb\textsuperscript{2+}</td>
<td>10</td>
<td>1</td>
<td>5.045 (0.366)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sc\textsuperscript{3+}</td>
<td>0</td>
<td>1</td>
<td>4.614 (0.337)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ti\textsuperscript{3+}</td>
<td>1</td>
<td>2</td>
<td>4.552 (0.349)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V\textsuperscript{3+}</td>
<td>2</td>
<td>3</td>
<td>4.498 (0.352)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V\textsuperscript{4+}</td>
<td>1</td>
<td>2</td>
<td>4.452 (0.346)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V\textsuperscript{5+}</td>
<td>0</td>
<td>1</td>
<td>4.433 (0.353)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr\textsuperscript{3+}</td>
<td>3</td>
<td>4</td>
<td>4.467 (0.359)</td>
<td>4.23 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn\textsuperscript{3+}</td>
<td>4</td>
<td>5</td>
<td>4.513 (0.398)</td>
<td>4.35 (0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fe\textsuperscript{3+}</td>
<td>5</td>
<td>6</td>
<td>4.509 (0.351)</td>
<td>4.39 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co\textsuperscript{3+}</td>
<td>6</td>
<td>1</td>
<td>4.367 (0.369)</td>
<td>4.19 (0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ga\textsuperscript{3+}</td>
<td>10</td>
<td>1</td>
<td>4.440 (0.353)</td>
<td>4.47 (0.05)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1: Calculated and experimental (EXAFS) average distances [Å] of the first, second and third coordination shells for metal-HDFOB complexes versus metal ionic radii [Å]. Linear correlations for the calculated values; 
\[ y = 1.01x + 1.35, \quad R^2 = 0.9468 \] (first coordination shell), \[ y = 0.968x + 2.18, \quad R^2 = 0.9188 \] (second coordination shell), \[ y = 0.966x + 3.87, \quad R^2 = 0.9459 \] (third coordination shell).

The average distances of the metal to the three proximal coordination shells trend with published crystallographic radii, which highlights the general agreement between calculated structures and experimental interatomic distances derived from EXAFS measurements (Figure 4.1 and Scheme 4.2). The first coordination shell of the complex is defined by the metal atom bound to six oxygen atoms in the hydroxamate moieties of HDFOB\(^{2-}\) (Scheme 4.1). The properties of these metal-oxygen bonds have been shown to be directly related to the strength and stability of the metal core, which are key chemical properties of the complex. As is expected from changes in ionic radii, the more highly-charged metal ions exhibit a smaller inner coordination shell and hence a stronger, more stable core. This trend is best illustrated in the vanadium ion series (Figure 4.2), where the average core size of the V(II) ion computed at 2.13 Å decreases to 2.01 Å for the V(III) ion, to 1.93 Å for V(IV)HDFOB\(^{2+}\) and to 1.89 Å for V(V)HDFOB\(^{3+}\). These highly charged vanadium complexes have been detected and studied in very acidic media.
For metal ions with the same charge, the average metal-oxygen distance in metal-HDFOB complexes is driven by the \( d \)-electron configuration. Among the series of M(II) first-row transition metal ions, the weakest core is found for Cr(II)HDFOB (calc. 2.20 Å) and Mn(II) HDFOB (calc. 2.19 Å and expt. 2.16 Å), whereas a triplet-Ni(II)HDFOB octahedral complex is predicted to have the strongest core (calc. 2.06 Å). The Ni(II)HDFOB complex is stabilized because its \( d^8 \) triplet-electronic structure occupies three binding \( d \)-orbitals. The metal core distance is larger for Mn(II)HDFOB because of its high-spin \( d^5 \) configuration, with two electrons in anti-bonding \( d \)-orbitals. The average core size for the high-spin \( d^4 \) configuration of Cr(II)HDFOB was found to be quite comparable to that of Mn(II)HDFOB, possibly because strong Jahn-Teller distortion of the axial ligands weakens the Cr(II)HDFOB core.

For complexes containing M(III) first-row transition metal ions, the average core distance is shorter than the M(II)DFOB complexes, and is also driven by the \( d \)-electron configuration. Among M(III) complexes, the smallest core (calc. 1.89 Å and expt. 1.88 Å) is found in the Co(III)HDFOB\(^+\) complex because of its low-spin \( d^6 \) singlet electronic structure. The core distance increases and reaches its maximum for the Fe(III)HDFOB\(^+\) complex (calc. 2.02 Å
and expt. 2.02 Å) with an electronic configuration equivalent to Mn(II). The comparison excludes the Ti(III) and Sc(III) complexes, for which no d-electrons are present and the core is even larger (2.05 Å and 2.12 Å, respectively). As expected, the HDFOB complexes with heavy metals Cd(II), Hg(II), and Pb(II) have the largest core size due to their increased ionic radii.

Complexes with Mn(III), Cr(II), or Cu(II) demonstrate strong structural deformation associated with the Jahn-Teller effect, with four short equatorial metal-oxygen distances and two long axial metal-oxygen distances. The Jahn-Teller distortion is driven by the d-electron configuration that has an odd number of electrons in the anti-bonding \( e_g \) orbitals.\(^{82}\) The effect is expected for the Cr(II) HDFOB and Mn(III)HDFOB\(^+\) complexes with a high-spin \( d^4 \) configuration, for a high-spin \( d^7 \) configuration of the Co(II)HDFOB complex, and for a \( d^9 \) configuration of the Cu(II)HDFOB complex. The \( d^7 \) configuration of Co(II)HDFOB, although Jahn-Teller-active, produces minimal elongation of the axial distances since the extra electron occupying \( d\pi \) orbital is only weakly anti-bonding.\(^{43}\) An interesting case is the diamagnetic singlet-Ni(II)HDFOB complex, with a \( d^8 \) electron configuration that adopts a square planar geometry by releasing the axial ligands. This large structural change between a square planar singlet-Ni(II)HDFOB complex and an octahedral triplet-Ni(II)HDFOB complex clearly illustrates a strong correlation between d-orbital occupation and structure of the metal core. Calculated structure, energy, and spectroscopy of both singlet and triplet Ni(II)HDFOB complexes are reported here to demonstrate these electronic effects. Comparison of the computed Ni(II) core with experimental results supports the triplet-Ni(II) state for the Ni(II)HDFOB complex in aqueous solution. The experimental core size of a Ni(II)HDFOB complex (2.04 Å) agrees well with the computed average found for an octahedral triplet-Ni(II) complex (2.06 Å)\(^{72}\), and it disagrees with the square planar geometry of the singlet-Ni(II) complex (2.13 Å).

Plotting the axial bond distances against equatorial distances allows for visualization of complex distortion. For the first coordination shell (Figure 4.3), the Jahn-Teller distorted
complexes exhibit axial bonds that are significantly longer than equatorial bonds. Additional distorted complexes are those that contain the toxic Hg(II) and Pb(II) ions, which display specific coordination preferences due to increased relativistic effects.\textsuperscript{69} The Pb(II) ion is known to display hemidirected coordination preferences to accommodate the stereochemically active orbital. To accommodate the orbital, the computed Pb(II)HDFOB complex adopts an asymmetric tri-coordination where among two distinct types of oxygen atoms (viz. C=O and N–O\textsuperscript{−}), only the N–O\textsuperscript{−} atoms are bound to the ion (Figure 4.4). This unusual Pb(II) coordination mode results in a bulky core with a very high standard deviation (see Figure 4.3A, Table 4.1), consistent with the complex coordination environment found in X-ray spectroscopic investigations of the Pb-DFOB system.\textsuperscript{69}

Computed coordination distances are slightly longer than the experimental values (Table 4.1 and Figure 4.3). For example, the computed metal-oxygen distances are approximately 0.05 Å longer than the experimentally-determined values for octahedral HDFOB complexes with Co(II), Mn(II), Ni(II), Zn(II), Cd(II) and Cr(III). The two experimentally observed M-O bond distances of the Jahn-Teller distorted Mn(III) HDFOB\textsuperscript{+} complex compare very well with the computed axial and equatorial bond averages. However, the computed distances for the Cu(II)HDFOB complex are approximately 0.05 Å longer than experimentally observed values. Nevertheless, these small structural deviations are consistent with previous DFT-reports\textsuperscript{43,73} and commonly result from basis set truncation and electron correlation errors.

Computed structures show that the terminal –NH\textsubscript{3}\textsuperscript{+} group forms hydrogen bonds to two N–O\textsuperscript{−} type oxygen atoms from the first coordination sphere in M(II)HDFOB, one in an axial and one in an equatorial position. For these complexes, the average O⋯HN distance is about 1.69 Å for both axial and equatorial oxygen atoms. However, as the oxidation state of the bound metal increases, the hydrogen bonding of NH\textsubscript{3}\textsuperscript{+} becomes weaker, likely due to increased Coulomb repulsion by the metal core. Consequently, in all of the structures of M(III)HDFOB\textsuperscript{+} complexes, the hydrogen bond interaction remains only for an equatorial
Figure 4.3: Calculated first-coordination-shell distances vs. experimental distances [Å] for M(II) metal (A) and for M(III), V(IV) and V(V) metal complexes (B). Error bars indicate calculated standard deviations for axial (vertical bars) and equatorial (horizontal bars) distances.
Figure 4.4: Calculated structure of the Pb(II)-HDFOB complex.

oxygen. The average O···HN distance elongates to 1.75 Å for equatorial oxygens and breaks up to a non-interacting distance of 2.23 Å for axial oxygens. For complexes with higher oxidation state metal ions, such as V(V), the NH$_3^+$ group is no longer H-bonded (O···HN distance about 4.5 Å), and the terminal chain is more exposed to the solvent. Structural fluctuation of the amine tail upon its deprotonation or interaction with solvent was previously reported in a computational study of the Fe(III)HDFOB$^+$ complex.$^83$ It has been proposed that the protonated pendant amine tail may play a role in the cellular recognition process based on crystallographic studies of ionophore-siderophore complexes, making the position of the protonated pendant amine tail in M-HDFOB complexes potentially important.$^{84,85}$ However, to our knowledge, the role of the interaction of the amine tail in the structure and stability of MHDFOB complexes has not been considered.

### 4.3 Binding Energy

The computational determination of an accurate log $K$ value for metal binding in a metal-HDFOB complex is very demanding, as an error in the $\Delta G$ value of only 1.36 kcal/mol is equivalent to a unit difference in the log $K$ value at room temperature. Various protonation states and conformational flexibility of the ligand as well as entropic effects greatly affect the binding energy of a complex. In general, these contributions could be as large as hundreds of kcal/mol and are difficult to account for in electronic structure calculations.$^{86-92}$ Computed
binding energies endure systematic errors, such as basis set truncation and electron correlation errors. For transition metal systems, the average mean absolute error with B3LYP density functional has been reported to be around 12 kcal/mol. At room temperature, for pK or pH estimation in acid-base chemistry, the value is equivalent to an unacceptably large error of 8.8 units. Furthermore, for calculations of aqueous complexes, large conformational flexibility, variable protonation states and various other entropic factors, inadequate estimation of solvent-reactants, and solvent-solvent interaction could significantly contribute to the accumulated error. It has been demonstrated that ion hydration models proposed on the basis of the long-range and short-range electrostatic interactions do not agree with experimental data. In particular, for transition metal ions, significant contributions of donor-acceptor interactions to the short-range interaction are neglected. Additionally, the contributions of the ligand-field stabilization in the hydration of a transition-metal ion are not considered, despite their recognized importance. For instance, in the reported DFT studies of hydration enthalpies of aqueous transition metal ions, an empirical correction of $n\Delta H_{vap}$ was applied to achieve reasonable accuracy with experimental data. The correction is considerable since the empirical vaporization energy ($\Delta H_{vap} = -10.67$ kcal/mol) has to be scaled by the number of water molecules that strongly interact with an aqueous ion. Most transition metal ions require between $n = 6$ and 18 water molecules, including interactions with the second coordination shell. Hence, for the chemistry of aqueous transition metals, a substantial, though justified, empirical correction is necessary to attain the desired accuracy.

A commonly applied solution to this problem is to approximate the reliable empirical log $K$ values by a linear function of the computed log $K$ values, where fitted parameters (the slope and intercept of the linear equation) can mitigate systematic errors of computational modeling and average conformational and entropic effects. This approach was employed to compute log $K$ values for the estimation of relative differences and tendencies in binding energy between various metals and oxidation states. The binding energy in the aqueous environment was estimated by the computed PCM energy difference between the complex
and the reactant fully protonated solitary ligand (H$_4$DFOB$^+$) and the metal ion. Additionally, the value has been corrected by thermal vibrational corrections (at 298 K) and by the zero-point energy (ZPE) contribution from the complex and the ligand.

The binding stability constants ($\log K_{Calc}$) for metal complexes were computed according to the following chemical reaction:

$$\text{M}^{n+}(\text{aq}) + \text{H}_4\text{DFOB}^+ \rightleftharpoons \text{MHDFOB}^{(n-2)+} + 3\text{H}^+(\text{aq}),$$

where the experimental value of a proton free hydration energy$^{100}$ was set at -262.18 kcal/mol, and computed PCM DFT energies for remaining species were employed. In order to assess the absolute error, the computed values for M(II)HDFOB complexes were compared to available experimental data.$^{43–45,47,101}$ On average, the computed metal(II) binding energies are overestimated by about 30 kcal/mol. The largest error is found for the Cu(II)HDFOB complex where the binding energy is overestimated by about 130 kcal/mol. Cd(II)HDFOB is the only complex in the set that underestimates the binding energy by about 27 kcal/mol (Table 4.4 and Figure 4.5). As pointed out previously, the generally observed overestimation of computed binding energy is likely due to several contributing factors: (1) the deficiency of the PCM model to reproduce the donor-acceptor and the ligand field interaction for metal ion-solvent; (2) large conformational flexibility of a free ligand, its various protonation states and its possible binding interaction with solvent; and (3) additional entropic factors.

A common approach to refine the energies of a series of computed compounds and to minimize any systematic errors is to apply a linear fit to a reliable set of experimental data to determine two empirical parameters (slope and intercept). We have used a set of experimentally derived constants$^{43–45,47,101}$ for metal(II) complexes to improve the computed data. The linear correlation (Figure 4.5) produced the empirical equation, $\log K_A = 0.0446 \log K_{Calc} + 9.5076$ which has been utilized to refine computed values (see Table 4.4).

The ion-solvent effects (i.e. donor-acceptor interactions, the ligand field stabilization, and entropy effects) are generally expected to be greater for M(III) ions than for M(II) ions.
Table 4.4: Calculated and experimental metal stability constants ($logK$)

<table>
<thead>
<tr>
<th>Metal</th>
<th>Complex $d^n$</th>
<th>2S + 1</th>
<th>Binding Energy</th>
<th>Calculated log $K$</th>
<th>Adjusted log $K$</th>
<th>Experiment log $K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{V}^{2+}$</td>
<td>3</td>
<td>4</td>
<td>41.94</td>
<td>11.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Cr}^{2+}$</td>
<td>4</td>
<td>5</td>
<td>30.79</td>
<td>10.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Mn}^{2+}$</td>
<td>5</td>
<td>6</td>
<td>11.19</td>
<td>10.01</td>
<td>7.68&lt;sup&gt;44&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Fe}^{2+}$</td>
<td>6</td>
<td>5</td>
<td>27.79</td>
<td>10.75</td>
<td>11.09&lt;sup&gt;102&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Co}^{2+}$</td>
<td>7</td>
<td>4</td>
<td>41.83</td>
<td>11.37</td>
<td>10.06&lt;sup&gt;43&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Ni}^{2+}$</td>
<td>8</td>
<td>1</td>
<td>27.48</td>
<td>10.73</td>
<td>11.78&lt;sup&gt;47&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Co}^{3+}$</td>
<td>8</td>
<td>3</td>
<td>48.21</td>
<td>11.66</td>
<td>11.78&lt;sup&gt;47&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Cu}^{2+}$</td>
<td>9</td>
<td>2</td>
<td>112.3</td>
<td>14.52</td>
<td>14.61&lt;sup&gt;45&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Zn}^{2+}$</td>
<td>10</td>
<td>1</td>
<td>35.12</td>
<td>11.07</td>
<td>11.95&lt;sup&gt;47&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Cd}^{2+}$</td>
<td>10</td>
<td>1</td>
<td>28.81</td>
<td>8.22</td>
<td>8.76&lt;sup&gt;47&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Hg}^{2+}$</td>
<td>10</td>
<td>1</td>
<td>45.27</td>
<td>7.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Pb}^{2+}$</td>
<td>10</td>
<td>1</td>
<td>17.02</td>
<td>10.27</td>
<td>10.88&lt;sup&gt;45&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Sc}^{3+}$</td>
<td>0</td>
<td>1</td>
<td>114.08</td>
<td>30.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Ti}^{3+}$</td>
<td>1</td>
<td>2</td>
<td>140.42</td>
<td>31.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{V}^{3+}$</td>
<td>2</td>
<td>3</td>
<td>163.57</td>
<td>32.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{V}^{4+}$</td>
<td>1</td>
<td>2</td>
<td>421.19</td>
<td>28.29</td>
<td>31.81&lt;sup&gt;81&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{V}^{5+}$</td>
<td>0</td>
<td>1</td>
<td>844.56</td>
<td>47.18</td>
<td>37.3&lt;sup&gt;81&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Cr}^{3+}$</td>
<td>3</td>
<td>4</td>
<td>193.43</td>
<td>33.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Mn}^{3+}$</td>
<td>4</td>
<td>5</td>
<td>157.27</td>
<td>32.22</td>
<td>29.88&lt;sup&gt;44&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Fe}^{3+}$</td>
<td>5</td>
<td>6</td>
<td>127.68</td>
<td>30.9</td>
<td>32.02&lt;sup&gt;47&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Co}^{3+}$</td>
<td>6</td>
<td>1</td>
<td>247.2</td>
<td>36.23</td>
<td>37.45&lt;sup&gt;43&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>$\text{Ga}^{3+}$</td>
<td>10</td>
<td>1</td>
<td>399.52</td>
<td>27.33</td>
<td>29.49&lt;sup&gt;101&lt;/sup&gt;</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.5: A linear fit of DFT computed ($\log K_{\text{Calc}}$) and experimental ($\log K_{\text{Expt}}$) stability constants based on available data$^{43-45,47,81,103}$ for metal(II)-HFOB complexes.
For instance, the experimentally determined entropy value of an aqueous Fe(III) ion at room temperature is $-63.3 \, \text{cal/K} \cdot \text{mol}$, over twice the value found for an aqueous Fe(II) ion $(-25.9 \, \text{cal/K} \cdot \text{mol})$. After the computed $\log K_{\text{Calc}}$ values were refined by the metal(II)-derived correlation, greater discrepancies were found for Mn(III)HDFOB$^+$ (expt. 29.88, calc. 16.52, an error of 13.36), Fe(III)HDFOB$^+$ (expt. 32.02, calc. 15.20, an error of 16.82), and Co(III)HDFOB$^+$ (expt. 37.45, calc. 20.53, an error of 16.92) complexes. Overall, these values were underestimated by a mean average of 15.70 (equivalent to $-21.42 \, \text{kcal/mol}$ or 71.85 $\text{cal/K} \cdot \text{mol}$ at room temperature). In order to account for the inherent energetic differences in the solvation of M(II) and M(III) ions, the intercept value was adjusted by +15.70 units while the original slope value of 0.0446 was kept constant. The empirical fit for M(III) ions is as follows: $\log K_A = 0.0446 \log K_{\text{Calc}} + 25.2076$. Both empirically fitted lines are shown in Figure 4.6. Both M(II) and M(III) sets of refined (log $K_A$) values agree well with experimental data as illustrated on the plot.

Higher oxidation metal complexes, represented by V(IV)HDFOB$^{2+}$ and V(V)HDFOB$^{3+}$ complexes did not exhibit oversized deviations from experimental values when the computed $\log K_{\text{Calc}}$ values were refined by the original empirical equations used for metals(II) (Figure 4.6). Therefore, for these two vanadium complexes reported the $\log K_A$’s are based on the metal(II) correlation fit.

Based on crystal field theory considerations, the $d$-orbital configuration of the metal ion is expected to contribute to the stability of octahedral complexes. The $d$-orbital stabilization energy should depend on the sequential occupation of three $t_{2g}$ bonding orbitals and two $e_g$ antibonding orbitals, which will stabilize and destabilize the complex, respectively. The effect arising from the electrostatic field predicts maxima at the $d^3$ and $d^8$ configurations, and minima for the spherically symmetrical ions, that is, $d^0$, high-spin $d^5$ and $d^{10}$ configurations. A plot of the experimental and computed stability constants for M(II) and M(III) metals as a function of $d$-orbital occupation is shown in Figure 4.7. Computed stability constants follow the expected trends, yielding the expected maxima for metals with $d^3$ configurations,
Figure 4.6: Comparison of experimental ($\log K_{\text{Expt}}$) and refined ($\log K_A$) stability constants. Solid line: $\log K_{\text{Expt}} = \log K_A$ where $\log K_A = 0.0446 \log K_{\text{Calc}} + 9.5076$ for metal(II) complexes and dashed line: $\log K_{\text{Expt}} = \log K_A + 15.70$ where $\log K_A = 0.0446 \log K_{\text{Calc}} + 25.2076$ for metal(III) complexes.
the Cr(III) HDFOB$^+$ and V(II)HDFOB complexes, and minima for the high-spin $d^5$ configurations, the Fe(III)HDFOB$^+$ and Mn(II)HDFOB complexes. It would appear from the foregoing consideration of the crystal field theory that the maximum for M(II)DFOB complexes should occur at $d^8$ (Ni(II)), not at $d^9$ (Cu(II)). However, as discussed earlier, the Cu(II) HDFOB complex adopts a distorted octahedral structure as a result of the Jahn-Teller effect, which leads to a greater stabilization than predicted by simple electrostatic theory.

Assuming the electrostatic interaction between the hard oxygen atoms and charge-dense metal ions contributes significantly to the stability of a complex, it is expected that stability of a M(III) (hard metals) metal complex will exceed that of a M(II) (soft metal) metal complex. Quantitative assessment of the relationship between structure and stability demonstrated that the density of the first coordination shell (metal-oxygen bonds)
could be directly related to complex stability. Figure 4.8 illustrates that the computed and experimental binding energies demonstrate comparable trends when plotted against charge-normalized ionic radii. For instance, the weak binding constants for Cd(II)HDFOB (expt. $\log K = 8.76^{47}$), Mn(II)HDFOB (expt. $\log K = 7.68^{44}$) are computationally predicted ($\log K_A$) to be 8.22 and 10.01, respectively. Similarly, the strongest binding constant for M(II) ions is found to be Cu(II) HDFOB (experimental $\log K = 14.61$, calculated $\log K_A = 14.52$). Also, the calculated stability constants for M(III) metal ions, Mn(III) HDFOB$^+$ (expt. 29.88$^{44}$, calc. $\log K_A = 32.22$), Fe(III)HDFOB$^+$ (expt. 32.02$^{47}$, calc. $\log K_A = 30.90$) and Co(III)HDFOB$^+$ (expt. 37.45$^{43}$, calc. $\log K_A = 36.23$), and for the tetravalent V(IV)HDFOB$^{2+}$ complex (expt. 31.81$^{105}$, calc. $\log K_A = 28.29$), agree quite well with experimental values. The greatest discrepancy between experimental and refined $\log K$ values is observed for the high charge V(V)HDFOB$^{3+}$ complex (expt. 37.37$^{105}$, calc. $\log K_A = 47.18$).

### 4.4 Spectroscopy

Vibrational frequencies were computed for structures at their fully optimized geometries using analytical derivative techniques. Negative frequencies were not found, indicating that the optimized geometries are at true energy minima. For a more accurate prediction and assignment of vibrational spectral modes, the computed B3LYP frequencies were refined by the scaled quantum-mechanical (SQM) formula.$^{16,17}$ The force-constant matrix for each metal complex was defined in a non-redundant set of Pulay’s internal coordinates$^{16,17}$ and scaled by three factors. Two of them were used to refine modes that are sensitive to aqueous environment: the CO stretching of amide groups scaled by a factor of 0.865 and the bending deformations of the terminal NH$_2^+$ group scaled by a factor of 0.875. All remaining elements of the force constant matrix were scaled by a uniform factor of 0.950. Previously published studies have shown that the SQM formula is expected to produce results that compare to
CHAPTER 4. SIDEROPHORE-METAL COMPLEXES

Figure 4.8: Calculated (log $K_A$) and experimentally determined (log $K_{Expt}$) metal-HDFOB complex stability constants$^{43–45,47,81,101,102}$ versus charged normalized ionic radii [Å]. Linear fits: $y = 76.2x + 43.9$, $R^2 = 0.776$ (expt, solid line), $y = 85.2x + 47.7$, $R^2 = 0.8003$ (calc, dashed line).
Experimental data within 12 cm\(^{-1}\) or less, allowing for confident assignments of observed bands.\(^{16–23}\)

Experimental and simulated infrared (IR) spectra of metal-HDFOB complexes are shown in Figure 4.9. The range of 1300 – 1650 cm\(^{-1}\) shown in Figure 4.9 has been identified as the most intense and most responsive to metal-ligand interactions, which is in accordance with previously published IR spectra for the Fe(III)HDFOB\(^+\) complex.\(^{73,83,106}\) Additionally, experimental spectra over the range of 900 – 1650 cm\(^{-1}\) along with computed spectra is shown in (Figure 4.10). Based on a normal mode analysis of computed frequencies, main spectral features of the region were assigned to (1) amide I and amide II bands of two amide groups that are exposed to the solvent, (2) bending deformations of the terminal NH\(_3^+\) group and (3) combinations of CN and CO stretches associated with atoms from the first and second coordination shells. Mode assignments and their computed and experimental frequencies for all metal complexes are listed in Table 4.5, Table 4.6 and Table 4.7.

The highest-frequency band in the experimental spectra is readily observed as a strong signal around 1630 cm\(^{-1}\), which was predicted to consist of two near-degenerate signals. The normal mode analysis of computed spectra identified two fundamentals also located at 1630 cm\(^{-1}\), corresponding to amide I bands of two solvent exposed amide groups (Table 4.5). The corresponding amide II modes for most metal complexes were predicted to be in the range of 1525 – 1540 cm\(^{-1}\). However, the experimental spectra did not clearly exhibit bands in that region. It is likely that the experimental bands are much weaker than the predicted signals and are therefore hidden under other broad bands in the region (see Figure 4.9). Both amide I and amide II frequencies remain constant for all metals except the V(V) ion, where one of the amide I frequencies shifts down by 15 cm\(^{-1}\). This is likely due to polarization of the amide group by the metal ion.
Figure 4.9: Calculated (black trace) and experimental (red trace) IR spectra (1300−1650 cm\(^{-1}\) frequency region). Trace of computed spectra based on the 12 − cm\(^{-1}\) bandwidth of the Lorentzian fit. Selected computed bands shown by vertical bars: amide I and amide II (black), the terminal NH\(^+_3\) deformations (green), amide I-like of oxyamide (blue), amide II-like of oxyamide (purple). Experimental assignments, marked by asterisk: amide I (black), amide I-like of oxyamide (blue), amide II-like of oxyamide (purple). For numerical values see Table 4.5, Table 4.6, and Table 4.7
Figure 4.10: Calculated (Black trace) and experimental (Red trace) IR spectra (900–1650 cm\(^{-1}\) frequency region). Trace of computed spectra based on the 12 cm\(^{-1}\) bandwidth of the Lorentzian fit. Selected computed bands shown by vertical bars: amide I and amide II (Black), the terminal \(-NH_3^+\) deformations (Green), amide I-like of oxyamide (Blue), amide II-like of oxyamide (Purple). Experimental assignments, marked by stars: amide I (Black), amide I-like of oxyamide (Blue), amide II-like of oxyamide (Purple). For numerical values see Table 4.5, Table 4.6, and Table 4.7
Table 4.5: Calculated and experimental frequencies \([cm^{-1}]\) for selected fundamentals of 1300 – 1650 \(cm^{-1}\) spectral region, infrared intensities \([km/mol]\), and vibrational mode assignments for amide I and amide II of the third coordination shell.

<table>
<thead>
<tr>
<th>Metal</th>
<th>Amide A (3rd Shell)</th>
<th>Amide B (3rd Shell)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A chain</td>
<td>B chain</td>
</tr>
<tr>
<td>V(^{2+})</td>
<td>1631</td>
<td>369</td>
</tr>
<tr>
<td>Cr(^{2+})</td>
<td>1630</td>
<td>358</td>
</tr>
<tr>
<td>Mn(^{2+})</td>
<td>1629</td>
<td>1631</td>
</tr>
<tr>
<td>Fe(^{2+})</td>
<td>1631</td>
<td>367</td>
</tr>
<tr>
<td>Co(^{2+})</td>
<td>1629</td>
<td>1631</td>
</tr>
<tr>
<td>Ni(^{2+})</td>
<td>1629</td>
<td>1631</td>
</tr>
<tr>
<td>Ni(^{3+})</td>
<td>1629</td>
<td>1631</td>
</tr>
<tr>
<td>Cu(^{2+})</td>
<td>1628</td>
<td>1631</td>
</tr>
<tr>
<td>Zn(^{2+})</td>
<td>1630</td>
<td>1632</td>
</tr>
<tr>
<td>Cd(^{2+})</td>
<td>1627</td>
<td>1635</td>
</tr>
<tr>
<td>Hg(^{2+})</td>
<td>1634</td>
<td>376</td>
</tr>
<tr>
<td>Pb(^{2+})</td>
<td>1634</td>
<td>373</td>
</tr>
<tr>
<td>Ave(^{2+})</td>
<td>1632</td>
<td>389</td>
</tr>
<tr>
<td>Sc(^{3+})</td>
<td>1630</td>
<td>366</td>
</tr>
<tr>
<td>Ti(^{3+})</td>
<td>1630</td>
<td>360</td>
</tr>
<tr>
<td>V(^{3+})</td>
<td>1630</td>
<td>377</td>
</tr>
<tr>
<td>V(^{4+})</td>
<td>1629</td>
<td>379</td>
</tr>
<tr>
<td>V(^{5+})</td>
<td>1614</td>
<td>512</td>
</tr>
<tr>
<td>Cr(^{3+})</td>
<td>1627</td>
<td>1630</td>
</tr>
<tr>
<td>Mn(^{3+})</td>
<td>1627</td>
<td>1631</td>
</tr>
<tr>
<td>Fe(^{3+})</td>
<td>1627</td>
<td>1631</td>
</tr>
<tr>
<td>Co(^{3+})</td>
<td>1630</td>
<td>1631</td>
</tr>
<tr>
<td>Ga(^{3+})</td>
<td>1631</td>
<td>1631</td>
</tr>
<tr>
<td>Ave(^{3+})</td>
<td>1631</td>
<td>350</td>
</tr>
</tbody>
</table>
Table 4.6: Calculated and experimental frequencies [cm\(^{-1}\)] for selected fundamentals of 1300–1650 cm\(^{-1}\) spectral region, infrared intensities [km/mol], and vibrational mode assignments for the NH\(_3\) bending of the third coordination shell.

<table>
<thead>
<tr>
<th>Metal</th>
<th>Umbrella</th>
<th>Terminal Chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>V(^{2+})</td>
<td>1532</td>
<td>188</td>
</tr>
<tr>
<td>Cr(^{2+})</td>
<td>1557</td>
<td>243</td>
</tr>
<tr>
<td>Mn(^{2+})</td>
<td>1539</td>
<td>153</td>
</tr>
<tr>
<td>Fe(^{2+})</td>
<td>1530</td>
<td>211</td>
</tr>
<tr>
<td>Co(^{2+})</td>
<td>1534</td>
<td>166</td>
</tr>
<tr>
<td>(^1)Ni(^{2+})</td>
<td>1534</td>
<td>919</td>
</tr>
<tr>
<td>(^3)Ni(^{2+})</td>
<td>1536</td>
<td>174</td>
</tr>
<tr>
<td>Cu(^{2+})</td>
<td>1543</td>
<td>169</td>
</tr>
<tr>
<td>Zn(^{2+})</td>
<td>1535</td>
<td>136</td>
</tr>
<tr>
<td>Cd(^{2+})</td>
<td>1557</td>
<td>241</td>
</tr>
<tr>
<td>Hg(^{2+})</td>
<td>1557</td>
<td>285</td>
</tr>
<tr>
<td>Pb(^{2+})</td>
<td>1557</td>
<td>518</td>
</tr>
<tr>
<td>Ave(^{2+})</td>
<td>1542</td>
<td>284</td>
</tr>
<tr>
<td>Sc(^{3+})</td>
<td>1484</td>
<td>130</td>
</tr>
<tr>
<td>Ti(^{3+})</td>
<td>1486</td>
<td>86</td>
</tr>
<tr>
<td>V(^{3+})</td>
<td>1485</td>
<td>108</td>
</tr>
<tr>
<td>V(^{4+})</td>
<td>1457</td>
<td>181</td>
</tr>
<tr>
<td>V(^{5+})</td>
<td>1455</td>
<td>166</td>
</tr>
<tr>
<td>Cr(^{3+})</td>
<td>1489</td>
<td>69</td>
</tr>
<tr>
<td>Mn(^{3+})</td>
<td>1527</td>
<td>148</td>
</tr>
<tr>
<td>Fe(^{3+})</td>
<td>1487</td>
<td>100</td>
</tr>
<tr>
<td>Co(^{3+})</td>
<td>1488</td>
<td>94</td>
</tr>
<tr>
<td>Ga(^{3+})</td>
<td>1486</td>
<td>166</td>
</tr>
<tr>
<td>Ave(^{3+})</td>
<td>1491</td>
<td>113</td>
</tr>
</tbody>
</table>
Table 4.7: Calculated and experimental frequencies [cm\(^{-1}\)] for selected fundamentals of 1300 – 1650 cm\(^{-1}\) spectral region, infrared intensities [km/mol], and vibrational mode assignments for the second coordination shell.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V(^{2+})</td>
<td>1581 576</td>
<td>1439 8</td>
<td>1568 651</td>
<td>1454 10</td>
<td>1574 609</td>
<td>1452 11</td>
<td>1574 612</td>
<td>1449 10</td>
<td></td>
</tr>
<tr>
<td>Cr(^{2+})</td>
<td>1593 676</td>
<td>1451 130</td>
<td>1610 235</td>
<td>1449 40</td>
<td>1581 795</td>
<td>1435 97</td>
<td>1590 569</td>
<td>1445 89</td>
<td></td>
</tr>
<tr>
<td>Mn(^{2+})</td>
<td>1586 1587 740</td>
<td>1459 1453 81</td>
<td>1573 602</td>
<td>1461 37</td>
<td>1579 596</td>
<td>1447 35</td>
<td>1586 1580 646</td>
<td>1459 1453 51</td>
<td></td>
</tr>
<tr>
<td>Fe(^{2+})</td>
<td>1587 669</td>
<td>1466 21</td>
<td>1571 515</td>
<td>1461 22</td>
<td>1576 581</td>
<td>1459 51</td>
<td>1579 588</td>
<td>1461 31</td>
<td></td>
</tr>
<tr>
<td>Co(^{2+})</td>
<td>1582 1593 635</td>
<td>1463 1466 30</td>
<td>1573 536</td>
<td>1462 27</td>
<td>1577 633</td>
<td>1456 16</td>
<td>1582 1581 601</td>
<td>1463 1462 24</td>
<td></td>
</tr>
<tr>
<td>Ni(^{2+})</td>
<td>1611 1611</td>
<td>1461 1450 48</td>
<td>1629 1628 450</td>
<td>1429 29</td>
<td>1587 1588 710</td>
<td>1428 193</td>
<td>1608 1605 497</td>
<td>1461 1432 90</td>
<td></td>
</tr>
<tr>
<td>Ni(^{3+})</td>
<td>1587 1595 775</td>
<td>1461 1453 79</td>
<td>1578 623</td>
<td>1462 32</td>
<td>1577 347</td>
<td>1456 17</td>
<td>1578 1586 626</td>
<td>1461 1456 43</td>
<td></td>
</tr>
<tr>
<td>Cu(^{2+})</td>
<td>1599 1600 830</td>
<td>1463 59</td>
<td>1593 450</td>
<td>1464 532</td>
<td>1581 673</td>
<td>1442 28</td>
<td>1599 1592 651</td>
<td>1464 1455 46</td>
<td></td>
</tr>
<tr>
<td>Zn(^{2+})</td>
<td>1586 1595 846</td>
<td>1466 24</td>
<td>1577 581</td>
<td>1464 25</td>
<td>1581 425</td>
<td>1462 35</td>
<td>1586 1586 617</td>
<td>1462 1463 28</td>
<td></td>
</tr>
<tr>
<td>Cd(^{2+})</td>
<td>1587 697</td>
<td>1437 23</td>
<td>1582 1575 709</td>
<td>1455 31</td>
<td>1582 1580 536</td>
<td>1446 1437 48</td>
<td>1582 1581 647</td>
<td>1446 1443 34</td>
<td></td>
</tr>
<tr>
<td>Hg(^{2+})</td>
<td>1581 665</td>
<td>1433 35</td>
<td>1576 556</td>
<td>1451 19</td>
<td>1573 685</td>
<td>1453 24</td>
<td>1577 636</td>
<td>1444 26</td>
<td></td>
</tr>
<tr>
<td>Pb(^{2+})</td>
<td>1604 795</td>
<td>1455 38</td>
<td>1587 486</td>
<td>1459 21</td>
<td>1599 428</td>
<td>1437 24</td>
<td>1598 570</td>
<td>1451 28</td>
<td></td>
</tr>
<tr>
<td>Ave(^{2+})</td>
<td>1593 686</td>
<td>1453 48</td>
<td>1584 533</td>
<td>1455 29</td>
<td>1581 596</td>
<td>1447 48</td>
<td>1586 605</td>
<td>1451 42</td>
<td></td>
</tr>
<tr>
<td>Sc(^{3+})</td>
<td>1583 455</td>
<td>1463 70</td>
<td>1566 412</td>
<td>1467 55</td>
<td>1573 532</td>
<td>1460 139</td>
<td>1574 466</td>
<td>1462 88</td>
<td></td>
</tr>
<tr>
<td>Ti(^{3+})</td>
<td>1579 401</td>
<td>1463 132</td>
<td>1559 348</td>
<td>1467 6</td>
<td>1559 409</td>
<td>1459 126</td>
<td>1570 386</td>
<td>1461 88</td>
<td></td>
</tr>
<tr>
<td>V(^{3+})</td>
<td>1596 349</td>
<td>1436 266</td>
<td>1568 610</td>
<td>1476 93</td>
<td>1559 262</td>
<td>1456 58</td>
<td>1574 407</td>
<td>1448 139</td>
<td></td>
</tr>
<tr>
<td>V(^{4+})</td>
<td>1559 279</td>
<td>1463 66</td>
<td>1554 153</td>
<td>1455 106</td>
<td>1549 53</td>
<td>1438 142</td>
<td>1557 162</td>
<td>1449 105</td>
<td></td>
</tr>
<tr>
<td>V(^{5+})</td>
<td>1497 162</td>
<td>1473 172</td>
<td>1477 133</td>
<td>1453 23</td>
<td>1480 159</td>
<td>1461 167</td>
<td>1485 151</td>
<td>1466 121</td>
<td></td>
</tr>
<tr>
<td>Cr(^{3+})</td>
<td>1594 1595 409</td>
<td>1466 1460 138</td>
<td>1594 1576 331</td>
<td>1462 44</td>
<td>1594 1582 398</td>
<td>1446 1455 96</td>
<td>1595 1585 379</td>
<td>1446 1458 93</td>
<td></td>
</tr>
<tr>
<td>Mn(^{3+})</td>
<td>1605 311</td>
<td>1443 1440 173</td>
<td>1576 1571 577</td>
<td>1443 1442 18</td>
<td>1576 1582 439</td>
<td>1425 349</td>
<td>1576 1582 443</td>
<td>1443 1430 180</td>
<td></td>
</tr>
<tr>
<td>Fe(^{3+})</td>
<td>1576 1565 394</td>
<td>1472 39</td>
<td>1554 343</td>
<td>1456 1471 158</td>
<td>1576 1559 421</td>
<td>1456 1463 127</td>
<td>1576 1560 386</td>
<td>1456 1468 108</td>
<td></td>
</tr>
<tr>
<td>Co(^{3+})</td>
<td>1607 1612 388</td>
<td>1442 1442 149</td>
<td>1579 1584 489</td>
<td>1442 1459 147</td>
<td>1597 450</td>
<td>1442 1452 68</td>
<td>1593 1597 442</td>
<td>1442 1451 121</td>
<td></td>
</tr>
<tr>
<td>Ga(^{3+})</td>
<td>1597 1599 524</td>
<td>1460 1468 200</td>
<td>1576 276</td>
<td>1476 30</td>
<td>1597 1585 405</td>
<td>1460 1462 104</td>
<td>1597 1589 402</td>
<td>1460 1467 111</td>
<td></td>
</tr>
<tr>
<td>Ave(^{3+})</td>
<td>1592 404</td>
<td>1455 146</td>
<td>1569 423</td>
<td>1465 69</td>
<td>1576 415</td>
<td>1454 133</td>
<td>1579 414</td>
<td>1457 116</td>
<td></td>
</tr>
</tbody>
</table>
Other IR intense modes of the third coordination shell are associated with bending of the terminal NH$_3^+$ group. The 1620 – 1630 cm$^{-1}$ frequency assigned to deformation B overlaps with and enhances the amide I signal (1630 cm$^{-1}$). The symmetric, umbrella-like bending and deformation A mode are predicted at lower frequencies and are more sensitive to the metal oxidation state. The symmetric bending mode is expected around 1540 cm$^{-1}$ for M(II) metals and around 1490 cm$^{-1}$ for M(III) metals. Deformation A is expected around 1605 cm$^{-1}$ and 1575 cm$^{-1}$, for M(II) and M(III) metals, respectively. The predicted sensitivity of these modes could be related to the formation of hydrogen bonds from the protonated pendant amine to the oxygen atoms, as briefly discussed in the Structural characterization section.

Table 4.7 lists computed and experimental IR bands in the region that are assigned to skeletal vibrations of the second coordination shell. The most prominent and characteristic signals are two broad bands; a strong signal around 1580 – 1590 cm$^{-1}$ and a moderate signal around 1445 – 1460 cm$^{-1}$. Assignment of these bands identifies a combination of symmetric and asymmetric CO and CN stretching modes of three ion coordinated oxyamide groups (A, B and C). Based on the structural similarity of the oxyamide and amide groups, the two skeletal vibrations of oxyamide groups could be characterized as an amide I-like (a symmetric stretch, around 1580 – 1590 cm$^{-1}$ assignable primary to oxyamide C=O stretching) and an amide II-like (asymmetric stretch around 1445 – 1460 cm$^{-1}$ assignable mainly to C–N stretching). Both bands contain three vibrators corresponding to three oxyamides that would group into A and (degenerate) E combinations, if there were strict threefold symmetry. This may explain the broadness of these bands in experimental spectra. For easier spectral assignments of these modes, Table 4.7 also compares the experimental data with the intensity-weighted, amide I-like and amide II-like frequencies. The variation of frequencies of the stronger and more easily-detectable 1580 cm$^{-1}$ signal (assignable to a primary C=O stretching mode) reveals variance in the M–O bond distance and could serve as a spectroscopic marker that reflects the nature of ion binding and kinematic and electronic coupling
among the C=O vibrators.

In order to illustrate the strong relationship between the C=O bond distances and metal binding stability, four panels of Figure 4.11 show the correspondence between average computed C=O distance of oxyamides and (A) the average of axial M–O distances, (B) the average of axial and equatorial M–O distances, (C) the difference in axial and equatorial distances, and (D) the average C–N bond distances. Panels (A) and (B) reflect the coupling of the C=O distance with the strength and nature of M–O binding, panel (C) demonstrates the Jahn-Teller distorted complexes, and panel (D) reveals substantial electron delocalization in the oxyamide rings modulated by the coordinated metal ion. As might be expected from the resonance structure of chelate rings (Scheme 4.3), the C=O distance and consequently the C=O vibrator is sensitive to kinematic and electronic changes in the rings. Generally, as inner coordination shell M–O distances increase, the C=O distance decreases. This correlates with an increased negative charge on the oxygen atom and the increased ionic nature of M–O(C) binding. M(II) complexes with weaker binding constants and longer M–O distances exhibit shorter C=O distances. The C=O distance gradually increases for M(III) and higher oxidation state metals. A remarkably short C=O distance has been found in the Co(III) complex, which correlates with the unusually strong binding constant of the complex. The greatest linear correlation ($R^2 = 0.7744$) was found between the C=O and the C–N bond distances (panel D of Figure 4.11). Generally, an inverse correlation between C=O distance and C–N distance might be expected, reflecting the resonant nature of the structure. However, some striking deviations from linearity demonstrate unexpectedly elongated C–N bonds, such as in the V(V)HDFOB$^{3+}$ and Cr(II)HDFOB complexes. The elongated C–N bond could be attributed to an extended delocalization of the double binding to alpha carbons (Scheme 4.3). This explanation is quite plausible for higher oxidation state metal ions such as V(V), but is not quite clear for softer ions such as Cr(II).

The corresponding panels of four correlational plots for the C=O marker vibrator (selected distances vs. computed frequency of the asymmetric, amide I-like mode) are presented
Figure 4.11: Correlation plots of computed values for selected bond distances [Å] vs. the CO bond length: M(II) ions (square) and M(III) ions (triangle). Vertical lines show the average CO bond distances for M(II) (black), M(III) (blue) and higher oxidation states (pink).

Scheme 4.3: The asymmetric (amide I-like) skeletal vibration of metal coordinated oxyamide for various ion coordination depicted by the Lewis structures. Legend: (1) Jahn-Teller distorted complexes (1600 cm\(^{-1}\), primary C=O stretching), (2) weak coordination of M(II) ions (1580 cm\(^{-1}\), primary C=O stretching), (3) strong coordination of M(III) ions (1580 cm\(^{-1}\), primary C=N\(^+\) stretching), (4) strong coordination of high oxidation metals (1500 cm\(^{-1}\)).
in Figure 4.12. Concomitant with decreasing C=O bond strength, the C=O frequency is expected to increase as the first-shell M–O distance increases (the weaker binding constants). In the case of the Jahn-Teller distorted complexes, such an increase would be expected due to the increased distance of at least one of the C=O fragments from the metal center, which would be less involved in the metal binding reaction. This expected trend is well replicated for M(II) complexes; however, for the much stronger M–O bonds observed in M(III) complexes, the frequency follows the reverse order. For example, the average marker frequency increases from 1560 cm$^{-1}$ for Fe(III)HFOB$^+$ to 1585 cm$^{-1}$ for Cr(III)HFOB$^+$ to 1597 cm$^{-1}$ for Co(III)HFOB$^+$, with the corresponding M–O bond distance decreasing from 2.018Å to 1.981Å to 1.893Å. This counterintuitive trend, both captured by DFT calculations and experimentally observed, is due to systematic changes in mode decomposition accompanied by changes in electron delocalization in the chelate rings. As the metal binding becomes stronger and more ionic in nature, shifting the electrons towards the oxygen atom, the composition of the amide I-like mode becomes primarily due to the C=N$^+$ stretching vibration (structure 3 of Scheme 4.3) resulting in a systematic increase of the frequency. It is worth noting that along with the composition changes in the amide I-like frequency for M(III) complexes, the corresponding lower frequency band (1450 cm$^{-1}$, amide II-like mode) becomes mainly due to the C–O$^-$ stretching vibration and therefore a noticeably stronger IR intensity than observed in the M(II) complexes (Table 4.7 and Figure 4.9). Remarkably, the predicted and experimentally observed substantial increase in IR intensity ratio of the amide II-like mode to the amide I-like band could be used for spectral recognition of M(II) vs. M(III) complexes. For example, based on the computed average intensities of the amide I and II-like modes, the IR intensity ratio is $41/605 = 0.067$ for M(II) ions and is expected to increase to $116/414 = 0.280$ for M(III) ions. Finally, for complexes with higher oxidation state metals such as V(V), both amide I-like and amide II-like modes are shifted to approximately 1470 cm$^{-1}$, which is consistent with elongated C–O$^-$ and C–N bond distances and with a plausible electronic configuration represented by Lewis structure 4 in Scheme 4.3.
Figure 4.12: Correlation plots of computed values for selected bond distances [Å] vs. the amide I-like frequency [cm⁻¹]: M(II) ions (square) and M(III) ions (triangle). The average amide I-like frequencies shown by vertical lines and the average CO bond distances shown by horizontal lines: for M(II) (black), M(III) (blue) and higher oxidation states (pink). Numerical labels in panel D correspond to the Lewis structure labels shown in Scheme 4.3. See text for more details.
It is also worth noting that in the Jahn-Teller distorted Mn(III) HDFOB$^+$ complex, the sensitivity of the CO vibration to increasing axial M–O(C) bond distance is greatly suppressed when compared to the sensitivity observed for Cu(II)HDFOB or singlet-Ni(II)HDFOB complexes. The average intensity of the amide I-like frequency is predicted to be 1592 $cm^{-1}$ for the Cu(II)HDFOB complex and 1605 $cm^{-1}$ for the singlet-Ni(II)HDFOB complex, which are noticeably higher than the predicted average of 1583 $cm^{-1}$ for the other M(II) metal complexes. The predicted maximum frequency of the CO vibration of the Mn(III)HDFOB$^+$ complex is 1582 $cm^{-1}$, which is only 3 $cm^{-1}$ higher than the average frequency of 1579 $cm^{-1}$ for the other M(III) complexes. This observation is consistent with the electronic and normal mode redistribution generally observed in M(III)HDFOB$^+$ complexes, where the kinematics of the vibration is shifted towards the C=N$^+$ stretching.

4.5 Conclusions

Density functional theory was employed to study the structure, binding, and vibrational spectra of metal desferrioxamine B complexes in aqueous solution. The DFT-computed structures agree well with structural data derived from EXAFS data collected for aqueous complexes. DFT binding energies have indicated systematic errors that were successfully extracted by an empirical fit. The refined energies agree well with and reproduce the general trends observed for measured binding constants, including dependence on the $d$-orbital electronic configuration and the charge-normalized ionic radii. Also, an interesting feature of the structural calculations is the hydrogen-bonding interaction of the amine tail with complex core, an interaction that becomes less important with increasing metal charge.

Normal mode analysis of the IR spectra in the 1300 – 1650 $cm^{-1}$ region showed that the dominating spectral mode in the region, the asymmetric combination of CO and CN stretching (amide I-like) is sensitive to the electronic properties of metal core. It was established that the normal mode composition of CO/CN stretchings and electron delocalization
in the chelated rings dramatically changes as a function of the metal charge and its frequency directly correlates with the strength and structural distortion of the core.

The average frequency for the marker band is estimated at 1582 $cm^{-1}$. Blue-shift of the band in M(II) metal complexes indicates breaking the octahedral symmetry of the metal core due to the Jahn-Teller or other effects. For example, for the Jahn-Teller distorted Cu(II) complex the band is observed at 1599 $cm^{-1}$ (calc. at 1592 $cm^{-1}$), for the square planar singlet-Ni(II) complex, the experimental band is at 1608 $cm^{-1}$ (calc. at 1605 $cm^{-1}$), and for Pb(II) complex the band is predicted at 1598 $cm^{-1}$. For the M(III) metal complexes the elevated frequency typically correlates with the increased strength of the complex. For example, for a stable Co(III)HDFOB$^+$ complex with log $K = 37.45^{43}$ (expt.) and 36.23 (calc.), the band is observed at 1593 $cm^{-1}$ (calc. at 1597 $cm^{-1}$). For higher oxidation state metals the asymmetric mode is expected to be red-shifted; for instance, the V(IV)HDFOB$^{2+}$ complex is predicted at 1557 $cm^{-1}$ and for the V(V)HDFOB$^{3+}$ complex at 1485 $cm^{-1}$. Remarkably, spectral segregation of M(II) vs. M(III) metals could relate to increased IR intensity ratio of the amide II-like band (typically 1450 $cm^{-1}$) to the marker band, amide I-like mode.
Chapter 5

Catalase Peroxidase (KatG)

E pur si muove.
– Galileo Galilei

Catalase-peroxidases (KatGs) are heme enzymes that exhibit broad peroxidase activity and unusually high catalase activity at rates approaching those of monofunctional catalases. Many aerobic microorganisms utilize this dual function enzyme to eliminate the potential damaging effects of hydrogen peroxide by converting it to water in the peroxidase pathway ($2AH_{\text{red}} + H_2O_2 \rightarrow 2H_2O + 2A_{\text{ox}}$), or oxygen and water during catalase turnover ($2H_2O_2 \rightarrow O_2 + 2H_2O$). The enzyme provides virulence to the pathogen causing tuberculosis (M. tuberculosis), a disease that continues to fatally infect millions of people each year.\textsuperscript{107,108} KatG is also required in the mechanism of action of the anti-tuberculosis pro-drug isoniazid, the activation of which involves the peroxidase activity of the enzyme.\textsuperscript{109,110} Mutations in this enzyme are the cause of widespread resistance to isoniazid,\textsuperscript{111,112} creating a serious worldwide obstacle to effective tuberculosis treatment. The available crystal structures of KatGs from H. marismortui (1ITK.pdb\textsuperscript{113}), B. pseudomallei (1MWV.pdb\textsuperscript{114}), and M. tuberculosis (2CCA.pdb\textsuperscript{115}, 1SJ2.pdb\textsuperscript{116}) reveal a conserved three-dimensional organization of the proximal and distal sides of the heme pocket that is structurally similar to well-characterized members of the Class I peroxidase superfamily, cytochrome c peroxidase (1CCA.pdb\textsuperscript{117}) and
ascorbate peroxidase (1IYN.pdb\textsuperscript{118}). However, for KatG, the X-ray structures also disclosed a novel amino acid adduct containing covalently linked side-chains of residues Trp-107, Tyr-229 and Met-255 on the distal side of the heme (\textit{M. tuberculosis} numbering, 2CCA.pdb\textsuperscript{115}). This distinctive feature (MYW) adds KatG to a growing list of metalloenzymes that have covalently modified amino acids in their active sites.\textsuperscript{119–122} The two covalent bonds required for formation of the adduct in KatG (\(C_\epsilon^2\) of Tyr to \(S_\delta\) of Met and \(C_\epsilon^1\) of Tyr to \(C_\eta^2\) of Trp) are generated post-translationally in auto-catalytic processing that has been partially characterized and was shown to require the heme,\textsuperscript{123,124} but is not yet clearly understood. It is clear however, that both the heme and MYW cofactors found in KatG are mutually dependent on one another to support catalase activity. The covalently linked adduct cannot be synthesized without the heme, and replacement of any of the three adduct residues by mutagenesis eliminates the enzyme’s catalase function.\textsuperscript{123,125} Additionally, replacement of Arg-418 (with Leu but not Lys) also eliminates catalase function, illustrating a unique role for this residue.\textsuperscript{126}

The turnover of the resting ferric enzyme with the first peroxide is considered to generate a heme intermediate (\([\text{Fe(I)}V=\text{OPor}^+\cdot]\)) typical of monofunctional peroxidase and catalase Compound I.\textsuperscript{127–130} Experimentally, the formation and characterization of a stable Compound I in monofunctional peroxidases is well established.\textsuperscript{131} However, in the case of typical catalases, this and also the subsequent steps of the catalase mechanism are experimentally challenging to observe because of very rapid turnover.\textsuperscript{132–134} Characterization of the Compound I species in KatG by extensive QM/MM studies\textsuperscript{127,135,136} is proposed to involve internal electron transfer processes producing protein based radicals on nearby residues. Depending on pH, two different pathways were computationally predicted. Both yield the ferryl heme intermediate (equivalent to peroxidase Compound II)\textsuperscript{131} with a radical localized on the MYW-adduct and on the proximal Trp-321 residue with relative spin proportions dependent on pH. The MYW-radical dominates at higher pH (MYW deprotonated at Y OH group) while the Trp-321 radical takes over at lower pH (\textit{M. tuberculosis} numbering, 2CCA.pdb\textsuperscript{115}).
Figure 5.1: Comparison of X-ray crystal structures of KatGs [gray 2CCA\(^{115}\), pink 1ITK\(^{113}\), blue 1SJ2\(^{116}\), yellow 1MWV\(^{114}\) showing the conserved MYW adduct nearby the heme group. Residue numbers from 2CCA.pdb. 1ITK: Trp95, Tyr218, Met244, His259, Arg409. 1SJ2: Trp107, Tyr229, Met255, His270, Arg418. 1MWV: Trp111, Tyr238, Met264, His279, Arg426.
While the latter species may form, it does not appear to be indispensable for catalase function in KatG, as replacement of the proximal Trp-321 with Phe does not eliminate catalase activity.\textsuperscript{137–139} In contrast, as mentioned above, the MYW adduct is essential for catalase turnover. Considering a radical localized on the MYW adduct (high pH) and the ferryl heme species generated after reaction with the first peroxide, two possible routes for turnover with the second peroxide may be considered, consistent with the known non-scrambling of peroxide oxygen atoms during catalase turnover by KatG:\textsuperscript{124} these include (pathway 1) the direct conversion of the ferryl heme/MYW-radical species to restore the resting enzyme and release molecular oxygen, and (pathway 2) the conversion of ferryl heme to oxyferrous heme (also known as peroxidase Compound III) in a reaction known in other peroxidases\textsuperscript{123,132–134,140–143} and a radical “quenching” step that restores the resting enzyme (ferric heme, neutral MYW) and releases molecular oxygen.

Under experimental conditions where the enzyme reacts with very high concentrations of $\text{H}_2\text{O}_2$ (250-400 mM), the radical and oxyferrous heme are present simultaneously (consistent with pathway 2) and persist long enough to be characterized.\textsuperscript{132} It is the mechanism of decay or quenching of this species that is of special interest, since oxyheme can be a catalytically dead-end species having the oxygen ligand tightly bound.\textsuperscript{144–147} In wild-type KatG, the catalytic ligand elimination rate from oxyheme must be at least as fast as the overall turnover rate of around 5000 $s^{-1}$.\textsuperscript{148} Of special importance to outlining a mechanism in which both oxyheme and the radical partner for a catalytic process is the finding that when oxyheme is generated in KatG without using peroxide, such that the MYW adduct is in its neutral (resting enzyme) form, oxyheme persists for minutes.\textsuperscript{133} Furthermore, oxyheme is very stable in the amino acid cofactor mutant KatG[\text{W107F}].\textsuperscript{133} Of additional relevance is that in the mutant KatG[\text{R418L}]\textsuperscript{126}, oxyheme and the MYW radical still accumulate, and these species persist during slow turnover of $\text{H}_2\text{O}_2$, illustrating first that the two species apparently reside together in this enzyme intermediate and second that structural features beyond the heme play a role in the catalase mechanism.
5.1 The Catalase Reaction - Dioxygen Release from the Oxyheme Intermediate

5.1.1 R(MYW)

Density functional theory was used to investigate the electronic and structural properties of the isolated MYW adduct along with Arg-418 (R(MYW)) in both the radical-cation (MYW) and net neutral (resting state) (MYW) adduct forms. The mobile Arg-418 side chain is incorporated in either the remote or vicinal positions found in the *M. tuberculosis* KatG X-ray crystal structure (2CCA.pdb\textsuperscript{115}).

All calculations were performed using density functional theory at the level of the B3LYP functional combined with 6-311G** basis sets enhanced for environmental effects by the Polarizable Continuum Model (PCM) as implemented in the Gaussian 09 program package\textsuperscript{78}. Systems of interest were built based on the coordinates taken from subunit A of the *M. tuberculosis* KatG structure (2CCA.pdb\textsuperscript{115}). Implicit hydrogens were added throughout the model except at the phenolic oxygen of Tyr-229, and the side chain of Arg-418 was kept in its guanidinium form, as these are the expected protonation states for these groups at neutral pH.

Structural models were generated for the neutral (MYW) and radical-cation (MYW)$^+$ form of the adduct combined with the side chain of the Arg-418 residue in either remote or vicinal positions as found in the KatG x-ray structure (2CCA.pdb\textsuperscript{115}). Each residue of the MYW adduct was terminated at its respective alpha carbon, which was formulated as a methyl group. The models include the side-chain and backbone carbonyl of Arg-418 and the peptide nitrogen and C atom of the adjacent Asp-419 residue. The Arg-418 carbonyl serves as a site for stabilization of the Arg-remote orientation because of a hydrogen bond with an N-bound hydrogen of the guanidinium group.

Structures and electronic features of the KatG active site were calculated using the cluster model approach,\textsuperscript{5,10,12,129,149} whereby the reacting part of the enzyme is isolated from the
macromolecule and is fully described at the quantum-mechanical level. Structural constraints were imposed on the system by fixing the backbone atoms (C) of amino acids Met-255, Tyr-229, Trp-107, Arg-418, and Asp-419. The dihedral angles C\(_\alpha\), C\(_\beta\), C\(_\gamma\), C\(_\delta\) for both Tyr-229 and Trp-107 residues were also fixed according to their crystallographic coordinates (2CCA.pdb\(^{115}\)).

Population analysis of computed spin densities provided confirmation that the MYW radical is essentially tyrosyl-like, with unpaired \(\pi\)-electron density localized mainly within the phenolic ring of the tyrosyl moiety, with some delocalization (30\%) to the covalently linked tryptophan ring, but no unpaired spin on the Met residue (Table 5.1, and Figure 5.2. This is in agreement with earlier calculations\(^{126,127}\) and reports of rapid freeze-quench EPR spectroscopy applied to trap intermediates during hydrogen peroxide turnover by KatG that revealed a tyrosyl-like radical later assigned to the MYW adduct.\(^{132,141}\) Computed charge distributions on the MYW adduct may be found in Table 5.2 and Figure 5.3.

Table 5.1: Calculated total spin per residue in the R(MYW)-radical models.

<table>
<thead>
<tr>
<th></th>
<th>Open R(MYW)(^{++})</th>
<th>S=1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arg-remote</td>
<td>Arg-vicinal</td>
</tr>
<tr>
<td>Arg</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Met</td>
<td>0.0047</td>
<td>0.0038</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.7355</td>
<td>0.6903</td>
</tr>
<tr>
<td>Trp</td>
<td>0.2598</td>
<td>0.3060</td>
</tr>
</tbody>
</table>

Previous calculations also showed that the stable radical form of MYW contains a neutral indole.\(^{132}\) Moreover, in the current study the influence of the distal His-108 versus a water molecule on the formation of an MYW-radical was assessed. Based on the energetics, it was found that the assistance of His-108 lowers the energy of the radical formation by 37 kcal/mole. However, formation of the radical on Tyr-229 rather than Trp-107 is still favored.
Figure 5.2: Calculated spin for the (MYW)$^{\bullet+}$ and net neutral (resting state) MYW-adduct in the R(MYW) models. For clarity the MYW drawings are only skeletal.

Table 5.2: Calculated total charge per residue in the R(MYW) models.

<table>
<thead>
<tr>
<th></th>
<th>Closed R(MYW) S=0</th>
<th>Open R(MYW) S=0</th>
<th>Open R(MYW)$^{\bullet+}$ S=1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arg-remote</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Met</td>
<td>0.8884</td>
<td>0.8908</td>
<td>0.9699</td>
</tr>
<tr>
<td>Tyr</td>
<td>-0.8618</td>
<td>-0.8702</td>
<td>-0.2362</td>
</tr>
<tr>
<td>Trp</td>
<td>-0.0267</td>
<td>-0.0206</td>
<td>0.2663</td>
</tr>
<tr>
<td><strong>Arg-vicinal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>0.9596</td>
<td>0.9574</td>
<td>0.9884</td>
</tr>
<tr>
<td>Met</td>
<td>0.9014</td>
<td>0.9060</td>
<td>0.9718</td>
</tr>
<tr>
<td>Tyr</td>
<td>-0.8538</td>
<td>-0.8558</td>
<td>-0.2732</td>
</tr>
<tr>
<td>Trp</td>
<td>-0.0132</td>
<td>-0.0076</td>
<td>0.3130</td>
</tr>
</tbody>
</table>
Figure 5.3: Bar representation of calculated charge distribution for the (MYW)$^{\bullet+}$ and net neutral (resting state) MYW-adduct in the R(MYW) models. For clarity the MYW drawings are only skeletal.
by 17 kcal/mole. Also, molecular orbital analysis showed that poor delocalization of spin density beyond tyrosine is expected given the anti-bonding nature of the singly occupied molecular orbital (SOMO) of the Y-W $\pi$-bond in the MYW radical. The Y-W $\pi$-bonding in the MYW radical is weak (bond order = 0.5) and the tyrosine-like topology of its SOMO is expected from the orbital energetics of the highest occupied molecular orbitals (HOMO) of the MY and W fragments (Figure 5.4).

Figure 5.4: Molecular orbital diagram of Y-W $\pi$-bonding and anti-bonding orbitals in the MYW-radical. Formed as symmetric and anti-symmetric combination of highest-occupied molecular orbitals for MY and W fragments.
Computed models of R(MYW) embraced two known Arg-418 side-chain conformers and two electronic states of the MYW adduct, namely, the radical cation and its neutral (resting state) form. These structures reveal two conformations of the adduct that differ in the relative orientation of the covalently linked Tyr-229 and Trp-107 rings. For simplicity, the disposition of the rings is best measured by the dihedral angle of the Y-W bond (Tyr Cζ, Cϵ1 and Trp Cη2, Cζ3). The MYW conformations having negative values for the Y-W dihedral angles will be referred to as closed and those with positive values will be referred to as open. The four distinct conformers are depicted in Scheme 5.1 and a scan of the potential energy of these models as a function of the Y-W dihedral angle is presented in Figure 5.5. Independent

![Diagram of Scheme 5.1](image)

Scheme 5.1: Representation of the four conformers of the isolated R(MYW) adduct, computed Y-W dihedral angles and their relative energies. (A) Arg-remote/MYW-closed, (B) Arg-vicinal/MYW-closed, (C) Arg-remote/MYW-open, (D) Arg-vicinal/MYW-open.

of the Arg-418 side-chain position, the most stable structures for both the radical and the neutral MYW adduct are open conformations. For the MYW radical, the calculated Y-W dihedral angle is optimal at around +32 degrees. At this open angle, the indole nitrogen is
oriented below the Tyr ring plane (as viewed in Scheme 5.1, panels C and D). In contrast to the single low energy conformer found for the MYW-radical species, two minima were found for the neutral (resting) MYW form. Of the two minima, the most stable conformation still occurs for positive angles (equal to +44.9 and +40.7 degrees for Arg-vicinal and Arg-remote, respectively) while a second, shallow minimum found about 4.0 kcal/mole higher than the first occurs for dihedral angles of -26.1 and -22.3 degrees (for Arg-vicinal and Arg-remote, respectively). In the latter closed conformers, the indole nitrogen lies above the Tyr ring plane (as viewed in Scheme 5.1, panels A and B). Interestingly, of the two minima for the neutral adduct, the closed conformation closely resembles the resting MYW adduct in the X-ray structure of the enzyme (2CCA.pdb).  

The computed energy barrier for the neutral MYW adduct is small and cannot be rationalized by an increased strength of the $\pi$-bond (Figure 5.4, $\pi$-bond order = 0). Instead, the barrier is likely due to enhanced steric and electrostatic interactions between the phenolate oxygen of Tyr-229 and the nearby hydrogen atom on C$_3$ of Trp-107 because the electron population and size of the lone-pair orbital on the oxygen atom in the neutral MYW adduct are greater than in the radical form. This also explains the slight shift of the optimal Y-W dihedral angle from about +32 degrees for the MYW-radical to about +43 degrees for the neutral MYW adduct. Furthermore, this also suggests that protonation of the phenolate oxygen of Tyr-229 could increase the rotational barrier and potentially differentiate either side of the coplanar Y-W orientations (open or closed).  

Calculations of R(MYW) models also disclose inter- and intra-residue hydrogen bonding for the Arg-vicinal and Arg-remote positions, respectively. For the neutral (resting) MYW-open adduct with Arg-vicinal (Scheme 5.1, panel D), the inter-residue salt bridge (Tyr O$^-$ ··· NH$_2^+$ Arg) stabilizes the orientation by 4.6 kcal/mole compared to the orientation with the intra-residue carbonyl O···NH hydrogen bond for Arg-remote (Scheme 5.1, panel C). This is in good agreement with the corresponding energy of 2.2 kcal/mole reported by the DFT QM/MM study. However, after the MYW-open radical cation is formed in which
Figure 5.5: The energy profiles for R(MYW) cofactor models as a function of the Y-W torsion angle. Neutral MYW (black) with Arg-remote (unfilled) and Arg-vicinal (filled). Radical-cation (MYW)^+ (red) with Arg-remote (unfilled) and Arg-vicinal (filled).
the Tyr O\(^*\) group is neutral, the energy difference between the two Arg side-chain positions is practically negligible, favoring the Arg-vicinal orientation by less than 0.2 kcal/mole. This suggests that for the enzyme in the MYW radical state, both Arg-418 side-chain positions can be equally populated. The consequence of the equal distribution of Arg-418 side-chain positions on the catalase mechanism is discussed later.

Interestingly, for the neutral-closed adduct, the energy cost to change from Arg-vicinal to Arg-remote positions (Scheme 5.1, panels A and B) is computed around 3.6 kcal/mole, which is quite comparable to the energy gained by the MYW opening. These energies are around -3.1 kcal/mole (Scheme 5.1, panels A and C) and around -4.1 kcal/mole (Scheme 5.1, panels B and D) for the adduct with Arg-remote and Arg-vicinal, respectively. Also, the energy difference between MYW-closed/Arg-vicinal (Scheme 5.1, panel B) and MYW-open/Arg-remote (Scheme 5.1, panel C) structures is small (0.5 kcal/mole). In addition, the ionization energies of the neutral MYW-open adduct to produce the MYW-open radical are calculated to be 112 kcal/mole and 104 kcal/mole with Arg-vicinal and Arg-remote, respectively. In comparison, other DFT gas-phase calculations reported these ionization energies at 191 kcal/mole and 132 kcal/mole, respectively.\(^{127}\) Both sets of calculations show a similar trend favoring formation of the MYW-radical with Arg-remote. However, this study only addresses the events occurring after the MYW-radical is formed.

### 5.1.2 R(MYW) Oxyheme

More comprehensive calculations employed the insights gained from the isolated R(MYW) models to create R(MYW)···oxyheme models including a designed 6-coordinate oxyheme group and the proximal His-270 residue. Although the X-ray structure (2CCA.pdb\(^{115}\)) was solved for the (ferric) resting state of the enzyme, the species shown experimentally to be present during catalysis (at neutral pH) is the oxyheme species. Thus, dioxygen was added as the sixth ligand in models assuming an Fe(II) oxidation state for iron. The calculated oxyheme unit closely resembles this structure in oxyAPX (2XIH.pdb\(^{150}\)), although with a
small misalignment of the ligand’s outer oxygen. Systems of interest were built based on the coordinates taken from subunit A of the *M. tuberculosis* KatG structure (2CCA.pdb) for the isolated R(MYW) models and the R(MYW)···oxyheme models that allowed the study of specific interactions between the R(MYW) cofactor and the nearby oxyheme. Implicit hydrogens were added throughout the model except at the two propionate groups of the oxyheme and the phenolic oxygen of Tyr-229. The side chain of Arg-418 was kept in its guanidinium form. A sketch of each residue and the relevant atom labels can be found in Scheme 5.2.

![Scheme 5.2: R(MYW)···oxyheme model showing a labeling scheme for selected carbon atoms; bold font indicates atoms with applied constraints.](image)

Structure optimization for larger models that include the distal His-108 residue found in KatG that could be involved in H-bond interactions with the outer oxygen of the Fe-O₂ complex were also performed at the same level but with the basis set reduced to 6-31G. These calculations were limited to assess the ground state and the catalytically relevant S=5/2 and S=7/2 spin states. Most importantly, these evaluations agreed with and confirmed results found in R(MYW)···oxyheme models presented here, greatly reducing ambiguities in the physical interpretation of the study and justifying the exclusion of His-108. Other
recent studies on diatomic ligands in 6-coordinate heme proteins via vibrational spectroscopy supported by DFT modeling illustrate H-bonding effects on the Fe-O$_2$ bond directed at both the outer and the inner oxygen.$^{151,152}$ It was demonstrated that H-bond donors increase the Fe-O and O-O distances in concert when they are directed at the outer oxygen (His-108 like), but decrease the Fe-O distance while increasing the O-O distance when they are directed at the inner oxygen (Trp-107 like). These studies agree with our computational assessments in cases that include H-bond interactions directed at both inner and outer oxygen of the Fe-O$_2$ complex.

The proximity of the MYW-adduct to the heme in KatG raises a question about its role in the catalase mechanism especially if it catalyzes oxygen release from the oxyheme intermediate. The oxyheme species in KatG may be considered analogous to oxymyoglobin, oxyhemoglobin and peroxidase Compound III formulated in limiting cases as an oxyferrous heme complex (Fe(II)-O$_2$Por) or a superoxyferric complex (Fe(III)-O$_2$•Por). The electronic structure of Fe-O$_2$ is the subject of long-standing and continuing debates that began with the Weiss-Pauling exchange published in Nature in 1964.$^{153,154}$ The debate centers on defining the extent of charge-transfer between iron and the ligand. Since the Weiss-Pauling exchange, an array of experimental and computational methods has been used to address the constitution of the Fe-O$_2$ bond in hemes. Three models are commonly considered: (1) the Pauling model$^{153}$ characterized by a low-spin (S=0) ferrous iron center with singlet O$_2$ bound, (2) the Weiss model$^{154}$ described as a low-spin (S=1/2) ferric iron center, antiferromagnetically coupled to superoxide anion radical (S = 1/2), and (3) the McClure-Goddard model$^{155}$ containing an intermediate-spin (S=1) ferrous ion antiferromagnetically coupled to triplet O$_2$ (S=1). The latter is commonly referred to as an ozone-like binding model that involves a three-center four-electron bond. The molecular orbital diagrams for these three named models are shown in Scheme 5.3. In a recent study by Shaik and coworkers, the constitution of the Fe-O$_2$ bond in oxymyoglobin was evaluated by comprehensive theoretical calculations, namely QM/MM supported by DFT and CASSCF (complete active
Scheme 5.3: Molecular orbital diagram of three models for oxyheme: (a) the Pauling model, \(^{153}\) \(\text{Fe}(S=0)\)-\(\text{O}_2(S=0)\) (b) the Weiss model, \(^{154}\) \(\text{Fe}(S=1/2)\)-\(\text{O}_2(S=1/2)\) and (c) the McClure-Goddard model, \(^{155}\) \(\text{Fe}(S=1)\)-\(\text{O}_2(S=1)\). Oxygen-like states (shown in bold) are relevant for the alternative valence bond (VB) viewpoint.

space self-consistent field) quantum mechanical methods. \(^{156}\) The study supports the idea that the protein environment exerts a major impact on the features of \(\text{Fe-O}_2\) in oxymyoglobin and favors the Weiss bonding model. This model is primarily a result of a \(\sigma(\text{Fe-O}_2)\) bond and a weak \(\pi(\text{Fe-O}_2)\) bond composed of two singly occupied and spin-paired electrons in the two bonding and anti-bonding \((3d_{xz}^\dagger\pi_\perp^\dagger)\) orbitals. A strong bonding interaction is also present and is the result of a doubly occupied bonding \((3d_{xz}^\dagger\pi_\parallel)\) orbital (Scheme 5.3). Despite the multi-configurational character of the \(\text{Fe-O}_2\) ground state described here, the applied density functional method proved to produce an accurate geometry for oxyheme and an electronic picture supporting the Weiss bonding model. \(^{130,156,157}\) For oxyheme in the intermediate of interest here, the presence of the MYW-radical cation introduces an additional close-lying, singly occupied orbital. Thus, for the \(\text{R(MYW)}\cdots\text{oxyheme}\) models, five spin states were considered as energetically accessible and were evaluated computationally: (1) \(\text{R(MYW)}(S=1/2)\cdots\text{Fe}(S=0)\)-\(\text{O}_2(S=0)\) (the Pauling model), with overall \(S=1/2\), (2) \(\text{R(MYW)}(S=1/2)\cdots\text{Fe}(S=1/2)\)-\(\text{O}_2(S=1/2)\), overall \(S=3/2\) and \(M_s=1/2\) state consistent with the Weiss model for oxyheme, (3) \(\text{R(MYW)}(S=1/2)\cdots\text{Fe}(S=1/2)\)-\(\text{O}_2(S=1/2)\),
overall S=3/2 and Ms=3/2 state, which is a spin-flipped Weiss model for oxyheme, (4) R(MYW)(S=1/2)···Fe(S=1)-O₂(S=1), overall S=5/2 and Ms=5/2 high-spin state, which corresponds to a double-spin-flipped (maximized Ms value) McClure-Goddard model, and finally (5) R(MYW)(S=1/2)···Fe(S=2)-O₂(S=1), overall S=7/2 and Ms=7/2 high-spin state. Models of the R(MYW)···oxyheme in each of these five spin states were calculated starting with the X-ray structure coordinates (2CCA.pdb) (closed MYW conformation) with dioxygen added as a sixth ligand to ferrous heme iron, and with the two alternate Arg side-chain conformations.

Relative energies for the five spin states (Table 5.3) identified the (S=3/2, Ms=1/2) state, featuring the Weiss model for oxyheme, as the ground state for the R(MYW)···oxyheme model. In contrast, the doublet state (S=1/2) (Pauling model for oxyheme), was the least favored, predicted about 18 kcal/mole higher than the ground state. Thus, the ground state Fe-O₂ complex modeled here is effectively a superoxide ion bound to low-spin ferric-heme iron, in agreement with QM/MM findings for oxyhemes in KatG and in oxymyoglobin. The energies for the two Arg-418 side-chain positions (Table 5.3) are practically equivalent for each of the five states, favoring the Arg-vicinal orientation by less than 0.2 kcal/mole, similar to results found for the isolated R(MYW) radical. Here, as mentioned above, the lack of a significant energy difference suggests that in the enzyme, there can be an equal population of Arg side-chain sites.

The Fe-O bond length averages around 1.91 Å for superoxy-ferric heme (the Weiss states), 1.75 Å for oxy-ferrous heme (the Pauling states), and about 2.16 Å for higher computed oxyheme spin states. The computed value of 1.91 Å found for Weiss states is in excellent agreement with expected values for oxyhemes and values reported by a QM/MM study. The O-O computed lengths reported here average to a typical value of 1.30 Å, also in good agreement with 1.35 Å computed in a QM/MM model and crystallographic value of 1.29 Å reported for oxy-cytochrome c peroxidase (1DCC.pdb).

In the computed models, the average calculated axial Fe-N_{His} bond length ranges from
Table 5.3: Comparison of the X-ray crystal structure (\textit{M. tuberculosis} KatG, 2CCA.pdb\textsuperscript{115}) with selected computed geometry parameters, computed relative energies\textsuperscript{a}, and S\textsuperscript{2} values of the five spin-states for the closed Arg-remote (MYW)···oxyheme models. Comparison to QM/MM values\textsuperscript{158} is given in parenthesis.

<table>
<thead>
<tr>
<th>X-ray</th>
<th>Arg-remote</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Closed R(MYW)\textsuperscript{++}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fe(III) KatG</th>
<th>Ms=1/2</th>
<th>Ms=3/2</th>
<th>Ms=3/2</th>
<th>Ms=5/2</th>
<th>Ms=7/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe-O (Å)</td>
<td>- 1.750</td>
<td>1.884 (1.91)</td>
<td>1.939 (1.95)</td>
<td>2.173</td>
<td>2.146</td>
<td></td>
</tr>
<tr>
<td>Fe-O-O Angle</td>
<td>- 123.9</td>
<td>120.4</td>
<td>118.9</td>
<td>118.1</td>
<td>115.9</td>
<td></td>
</tr>
<tr>
<td>Fe-N\textsubscript{His} (Å)</td>
<td>2.22</td>
<td>2.070</td>
<td>2.057</td>
<td>2.022</td>
<td>2.325</td>
<td>2.258</td>
</tr>
<tr>
<td>Heme Core Size (Å)</td>
<td>2.07</td>
<td>2.016</td>
<td>2.014</td>
<td>2.012</td>
<td>2.011</td>
<td>2.069</td>
</tr>
<tr>
<td>O-O (Å)</td>
<td>- 1.281</td>
<td>1.300 (1.35)</td>
<td>1.308 (1.36)</td>
<td>1.315</td>
<td>1.308</td>
<td></td>
</tr>
<tr>
<td>W\textsubscript{\textsuperscript{N}}-O dist. (Å)</td>
<td>- 3.036</td>
<td>2.931</td>
<td>2.890</td>
<td>2.809</td>
<td>2.845</td>
<td></td>
</tr>
<tr>
<td>Y-W dih (degrees)</td>
<td>-12.3</td>
<td>-15.0</td>
<td>-15.1</td>
<td>-15.1</td>
<td>-18.2</td>
<td>-17.3</td>
</tr>
<tr>
<td>Rel. Energy (kcal/mol)</td>
<td>- 27.9</td>
<td>9.6</td>
<td>12.0</td>
<td>16.5</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>Computed S\textsuperscript{2} [exact]</td>
<td>- 0.750 [0.75]</td>
<td>0.918 [0.75]</td>
<td>3.751 [3.75]</td>
<td>8.751 [8.75]</td>
<td>15.750 [15.75]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X-ray</th>
<th>Arg-vicinal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Closed R(MYW)\textsuperscript{++}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fe(III) KatG</th>
<th>Ms=1/2</th>
<th>Ms=3/2</th>
<th>Ms=3/2</th>
<th>Ms=5/2</th>
<th>Ms=7/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe-O (Å)</td>
<td>- 1.749</td>
<td>1.884 (1.91)</td>
<td>1.940 (1.95)</td>
<td>2.175</td>
<td>2.148</td>
<td></td>
</tr>
<tr>
<td>Fe-O-O Angle</td>
<td>- 123.8</td>
<td>120.2</td>
<td>119.0</td>
<td>117.6</td>
<td>115.4</td>
<td></td>
</tr>
<tr>
<td>Fe-N\textsubscript{His} (Å)</td>
<td>2.22</td>
<td>2.069</td>
<td>2.058</td>
<td>2.022</td>
<td>2.325</td>
<td>2.257</td>
</tr>
<tr>
<td>Heme Core Size (Å)</td>
<td>2.07</td>
<td>2.016</td>
<td>2.014</td>
<td>2.012</td>
<td>2.011</td>
<td>2.069</td>
</tr>
<tr>
<td>O-O (Å)</td>
<td>- 1.281</td>
<td>1.300 (1.35)</td>
<td>1.309 (1.36)</td>
<td>1.315</td>
<td>1.307</td>
<td></td>
</tr>
<tr>
<td>W\textsubscript{\textsuperscript{N}}-O dist. (Å)</td>
<td>- 3.049</td>
<td>2.941</td>
<td>2.890</td>
<td>2.809</td>
<td>2.848</td>
<td></td>
</tr>
<tr>
<td>Y-W dih (degrees)</td>
<td>-12.3</td>
<td>-14.7</td>
<td>-14.7</td>
<td>-14.4</td>
<td>-17.4</td>
<td>-17.0</td>
</tr>
<tr>
<td>Rel. Energy (kcal/mol)</td>
<td>- 27.9</td>
<td>9.6</td>
<td>11.9</td>
<td>16.5</td>
<td>20.2</td>
<td></td>
</tr>
<tr>
<td>Computed S\textsuperscript{2} [exact]</td>
<td>- 0.750 [0.75]</td>
<td>0.912 [0.75]</td>
<td>3.751 [3.75]</td>
<td>8.751 [8.75]</td>
<td>15.750 [15.75]</td>
<td></td>
</tr>
</tbody>
</table>
2.05 Å in low-spin states to 2.30 Å for high-spin states. The longer bonds are expected for high spin states because of the occupancy of the iron 3d_{z^2} antibonding orbital. For S=7/2 states, the occupancy of the iron d_{xy} antibonding orbital is reflected by a 0.06 Å increase of the heme core size (average Fe-N_pyrrole length). Thus, calculated oxyheme parameters are consistent with expected trends and correlate with electron population in computed spin states.

A general observation about the R(MYW)···oxyheme models is that, unlike the isolated R(MYW) radical models, all structures initially parameterized in the closed conformation of the MYW radical (negative Y-W dihedral angle) remain closed at their optimized energy minima. The computed models have an average Y-W dihedral angle around -15.9 degrees, which compares well with -12.3 degrees found in the X-ray crystal structure of the resting state MYW-adduct in KatG (2CCA.pdb). In our models, a key factor stabilizing these closed forms is the presence of a hydrogen bond between the indole ring NH of Trp-107 and the inner oxygen atom of the bound O_2 ligand. The average computed hydrogen bond length (measured as a donor-acceptor, N···O distance) is around 2.91 Å for the low spin (S = 3/2) states, and around 2.83 Å for high spin states, which indicates a stronger hydrogen bond for the latter case in accord with the occupancy of the Fe 3d_{z^2} orbital and elongated Fe-O bond distance (see below). These values compare well with the corresponding distance of 2.74 Å reported from the X-ray structure of oxy-cytochrome c peroxidase (1DCC.pdb) in which the distal Trp residue is conserved and the H-bond with inner oxygen is also identified.

It should be pointed out here that the computed stable oxygen bound R(MYW)···oxyheme models support experimental evidence that the state exists in the enzyme and can be identified and characterized. Independent of the Arg-418 position, vicinal or remote, the energy of the oxygen bound ground state is computed around 9.6 kcal/mole above the energy of the catalase reaction products: ferric heme, neutral MYW adduct, and molecular oxygen. This energy value is in agreement with the QM(B3LYP)/MM studies by Rovira that identified a KatG oxyheme···MYW-radical state around 11 kcal/mol above the energy of the
enzyme’s resting state and free oxygen. In addition, structural characteristics (Table 5.3) of the ground state oxyheme are comparable in both the QM/MM and our cluster models; however, two highly energetic high-spin states found by QM(B3LYP)/MM at 41.4 kcal/mol (S=5/2) and 83 kcal/mol (S=7/2) above the ground state indicate unexpected spin density on the porphyrin ring. This density is inconsistent with the expected occupation of 3d^*_{z^2} and 3d_{xy} orbitals (see Scheme 5.3) that would increase spin on the Fe atom, which requires much less energy. These two high spin states with expected d-orbital occupancy were computed in our models at 16.5 and 20.2 kcal/mol for S=5/2 and S=7/2, respectively.

Comparing details of the H-bonding interaction between the O_2 ligand and the MYW-radical, the QM/MM and our cluster models differ. In our models, the H-bond occurs with the inner oxygen atom, similar to the H-bonding found in the oxyferrous ascorbate peroxidase (2XIH.pdb\textsuperscript{150}) and the oxyferrous cytochrome c peroxidase (1DCC.pdb\textsuperscript{159}) crystal structures in which the conserved Trp-41 and Trp-51 residues participate (see Figure 5.6). In the QM(B3LYP)/MM calculations, H-bonding involves the outer oxygen atom and both the MYW tryptophan and water. Without definite experimental evidence, both structural models of H-bonding interactions should be considered plausible. In our models, the hydrogen bonding effects on the Fe-O_2 complex play a critical role, since the lone pair orbital (sp^2 hybrid) on the Fe-O_2 inner oxygen, when polarized by H-bonding, contributes to increased (3d_{z^2-\sigma}) and (3d_{xz}) mixing that stabilizes the complex. This is due to increased synergic bonding forces where the O_2 ligand is forced to bend in order to decrease the antibonding interaction with the 3d_{yz} orbital, and at the same time, increase back donation to the (3d_{xz}-\pi^*_z) orbital.

Ultimately, the catalytic process for catalase turnover in KatG requires the dissociation of oxygen, which is predicted to involve an electron transfer to re-establish the resting state enzyme containing ferric heme. It is proposed here that in order to activate the complex, the hydrogen bond between the indole nitrogen of Trp-107 and the bound oxygen must be destabilized. First, it should be recalled that for the isolated R(MYW) models, inde-
Figure 5.6: Comparison of X-ray crystal structures of the oxyferrous ascorbate peroxidase, oxy-APX, (2XIH.pdb\textsuperscript{150}) (Panel A) and the oxy-cytochrome c peroxidase, oxy-CCP (1DCC.pdb\textsuperscript{159}) (Panel B) and \textit{M. tuberculosis} KatG (2CCA.pdb\textsuperscript{115}), showing H-bonding of conserved residues. Color scheme for oxy-APX and oxy-CCP by elements (C gray, N blue, O red and Fe orange), and for KatG by a uniform orange.

independent of the Arg-418 side-chain position, the most stable structures for both the radical and the neutral MYW adduct are open conformations (Figure 5.5). In contrast, for the R(MYW)\cdots oxyheme models, the stable structures remain in the closed conformation and maintain the MYW-radical. For these R(MYW)\cdots oxyheme closed models, increasing the Y-W dihedral angle will bring the Trp-107 indole nitrogen towards the heme plane. Based on structural considerations alone, this motion would break the hydrogen bond to the oxygen atom bound directly to iron, and also internally stabilize the MYW structure (MYW inherent stability described above). Thus, the stable R(MYW)-open conformation of the isolated models was imposed on the oxyheme structure to assess its stability and potential significance. Consideration of both remote and vicinal Arg side-chain sites and the five spin states resulted in 10 additional successfully optimized minima for the MYW-open conformation (Table 5.4). Similar to the procedure applied for the MYW-closed oxyheme states, all initial MYW-open structures included oxygen bound to iron prior to optimization. A new finding emerging from the optimized open models is that dissociated molecular oxygen is
found but only for the two high spin states ($S=5/2$ and $S=7/2$) with Arg-vicinal. In these
two states, the MYW-radical is reduced and $O_2$ is found at a distance greater than 4 Å from
iron (Table 5.4). For clarity, the four stable conformations of the system with bound oxygen
(A-D) and the conformation with dissociated oxygen (E) are shown in Scheme 5.4. The rel-
ative spin-state energies for optimized R(MYW)···oxyheme structures (A-E) are presented
graphically in Figure 5.7. The relative energies for the highest computed spin state ($S=1/2,$
Pauling oxyferrous model) are listed in Table 5.3 and Table 5.4 but are excluded from the

graph.

Figure 5.7: Relative energies for stable oxygen bound (A-D) and oxygen dissociated (E) structures in four \( \text{R(MYW)} \cdot \cdot \cdot \text{oxyheme} \) spin states (shown as horizontal bars). Labels A-E are adapted from Scheme 5.4: (A) Arg-remote/MYW-closed, (B) Arg-vicinal/MYW-closed, (C) Arg-remote/MYW-open, (D) Arg-vicinal/MYW-open, and (E) the \( \text{O}_2 \) released conformation Arg-vicinal/MYW-open. Configurations of the four shown spin states are: \( \text{R(MYW)}(S=1/2) \cdot \cdot \cdot \text{Fe}(S=2)-\text{O}_2(S=1) \) configuration of the overall \( S=7/2 \) spin state (green); \( \text{R(MYW)}(S=1/2) \cdot \cdot \cdot \text{Fe}(S=1)-\text{O}_2(S=1) \) configuration of the overall \( S=5/2 \) spin state (blue), and two spin states of \( \text{R(MYW)}(S=1/2) \cdot \cdot \cdot \text{Fe}(S=1/2)-\text{O}_2(S=1/2) \) configuration of the overall \( S=3/2 \) and \( M_s=3/2 \) (gray) and \( M_s=1/2 \) (black). Computed high energy \( S=1/2 \) states, representing the Pauling model, \( \text{R(MYW)}(S=1/2) \cdot \cdot \cdot \text{Fe}(S=0)-\text{O}_2(S=0) \) are omitted. The energy scale is relative to the oxygen released \( S=5/2 \) state (E, blue bar).
Table 5.4: Comparison of the X-ray crystal structure (\textit{M. tuberculosis} KatG, 2CCA.pdb\textsuperscript{115}) with selected computed geometry parameters, computed relative energies\textsuperscript{a} and \(S^2\) values of the five spin-states for the open Arg-remote (MYW)\(\cdots\)oxyheme models. Comparison to QM/MM values\textsuperscript{158} is given in parenthesis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fe(III)</th>
<th>S=1/2</th>
<th>S=3/2</th>
<th>S=3/2</th>
<th>S=5/2</th>
<th>S=7/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray</td>
<td>KatG</td>
<td>Ms=1/2</td>
<td>Ms=1/2</td>
<td>Ms=3/2</td>
<td>Ms=5/2</td>
<td>Ms=7/2</td>
</tr>
<tr>
<td>Fe-O (Å)</td>
<td>-</td>
<td>1.748</td>
<td>1.872</td>
<td>1.922</td>
<td>2.150</td>
<td>2.210</td>
</tr>
<tr>
<td>Fe-O-O Angle</td>
<td>-</td>
<td>123.5</td>
<td>119.8</td>
<td>119.3</td>
<td>117.2</td>
<td>131.7</td>
</tr>
<tr>
<td>Fe-N\textsubscript{His} (Å)</td>
<td>2.22</td>
<td>2.071</td>
<td>2.069</td>
<td>2.034</td>
<td>2.393</td>
<td>2.237</td>
</tr>
<tr>
<td>Heme Core Size (Å)</td>
<td>2.07</td>
<td>2.015</td>
<td>2.013</td>
<td>2.011</td>
<td>2.010</td>
<td>2.075</td>
</tr>
<tr>
<td>O-O (Å)</td>
<td>-</td>
<td>1.275</td>
<td>1.295</td>
<td>1.303</td>
<td>1.302</td>
<td>1.263</td>
</tr>
<tr>
<td>W\textsubscript{N}-O dist. (Å)</td>
<td>-</td>
<td>4.704</td>
<td>4.645</td>
<td>4.619</td>
<td>4.506</td>
<td>4.590</td>
</tr>
<tr>
<td>Y-W dih (degrees)</td>
<td>-12.3</td>
<td>29.0</td>
<td>29.0</td>
<td>28.9</td>
<td>28.8</td>
<td>28.9</td>
</tr>
<tr>
<td>Rel. Energy (kcal/mol)</td>
<td>-</td>
<td>28.2</td>
<td>11.3</td>
<td>14.3</td>
<td>20.7</td>
<td>22.6</td>
</tr>
<tr>
<td>Computed (S^2) [exact]</td>
<td>-</td>
<td>0.750</td>
<td>0.935</td>
<td>3.751</td>
<td>8.751</td>
<td>15.750</td>
</tr>
</tbody>
</table>
MYW-radical opening coordinate

Deeper understanding of the role that the MYW-radical plays in the mechanism of catalytic activation of the oxyheme and O₂ release requires examination of the conformational change from the MYW-closed to MYW-open structure. Both structures with Arg-vicinal and Arg-remote are considered for this detailed examination. However, only structures with Arg-vicinal are discussed and presented in detail here as both MYW-closed and MYW-open oxyheme models with Arg-remote are shown to be catalytically inert (resulting in stable oxygen-bound structures, A and C). As discussed above, the conformational flexibility is inherent to the MYW-adduct radical, and is referred to here as the MYW-radical opening coordinate. The path of the MYW opening is monitored by imposing incremental changes on the Y-W dihedral angle. The resulting energy profiles along the MYW-opening coordinate with Arg-vicinal are shown in Figure 5.8 for the four lowest energy spin states. The plots of Figure 5.8 show a relatively flat potential energy gradient for the two low-spin states, (S=3/2, Ms=1/2) and (S=3/2, Ms=3/2). For these two spin states, increasing the Y-W dihedral angle eventually eliminates the H-bond to the ligand, but oxygen remains bound and MYW retains the radical form. Thus, these two states are not catalytically activated by the MYW opening. Their energy minima are found for Y-W dihedral angles of approximately +30 degrees, similar to optimal structures found for the isolated MYW-radical (+32 degrees). These open conformations are stable because of an inherent preference of the MYW radical as discussed above. The energy barrier for the transition from the closed to open conformation (shown in Figure 5.8) is 1.8 kcal/mol and 2.4 kcal/mol, when the system is on the (S=3/2, Ms=1/2) and (S=3/2, Ms=3/2) potential energy surfaces, respectively. For these states, the reverse process that would regenerate the H-bond requires significantly less energy (≈0.5 kcal/mol). Thus, the closed conformation (stabilized by H-bonding with bound O₂ ligand) is expected to be more probable than the open conformation.

Considering a possible spin transition from the Ms=1/2 state (Weiss model) to the Ms=3/2 state (spin-flipped Weiss model), the process is slightly more accessible when the
Figure 5.8: Potential energy profiles as a function of the Y-W torsional angle (MYW opening coordinate) for Arg-vicinal R(MYW)···oxyheme states. Four computed spin states for the stable MYW-closed oxyheme structure (B): two $S=3/2$ (black and gray), $S=5/2$ (blue) and $S=7/2$ (green) are shown on the left (negative Y-W angle). Opening of the MYW-radical leads to the stable MYW-open oxyheme structure (D) for $S=3/2$ states and the oxygen released structure (E) for $S=7/2$ and $S=7/2$. The yellow area indicates a crucial transition region where at about +25 degrees, for both $S=5/2$ and $S=7/2$ spin states, electron reconfiguration (electron transfer from oxyheme to the MYW-radical) and O$_2$ dissociation occur (shown by red arrows). The relative energy scale and color schemes are adapted from Figure 5.7.
radical is in the closed conformation with the H-bond. The energy cost of such a spin flip is about 2.3 kcal/mol (Figure 5.8 and Table 5.3), which is comparable to the energy for breaking the H-bond due to opening of the MYW-radical. Both of these changes may therefore occur in concert. However, O\textsubscript{2} release and production of ferric iron is not computationally predicted upon opening of the MYW-radical for either of the two S=3/2 states (conformer D).

For the higher spin states (S=5/2 and S=7/2), the potential energy surfaces indicate a steeper gradient. The MYW opening to reach a Y-W dihedral angle of +20 degrees, the region for breaking the H-bond (as identified for the lower spin states), requires about 4.1 kcal/mol for the S=5/2 state and 3.0 kcal/mol for the S=7/2. Clearly, the increased energy requirement for these two high spin states reflects a stronger H-bond manifested by an elongated Fe-O bond (electron donation to the Fe 3d\textsuperscript{*}z\textsuperscript{2} orbital) resulting in an increased polarization of the O\textsubscript{2} ligand (Table 5.5). Interestingly, these states are more suited to donate an electron to the nearby (MYW)\textsuperscript{•+} radical. As illustrated by the red arrows in Figure 5.8, this crucial electron reconfiguration was found at a Y-W dihedral angle of +25 degrees. At this point, R(MYW)···oxyheme is catalytically activated, reduction of the MYW radical occurs, and triplet oxygen dissociates, producing five-coordinate ferric iron. Furthermore, the energy minima for these oxygen released states (E) are found with a Y-W dihedral angle around +55 degrees or greater, in agreement with the angle in the open neutral MYW adduct (+43 degrees) as discussed above.
Table 5.5: Calculated total spin per residue in the R(MYW)···oxyheme models. Heme = O₂ + Fe + Por + His-270.

<table>
<thead>
<tr>
<th></th>
<th>Arg-remote</th>
<th>Arg-vicinal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O₂ bound</td>
<td>O₂ bound</td>
</tr>
<tr>
<td><strong>Closed R(MYW)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S=1/2</td>
<td>S=3/2</td>
<td>S=5/2</td>
</tr>
<tr>
<td>Arg</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Met</td>
<td>0.0034</td>
<td>0.0033</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.6465</td>
<td>0.6329</td>
</tr>
<tr>
<td>Trp</td>
<td>0.3466</td>
<td>0.3609</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.9965</td>
<td>0.9970</td>
</tr>
<tr>
<td>O₂</td>
<td>0.0001</td>
<td>-1.0240</td>
</tr>
<tr>
<td>Fe</td>
<td>-0.0001</td>
<td>1.0685</td>
</tr>
<tr>
<td>Por + His²⁷⁰</td>
<td>0.0035</td>
<td>-0.0415</td>
</tr>
<tr>
<td>Heme</td>
<td>0.0035</td>
<td>0.0030</td>
</tr>
<tr>
<td><strong>Open R(MYW)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S=1/2</td>
<td>S=3/2</td>
<td>S=5/2</td>
</tr>
<tr>
<td>Arg</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Met</td>
<td>0.0044</td>
<td>0.0044</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.7187</td>
<td>0.7189</td>
</tr>
<tr>
<td>Trp</td>
<td>0.2767</td>
<td>0.2765</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.9998</td>
<td>0.9999</td>
</tr>
<tr>
<td>O₂</td>
<td>0.0000</td>
<td>-1.0340</td>
</tr>
<tr>
<td>Fe</td>
<td>0.0000</td>
<td>1.0780</td>
</tr>
<tr>
<td>Por + His²⁷⁰</td>
<td>0.0002</td>
<td>-0.0438</td>
</tr>
<tr>
<td>Heme</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
Table 5.6: Calculated total charge per residue in the R(MYW)···oxyheme models. Heme = $O_2 + Fe + Por + His-270$.

<table>
<thead>
<tr>
<th>S=1/2</th>
<th>S=3/2</th>
<th>S=5/2</th>
<th>S=7/2</th>
<th>S=1/2</th>
<th>S=3/2</th>
<th>S=5/2</th>
<th>S=7/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
<td>Arg-remote</td>
</tr>
<tr>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
<td>O$_2$ bound</td>
</tr>
<tr>
<td>Arg</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9909</td>
<td>0.9907</td>
<td>0.9906</td>
</tr>
<tr>
<td>Met</td>
<td>0.9586</td>
<td>0.9567</td>
<td>0.9558</td>
<td>0.9552</td>
<td>0.9558</td>
<td>0.9583</td>
<td>0.9565</td>
</tr>
<tr>
<td>Tyr</td>
<td>-0.3040</td>
<td>-0.3159</td>
<td>-0.3219</td>
<td>-0.3260</td>
<td>-0.3212</td>
<td>-0.3518</td>
<td>-0.3643</td>
</tr>
<tr>
<td>Trp</td>
<td>0.3297</td>
<td>0.3372</td>
<td>0.3398</td>
<td>0.3305</td>
<td>0.3320</td>
<td>0.3841</td>
<td>0.3928</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>1.9843</td>
<td>1.9781</td>
<td>1.9737</td>
<td>1.9596</td>
<td>1.9666</td>
<td>1.9815</td>
<td>1.9757</td>
</tr>
<tr>
<td>$O_2$-bound</td>
<td>-0.4284</td>
<td>-0.5628</td>
<td>-0.6283</td>
<td>-0.7251</td>
<td>-0.6847</td>
<td>-0.4284</td>
<td>-0.5628</td>
</tr>
<tr>
<td>Fe</td>
<td>1.2989</td>
<td>1.4319</td>
<td>1.4507</td>
<td>1.6806</td>
<td>1.9762</td>
<td>1.2988</td>
<td>1.4320</td>
</tr>
<tr>
<td>Por + His$^{270}$</td>
<td>-2.8548</td>
<td>-2.8472</td>
<td>-2.7961</td>
<td>-2.9152</td>
<td>-3.2581</td>
<td>-2.8519</td>
<td>-2.8450</td>
</tr>
<tr>
<td>Heme</td>
<td>-1.9843</td>
<td>-1.9781</td>
<td>-1.9737</td>
<td>-1.9597</td>
<td>-1.9666</td>
<td>-1.9815</td>
<td>-1.9757</td>
</tr>
<tr>
<td>Open R(MYW)$^{++}$</td>
<td>O$_2$ bound</td>
<td>Open R(MYW)$^{++}$</td>
<td>O$_2$ release</td>
<td>Open R(MYW)$^{++}$</td>
<td>O$_2$ release</td>
<td>Open R(MYW)$^{++}$</td>
<td>O$_2$ release</td>
</tr>
<tr>
<td>Arg</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9881</td>
<td>0.9883</td>
<td>0.9880</td>
</tr>
<tr>
<td>Met</td>
<td>0.9685</td>
<td>0.9685</td>
<td>0.9685</td>
<td>0.9685</td>
<td>0.9704</td>
<td>0.9700</td>
<td>0.9705</td>
</tr>
<tr>
<td>Tyr</td>
<td>-0.2511</td>
<td>-0.2509</td>
<td>-0.2514</td>
<td>-0.2507</td>
<td>-0.2577</td>
<td>-0.2849</td>
<td>-0.2888</td>
</tr>
<tr>
<td>Trp</td>
<td>0.2776</td>
<td>0.2774</td>
<td>0.2779</td>
<td>0.2763</td>
<td>0.2794</td>
<td>0.3215</td>
<td>0.3248</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>1.9950</td>
<td>1.9950</td>
<td>1.9950</td>
<td>1.9941</td>
<td>1.9894</td>
<td>1.9951</td>
<td>1.9943</td>
</tr>
<tr>
<td>$O_2$-bound</td>
<td>-0.3937</td>
<td>-0.5322</td>
<td>-0.6001</td>
<td>-0.6681</td>
<td>-0.4428</td>
<td>-0.3938</td>
<td>-0.5318</td>
</tr>
<tr>
<td>Fe</td>
<td>1.2860</td>
<td>1.4188</td>
<td>1.4381</td>
<td>1.6480</td>
<td>1.8504</td>
<td>1.2860</td>
<td>1.4185</td>
</tr>
<tr>
<td>Por + His$^{270}$</td>
<td>-2.8873</td>
<td>-2.8816</td>
<td>-2.8329</td>
<td>-2.9741</td>
<td>-3.3971</td>
<td>-2.8873</td>
<td>-2.8810</td>
</tr>
<tr>
<td>Heme</td>
<td>-1.9950</td>
<td>-1.9950</td>
<td>-1.9950</td>
<td>-1.9942</td>
<td>-1.9894</td>
<td>-1.9951</td>
<td>-1.9943</td>
</tr>
</tbody>
</table>
For comparison, a similar energy profile with the catalytically inert Arg-remote configuration is presented in Figure 5.9 (Both graphs exclude the oxyferrous state, Pauling model.) It is clear from this analysis that opening of the MYW adduct radical which Arg-418 assumes a remote position will not lead to the catalase products.

Figure 5.9: Potential energy profiles as a function of the Y-W torsional angle (MYW opening coordinate) for Arg-remote R(MYW)···oxyheme states. Four computed spin states for the stable MYW-closed oxyheme structure (B): two S=3/2 (black and gray), S=5/2 (blue) and S=7/2 (green) are shown on the left (negative Y-W angle). The relative energy scale and color schemes are adapted from Figure 5.7.

To understand the role of MYW-radical opening in the electron transfer process, the relative alignment of three key MYW-oxyheme orbitals is considered: (1) the SOMO of π-character (π*Y-W), (2) the inner oxygen lone pair orbital (sp² hybrid), and (3) a hydrogen
s-orbital involved in H-bonding of σ-character to the Fe-O₂ inner oxygen lone pair orbital (Scheme 5.5). For the closed MYW conformation with an optimal H-bonding contact, the overlap of the inner oxygen lone pair with the MYW radical SOMO at the indole nitrogen is negligible (insignificant sp² - π*⁻Y⁻W orbital mixing). As a result, electron transfer from the sp² hybrid to the MYW-radical is not expected (Closed stage in Scheme 5.5). As the Trp-107 indole nitrogen moves towards the heme plane, the H-bond donation weakens while the overlap of the Fe-O₂ inner oxygen lone pair orbital and the SOMO of MYW increases (increased mixing of sp² - π*⁻Y⁻W orbitals), permitting an effective electron transfer (TS stage in Scheme 5.5) for S=5/2 and S=7/2 states. For these electronic states, an electron re-configuration indicating electron transfer to the MYW radical is predicted at the partially-open conformation of around +25 degrees, which will be examined in further detail below. The post-electron-transfer system, with total spin conserved in the S=5/2 or S=7/2 configuration, falls into a lower energy state with an open Y-W dihedral angle, as molecular oxygen diffuses away from the ferric heme (Open stage in Scheme 5.5).

Scheme 5.5: Proposed stages of the MYW-radical opening process illustrating the effects of reorientation of the SOMO (π*⁻Y-W) orbital and a hydrogen 1s orbital (both in red) with respect to the sp² hybrid (in blue) of the inner oxygen lone pair of oxyheme. (Closed) Optimal (1s-sp²) orbital overlap and strong H-bonding for the closed MYW-radical, (TS) development of (π*⁻Y-W-sp²) orbital mixing and weakened H-bonding for the transitional partially-open MYW-radical that leads to electron transfer for catalytically active spin states: S=5/2 and S=7/2, (Open) catalase reaction products: neutral open MYW adduct, 5-coordinate ferric-heme and molecular oxygen.
While it is now clear that the MYW-opening process plays a key role in the \( \text{O}_2 \) release mechanism, the transition also strongly couples with electron transfer and spin-crossing processes. Thus, the computed one-dimensional energy profile along the opening coordinate alone is not sufficient to provide a realistic physical view of the mechanism of oxygen release. Monitoring the coordinate along the \( \text{Fe-O}_2 \) bond length is required to complete the energy landscape of the MYW-radical opening. For the size of this molecular system and level of theory, sampling a comprehensive two-dimensional representation of the energy surface along the two coordinates is exceptionally expensive and has not been attempted here. Instead, a detailed examination of the one-dimensional energy landscape along the \( \text{Fe-O} \) dissociation coordinate is presented next, revealing a number of new findings about structural and electronic changes along this catalytic coordinate.

**Oxygen dissociation coordinate**

Analogous to the MYW-radical opening coordinate, the energy landscape along the oxygen dissociation coordinate (\( \text{Fe-O} \) distance) was explored and is presented in Figure 5.10. The \( \text{Fe-O} \) energy profiles are shown for Arg-vicinal structures in the four lowest energy states (both \( S=3/2 \) spin states and two high spin states, \( S=5/2 \) and \( S=7/2 \)) and two relevant conformations of the MYW-radical. The first conformation is the energetically stable closed conformation (solid lines and full circles), and the second is the transitional partially-open structure with a Y-W torsional angle of +25-degrees (dashed lines and half-filled circles) shown above to be critical for initiation of the electron transfer process (Scheme 5.5 and Figure 5.8). Starting with the energy minimum found for each spin state, the structures were perturbed by imposing incremental changes on the \( \text{Fe-O} \) bond, and re-optimized with a fixed \( \text{Fe-O} \) bond length. The same protocol was implemented to explore the energy landscape in the transitional partially-open structure, but re-optimization was performed with an additional constraint that fixed the Y-W torsional angle at the critical opening stage (+25 degrees).
Figure 5.10: Potential energy profiles as a function of Fe-O₂ distance (oxygen-dissociation coordinate) for Arg-vicinal R(MYW)···oxyheme spin states for two conformations of the MYW-radical: (1) closed conformation (solid lines and filled circles), and (2) transitional structure (partially-opened) with Y-W torsion angle fixed at +25 degrees (dashed lines and half-filled circles). The yellow area indicates a crucial region of Fe-O distance where spin-crossing between the ground and the high-spin state accompanied by electron transfer occur. Spin-crossings for the S=5/2 state are indicated for closed MYW-radical conformations (large circle, solid blue line) and for partially-opened MYW-radical conformations (large circle, dashed blue line). The relative energy scale and color schemes are adapted from Figure 5.7.
For the energy profiles of the closed MYW-radical conformation with Arg-vicinal (Figure 5.10, solid lines with filled circles), the catalytically inactive, oxygen bound structures as described above (structures B) correspond to the local minima found for each of the four spin states. In the considered range of Fe-O distances (1.7 Å - 2.4 Å), both S=3/2 states (black and gray) show steep energy gradients, while the energetically higher S=5/2 and S=7/2 states show relatively flatter energy gradients, especially as the Fe-O bond length is stretched. Thus, at a sufficiently elongated Fe-O bond length, the S=3/2 states are expected to cross with the higher spin states. In particular, such a crossing point between the ground state and the S=5/2 state is seen at an Fe-O distance of 2.25 Å (Figure 5.10, solid blue large circle). This intersection is accompanied by an electron re-configuration (electron transfer from oxyheme to the MYW radical, indicated by a solid red arrow), that weakens O₂ binding as the system falls into a new (spin-conserved) potential energy profile leading to the minimum at which dissociated oxygen, ferric heme and the net neutral MYW adduct is found. An analogous spin crossing between the ground state and the energetically higher S=7/2 state is expected at an even longer Fe-O distance. At the limits of the plotted Fe-O distance (2.4 Å), the energy profile for the S=7/2 state still displays no clear signs of the electron re-configuration, and the potential energy profile maintains the MYW-radical and weakly bound oxygen. In this case, the spin crossing point is expected to fall outside the plotted range that can be extrapolated from the two energy profiles to occur at or greater than 2.5 Å. These estimations are in good agreement with other extensive DFT reports on the Fe-O multi-spin energy profiles for O₂ binding to ferrous heme.¹⁶⁰

Thus, retaining the closed conformation of the MYW-radical, the lowest energy barrier to release molecular oxygen from the system (catalytically activate the complex) is computed at about 7.5 kcal/mol. The dissociation process requires stretching of the Fe-O bond by a large amount, from 1.91 Å (the ground state minimum) to at least 2.25 Å where the ground state crosses with the S=5/2 spin state and couples with an electron transfer from oxyheme to the MYW radical.
The Fe-O energy profiles for the transitional partially-open structure with the Y-W torsion fixed at +25 degrees (Figure 5.10, dashed lines with half-filled circles) disclose similar trends when compared to the corresponding curves for the closed conformation. These profiles highlight the significance of MYW-radical opening for catalytic oxygen release from the oxyheme intermediate. In the range of Fe-O distance shorter than 2.1 Å, all computed energy profiles support the electronic configurations that maintain the MYW-radical. As expected from the MYW opening coordinate, the energy profiles for the partially-open MYW structure lie above the energy profiles for the closed conformation when the MYW-radical is present. In this range of Fe-O distance for the transitional partially-open structure, consistent with MYW-opening profiles, stable oxygen bound states are identified only for the two S=3/2 spin states. For the other two energy profiles of the partially-open structure (S=5/2 and S=7/2 spin states), before the stable oxygen bound states are not found, instead the system reaches the stable oxygen bound minima at Fe-O distance around 2.1 Å of Fe-O bond length, an electron re-configuration (electron transfer to the MYW-radical, indicated by a dashed red arrow) is observed. This electron re-configuration occurs at a significantly shorter distance than the 2.25 Å observed for the closed MYW conformation described above. These observations can be traced to the favorable orientation of the SOMO ($\pi^*_{Y-W}$) orbital in the partially-open conformation that facilitates electron transfer. With an electron transferred to the MYW adduct, the new spin-conserved electronic configuration for both S=5/2 and S=7/2 spin states lowers the energy of the system and leads to the minima with dissociated oxygen, ferric heme and the net neutral MYW adduct. Besides the electron re-configuration at a shorter Fe-O distance, the energy barrier for the catalytic process is also lower for the partially-open conformation. In this case, the lowest energy barrier identified by the spin crossing point (Figure 5.10, dashed large blue circle) between the ground state and the S=5/2 state, is found at about 5.6 kcal/mol above the ground state. These findings strongly suggest that the opening of the MYW radical effectively decreases the requirement for Fe-O bond stretching and reduces the energy barrier for the catalytic event. Thus, the final step in
the proposed catalase mechanism that involves oxyheme as an intermediate in KatG likely depends on both Fe-O dissociation and MYW-radical opening coordinates.

Another important outcome is observed for compression of the Fe-O bond when comparing profiles for closed and critical partially-open MYW structures. As the Fe-O bond length decreases, the energy difference between a pair of curves (solid vs. dashed lines of the same color/spin state) also decreases, illustrating that the small energy barrier to open the MYW radical is further reduced. This trend can be traced to a concerted motion along MYW opening and Fe-O dissociation coordinates. As the radical opens and the indole NH of Trp moves towards the heme plane, maintenance of the H-bond with oxygen brings the ligand closer to iron. The shortened Fe-O bond shifts the electron configuration of oxyheme towards the Pauling-like Fe(II)-O$_2$ model that decreases the total charge on oxygen. The decreased charge weakens the H-bonding interaction and reduces the barrier to open the radical. In fact, this concerted motion was identified as a low frequency vibrational mode computed to be around 224 cm$^{-1}$ for the ground state. The energy of this vibrational mode that symmetrically couples the two deformations, Fe-O stretching and Y-W torsion, can assist the transition of the system from the MYW-$\text{closed}$ to MYW-$\text{open}$ conformation, which will be discussed in greater detail below.

For comparison, a similar energy profile with the catalytically inert Arg-$\text{remote}$ conformation is presented in Figure 5.11 (excluding the oxyferrous state, Pauling model). The Fe-O energy profiles are shown for Arg-remote structures in the energetically stable closed conformation (solid lines and full square), and in the transitional partially-open conformation with a Y-W torsional angle of +25-degrees (dashed lines and half-filled squares). As expected from Figure 5.7, all energy profiles with the closed conformation are nearly equal to those of the Arg-vicinal low spin states in both the corresponding closed conformations. The release of molecular oxygen is observed for the two high spin states (S=5/2 and S=7/2) but only when the Fe-O bond is stretched to a large distance: 2.2Å for the S=5/2 state and 2.4Å for the S=7/2 state.
Figure 5.11: Potential energy profiles as a function of Fe-O₂ distance (oxygen-dissociation coordinate) for Arg-remote R(MYW)···oxyheme spin states for two conformations of the MYW-radical: (1) closed conformation (solid lines and filled squares), and (2) transitional structure (partially-opened) with Y-W torsion angle fixed at +25 degrees (dashed lines and half-filled square). The yellow area indicates a crucial region of Fe-O distance where spin-crossing between the ground and the high-spin state accompanied by electron transfer occur. The relative energy scale and color schemes are adapted from Figure 5.7.
**Electron Transfer Process**

The structural and electronic features of the electron transfer followed by the oxygen release process can be comprehensively viewed in terms of the changes in the molecular orbital interactions contributing to the covalency of the Fe-O$_2$ bond (Scheme 5.3). In the closed conformation of the R(MYW)···oxyheme model, these orbitals are expected to be strongly polarized by H-bonding with the inner oxygen of the Fe-O$_2$ complex. The polarization effect stabilizes the sp$^2$ hybrid formed on the inner oxygen atom and destabilizes the bonding ($3d^*_{z^2}$-$\pi^*_x$) orbital by decreasing the electron donation to the $3d^*_z$ orbital. This polarization effect is schematically illustrated by the blue arrows in Figure 5.12 where the left and right panels represent the closed and open conformations, respectively. These changes are also reproduced in the computed average natural charges for atoms in the oxyheme group (Table 5.6). The polarization of charge occurs mainly along the Fe-O$_2$ bond, as the average charge on Fe changes from +1.57e to +1.47e, and on the O$_2$ ligand from -0.61e to -0.52e for closed and open conformations, respectively. Thus, the increased electron donation to the Fe $3d^*_z$ orbital effectively influences the energetics of the ($3d^*_{z^2}$-$\pi^*_x$) bonding (shown in Figure 5.12) and anti-bonding (not shown) orbitals. Also, as described in detail above, polarization of the MYW-radical by the Arg-418 vicinal interaction (Table 5.5 and Table 5.6, Figure 5.13 and Figure 5.14) is expected to stabilize the SOMO orbital ($\pi^*_{Y-W}$).

![Figure 5.12: Stabilization of the ($d^*_{z^2}$-$\pi^*_x$) bonding orbital of the Fe-O$_2$ complex modulated by the MYW-radical opening process in KatG.](image-url)
Figure 5.13: Bar representation of calculated atomic charge values for the MYW-radical in R(MYW)···oxyheme models for both $S=3/2$ spin states; ($S=3/2$, $M_s=3/2$) and ($S=3/2$, $M_s=1/2$). For clarity the MYW drawings are only skeletal.
Figure 5.14: Representation of calculated atomic charge values for the MYW-radical in R(MYW)···oxyheme models for two high spin states; S=5/2 and S=7/2. For clarity the MYW drawings are only skeletal.
A complete energy level diagram for the molecular orbitals of the \( R(MYW) \cdots \text{oxyheme} \) model (eight orbitals and nine electrons) is schematically shown in Scheme 5.6 and Scheme 5.7. Two spin states are shown: the catalytically inert \( S=3/2 \) ground state (Scheme 5.6), and the catalytically active \( S=5/2 \) high spin state (Scheme 5.7). Both bonding and anti-bonding \( (3d^{*}_{2z}-\pi^{*}_{\parallel}) \) orbital energy changes are indicated by dashed lines between the Fe-O\(_{2}\) orbitals for closed and open conformations (left to right) and between the energy of the \( (\pi Y-W^{*}) \) orbital with and without the Arg interaction, which is shown here as occupied only for Arg-vicinal cases. These changes arise from the H-bonding interaction discussed earlier in the text.

As discussed above, the three unpaired electrons of the \( S=3/2 \) low spin state populate the valence orbitals according to the Weiss model \( (R(MYW)(S=1/2) \cdots Fe(S=1/2)-O_{2}(S=1/2)), \) and effectively bind the superoxide ion to the ferric heme (Scheme 5.6). For this spin state, the orbital changes comparing closed (H-bond) to open (no H-bond) cases do not impact the electron configuration and the Fe-O bond remains intact. For the open conformation, the Fe-O\(_{2}\) complex is slightly destabilized due to reduced synergic bonding forces (decreased \( (3d^{*}_{2z}-\pi^{*}_{\parallel}) \) and \( (3d_{xz}-\pi^{*}_{\perp}) \) orbital mixing). In addition, DFT computed orbital energies and
their electron occupations are shown for these two $S=3/2$ spin states in Figure 5.15 and Figure 5.16.

Figure 5.15: Computed orbital energies, electron occupations and spin density drawings (green for positive and purple for negative spin) for four adduct conformations (MYW-closed/Arg-remote; MYW-open/Arg-remote; MYW-closed/Arg-vicinal; MYW-open/Arg-vicinal) in the $(S=3/2, M_s=1/2)$ spin state.
Figure 5.16: Computed orbital energies, electron occupations and spin density drawings (green for positive and purple for negative spin) for four adduct conformations (MYW-closed/Arg-remote; MYW-open/Arg-remote; MYW-closed/Arg-vicinal; MYW-open/Arg-vicinal) in the (S=3/2, Ms=3/2) spin state.
The S=5/2 high-spin state \((R(MYW)(S=1/2)\cdots Fe(S=1)-O_2(S=1))\) can be viewed as a product of a single electron excitation from the S=3/2 spin state, altering the configuration from \((3d_{z^2}^*-\pi_{\parallel}^*)^2\) to \((3d_{z^2}^*-\pi_{\parallel}^*)^1(3d_{z^2}^*-\pi_{\parallel}^*)^1\) (Scheme 5.7, left panel, blue arrow). This state corresponds to the double-spin flipped McClure-Goddard, ozone-like binding model containing an intermediate-spin (S=1) ferrous ion, ferromagnetically coupled to \(O_2\) (S=1). For this electron configuration with the closed conformation, the Fe-O\(_2\) complex weakens but the ligand remains bound mainly due to the sustained \((3d_{z^2}-\sigma)\) and \((3d_{xz}-\pi^*_\perp)\) orbital mixing and reduced anti-bonding interactions with the \(d_{yz}\) orbital when the lone pair orbital (sp\(^2\) hybrid) is polarized by H-bonding. However, for the open MYW-radical conformation (Scheme 5.7, right panel), the increased electron donation to \(3d_{z^2}^*\) destabilizes the singly-occupied anti-bonding \((3d_{z^2}^*-\pi_{\parallel}^*)^*\) orbital in the absence of H-bonding. Consequently, the \((3d_{z^2}^*-\pi_{\parallel}^*)^*\) orbital energy is greater than the \((\pi_{Y-W}^*)\) orbital, which is stabilized by the Arg-vicinal interaction. In addition, the SOMO orbital is now favorably oriented to maximize its overlap with the lone pair orbital of the inner oxygen (as described earlier). Thus, as the MYW-radical opens, the initial \((3d_{z^2}^*-\pi_{\parallel}^*)^1(\pi_{Y-W}^*)^1\) configuration becomes less stable and changes to the \((\pi_{Y-W}^*)^2\) configuration indicative of an electron transfer from the oxyheme to the MYW-radical (Scheme 5.7, right panel, red arrow). At the same time, in order to conserve the S=5/2 spin state, the electron must be promoted from \((d_{yz})^2\) to the \((d_{yz})^1(3d_{z^2}^*-\pi_{\parallel}^*)^1\) symmetry conserved configuration (Scheme 5.7, right panel, blue arrow). As a result, the Fe-O\(_2\) bond is unstable and molecular oxygen (S=1) dissociates, leaving the ferric-heme resting state (S=3/2) that poises the enzyme for another catalytic cycle. An analogous mechanism can be depicted for the S=7/2 state \((R(MYW)(S=1/2)\cdots Fe(S=2)-O_2(S=1))\), see Scheme 5.8, which also leads to the release of molecular \(O_2\) (S=1) and the resting high-spin ferric-heme \((S=5/2)\). DFT computed orbital energies and electron occupation for the two S=5/2 and S=7/2 high spin states are also shown in Figure 5.17 and Figure 5.18, respectively. It is worth mentioning that experimentally, the resting enzyme exhibits both an intermediate (S=3/2) and high-spin (S=5/2) state for 5-coordinate heme and may be viewed as a quantum-mixed state.
spin heme. Detailed constitution of the mixed spin state of the resting ferric-heme was not explored in this work.

Scheme 5.7: Molecular orbital diagram of R(MYW)···oxyheme models in high ($S=5/2$) spin states. The dotted lines show two relevant effects: (1) (de)stabilization of the two (3$d_{z^2}$-$\pi^*$) orbitals for closed and open MYW-radical, (2) stabilization of the $\pi^*_Y-W$ orbital for the Arg-vicinal compared to Arg-remote positions. See text for more details.

Scheme 5.8: Molecular orbital diagram of R(MYW)···oxyheme models in the $S=7/2$ spin state. The dotted lines show: (1) (de)stabilization of the Fe-O orbitals for closed and open MYW-radicals, (2) stabilization of the $\pi^*_Y-W$ orbital for the Arg-vicinal compared to the Arg-remote positions. See text for more details.
Figure 5.17: Computed orbital energies, electron occupations and spin density drawings (green for positive and purple for negative spin) for four adduct conformations (MYW-closed/Arg-remote; MYW-open/Arg-remote; MYW-closed/Arg-vicinal; MYW-open/Arg-vicinal) in the \((S=5/2, M_s=5/2)\) spin state.
Figure 5.18: Computed orbital energies, electron occupations and spin density drawings (green for positive and purple for negative spin) for four adduct conformations (MYW-closed/Arg-remote; MYW-open/Arg-remote; MYW-closed/Arg-vicinal; MYW-open/Arg-vicinal) in the \((S=7/2, Ms=7/2)\) spin state.
5.1.3 Catalase Reaction - Mechanism of Dioxygen Release

The DFT study of the isolated R(MYW) and R(MYW)···oxyheme models, incorporating geometry optimizations and sampling the potential energy surfaces, provided characterization of a novel catalytic step that involves both an MYW-radical and oxyheme in the KatG catalase reaction. The intrinsic structural plasticity of the MYW-radical allowing the closed and open conformations contributes to catalytic oxygen release in KatG.

In the proposed oxygen release mechanism, the oxyheme intermediate cannot complete the pathway when it is partnered with the closed MYW-radical conformation; the MYW-radical must open to accept an electron. Two key factors apparently assist the MYW-radical opening process. First, the H-bonding interaction of the tyrosyl oxygen with the Arg-418 side-chain, and second, the low frequency vibrational mode of the Fe-O stretch coupled with the Y-W torsional deformation aligned with the MYW opening coordinate.

The results showed that after the MYW-radical is formed with the Arg in the remote position, the energy difference for the two Arg-418 side-chain positions is very small (computed around 0.2 kcal/mole), suggesting a nearly equal population of Arg side-chain sites in the enzyme. Thus, these calculations support the view that while almost the entire enzyme population can harbor oxyheme, only half the enzyme population may be poised for the catalytic event we describe. This agrees with the quantification of the MYW-adduct radical by EPR spectroscopy, which showed approximately 0.5 spins per heme (one-half equivalent of MYW-radical) present during steady-state catalase turnover by KatG. This is expected if the MYW radical does not persist in enzyme molecules having the Arg side-chain in the vicinal position and further suggests that the MYW-radical species observed experimentally represents those enzyme molecules awaiting Arg side-chain reorientation.

After the MYW-radical is formed with the Arg in the remote position, an equal distribution of Arg remote ⇌ vicinal is attained. Based on these results, a mechanism starting with Arg in the vicinal position can be proposed that includes MYW-radical opening leading to electron transfer and molecular oxygen release (Scheme 5.9). Here, considering the closed
MYW-radical conformation in the ground state (S=3/2), the vibrational mode of concerted motion will induce both compression of Fe-O and the MYW opening. The compression of Fe-O (that promotes the Pauling model, Fe(II)-O$_2$) reduces the negative charge on the O$_2$ ligand, and weakens the H-bonding with Trp-107. At the stage of an H-bond breakdown with the oxyhemoglobin group, the Fe-O compression and Y-W torsion decouple, and the MYW-radical continues along a favorable trajectory, increasing the Y-W dihedral angle. With the H-bond now released, the Fe-O bond decompresses, and can achieve a larger stretching amplitude due to a change in the reduced mass of the vibrational mode. When the Fe-O bond is stretched to about 2.1 Å (that promotes the Weiss model, Fe(III)-O$_2^-$), the negative charge on the O$_2$ ligand increases to a critical amount allowing the ground state (S=3/2) to spin-cross with the high spin states, particularly with the S=5/2 state. In the S=5/2 state, at around the +25-degree Y-W torsional angle, the MYW-radical can accept an electron, and molecular oxygen is released as a product, completing the catalytic event. The system requires about 5.6 kcal/mole starting from the closed-radical ground state to rotate the Trp-107 plane, displace the Fe-O distance to about 2.1Å, and release the products.

Scheme 5.9: Pathway for MYW-radical opening assisted by Arg-vicinal and the vibrational mode of Fe-O stretching (blue arrow) coupled with the Y-W torsion (blue curved arrow) leading to H-bond breaking (wavy line) and decoupling of Fe-O stretching with Y-W torsion, followed by electron transfer (red arrow), and oxygen release.

Finally, it is important to note that the computational study presented here suggests that the neutral (resting state) MYW adduct after the oxygen release would maintain its open conformation. This is in disagreement with most available crystallographic structures.
of KatG (see Table 5.7) and could be considered a limitation of the current model that needs to be addressed in the future. However, a reasonable picture of the closing event could be put forward based on current limited results. It was found that the energy required to close the MYW-radical (around 4 kcal/mole) is half the energy gained during the oxygen release (around 9.6 kcal/mole) and certainly a part of the released energy could be distributed to close the neutral MYW adduct. The MYW adduct may remain closed due to a metastable state (the shallow minimum) found exclusively for the neutral (resting) MYW adduct. Furthermore, the stability of the closed conformation is expected to increase with protonation of the MYW adduct which could lock the MYW conformation, in agreement with crystallographic data.

Table 5.7: Comparison of torsional angles in X-ray crystal structures of KatGs: *M. tuberculosis* (2CCA, 1SJ2), *S. elongatus* (1UB2), *H. marismortui* (1ITK) and *B. pseudomallei* (3N3P, 1MWV, 2FXG, 2FXH, 2FXJ). a) torsion measured between C$_\alpha$, C$_\beta$, C$_\gamma$, C$_\delta$; b) dihedral angle measured between Tyr C$_\zeta$, C$_\epsilon_1$ and Trp C$_\eta_2$, C$_\zeta_3$.  

<table>
<thead>
<tr>
<th>pdb code</th>
<th>Resolution (Å)</th>
<th>W C$_\alpha$ tors$^a$ (degrees)</th>
<th>Y C$_\alpha$ tors$^a$ (degrees)</th>
<th>Y-W dih$^b$ (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2CCA</td>
<td>2.00</td>
<td>-116.7</td>
<td>119.6</td>
<td>-12.3</td>
</tr>
<tr>
<td>1SJ2</td>
<td>2.41</td>
<td>-117.3</td>
<td>119.4</td>
<td>9.3</td>
</tr>
<tr>
<td>1UB2</td>
<td>2.40</td>
<td>-138.1</td>
<td>108.8</td>
<td>-1.4</td>
</tr>
<tr>
<td>1ITK</td>
<td>2.00</td>
<td>-121.4</td>
<td>124.3</td>
<td>-19.1</td>
</tr>
<tr>
<td>3N3P</td>
<td>1.89</td>
<td>-110.2</td>
<td>118.9</td>
<td>-20.2</td>
</tr>
<tr>
<td>1MWV</td>
<td>1.70</td>
<td>-120.9</td>
<td>118.5</td>
<td>-17.3</td>
</tr>
<tr>
<td>2FXG</td>
<td>2.00</td>
<td>-116.0</td>
<td>122.7</td>
<td>-26.2</td>
</tr>
<tr>
<td>2FXH</td>
<td>1.90</td>
<td>-116.3</td>
<td>115.5</td>
<td>-10.1</td>
</tr>
<tr>
<td>2FXJ</td>
<td>1.95</td>
<td>-120.1</td>
<td>108.6</td>
<td>-3.1</td>
</tr>
<tr>
<td>ave</td>
<td></td>
<td>-119.7</td>
<td>117.4</td>
<td>-11.1</td>
</tr>
<tr>
<td>stdev</td>
<td></td>
<td>7.7</td>
<td>5.5</td>
<td>11.1</td>
</tr>
</tbody>
</table>
5.2 The g-tensor as a function of the MYW opening coordinate

While the comprehensive density functional theory study in this chapter provides strong evidence that the opening of the MYW adduct could be involved in the oxygen release mechanism of KatG, experimental evidence is needed to confirm the role of the MYW adduct. As the catalase cycle involves a radical on the MYW adduct, EPR spectroscopy is well suited to monitor the MYW radical, and could be used to provide evidence of the MYW-opening. Electron paramagnetic resonance spectroscopy (EPR) has become an increasingly important tool for studying radicals in proteins. Measurement of EPR spectra depends on the determination of the g-values for the unpaired electrons in a sample, which differ from those of free electrons and are sensitive to the chemical environment of the paramagnetic atom. Thus, the g-values are sensitive to the electronic structure of radicals, and hence, they provide a valuable probe for identifying radicals within proteins.\textsuperscript{163} For tyrosyl radicals, the g-tensor component in the direction of the C-O bond ($g_{xx}$) is commonly used to probe a tyrosyl radical’s protein environment, and is particularly sensitive to hydrogen bonding.\textsuperscript{164,165} The D-band EPR spectrum of the KatG enzyme (frozen using rapid-freeze-quench approach in the presence of mM H$_2$O$_2$), shown in Figure 5.19, exhibits a rhombic g-tensor with dual $g_{xx}$ values (2.00550 and 2.00606) and unique $g_{yy}$ (2.00344) and $g_{zz}$ (2.00186) values.\textsuperscript{132}

![Figure 5.19: The D-band EPR spectrum of the MYW radical in KatG.\textsuperscript{132}](image-url)
The g-values for Arg-remote and Arg-vicinal MYW and YW models have been computed using the B3LYP functional and 6-311G** basis set. The g-values were calculated for optimized geometries that spanned the YW dihedral angle from -60 degrees to +60 degrees in 5 degree increments. Relative energies, $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for both Arg-remote and Arg-vicinal MYW models can be found in Table 5.8, and for YW models in Table 5.8.

A scan of the opening coordinate as a function of $g_{xx}$ value for Arg-remote and Arg-vicinal MYW and YW models is shown in Figure 5.20. For all models, the $g_{xx}$ value is highest at both extremes of the scan, which represents the maximum values plotted for the open YW-dihedral angle (+60 degrees) and the closed YW-dihedral angle (-60 degrees). The $g_{xx}$ values are at a minimum when the tyrosine and tryptophan rings are nearly planar, which is likely due to an increased electrostatic interaction between the hydrogen connected to $C_3$ of Trp and the Tyr-O. This interaction polarizes the spin on the tyrosyl oxygen and modulates the $g_{xx}$ value.

![Figure 5.20: The $g_{xx}$ value as a function of the MYW opening coordinate for Arg-remote MYW (purple) and Arg-vicinal MYW (orange) models.](Image)
Table 5.8: Relative energies, $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for the scan of the MYW opening coordinate of Arg-remote and Arg-vicinal MYW models. The numbers in bold are the energetic minima (0 kcal/mol) for each structure.

<table>
<thead>
<tr>
<th>Y-W dih. (degrees)</th>
<th>Arg-remote MYW</th>
<th>Arg-remote MYW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rel. Energy (kcal/mol)</td>
<td>$g_{xx}$</td>
</tr>
<tr>
<td>60.0</td>
<td>2.46</td>
<td>2.00787</td>
</tr>
<tr>
<td>55.0</td>
<td>1.72</td>
<td>2.00779</td>
</tr>
<tr>
<td>50.0</td>
<td>1.09</td>
<td>2.00770</td>
</tr>
<tr>
<td>45.0</td>
<td>0.59</td>
<td>2.00760</td>
</tr>
<tr>
<td>40.0</td>
<td>0.23</td>
<td>2.00748</td>
</tr>
<tr>
<td>35.0</td>
<td>0.04</td>
<td>2.00736</td>
</tr>
<tr>
<td><strong>32.5</strong></td>
<td><strong>0.00</strong></td>
<td><strong>2.00726</strong></td>
</tr>
<tr>
<td>30.0</td>
<td>0.01</td>
<td>2.00722</td>
</tr>
<tr>
<td>25.0</td>
<td>0.12</td>
<td>2.00707</td>
</tr>
<tr>
<td>20.0</td>
<td>0.37</td>
<td>2.00692</td>
</tr>
<tr>
<td>15.0</td>
<td>0.72</td>
<td>2.00678</td>
</tr>
<tr>
<td>10.0</td>
<td>1.15</td>
<td>2.00664</td>
</tr>
<tr>
<td>5.0</td>
<td>1.58</td>
<td>2.00652</td>
</tr>
<tr>
<td>0.0</td>
<td>1.95</td>
<td>2.00642</td>
</tr>
<tr>
<td>-5.0</td>
<td>2.18</td>
<td>2.00638</td>
</tr>
<tr>
<td>-10.0</td>
<td>2.31</td>
<td>2.00643</td>
</tr>
<tr>
<td>-15.0</td>
<td>2.41</td>
<td>2.00650</td>
</tr>
<tr>
<td>-20.0</td>
<td>2.51</td>
<td>2.00663</td>
</tr>
<tr>
<td>-25.0</td>
<td>2.66</td>
<td>2.00678</td>
</tr>
<tr>
<td>-30.0</td>
<td>2.88</td>
<td>2.00694</td>
</tr>
<tr>
<td>-35.0</td>
<td>3.22</td>
<td>2.00710</td>
</tr>
<tr>
<td>-40.0</td>
<td>3.66</td>
<td>2.00726</td>
</tr>
<tr>
<td>-45.0</td>
<td>4.21</td>
<td>2.00740</td>
</tr>
<tr>
<td>-50.0</td>
<td>4.86</td>
<td>2.00753</td>
</tr>
<tr>
<td>-55.0</td>
<td>5.58</td>
<td>2.00766</td>
</tr>
<tr>
<td>-60.0</td>
<td>6.36</td>
<td>2.00777</td>
</tr>
</tbody>
</table>
Table 5.9: Relative energies, $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for the scan of the YW opening coordinate of Arg-remote and Arg-vicinal YW models. The numbers in bold are the energetic minima (0 kcal/mol) for each structure.

<table>
<thead>
<tr>
<th>Y-W dih. (degrees)</th>
<th>Arg-remote YW</th>
<th>Arg-vicinal YW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rel. Energy (kcal/mol)</td>
<td>$g_{xx}$</td>
</tr>
<tr>
<td>60</td>
<td>1.43</td>
<td>2.00949</td>
</tr>
<tr>
<td>55</td>
<td>0.94</td>
<td>2.00945</td>
</tr>
<tr>
<td>50</td>
<td>0.53</td>
<td>2.00941</td>
</tr>
<tr>
<td>45</td>
<td>0.21</td>
<td>2.00934</td>
</tr>
<tr>
<td>40</td>
<td>0.05</td>
<td>2.00926</td>
</tr>
<tr>
<td>38.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.725</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>0.08</td>
<td>2.00903</td>
</tr>
<tr>
<td>30</td>
<td>0.30</td>
<td>2.00890</td>
</tr>
<tr>
<td>25</td>
<td>0.62</td>
<td>2.00875</td>
</tr>
<tr>
<td>20</td>
<td>1.02</td>
<td>2.00860</td>
</tr>
<tr>
<td>15</td>
<td>1.43</td>
<td>2.00844</td>
</tr>
<tr>
<td>10</td>
<td>1.82</td>
<td>2.00832</td>
</tr>
<tr>
<td>0</td>
<td>2.09</td>
<td>2.00825</td>
</tr>
<tr>
<td>-5</td>
<td>2.16</td>
<td>2.00826</td>
</tr>
<tr>
<td>-10</td>
<td>2.07</td>
<td>2.00835</td>
</tr>
<tr>
<td>-15</td>
<td>1.89</td>
<td>2.00848</td>
</tr>
<tr>
<td>-20</td>
<td>1.71</td>
<td>2.00861</td>
</tr>
<tr>
<td>-25</td>
<td>1.57</td>
<td>2.00875</td>
</tr>
<tr>
<td>-30</td>
<td>1.54</td>
<td>2.00887</td>
</tr>
<tr>
<td>-35</td>
<td>1.61</td>
<td>2.00898</td>
</tr>
<tr>
<td>-40</td>
<td>1.81</td>
<td>2.00906</td>
</tr>
<tr>
<td>-45</td>
<td>2.14</td>
<td>2.00913</td>
</tr>
<tr>
<td>-50</td>
<td>2.60</td>
<td>2.00919</td>
</tr>
<tr>
<td>-55</td>
<td>3.16</td>
<td>2.00924</td>
</tr>
<tr>
<td>-60</td>
<td>3.78</td>
<td>2.00928</td>
</tr>
</tbody>
</table>
For all considered YW-dihedral angles, it is also interesting to note the difference in \( g_{xx} \) values between Arg-remote and Arg-vicinal pairs. For the MYW model, the difference is small, an average of \( 6 \times 10^{-4} \). The gap between the Arg-remote and Arg-vicinal positions of the YW model is larger, an average of \( 1 \times 10^{-3} \). This indicates that the arginine vicinal interaction that forms a hydrogen bond with the tyrosine oxygen is more effective at polarizing the tyrosine when the methionine is not present. One possible explanation for this is a contribution from a 1,4- interaction between the sulfur of methionine and the oxygen of tyrosine in the MYW models. Crystallographic studies have shown a wide occurrence 1,3-, 1,4-, and 1,5- interactions between non-bonded sulfur and oxygen atoms in biological systems.\(^{166,167}\) It has also been shown that the 1,4- type compounds exhibit strong intermolecular electrostatic attractive sulfur-oxygen interactions, which is directly influenced by the atom bonded to the sulfur. For example, electronegative counter atoms (X) in an X–S···O moiety increase the strength of the S-O interaction, and decrease the distance between the two atoms.\(^{166}\) This can be explained in terms of molecular orbital theory by the HOMO(\( \pi \))-LUMO(\( \sigma^* \)) orbital interaction. Several computational studies have also confirmed this finding.\(^{168–170}\) The positively charged sulfur of the -CH\(_2\)-S(CH\(_3\))\(^+\)···O\(^•\) moiety of MYW is a strongly deactivating group, and thus is expected to cause a strong interaction between the oxygen of tyrosine and the sulfur of methionine.
Chapter 6

A Unique Tyrosyl Radical: (MYW)$^\bullet+$$^\bullet$

Somewhere, something incredible is waiting to be known.

– Carl Sagan

The g-tensor component in the direction of the C-O bond ($g_{xx}$) is commonly used to probe a tyrosyl radical’s protein environment because it is most sensitive to local electrostatic effects.\textsuperscript{164} Previous experimental studies of tyrosyl radicals in several proteins have found that the g-values show a large variation in the g-tensor component in the direction of the C-O bond ($g_{xx}$), from about 2.00550 to 2.00912 (Table 6.1). Several computational studies have shown that these differences are correlated to variations in hydrogen bonding to the tyrosyl-oxygen.\textsuperscript{164,165} Gamma-radiation generated Tyr radicals in HCl crystals represent an experimental example of a hydrogen bonded tyrosyl radical with a $g_{xx}$ value around 2.00670.\textsuperscript{171} The tyrosyl radical of E. coli ribonucleotide reductase (RNR) is relatively isolated within a hydrophobic pocket,\textsuperscript{172,173} and serves as an example of a non-hydrogen bonded tyrosyl radical (2.00866).\textsuperscript{164} The tyrosyl radical formed in photosystem II (PSII) has also been shown to be hydrogen bonded,\textsuperscript{174–177} ($g_{xx} = 2.00745$), and represents an intermediate situation between the non-hydrogen-bonded case of RNR and the strongly hydrogen bonded situation in Tyr-HCl. The D-band EPR spectrum of the Catalase-Peroxidase (KatG) enzyme exhibits a rhombic g-tensor with dual $g_{xx}$ values (2.00550 and 2.00606) and unique $g_{yy}$
The two dual g values are likely due to the mobile Arg side-chain that forms a hydrogen bond with the tyrosyl-oxygen, described earlier in Chapter 5. Without the arginine interaction, one would expect the g value to be similar to that of a non-hydrogen bonded tyrosyl radical such as E. coli RNR. However, both values (2.00550 and 2.00606) would suggest strong hydrogen bonding with the tyrosyl radical. Both of these values are lower than the g value of an experimentally demonstrated strong hydrogen bond to a tyrosyl radical (Tyr-HCl, g = 2.0067), and deviate from the typical behavior of tyrosyl radicals. Other covalently modified tyrosyl radicals do not show as great of a deviation away from typical values. For example, high-frequency EPR spectroscopy of the Cys-Tyr radical in Galactose Oxidase exhibits a g value of 2.00741. This suggests that factors beyond local electrostatic effects along the C-O bond impact the g value of the MYW radical found in the catalase intermediate of KatG.

6.1 Modeling of the g-tensor in MYW

The goal of this study is not to reproduce experimentally determined g-values, but to reproduce trends in experimental g-values, particularly those with hydrogen bonding effects, and to ultimately understand their origins. To do so, a comprehensive set of density functional theory calculations have been performed to extract g-values from new models created from fragments of the MYW adduct. All structures were optimized at the level of the B3LYP functional combined with 6-311G** basis sets. The models considered are the R(MYW) and MYW radicals, the YW radical fragment, the MY radical fragment, and tyrosine with and without the arginine interaction, R(Y) and Y, respectively. As discussed in Chapter 5.1.1, the optimized geometries of R(MYW) and MYW correspond to an open (positive) YW-dihedral angle. Computed g values for both gas phase and PCM models are shown in Figure 6.1. As discussed above, the greatest deviation from the free electron value is expected for the g parameter. For gas-phase models, the R(MYW) g value is in excellent agreement with
Table 6.1: Experimentally and computationally determined g-values for tyrosyl radicals.

<table>
<thead>
<tr>
<th>Species</th>
<th>$g_{xx}$</th>
<th>$g_{yy}$</th>
<th>$g_{zz}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalase Peroxidase [MYW]\textsuperscript{132}</td>
<td>2.0055/2.00606</td>
<td>2.00344</td>
<td>2.00186</td>
</tr>
<tr>
<td>E. coli RNR\textsuperscript{164}</td>
<td>2.00866</td>
<td>2.00423</td>
<td>2.002</td>
</tr>
<tr>
<td>E. coli RNR WT\textsuperscript{179}</td>
<td>2.00895</td>
<td>2.00438</td>
<td>2.00208</td>
</tr>
<tr>
<td>E. coli RNR I74R\textsuperscript{179}</td>
<td>2.00888</td>
<td>2.0044</td>
<td>2.00205</td>
</tr>
<tr>
<td>E. coli RNR I74K\textsuperscript{179}</td>
<td>2.00884</td>
<td>2.00438</td>
<td>2.00203</td>
</tr>
<tr>
<td>E. coli RNR I74F\textsuperscript{179}</td>
<td>2.00891</td>
<td>2.00438</td>
<td>2.00203</td>
</tr>
<tr>
<td>E. coli RNR L77F\textsuperscript{179}</td>
<td>2.00884</td>
<td>2.0043</td>
<td>2.00205</td>
</tr>
<tr>
<td>E. coli RNR V136R\textsuperscript{179}</td>
<td>2.009</td>
<td>2.00444</td>
<td>2.00209</td>
</tr>
<tr>
<td>E. coli RNR V136N\textsuperscript{179}</td>
<td>2.00898</td>
<td>2.00449</td>
<td>2.00213</td>
</tr>
<tr>
<td>E. coli RNR (calc)\textsuperscript{180}</td>
<td>2.01206</td>
<td>2.00496</td>
<td>2.00219</td>
</tr>
<tr>
<td>E. coli RNR (calc)\textsuperscript{180}</td>
<td>2.01012</td>
<td>2.00486</td>
<td>2.0022</td>
</tr>
<tr>
<td>E. coli RDPR\textsuperscript{181}</td>
<td>2.00912</td>
<td>2.00457</td>
<td>2.00225</td>
</tr>
<tr>
<td>mouse RNR\textsuperscript{182}</td>
<td>2.0076</td>
<td>2.0043</td>
<td>2.0022</td>
</tr>
<tr>
<td>mouse RNR (calc)\textsuperscript{180}</td>
<td>2.01173</td>
<td>2.00495</td>
<td>2.00215</td>
</tr>
<tr>
<td>S. typhimurium RNR\textsuperscript{183}</td>
<td>2.0089</td>
<td>2.0043</td>
<td>2.0021</td>
</tr>
<tr>
<td>S. typhimurium RNR (calc)\textsuperscript{180}</td>
<td>2.01012</td>
<td>2.00486</td>
<td>2.0022</td>
</tr>
<tr>
<td>HSV1 RNR\textsuperscript{182}</td>
<td>2.0076</td>
<td>2.0043</td>
<td>2.0022</td>
</tr>
<tr>
<td>PSII\textsuperscript{164}</td>
<td>2.00745</td>
<td>2.00422</td>
<td>2.00211</td>
</tr>
<tr>
<td>PSII\textsuperscript{184}</td>
<td>2.00745</td>
<td>2.00422</td>
<td>2.00212</td>
</tr>
<tr>
<td>hPGHS-1/indomethacin (NS)\textsuperscript{185}</td>
<td>2.0059</td>
<td>2.004</td>
<td>2.023</td>
</tr>
<tr>
<td>hPGHS (WS1)\textsuperscript{185}</td>
<td>2.0074</td>
<td>2.004</td>
<td>2.003</td>
</tr>
<tr>
<td>PGH Synthase (wide doublet)\textsuperscript{186}</td>
<td>2.0089</td>
<td>2.0044</td>
<td>2.0023</td>
</tr>
<tr>
<td>PGH Synthase (wide singlet)\textsuperscript{186}</td>
<td>2.0064</td>
<td>2.0044</td>
<td>2.0023</td>
</tr>
<tr>
<td>PGH Synthase (narrow singlet)\textsuperscript{186}</td>
<td>2.0074</td>
<td>2.0044</td>
<td>2.0023</td>
</tr>
<tr>
<td>Tyr-HCl\textsuperscript{171}</td>
<td>2.0067</td>
<td>2.0042</td>
<td>2.0023</td>
</tr>
<tr>
<td>Galactose Oxidase [CY]\textsuperscript{187}</td>
<td>2.00741</td>
<td>2.00641</td>
<td>2.00211</td>
</tr>
<tr>
<td>phenoxyl (calc)\textsuperscript{165}</td>
<td>2.0087</td>
<td>2.005</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.95 (calc)\textsuperscript{165}</td>
<td>2.0081</td>
<td>2.0049</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.70 (calc)\textsuperscript{165}</td>
<td>2.0079</td>
<td>2.0048</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.50 (calc)\textsuperscript{165}</td>
<td>2.0076</td>
<td>2.0047</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.95 180 degrees (calc)\textsuperscript{165}</td>
<td>2.0081</td>
<td>2.0049</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.95 135 degrees (calc)\textsuperscript{165}</td>
<td>2.0078</td>
<td>2.0049</td>
<td>2.0025</td>
</tr>
<tr>
<td>phenoxyl + water R(O–H) ) 1.95 90 degrees (calc)\textsuperscript{165}</td>
<td>2.0074</td>
<td>2.0049</td>
<td>2.0025</td>
</tr>
<tr>
<td>Bovine Catalase (calc)\textsuperscript{180}</td>
<td>2.00699</td>
<td>2.00463</td>
<td>2.00223</td>
</tr>
</tbody>
</table>
one of the two dual experimental \( g_{xx} \) values (2.0060 calc vs 2.00606 expt). The \( g_{xx} \) values for two structures, MYW and R(Y), are in closest agreement with the tyrosyl radical formed in PSII, the intermediate situation between a hydrogen-bonded and non-hydrogen-bonded tyrosyl radical. These values are 2.0073 for the MYW structure and 2.0078 for the R(Y) structure. However, the calculated free gas-phase tyrosine radical deviates considerably from E. coli RNR, the experimental example of a non-hydrogen bonded tyrosyl radical. In this case, the \( g_{xx} \) value for a calculated free tyrosine is around 2.0121 while the experimental \( g_{xx} \) value for E. coli RNR is around 2.0091. The fragments of MY and YW also deviate from the expected non-hydrogen bonded experimental example of a tyrosyl radical (E. coli RNR). These deviations from experimental values are likely due to basis set deficiencies of the 6-311G** basis set. Generally, the most popular basis set for EPR calculations is that of EPR-II and EPR-III\(^{188}\), which are optimized for the computation of EPR parameters by DFT methods (particularly B3LYP). However, these basis sets are not available for sulfur atoms, and thus could not be used in this study due to the sulfur of methionine. Another source of error could be the polarization due to solvent. The experimental examples provided here were conducted in aqueous solution; thus, the calculated \( g_{xx} \) values may improve when embedded in a polarizable continuum model to account for solvent effects.

An aqueous solvent can be accounted for by embedding the system in the polarizable continuum model (PCM). As mentioned above, a tyrosyl radical is polarized in the presence of a hydrogen bond, and consequently the magnitude of the \( g_{xx} \) value is reduced. Similarly, this lowering of the \( g_{xx} \) value is also predicted for tyrosine models embedded in PCM that mimics an aqueous solvent. The \( g_{xx} \) value of tyrosine embedded in PCM is computed around 2.010, which is consistent with a lowering of the \( g_{xx} \) value expected from an aqueous environment (compared to 2.0121 in the gas phase). In the cases where hydrogen bonding between arginine and the phenoxy group of tyrosine is present (R(Y) and R(MYW)), the opposite effect occurs; introduction of the polarizable continuum model increases the \( g \)-value. This indicates that in comparison to an explicit modeling of a hydrogen bond, models embedded
in PCM underestimate the strength of the interaction. The R(MYW) calculated value shifts from \( g_{xx} = 2.0060 \) computed in the gas phase to \( g_{xx} = 2.0068 \) computed with PCM(aqueous). The R(Y) calculated value also shifts from \( g_{xx} = 2.0078 \) (gas phase) to \( g_{xx} = 2.0084 \) (pcm, aq). Importantly, and in agreement with experimental data, the MYW \( g_{xx} \) value between gas phase and aqueous models remains unchanged. The gas phase \( g_{xx} \) value for MYW was calculated at 2.0073 while the \( g_{xx} \) value in PCM aqueous solution was also calculated at 2.0073. This demonstrates that the tyrosyl radical of MYW is not sensitive to environmental effects and must be explored further to understand why the \( g_{xx} \) value is not sensitive to its environment.

To further justify the choice in computational theory, additional reliability tests of the methodology were performed using different density functional theory packages, basis sets and solvation models. The geometries of four structures, Y, R(Y), MYW, and R(MYW), were optimized using the B3LYP hybrid functional, 6-311G** basis set, and PCM with water
as a solvent, as previously stated. Subsequent calculations on the optimized geometries were performed with two different density functional theory packages (Gaussian 09 Rev. C.02 vs. Orca Version 3.0), two basis sets were used (6-311G** vs. IGLO-III), and two solvation models (PCM=water vs. the conductor like screening model (COSMO)=water). The $g_{xx}$, $g_{yy}$, and $g_{zz}$ values for these structures with multiple methodologies may be found in Table 6.2. For gas phase models, the difference between $g$-tensor values among methods is very small. The greatest standard deviation between $g_{xx}$ gas phase models is only about $\pm 0.0006$ for the Y structure, $\pm 0.0003$ for the R(Y) structure, and $\pm 0.0002$ for both the MYW and R(MYW) structures. For aqueous models, the standard deviation between $g_{xx}$ values increases to about $\pm 0.0010$ for Y, $\pm 0.0004$ for R(Y), $\pm 0.0003$ for MYW, and $\pm 0.0002$ for the R(MYW) structure. These deviations are all small, and thus justify our use of the Gaussian 09 program package, the 6-311G** basis, and the polarizable continuum model.

### 6.1.1 The $g_{xx}$ Value as a Function of PCM Environment

In contrast to typical tyrosyl radicals, the $g_{xx}$ value for MYW exhibits no solvent dependence experimentally. The hydrogen bonding effect on the tyrosyl-like radical of MYW was further explored by varying the solvent defined by the polarizable continuum model on fragments of the MYW adduct including Y, MY, YW, and MYW. The solvent was varied by implementing the following values for the permittivity of environment in the polarizable continuum model: vacuum ($\varepsilon = 1$), argon ($\varepsilon = 1.430$), carbontetrachloride ($\varepsilon = 2.2280$), diethylether ($\varepsilon = 4.2400$), dichloroethane ($\varepsilon = 10.125$), ethanol ($\varepsilon = 24.852$), DMSO ($\varepsilon = 46.826$), and water ($\varepsilon = 78.3553$).

The $g_{xx}$ value as a function of the environment is shown in Figure 6.2. As expected, the free tyrosine calculation shows a strong dependence on solvent as the change from aqueous to vacuum is on the order of $2.1 \times 10^{-3}$. The greatest shift in $g_{xx}$ value occurs between $\varepsilon = 1$ and $\varepsilon = 5$. After which, the change in $g_{xx}$ gradually levels off. Similarly, YW and MY also show a dependence on solvent, the change being on the order of $1 \times 10^{-3}$. Again, a sharp
Table 6.2: Comparison of g-values from multiple methodologies for Y, R(Y), MYW, and R(MYW) structures. Two density functional theory packages were used, Gaussian 09 Rev. C.02 and Orca Version 3.0. Two basis sets were used, 6-311G** and IGLO-III. And two solvation models were used, PCM=water and COSMO=water.

<table>
<thead>
<tr>
<th>Method</th>
<th>Basis Set</th>
<th>Solvation</th>
<th>Y</th>
<th>R(Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_{xx}$</td>
<td>$g_{yy}$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>6-311G**</td>
<td>gas phase</td>
<td>2.0121</td>
<td>2.0051</td>
</tr>
<tr>
<td>Gaussian</td>
<td>IGLO-III</td>
<td>gas phase</td>
<td>2.0117</td>
<td>2.0051</td>
</tr>
<tr>
<td>Orca</td>
<td>IGLO-III</td>
<td>gas phase</td>
<td>2.0109</td>
<td>2.0050</td>
</tr>
<tr>
<td>Gaussian</td>
<td>6-311G**</td>
<td>PCM=water</td>
<td>2.0101</td>
<td>2.0049</td>
</tr>
<tr>
<td>Gaussian</td>
<td>IGLO-III</td>
<td>COSMO=water</td>
<td>2.0087</td>
<td>2.0048</td>
</tr>
<tr>
<td>Orca</td>
<td>IGLO-III</td>
<td>COSMO=water</td>
<td>2.0082</td>
<td>2.0047</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Basis Set</th>
<th>Solvation</th>
<th>MYW</th>
<th>R(MYW)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$g_{xx}$</td>
<td>$g_{yy}$</td>
</tr>
<tr>
<td>Gaussian</td>
<td>6-311G**</td>
<td>gas phase</td>
<td>2.0073</td>
<td>2.0039</td>
</tr>
<tr>
<td>Gaussian</td>
<td>IGLO-III</td>
<td>gas phase</td>
<td>2.0073</td>
<td>2.0039</td>
</tr>
<tr>
<td>Orca</td>
<td>IGLO-III</td>
<td>gas phase</td>
<td>2.0070</td>
<td>2.0039</td>
</tr>
<tr>
<td>Gaussian</td>
<td>6-311G**</td>
<td>PCM=water</td>
<td>2.0073</td>
<td>2.0041</td>
</tr>
<tr>
<td>Gaussian</td>
<td>IGLO-III</td>
<td>COSMO=water</td>
<td>2.0070</td>
<td>2.0040</td>
</tr>
<tr>
<td>Orca</td>
<td>IGLO-III</td>
<td>COSMO=water</td>
<td>2.0067</td>
<td>2.0040</td>
</tr>
</tbody>
</table>

|               |               |             | stdev (gas phase) |         |         |         |         |         |         |         |         |         |         |
|               |               |             | 0.0006        | 0.0001   | 0.0000  | 0.0003  | 0.0001  | 0.0000  |         |         |         |         |
|               |               |             | stdev (water)  | 0.0010   | 0.0001  | 0.0000  | 0.0004  | 0.0001  | 0.0000  |         |         |         |         |

|               |               |             | stdev (gas phase) |         |         |         |         |         |         |         |         |         |         |
|               |               |             | 0.0002        | 0.0000   | 0.0000  | 0.0002  | 0.0000  | 0.0000  |         |         |         |         |
|               |               |             | stdev (water)  | 0.0003   | 0.0000  | 0.0000  | 0.0002  | 0.0000  | 0.0000  |         |         |         |         |
decrease is observed in the range of \( \varepsilon = 1 \) and \( \varepsilon = 5 \). In each of these three cases, the change in \( g_{xx} \) is consistent with the expected trend in tyrosyl radicals in various environments. It also shows that the expected trend is stronger at lower permittivity levels (between \( \varepsilon = 1 \) and \( \varepsilon = 5 \)), indicating that the polarization effect is sufficient at this level. This finding is consistent with previous studies that have shown a dielectric constant of \( \varepsilon = 4 \) is considered to be an accurate representation of a protein environment (see Chapter 2.7).\(^3\)–\(^5\),\(^11\)–\(^13\) However, MYW remains unchanged even with variation of solvent; the difference in \( g_{xx} \) value between vacuum and aqueous is only \( 7.8 \times 10^{-5} \). This result is significant, as it provides further evidence that the MYW-radical is not sensitive to environmental effects.

Figure 6.2: Effect of varying solvents from \( \varepsilon = 1 \) (vacuum) to \( \varepsilon = 78 \) (water) on five computed tyrosyl radicals: MYW (red), MY (violet), YW (cyan), and Y (blue). The difference in \( g_{xx} \) value from aqueous solvent to gas phase is \( 7.8 \times 10^{-5} \) for MYW, \( 1.2 \times 10^{-3} \) for MY, \( 1.4 \times 10^{-3} \) for CYW, \( 1.4 \times 10^{-3} \) for YW, and \( 2.1 \times 10^{-3} \) for Y.
As mentioned in Chapter 5.2, a possible explanation for the atypical behavior of the MYW $g_{xx}$ value is the 1,4- interaction between the sulfur of methionine and the oxygen of tyrosine in the MYW models. Interpretations of the S···O interaction are based on both electrostatic and frontier orbital arguments.\textsuperscript{189,190} Using simple molecular orbital theory, the S···O interaction in these compounds is due to a $\sigma$-type interaction between the oxygen $p$ and the sulfur $p$ and $d$ orbitals. Variation in the S···O distance corresponds with the nature of the atom bonded to S, and can be explained in terms of the strength of the coupling between the X-S antibonding orbital and the oxygen lone-pair orbitals.\textsuperscript{190} Furthermore, computational studies on XSCH$_2$CO$_2^-$ and XSCH$_2$CHO compounds where X = H, F, Cl, CH$_3$, Li, Na, and BeH showed that electronegative atoms promote a greater electrostatic attraction between the sulfur and oxygen, especially in the XSCH$_2$CO$_2^-$ case.\textsuperscript{168} It is possible, that the -CH$_2$S(CH$_3$)$_2^+$···O• in MYW causes an even greater electrostatic attraction between the oxygen radical and the positively charged sulfur atom. This could potentially be the distinguishing factor that explains why the MYW $g_{xx}$ factor is not dependent on the solvent environment.

The proposed effect of the S$^+$···O• interaction is put forward in Scheme 6.1. For a tyrosine radical, polarization of the C-O bond due to PCM modeling or an explicit hydrogen bond will cause electron density to flow from the tyrosine ring toward the tyrosine oxygen. This will effectively increase the amount of spin on the tyrosine ring and decrease the amount of spin on the tyrosine oxygen. The decrease in spin along the C-O corresponds to a decrease in the $g_{xx}$ values as expected from experimental results. For a MY radical, the S$^+$···O• interaction will direct electron density away from the tyrosine-oxygen towards the sulfur. Based on inductive effects, the electron density will be donated back to the tyrosine ring. Again, the spin on the tyrosine ring is expected to increase and the spin on the tyrosine oxygen is expected to decrease. For a MYW radical, the divergence of the electron density through the S$^+$···O• interaction will continue; however, instead of the inductive effects spreading back through the tyrosine oxygen, the electron density will be pushed towards the more positively charged tryptophan (refer Table 5.2 for calculated total charge per residue).
This polarization effect in MYW should result in a net loss of electron density from tyrosine, which corresponds to an increase in spin density on the tyrosine. Thus, any spin that is lost on the tyrosine due to the environment modeled by the polarizable continuum (or by a hydrogen bond) can be replaced, and the $g_{xx}$ value will remain constant. Based on the aromaticity of the system, the $\pi$-interaction is likely to be important; however, based on the proximity of the sulfur to the oxygen, the $\sigma$-interaction is also likely to be important. A full orbital analysis is necessary to distinguish between $\sigma$- and $\pi$ interactions on the tyrosine oxygen and tyrosine/tryptophan rings.

Scheme 6.1: The $S^+ \cdots O^*$ interaction and the polarization of electron density in three models: Y, MY, and MYW. The blue arrows indicate the direction of electron density moving in the adduct.

Evidence of the increase in spin density on tyrosine for MYW models can be provided by computing the percentage of spin on the tyrosine oxygen and the tyrosine ring in gas phase and aqueous models, Table 6.3. (Total spin and charge per residue is shown in Table 6.4 and Table 6.5, respectively). The percentage of spin on the tyrosine-ring is calculated as the sum of spin values for all carbon and hydrogen atoms of the tyrosine over the total spin of the molecule. The percentage of spin on the tyrosine-ring is calculated as the spin value for oxygen over the total spin of the molecule. The difference between aqueous and gas phase models is shown as $\Delta$Spin in Table 6.3 to demonstrate the net increase/decrease of spin on the tyrosine when embedding the molecule in the polarizable continuum model. These values are also plotted in Figure 6.3 for $\Delta$Spin-Ring (x-axis) and $\Delta$Spin-oxygen (y-axis). A negative value of $\Delta$Spin indicates a net loss of spin density, and a positive value of
\( \Delta \text{Spin} \) indicates a net gain of spin density on the component (ring or oxygen) of the tyrosine radical. For example, the tyrosine model gains 3.7% spin on the Y-ring and loses 3.7% spin on the Y-oxygen when embedded in PCM. This is consistent with a polarization of the C–O bond due to the solvation model, and in agreement with experimental results on hydrogen bonding environments with tyrosyl radicals. The MYW and R(MYW) models both show a net increase in spin on both the tyrosine-ring and tyrosine-oxygen. This result provides evidence for the polarization effect due to the \( S^+ \cdots O^* \) interaction that was put forward earlier. A net increase in spin on all components of the tyrosine would replace any spin lost due to polarization from a hydrogen bond, keeping the \( g_{xx} \) value of an MYW-radical constant. To further test this idea, additional structures that will test both the substituent position and substituent identity will be explored in the proceeding sections.

Table 6.3: Percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for combinations of M, Y, and W. Note: \( \Delta \text{Spin} \) (in bold) is the difference between the percentage of aqueous spin and the percentage of gas phase spin. These values are plotted as x (Y-ring) and y (Y-oxygen) coordinates in Figure 6.3.

<table>
<thead>
<tr>
<th>model</th>
<th>Percent Spin Y-ring</th>
<th>Percent Spin Y-oxygen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gas phase</td>
<td>aqueous</td>
</tr>
<tr>
<td>Y</td>
<td>58.0</td>
<td>61.6</td>
</tr>
<tr>
<td>MY</td>
<td>60.2</td>
<td>63.1</td>
</tr>
<tr>
<td>YW</td>
<td>52.0</td>
<td>53.8</td>
</tr>
<tr>
<td>MYW</td>
<td>39.8</td>
<td>47.0</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>38.4</td>
<td>44.9</td>
</tr>
</tbody>
</table>

Table 6.4: Calculated total spin per residue in gas phase and aqueous models for combinations of M, Y, and W.
Figure 6.3: Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for MYW fragments.

Table 6.5: Calculated total charge per residue in gas phase and aqueous models for combinations of M, Y, and W.

<table>
<thead>
<tr>
<th></th>
<th>Gas Phase</th>
<th>Aqueous</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>Y</td>
</tr>
<tr>
<td>Y</td>
<td>-</td>
<td>0.0000</td>
</tr>
<tr>
<td>MY</td>
<td>0.9781</td>
<td>0.0219</td>
</tr>
<tr>
<td>YW</td>
<td>-</td>
<td>-0.1001</td>
</tr>
<tr>
<td>MYW</td>
<td>0.9374</td>
<td>-0.3058</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.9435</td>
<td>-0.3541</td>
</tr>
</tbody>
</table>
Substituent Position Effect

It is apparent from the foregoing discussion that having both Met and Trp covalently linked to Tyr contributes to the constant behavior of the $g_{xx}$ tensor as a function of the C−O environment. The orientation of substituents with respect to the tyrosine ring has been explored to verify the dependence of the S$^+\cdots O^*$ interaction, and to determine the role of the Trp position. The following structures were calculated for both gas phase and aqueous models: MYW$_m$, M$_m$YW, and M$_m$YW$_m$. Unless otherwise noted (m=meta) the substituent assumes the original ortho position. It is worth noting that computationally, the MYW (with each substituent in the ortho position, as found in KatG) is energetically favored. Relative energies of the four structures are shown in Scheme 6.2.

Scheme 6.2: Representation of the four structures calculated to test the effect of the Met and Trp positions on the MYW adduct.

Analogous to Figure 6.2, the $g_{xx}$ value of each model was computed in various PCM environments, as shown in Figure 6.4. All three newly computed models show dependence on the solvent, particularly from $\epsilon = 1$ to $\epsilon = 5$, similar to the Y-radical. This is a first indication that the ortho positions of Met and Trp in MYW are significant; however, the percentage of spin on the tyrosine must also be explored.
CHAPTER 6. A UNIQUE TYROSYL RADICAL: (MYW)$^{+}$

Figure 6.4: Effect of varying solvents from $\varepsilon = 1$ (vacuum) to $\varepsilon = 78$ (water) on models that varied the Met and Trp positions on the MYW adduct: MYW (red), Y (blue), MYW$_m$ (black), M$_m$YW (black), and M$_m$YW$_m$ (black).

The percentage of spin on the tyrosine oxygen and the tyrosine ring in gas phase and aqueous models for M$_m$YW, MYW$_m$, and M$_m$YW$_m$ models is shown in Table 6.6 and Figure 6.5. For all models with Met in a meta position (M$_m$YW and M$_m$YW$_m$) the percentage of spin increases on the ring and decreases on the oxygen, as expected for a typical tyrosine radical. This provides further evidence that the 1,4- S$^+$···O$^*$ interaction plays an important role in MYW. For the MYW$_m$ model, the percentage of spin on the ring decreases while the percentage of spin on the oxygen increases. This means that for the electron density to distribute on the Trp, as shown in Scheme 6.1, the residue must be in the meta position. These results suggest that both the Met and Trp residues are required to be in the ortho position on the tyrosine ring in MYW. Calculated total spin and total charge per residue may also be found in Table 6.7 and Table 6.8, respectively.
Table 6.6: Percentage of spin distributed on the Tyr-ring vs Tyr-O when structures are computed as gas-phase models vs. aqueous models for variations of Met and Trp positions on the MYW adduct. Note: ∆Spin (in bold) is the difference between the percentage of aqueous spin and the percentage of gas phase spin. These values are plotted as x (Y-ring) and y (Y-oxygen) coordinates in Figure 6.5.

<table>
<thead>
<tr>
<th>model</th>
<th>Percent Spin Y-ring</th>
<th>Percent Spin Y-oxygen</th>
<th>∆Spin</th>
<th>∆Spin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gas phase</td>
<td>aqueous</td>
<td>gas phase</td>
<td>aqueous</td>
</tr>
<tr>
<td>MYWₘ</td>
<td>65.5</td>
<td>63.9</td>
<td>-1.6</td>
<td>33.0</td>
</tr>
<tr>
<td>MₘYW</td>
<td>38.5</td>
<td>46.4</td>
<td>7.9</td>
<td>32.5</td>
</tr>
<tr>
<td>MₘₘₘYWₘ</td>
<td>55.2</td>
<td>60.5</td>
<td>5.3</td>
<td>44.9</td>
</tr>
<tr>
<td>MYW</td>
<td>39.8</td>
<td>47.0</td>
<td>7.1</td>
<td>24.9</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>38.4</td>
<td>44.9</td>
<td>6.5</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Figure 6.5: Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for variations of Met and Trp positions on the MYW adduct (pink dots). R(MYW) and MYW are shown as blue dots.
Table 6.7: Calculated total spin per residue in gas phase and aqueous models for variations of Met and Trp positions on the MYW adduct.

<table>
<thead>
<tr>
<th>model</th>
<th>Gas Phase M</th>
<th>Y</th>
<th>W</th>
<th>Aqueous M</th>
<th>Y</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYW&lt;sub&gt;m&lt;/sub&gt;</td>
<td>0.0100</td>
<td>0.9848</td>
<td>0.0053</td>
<td>0.0113</td>
<td>0.9922</td>
<td>-0.0035</td>
</tr>
<tr>
<td>M&lt;sub&gt;m&lt;/sub&gt;YW</td>
<td>-0.0041</td>
<td>0.7105</td>
<td>0.2937</td>
<td>-0.0030</td>
<td>0.7847</td>
<td>0.2184</td>
</tr>
<tr>
<td>M&lt;sub&gt;m&lt;/sub&gt;YW&lt;sub&gt;m&lt;/sub&gt;</td>
<td>-0.0039</td>
<td>1.0010</td>
<td>0.0029</td>
<td>-0.0026</td>
<td>1.0027</td>
<td>-0.0001</td>
</tr>
<tr>
<td>MYW</td>
<td>0.0044</td>
<td>0.6470</td>
<td>0.3486</td>
<td>0.0047</td>
<td>0.7357</td>
<td>0.2596</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.0028</td>
<td>0.5803</td>
<td>0.4170</td>
<td>0.0038</td>
<td>0.6903</td>
<td>0.3060</td>
</tr>
</tbody>
</table>

Table 6.8: Calculated total charge per residue in gas phase and aqueous models for variations of Met and Trp positions on the MYW adduct.

<table>
<thead>
<tr>
<th>model</th>
<th>Gas Phase M</th>
<th>Y</th>
<th>W</th>
<th>Aqueous M</th>
<th>Y</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYW&lt;sub&gt;m&lt;/sub&gt;</td>
<td>0.9623</td>
<td>-0.0828</td>
<td>0.1205</td>
<td>0.9968</td>
<td>-0.0514</td>
<td>0.0546</td>
</tr>
<tr>
<td>M&lt;sub&gt;m&lt;/sub&gt;YW</td>
<td>0.9092</td>
<td>-0.2453</td>
<td>0.3361</td>
<td>0.9586</td>
<td>-0.1953</td>
<td>0.2367</td>
</tr>
<tr>
<td>M&lt;sub&gt;m&lt;/sub&gt;YW&lt;sub&gt;m&lt;/sub&gt;</td>
<td>0.9243</td>
<td>-0.0030</td>
<td>0.0787</td>
<td>0.9703</td>
<td>-0.0223</td>
<td>0.0520</td>
</tr>
<tr>
<td>MYW</td>
<td>0.9374</td>
<td>-0.3058</td>
<td>0.3684</td>
<td>0.9699</td>
<td>-0.2362</td>
<td>0.2663</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.9435</td>
<td>-0.3541</td>
<td>0.4254</td>
<td>0.9718</td>
<td>-0.2732</td>
<td>0.3130</td>
</tr>
</tbody>
</table>

**Substituent Identity Effect**

The identity of the substituent in the methionine position has also been explored. New combinations of covalent linkages to the Tyr-Trp were explored by replacing Met with either cysteine or histidine. Cysteine is also expected to have a 1,4- S···O interaction with tyrosine, and a CY covalent bond is found in galactose oxidase (1GOF.pdb<sup>120</sup>), another example of a covalently linked tyrosine adduct. Two structures that included cysteine in place of methionine were calculated for both gas phase and aqueous models: CY and CYW (Scheme 6.3). Histidine was chosen due to its various protonation states that would allow the formation of +1, 0, and −1 charged systems. Histidine is also experimentally found as a HY linkage in cytochrome c oxidase (1V54.pdb<sup>191</sup>). In the case of histidine, two positions for the linkage were explored: 1) through nitrogen (as found in cytochrome c oxidase) and 2) through carbon, which will be referred to as N and C respectively. Two overall charges are considered for the nitrogen connectivity (neutral and +1 charge) and three charges are considered for
the carbon connectivity (+1, neutral and -1). Thus, a total of 10 structures that replace methionine with histidine were explored: H$_N^+$Y, H$_N$Y, H$_C^+$Y, H$_C$Y, H$_C^-$Y, H$_N^+$YW, H$_N$YW, H$_C^+$YW, H$_C$YW, and H$_C^-$YW (Scheme 6.3).

Scheme 6.3: Representation of the structures calculated to test the effect of the identity of the substituent in the Met position on the MYW adduct. Note: structures that omit the Trp residue are not shown here.

First, the $g_{xx}$ value of each model was computed in various PCM environments, and is shown in Figure 6.6 for H$_N$-Y and C-Y models and in Figure 6.7 for H$_C$-Y models. In both of these plots, it is important to notice that, consistent with MYW, only the $g_{xx}$ values of structures with a positive charge (H$_N^+$YW and H$_C^+$YW) remain constant, and are therefore not dependent on the C−O environment. It is also important to notice that the $g_{xx}$ values of the CYW model, which is expected to have a 1,4- S···O interaction, is dependent on the solvent. This suggests that the positive charge of Met is the origin of the behavior of the $g_{xx}$ value of MYW. While a 1,4- S···O interaction is present in CYW, it is not sufficient to polarize electron density away from the tyrosine oxygen. To further confirm this finding, the percentage of spin on the tyrosine will be discussed.
Figure 6.6: Effect of varying solvents from $\varepsilon = 1$ (vacuum) to $\varepsilon = 78$ (water) on models that replaced Met with either Cys or His (N connectivity): MYW (red), Y (blue), $H_N^+Y$ (black), $H_NY$ (black), $H_N^+YW$ (black), $H_NYW$ (black), and CYW (black).

Figure 6.7: Effect of varying solvents from $\varepsilon = 1$ (vacuum) to $\varepsilon = 78$ (water) on models that replaced Met with His (C connectivity): MYW (red), Y (blue), $H_C^+Y$ (black), $H_CY$ (black), $H_CY$ (black), $H_C^+YW$ (black), $H_CYW$ (black), and $H_C^-YW$ (black).
The percentage of spin on the tyrosine oxygen and the tyrosine ring in gas phase and aqueous models for Histidine and Cysteine structures are shown in Table 6.9 and Figure 6.8. (Total spin and charge per residue is shown in Table 6.10 and Table 6.10, respectively). All structures with a neutral or negative charge behave similar to a tyrosine radical, with the characteristic increase of spin on the ring and decrease of spin on the oxygen when embedded in the polarizable continuum model. However, consistent with MYW, the structures with positive charges ($\text{H}_x^\pm \text{YW}$ and $\text{H}_n^\pm \text{YW}$) show an increase in both the percentage of spin on the tyrosine-ring and on the tyrosine-oxygen. This result provides further evidence that the positive charge is an essential component to explaining the constant behavior of the $g_{xx}$ value of MYW. Revisiting the polarization effect described in Scheme 6.1, the positive charge of histidine will polarize the electron density of tyrosine-oxygen. $\text{His}^+$ can then donate the electron density back to the tyrosine ring, which can be distributed on the more positively charged tryptophan.

Table 6.9: Percentage of spin distributed on the Tyr-ring vs Tyr-O when structures are computed as gas-phase models vs. aqueous models for variations of the MYW adduct that replaced Met with either Cys or His. Note: $\Delta \text{Spin}$ (in bold) is the difference between the percentage of aqueous spin and the percentage of gas phase spin. These values are plotted as x (Y-ring) and y (Y-oxygen) coordinates in Figure 6.8.

<table>
<thead>
<tr>
<th>model</th>
<th>Percent Spin Y-ring</th>
<th>Percent Spin Y-oxygen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gas phase</td>
<td>aqueous</td>
</tr>
<tr>
<td>CY</td>
<td>51.4</td>
<td>54.4</td>
</tr>
<tr>
<td>CYW</td>
<td>53.4</td>
<td>54.9</td>
</tr>
<tr>
<td>$\text{H}_n^+ \text{Y}$</td>
<td>58.7</td>
<td>60.7</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{Y}$</td>
<td>53.4</td>
<td>56.7</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{YW}$</td>
<td>57.5</td>
<td>59.0</td>
</tr>
<tr>
<td>$\text{H}_n^+ \text{YW}$</td>
<td>50.0</td>
<td>53.2</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{YW}$</td>
<td>34.4</td>
<td>35.9</td>
</tr>
<tr>
<td>$\text{H}_n^+ \text{YW}$</td>
<td>43.0</td>
<td>49.9</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{YW}$</td>
<td>50.1</td>
<td>52.3</td>
</tr>
<tr>
<td>$\text{H}_n^+ \text{YW}$</td>
<td>40.3</td>
<td>47.2</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{YW}$</td>
<td>48.5</td>
<td>51.1</td>
</tr>
<tr>
<td>$\text{H}_n^\pm \text{YW}$</td>
<td>33.6</td>
<td>36.0</td>
</tr>
<tr>
<td>MYW</td>
<td>39.8</td>
<td>47.0</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>38.4</td>
<td>44.9</td>
</tr>
</tbody>
</table>
Figure 6.8: Plot of the change that occurs in the percentage of spin distributed on the Tyr-ring vs Tyr-oxygen when structures are computed as gas-phase models vs. aqueous models for variations of the MYW adduct that replaced Met with either Cys or His (black dots). R(MYW) and MYW are shown as blue dots.

Table 6.10: Calculated total spin per residue in gas phase and aqueous models for variations of the MYW adduct that replaced Met with either Cys or His.

<table>
<thead>
<tr>
<th>C or H</th>
<th>Y</th>
<th>W</th>
<th>C or H</th>
<th>Y</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>CY</td>
<td>0.1409</td>
<td>0.8591</td>
<td>-</td>
<td>0.1461</td>
<td>0.8540</td>
</tr>
<tr>
<td>CYW</td>
<td>0.0134</td>
<td>0.8749</td>
<td>0.1117</td>
<td>0.0109</td>
<td>0.8610</td>
</tr>
<tr>
<td>H+ Y</td>
<td>0.0385</td>
<td>0.9615</td>
<td>-</td>
<td>0.0404</td>
<td>0.9596</td>
</tr>
<tr>
<td>H++ Y</td>
<td>0.1008</td>
<td>0.8992</td>
<td>-</td>
<td>0.0742</td>
<td>0.9258</td>
</tr>
<tr>
<td>H+ Y</td>
<td>0.0721</td>
<td>0.9280</td>
<td>-</td>
<td>0.0786</td>
<td>0.9214</td>
</tr>
<tr>
<td>H++ Y</td>
<td>0.1208</td>
<td>0.8792</td>
<td>-</td>
<td>0.1331</td>
<td>0.8669</td>
</tr>
<tr>
<td>H- Y</td>
<td>0.3958</td>
<td>0.6042</td>
<td>-</td>
<td>0.4101</td>
<td>0.5899</td>
</tr>
<tr>
<td>H++ YW</td>
<td>0.0121</td>
<td>0.6979</td>
<td>0.2900</td>
<td>0.0141</td>
<td>0.7899</td>
</tr>
<tr>
<td>H+ YW</td>
<td>0.0521</td>
<td>0.8393</td>
<td>0.1085</td>
<td>0.0412</td>
<td>0.8381</td>
</tr>
<tr>
<td>H++ YW</td>
<td>0.0303</td>
<td>0.6325</td>
<td>0.3372</td>
<td>0.0385</td>
<td>0.7316</td>
</tr>
<tr>
<td>H+ YW</td>
<td>0.0775</td>
<td>0.8336</td>
<td>0.0889</td>
<td>0.0782</td>
<td>0.8256</td>
</tr>
<tr>
<td>H++ YW</td>
<td>0.4149</td>
<td>0.5668</td>
<td>0.0184</td>
<td>0.3913</td>
<td>0.5863</td>
</tr>
<tr>
<td>MYW</td>
<td>0.0044</td>
<td>0.6470</td>
<td>0.3486</td>
<td>0.0047</td>
<td>0.7357</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.0028</td>
<td>0.5803</td>
<td>0.4170</td>
<td>0.0038</td>
<td>0.6903</td>
</tr>
</tbody>
</table>
Table 6.11: Calculated total charge per residue in gas phase and aqueous models for variations of the MYW adduct that replaced Met with either Cys or His.

<table>
<thead>
<tr>
<th></th>
<th>Gas Phase</th>
<th>Aqueous</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C or H</td>
<td>Y</td>
</tr>
<tr>
<td>CY</td>
<td>0.3056</td>
<td>-0.3056</td>
</tr>
<tr>
<td>CYW</td>
<td>0.1490</td>
<td>-0.2484</td>
</tr>
<tr>
<td>H_2^+Y</td>
<td>0.7012</td>
<td>0.2989</td>
</tr>
<tr>
<td>H_2-Y</td>
<td>0.0861</td>
<td>-0.0861</td>
</tr>
<tr>
<td>H_2^+Y</td>
<td>0.9047</td>
<td>0.0954</td>
</tr>
<tr>
<td>H_2-Y</td>
<td>0.1043</td>
<td>-0.1043</td>
</tr>
<tr>
<td>H_2^-Y</td>
<td>-0.5431</td>
<td>-0.4569</td>
</tr>
<tr>
<td>H_2^+YW</td>
<td>0.6633</td>
<td>0.0284</td>
</tr>
<tr>
<td>H_2-YW</td>
<td>-0.1661</td>
<td>0.0608</td>
</tr>
<tr>
<td>H_2^+YW</td>
<td>0.8463</td>
<td>-0.2006</td>
</tr>
<tr>
<td>H_2-YW</td>
<td>0.0726</td>
<td>-0.1532</td>
</tr>
<tr>
<td>H_2^-YW</td>
<td>-0.5163</td>
<td>-0.4286</td>
</tr>
<tr>
<td>MYW</td>
<td>0.9374</td>
<td>-0.3058</td>
</tr>
<tr>
<td>R(MYW)</td>
<td>0.9435</td>
<td>-0.3541</td>
</tr>
</tbody>
</table>

6.1.2 Summary

The gy value of the MYW-radical found in the catalase intermediate KatG was calculated to be independent of the C−O environment, which is atypical of tyrosyl radicals. Two main features contribute to this behavior. First, the interaction between the positively charged sulfur on methionine and the oxygen of tyrosine effectively redistribute electron density from the Tyr-oxygen back to the Tyr ring. This positive charge is only effective when the methionine is covalently linked to an ortho position on the tyrosine ring. Second, the Trp that is covalently linked to the other ortho position on the tyrosine ring effectively redistributes electron density away from the tyrosine to regenerate spin on the tyrosine. Both of these factors work simultaneously to maintain the spin of tyrosine as constant. Further studies are needed to confirm and elucidate the exact details of this effect.
Chapter 7

Concluding Remarks

Density functional theory was used to investigate the properties of transition metal complexes that are geochemically and biologically relevant as a function of the occupation of $d$-orbital electrons. Two systems were explored: siderophore-transition metal complexes, and the catalase intermediate of the Catalase-Peroxidase (KatG) enzyme. This dissertation has demonstrated that density functional theory is a useful and valuable method to gain insights into transition metal complexes.

The first comprehensive study described in Chapter 4 was used to assess the structural parameters, binding energies, and vibrational spectra of metal-siderophore complexes in aqueous solution. The study was supported by structural data from extended X-ray absorption fine structure (EXAFS) spectroscopy and vibrational data from infrared spectroscopy. The structural parameters derived from EXAFS analyses were well reproduced and validated by the applied computational model. The DFT binding energies indicated systematic errors that were successfully extracted by an empirical fit. The refined energies agree with and reproduce the general trends observed for measured binding constants, including dependence on the $d$-orbital electronic configuration and the charge-normalized ionic radii. The normal mode analysis of the IR spectra revealed how structure and vibrational spectra of aqueous desferrioxamine B (DFOB) metal complexes vary with the metal ion identity. In conclu-
sion, the density functional theory study of metal-siderophore complexes demonstrated that density functional theory is a powerful method to gain insights into metal binding complexes.

The second application of density functional theory described in Chapter 5 revealed the role of a unique tyrosyl radical in oxygen release from the catalase intermediate of the Catalase-Peroxidase (KatG) enzyme. The DFT study provided characterization of a novel catalytic step that involves both an MYW-radical and oxyheme in the KatG catalase reaction. It was established that the Met-Tyr-Trp radical is flexible such that a hinge-like opening rotation of the Trp-107 ring with respect to the Tyr-229 ring along their covalent C-C bond, is an inherent feature of its catalytic properties. Also, an H-bond between the Tyr-229 and the mobile side-chain of Arg-418 further enables the catalytic events. The opening process breaks an H-bond between the N-H of Trp-107 and the inner oxygen of the Fe-O$_2$ complex present in the closed conformation of the MYW-radical. This motion lowers the spin-crossing energy barrier between the ground state and the catalytically active high-spin states and enables electron transfer from the oxyheme group to the MYW-radical. The release of molecular oxygen is thereby catalyzed and leaves ferric-heme poised for another catalytic cycle. These conclusions based on the calculated models are not presented as the definitive analysis of the mechanism but rather one that yielded a plausible explanation for the interactions between two species experimentally identified in the enzyme during steady-state catalytic turnover. Thus, density functional theory has been shown here to provide insights into a catalytic step of a complicated enzyme process.

The third application of density functional theory described in Chapter 6 provided insights into the radical character and $g_{xx}$ value of the MYW-radical in the Catalase-Peroxidase (KatG) enzyme. It was shown that the $g_{xx}$ value of the MYW-radical found in the catalase intermediate of KatG is independent of the environment, which is atypical of tyrosyl radicals. The 1,4-sulfur-oxygen interaction between methionine and tyrosine was put forward as a possible explanation for the non-environment dependent $g_{xx}$ value of MYW. Preliminary results of spin distributions on the radical from the DFT calculations were assessed and
confirm the hypothesis. A deeper analysis of the orbitals using density functional theory could provide greater insight and further evidence of this effect.

In conclusion, density functional theory applied to the two systems, metal-siderophore complexes and the KatG enzyme, has demonstrated the validity of the computational approach. As density functional theory continues to advance and additional experimental data is obtained, it is with hope that greater details stemming from those presented here will be elucidated.
Appendices
Appendix A

Calculation of the J and K Integrals

Considering the helium atom, a three-body system consisting of two electrons and a nucleus at rest, the origin of the coordinate system will be placed at the nucleus, generating the coordinate of electron 1 as \((x_1, y_1, z_1)\) and of electron 2 as \((x_2, y_2, z_2)\).

If the nuclear charge is taken as \(+Ze\), the Hamiltonian is given by:

\[
\hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} \quad (A.1)
\]

\[
\hat{H}_0 = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{r_1}, \quad \hat{H}' = \frac{e^2}{|r_1 - r_2|} \quad (A.2)
\]

where \(m\) is the mass of the electrons, \(r_1\) and \(r_2\) are the distances of electrons 1 and 2 from the nucleus, and \(r_{12}\) is the distance between electron 1 and electron 2. The first two terms are the kinetic energy operators, while the third and fourth terms are the potential energy of inter-electronic repulsion.

The Schrödinger equation involves six independent variables, three coordinates for each electron. In spherical coordinates, \(\psi = \psi(r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2)\). The operator \(\hat{p}_1^2 \cong -\hbar^2 \nabla_1^2\) is given by the Laplacian spherical coordinates with \(r_1, \theta_1, \phi_1\) replacing \(r, \theta, \phi\). The variable \(r_{12}\) is \(r_{12} = [(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]^{1/2}\), and by using the relations between Cartesian and spherical coordinates, \(r_{12}\) can be expressed in terms of \(r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2\).
Suppose $e^2/|\mathbf{r}_1 - \mathbf{r}_2| \equiv e^2/r_{12}$. Because of this term, the Schrödinger equation for helium cannot be separated and approximation methods must be used. The perturbation method separates the Hamiltonian (A.1) into two parts, $\hat{H}_0$ and $\hat{H}'$, where $\hat{H}_0$ is the Hamiltonian of an exactly solvable problem and $\hat{H}'$ is the electron-electron interaction. The wave function would then be the product of two hydrogen atom wave functions, with $Z = 2$. The total spin remains constant so the state is either singlet or triplet.

The spatial part of the wave function where one electron is in the ground state and the other is in a higher excited state characterized by $(n,l,m)$ is

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\psi_{100}(\mathbf{r}_1)\psi_{n,l,m}(\mathbf{r}_2) \pm \psi_{100}(\mathbf{r}_2)\psi_{n,l,m}(\mathbf{r}_1)] \quad (A.3)$$

From the expectation value,

$$\langle f \rangle = \int \psi^*(x,t)\hat{f}\psi(x,t) dx \quad (A.4)$$

And the expectation values of $\hat{H}'$ are defined as:

$$\left\langle \frac{e^2}{r_{12}} \right\rangle = \int_{\mathbf{r}_2} \int_{\mathbf{r}_1} \psi^*(\mathbf{r}_1, \mathbf{r}_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi(\mathbf{r}_1, \mathbf{r}_2) d^3r_1 d^3r_2 \quad (A.5)$$

Next, the first-order perturbation correction to the energy must be evaluated.

$$\Delta E^{(1)} = \langle H' \rangle = \left\langle \frac{e^2}{r_{12}} \right\rangle \quad (A.6)$$

By substituting (A.3) into (A.6), the expectation value becomes,

$$\left\langle \frac{e^2}{r_{12}} \right\rangle = \frac{e^2}{2} \int_{\mathbf{r}_2} \int_{\mathbf{r}_1} [\psi_{100}(\mathbf{r}_1)^2|\psi_{n,l,m}(\mathbf{r}_2)|^2 \pm \psi_{100}^*(\mathbf{r}_1)|\psi_{n,l,m}(\mathbf{r}_2)|^2]$$
APPENDIX A. CALCULATION OF THE J AND K INTEGRALS

\[ \pm \psi_{100}^*(r_2) \psi_{n\ell m}(r_1) \psi_{100}(r_1) \psi_{n\ell m}(r_2) + |\psi_{100}^*(r_2)|^2 |\psi_{n\ell m}(r_1)|^2 \frac{d^3r_1 d^3r_2}{|r_1 - r_2|}. \]

Considering the symmetry between \( r_1 \) and \( r_2 \), the equation can be rewritten as

\[
\left\langle \frac{e^2}{r_{12}} \right\rangle = \int_{r_2} \int_{r_1} |\psi_{100}(r_1)|^2 \cdot \frac{e^2}{|r_1 - r_2|} \cdot |\psi_{n\ell m}(r_2)|^2 d^3r_1 d^3r_2 \pm Re
\]

\[ \int_{r_2} \int_{r_1} \psi_{100}(r_2) \psi_{n\ell m}^*(r_1) \cdot \frac{e^2}{|r_1 - r_2|} \cdot \psi_{100}(r_1) \psi_{n\ell m}(r_2) d^3r_1 d^3r_2. \quad (A.7) \]

The first term of this equation is the **Coulomb integral**:

\[
J = \left\langle \frac{e^2}{r_{12}} \right\rangle = \int_{r_2} \int_{r_1} |\psi_{100}(r_1)|^2 \cdot \frac{e^2}{|r_1 - r_2|} \cdot |\psi_{n\ell m}(r_2)|^2 d^3r_1 d^3r_2 \quad (A.8)
\]

The second term is the **Exchange integral**:

\[
K = \int_{r_2} \int_{r_1} \psi_{100}(r_2) \psi_{n\ell m}^*(r_1) \cdot \frac{e^2}{|r_1 - r_2|} \cdot \psi_{100}(r_1) \psi_{n\ell m}(r_2) d^3r_1 d^3r_2. \quad (A.9)
\]

Consider an electron in the \( 1s \) orbital. The energy of this state can be written as

\[ E = E_{100} + E_{n\ell m} + \Delta E. \quad (A.10) \]

In first-order perturbation theory, \( \Delta E \) is obtained by evaluating the expectation value of \( e^2/r_{12} \)

\[
\left\langle \frac{e^2}{r_{12}} \right\rangle = (J \pm K). \quad (A.11)
\]

The total energy of the helium atom is then obtained by substituting A.11 into A.10.

\[ E = E_{100} + E_{n\ell m} + J \pm K. \]
APPENDIX A. CALCULATION OF THE J AND K INTEGRALS

The lower energy level is therefore,

$$E_t = E_{100} + E_{nlm} + (J - K), \quad (A.12)$$

a triply degenerate state, thus called a **triplet**. The higher energy level,

$$E_t = E_{100} + E_{nlm} + (J + K) \quad (A.13)$$

is non-degenerate, and is therefore a **singlet**. In the singlet case A.13 the spatial function is symmetric and the electrons have the tendency to be in close proximity to each other. Therefore, the effect of electrostatic repulsion is greater, and a higher energy results. In the triplet case A.12, the spatial function is antisymmetric and the electrons tend to avoid each other.

Using the addition theorem of spherical harmonics, the Coulomb integral A.8 can be written as

$$\frac{1}{|r_1 - r_2|} = \sum_{l' = 0}^{\infty} \sum_{m' = -l'}^{l'} \left( \frac{4\pi}{2l' + 1} \right) \cdot \frac{r_<}{r_>}^{l'} Y_{l'm'}^*(\Omega_1) Y_{l'm'}(\Omega_2). \quad (A.14)$$

where $r_>(r_<)$ is the larger (or smaller) of $r_1$ and $r_2$, respectively. Thus a different expansion is required for each region in which $r_>$ represents the vector indicating the position of the electron far away from the nucleus, and $r_<$ represents the vector indicating the position of the electron near the nucleus.

Substituting A.14 into A.8 yields:

$$J = e^2 \sum_{l' = 0}^{\infty} \sum_{m' = -l'}^{l'} \left( \frac{4\pi}{2l' + 1} \right) \int_{\Omega_2} \int_{\Omega_1} \int_{r_2} \int_{r_1} dr_1 dr_2 d\Omega_1 d\Omega_2 r_1^2 R_1 0^2(r_1) Y_{00}^*(\Omega_1)$$

$$Y_{00}(\Omega_1) \left( \frac{r_<^{l'}}{r_>^{l'+1}} \right) Y_{l'm'}^*(\Omega_1) Y_{l'm'}(\Omega_2) r_2^2 R_{nl}^2(r_2) Y_{lm}^*(\Omega_2) \quad (A.15)$$
Substituting $Y_{00}(\Omega_1)Y_{00}^*(\Omega_1) = (4\pi)^{-1/2}$ into Eq. (A.15) gives

$$J = e^2 \sum_{l'v=0}^{\infty} \sum_{m'==-l'}^{l'} \frac{1}{2l' + 1} \int_{\Omega_2} \int_{\Omega_1} d\Omega_1 d\Omega_2 Y_{l'm'}(\Omega_1)Y_{l'm'}^*(\Omega_2)Y_{lm}^*(\Omega_2)$$

\[ \int_{r_2} r_1 \int_{r_1} dr_1 dr_2 r_2^2 R_{10}^2(r_1) r_2^2 R_{nm}(r_2) \left( \frac{r_{l'}^{<}}{r_{l'+1}^{>}} \right) \] \hspace{1cm} (A.16)

A new variable, $C_{l'}$, can be defined such that

$$C_{l'} = \int_{r_2} r_1 \int_{r_1} dr_1 dr_2 r_2^2 R_{10}^2(r_1) r_2^2 R_{nm}(r_2) \left( \frac{r_{l'}^{<}}{r_{l'+1}^{>}} \right).$$ \hspace{1cm} (A.17)

If the definition of $\left( \frac{r_{l'}^{<}}{r_{l'+1}^{>}} \right)$ is

$$\left( \frac{r_{l'}^{<}}{r_{l'+1}^{>}} \right) = \begin{cases} \frac{r_{l'}^{1}}{r_{l'+1}^{2}} & \text{for } r_1 < r_2 \\ \frac{r_{l'}^{2}}{r_{l'+1}^{1}} & \text{for } r_1 > r_2 \end{cases}$$

Then A.17 becomes

$$C_{l'} = \int_{0}^{\infty} \int_{0}^{r_2} r_2^2 \frac{r_{l'}^{1+2}}{r_{l'+1}^{2}} R_{nm}(r_2) R_{10}^2(r_1) dr_1 dr_2 + \int_{0}^{\infty} \int_{r_2}^{\infty} r_1^2 \frac{r_{l'}^{1+2}}{r_{l'+1}^{2}} R_{nm}(r_2) R_{10}^2(r_1) dr_1 dr_2 \hspace{1cm} (A.18)$$

The first term on the right-hand side of A.18 is $C_{l'}$ and the second term is $C_{l'}^*$. By inserting Eq. (A.18) into Eq. (A.16) the Coulomb integral becomes

$$J = e^2 \sum_{l'v=0}^{\infty} \sum_{m'==-l'}^{l'} \frac{C_{l'}}{2l' + 1} \int_{\Omega_2} \int_{\Omega_1} Y_{l'm'}(\Omega_1)Y_{l'm'}^*(\Omega_2)Y_{lm}^*(\Omega_2)$$ \hspace{1cm} (A.19)

Consider the first excited state ($1s2p$) of helium, in which $n = 2, l = 1, m = \pm 1, 0$. Eq.
A.19 becomes

\[
J = e^2 \sum_{l' = 0}^{\infty} \sum_{n' = -l'}^{l'} \frac{C_{l'}}{2l' + 1} \int_\Omega \int \Omega_1 \int \Omega_2 \int \Omega_1 \int \Omega_2 \left\{ \int_\Omega_1 \int \Omega_1 \int \Omega_2 \int \Omega_2 \int \Omega_1 \int \Omega_2 \right\} \cdot d\Omega_1 \cdot d\Omega_2 \tag{A.20}
\]

Based on the orthonormality of spherical harmonics

\[
J = e^2 C_{l'} \tag{A.21}
\]

By substituting A.18 into , the Coulomb integral becomes

\[
J = e^2 \int_0^\infty r_2^2 R_{l'1}^2(r_2) \left[ \int_0^{r_2} r_1^2 R_{l10}^2(r_1) dr_1 \right] dr_2 + \int_0^\infty r_2^2 R_{l'1}^2(r_2) \left[ \int_0^\infty r_1^2 R_{l10}^2(r_1) dr_1 \right] dr_2 \tag{A.22}
\]

The exchange integral can be written as

\[
K = e^2 \int_{r_2} \int_{r_1} \psi_{100}^*(r_2) \psi_{nltm}^*(r_1) \cdot \frac{1}{|r_1 - r_2|} \cdot \psi_{100}(r_1) \psi_{nltm}(r_2) d^3r_1 d^3r_2 \tag{A.23}
\]

Substituting the addition theorem of a spherical harmonics series into A.23, the exchange integral becomes

\[
K = e^2 \sum_{l' = 0}^{\infty} \sum_{m' = -l'}^{l'} \left( \frac{4\pi}{2l' + 1} \right) \int_{r_2} \int_{r_1} \int_{r_2} \int_{r_1} \psi_{100}^*(r_2) \psi_{nltm}^*(r_1) \left( \frac{r_1^{l'} r_2^{l'+1}}{r_1^{l'+1}} \right) Y_{l'm'}(\Omega_1) Y_{l'm'}(\Omega_2) \psi_{nltm}(r_1) \psi_{nltm}(r_2) \tag{A.24}
\]

Considering the first excited state (1s2p) of helium, in which \( n = 2, l = 1 \), and \( m = \pm 1 \), A.24 becomes

\[
K = e^2 \sum_{l' = 0}^{\infty} \sum_{m' = -l'}^{l'} \left( \frac{4\pi}{2l' + 1} \right) \int_{r_2} \int_{r_1} \int_{r_2} \int_{r_1} \psi_{100}^*(r_2) \psi_{21m}^*(r_1) \left( \frac{r_1^{l'} r_2^{l'+1}}{r_1^{l'+1}} \right) Y_{l'm'}(\Omega_1) Y_{l'm'}(\Omega_2) \psi_{21m}(r_1) \psi_{21m}(r_2) \tag{A.25}
\]
Now considering the wave function to be $\psi_{2,1,-1}(r) = R_{2,1}(r)Y_{1,-1}(\Omega)$ and using the properties of the orthonormality relations of spherical harmonics, the integral becomes

$$K = \frac{e^2}{3} \int_{r_2} \int_{r_2} r_2^2 R_{10}(r_2) \left\{ \int_{r_1} r_1^2 R_{10}(r_1) R_{21}(r_1) \left( \frac{r_1^{\prime\prime}}{r_1^{\prime\prime} + 1} \right) dr_1 \right\} dr_2. \quad (A.26)$$

Substituting the definition

$$\left( \frac{r_1^{\prime\prime}}{r_1^{\prime\prime} + 1} \right) = \begin{cases} \frac{r_1^{\prime\prime}}{r_1^{\prime\prime} + 1} & \text{for } r_1 < r_2 \\ \frac{r_2^{\prime}}{r_2^{\prime} + 1} & \text{for } r_1 > r_2 \end{cases}$$

into Eq. (A.26), the equation becomes

$$K = \frac{e^2}{3} \int_0^\infty \int_0^{r_2} r_2^2 r_1^3 R_{10}(r_2) R_{21}(r_2) R_{10}(r_1) R_{21}(r_1) dr_1 dr_2$$

$$+ \frac{e^2}{3} \int_0^\infty \int_{r_2}^\infty r_1^2 r_2^3 R_{10}(r_2) R_{21}(r_2) R_{10}(r_1) R_{21}(r_1) dr_1 dr_2 \quad (A.27)$$

The first term on the right hand side of A.27 is $K_{1a}$ and the second term is $K_{2a}$. Because $K_{1a} = K_{2a}$, the integral can be rewritten as

$$K = \frac{2e^2}{3} \int_0^\infty r_2^3 R_{10}(r_2) R_{21}(r_2) \left\{ \int_{r_2}^\infty R_{10}(r_1) R_{21}(r_1) dr_1 \right\} dr_2. \quad (A.28)$$
Appendix B

Theoretical Basis for DFT

Density functional theory is based on three important theorems: the Hohenberg-Kohn theorem, the Hohenberg-Kohn variational theorem, and the Kohn-Sham theorem. The first two provide the theoretical basis for DFT, and the third provides a method for practical implementation.

B.0.3 The Hohenberg-Kohn Theorem

The ground state probability density is uniquely determined by the external potential (up to an arbitrary energy shift).

In 1962, Pierre Hohenberg and Walter Kohn proved that the electronic properties of a molecule are uniquely determined by the ground-state electron probability density $\rho_0(x, y, z)$. More specifically, Hohenberg and Kohn proved that for systems with a non-degenerate ground state, the ground-state electron probability density $\rho_0(r)$ determines the external potential and the number of electrons. The ground-state electronic wave function $\psi_0$ of an $N$-electron molecule is an eigenfunction of the electronic Hamiltonian

$$H = -\frac{1}{2} \sum_{i=1}^{n} \nabla_i^2 + \sum_{i=1}^{n} v(r_i) + \sum_{j} \sum_{i>j} \frac{1}{r_{ij}}$$
APPENDIX B. THEORETICAL BASIS FOR DFT

where $v(r_i)$ is the potential energy of interaction between electron $i$ and the nuclei, $v(r_i) = -\sum_{\alpha} \frac{Z_{\alpha}}{r_{i\alpha}}$, also known as the external potential. Consider two different external potentials ($v_1$ and $v_2$) that give rise to the same ground-state electron density $\rho_0$, where $\psi_1$ is the unique ground state of the first Hamiltonian and $\psi_2$ is the unique ground state of the second Hamiltonian. The ground state energy of the first is then

$$E_1 = \langle \psi_1 | \hat{H}_1 | \psi_1 \rangle = \langle \psi_1 | T + V_{ee} | \psi_1 \rangle + \int \rho_0(r)V_{en}^{(1)}(r)dr \quad (B.1)$$

And the second is

$$E_1 = \langle \psi_2 | \hat{H}_2 | \psi_2 \rangle = \langle \psi_2 | T + V_{ee} | \psi_2 \rangle + \int \rho_0(r)V_{en}^{(1)}(r)dr \quad (B.2)$$

The variational principle and the assumption that each have a unique ground state implies

$$E_1 < E'_2 = \langle \psi_1 | T + V_{ee} | \psi_1 \rangle + \int \rho_0(r)V_{en}^{(1)}(r)dr = E_2 + \int \rho_0(r) [V_{en}^{(1)}(r) - V_{en}^{(2)}(r)] \quad (B.3)$$

And

$$E_2 \leq E'_2 = E_1 + \int \rho_0(r) [V_{en}^{(2)}(r) - V_{en}^{(1)}(r)] \quad (B.4)$$

Adding these two equations yields

$$E_1 + E_2 < E_1 + E_2 \quad (B.5)$$

Clearly, this is a contradictory statement. The assumption that two different external potentials could produce the same ground-state electron density must be false. Thus, the ground-state density uniquely determines the potential, up to an arbitrary constant. The external potential determines the Hamiltonian which, in turn, determines the ground-state wave function; therefore, the ground-state wave function must also be determined uniquely by the ground-state density. In other words, the ground-state wave function is a functional
of the ground-state density.

$$\psi(x_1, ..., x_n) = f[\rho(x)]$$

Similarly, the ground-state total energy must also be a functional of the ground-state density:

$$E = W[\rho(x)]$$

The sum of the kinetic energy and electron-repulsion energies must also be a functional of the density:

$$\int \psi^*(\hat{T} + \hat{G})\psi d\tau = F[\rho(\vec{x})]$$

The total ground-state energy can be written as a functional of the ground-state density:

$$E = W[\rho(\vec{x})] = F[\rho(\vec{x})] + \int \left( v(\vec{x})\rho(\vec{x}) \right) d\tau$$

This function is a universal functional that is independent of the external potential.

Although the functional $F[\rho(\vec{x})]$ is unknown, it must contain the Coulomb repulsion energy of the total electron density:

$$\int \int \rho(\vec{x}_1) \frac{1}{r_{12}} \rho(\vec{x}_2) d\tau_1 d\tau_2$$

This Coulomb expression contains the electronic self-interaction energies which are canceled by the exchange terms. Therefore,

$$F[\rho(\vec{x})] = G[\rho(\vec{x})] + \int \int \rho(\vec{x}_1) \frac{1}{r_{12}} \rho(\vec{x}_2) d\tau_1 d\tau_2$$

which concentrates everything that is unknown about the functional into $G[\rho(\vec{x})]$, giving

$$E = G[\rho(\vec{x})] + \int v(\vec{x})\rho(\vec{x}) d\tau + \int \int \rho(\vec{x}_1) \frac{1}{r_{12}} \rho(\vec{x}_2) d\tau_1 d\tau_2$$
where the functional $G[\rho(\vec{x})]$ is given by:

$$G[\rho(\vec{x})] = \int \Psi^* \hat{T} \Psi d\tau + \int \Psi^* \hat{U} \Psi d\tau - \int \int \rho(\vec{x}_1) \frac{1}{r_{12}} \rho(\vec{x}_2) d\tau_1 d\tau_2.$$ 

Thus, from the one-electron density $\rho(\vec{x})$, the functional $G[\rho(\vec{x})]$ generates the kinetic energy ($T[\rho(\vec{x})]$), and the non-Coulomb part of the electron-repulsion energy, the exchange-correlation energy ($T_{XC}[\rho(\vec{x})]$).

$$G[\rho(\vec{x})] = T[\rho(\vec{x})] + T_{XC}[\rho(\vec{x})].$$

### B.0.4 The Hohenberg-Kohn Variational Theorem

*The density corresponding to the lowest energy, corresponds to the density obtained from the exact wave function.*

A second theorem by Hohenberg and Kohn states that the density corresponding to the lowest energy also corresponds to the density obtained from the exact wave function. This generates a practical method for the computation of molecular electron distributions and their energies.

A given electron density function $\rho(\vec{x})$ determines the external potential $v(\vec{x})$ in which the electrons move and hence both the Hamiltonian ($\hat{H}$) and the solution of the associated Schrödinger equation. If the density function is used to evaluate the energy of a system of electrons with a different, known external potential associated with a known Hamiltonian, then this energy is known in terms of the conventional wave function expectation value and as a functional of $\rho(\vec{x})$:

$$E = \int \Psi \hat{H} \Psi d\tau = W[\rho(\vec{x})].$$

By the variational theorem,

$$\tilde{E} \geq E = \int \Psi^* \hat{H} \Psi d\tau.$$
where $E$ is the eigenvalue of the Schrödinger equation

$$
\hat{H} \Psi = E \Psi.
$$

This energy eigenvalue is guaranteed to be the same functional of the true density:

$$
E = \int \Psi^* \hat{H} \Psi d\tau = W[\rho(\vec{x})],
$$

that is

$$
W[\tilde{\rho}(\vec{x})] \geq W[\rho(\vec{x})],
$$

for any $\tilde{\rho}(\vec{x}) \neq \rho(\vec{x})$ which establishes the required variational condition.
Appendix C

Calculation of EPR properties

For organic radicals including amino acid radicals, accurate results are usually obtained for g-tensors, hyperfine and quadrupole couplings and zero-field splittings. Basis sets such as Barone’s EPR-II and EPR-III\(^{192}\) have been developed particularly for DFT investigations of EPR parameters and are characterized by extra flexibility in the core region.

EPR experiments are parameterized by an effective spin-Hamiltonian (SH) that only contains spin-degrees of freedom. Computational chemistry relates the spin-Hamiltonian parameters to the microscopic interactions that occur in the full (relativistic) molecular Hamiltonian.

There are four contributions to the g-tensor:

\[
g_{kl} = g_{e} \delta_{kl} + \Delta g^{RMC}_{kl} + \Delta g^{GC}_{kl} + \Delta g^{OZ/SOC}_{kl}. \tag{C.1}\]

The first three terms are first-order contributions given by:

\[
\Delta g^{RMC}_{kl} = \frac{\alpha^2}{S} \sum_{\mu,\nu} P^{\alpha\beta}_{\mu\nu} \left\langle \varphi_{\mu} \middle| \frac{1}{2} \nabla^2 | \varphi_{\nu} \right\rangle, \tag{C.2}\]

and

\[
\Delta g^{GC}_{kl} = \frac{\alpha^2}{4S} \sum_{\mu,\nu} P^{\alpha-\beta}_{\mu\nu} \left\langle \varphi_{\mu} \left| \sum_{A} Z_{AR} r_{AR}^3 \left[ r_{AR} - r_{Akr} \right] \right| \varphi_{\nu} \right\rangle. \tag{C.3}\]
APPENDIX C. CALCULATION OF EPR PROPERTIES

The dominant contribution is provided by the second-order contribution of the fourth term.

\[
\Delta g_{kl}^{(OZ/SOC)} = \frac{1}{2S} \sum_{\mu\nu} \frac{P_{\alpha-\beta}^{\mu\nu}}{B_k} \langle \varphi_\mu | z_l^{SOMF} | \varphi_\nu \rangle
\]  

(C.4)

This term involves the derivative of the spin-density matrix with respect to a component of the magnetic field.

Hyperfine couplings consists of three expressions:

\[
A_{kl}^{(A;c)} = \delta_{kl} \frac{8\pi P_A}{3} \frac{\rho^{\alpha-\beta}(R_A)}{2S} 
\]  

(C.5)

\[
A_{kl}^{(A;d)} = \frac{P_A}{2S} \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} \langle \varphi_\kappa | z_l^{SOMF} | \varphi_\tau \rangle 
\]  

(C.6)

\[
A_{kl}^{(A;SO)} = -\frac{P_A}{S} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial I_k^{(A)}} \langle \varphi_\mu | z_l^{SOMF} | \varphi_\nu \rangle 
\]  

(C.7)

where \( P_A = g_e g_N \beta_e \beta_N \). The first two terms are expectation values that represent the Fermi contact interaction and the electron-spin nuclear spin dipolar interactions.
Bibliography


BIBLIOGRAPHY


(75) Frisch, M. J. et al., Gaussian09 Revision B.01, Revision B; Gaussian, Inc.: Wallingford, CT, 2009.


(137) Regelsberger, G; Jakopitsch, C; Rüker, F; Krois, D; Peschek, G. a.; Obinger, C J. Biol. Chem. 2000, 275, 22854–61.


BIBLIOGRAPHY


