Computational Investigations on Enzymatic Catalysis and Inhibition

Daniel James Simard
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
Simard, Daniel James, "Computational Investigations on Enzymatic Catalysis and Inhibition" (2015). Electronic Theses and Dissertations. 5440.
https://scholar.uwindsor.ca/etd/5440

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Computational Investigations on Enzymatic Catalysis and Inhibition

By

Daniel Simard

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Chemistry and Biochemistry
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

© 2015 Daniel J. Simard
Computational Investigations on Enzymatic Catalysis and Inhibition

by

Daniel Simard

APPROVED BY:

L. Porter
Department of Biological Sciences

C. MacDonald
Department of Chemistry & Biochemistry

J. Gauld, Advisor
Department of Chemistry & Biochemistry

September 03, 2015
Declaration of Co-Authorship

I. Co-Authorship Declaration

I hereby declare that this thesis material incorporates material that is a result of joint research as follows:

**Chapter 3** was done in collaboration with Hisham Dokainish under the supervision of Dr. James W. Gauld.

**Chapter 4** was done in collaboration with Rami Gerib under the supervision of Dr. James W. Gauld.

**Chapter 5** was done in collaboration with Grant Fortosky under the supervision of Dr. James W. Gauld.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-authors to include the above material in my thesis.

I certify that, with the above qualifications, this thesis, and the research to which it refers, is the product of my own work.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the
bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
Abstract

Enzymes are the bimolecular “workhorses” of the cell due to their range of functions and their requirement for cellular success. The atomistic details of how they function can provide key insights into the fundamentals of catalysis and in turn, provide a blueprint for biotechnological advances. A wide range of contemporary computational techniques has been applied with the aim to characterize recently discovered intermediates or to provide insights into enzymatic mechanisms and inhibition. More specifically, an assessment of methods was conducted to evaluate the presence of the growing number 3– and 4–coordinated sulfur intermediates in proteins/enzymes. Furthermore, two mechanisms have been investigated, the μ-OH mechanism of the hydrolysis of dimethylphosphate in Glycerophosphodiesterase (GpdQ) using five different homonuclear metal combinations Zn(II)/Zn(II), Co(II)/Co(II), Mn(II)/Mn(II), Cd(II)/Cd(II) and Ca(II)/Ca(II) as well as a preliminary study into the effectiveness of boron as an inhibitor in the serine protease reaction of class A TEM-1 β-lactamases.
Dedication

I dedicate this work to my family.
Acknowledgements

I would like to start by expressing my appreciation to a wonderful supervisor in Dr. James Gauld. Over the past three years, he has provided me with the opportunity to conduct research both during my undergraduate honours thesis project and as a Master’s student. During this time I have learned much about chemistry, enzymology and computation as well as gained invaluable interpersonal skills that I have no doubt will assist me in my future endeavours.

I would like to extend my appreciation to both Prof. Charles Macdonald and Prof. Lisa Porter for taking the time to read and evaluate my thesis.

I am also grateful to the Gauld Group, all the members both past and present, that I had the pleasure to work with. I strongly believe that the mischief and eccentric events that occurred on a daily basis in the lab would have given the humorous sitcom “The Office” a run for its money. On a more serious note, I am very appreciative to Dr. Hisham Dokainish who supervised me as an undergrad and made the transition into a Master’s student as seamless as possible. To all the other members, Dr. Eric Bushnell, Grant Fortosky, Bogdan Ion, Rami Gherib, Mohamed Aboelnga and Wanlei Wei as well as the numerous undergraduate students, thank you for all the support and laughs.

Finally, I would like to thank my family for all the support they have always provided me.
# Table of Contents

Declaration of Co-Authorship ........................................................................ iii

Abstract ........................................................................................................ v

Dedication ....................................................................................................... vi

Acknowledgements ......................................................................................... vii

List of Tables ................................................................................................ xi

List of Figures ................................................................................................ xii

List of Appendices ........................................................................................ xvi

List of Abbreviations/Symbols ..................................................................... xvii

**Chapter 1  Introduction** .............................................................................. 1

1.1 Enzymes ................................................................................................... 2

1.2 Enzymatic Protection .............................................................................. 2
    1.2.1 Protection from Oxidative Stress Using Sulfur ................................ 2
    1.2.2 Bioremediation Using GpdQ .......................................................... 4
    1.2.3 Inhibition of a Bacterial Resistance Mechanism ......................... 5

1.3 Computational Enzymology ................................................................... 5

1.4 References ............................................................................................... 6

**Chapter 2  Theoretical Methods** .............................................................. 10

2.1 Introduction ............................................................................................. 11

2.2 Computational Chemistry in Biological Systems ................................. 11

2.3 Molecular Mechanics ........................................................................... 12
    2.3.1 Molecular Dynamics ................................................................... 14
    2.3.2 Molecular Docking ...................................................................... 16

2.4 Quantum Mechanics ............................................................................. 19
    2.4.1 Higher Order Correlation Methods ........................................... 21
    2.4.2 Density Functional Theory ......................................................... 22

2.5 Quantum Mechanics/Molecular Mechanics ......................................... 23

2.6 Technical Aspects and Units ................................................................. 23
Chapter 3  A Density Functional Theory Assessment of 3– and 4– Coordinated Sulfur Species in Biological Systems ................................................................................. 28
3.1 Introduction .............................................................................................................. 29
3.2 Computational Methods ............................................................................................. 31
3.3 Results and Discussion .............................................................................................. 34
   3.3.1 Geometrical Assessment of Sulfenic Acid and Sulfoxide Derivatives ............... 35
      3.3.1.1 Series 1: Sulfenic Acid Derivatives ................................................................. 35
      3.3.1.2 Series 2: Sulfoxide Derivatives ....................................................................... 37
      3.3.1.3 Determination of Relative Energy Trends Between Sulfur Tautomers ....... 39
   3.3.2 Applications Into Biologically Relevant Intermediates ...................................... 41
      3.3.2.1 Archaeal Thioredoxin Peroxidase (ApTPx) ..................................................... 41
         3.3.2.1.1 The Sulfurane with a Neutral Histidine .................................................. 41
         3.3.2.1.2 The Sulfurane with a Protonated Histidine ............................................ 42
      3.3.2.2 Suliredoxins (Srx) .......................................................................................... 44
         3.3.2.2.1 Sulfinic acid ............................................................................................ 44
         3.3.2.2.2 Sulfinic Phosphoryl Ester ....................................................................... 45
         3.3.2.2.3 Thiosulfinate .......................................................................................... 45
      3.3.2.3 Ubiquitin-dependent Proteasomes ................................................................ 47
      3.3.2.4 Methionine Sulfoxide Reductase (Msr A & B) ............................................. 47
      3.3.2.5 Collagen IV ................................................................................................ 49
      3.3.2.6 Myelpersoxidase ........................................................................................... 50
   3.4 Conclusions ............................................................................................................. 51
3.5 References ............................................................................................................... 53

Chapter 4  Computational Insights into the Catalytic Mechanism of Homonuclear Glycerophosphodiesterase (GpdQ) ................................................................. 56
4.1 Introduction ............................................................................................................... 57
4.2 Computational Methods ............................................................................................ 60
   4.2.1 Preparation of the QM/MM Model ..................................................................... 60
   4.2.2 QM/MM Calculations ....................................................................................... 62
   4.2.3 Quantum Mechanical Cluster Calculation of Cd(II)/Cd(II) Complexes .......... 62
4.3 Results and Discussion

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.1 Reproducing the Catalytic Mechanism</td>
<td>63</td>
</tr>
<tr>
<td>4.3.2 DMP Binding to Form the Reactive Complex</td>
<td>66</td>
</tr>
<tr>
<td>4.3.3 DMP Transition State Structural Parameters</td>
<td>69</td>
</tr>
<tr>
<td>4.3.4 Metal Induced Phosphate Strain</td>
<td>70</td>
</tr>
<tr>
<td>4.3.5 Natural Bond Order (NBO) Analysis</td>
<td>72</td>
</tr>
<tr>
<td>4.3.6 Substituting Calcium Ions for Mα and Mβ</td>
<td>75</td>
</tr>
<tr>
<td>4.3.7 Mα-ligand Substitutions for Homonuclear Cd(II)/Cd(II)</td>
<td>76</td>
</tr>
<tr>
<td>4.4 Conclusions</td>
<td>79</td>
</tr>
<tr>
<td>4.5 References</td>
<td>80</td>
</tr>
</tbody>
</table>

Chapter 5  Insights into the Use of Boron for Inhibiting β-lactamases

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>85</td>
</tr>
<tr>
<td>5.2 Computational Methods</td>
<td>88</td>
</tr>
<tr>
<td>5.2.1 Molecular Docking and Molecular Dynamic Simulations</td>
<td>88</td>
</tr>
<tr>
<td>5.2.2 The QM/MM Model and Calculations</td>
<td>89</td>
</tr>
<tr>
<td>5.3 Results and Discussion</td>
<td>90</td>
</tr>
<tr>
<td>5.3.1 Selection of the Benzylpenicillin and the Boron Substitute</td>
<td>90</td>
</tr>
<tr>
<td>5.3.2 Molecular Docking Analysis</td>
<td>91</td>
</tr>
<tr>
<td>5.3.3 Molecular Dynamic Analysis</td>
<td>92</td>
</tr>
<tr>
<td>5.3.4 Reproducing the Catalytic Mechanism of Benzylpenicillin</td>
<td>93</td>
</tr>
<tr>
<td>5.3.5 The Catalytic Mechanism of Boron Substituted Benzyl Penicillin</td>
<td>97</td>
</tr>
<tr>
<td>5.3.6 Implications of the Preliminary Results</td>
<td>101</td>
</tr>
<tr>
<td>5.4 Conclusions</td>
<td>102</td>
</tr>
<tr>
<td>5.5 References</td>
<td>103</td>
</tr>
</tbody>
</table>

Chapter 6  Conclusions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Conclusions</td>
<td>109</td>
</tr>
<tr>
<td>6.2 References</td>
<td>111</td>
</tr>
</tbody>
</table>

Vita Auctoris                                                                                           113
List of Tables

Table 4.1. The transition state energies calculated for the hydrolysis of DMP for the four metals combinations using five different exchange-correlation functionals. .......................... 66

Table 4.2 The geometric parameters for reactive complex (RC), transition state (TS) and product complex (PC) for the cleavage of DMP ................................................................. 67

Table 4.3. The metal separation in the active site of GpdQ in Angstroms (Å) and phosphate angle in degrees (°) .................................................................................................................. 72

Table 4.4. NBO calculated natural charges for both reactant complex (RC) and transition state structures (TS) ........................................................................................................... 74

Table 4.5. Selected parameters of the Ca(II)/Ca(II) derivative ............................................ 76

Table 4.6 Selected parameters in the Mα substituted ligands ............................................. 78

Table 5.1. Key average distances obtained throughout the MD simulation .................... 93
List of Figures

Figure 1.1. Pictorial illustration of the sulfur intermediates that have been discovered or proposed to occur in various enzymes. Atom colours: yellow=S, orange=P, blue=N, red = O, grey=C and white=H. ................................................................. 3

Figure 1.2. Select organophosphates that can be cleaved using bioremediation. Atom colours: pink=P, yellow=S, blue=N, red=O, grey=C and white=H. ........................................... 5

Figure 2.1. A pictorial illustration of a wide range of timescales that is relevant to studying biological systems. .............................................................. 12

Figure 2.2. A ball and stick representation of the general terms that are considered in molecular mechanics.......................................................... 14

Figure 2.3. The total number of publications that involve a molecular docking study as a stand-alone or a complement to experimental work. Data collected from Web of Science using the criteria Topic = molecular docking. .............................................................. 17

Figure 3.1. Illustrations of several 3– and 4–coordinated sulfur intermediates found in various proteins that were investigated in this study. ............................................. 30

Figure 3.2. Schematic illustrations of the compounds used in the systematic study of a) sulfenic acid derivatives and b) sulfoxides derivatives................................................... 33

Figure 3.3. Schematic illustrations of the eleven modeled 3– and 4–coordinated sulfur compounds found in 6a,b) ApTPx, 7a–c) SrX, 8) Ubiquitin–dependent proteasomes, 9a–b) Msr 10) Collagen IV and 11) Myeloperoxidase. .............................................................. 34

Figure 3.4. Plot of the optimized r(S–OH) bond distances in Å for ethyl sulfenic acid (EtSOH) Figure 3.2, 3a. Color-coding Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively. ............................................. 37
Figure 3.5. Plot of the optimized r(S–OH) bond distances in Å for ethyl sulfoxide (EtSO) Figure 3.2, 3b. Benchmark: Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.

Figure 3.6. a) Schematic showing an example of the difference between one of the five tautomers used to compare relative energies. b) Plot of the relative energy values after optimization for the tautomers of ethyl sulfenic acid and ethyl sulfoxide. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.

Figure 3.7. Plots for optimized bond lengths of a) S–OH, b) S–N of un-protonated for hydroxy-imidazolyl-λ4-sulfanyl methane and c) S–OH and d) S–N of protonated for hydroxy-imidazolyl-λ4-sulfanyl methane from Figure 3.3, 6a and b. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.

Figure 3.8. Plot of the optimized for a) the SO–PO–2 bond and b) S–OPO–2 bond distances in Å for ethyl sulfinic phosphoryl ester representing the first intermediate in Srx from Figure 3.3, 7b and c) the S–S bond distances of thiosulfinate representing the second intermediate in Srx Figure 3.3, 7c. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.

Figure 3.9. Plots of the optimized for S–S bond distances in Å for a) sulfurane, from Figure 3.3, 9b. b) sulfonium cation, from Figure 3.3, 9c from Msr A and B. c) S–OH bond distances in the sulfurane intermediate. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.
Figure 3.10. Plot of the optimized for S=N bond distances in Å of the sulfimine protein cross linker found in collagen IV. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311+G(2df,p) and MP2/6-311++G(3df,3pd) in black and red respectively. .............................................. 50

Figure 3.11. Plot of the optimized for S–N bond distances in Å of the sulfinamide protein cross linker found in Myeloperoxidase. Benchmark QCISD 6-311++G(3df,3pd) from Figure 3.3, 11. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively. ................................................................. 51

Figure 4.1. Comparing the similar active site of three different classes. A) OdpA/PTE, b) PAP and c) GpdQ with their presumed native metal composition ................................................. 58

Figure 4.2. The mechanism for the a) terminal hydroxide based hydrolysis of a general diphosphate by GpdQ vs. b) the bridging hydroxide mechanism proposed mechanism of PAP. ................................................................................................................................. 59

Figure 4.3. The QM/MM model used in the study of GpdQ. The QM layer is shown in the foreground and the MM layer is faded in the background for clarity. Atom colours, Red=O, Blue=N, Orange=P, Dark Yellow=S, Grey=C, White=H, Pale Yellow=general metal ......................................................................................................................... 61

Figure 4.4. Pictorial representation of the reactant complex RC, transition state TS and product complex PC for DMP cleavage. Only the QM layer is shown with the catalytically relevant atoms enlarged. Atom colours, Red = O, Blue = N, Orange = P, Dark Yellow = S, Grey = C, White = H, Pale Yellow = general metal ................................................................. 64

Figure 4.5. Pictorial representation of the substitution on the Ma-site. Blue font colour represents the change from the wild type (top). ........................................................................................................ 77

Figure 5.1. The proposed mechanism for the acylation step with Glu166 as the catalytic base. A) Activation of serine B) the unstable tetrahedral intermediate C) Protonation of amine and ring opening D) Acyl-enzyme ......................................................................................... 86
**Figure 5.2.** The four currently approved β-lactamases inhibitors and the most promising boronic acid inhibitors. .......................................................... 87

**Figure 5.3.** Comparative models of A) benzylpenicillin and B) boron substitute used in this study. .......................................................... 89

**Figure 5.4.** Pictorial representation of the QM layer modeled in the study, the MM was excluded for clarity. Atom colours, Red=O, Blue=N, Yellow=S, Grey=C, White=H, .... 90

**Figure 5.5.** An pictorial overlay of the docked structures of benzylpenicillin and the boron analogue. Atom colours: grey=C, red=O, blue=N yellow=S orange=B and white=H .......................................................... 92

**Figure 5.6.** Key structural parameters for the acylation step of benzyl penicillin. The first part of the mechanism, the nucleophilic attack, is highlighted in red. The main atoms involved in the second part, the C-N cleavage are highlight in blue. All bond lengths are in Angstroms (Å). .......................................................... 94

**Figure 5.7.** Potential energy surface representing the acylation of benzylpenicillin in TEM-1. Atom colours: grey=C, red=O, blue=N and white=H........................................... 96

**Figure 5.8.** Key structural parameters for the acylation step of benzyl penicillin. The first part of the mechanism, the nucleophilic attack, is highlighted in red. The main atoms involved in the second part, the C-N cleavage are highlight in blue. All bond lengths are in Angstroms (Å). .......................................................... 99

**Figure 5.9** Potential energy surface representing the acylation of benzylpenicillin in TEM-1. Atom colours: grey=C, red=O, blue=N and white=H................................. 100
List of Appendices

Table A1. Cartesian coordinates for optimized small model sulfenic acid derivatives using eight basis sets and six functionals (see text).

Table A2. Cartesian coordinates for sulfenic acid tautomer derivatives using eight basis sets and six functionals (see text).

Table A3. Cartesian coordinates for all the biologically characterized 3- and 4-coordinated sulfur species using eight basis sets and six functionals (see text).

Table B1. Cartesian coordinates the reactant complex, transition state and product complex for Zn(II)/Zn(II) combinations using all five functionals with the 6-31G(d).

Table B2. Cartesian coordinates the reactant complex, transition state and product complex for Co(II)/Co(II) combinations using all five functionals with the 6-31G(d).

Table B3. Cartesian coordinates the reactant complex, transition state and product complex for Mn(II)/Mn(II) combinations using all five functionals with the 6-31G(d).

Table B4. Cartesian coordinates the reactant complex, transition state and product complex for Cd(II)/Cd(II) combinations using all five functionals with the 6-31G(d).

Table B5. Cartesian coordinates the reactant complex, transition state and product complex for Ca(II)/Ca(II) combinations using all five functionals with the 6-31G(d).

Table B6. Cartesian coordinates the reactant complex, transition state for the mutated Cd(II)/Cd(II) combinations using M06 and 6-31G(d,p) basis set.

Table C1. Cartesian coordinates for the hydrolysis of benzylpenicillin at the B3LYP/6-31G(d) level of theory.

Table C2. Cartesian coordinates for the hydrolysis of boron substituted benzylpenicillin at the B3LYP/6-31G(d) level of theory.

**For access to appendices, please contact Dr. James W. Gauld
## List of Abbreviations/Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AO</td>
<td>Atomic orbital</td>
</tr>
<tr>
<td>ApTPx</td>
<td>Archaeal thioredoxin peroxidase</td>
</tr>
<tr>
<td>BO</td>
<td>Born-Oppenheimer</td>
</tr>
<tr>
<td>bpNPP</td>
<td>Bis(p-nitrophenyl)phosphate</td>
</tr>
<tr>
<td>DFT</td>
<td>Density functional theory</td>
</tr>
<tr>
<td>DMP</td>
<td>Dimethylphosphate</td>
</tr>
<tr>
<td>FF</td>
<td>Force field</td>
</tr>
<tr>
<td>GGA</td>
<td>Generalized gradient approximation</td>
</tr>
<tr>
<td>GpdQ</td>
<td>Glycerophosphosdiesterase</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree-Fock</td>
</tr>
<tr>
<td>IEFPCM</td>
<td>Integral equation formalism polarization continuum method</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MM</td>
<td>Molecular mechanics</td>
</tr>
<tr>
<td>MOE</td>
<td>Molecular operating environment</td>
</tr>
<tr>
<td>MP2</td>
<td>Möller Plesset perturbation doubles</td>
</tr>
<tr>
<td>Msr</td>
<td>Methionine sulfoxide reductase</td>
</tr>
<tr>
<td>NAMD</td>
<td>Nanoscale molecular dynamics</td>
</tr>
<tr>
<td>ONIOM</td>
<td>Our Own N-Integrated Molecular Orbital Molecular Mechanics</td>
</tr>
<tr>
<td>OP</td>
<td>Organophosphate</td>
</tr>
<tr>
<td>PAP</td>
<td>Purple acid phosphates</td>
</tr>
<tr>
<td>PBP</td>
<td>Penicillin binding proteins</td>
</tr>
<tr>
<td>PC</td>
<td>Product complex</td>
</tr>
<tr>
<td>PES</td>
<td>Potential energy surface</td>
</tr>
<tr>
<td>Prx</td>
<td>Peroxiredoxins</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Srx</td>
<td>Sulfiredoxins</td>
</tr>
<tr>
<td>QCISD</td>
<td>Quadratics configuration interaction of singles and doubles</td>
</tr>
<tr>
<td>QM</td>
<td>Quantum mechanics</td>
</tr>
<tr>
<td>QM/MM</td>
<td>Quantum mechanics/molecular mechanics</td>
</tr>
<tr>
<td>RC</td>
<td>Reactant complex</td>
</tr>
<tr>
<td>Trx</td>
<td>Thioredoxins</td>
</tr>
</tbody>
</table>
Chapter 1  Introduction
1.1 Enzymes

The importance of enzymes to life cannot be understated because they are able to facilitate chemical reactions at life sustainable rates.\textsuperscript{1} They are by definition biocatalysts therefore they perform their function without being consumed in the reaction.\textsuperscript{1} What makes enzymes remarkable is that they are able to significantly enhance catalysis often at physiological conditions.\textsuperscript{2} A well-illustrated example of this is in the decarboxylation reaction of orotidine 5-phosphate. In absence of the enzyme, the decarboxylation process has half life of 78 million years, however, when orotidine 5-phosphate decarboxylase (ODCase) is present there is a $10^{17}$-fold rate enhancement that decreases the half-life to 18 milliseconds.\textsuperscript{3} A second advantage of enzymes over traditional catalysts is its high selectivity. This may be best demonstrated in the family of aminoacyl-tRNA synthetases that can discriminate between serine and cysteine (differing by only a single atom) by a factor of $10^{8}-5$. It is clear that enzymes have inherent properties that are important to deeply understand.

1.2 Enzymatic Protection

Enzymes have long list of functions,\textsuperscript{6} however protecting the cell from cellular stress can connect the work performed herein. There are three permutations of how an enzyme can alter a toxin to a less toxic forms: (1) they can catabolically break down the pathogen,\textsuperscript{7} (2) anabolically build a larger molecule,\textsuperscript{8,9} or (3) isomerize.\textsuperscript{10} Three topics where the enzyme’s role can be attributed to the breakdown down harmful toxins are described in the following sections.

1.2.1 Protection from Oxidative Stress Using Sulfur

Possibly the most unavoidable stressor that occurs in our cells is oxidative stress.\textsuperscript{11} This is partly due to the necessity of oxygen to produce energy.\textsuperscript{12} In addition, intrinsic oxidative stress can be produced by enzymes themselves such as in the peroxisomes, during phagocytosis or as byproducts of standard P450 reactions.\textsuperscript{13-14} Oxidative stress can
also be caused from extrinsic factors such as from UV radiation, exercise or toxins from air pollution or smoking. Consequently, the cellular defense mechanism is to produce antioxidants to remove them. Glutathione (GSH), lipoic acid, and coenzyme Q are small molecules natural antioxidants. However, several enzymes have evolved that can efficiently remove ROS as seen in catalase (CAT), glutathione peroxidases (GPx), thioredoxins (Trx), peroxiredoxins (Prx) and methionine (MSR) where more than 10^8 ROS can be removed per second.

**Figure 1.1.** Pictorial illustration of the sulfur intermediates that have been discovered or proposed to occur in various enzymes. Atom colours: yellow=S, orange=P, blue=N, red = O, grey=C and white=H.
Chapter 1: Introduction

Understandably, the source of enzyme’s catalytic power is important to study due to its efficiency and perhaps more importantly, the variability of the mechanism of action. For instance, CAT reduces ROS with an iron-heme group, Gpx uses selenium redox center, and the remaining enzymes: Trx, Prx, Srx and Msr utilize sulfur to rid of ROS. Due to the variability of the possible redox states for sulfur, nature utilized several unique 3– and 4–coordinated sulfur intermediates to perform this function (Figure 1.1). These include: sulfuranes in archaeal thioredoxin peroxidase (ApTPx),17 methionine sulfoxide reductase (Msr),18-19 In addition, sulfoxides,18,20 sulfonic acid,21 sulfinamides,22-23 sulfenic acid phosphoryl esters,20 and sulfilimines24 have been found. These intermediates were assessed in Chapter 3 to set a foundation for future studies.

1.2.2 Bioremediation Using GpdQ

Organophosphates (OPs) are commonly used as insecticides to minimize crop damage.25-26 Unfortunately after a heavy rainfall, the OPs are washed away into the waterways where they can cause severe ailments in humans.27-28 It is estimated that 3 million humans are affected from OP poisoning annually, therefore innovative solutions are required.29 One possibility is to use enzymes as a form of bioremediation because it is well known that certain enzymes have the inherent ability to break down phosphates.30-31 The most ideal candidate would then have two properties: high catalytic rate and promiscuity.32 The latter is a necessary because OPs come in all different shapes and sizes Figure 1.2. A class of enzymes that fit this description is binuclear hydrolases. These include: OPH from Pseudomonas diminuta (PTE),33-34 purple acid phosphates (PAP),35-36 OpdA from Agrobacterium radiobacter, and glycerophosphodiesterase from Enterbacter aerogene (GpdQ).37-38 GdpQ in particular, has the ability to decompose mono, di and triphosphoesters.39 However, the atomistic mechanistic details are not well defined. In Chapter 4, a possible catalytic mechanism for GdpQ was explored to gain a better understanding of how it performs its function.
Chapter 1: Introduction

Figure 1.2. Select organophosphates that can be cleaved using bioremediation. Atom colours: pink=P, yellow=S, blue=N, red=O, grey=C and white=H.

1.2.3 Inhibition of a Bacterial Resistance Mechanism

In some cases, enzymatic protection mechanisms work against us. This is true in case of bacterial resistance. Bacteria have evolved to protect themselves from potent external toxins, i.e. antibiotics. Some antibiotics function by blocking bacterial transpeptidases activity in penicillin-binding proteins (PBPs) that are responsible for building their cell walls. This has been an extremely successful target for inhibition due to the lack of cell walls in mammals. However, enzymes known as β-lactamases have developed over their evolution that are able to hydrolyze these antibiotic before they can inhibit the PBPs. Today there have been about 1300 distinct subclasses have been identified and various approaches for inhibiting these enzymes are required. One approach is to use boron as a Lewis acid because β-lactamases are often serine proteases that initiate through a nucleophilic attack. Boron’s empty p-orbital would, in theory, be a great electrophile to increase reactivity. This feature would commence a unique reversible mechanism that has not been energetically characterized. Chapter 5 studied this possibility.

1.3 Computational Enzymology

Computation is a cost efficient strategy to understand properties of enzymes. Perhaps one of the strongest advantages of this field is its ability to characterize short-lived
intermediates and transition states that would be extremely difficult or impossible to do experimentally. Such atomistic details can provide an unprecedented foundation for rational design of biomimics or therapeutics. A multi-scale computational enzymology approach is used in this thesis. This includes the combination of small models, quantum mechanical clusters, quantum mechanics/molecular mechanics, molecular dynamics as well as docking to investigate enzymes in the various systems that have been discussed.

1.4 References

Chapter 1: Introduction

Chapter 1: Introduction


Chapter 2  Theoretical Methods

MOLECULAR DOCKING

MOLECULAR DYNAMICS

QM CLUSTER

QM/MM
2.1 Introduction

Computational chemistry can be defined as a toolbox of computer-facilitated-techniques with the purpose to understand numerous chemical phenomena.\(^1\) These include but are not exclusive to: molecular geometries, energetics, reactivity, IR, UV, NMR and EPR.\(^1\) One particular field in computation chemistry that has gained recent global recognition is multi-scale modeling of complex systems.\(^2\) Contributions to its development awarded Karplus, Levitt and Warshel the Nobel Prize in chemistry in 2013.\(^2-3\) In this thesis, contemporary multi-scale modeling techniques were used to study biological enzymes and the underlying concepts for these methods. Specifically, molecular and quantum mechanics are described in this chapter.

2.2 Computational Chemistry in Biological Systems

The focus of this work is computational investigations into biological systems, i.e. understanding the atomistic details that occurs within living organisms. Biological systems contain numerous macromolecules such as lipids, nucleic acids, carbohydrates and proteins that can sometimes be very difficult to study experimentally.\(^4\) For example, it can be difficult to determine a catalytic mechanism of a protein (enzyme) or to monitor the dynamics of a lipid membrane without the aid of calculation due to their size and the time scale of which they occur.\(^5-6\) Computational chemistry has the ability to provide insight into these features; however, there are still numerous challenges that must be addressed. A marriage of accuracy and efficiency must be sought after to provide meaningful results in a timely matter. For example, it is now possible to accurately calculate properties dipeptides using high-order correlated wave function based methods to excruciating accuracy.\(^7\) However, an enzyme can be 1000+ fold larger and it is just not feasible to take on that approach. In addition to size, biological processes occur over a large time scale, ranging from femtoseconds to seconds, Figure 2.1.\(^6-8-10\) Strategic approaches must be used to overcome such chemical diversity to pull out useful valuable information. One approach, for example, is to use molecular dynamics simulations that
use simpler mathematical formulae such as classical mechanics to evaluate how protein motion over longer periods of time. Moreover, more complex quantum mechanical methods can look at only the specific atoms involved in catalysis that occurs over picoseconds.\(^\text{11}\) Combining two or more methods to attempt to fully describe these systems are commonly referred to as multi-scale enzymology.\(^\text{12-13}\)

Figure 2.1. A pictorial illustration of a wide range of timescales that is relevant to studying biological systems.\(^\text{10}\)

**PART 1**

**2.3 Molecular Mechanics**

Molecular mechanics (MM) is the simplest method to describe how a molecular system behaves. An analogy that is often used to describe MM is that it is a “ball & spring” approach.\(^\text{14}\) Atoms are treated as spherical point charges that are connected through a potential “spring”.\(^\text{14}\) Consequently, this method cannot model a chemical reaction because the behavior of individual electrons are not properly described.\(^\text{14-15}\) However, this method does have great application because several chemical processes do not require bond orders to change in order to characterize a system. Examples include: energy minimizations, protein folding over time or quantifying binding interactions.\(^\text{16}\)
Chapter 2: Theoretical Methods

The driving force of molecular mechanics can be summed into three terms; atomic position \( r \), force \( F(r) \) and potential energy \( U(r) \). The relationship between these variables is shown in Equation 2.1.

\[
F(r) = - \frac{\partial U(r)}{\partial r} \tag{2.1}
\]

In a typical molecular mechanics calculation, the initial atomic coordinates \( r \) are required. When studying proteins, the source of these coordinates is often X-ray crystallographic or NMR derived data that can be found in the protein data bank (PDB). Next, an approximation of the potential energy \( U(r) \) is calculated. This is possible by using an empirical function known as a force field (FF), Equation 2.2. The FF can be thought of mathematical equivalent of “a bridge” between atomic position and how they influence one another. Unfortunately, a universal formula does not exist therefore, the one selected for calculation is chosen based on the system that for which it was built. CHARMM,\textsuperscript{17-18} AMBER\textsuperscript{19-20} and MMFF94\textsuperscript{21} are good FFs to use when studying proteins. All these force field have the general equation, Equation 2.2:

\[
U_{total} = U_{bond} + U_{angle} + U_{dihedral} + U_{vdW} + U_{Coulomb} \tag{2.2}
\]
Figure 2.2. A ball and stick representation of the general terms that are considered in molecular mechanics.

Each potential (U) term in Equation 2.2 houses a different internal equation. The first three terms are intramolecular parameters. $U_{bond}$ is the potential energy stored in covalent bonds. $U_{angle}$ represents three-center bond angle. $U_{dihedral}$ represents the torsion (dihedral) angle that describes relationship between two atoms that are three covalent bonds apart where the central bond is free to rotate. The final two terms in Equation 2.2 are intermolecular interactions. $U_{vdW}$ represents the van der Waal’s forces and $U_{Coulomb}$ represents the electrostatic interactions. The solution of $U_{total}$ will give a potential energy that is representative of the starting atomic coordinates. Finally, since the potential and coordinates are now known, the force applied on each atom can be determined as per Equation 2.1. From this point, Newton’s second law can relate position (r) to time (t) to create a trajectory for the atoms. This is the foundation of molecular dynamics.

2.3.1 Molecular Dynamics

Molecular dynamics (MD) contains the mathematical equations to predict how molecules in a system move over a period of time. Alder and Wainwright were the founders of modern day MD.22-23 However, the first molecular dynamic simulation performed on biological relevant molecules occurred in 1977 when Nobel Prize laureate Karplus and coworkers studied bovine pancreatic trypsin inhibitors (BTPI, 910 atoms).24
This simulation was performed over 9.2 ps. Today, a 10³-fold to 10⁶-fold increase (nanoseconds to microseconds) can be performed in a simulation.

A general scheme to a MD simulation is outlined below:  

1. **Assign initial position to atoms r (t = 0).** This usually comes from crystallographic data.

2. **Prepare and parameterize the system.** Crystallographic data does not include the position of hydrogen atoms and the structure is generally strained in order to produce the crystal. Therefore the valence is satisfied by the addition of the appropriate number of hydrogen atoms, re-solvated, and energy minimized using molecular mechanics.

3. **Choose an appropriate ∆t and total time-scale.** Generally speaking, for enzymatic systems, ∆t is between 1-2 fs and total time is between 1 - 1000 ns. The total will vary based on what is being calculated. The reason for a seemingly small ∆t is that the error is proportional to ∆t² thus cannot be too large or the calculation becomes unreliable.

4. **Calculate the potential energy function.** Using an empirical force field and convert this energy into a force function, Equation 2.1.

5. **Move the atoms.** Determine new atomic positions using Newtonian equations, Equations 2.3 and 2.4.

\[
F(r) = m a(r) \tag{2.3}
\]

\[
r( t + \Delta t) = r( t) + v(t)\Delta t + \frac{1}{2} a\Delta t^2 \tag{2.4}
\]

where \( t = \sum \Delta t \)
6. **Repeat.** Recalculate potential as in step 4 continue until total time-scale is met.

There are different algorithms that exist for MD simulations that all use, in some form, this general outline. In this thesis, the Nanoscale Molecular Dynamics (NAMD) program is used to perform the simulations.\(^{27}\) NAMD is tailored to study large bimolecular systems such as proteins, as well as lipid membranes.\(^{27}\) Another benefit is that it has efficiently integrated the Newtonian equations while adequately controlling the temperature and pressure of the system. This is a necessity for accuracy in physiological conditions.\(^{27-28}\) It is accepted that NAMD is well-suited for running such simulations.\(^{28}\)

### 2.3.2 Molecular Docking

A second common application of MM in studying biological systems is molecular docking. This allows one to explore several conformations of a ligand in a receptor-binding site. The purpose of docking is often to find and rank the optimal binding geometries and associated energies with the use of a scoring function.\(^{29}\) As a computational chemist, there are several efficient docking programs that exist such as DOCK,\(^{30}\) AutoDock,\(^{31}\) FlexX,\(^{32}\) GOLD,\(^{33}\) and GLIDE.\(^{34}\) Docking has gathered great interest for both experimentalist and computational chemist alike. This is illustrated by the massive increase in publications that contain a docking protocol, Figure 2.3. For instance, since the turn of the century, there have been over 35000 publications that include molecular docking.
Figure 2.3. The total number of publications that involve a molecular docking study as a stand-alone or a complement to experimental work. Data collected from Web of Science using the criteria Topic = molecular docking.

There are three structural components to consider in a docking simulation:\textsuperscript{35}

1. Receptor
2. Pocket
3. Ligand

In the case of proteins, the receptor is the entire protein. The pocket is the specific region that the ligand directly interacts with. This can be a pre-defined pocket or the entire surface of the protein if it is not known. Finally, the ligand is the molecule whose binding is of interest. In general, there are five steps that occur during a docking procedure that have been used herein that are outlined below.\textsuperscript{4,35-37}

1. Conformational analysis is to explore different intramolecular conformations that are favourable in energy. This will scan favourable torsion angles and will not alter the bond lengths or the 3-center bond angles (this may be altered in the
refinement stage). There may be tens of thousands different possibilities found and any duplicates and ones that have steric clash are deleted.

2. **The placement stage** involves putting the resulting conformers into a predefined receptor pocket in several orientations that fit. Here, hundreds of structures are generated.

3. **The first scoring step**: is then used to rank all the determined structures. There are several different types of scoring functions that exist. The one that is used in this thesis is the London dG scoring function (Equation 2.5).

\[
\Delta G = c + E_{\text{flex}} + \sum_{h-\text{bonds}} c_{HB}f_{HB} + \sum_{m-\text{lig}} c_{M}f_{M} + \sum_{\text{atoms } i} \Delta D_i
\]  

(2.5)

where c is the average gain or loss of entropy from rotation and translation. This allows the calculation to be \( \Delta G \) instead of just the potential enthalpy \( \Delta H \). \( E_{\text{flex}} \) is the flexibility of the ligand. \( c_{HB} \) is the energy of an ideal hydrogen bond whereas \( f_{HB} \) calculates the imperfections that will deviate it from the ideal (coefficient would be between 0 and 1). \( \sum_{m-\text{lig}} c_{M}f_{M} \) deals with metals and \( \Delta D_i \) measures the desolvation of an atom.

4. **Refinement**: Once all the poses are scored, the user defines the number of poses that are allowed to enter the second stage of docking. The pocket up to this point is strictly rigid and has not adjusted to this ligand. The refinement uses a MM force field minimization of the protein receptor side chains to help relax the side chains to fit that pose.

5. **The second rescoring** will then calculate the energy of binding again and rank the new list after minimization giving an output of the top pose.
All the aforementioned methods use classical mechanics as the foundation of their calculations. When the chemistry of bond formation and breaking is important to the chemical system, a more sophisticated theory is required. This is known as quantum mechanics.

\section*{PART 2}

\subsection*{2.4 Quantum Mechanics}

A brief introduction into the foundational equations of quantum mechanics is discussed. To start, the most famous equation that is associated with the words “quantum mechanics” is no doubt the Schrödinger equation (Equation 2.6).

\begin{equation}
\hat{H}\psi = E\psi \tag{2.6}
\end{equation}

where $\hat{H}$ is the Hamiltonian operator, $\psi$ is the wave equation and $E$ is the energy eigenfunction. The wave equation, which cannot be experimentally measured, is postulated to contain all possible information regarding the chemical system.\textsuperscript{38} Thus, a Hamiltonian operator can act on the wave equation to extract specific energies about that system. The total molecular Hamiltonian operator is:

\begin{equation}
\hat{H}_{tot} = -\frac{1}{2} \sum_{i=1}^{e} \nabla_i^2 - \frac{1}{2} \sum_{A=1}^{N} \frac{1}{M_A} \nabla_A^2 - \sum_{i=1}^{e} \sum_{A=1}^{N} \frac{Z_A}{r_{iA}} + \sum_{A=1}^{N} \sum_{B>A}^{N} \frac{Z_A Z_B}{r_{AB}} + \sum_{i=1}^{e} \sum_{j>i}^{e} \frac{1}{r_{ij}} \tag{2.7}
\end{equation}

where $i$ and $j$ represent different electrons and $A$ and $B$ represent different nuclei. The first two terms are kinetic terms of the electrons and nuclei respectively. The next three
terms are the potential energy terms for electron-nuclear, nuclear-nuclear and electron-electron interactions respectively. Unfortunately, the total molecular Hamiltonian is too difficult to solve exactly therefore approximations are introduced. The first of which is the Born-Oppenheimer (BO) approximation. Since a nucleus is about $10^3$ fold more massive than an electron, the kinetic energy is negligible for the nuclei with respect to the electrons. This can set the nuclear kinetic energy to 0. In addition, the nuclear-nuclear repulsion potential and can be approximated to a constant. After these two simplifications, the resultant is the electronic Hamiltonian, Equation 2.8.

$$
\tilde{H}_{\text{elec}} = -\frac{1}{2} \sum_{i=1}^{e} \nabla_i^2 - \sum_{i=1}^{e} \sum_{A=1}^{N} \frac{z_A}{r_{iA}} + \sum_{i=1}^{e} \sum_{j>i}^{e} \frac{1}{r_{ij}}
$$

(2.8)

This the starting point to solve for the energy of a system. The only term that is not straightforward to solve is the final electron-electron repulsion term in a many electron system. An accepted way to treat this is to assume that the each electron moves independently. This gives rise to a series of one-electron spin orbitals, Hartree-product, that are much easier to treat mathematically. There are several different ways to further describe the electrons in a QM calculation that varies based on the approximations used. Nevertheless, at the foundation is the Hartree-Fock method that treats the electrons as an average “smeared” electrostatic potential. As one can imagine, averaging electrons over space can lead to large errors because in a real system the motion of each individual electron influences each other.
2.4.1 Higher Order Correlation Methods

The difference in energy between the HF electron averaging, $E_{HF}$, and the real system, $E_{exact}$ is known as electron correlation $E_{corr}$, Equation 2.9. Higher order correlation methods have been developed to minimize the discrepancy between these two terms.\textsuperscript{40-41}

$$E_{exact} = E_{HF} + E_{corr} \quad (2.9)$$

There are two high order correlation methods, out of many, that are used in thesis, quadratic configuration interaction of singles and doubles (QCISD) and Möller Plesset perturbation doubles (MP2).\textsuperscript{38,40-41} In the former, new determinates are added to the HF wavefunction as a means to promote electrons to virtual orbitals. If all the possible excitations are included this would be called full-CI and the exact energy of the system would be known.\textsuperscript{42} As described by the name, QCISD allows for only single and double excitations.

In the latter case, MP2 uses perturbations to treat the electron correlation. They are often referred to as MPn methods where n is the order.\textsuperscript{38} Furthermore, the order represents the number of terms in the Taylor Series that are calculated. The pitfall for solving both of these post Hartree-Fock methods is that they are very costly in terms of the time it takes to achieve this accuracy. For example, QCISD scales $N^6$ and MP2 ~ $N^5$, where N is size of the system.\textsuperscript{40} Therefore, higher order correlation methods are often restricted to studying small systems.
2.4.2 Density Functional Theory

Fortunately, there exists an alternate theory that can synergize efficiency with accuracy known as density functional theory (DFT). DFT is different from wavefunction-based methods because it uses an electron probability density function, an observable property. Its aim is to calculate all physical ground state chemical properties from only the electron density.\(^{38}\) The electron probability density function is the chance of finding electron in a predefined space, Equation 2.10.\(^{43}\)

\[
\rho (x,y,z) d x d y d z
\]  \hspace{1cm} (2.10)

This equation comes with benefits compared to conventional wavefunction methods. The function is dependent on a total of three spatial components, \(x, y\) and \(z\) regardless of the size of the molecule. In contrast, wavefunction methods are dependent on three spatial and one spin component. This reduces the complexity of the equations used and in result speeds up calculations considerably.

The pitfall for DFT is the relationship between the electron density and the energy of the system is not known. However, it was eventually shown that the exact energy can be divided into five terms Equation 2.11:\(^{44}\)

\[
\]  \hspace{1cm} (2.11)

Where \(E^T\) is the kinetic energy of non interacting electrons, \(E^V\) is the total nuclear potential that includes nuclear-electron attraction and nuclear-nuclear repulsion, \(E^I\) is the electron potential that includes the electron-electron repulsion and \(E^{XC}\) is the exchange
correlation. Unfortunately, there is no known solution for the $E^{xc}$ term therefore it is approximated.\textsuperscript{44} This is why there are several functionals that exist that are trained against certain types of systems. It is important to know the limitation of each functional used to gauge the accuracy of the results produced.

### 2.5 Quantum Mechanics/Molecular Mechanics

Finally, the pioneer work combining quantum mechanics and molecular mechanics (QM/MM) was done by the Nobel laureates Warshel and Levitt in 1976.\textsuperscript{45} To illustrate the growth of the QM/MM method, there have been over 100 unique studies on different enzymes every year since 2007.\textsuperscript{46} The aim of quantum mechanics/molecular mechanics is to combine two theories to use large models with high accuracy without substantially increasing the amount of time it takes to run a calculation. This method brings in the best of both methods.

To briefly describe the QM/MM model, the inner subsection is treated with quantum mechanical methods that adequately describe how the electrons behave, commonly representing the active site of enzymes. Molecular mechanics is used as a cheap method to incorporate steric effects of thousands of atoms that surround the active site in the outer subsection. By using this method, proteins that are thousands of atoms in size can now be submitted to calculation.

### 2.6 Technical Aspects and Units

All calculations in this thesis were performed with the Gaussian 09 suites of programs,\textsuperscript{47} NAMD\textsuperscript{27} or within the Molecular Operating Environment (MOE) software.\textsuperscript{35} The relative energies are reported as kilojoules per moles (kJ mol$^{-1}$). The conversation factor from hartree (h) is

\[ 1 \, \text{h} = 2625.5 \, \text{kJ mol}^{-1} \]
2.7 References


Chapter 2: Theoretical Methods


Chapter 2: Theoretical Methods


35. *Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group Inc.: Montréal, QC, Canada, **2015**.


Chapter 3  A Density Functional Theory Assessment of 3– and 4– Coordinated Sulfur Species in Biological Systems
3.1 Introduction

The sulfur atom that is found in two of the amino acids, cysteine and methionine, predominately exists in a mono-coordinated (anionic cysteine) or di-coordinated (neutral cysteine, methionine or in disulfide bridges) state. However, a modest number of naturally occurring 3– and 4–coordinated sulfur intermediates have been revealed experimentally by spectroscopy or theoretically using computational methods.\(^1\text{-}\text{6}\) These include unique 4-coordinate organosulfuranes found in archaeal thioredoxin peroxidase (ApTPx)\(^1\) and methionine sulfoxide reductase (Msr).\(^2\) In addition, various 3-coordinate sulfur centers have been also observed, including sulfoxides (RS(O)R’),\(^2,\text{7}\) sulfonic acid (RSO\(_3\)H) in ubiquitin-dependent proteasomes,\(^5\) sulfinamides (RS(O)NR’) in myeloperoxidase,\(^6,\text{8}\) sulfenic acid phosphoryl ester in sulfiredoxins (Srx),\(^7\) and sulfilimines (RS=NR’) collagen IV,\(^9\) see Figure 3.1. Several studies have examined these structures experimentally to an extent and some insights have been determined towards their characterization.\(^1,\text{5}-\text{7},\text{9}\) Nevertheless, out of the aforementioned list of enzymes, only two have been examined computationally: ApTPx\(^1\) and Msr.\(^2,\text{10}\)

In ApTPx, a novel sulfurane was crystalized for the first time at a resolution of 1.77 Å and an electron density suggested a covalent bond between γ-S\(_{\text{Cys50}}\) and histidine (His), δ1-N\(_{\text{His42}}\).\(^1\) The distance between the two central atoms was found to be 2.21 Å. The authors performed quantum mechanical (QM) calculations to characterize the nature of this bond. Using small models at the B3LYP/6-31G* level of theory, they concluded the formation of a sulfurane with two possible protonation states for His42 (see Figure 3.1).\(^1\)
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

Figure 3.1. Illustrations of several 3– and 4–coordinated sulfur intermediates found in various proteins that were investigated in this study.

The second sulfurane intermediate found in Msr was characterized exclusively by computation. This intermediate was first shown in the subclass MsrB using a QM cluster approach. Various models with small to modest basis sets, 3-21G* and 6-31G(d) for the large and smaller model respectively, were employed using the B3LYP functional. They showed that the sulfurane was composed of a methionine group bound to a cysteiny1 S– and a hydroxyl group (see Figure 3.1). Similar results were obtained for the sister enzyme MsrA where a more extensive ONIOM (QM/MM) approach at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96)//ONIOM(B3LYP/6-31G(d):AMBER96) level of theory.

It is important to note all previous computational studies on these species have chosen the B3LYP functional at a modest basis set for their investigation. Although this
functional is extensively used and is generally reliable for a wide range of systems, to our knowledge, its behavior when studying these sulfur intermediates has never been explicitly assessed. Previously, Pearson et al. assessed a series of organoselenium, selenoxides and some divalent sulfur compounds using various density functional theory (DFT) methods.\textsuperscript{11-12} They showed that B3PW91 at 6-311(2df,p) was the best level of theory to use and smaller basis sets such as 6-31G(d,p) predicted geometries in almost equal accuracy as larger basis sets.\textsuperscript{11-12} Herein we used various common DFT functionals including the newer hybrid meta–GGA functionals with respect to 3– and 4–coordinated sulfur species found in biological systems. This will provide a reliable level of theory that can be used as a reference for any further computational studies to describe the diversity in sulfur chemical bonding.

3.2 Computational Methods

All calculations were performed with the Gaussian 09 suite of programs.\textsuperscript{13} A series of commonly used DFT functionals in studying enzymatic catalysis were considered. Specifically, the hybrid Becke’s three-parameter hybrid exchange functional\textsuperscript{14-15} with either the Lee–Yang-Parr\textsuperscript{16} (B3LYP) or the Perdew–Wang–1991\textsuperscript{17-18} (B3PW91) correlation functional were tested. In addition, the M06 suite of functionals meta-GGA M06L\textsuperscript{19} and the meta-hybrid-GGA functionals MO6\textsuperscript{20-21}, MO62X\textsuperscript{20} and MO6HF\textsuperscript{20} were also considered. These functionals were evaluated systematically using a wide range of Pople basis sets: 6-31G(d), 6-31G(d,p), 6-311G(d,p), 6-311+G(d,p), 6-311G(2d,p), 6-311G(df,p), 6-311+G(2df,p), 6-311G(d,p) to 6-311+G(3df,3pd). The employed basis sets represent a single change from a previous one, for instance, the addition of a d-function or an f-function to heavy atoms or diffuse functions. The exceptions are 6-311+G(2df,p) and 6-311+G(3df,3pd), which represents the cumulative expansion of all previously listed. All benchmark calculations were performed using high level \textit{ab initio}
methods, quadratic configuration interaction with singles and doubles excitations (QCISD) and the second–order Møller–Plesset perturbation (MP2) at the 6-311+G(2df,p) or 6-311++G(3df,3pd) level of theory as indicated. All calculations were performed in gas phase with a dielectric of 1 with the exception of the ethyl sulfinic phosphoryl ester found in Srx. In this case, the calculations were performed using a dielectric of 4 (IEFPCM) to increase the stability of the existing two negative charges in the absence of any stabilizing group. All bond lengths and obtained energetics are reported in Å and kJ mol⁻¹, respectively.

Prior to exploring the previous discussed 3– and 4–coordinated sulfur species, we constructed two series of small molecules starting from sulfenic acid and sulfoxide tautomeric structures (Figure 3.2) to elucidate the effect of chosen level of theory with respect to geometries and energetics. In the former series, sulfenic acid derivatives were increased in size and conjugation, Figure 3.2a. The latter series was mainly used to allow us to assess the effect of changing the level of theory on relative energies as well as introducing 3–coordinated sulfur species, Figure 3.2b.

Based on the results of the two aforementioned series, eleven 3– and 4–coordinated sulfur based intermediates were modeled from their native structures in proteins including truncated models of the involved amino acids, Figure 3.3. For example, an imidazole ring was used represent the histidyl residues in ApTPx. Additionally, cysteine and methionine were truncated to include only ethylthiol and ethylmethylsulfide moiety, respectively. Finally, in collagen IV, lysine was represented by a methylamine moiety.
Chapter 3: Assessment of 3– and 4-coordinated Sulfur Species in Biological Systems

Figure 3.2. Schematic illustrations of the compounds used in the systematic study of a) sulfenic acid derivatives and b) sulfoxides derivatives.
Figure 3.3 Schematic illustrations of the eleven modeled 3– and 4–coordinated sulfur compounds found in 6a,b) ApTPx, 7a–c) Srx, 8) Ubiquitin–dependent proteasomes, 9a–b) Msr 10) Collagen IV and 11) Myeloperoxidase.

3.3 Results and Discussion

One of the main challenges in any computational study is to know the limitations of the employed methods. The use of small models helps to overcome this obstacle by allowing for highly accurate computational methods to be used as a reference to assess
different DFT functionals. These are commonly applied to model large biological systems with lower computational costs. Therefore, the correlation methods, QCISD and MP2 calculations at 6-311+G(2df,p) or higher (as indicated) were used to reference energetics and geometric parameters.

In general, from all structures considered, the main differences in geometries were found to occur in bonds that have atoms that have equal or greater electronegativity than sulfur. Interestingly, the bond lengths in some of these cases deviated from the benchmark calculations up to and occasionally exceeding 0.1 Å. In contrast, sulfur–carbon and sulfur–hydrogen bonds deviated from the benchmark calculations minimally, rarely exceeding 0.01 Å. For this reason, only the former bond types will be discussed in the following sections.

3.3.1 Geometrical Assessment of Sulfenic Acid and Sulfoxide Derivatives

3.3.1.1 Series 1: Sulfenic Acid Derivatives

Sulfenic acid has been shown to be a common starting point for cysteine derived 3– and 4–coordinated structures found in biological systems.\textsuperscript{1,22} In this series, five different derivatives of sulfenic acid were assessed (Figure 3.2a). Notably, irrespective of the structure used, we noticed a specific ordering of the employed functionals with respect to the S–O distance. This order did not change upon the increase of the basis set. Therefore in the following section we discuss the ethyl sulfenic acid as it is represents the entire series well and is most similar to cysteine sulfenic acid (Figure 3.4).

In general, the S–OH bond is described well independently of the choice of the functionals considered. Indeed, the maximum difference in S–OH distance between these functions is less than 0.04 Å, Figure 3.4. Although this may apply to small models, in proteins this may not the case and the differences are expected to be larger due to the surrounding interactions. Comparing our two–benchmark methodologies, QCISD and
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

MP2 at the 6-311++G(3df,3pd), the difference between them is minimal at 0.005 Å. With respect to the benchmark, our calculations show that in general, B3LYP always overestimates the S–OH bond distance. Unlike B3LYP, other functionals were found to approach the benchmark as we increase the basis set. In particular, B3PW91 and M06L were found to overestimate the distance at smaller basis set, however at 6-311+G(2df,p) and 6-311++G(3df,3pd) give a similar results to the benchmark. Interestingly, M06, M062X and M06HF approach the benchmark at a moderate basis set, 6-311G(2d,p) and 6-311G(df,p), At smaller and larger basis sets, they were found to overestimates or underestimate the bond distance respectively. It is important to mention that in small models the application of a high basis set is feasible whereas in larger models the computational cost limits the size of the employed basis set to a moderate size such as: 6-31G(d) or 6-31G(d,p). Therefore, the application of DFT in computational enzymology depends on finding a exchange correlation functional that behaves close to the benchmark at moderate basis set.

These results highlight the applicability of these DFT functionals in small as well as larger protein models. For instance the B3PW91 and M06L are best in comparison to the benchmark at a high basis set. However, the results show that the hybrid meta-GGA Minnesota functionals: M06, M062X, M06HF would be better suited for modeling at a smaller basis set, one that is appropriate for studying protein system. Specifically, the difference from 6-31G(d,p) and the benchmark to these functionals is minimal.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

3.3.1.2 Series 2: Sulfoxide Derivatives

In general, methionine sulfoxide commonly produced in proteins as a result of methionine’s reaction with reactive oxygen species and is involved in protein regulation.23 Here, as indicated in the methodology, for general assessment, a second series of a simple sulfoxide (ethyl sulfoxide) derivative was investigated. As seen in Figure 3.5, unlike sulfenic acid, there are no distinct groupings of functionals, GGA, meta-GGA and hybrid meta-GGA, with respect to their performance. Similar to sulfenic acid, the S=O difference was used as the main criteria to assess the functionals geometrical parameters. The maximum difference in distance from the benchmark is 0.033 Å. As in previous the case, both QCISD and MP2 benchmark calculations are very similar, 0.003 Å. Interestingly, B3LYP and M06HF are now almost identical,

**Figure 3.4.** Plot of the optimized r(S–OH) bond distances in Å for ethyl sulfenic acid (EtSOH) Figure 3.2, 3a. Color-coding Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.

![Graph showing bond distances for ethyl sulfenic acid](image)
overestimating the S=O length to the greatest extent. The M06 suite of functionals M06, M062X and M06L, performed best at moderate basis sets. By increasing the size of the basis set, more fluctuation occurs during the systematic improvement of the optimized geometries. For instance, adding an extra d-functions from 6-311G(d,p) to 6-311G(2d,p) showed larger relative improvements for all functionals than adding a diffuse or f-functions. This improvement has been suggested before because adding d-functions allows for a better description of the interaction between the S=O bond.²⁴

![Graph showing bond distances](image)

**Figure 3.5.** Plot of the optimized r(S–OH) bond distances in Å for ethyl sulfoxide (EtSO) Figure 3.2, 3b. Benchmark: Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.

Collectively, B3PW91, B3LYP and M06HF all behave well with respect to modeling small molecules at a very high basis set. However, at a smaller basis set, 6-311G(2d,p), M06L, M062X and B3PW91 would be suitable for small model. At a moderate basis set, 6-31G(d,p) and 6-311G(d,p), M06L, M06 and M062X are best feasible choice for modeling enzymes.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

In summary considering functional behaviors in both the sulfenic acid and sulfoxide series, modeling enzymatic catalysis of sulfur containing species would be best performed with meta-hybrid-GGA M06 and M062X at a moderate basis set of 6-31G(d,p) for geometrical optimization.

3.3.1.3 Determination of Relative Energy Trends Between Sulfur Tautomers

Since the previous section provided a reference for calculating geometries in sulfur containing species, our second goal is to determine how the relative energy would change with respect to the level of theory used. This carried out by calculating the differences in energies between the sulfenic acid derivatives and their corresponding sulfoxides. For the two previous discussed derivatives, ethyl sulfenic acid and ethyl sulfoxide (Figure 3.6,a), the benchmark results showed a relative energy difference of 32.9 and 41.7 kJ mol\(^{-1}\) for MP2 and QCISD, respectively. Unlike the geometrical parameters, the choice of DFT functionals and basis sets has considerable implications on energetics, as shown in Figure 3.6b.

Using a small basis set, except M06L, can lead to large over approximations in the obtained relative energies. Even the best performing functional, B3PW91, deviates from the QCISD benchmark by at least 42.4% (17.7 kJ mol\(^{-1}\)) and the worst, M06HF, differs by 175.1% (73.0 kJ mol\(^{-1}\)). Although M06L differs by from the QCISD benchmark by only 0.7% (0.3 kJ mol\(^{-1}\)), there is a general decline in accuracy upon increasing the size of basis set. This is contrasted by the remaining functionals that improve considerably upon increasing basis sets. Specifically, adding extra d-functions (2d) provides the greatest improvements and any basis set that does not include at least extra d-functions would not be recommended.
Notably, the percentage of Hartree-Fock (HF) appears to systematically affect the functional's performance. As in M06HF which contains 100% HF, energy values have the largest deviations whereas M06L with zero HF is the closest to the benchmark at moderate basis sets. In addition, B3LYP, B3PW91 and M06, all have between 20-27% HF, performs similarly and very well at a high basis set, 6-311+G(2df,p) with respect to the benchmark. Therefore, they are better suited for energies in enzymatic catalysis.

Figure 3.6. a) Schematic showing an example of the difference between one of the five tautomers used to compare relative energies. b) Plot of the relative energy values after optimization for the tautomers of ethyl sulfenic acid and ethyl sulfoxide. Color-coding-Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.


3.3.2 Applications Into Biologically Relevant Intermediates

It must be noted the results of the simple di and tri-coordinated sulfur are proposed precursors but may not translate to the larger and more complex systems situated in proteins where the bond types may vary greatly. Therefore, herein, all suggested sulfur species have been examined, highlighting the new types of bonds when appropriate.

3.3.2.1 Archaeal Thioredoxin Peroxidase (ApTPx)

The first biological system subject to analysis is from archaeal thioredoxin peroxidase (ApTPx). This enzyme is mainly responsible for reducing oxidative stress via the reduction of hydrogen peroxide molecules.\(^1\) Here, a unique covalent bond between the \(\gamma_{\text{Cys50}}\) and the neighboring histidine (His), \(\delta_{1-\text{N_{His42}}}\) was proposed by Nakamura et al.\(^1\) according to the obtained electron density from the crystal structure. Based on B3LYP/6-31G*, the authors suggested two possible “hypervalent” sulfuranes to be a part of the enzyme's mechanism (Figure 3.3, 6 a,b). The sulfur has a hydrogen atom bound to its center and they differ only upon the protonation state of His42.

3.3.2.1.1 The Sulfurane with a Neutral Histidine

Both QCISD and MP2 calculations gave similar results with respect to the distance between S–OH and S–N. In addition, the performances of variant DFT functionals are in agreement with the benchmarks. More specifically, for the S–OH bond, the order of the functionals is consistent with the S–OH in series 1 (Figure 3.4 and 3.7a); B3LYP, B3PW91 and M06L resulted in longer distances with respect to the hybrid meta-GGA M06 functionals. However in this case, increasing the size of the basis set does not systemically improve the obtained distances. Notably, smaller basis set performs at the same accuracy as the largest basis sets. Increasing the valence description, using triple \(\zeta\), and adding diffuse functions had the largest negative effect on the bond description by about 0.02 Å each. A better performance could be achieved by adding more polarized functions to the heavy atoms.
For the S–N interaction, similar to the S–OH bond, the variations were minute with respect to the employed functionals (Figure 3.7b). In general, with the exception of M06HF, there is a positive correlation between size of basis sets and overall performance. B3LYP and M06L predict the longest distance. M06 and M062X are the closest to the benchmark. M06HF performs well at small basis set however at higher basis sets it grossly underestimates the bond distance and would not be recommended.

3.3.2.1.2 The Sulfurane with a Protonated Histidine

Here, the MP2 calculations do not match the QCISD benchmark (Figure 3.7 c,d). Specifically, they deviate by 1.4% (0.023) and 6.4% (0.143 Å) with S–OH and S–N bonds, respectively. Thus, a third ab initio method CCSD/6-311G(d,p) was tested as a reference, showing an agreement with QCISD results. Therefore, in the following discussion we used the results of QCISD as the reference. In contrast to the neutral sulfurane, the overall increase of the basis set improved the S–OH bond length by about 0.02 Å in total towards the QCISD benchmark (Figure 3.7b). The ordering of the functionals for S–OH is similar to the previous case. Again, the best performing functionals are the hybrid meta-GGA M06 functionals. For S–N interaction, significant differences were found to occur with respect to the employed functionals. However, the M06 functional was found to be agreement with the QCISD benchmark. The B3LYP and M06L slightly overestimate the bond distance and in contrast, the M062X and B3PW91 slightly underestimate the bond distance. Notably, increasing the size of the used basis set has a minimal effect on the obtained results. In M06HF, the bond is grossly underestimated and closely resembles the MP2 results.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

Figure 3.7. Plots for optimized bond lengths of a) S–OH, b) S–N of un-protonated for hydroxy-imidazolyl-λ4-sulfanylmethane and c) S–OH and d) S–N of protonated for hydroxy-imidazolyl-λ4-sulfanylmethane from Figure 3.3, 6a and b. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

Overall, with the two different bond types considered the best preforming functional is either M06 or M062X, as it consistently lies relatively close or closest to the benchmark calculations. In this case when the imidazole is protonated the functionals shows more discrepancy when describing the structure, particularly in the S–N interaction. Small basis sets such as 6-31G(d) and 6-31G(d,p) perform nearly as well as the extremely large, 6-311++G(3df.3pd). This is a desirable feature when using large models. Unfortunately, these results cannot be directly compared to the crystal structure S–N bond length of 2.21 +/- 0.13 Å due to the fact that the distances would likely differ based on the surrounding environment.

3.3.2.2 Sulfiredoxins (SrX)

This enzyme is of particular significance because the substrate and two of its intermediates are all 3–coordinated. SrX catalyzes the reduction of cysteine sulfenic acid (CysSO₂H) to sulfenic acid. The two proposed intermediates along the reaction coordinate are sulfenic phosphoryl ester and thiosulfinate (Figure 3.1). Both of these incorporate new bond types not explored thus far, these include: S–OPO₃⁻², SO–PO₃⁻² and S–S(O).

3.3.2.2.1 Sulfinic acid

This molecule is generally formed as a result of the over-oxidation of the cysteiny1 residue(s) in proteins. This molecule contains two bond types that have already been evaluated in the general assessment, S=O and S–OH. The only difference was that the sulfur oxidation state changed from 0 in sulfenic acid to +2 in sulfinic acid. The behaviors of these bond types are virtually identical to the trends outlined in the general assessment (refer back to Figure 3.4 and 3.5). This indicates that different sulfur redox states are described consistently between functionals for these small molecules.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

3.3.2.2 Sulfinic Phosphoryl Ester

For this intermediate, the S–OPO$_3^{-2}$ bond and the SO–PO$_3^{-2}$ bond are of interest (Figure 3.3, 7b). Unlike previous gas phase calculations, the sulfinic phosphoryl ester was optimized in a dielectric of 4 due to the presence of two negative charges. A dielectric of 4 was used because it represents a reasonable polarity of an active site environment.$^{25}$

First, the S–OPO$_3^{-2}$ bond was found to be described in a similar way to the sulfoxides of the S=O bond with respect to functional behavior as we increase the basis set (Figure 8a). The inclusion of more d-functions (2d) and removing the diffuse functions improves the bond description by 0.015 Å. Contrariwise, the oxygen/phosphate, SO–PO$_3^{-2}$ bond has a mirrored relationship of the sulfur/oxygen, S–OPO$_3^{-2}$, bond (Figure 3.8b) and notably, the inclusion of (2d) now worsen the results for the SO–PO$_3^{-2}$. Aside from that, increasing the size of the basis was found to overall improve the obtained geometries. As previous, the M06, M062X and M06HF performed best to the benchmark.

3.3.2.3 Thiosulfinate

This is similar to a disulfide bond with the difference of one the sulfurs is oxidized to a sulfoxide. (Figure 3.3, 7c). One sulfur has a +1 oxidation state next to a sulfur with a -1 oxidation state, the R–S(O)–SR’ functional group. Since S=O behaves similarly with all prior S=O examples, only the S–S bond will be discussed herein.

A larger variation was found to occur to with respect to the S–S bond, Figure 3.8c. The maximum deviation from the benchmark is found using B3LYP at all basis sets. The best performing functional at smaller basis sets are M062X.

Overall to generalize the choice of functional to use with the these types of intermediates in enzymes, it is found that the M062X functional consistently falls closest
or among the closest at a level of theory that is appropriate for studying proteins, 6-31G(d,p).

**Figure 3.8.** Plot of the optimized for a) the SO–PO\(^{-2}\) bond and b) S–OPO\(^{-2}\) bond distances in Å for ethyl sulfinic phosphoryl ester representing the first intermediate in Srx from Figure 3.3, 7b and c) the S–S bond distances of thiosulfinate representing the second intermediate in Srx Figure 3.3, 7c. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.
3.3.2.3 Ubiquitin-dependent Proteasomes

Ubiquitin-dependent proteasomes are responsible for the degradation of proteins in the cell by recognizing the irreversible over oxidation of cysteiny1 residue(s) into sulfonic acid.\textsuperscript{26,27} In sulfonic acid, the sulfur center has an oxidation state of +4 and found to behave similarly to its precursor sulfinic acid intermediate. The S–OH bond displayed a similar trends with respective to increasing basis set as the S=O bond, in series 2 (Figure 3.5). The basis set 6-311G(2d,p) performs best compared other than the two highest basis sets studied.

3.3.2.4 Methionine Sulfoxide Reductase (Msr A & B)

Msr is the first system under investigation whose precursor is methionine. Methionine in biological systems is readily oxidized to methionine sulfoxide and it is thought to be important in protein regulation as well as an oxidative defense method.\textsuperscript{2} The two enzymes MsrA and MsrB catalyze the reduction of S– and R–epimers of methionine sulfoxide respectively.\textsuperscript{2} For this enzyme, there are two molecules under investigation; a sulfurane and the other is a unique sulfonium cation intermediate that is formed along the proposed reaction pathway.\textsuperscript{2} Here, two bond types are of interest, first the S–S bond in both intermediates and second, the S–OH bond in the sulfurane. In general, the bond lengths in the sulfurane were found to be more dependent on the functional used than in the sulfonium cation and will be emphasized more extensively (Figure 3.9).

For the sulfurane, for S–S bond B3LYP as seen before, predicts lengths furthest from the benchmark (Figure 3.9a). Furthermore, there was a general systematic increase with performance with respect to increasing the basis set. The M06 and M062X functional are the closest at smaller basis sets. For the sulfonium cation, at a small basis set M062X, M06HF and M06L are very close to the QCISD benchmark (Figure 3.9b). Increasing the
valence description and adding diffuse functions without the incorporation of extra d or f-functions were found to worsen the results with respect to the smaller basis sets.

Figure 3.9. Plots of the optimized for S–S bond distances in Å for a) sulfurane, from Figure 3.3, 9b. b) sulfonium cation, from Figure 3.3, 9c from Msr A and B. c) S–OH bond distances in the sulfurane intermediate. Color-coding- Blue: B3LYP, Dark Red:
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems

B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6-311+G(2df,p) and MP2/6-311+G(2df,p) in black and red respectively.

Unlike the sulfurane, M06 was not amongst the top the performing functionals. As with the sulfurane S–S bond, the basis set can have a considerable impact on the S–OH bond lengths (Figure 3.9c). As before, it was beneficial to add diffuse functions however in this case, their addition elongates the S–OH by near 0.085 Å, or 4.4%. Although the smaller basis set describes the S–OH bond accurately, it does poorly for the S–S bond. Moreover, modest size basis set, 6-311G(d,p) to 6-311+G(d,p) describes the S–S bond well, they do not perform well for the S–OH. Thus, a larger basis set is required when doing characterization of this intermediate, such as 6-311G(2d,p). Unfortunately, today this size of basis set is not yet feasible to study enzymatic systems. Overall, the M062X functional collectively performs best with the two bond types and would be appropriate for studying enzymatic systems

3.3.2.5 Collagen IV

Recently a new type of bond was found for protein crosslinking in collagen introducing the R=S=NR. Collagen is required for the integrity of tissues in a wide variety of organisms. This structure has a double bond between the sulfur in methionine with the nitrogen of a nearby lysine residue. Consequently, this preforms comparably to the S=O bond of the sulfoxides. Interestingly, in this case all functionals studied with the exception of B3LYP, behave virtually identically, Figure 3.10 Thus solely out of looking at the S=N bond, out of the functionals studied, the choice is not as significant for this system. It is interesting that the all the DFT functionals do not approach the QCISD benchmark at the highest level of theory as it often does in the previous cases.
Figure 3.10. Plot of the optimized for S=N bond distances in Å of the sulfimine protein cross linker found in collagen IV. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311+G(2df,p) and MP2/6-311++G(3df,3pd) in black and red respectively.

3.3.2.6. Myelpersoxidase

Finally, when looking at myelpersoxides, there is a 3–coordinated sulfur with a S–N single bond, sulfinamide. Myelpersoxides are heme-enzymes that are released by pathogen combatting enzymes to form hypohalous acids (HOX, X is a halogen). The hypohalous acids can then produce sulfinamide or sulfonamide.

In this structure, we are only considering the behavior of functionals with respect to the S–N as the S–O bond is quite similar to previous cases, Figure 3.11. Both B3LYP and M06L are similar and the furthest from the benchmark. M06HF again found to deviate from the benchmark as we increase the basis set. Furthermore, M062X was found to be the best with a systematic increase in the performance with respect to basis set.
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems


![Diagram of bond distances](image.png)

**Figure 3.11.** Plot of the optimized for S–N bond distances in Å of the sulfinamide protein cross linker found in Myelopersoxidase. Benchmark QCISD 6-311++G(3df,3pd) from Figure 3.3, 11. Color-coding- Blue: B3LYP, Dark Red: B3PW91, Purple: M06HF, Green: M06, Orange: M062X and Teal: M06L. QCISD/6311++G(3df,3pd) and MP2/6-311++G(3df,3pd) in black and red respectively.

### 3.4 Conclusions

In this study the performance of several DFT functionals with different sized basis sets were assessed with respect to higher *ab initio* QCISD and MP2 benchmarks calculations dealing with several proposed 3– and 4–coordinated sulfur species in biochemistry. First, two series of molecules were evaluated including sulfinic acid/sulfoxide derivatives in which the geometric parameters and energetics were used as the criteria for the performance of the level of theory. Second, the previously purposed 3– and 4–coordinated species in ApTPx, Srx, Ubiquitin-dependent proteasomes, MsrA and B, Collagen IV and Myeloperoxidase were evaluated and found in some cases to be
highly dependent on the level of theory used. General remarks upon the performance of each functional are summarized below.

- **B3LYP**, in general was found to have the greatest overestimation of the investigated bond distances especially at small to moderate size basis sets in virtually all cases. In particular, very large basis sets are often required to approach to the benchmark calculations with respect to geometric parameters. However with respect to energetics, B3LYP performed on par with the other functionals. This highlights its limitation in describing 3– and 4– coordinated sulfur species and thus would not be recommended for geometric optimizations.

- **B3PW91** was found to perform slightly better than B3LYP. Furthermore, with respect to relative energies, it performs almost identically to B3LYP.

- **M06HF** performance varied depending on the system under study. For instance, in sulfenic acid it was found to the best functional to predict the S–OH bond at a small basis set. However, for the sulfoxide, it was found to be the worst functional to describe the S=O bond. In addition, the predicted relative energies were found to be the highest with a great deviation from the benchmark. Thus, although it behaves well with respective to certain species, it is often not recommend for studying these types of bonds.

- **M06L** was one of the best performing functionals with respect with S=O bond. However, for S–OH and S–S the results are moderate and similar to B3PW91. With respect to energy evaluation, M06L predicted values close to the benchmark at a smaller basis set and deviated away from the benchmark as the level of theory increased. In general its performance is moderate to poor with respect to geometric optimization.
M06 and M062X, both functionals were always found to be the best in describing the investigated molecules. In particular at smaller basis sets, applicable for studying large systems, they were found to be the closest and reliable to the ab initio benchmark calculations. Similarly for energetics, they are among the best functionals to be used. Thus both functionals are recommended for studying 3– and 4–coordinated sulfur systems.

In general our results show the percentage of HF in each functional appears to be correlated to its performance; this is especially evident when looking at energetics with 100% or 0 HF. The largest deviations occur with M06L and M06HF. In addition, B3LYP, B3PW91 and M06, all have between 20-27% HF, performs similarly and very well at a high basis set, 6-311+G(2df,p) with respect to the benchmark. Therefore, they are more suited for energies in enzymatic catalysis.

3.5 References
Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems


Chapter 3: Assessment of 3– and 4–coordinated Sulfur Species in Biological Systems


Chapter 4  Computational Insights into the Catalytic Mechanism of Homonuclear Glycerophosphodiesterase (GpdQ)

\[ \text{GpdQ} + \text{H}_2\text{O} \]
4.1 Introduction

Organophosphate (OP) poisoning is a global health concern that affects more than 3-million people in the developing world due to its high neurotoxicity.\textsuperscript{1} Exposure leads to several ailments such as seizures, paralysis, and cardiac conduction disorders that can ultimately lead to death.\textsuperscript{1-3} To cease the use of OPs is not likely due to several industrial and military applications.\textsuperscript{4-6} For example, they are one of the most commonly used types of pesticides marketed worldwide; present in about one-third of all pesticides.\textsuperscript{4,6} Unfortunately, OPs often unintentionally end up in water supplies and are difficult to remove.\textsuperscript{5} There are currently a few common treatment options to degrade phosphates into smaller and usually unharmed forms these include: chemical hydrolysis/oxidation by advanced oxidation processes (AOPs)\textsuperscript{7-8} or the use of UV radiation.\textsuperscript{9} However in large bodies of water, these methods can be expensive, non-specific and risk secondary contamination.\textsuperscript{10} An alternative strategy is bioremediation, which uses enzymes to breakdown these harmful agents. This method is thought to have large potential because the degradation process is catalytic and specific to OPs. Furthermore, they are considered “greener” because biocatalysts are readily biodegradable.\textsuperscript{11}

There have been several enzymes that have been proposed to be potential bioremediators for OPs and they belong to the family of binuclear metallohydrolases\textsuperscript{12-15} These include: OP hydrolase from Agrobacterium radiobacter (OpdA)\textsuperscript{12} and from Pseudomonas diminuta (OPH/PTE),\textsuperscript{13} purple acid phosphatase (PAP)\textsuperscript{14-15} from mammals and plants, and Glycerophosphodiesterase (GdpQ) from Enterobacter aerogenes.\textsuperscript{13, 16} Interestingly, despite having remarkably similar catalytic active sites, they have very different catalytic rates and degrees of promiscuity.\textsuperscript{17-20} OpdA/PTE are the most effective with a natural turnover number (k\textsubscript{cat}) close to 10\textsuperscript{4} s\textsuperscript{-1} and an catalytic efficiency (k\textsubscript{cat}/K\textsubscript{m}) value of 10\textsuperscript{8} M\textsuperscript{-1} s\textsuperscript{-1}.\textsuperscript{17-18} PAP have a k\textsubscript{cat} of about 400 s\textsuperscript{-1} with a k\textsubscript{cat}/K\textsubscript{m} value of 10\textsuperscript{3} M\textsuperscript{-1} s\textsuperscript{-1},\textsuperscript{20} whereas GdpQ is the least efficient with both catalytic turnover and k\textsubscript{cat}/K\textsubscript{m} of about
In terms of promiscuity, however, GpdQ is the only one that is readily reactive against all three types of phosphoesters (mono, di and tri) whereas the others tend to have only monoesterase with slight diester activity (PAP) or triesterase activity (OpdA).\textsuperscript{20-21} This feature has promoted GpdQ as a high potential candidate for bioremediation especially if the catalytic activity is enhanced. There has been an effort to achieve this; a 500-fold increase in catalytic activity was achieved by altering only its oligiometric structure.\textsuperscript{22}

![Figure 4.1](image-url)

**Figure 4.1.** Comparing the similar active site of three different classes. A) OdpA/PTE\textsuperscript{12-13}, b) PAP\textsuperscript{14-15} and c) GpdQ\textsuperscript{13,16} with their presumed native metal composition.

GpdQ’s reactivity can also be enhanced by altering the metal combination present in the active site.\textsuperscript{19} It has been suggested that the native composition is heteronuclear and most likely contains Fe(II) in the $\alpha$-site and the $\beta$-site is Zn(II).\textsuperscript{19,23} A study performed by Daumann et al. investigated the effect of various metal combination on the catalytic rates using Bis(p-nitrophenyl) phosphate (bpNPP).\textsuperscript{19} They determined an empirical reactivity series of Zn(II) < Co(II) < Mn(II) < Cd(II).\textsuperscript{19} Furthermore, both $k_{\text{cat}}$ and $k_{\text{cat}}/k_M$ are higher.
when the metals in the active site are the same, M(II)/M(II), opposed to the Fe(II)/M(II) combinations.\textsuperscript{19}

A single mechanism is not likely due to GpdQ’s activity in a large pH range, 3.0–11.0; an optimum pH that varies from 5.0 to 9.0 depending on the substrate used.\textsuperscript{16,21-23-24} Experimental studies performed by Hadler \textit{et al.} has proposed that there is terminal water/hydroxide located on the α-metal that acts as a nucleophile (Figure. 4.2).\textsuperscript{22} However, it should be noted that the pKa of the terminal water was found to be between 9 and 10.\textsuperscript{19} This mechanism is plausible at alkaline pHs but may not be the predominate route at a neutral or acidic pH. An alternative mechanism where the bridging µ–OH as the nucleophile has been proposed to occur in PAPs Figure 4.2 which has a similar active site composition.\textsuperscript{25} Therefore, it is possible that a single mechanism is not sufficient to explain the full range of activity associated with this enzyme.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure42.png}
\caption{The mechanism for the a) terminal hydroxide based hydrolysis of a general diphosphate by GpdQ\textsuperscript{21-22} vs. b) the bridging hydroxide mechanism proposed mechanism of PAP.\textsuperscript{25}}
\end{figure}
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

Recently, Gherib et al. begun to provide insights into the terminal water vs. μ–OH nucleophile hypothesis of dimethylphosphate (DMP) via computation.\textsuperscript{26} Using quantum mechanical cluster calculation B3LYP/6-31G(d) level of theory, they showed a large gain in entropy drove DMP to bind bidentate displacing the terminal water.\textsuperscript{26} The change in Gibbs free energy of the bidentate ligation process was -33.2 kJ mol\textsuperscript{-1} more favourable than when the terminal water was present.\textsuperscript{26} In addition, they suggested the critical role of an acidic His81 in the second coordination sphere to facilitate catalysis.

Herein we used a quantum mechanics/molecular mechanics (QM/MM) approach to continue the investigation the possibility of a μ–OH mechanism in GpdQ. Since the approximate homonuclear $k_{\text{cat}}$ are known for the phosphodiester bpNPP, an investigation its catalytic source was performed. In addition, a QM cluster approach investigated the roles of the residues in the α-site by \textit{in silico} mutagenesis on most efficient metal combination Cd(II)/Cd(II). The aim was to quantify the changes in structural parameters and the activation barriers to provide useful information for the rationale design of biomimics.

4.2 Computational Methods

4.2.1 Preparation of the QM/MM Model

The crystal structure PDB ID: 3D03\textsuperscript{27} was used as the starting point for the for the GdpQ model. This structure did not have a substrate bound in the active site therefore a molecular docking protocol was used within the molecular operating environment (MOE) software.\textsuperscript{28} The reason for choosing dimethylphosphate (DMP) as the substrate was three-fold, (1) the experimental $k_{\text{cat}}$ is for a phosphodiester (2) its small size is more cost effective for the quantity of QM/MM calculation performed and (3) a study for Co(II)-GpdQ have shown that rate limiting step for phosphodiester cleavage is dictated, at least partially, by cleavage.\textsuperscript{21}
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

The specific docking protocol used the induced fit formalism, specifically with the London dG scoring function and AMBER12EHT force field for minimization.\(^{28}\) This generated the Michaelis complex for which the quantum mechanics/molecular mechanism (QM/MM) model was built. The quantum mechanical (QM) layer of the QM/MM model consisted of the full substrate, metals, the first coordination sphere and two residues from the second coordination sphere, a methionine and histidine for a total of 101 atoms Figure 4.3. The molecular mechanics layer contained 933 atoms that surrounded the QM layer. The \(\alpha\)-carbons in the MM region were fixed outside the third coordination sphere to allow for sufficient flexibility in the active site and to keep the integrity of the overall enzyme structure.

![Figure 4.3](image.png)

**Figure 4.3.** The QM/MM model used in the study of GpdQ. The QM layer is shown in the foreground and the MM layer is faded in the background for clarity. Atom colours, Red=O, Blue=N, Orange=P, Dark Yellow=S, Grey=C, White=H, Pale Yellow=general metal.
4.2.2 QM/MM Calculations

All calculations were performed using the Gaussian 09 suite of programs. The geometry calculations were completed using several different functionals that have been used to describe metals in enzymes. In particular the hybrid functionals B3LYP, \( \text{B3LYP}^* \) (15% Hartree-Fock contribution) were used, as well as the meta-GGA functional M06L, and hybrid meta-GGA functionals M06 and MPWB1K. All optimizations used a 6-31G(d) basis set. Frequency calculations were performed on the reactant, transition state and product complexes to ensure that the reactants and products were at a minimum and that the transition state was at a maximum saddle point, (indicated by only positive and a single negative frequency value(s), respectively). The optimized structures were submitted to single points calculations using a much larger 6-311G(2d,2p) basis set. For the Mn(II) and Co(II) metals that can occupy several spin states, previous preliminary work in our group has shown that the lowest energy electronic configuration is when both metals occupy the high spin state. Finally, natural bond order (NBO) analysis was performed at the same level of theory as the optimizations to get the partial charges of each atom after truncating the QM/MM model to remove the MM layer and replacing the boundary with hydrogen caps.

4.2.3 Quantum Mechanical Cluster Calculation of Cd(II)/Cd(II) Complexes

Based off of the results from the first part of the study, the hybrid meta-GGA M06 functional for predicting activation barriers was used. The model was identical to the QM layer of the QM/MM model and the boundaries between the QM and MM were capped similarly to the NBO calculations. The \( \alpha \)-carbons were then fixed to maintain the active site integrity. In the absence of a MM layer, single points used IEFPCM solvation method with a dielectric constant of 4.0 to simulate the presence of a protein environment. In general, the native \( \alpha \)-site consists of Asp and His residues. A stepwise systematic change
of single, double and triple mutations from Asp to His and vice versa were performed to evaluate how each change can influence catalysis between each variant.

4.3 Results and Discussion

4.3.1 Reproducing the Catalytic Mechanism

Previously, our group has investigated the catalytic mechanism of the heteronuclear combinations of Fe(II)/M(II)-GpdQ, where M(II) can be Zn(II), Co(II), Mn(II) or Cd(II), using a quantum mechanical cluster approach.\textsuperscript{26} It was suggested that the mechanism was energetically favoured to undergo the $\mu$–OH nucleophilic mechanism with the requirement that His81 acts as a catalytic acid.\textsuperscript{26} Herein, we continued to explore this mechanism with homonuclear M(II)/M(II) combinations. It has been experimentally shown that the homonuclear derivatives have much higher catalytic turnover rates than the Fe(II)/M(II) counterpart and the reason behind its rate should be explored.\textsuperscript{19} Several DFT methods were used in an effort to remove any systematic anomalies that could arise from a single method interpretation. This approach was similar to what was done in the study of PAP\textsuperscript{25} and OpdA.\textsuperscript{35} Furthermore, this study used a quantum mechanical/molecular mechanics model. This may enhance the study because it has been shown that the third coordination sphere is important structural integrity of the enzyme and hydrogen-bonding networks may facilitate catalysis.\textsuperscript{19,27}

There are three possibilities for the cleavage of DMP using the $\mu$–OH nucleophile: (1) a single concomitant step where the nucleophilic attack of $\mu$–OH on the phosphate occurs as the OR leaving group is being protonated, (2) the $\mu$–OH attack occurs before the protonation of the leaving group forming a pentacoordinate intermediate and (3) the protonation of the leaving group occurs before $\mu$–OH attack. All three possibilities were attempted and the modeled catalytic mechanism was found to only occur in a single step
as shown in Figure 4.5. All attempts to form intermediates in the scenarios (2) and (3) resulted in collapses of the intermediate back to the RC or PC.

Figure 4.4 Pictorial representation of the reactant complex RC, transition state TS and product complex PC for DMP cleavage. Only the QM layer is shown with the catalytically relevant atoms enlarged. Atom colours, Red = O, Blue = N, Orange = P, Dark Yellow = S, Grey = C, White = H, Pale Yellow = general metal.

Notably, the previous QM cluster studies on GpdQ and PAP yielded a pentacoordinate intermediate, scenario (2).\textsuperscript{25-26} In both cases, the first step was rate limiting and the energy profiles showed the metastable intermediate almost equivalent to the barrier for the proceeding proton transfer.\textsuperscript{25-26} Notably, a one step mechanism is supported in the literature when the leaving group is stabilized.\textsuperscript{36-37} Therefore, we can hypothesize that the
inability to find a metastable pentacoordinate may be due to the QM/MM model holding His81 close to the leaving group, greatly stabilizing it. Repetition of the three remaining metals (Co(II)/Co(II), Mn(II)/Mn(II) and Cd(II)/Cd(II)) did not produce these intermediates therefore only the concerted mechanism was modeled throughout.

Daumann et al. has shown experimentally that when Zn(II)/Zn(II), Co(II)/Co(II), Mn(II)/Mn(II) and Cd(II)/Cd(II) are placed in the active site, the catalytic rates $k_{\text{cat}}$ for the phosphodiester bpNPP was 0.11, 1.62, 4.86 and 15.0 s$^{-1}$ respectively.$^{19}$ The calculated barriers for the modeled $\mu$–OH attack for each method and metal are shown in Table 4.1. It should be noted that the Mn(II)/Mn(II) transition state using M06L was located in a saddle point, as indicated by a frequency calculation, however it would not fully converge thus a reasonable approximation of the transition state is provided. The hybrid functional B3LYP and B3LYP* gave generally higher barriers than the meta-GGA functional M06L and hybrid-meta GGA functional MPW1BK. This is consistent with the previous study performed on PAPs.$^{25}$ Interestingly, the barriers closely or in some cases, exactly followed the ordering of the experimental trend. This occurred in B3LYP*, M06 and MPW1BK. Unfortunately the proportionality of the calculated barriers between each metal is difficult to quantitatively compare to the value of the $k_{\text{cat}}$ due to the inability to determine exactly how much of the enzyme contained homonuclear compositions vs. heteronuclear.$^{19}$ Therefore comment of the performance must be done with caution.

Ultimately, if modeling DMP is representative enough of bpNPP cleavage and it does undergo the proposed mechanism, we can suggest that either the hybrid meta-GGA M06 and MPW1BK functional or the hybrid B3LYP* functional produced barriers that are best in agreement with the experimental results.
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

Table 4.1. The transition state energies calculated for the hydrolysis of DMP for the four metals combinations using five different exchange-correlation functionals.

<table>
<thead>
<tr>
<th></th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zn(II)/Zn(II)</td>
<td>118.7</td>
<td>112.2</td>
<td>107.7</td>
<td>97.9</td>
<td>97.9</td>
</tr>
<tr>
<td>Co(II)/Co(II)</td>
<td>122.5</td>
<td>101.4</td>
<td>96.3</td>
<td>86.2</td>
<td>66.1</td>
</tr>
<tr>
<td>Mn(II)/Mn(II)</td>
<td>81.5</td>
<td>74.1</td>
<td>74.9</td>
<td>75.1*</td>
<td>48.5</td>
</tr>
<tr>
<td>Cd(II)/Cd(II)</td>
<td>69.0</td>
<td>64.4</td>
<td>64.7</td>
<td>71.3</td>
<td>30.6</td>
</tr>
</tbody>
</table>

*An approximate TS value

4.3.2 DMP Binding to Form the Reactive Complex

To begin the investigation of elucidating the source of the metal-dependent rate enhancement we examined the reactive complex (RC) structural parameters. The binding to the β metal varied significantly depending on the metal and functional used. For example, from Zn(II)/Zn(II), Co(II)/Co(II), Mn(II)/Mn(II) to Cd(II)/Cd(II) the $\text{phosO}\beta\cdots M\beta$ decreases from 4.16 Å, 4.09Å, 3.10 Å and 2.56 Å respectively with B3LYP, 3.48, 3.45, 3.09 and 2.56 with B3LYP*. However with the Truhlar functionals, the binding is more consistent regardless of the metal of interest with no obvious ordering; recording averages of 2.39 +/- 0.045 Å, 2.36 +/- 0.052 Å, 2.56 +/- 0.036 Å and 2.47 +/- 0.019 Å for Zn(II)/Zn(II), Co(II)/Co(II), Mn(II)/Mn(II) and Cd(II)/Cd(II) respectively (Table 4.2). Therefore for Zn(II)/Zn(II), B3LYP and M06, for example, have a 1.72 Å discrepancy in binding strength between them. A second observation is that the $r_{\text{phosO}\alpha\cdots M\alpha}$ interaction is generally stronger than the $\text{phosO}\beta\cdots M\beta$ interaction indicated by the shorter bond lengths with the exception of Cd(II)/Cd(II) whereas $\mu$–OH is bound consistently stronger to the Mα, that is, Mα$\cdots \mu$–OH is smaller in the case of Zn(II)/Zn(II), Co(II)/Co(II) and Mn(II)/Mn(II) and in the case of Cd(II)/Cd(II) it is equidistant between the two (Table 4.3).
Table 4.2 The geometric parameters for reactive complex (RC), transition state (TS) and product complex (PC) for the cleavage of DMP

<table>
<thead>
<tr>
<th></th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>phos{Oa}−Ma</td>
<td>2.38</td>
<td>2.09</td>
<td>2.01</td>
<td>2.78</td>
<td>2.09</td>
</tr>
<tr>
<td>phos{Ob}−Mb</td>
<td>4.16</td>
<td>2.22</td>
<td>2.08</td>
<td>3.48</td>
<td>2.22</td>
</tr>
<tr>
<td>µ-O(H)−Ma</td>
<td>2.00</td>
<td>2.42</td>
<td>3.75</td>
<td>2.07</td>
<td>2.41</td>
</tr>
<tr>
<td>µ-O(H)−Mb</td>
<td>2.01</td>
<td>2.32</td>
<td>3.44</td>
<td>2.01</td>
<td>2.31</td>
</tr>
<tr>
<td>His{82−H}−phos{OR}</td>
<td>1.90</td>
<td>1.32</td>
<td>0.98</td>
<td>1.78</td>
<td>1.33</td>
</tr>
<tr>
<td>phos{OR}−phos{P}</td>
<td>1.62</td>
<td>1.93</td>
<td>-</td>
<td>1.63</td>
<td>1.93</td>
</tr>
<tr>
<td>µ-O(H)−phos{P}</td>
<td>3.84</td>
<td>1.87</td>
<td>1.65</td>
<td>3.82</td>
<td>1.88</td>
</tr>
<tr>
<td>Co</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>phos{Oa}−Ma</td>
<td>2.38</td>
<td>2.06</td>
<td>1.99</td>
<td>2.16</td>
<td>2.06</td>
</tr>
<tr>
<td>phos{Ob}−Mb</td>
<td>4.09</td>
<td>2.18</td>
<td>2.05</td>
<td>3.45</td>
<td>2.18</td>
</tr>
<tr>
<td>µ-O(H)−Ma</td>
<td>1.99</td>
<td>2.51</td>
<td>3.86</td>
<td>2.09</td>
<td>2.55</td>
</tr>
<tr>
<td>µ-O(H)−Mb</td>
<td>1.97</td>
<td>2.23</td>
<td>3.55</td>
<td>1.96</td>
<td>2.22</td>
</tr>
<tr>
<td>His{82−H}−phos{OR}</td>
<td>1.86</td>
<td>1.31</td>
<td>0.98</td>
<td>1.77</td>
<td>1.33</td>
</tr>
<tr>
<td>phos{OR}−phos{P}</td>
<td>1.68</td>
<td>1.93</td>
<td>-</td>
<td>1.69</td>
<td>1.92</td>
</tr>
<tr>
<td>µ-O(H)−phos{P}</td>
<td>3.84</td>
<td>1.88</td>
<td>1.64</td>
<td>3.83</td>
<td>1.88</td>
</tr>
<tr>
<td>Mn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>phos{Oa}−Ma</td>
<td>2.37</td>
<td>2.15</td>
<td>2.04</td>
<td>2.37</td>
<td>2.15</td>
</tr>
<tr>
<td>phos{Ob}−Mb</td>
<td>3.10</td>
<td>2.27</td>
<td>2.09</td>
<td>3.09</td>
<td>2.27</td>
</tr>
<tr>
<td>µ-O(H)−Ma</td>
<td>2.15</td>
<td>2.40</td>
<td>3.97</td>
<td>2.16</td>
<td>2.40</td>
</tr>
<tr>
<td>µ-O(H)−Mb</td>
<td>2.03</td>
<td>2.33</td>
<td>3.74</td>
<td>2.03</td>
<td>2.33</td>
</tr>
</tbody>
</table>
## Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

<table>
<thead>
<tr>
<th></th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hist82-H\cdots,\text{phos,OR}</strong></td>
<td>1.75</td>
<td>1.30</td>
<td>0.98</td>
<td>1.74</td>
<td>1.30</td>
<td>0.98</td>
<td>1.80</td>
<td>1.32</td>
<td>0.97</td>
<td>1.77</td>
<td>1.31</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>phos,OR\cdots,\text{phos,P}</strong></td>
<td>1.67</td>
<td>1.94</td>
<td></td>
<td>1.68</td>
<td>1.94</td>
<td></td>
<td>1.67</td>
<td>1.91</td>
<td></td>
<td>1.67</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{phos,P}</strong></td>
<td>3.48</td>
<td>1.90</td>
<td>1.64</td>
<td>3.48</td>
<td>1.90</td>
<td>1.64</td>
<td>3.44</td>
<td>1.87</td>
<td>1.65</td>
<td>3.31</td>
<td>1.86</td>
<td>1.67</td>
</tr>
<tr>
<td><strong>Ca</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hist82-H\cdots,\text{phos,OR}</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>phos,OR\cdots,\text{phos,P}</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{phos,P}</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Cd

<table>
<thead>
<tr>
<th></th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>phos,Oa\cdots,\text{Ma}</strong></td>
<td>2.66</td>
<td>2.31</td>
<td>2.18</td>
<td>2.65</td>
<td>2.31</td>
<td>2.27</td>
<td>2.61</td>
<td>2.33</td>
<td>2.18</td>
<td>2.56</td>
<td>2.32</td>
<td>2.37</td>
</tr>
<tr>
<td><strong>phos,Ob\cdots,\text{Mb}</strong></td>
<td>2.56</td>
<td>2.38</td>
<td>2.22</td>
<td>2.56</td>
<td>2.38</td>
<td>2.30</td>
<td>2.46</td>
<td>2.31</td>
<td>2.21</td>
<td>2.47</td>
<td>2.32</td>
<td>2.28</td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{Ma}</strong></td>
<td>2.20</td>
<td>2.47</td>
<td>3.99</td>
<td>2.20</td>
<td>2.47</td>
<td>3.27</td>
<td>2.20</td>
<td>2.42</td>
<td>3.85</td>
<td>2.22</td>
<td>2.44</td>
<td>2.81</td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{Mb}</strong></td>
<td>2.19</td>
<td>2.43</td>
<td>3.90</td>
<td>2.19</td>
<td>2.42</td>
<td>3.08</td>
<td>2.20</td>
<td>2.42</td>
<td>3.66</td>
<td>2.22</td>
<td>2.44</td>
<td>2.87</td>
</tr>
<tr>
<td><strong>Hist82-H\cdots,\text{phos,OR}</strong></td>
<td>1.77</td>
<td>1.30</td>
<td>0.98</td>
<td>1.76</td>
<td>1.30</td>
<td>0.98</td>
<td>1.79</td>
<td>1.34</td>
<td>0.98</td>
<td>1.76</td>
<td>1.32</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>phos,OR\cdots,\text{phos,P}</strong></td>
<td>1.69</td>
<td>1.96</td>
<td></td>
<td>1.70</td>
<td>1.96</td>
<td></td>
<td>1.67</td>
<td>1.90</td>
<td></td>
<td>1.69</td>
<td>1.94</td>
<td></td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{phos,P}</strong></td>
<td>3.67</td>
<td>1.89</td>
<td>1.65</td>
<td>3.67</td>
<td>1.90</td>
<td>1.67</td>
<td>3.52</td>
<td>1.89</td>
<td>1.64</td>
<td>3.51</td>
<td>1.88</td>
<td>1.69</td>
</tr>
</tbody>
</table>

### Ca

<table>
<thead>
<tr>
<th></th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
<th>RC</th>
<th>TS</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>phos,Oa\cdots,\text{Ma}</strong></td>
<td>2.54</td>
<td>2.36</td>
<td>2.24</td>
<td>2.59</td>
<td>2.37</td>
<td>2.25</td>
<td>2.57</td>
<td>2.42</td>
<td>2.24</td>
<td>2.55</td>
<td>2.44</td>
<td>2.24</td>
</tr>
<tr>
<td><strong>phos,Ob\cdots,\text{Mb}</strong></td>
<td>2.58</td>
<td>2.44</td>
<td>2.31</td>
<td>2.54</td>
<td>2.44</td>
<td>2.31</td>
<td>2.49</td>
<td>2.36</td>
<td>2.31</td>
<td>2.48</td>
<td>2.37</td>
<td>2.30</td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{Ma}</strong></td>
<td>2.31</td>
<td>2.48</td>
<td>4.01</td>
<td>2.31</td>
<td>2.46</td>
<td>4.01</td>
<td>2.29</td>
<td>2.42</td>
<td>3.93</td>
<td>2.29</td>
<td>2.43</td>
<td>3.97</td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{Mb}</strong></td>
<td>2.32</td>
<td>2.46</td>
<td>4.06</td>
<td>2.32</td>
<td>2.48</td>
<td>4.07</td>
<td>2.31</td>
<td>2.51</td>
<td>3.99</td>
<td>2.31</td>
<td>2.51</td>
<td>4.03</td>
</tr>
<tr>
<td><strong>Hist82-H\cdots,\text{phos,OR}</strong></td>
<td>1.80</td>
<td>1.26</td>
<td>0.98</td>
<td>1.73</td>
<td>1.26</td>
<td>0.98</td>
<td>1.76</td>
<td>1.33</td>
<td>0.98</td>
<td>1.73</td>
<td>1.31</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>phos,OR\cdots,\text{phos,P}</strong></td>
<td>1.68</td>
<td>1.94</td>
<td></td>
<td>1.69</td>
<td>1.94</td>
<td></td>
<td>1.67</td>
<td>1.87</td>
<td></td>
<td>1.69</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td><strong>\mu,-O(H)\cdots,\text{phos,P}</strong></td>
<td>3.46</td>
<td>1.97</td>
<td>1.65</td>
<td>3.66</td>
<td>1.97</td>
<td>1.65</td>
<td>3.52</td>
<td>1.96</td>
<td>1.63</td>
<td>3.49</td>
<td>2.00</td>
<td>1.65</td>
</tr>
</tbody>
</table>
This highlights a consequence of single functional analysis. For example, by looking at just the B3LYP and B3LYP* functionals, \( \text{phos} \cdot \text{O} \cdots \text{M} \) appears to follow the energy trends. The \( \text{phos} \cdot \text{O} \cdots \text{M} \) length using B3LYP for Zn(II)/Zn(II), Co(II)/Co(II), Mn(II)/Mn(II) and Cd(II)/Cd(II) was 4.16 Å, 4.09 Å, 3.10 Å and 2.56 Å, and the associated energy barrier are 118.7, 122.5, 81.5, 69.0 kJ mol\(^{-1}\) and for B3LYP*, 3.48, 3.45, 3.09 and 2.56 Å yielded 112.2, 101.4, 74.1, 64.4, 54.5 kJ mol\(^{-1}\). This could lead to a hypothesis that by binding to two Lewis acids, the phosphate would be more polarized and thus more susceptible to nucleophilic attack by \( \mu-\text{OH} \). Furthermore, it could be stated that the substrate binding of Cd(II)/Cd(II) and Mn(II)/Mn(II) mimics more of the transition state structure than Zn(II)/Zn(II) and Co(II)/Co(II), reducing the energy gap between them. However, by including more functionals in the study such as M06, M06L and MPW1BK this hypothesis cannot explain the source of catalytic activity because these functionals predict the same relative energy trends in the absence of those structural features. This indicates that there are more factors that are contributing to catalysis.

### 4.3.3 DMP Transition State Structural Parameters

Overall, unlike the RC, the general structures of the TS appear to be more consistent across all functionals for each metal. In phosphate hydrolysis, in general, there are two definitions of \( S_N2(P) \) type reactions to describe the transition state: associative and dissociative.\(^{38-40}\) If the degree of bond cleavage \( \text{phos} \cdot \text{OR} \cdots \text{phos} \cdot \text{P} \) is greater than the degree of bond formation (\( P \cdots \mu-\text{OH} \)) then the resulting net loss in bond order means the transition state has dissociative character.\(^{38-39}\) In contrast, if the \( P \cdots \mu-\text{OH} \) comes together before cleavage then it is denoted an associative mechanism. It has been determined using the kinetic isotope effect that phosphodiester in solution utilize a concerted associative transition state.\(^{36}\) This has energetic implications because it is thought the dissociative is predicted to be up to 1000-fold faster than the associative pathway.\(^{38}\)
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

With this knowledge, we explored the possibility that the larger metals, Mn(II)/Mn(II) and Cd(II)/Cd(II) may be increasing the proton transfer component to enhance the rate of reaction. To test this hypothesis, the bond distances in the transition state particularly the distance of \( \text{phosOR}^{-}\ldots\text{phosP} \) and \( \text{His82}-\text{H}^{-}\ldots\text{phosOR} \) can be examined, Table 4.2. In theory, an increase in \( \text{phosOR}^{-}\ldots\text{phosP} \) would decrease the bond order and shift to be a slightly more dissociative. This is observed where \( \text{phosOR}^{-}\ldots\text{phosP} \) averages about 1.90 for Zn(II)/Zn(II) and Co(II)/Co(II), 1.92 Å for Mn(II)/Mn(II) and 1.94 for Cd(II)/Cd(II) in correlation with the rates/barriers. Consequently, the proton from His81 is consistently closer to the OR leaving group, by about 0.02 Å closer for Cd(II)/Mn(II) than it is for Zn(II)/Co(II) which stabilizes the leaving group Table 4.2.

Additional transition state parameters include the \( \text{phosO}^{-}\ldots\text{M}^{-}\beta \) length ranges from about 2.20 Å to 2.38 Å from Zn(II)/Zn(II) to Cd(II)/Cd(II). The \( \text{phosO}^{-}\ldots\text{M}^{-}\alpha \) interaction is stronger for each metal with Zn(II)/Zn(II) and Co(II)/Co(II) being 2.05–1.10 Å, Mn(II)/Mn(II) about 2.15 Å and Cd(II)/Cd(II) 2.31 Å. Similar to the RC, the \( \mu^{-}\text{OH} \) is shifted closer to the M\( \alpha \) than the M\( \beta \) for all metals. For Cd(II)/Cd(II) and Mn(II)/Mn(II) the hydroxyl more centered than for Zn(II)/Zn(II) and Co(II)/Co(II), Table 4.2

4.3.4 Metal Induced Phosphate Strain

Gherib et al. suggested that phosphate strain was a possible reason for the different rates among metals, specifically for Cd(II)/Cd(II). A size-dependent theory is possible because the ionic size does generally follow the trend of the experimental observed rates, albeit not proportional. For instance, the ionic radii of M\(^{2+}\) in an octahedral complex are 74, 74.5, 83, 95 pm for Zn(II), Co(II), Mn(II), and Cd(II) respectively.\(^{41}\) Looking at the M\( \alpha \ldots\text{M}^{-}\beta \) separation, it is clear that due to their size Zn(II)/Zn(II) and Co(II)/Co(II) are closer together averaging 3.07 +/- 0.02 Å and 3.09 +/- 0.02 Å respectively, Mn(II)/Mn(II) and Cd(II)/Cd(II) are further apart at 3.15 +/- 0.01 Å and 3.28 +/- 0.03 Å respectively (Table 4.3). Surprisingly, Cd(II)/Cd(II) being furthest apart did not lead to increase in
\( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \), in fact for the RC, the \( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \) angle is the smallest, averaging 116.2° and there is no distinct trend between the \( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \) and M\( \alpha \cdots \) M\( \beta \). To compare the natural structural parameters of DMP in solution, absent of the active site, DMP was optimized using an IEFPCM solvation representative of water at the M06/6-31G(d) level of theory. This resulted in a much larger \( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \) of 125.9°. For a direct comparison, looking at the M06 calculations in the active site was 116.2°, a difference of 9.5° (Table 4.3). A single point calculation of both phosphates showed that the enzyme destabilizes the substrate by 39.9 kJ mol\(^{-1}\), illustrating reactant destabilization by angle compression. For the remaining metals, Zn(II)/Zn(II), Co(II)/Co(II) and Mn(II)/Mn(II), averages are similar of 116.6°, 116.8° and 117.6° making a difference of 9.1°, 8.9° and 8.1° respectively. This provides an explanation of how enzymes can significantly enhance reaction rates over un-catalyzed reactions but cannot be the source of the fine-tuning.

The transition state bond angles averages may provide a portion of the fine-tuning. During the transition state, the \( \text{phos}O\alpha\cdots\text{M}\alpha \) and \( \text{phos}O\beta\cdots\text{M}\beta \) bond lengths decrease significantly along with the large increase in the M\( \alpha \cdots \) M\( \beta \) distance, Table 4.3. At this point, the phosphate/metal interaction appears to be strong enough systematically influence \( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \). The relative energies between transition states were approximated by single point calculations with \( \epsilon=4 \). Referenced against Zn(II)/Zn(II)'s TS, the relative energy of Mn(II)/Mn(II), Co(II)/Co(II) and Cd(II)/Cd(II) are more stable by 3.1, 3.0 and 11.0 kJ mol\(^{-1}\) respectively. This correlates with the \( \angle(\text{phos}O\alpha-P\text{phos}O\beta) \) for Cd(II)/Cd(II), averaging 122.6° +/- 0.7° and Mn(II)/Mn(II), Co(II)/Co(II) and Zn(II)/Zn(II) from largest to smallest with 121.5 +/- 0.4°, 121.1 +/- 0.4° and 120.9 +/- 0.3° respectively (Table 4.3).
Table 4.3. The metal separation in the active site of GpdQ in Angstroms (Å) and phosphate angle in degrees (°).

<table>
<thead>
<tr>
<th></th>
<th>Zn(II)</th>
<th></th>
<th></th>
<th>Co(II)</th>
<th></th>
<th></th>
<th>Mn(II)</th>
<th></th>
<th></th>
<th>Cd(II)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
<td>PC</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>B3LYP</td>
<td>3.04</td>
<td>3.50</td>
<td>3.87</td>
<td>117.7</td>
<td>121.2</td>
<td>115.9</td>
<td>3.06</td>
<td>3.46</td>
<td>3.76</td>
<td>117.8</td>
<td>121.8</td>
</tr>
<tr>
<td>B3LYP*</td>
<td>3.08</td>
<td>3.49</td>
<td>3.87</td>
<td>117.7</td>
<td>121.2</td>
<td>116.0</td>
<td>3.09</td>
<td>3.49</td>
<td>3.75</td>
<td>118.0</td>
<td>120.9</td>
</tr>
<tr>
<td>M06</td>
<td>3.07</td>
<td>3.45</td>
<td>3.85</td>
<td>116.0</td>
<td>120.6</td>
<td>117.0</td>
<td>3.09</td>
<td>3.42</td>
<td>3.72</td>
<td>115.9</td>
<td>121.1</td>
</tr>
<tr>
<td>M06L</td>
<td>3.08</td>
<td>3.45</td>
<td>3.85</td>
<td>116.2</td>
<td>121.2</td>
<td>118.9</td>
<td>3.12</td>
<td>3.43</td>
<td>3.75</td>
<td>116.5</td>
<td>121.0</td>
</tr>
<tr>
<td>MPW1BK</td>
<td>3.07</td>
<td>3.49</td>
<td>3.86</td>
<td>115.6</td>
<td>120.5</td>
<td>117.3</td>
<td>3.10</td>
<td>3.44</td>
<td>3.77</td>
<td>115.7</td>
<td>120.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3.5 Natural Bond Order (NBO) Analysis

NBO calculations have been shown to be effective at estimating rate enhancement in phosphate cleavage reactions\(^{39,42}\). This can help to explain if the particular metals polarizes the substrate to facilitates catalyst. The natural partial charges of oxygen of the nucleophile O(H), phosphorus P, leaving group OR two binding oxygens \(\text{phosO}_\alpha\), \(\text{phosO}_\beta\) and both metal \(M\alpha\) and \(M\beta\) are shown in Table 4.4. These results were roughly compared
to a similar concerted mechanism of the cleavage of the RNA analogue (HpPNP) who compared Zn(II)/Zn(II) mediated cleavage vs. uncatalyzed reaction. They showed the NBO charge of the phosphate in the TS increased from +2.48 to +2.61, from uncatalyzed to catalyzed coupled with the build up of charge of both the nucleophile and the leaving group. If the catalyzed reaction is many folds faster than the uncatalyzed, then changes in these parameters may enhance the reaction.

In agreement with their Zn(II)/(II) assisted mechanism, the natural charges of our phosphate center are around +2.60. The natural charges of the metals increases in positive charge where Co(II) < Mn(II) < Zn(II) ≈ Cd(II), Table 4.4. This has large impact on the μ-OH in the RC and TS. The natural charge of μ–OH averages in the RC -1.08, -1.11, -1.25 and 1.23 Co(II)/Co(II), Mn(II)/Mn(II), Zn(II)/Zn(II), Cd(II)/Cd(II) respectively. Furthermore, in the TS they are -1.02, -1.04, -1.08 and -1.09 of Co(II)/Co(II), Mn(II)/Mn(II), Zn(II)/Zn(II), Cd(II)/Cd(II) respectively. Here there is a relationship between the metals natural charge and build up of charge on the μ–OH. The positive charge is not greatly affected, indicated by a decrease of 0.01 following that same series. Surprisingly, we do not see a build up of negative charge on the leaving group from the increase μ–OH but we speculate that this may due to the proton neutralizing this charge where in their study a proton source was not present.
Table 4.4. NBO calculated natural charges for both reactant complex (RC) and transition state structures (TS).

<table>
<thead>
<tr>
<th></th>
<th>Zn$^{2+}$</th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>O(H)</td>
<td>-1.25</td>
<td>-1.07</td>
<td>-1.24</td>
<td>-1.07</td>
<td>-1.25</td>
<td>-1.09</td>
</tr>
<tr>
<td>P</td>
<td>2.60</td>
<td>2.59</td>
<td>2.57</td>
<td>2.56</td>
<td>2.65</td>
<td>2.65</td>
</tr>
<tr>
<td>OR</td>
<td>-0.88</td>
<td>-0.84</td>
<td>-0.87</td>
<td>-0.83</td>
<td>-0.89</td>
<td>-0.83</td>
</tr>
<tr>
<td>O$\alpha$</td>
<td>-1.16</td>
<td>-1.19</td>
<td>-1.14</td>
<td>-1.17</td>
<td>-1.16</td>
<td>-1.21</td>
</tr>
<tr>
<td>O$\beta$</td>
<td>-1.14</td>
<td>-1.20</td>
<td>-1.12</td>
<td>-1.19</td>
<td>-1.15</td>
<td>-1.22</td>
</tr>
<tr>
<td>M$\alpha$</td>
<td>1.40</td>
<td>1.40</td>
<td>1.38</td>
<td>1.39</td>
<td>1.39</td>
<td>1.40</td>
</tr>
<tr>
<td>M$\beta$</td>
<td>1.40</td>
<td>1.39</td>
<td>1.38</td>
<td>1.37</td>
<td>1.38</td>
<td>1.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Co$^{2+}$</th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>O(H)</td>
<td>-1.05</td>
<td>-1.01</td>
<td>-1.04</td>
<td>-1.00</td>
<td>-1.08</td>
<td>-1.03</td>
</tr>
<tr>
<td>P</td>
<td>2.57</td>
<td>2.60</td>
<td>2.57</td>
<td>2.57</td>
<td>2.66</td>
<td>2.65</td>
</tr>
<tr>
<td>OR</td>
<td>-0.88</td>
<td>-0.83</td>
<td>-0.87</td>
<td>-0.83</td>
<td>-0.89</td>
<td>-0.86</td>
</tr>
<tr>
<td>O$\alpha$</td>
<td>-1.11</td>
<td>-1.09</td>
<td>-1.10</td>
<td>-1.11</td>
<td>-1.12</td>
<td>-1.15</td>
</tr>
<tr>
<td>O$\beta$</td>
<td>-1.15</td>
<td>-1.13</td>
<td>-1.12</td>
<td>-1.14</td>
<td>-1.14</td>
<td>-1.18</td>
</tr>
<tr>
<td>M$\alpha$</td>
<td>1.14</td>
<td>1.12</td>
<td>1.07</td>
<td>1.05</td>
<td>1.08</td>
<td>1.09</td>
</tr>
<tr>
<td>M$\beta$</td>
<td>1.10</td>
<td>1.10</td>
<td>1.06</td>
<td>1.05</td>
<td>1.08</td>
<td>1.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mn$^{2+}$</th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>O(H)</td>
<td>-1.11</td>
<td>-1.04</td>
<td>-1.09</td>
<td>-1.03</td>
<td>-1.10</td>
<td>-1.05</td>
</tr>
<tr>
<td>P</td>
<td>2.58</td>
<td>2.59</td>
<td>2.55</td>
<td>2.56</td>
<td>2.63</td>
<td>2.65</td>
</tr>
<tr>
<td>OR</td>
<td>-0.89</td>
<td>-0.84</td>
<td>-0.88</td>
<td>-0.83</td>
<td>-0.89</td>
<td>-0.86</td>
</tr>
<tr>
<td>O$\alpha$</td>
<td>-1.13</td>
<td>-1.14</td>
<td>-1.11</td>
<td>-1.13</td>
<td>-1.14</td>
<td>-1.17</td>
</tr>
<tr>
<td>O$\beta$</td>
<td>-1.12</td>
<td>-1.17</td>
<td>-1.11</td>
<td>-1.15</td>
<td>-1.14</td>
<td>-1.18</td>
</tr>
<tr>
<td>M$\alpha$</td>
<td>1.21</td>
<td>1.20</td>
<td>1.17</td>
<td>1.16</td>
<td>1.17</td>
<td>1.16</td>
</tr>
<tr>
<td>M$\beta$</td>
<td>1.20</td>
<td>1.21</td>
<td>1.16</td>
<td>1.18</td>
<td>1.17</td>
<td>1.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Cd$^{2+}$</th>
<th>B3LYP</th>
<th>B3LYP*</th>
<th>M06</th>
<th>M06L</th>
<th>MPW1BK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
<td>RC</td>
<td>TS</td>
</tr>
<tr>
<td>O(H)</td>
<td>-1.23</td>
<td>-1.09</td>
<td>-1.22</td>
<td>-1.08</td>
<td>-1.24</td>
<td>-1.10</td>
</tr>
<tr>
<td>P</td>
<td>2.59</td>
<td>2.58</td>
<td>2.57</td>
<td>2.55</td>
<td>2.64</td>
<td>2.64</td>
</tr>
<tr>
<td>OR</td>
<td>-0.88</td>
<td>-0.84</td>
<td>-0.87</td>
<td>-0.83</td>
<td>-0.89</td>
<td>-0.86</td>
</tr>
<tr>
<td>O$\alpha$</td>
<td>-1.13</td>
<td>-1.18</td>
<td>-1.12</td>
<td>-1.17</td>
<td>-1.15</td>
<td>-1.20</td>
</tr>
<tr>
<td>O$\beta$</td>
<td>-1.15</td>
<td>-1.19</td>
<td>-1.13</td>
<td>-1.18</td>
<td>-1.16</td>
<td>-1.21</td>
</tr>
<tr>
<td>M$\alpha$</td>
<td>1.41</td>
<td>1.44</td>
<td>1.39</td>
<td>1.42</td>
<td>1.41</td>
<td>1.44</td>
</tr>
<tr>
<td>M$\beta$</td>
<td>1.41</td>
<td>1.43</td>
<td>1.39</td>
<td>1.41</td>
<td>1.41</td>
<td>1.44</td>
</tr>
</tbody>
</table>
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

To summarize, the characterization of mechanistic and structural parameters have led to several possibilities to enhance catalysis with the different metals combinations.

- Cd(II)/Mn(II) shift the underlying mechanism more towards the dissociative mechanism over the associative than Zn(II)/Co(II) illustrated by the increase in $\text{phosOR}^{\cdots}\text{phosP}$ and decrease in $\text{His}_{82}\cdots\text{H}^{\cdots}\text{phosOR}$ stabilizing the leaving group.
- Reactant destabilization by phosphate angle strain activates the phosphate for enzymatic catalysis. This would reduce the barrier of reaction for Cd(II)/Cd(II) over the uncatalyzed reaction by 39.9 kJ mol$^{-1}$.
- Transition state stabilization via phosphate angles contributes a small degree to fine-tuning up to 11 kJ mol$^{-1}$.
- The natural charges of the metals strongly influence the negative charge on the $\mu$–OH enhancing its nucleophilic ability.

4.3.6 Substituting Calcium Ions for $\text{M}_\alpha$ and $\text{M}_\beta$

To test our theory that was built from examining the first four metals, Ca(II)/Ca(II) homonuclear compositions were used to examine any similarities and differences. Octahedral Ca(II) has an ionic radii of about 100 pm making it marginally larger than Cd(II) ~5 pm.$^{41}$ The structural parameters of interests are shown in Table 4.5. The larger ionic radii continues to widen the active site indicated by an $\text{M}_\alpha$-$\text{M}_\beta$ in the RC of about 3.40 Å (up from about 3.30 Å in Cd(II)). For the first four metals, the increase in $\text{M}_\alpha$-$\text{M}_\beta$ correlated with the decrease in $\angle_{\text{phosOC}–\text{P}–\text{phosO}}$. This was observed in Ca(II)/Ca(II) as well with an $\angle_{\text{phosOC}–\text{P}–\text{phosO}}$ of 115.3$^\circ$. This difference was quantified by preforming single point calculations on DMP and compared the results with the DMP of Cd(II)/Cd(II). The calculated relative energy difference was +12.1 kJ mol$^{-1}$ higher for Ca(II)/Ca(II) than Cd(II)/Cd(II). Remarkably, the phosphate strain is becoming large enough that a small change of about one degree causes 10+ kJ mol$^{-1}$ destabilization of the
reactant. The calculated barrier between Cd(II)/Cd(II) and Ca(II)/Ca(II) for M06 was about 6 kJ mol\(^{-1}\). The transition state structure stabilization was also quantified, the difference was minimal; 1.35 kJ mol\(^{-1}\) less stable than Cd(II)/Cd(II). Therefore, it is likely that reactant destabilization is accounting for the majority of the enhanced rate predicted by DFT calculation.

**Table 4.5.** Select parameters of the Ca(II)/Ca(II) derivative.

<table>
<thead>
<tr>
<th>Ca(II)</th>
<th>Barrier kJ mol(^{-1})</th>
<th>(r(M\alpha-M\beta))</th>
<th>(\angle_{\text{phos O} \alpha - \text{p} - \text{phos O} \beta})</th>
<th>(\text{His}_{82}\text{H} - \text{phos OR} \cdots \text{phos OR} \cdots \text{P})</th>
</tr>
</thead>
<tbody>
<tr>
<td>B3LYP</td>
<td>59.7</td>
<td>3.43</td>
<td>3.70</td>
<td>115.1 124.0 1.26 1.94</td>
</tr>
<tr>
<td>B3LYP*</td>
<td>54.5</td>
<td>3.43</td>
<td>3.69</td>
<td>115.1 124.0 1.26 1.94</td>
</tr>
<tr>
<td>M06</td>
<td>59.0</td>
<td>3.38</td>
<td>3.75</td>
<td>115.6 122.3 1.33 1.87</td>
</tr>
<tr>
<td>M06L</td>
<td>42.3</td>
<td>3.40</td>
<td>3.76</td>
<td>115.5 122.1 1.31 1.90</td>
</tr>
<tr>
<td>MPW1BK</td>
<td>42.6</td>
<td>3.38</td>
<td>3.73</td>
<td>115.1 122.5 1.25 1.88</td>
</tr>
</tbody>
</table>

It is important to note that binuclear serum paraoxonase-1 (PON) found in humans (PBD: 3SRE) use catalytic calcium in its active site.\(^{43-44}\) Notably, the name paraoxonase comes from its ability to degrade the organophosphate paraoxon.\(^{43}\) The surround residues around the metals differ from GpdQ,\(^{43}\) it would be interesting to design active site in GdpQ that would utilize calcium due its low toxicity in high concentrations and good tolerance towards air and moisture.

### 4.3.7 \(M\alpha\)-ligand Substitutions for Homonuclear Cd(II)/Cd(II)

Here we show preliminarily data on the result of mutating the surrounding ligands around the \(\alpha\)-ligands as this may affect the rate of phosphate cleavage. We tested single, double and triple mutations on the \(M\alpha\) switching aspartate with histidine and vise versa. Figure 4.4. Out of the seven variants tested, only two had a stable RC. Unfortunately,
none had a barrier that was less than the calculated wild type (Table 4.5). It is interesting to note we see that the difference in barrier do, in general, correspond to the strain in the RC. This is consistent to what we shown is a major factor in rate enhancement from the Zn, Co, Mn, Cd and Ca series. The barriers calculated were 73.6 (WT), 85.3 (H8D) and 94.9 kJ mol\(^{-1}\) (D10H, D197H) and had \(\angle_{\text{phos}\alpha-P\text--phos}\beta\) angles of 118.8°, 121.1° and 120.1° respectively. The difference in the latter two may be explained by the overstretch of the TS of 127.6° and 129.7°.

**Figure 4.5.** Pictorial representation of the substitution on the M\(\alpha\)-site. Blue font colour represents the change from the wild type (top).
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

**Table 4.6** Selected parameters in the Mα substituted ligands.

<table>
<thead>
<tr>
<th>Substitution</th>
<th>Relative TS Energy (kJ mol$^{-1}$)</th>
<th>$\angle_{\text{phosO} \alpha - P - \text{phosO} \beta}$ (°)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RC</td>
<td>TS</td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>73.6</td>
<td>118.8</td>
<td>126.9</td>
</tr>
<tr>
<td>H8D</td>
<td>85.3</td>
<td>121.1</td>
<td>127.6</td>
</tr>
<tr>
<td>D197H</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D10H, D197H</td>
<td>94.9</td>
<td>120.1</td>
<td>129.7</td>
</tr>
<tr>
<td>H8D, D10H, D197H</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>D10H</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>H8D, D10H</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>H8D, D197H</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
4.4 Conclusions

Herein we explored the $\mu$–OH directed mechanism in homonuclear Glycerophosphodiesterase (GpdQ). The mechanism was found to occur in a single step with the concomitant nucleophilic attack and protonation of the leaving group. The ordering of the calculated barriers was in general agreement for all functionals and exact for B3LYP*, M06 and MPW1BK.

The driving force behind the mechanism was explored. The reactivity of GpdQ over uncatalyzed reactions can be contributed greatly to reactant destabilization. It was calculated that the strain of phosphate indicated by a compressed $\angle(\text{phos}O\alpha–\text{P}–\text{phos}O\beta)$ activates the substrate. For Cd(II)/Cd(II), Zn(II)/Zn(II), Co(II)/Co(II) and Mn(II)/Mn(II), the averages $\angle(\text{phos}O–\text{P}–\text{phos}O\beta)$ is 116.2, 116.6, 116.8 and 117.6 making a difference for DMP in water of 9.5, 9.1°, 8.9° and 8.1° respectively. For Cd(II)/Cd(II), reactant destabilization was calculated to be 39.9 kJ mol$^{-1}$.

The source of the fine-tuning between the metals was also explored. First, it was hypothesized that Cd(II)/Cd(II) and Mn(II)/Mn(II) may stabilize the leaving group of the transition state. This was characterized by a systematic increase in $\text{phos}OR^–\text{phos}P$ in the transition states complexes coupled with a decrease in $\text{His}_{82}^–\text{H}^–\text{phos}OR$. Furthermore, due to the increase in ionic radii between the metals, the strong interactions between the phosphate and the metals then stretched the transition state. The relative energy of difference contributed to the TS $\angle(\text{phos}O\alpha–\text{P}–\text{phos}O\beta)$ was compared to Zn(II)/Zn(II) structure. It was found that Mn(II)/Mn(II), Co(II)/Co(II) and Cd(II)/Cd(II) are more stable by 3.1, 3.0 and 11.0 kJ mol$^{-1}$ respectively.

A Ca(II)/Ca(II) derivative was explored in an attempt to extrapolate the previous results in to a larger metal. It was found that reactant destabilization dictated the lower barrier. A $\angle\text{phos}O\alpha–\text{P}–\text{phos}O\beta$ of 115.3° contributed to a 12.1 kJ mol$^{-1}$ increase in energy.
over Cd(II)/Cd(II). Finally, we attempted to increase the rate of reaction by substituting ligand on Mα-site. Unfortunately, none of the mutation resulted in a barrier lower than the wild type however $\angle_{\text{phos} \alpha - \text{P} - \text{phos} \beta}$ angle did agree with our previous predictions.

Overall, we provide several contributing factors to metal dependent phosphate hydrolysis in GdpQ. Future studies must now revisit the terminal water mechanism to determine if it is competitive with the mechanism shown in this study.

4.5 References
Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ

Chapter 4: Computational Insights into the Catalytic Mechanism of Homonuclear GpdQ


Chapter 5   Insights into the Use of Boron for Inhibiting β-lactamases
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

5.1 Introduction

The increasing frequency of antibiotic resistance bacteria is perhaps one of the greatest health related challenges of this century.\(^1\) This has been magnified due to the overuse/misuse of antibiotics dating back to when penicillin mass production was first made possible in the 1940s.\(^2\)\(^3\) This is problematic because antibiotics are vital to combat infections in endless medical procedures from surgery to cancer chemotherapy.\(^1\)\(^4\) In addition, there is also an astonishing economic consequence, for example it cost both Europe\(^5\)\(^6\) and the United States\(^7\) several billion dollars annually.\(^8\) In result, there is an urgency to react.

One of the main modes of antibiotic resistance is through the class of enzymes known as β-lactamases.\(^9\) Bacteria have an enzyme that has glycosyltransferase and transpeptidases activity called penicillin binding proteins, PBPs that are responsible for the construction of cell walls.\(^10\)\(^11\) PBPs are a bacterial specific target for antibiotics because eukaryotes do not have cell walls.\(^10\) β-lactam antibiotics, in particular, function by inhibiting specifically the transpeptidases activity of PBPs.\(^10\) Unfortunately, bacteria then counter this inhibition via β-lactamases that hydrolyze the β-lactam ring before it can perform its function. Collectively, there are four different classes of β-lactamases by the Ambler classification system that places them, based on there sequence, in classes A through D.\(^12\) Classes A, C and D are mechanistically similar as they are serine driven hydrolases. Class B enzymes are the exception as they are metallo-β-lactamases (MBL).\(^12\)\(^13\)

Interest here is in the TEM-1 class A β-lactamase. It is well characterized due to its high frequency in infections therefore its mechanism has been extensively studied.\(^14\)\(^19\) The mechanism can be summarized into two half reactions: acylation and hydrolysis/deacylation. In the former, Ser70 is acylated by the β-lactam’s carbonyl carbon. This occurs through the formation of an unstable oxyanion that is stabilized by an “oxyanion
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

pocket” that is composed of the backbone hydrogen bonds from Ala237 and Ser70.\textsuperscript{19} However, how this oxyanion itself is formed has been greatly debated. There have been three different proposals for the general base for this step.\textsuperscript{18,20-21} First, Lys73 is proposed to act as the catalytic base to activate Ser70 for nucleophilic attack.\textsuperscript{21} Second, a substrate-assisted mechanism has been proposed by Díaz et al.\textsuperscript{18} Finally, Glu166 has been proposed as the mechanistic base either directly deprotonating Ser70 or indirectly through a conserved water.\textsuperscript{18-19} It is important to note that it is likely that the enzyme could undergo all the aforementioned pathways because site-directed mutagenesis of Lys73, Glu166 or both only reduces acylation activity.\textsuperscript{22-23} In contrast to the first step, the second step is not generally debated. The deacylation of the substrate is facilitated by a conserved water using Glu166 to release the product.\textsuperscript{24}

\textbf{Figure 5.1.} The proposed mechanism for the acylation step with Glu166 as the catalytic base. A) Activation of serine B) the unstable tetrahedral intermediate C) Protonation of amine and ring opening D) Acyl-enzyme.\textsuperscript{25}

Currently, there are four commercially available β-lactamases inhibitors: clavulanic acid, sulbactam and tazobactum and most recently avibactam Figure 5.1. They all have the β-lactam functional group with the exception of avibactam.\textsuperscript{26} They are reasonably effective against class A and to a lesser extent against class C.\textsuperscript{27-28} A strategy that has been proven to work is to combine β-lactamase inhibitors with β-lactam antibiotic to enhance potency.\textsuperscript{29} For example, clavulanic acid matched with amoxicillin, sulbactam
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

with ampicillin, and tazobactam with piperacillin are all approved for use in the U.S.A.\textsuperscript{27} As time progresses, mutations have allowed β-lactamases to confer more resistance to these cocktails and it is important to continually be innovating new strategies.\textsuperscript{30}

![Figure 5.2](image_url)

**Figure 5.2.** The four currently approved β-lactamases inhibitors and the most promising boronic acid inhibitors.

One strategy that has recently received increased attention is revisiting the use of boronic acid transition state inhibitors (BATSIs).\textsuperscript{31} In principle, BATSIs are advantageous mainly due to two chemical properties of boron (1) when in its neutral form; an empty p-orbital is present that is a strong Lewis acid. (2) The resulting intermediate mimics a more stable version of the original β-lactam transition state complex.\textsuperscript{32} Interestingly, very little is known about boron in therapeutics mostly because there is very few naturally occurring boron containing compounds.\textsuperscript{32} This may be due that the fact that many antibiotics are often modeled after natural substrates and the synthetic methods to build complex organic boron containing molecule are not as advanced.

The most common type of boron-containing β-lactamases inhibitors are derivatives of boronic acid.\textsuperscript{33-35} The pK\textsubscript{a} often fall between 8-10 therefore they are three-coordinated and pronated at physiological pH.\textsuperscript{36} Notably, boronic acids are susceptible to reversible
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

nucleophile attack in water therefore they are in an equilibrium between the anionic and natural state. Currently, most boronic acid antibiotics are in its discovery stage (Pre phase 1 trial) however there is one combination of carbapenem/boronic acid that is currently in phase II trial.\textsuperscript{37} RPX7009 (Figure 5.2) is a very promising β-lactamase inhibitor with activity against serine carbapenemase.\textsuperscript{13,38}

Computational chemistry allows for the determination of different functional groups that would be difficult to synthesize experimentally to gain knowledge about its fundamental properties. This allows for cost friendly way to rationally design new inhibitors. In this study, we explored the thermodynamic stability of the boron inhibition pathway benchmarked against a common antibiotic, benzylpenicillin.

5.2 Computational Methods

5.2.1 Molecular Docking and Molecular Dynamic Simulations

The crystal structure PDB ID: 1BTL\textsuperscript{39} from Class A Escherichia coli TEM1 high-resolution 1.8 Å was used as the template for the investigation. First, the initial crystallographic data was prepared using the molecular operating environment (MOE) software by removing the waters and resolving the system.\textsuperscript{40} The structure was then allowed to relax using the AMBER12:EHT force field to remove any strain from the preparation of the protein structure. The substrate benzylpenicillin was then docked in its known active site\textsuperscript{19} using the induced fit formulism.\textsuperscript{40} The to evaluate the top poses, the London dG scoring function was used with the AMBER12:EHT force field refinement keeping the top 10 scores. The top structure that was oriented in a matter that correlated well with the known binding mode of benzylpenicillin was selected. A boron-substituted substrate was created by mutating the carbonyl carbon to neutral boron, which sequentially required the protonation of the adjacent oxygen (Figure 5.3). The substitute
was then docked and scored to compare the relative binding mode. This hypothetical compound would allow for the testing of boron as a Lewis acid compared to the natural substrate with directly comparable binding orientations.

Figure 5.3. Comparative models of A) benzylpenicillin and B) boron substitute used in this study.

Two one-nanosecond simulations were performed on both complexes to find a relaxed minimum energy conformation with the substrate present. In the MD simulation, all atoms were free to move and a time step of 2 fs was used. The Columbic interactions were calculated with the PME method and 8-10 Å were the upper limit cutoff for the van der Waals interactions. The system was first annealed with a temperature built from 150-300 K and equilibrated for 100 ps. The production run was carried out at 300 K for 1000 ps. An estimated average structure found in the production run was then selected and minimized with the AMBER12:EHT force field. This structure was the starting structure to build the QM/MM model.

5.2.2 The QM/MM Model and Calculations.

The quantum mechanical layer of the QM/MM model consisted of the full substrate; and truncated surrounding residues, Ser70, Lys73, Asn170, Glu166, Ser130, Lys234, Ala237, Gly236, and Arg244 for a total of 150 atoms. The molecular mechanics layer contained 2806 atoms that surrounded the QM layer. Atoms 10 angstroms from the boundary of the QM layer were free to move whereas all atoms beyond that point were fixed. The calculations were performed using the Gaussian 09 suite of programs.$^{41}$
Optimizations were performed using the ONIOM formulism. The hybrid exchange functional B3LYP using the 6-31G(d, p) Pople basis set was used to calculate the QM atoms geometries and energies. The MM region was calculated using AMBER96 as implemented in Gaussian 09. This functional has been used previous when studying this system and results were comparable with experimental work.\textsuperscript{19,25}

\textbf{Figure 5.4.} Pictorial representation of the QM layer modeled in the study, the MM was excluded for clarity. Atom colours, Red=O, Blue=N, Yellow=S, Grey=C, White=H.

\section*{5.3 Results and Discussion}

\subsection*{5.3.1 Selection of the Benzylpenicillin and the Boron Substitute}

The focus of this study was to make a single variable change to quantify the use of boron as a Lewis acid in serine protease reaction. In order to do this, two scenarios must
be achieved, (1) the system must be standardized to a well studied β-lactamases and its associated mechanism and (2) the binding must be very similar because this would have an effect on the thermodynamics of the mechanism.\textsuperscript{19} Experimentally, TEM-1 β-lactamases have been found to strongly bind and cleave benzylpenicillin, \(k_{\text{cat}} = 1200 \text{ s}^{-1}\), \(K_m = 22 \mu\text{M}\), \(k_{\text{cat}}/K_m = 5.5 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}\).\textsuperscript{43} A small \(K_m\) and a large \(k_{\text{cat}}\) indicates that the structure of penicillin is optimal for TEM-1’s active site. Furthermore, there have been previous computational studies on benzylpenicillin\textsuperscript{19} and the similar penicillanic acid\textsuperscript{21} to provide a means to benchmark our model. Therefore, boron-substituted benzylpenicillin (Figure 5.4, B), was constructed where the boron was replaced the carbonyl of the amide in the lactam ring and the carbonyl was protonated. This model, in theory, would allow for direct comparison of the thermodynamics of the natural substrate with that of boron. Before this compound was studied in an enzyme environment, it was tested for general thermodynamic stability. Both the neutral and anionic form was optimized at the B3LYP/6-31G(d,p) level of theory and the anionic state was +10.1 kJ mol\(^{-1}\) higher in energy than the neutral form. Therefore in solution, it would be more favourable for this compound to be in the neutral state vs. the anionic state, which is consistent with the pKa determined for boronic acids.\textsuperscript{36}

5.3.2. Molecular Docking Analysis

A molecular docking protocol was performed to both substrates in the active site before the molecular dynamic (MD) simulation. The dock was guided with the aid of previous computational work,\textsuperscript{19,21} by disregarding scores that did not bind benzylpenicillin correctly into the active site. The Gibbs free energy of benzylpenicillin corresponding to the top score, in the proper orientation, was -62.6 kJ mol\(^{-1}\). This large exergonic binding energy corresponds with the low \(K_m\) value found in class A β-lactamases.\textsuperscript{43} The dock protocol was then carried out on the boron substitute and the pose selected overlaid virtually identically with benzylpenicillin. Interestingly, the docking
score for the boron substitute was slightly greater at -66.5 kJ mol\(^{-1}\). This small difference may be due to the slightly stronger interaction between Arg244 and the carboxylic acid on the substrate.

**Figure 5.5.** An pictorial overlay of the docked structures of benzylpenicillin and the boron analogue. Atom colours: grey=C, red=O, blue=N yellow=S orange=B and white=H.

### 5.3.3 Molecular Dynamic Analysis

A one-nanosecond molecular dynamic simulation was performed on both docked models to allow the substrate to relax in the environment of the active site residues. Due to the nature of the mechanism, a few key interactions were recorded through the simulation, Table 5.1. First and foremost was the comparison between the electrophilic carbonyl carbon of penicillin and the nucleophilic oxygen of Ser70 vs. the analogous boron and Ser70 as this interaction is vital of the initiation of the mechanism. The
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

distances between C...OSer70 and B...OSer70 averaged 3.16 +/- 0.17 Å and 3.14 +/- 0.16 Å respectively. Second, Ala237 has been shown to be important for the stabilization of the substrates oxyanion, therefore this interaction must be present to make the mechanism energetically feasible. As expected, there is a slightly stronger interaction between the benzylpenicillin’s carbonyl Ala237-H...O=C of 1.87 +/- 0.11 Å than the hydroxyl group on the boron analogue Ala237-H...O(H)B of 2.00 +/- 0.16 Å. However, it is clear that in both cases, the substrate fits directly in the pocket. Notably, in the boron-substituted model, the conserved water immediately leaves the active site and serine is strongly H-bonded to Glu166, 1.36 Å apart.

Table 5.1. Key average distances obtained throughout the MD simulation.

<table>
<thead>
<tr>
<th></th>
<th>Benzylpenicillin</th>
<th>Boron-Substitute</th>
</tr>
</thead>
<tbody>
<tr>
<td>r(X...OSer70)</td>
<td>1.87 0.11</td>
<td>2.00 0.16</td>
</tr>
<tr>
<td>r(N_{sub}...OSer130)</td>
<td>3.16 0.17</td>
<td>3.14 0.21</td>
</tr>
<tr>
<td>r(Ala237-H...O-X*)</td>
<td>3.46 0.42</td>
<td>3.01 0.44</td>
</tr>
</tbody>
</table>

*Where X is either the carbonyl carbon in benzylpenicillin or boron in the derivative

5.3.4 Reproducing the Catalytic Mechanism of Benzylpenicillin

Hermman and coworkers first performed a QM/MM study of the acylation half reaction of benzylpenicillin with Glu166 as the base. Meroueh and coworkers re-evaluated that mechanism along with Lys73 as the catalytic base with penicillanic acid. In both cases, the first acylation step was rate limiting followed by a rapid protonation of the nitrogen of the amide group to create the acyl enzyme. Here, we modeled the former with Glu166 as the base, acknowledging that both mechanisms could be competitive.
Figure 5.6. Key structural parameters for the acylation step of benzyl penicillin. The first part of the mechanism, the nucleophilic attack, is highlighted in red. The main atoms involved in the second part, the C-N cleavage are highlight in blue. All bond lengths are in Angstroms (Å).

The initial optimization of the reactive complex (RC) from the MD simulation reduced C·-OSer70 from 3.16 Å to 3.10 Å, Figure 5.4. A water (WAT_cat) molecule was found to remain in the active site, between Glu166 and Ser70 throughout the simulation. This water was found to be in a position that is optimal to act as a proton shuttle positioned 1.75 Å between HO-H·O(H)Ser70 and 1.81 Å from HO-H·OOC-Glu166. The activation of Ser70 has been shown to occur concomitantly with the nucleophilic
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

attack on the substrate whereas a step wise mechanism is not likely due to the formation of an unstable positively charged alcohol functional group.\textsuperscript{19} The concomitant reaction is shown in Figure 5.7 and is found to occur at a barrier of 92.6 kJ mol\textsuperscript{-1}. This is comparable to the re-evaluation where the calculated mechanism of penicillanic acid first step was found to be 108.7 kJ mol\textsuperscript{-1}.\textsuperscript{21} Furthermore, this is consistent with experimental results that show that penicillanic acid k\textsubscript{cat} to be 38 s\textsuperscript{-1} nearly a thousand-fold decrease from benzylpenicillin.\textsuperscript{43} This would predict a theoretical reduction of about 20 kJ mol\textsuperscript{-1} using the Arrhenius equation.

Stabilization of the transition state is achieved by both the backbone N-H of both Ala237 and Ser70 (Figure 5.6). The increase in negative charge on the carbonyl oxygen is inferred by the shortening of the hydrogen bond between $^p$RC to $^p$TS\textsubscript{1} from 2.04 to 1.74 and 3.83 to 2.26 Å for Ala237 and Ser70 respectively. The C\textsuperscript{−}\textsuperscript{−}O Ser70 length is 1.94 Å in $^p$TS\textsubscript{1} with both hydrogen atoms proportionately between Ser70 and WAT\textsubscript{cat}, and WAT\textsubscript{cat} Glu166 (Figure 5.6). This results in the formation of the metastable tetrahedral intermediate $^p$IC\textsubscript{1}. This intermediate was found to be 48.3 kJ mol\textsuperscript{-1} higher than the $^p$RC (Figure 5.7). The carbonyl double bond has lengthen, C–O is now 1.30 Å from 1.21 in the $^p$RC. The oxygen is partially anionic that is stabilized by the backbone hydrogen bonds of Ala237 and Ser70 of 1.68 and 2.03 Å respectively.
Chapter 5: Computational Insights into the use of Boron for Inhibiting $\beta$-lactamases

Figure 5.7. Potential energy surface representing the acylation of benzylpenicillin in TEM-1. Atom colours: grey=C, red=O, blue=N and white=H.

The second part of the acylation half reaction is to collapse the meta-stable tetrahedral intermediate by cleaving the carbon-nitrogen bond. Ser130 functions as a catalytic acid by protonating the nitrogen amide of the substrate. In a similar manner to the $^\text{P}^\text{TS1}$, Ser130 must receive a proton simultaneously with its deprotonation because anionic Ser130 would be unstable. The proton source has been proposed to be Lys73.$^{19,21}$ Interestingly, upon optimization Lys73 and Lys234 are both positioned in close proximity to do this task. Since $^\text{P}^\text{IC1}$ had Lys234 already strongly polarized to Ser130, it was used as the base in the reaction; either would be sufficient since its known not to be the rate-limiting step. The proton transfer was found to occur at a low barrier of 26.4 kJ mol$^{-1}$ forming the acyl enzyme $^\text{P}^\text{IC2}$ Figure 5.7. Finally, the tetrahedral intermediate collapse
formed the acyl-enzyme (IC3) that was 83.3 kJ mol\(^{-1}\) more stable than the metastable \(^{p}\)IC1 and overall -35 kJ mol\(^{-1}\) less than the \(^{p}\)RC. This emphasizes the irreversibility and therefore the necessity of the second half reaction for \(\beta\)-lactamases, deacylation by WAT\(_{\text{cat}}\) whose mechanism is beyond the scope of this preliminary study.

5.3.5 The Catalytic Mechanism of Boron Substituted Benzyl Penicillin

The boron derivative was modeled using QM/MM in a similar manner to penicillin. However, as mentioned, there were key differences in MD average structure. As previously stated, the water left the active site pocket during the molecular dynamic simulation, which resulted in the Glu166 to be in close proximity to Ser70. This conformation has been observed before in TEM-1 as the \(\Omega\) loop that contains Glu166 has been shown to have high flexibility and mobility.\(^{44}\) After optimization, the \(^{B}\)RC had Ser70 strongly hydrogen bound to Glu166 with 1.80 Å separating the two moieties. It is important to note that there may be a direct energetic consequence due to this change. The pKa of water is greater than that of methanol (15.7 vs. 15.5) therefore the conjugate base of a carboxylic acid must abstract a proton through two moieties that are unfavourable to be deprotonated. Fortunately, Meroueh and coworkers also carried out the mechanism of Lys73 as the base where Lys73 (a stronger base than Glu166) directly abstracted a proton from Ser70. The difference between their Glu166 mechanism and Lys73 was about 17 kJ mol\(^{-1}\).\(^{21}\) This would roughly represent the upper and lower bound to direct proton abstraction and should be considered if directly comparing benzylpenicillin and boron’s mechanisms.

Further examination the structural parameters calculated in the \(^{B}\)RC showed that B\(^{-}\)OSer70 lengthened slightly from the MD simulation from 3.16 to 3.24 Å Figure 5.8. The oxyanion pocket that consists of the backbone Ala237 and Ser70 still interacts with the hydroxyl group, Ala237-H\(^{-}\)-O(H)B distance was 2.21 Å. The B–N and B–O bond in the \(^{B}\)RC was 1.44 Å and 1.35 Å. In contrast to benzylpenicillin’s mechanism, the first
transition state did not require the activation of serine. This occurred at a barrier of 36.4 kJ mol\(^{-1}\) to yield an intermediate complex, \(^{\text{B}}\text{IC1}\) that lies 7.9 kJ mol\(^{-1}\) higher than the reactant, Figure 5.9. This is compared to the difference between the \(^{\text{P}}\text{RC}\) and \(^{\text{P}}\text{TS}\) and \(^{\text{P}}\text{RC}\) and \(^{\text{P}}\text{IC1}\) of +92.6 and +48.3 kJ mol\(^{-1}\), Figure 5.7. This explicitly shows theoretical application of boronic acid transition state inhibitors (BATSIs) that mimic the tetrahedral transition state of the natural reaction but offers much higher stability in both the TS and IC. The electron density approaching the boron atom does increase the H-bonding in the pocket Ala237-H\(\cdots\)O(H)B by 0.24 Å to 1.97 Å and also elongated the B–N bond by 0.05 Å to 1.49 Å.

\(^{\text{B}}\text{IC1}\) differs from \(^{\text{P}}\text{IC1}\) in that the Ser70 is positively charged (Figure 5.8). In \(^{\text{B}}\text{IC1}\), B\(^{-}\)Oser70 distance is 1.68 Å, indicating a weak covalent/dative bond therefore boron is now tetrahedral and has a formal negative charge. The electrons from nitrogen lone pair can no longer semi-overlap with boron once empty p-orbital and we see that the B–N lengthens to 1.55 Å, a 0.11 Å increase from the \(^{\text{B}}\text{RC}\). The next likely step is the abstraction of the proton from Ser70. The Glu166 is in a moderately strong hydrogen bond distance from H\(-\)Oser70, 2.01 Å apart, and this abstraction was modeled as \(^{\text{B}}\text{TS2}\). In \(^{\text{B}}\text{TS2}\), the proton is positioned linearly between Glu166 and Ser70 of 1.26 Å and 1.18 Å respectively, Figure 5.8. In \(^{\text{B}}\text{IC2}\), B-Oser70 is reduced to 1.49 Å. This is a much stronger bonding interaction that is comparable to the B–OH in the “oxyanion” hole that has lengthened of 1.46 Å Figure 5.8. The increase in electron density on boron increased B–N further to 1.64 Å.
Figure 5.8. Key structural parameters for the acylation step of benzyl penicillin. The first part of the mechanism, the nucleophilic attack, is highlighted in red. The main atoms involved in the second part, the C-N cleavage are highlight in blue. All bond lengths are in Angstroms (Å).

The second part of the acylation half reaction is the protonation of the adjacent nitrogen group to cleave the B–N interaction, which is now completely analogous to benzylpenicillin. In the $\text{^BIC}{\text{}}_2$ optimized structure, Ser130 is already strongly hydrogen bonded to the nitrogen hydrogen acceptor of the substrate, separated by 1.63 Å, Figure
5.8. For direct comparison, we chose to use Lys234 to protonate Ser130 concomitantly with the protonated of the substrate's nitrogen atom, $^B$TS3. As with penicillin, this reaction is rapid in comparison to the rate-limiting step illustrated by a relative difference of only 4.0 kJ mol$^{-1}$. The final intermediate, $^B$IC3 has the proton completely transferred to the nitrogen, this lengthened the B–N distance to 1.75 Å. Surprisingly, the interaction did not break as in penicillin.

![Potential energy surface](image)

**Figure 5.9** Potential energy surface representing the acylation of benzylpenicillin in TEM-1. Atom colours: grey=C, red=O, blue=N and white=H.

In an alternative pathway, we tested if after the $^B$IC1 is formed, could the nitrogen be protonated before glutamate abstracts the proton from Ser70? The barrier of this process was 28.0 kJ mol$^{-1}$ compared to the 16.3 kJ mol$^{-1}$ of proton abstraction in the prior
mechanism. The following proton abstraction from Ser70 by Glu166 was only 4.4 kJ mol\(^{-1}\). This shows that once the tetrahedral complex is formed, it is likely that the proton is abstracted from serine followed by the protonation of nitrogen by Ser70. However these results show that both mechanisms are energetically feasible.

**5.3.6 Implications of the Preliminary Results**

There are considerable differences between the acylation of the substrate through a carbonyl carbon and boron. In benzylpenicillin, the barrier was calculated to be 92.6 kJ mol\(^{-1}\) compared to 36.4 kJ mol\(^{-1}\) for boron. This can be partly attributed to boron having an empty p-orbital as well that Ser70 did not need to be activated prior to nucleophilic attack. In addition, the carbonyl must be stabilized by the oxyanion hole whereas the hydroxyl group does not require significant stabilization. While it is beneficial to have a low barrier for nucleophilic attack there are also consequences. One of the benefits of using \(\beta\)-lactams as an antibiotic inhibitor is that once formed, the reaction is not reversible. This is highlighted in the benzylpenicillin reaction potential energy coordinate (Figure 5.7). The energy difference between the \(^{p}\text{IC2}\), acyl-enzyme and the rate-limiting step is 127.6 kJ mol\(^{-1}\). This is would be difficult to overcome. In contrast, the difference between \(^{B}\text{IC3}\), acyl enzyme, and the rate-limiting step \(^{B}\text{TS1}\) is only 47.9 kJ mol\(^{-1}\). The potential energy curve is also much flatter in the case of boron making it much more reversible; this has been suggested and observed experimentally for boric acid derivatives.\(^{45}\)

Although the second step of \(\beta\)-lactamases was not explicitly explored in this study inferences can be made because it well established.\(^{46}\) The consensus of this reaction is that Glu166 uses a conserved water to hydrolyze the Ser70-subtrated interaction. This is possible because the acyl enzyme has an electrophilic carbon susceptible to nucleophilic attack. In the case tetrahedral boron, nucleophilic attack cannot take place because it is no longer electrophilic and is formally anionic.\(^{28}\)
Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

Another major difference in the two mechanisms was the B–N bond did not break upon protonation. This is an interesting aspect that has potential implications as it is thought that cyclical boron variant would be more selective to β-lactamase as its pocket is designed to hold cyclic transition states and intermediates.\(^{28-29}\) Recently, a crystal structure of RPX7009 (Figure 5.2) was resolved in the active site of a Class A.\(^{29}\) This crystal structure suggests that Ser130 is in close proximity to the adjacent oxygen to boron in the 6-membered ring. This shows that cyclic boron derivatives are orientated similarly as 4-member β-lactam.

Although this study uses the boron-nitrogen combination, its use in cells would be difficult. B-N bonds are generally not hydrolytically stable unless modifications are made to protect boron.\(^{47}\) 1,2-azaborines, on the other hand, are water stable conjugated ring structures.\(^{48-49}\) Furthermore, the boron on 1,2-azaborines also has been shown to have electrophilic character and found to inhibit Enoyl reductase (ENR) in the same way as it would inhibit a serine protease.\(^{50-51}\) It would be interesting to test if a 1,2-azaborines analogue could properly bind to the active site of TEM-1 β-lactamase.

5.4 Conclusions

In this study, preliminary calculations were performed to investigate the effectiveness of using boron as a Lewis acid for the inhibition of class A TEM-1 β-lactamase. This was accomplished by comparing the well-characterized mechanism of benzylpenicillin/penicillanic acid with a single variable change to make a cyclic boron substitute. This allowed for the quantification of using boron as a Lewis acid in a serine protease type reaction.

In the control acylation of TEM-1 with benzylpenicillin, the rate limiting first step was the nucleophilic attack of Ser70 on the substrate with a calculated barrier of 94.5 kJ
mol\(^{-1}\). This was compared and found to be consistent with previous work on this enzyme.\(^2\)\(^1\) This validated that the QM/MM model was satisfactory to quantify the boron substitute. The 4-member hypothetical compound conferred virtually identical binding to the TEM-1 active site as confirmed by molecular docking and molecular dynamic simulation. The rate-limiting step for the nucleophilic attack on boron was substantially less than that of β-lactam, a barrier of 36.4 kJ mol\(^{-1}\). The nature of the mechanism shows why the barrier is lower, that is, Ser70 did not need to be activated in order to covalently bind to the substrate nor did it form an oxyanion intermediate. This exemplifies the power of using a transition state analogue. A second major difference is in the second part of the first half reaction. In benzylpenicillin, this step causes the cleavage of the amide bond whereas in the boron analogue, the B–N bond elongates but stays intact. This has mechanistic consequences for the second half reaction the deacylation. A water molecule would not be able to cleave the “acyl-enzyme” effectively halting the second half reaction.

However, the major drawback to using boron as an inhibitor is its reversibility. This has now been quantified where the potential energy of the reverse reaction, going from the acyl enzyme and the rate-limiting step, is only 47.9 kJ mol\(^{-1}\) apart. This can be compared to benzylpenicillin that would have to overcome a much larger barrier of 127.6 kJ mol\(^{-1}\). Future work requires stabilizing the structure of the “acyl-enzyme” complex to minimize the reversibility. This may be aided by the recently published crystal structure of the cyclic RPX7009 bound in the active site of a class A and class C β-lactamase reveling of a 6-member ring would orientate itself in the active site.\(^2\)\(^9\)

### 5.5 References


Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases

Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases


40. *Molecular Operating Environment (MOE)*, 2013.08; Chemical Computing Group Inc.: Montréal, QC, Canada, 2015.


Chapter 5: Computational Insights into the use of Boron for Inhibiting β-lactamases


Chapter 6  Conclusions
6.1 Conclusions

In this thesis, several computational methods have been used to investigate properties of various enzymes and intermediates. In particular, we investigated 3– and 4–coordinated sulfur intermediates that have been proposed/shown to occur in biological systems. Furthermore, we investigated two catalytic mechanisms (1) the phosphodiesterase activity in GpdQ and (2) the hydrolysis activity of TEM-1 β-lactamase with emphasis on its inhibition.

In Chapter 3, we provided a foundation for selecting a proper level of theory for futures enzymatic studies that contain 3– and 4–coordinated sulfur intermediates. In particular, systems that were explicitly investigated were sulfuranes from archaeal thioredoxin peroxidase (ApTPx)\(^1\) and methionine sulfoxide reductase (Msr A&B),\(^2-3\) sulfinic acid, thiosulfinate and sulfinic acid phosphoryl esters from sulfuredxodin,\(^4\) sulfonic acid from myeloperoxidase,\(^5\) sulfinamides and sulfinimines from myeloperoxidase\(^6-7\) and collagen IV\(^8\) respectively. In addition, we also investigated the relative energy difference of the common precursor sulfenic acid and its tautomer. Throughout our study the geometric parameters and energy values were benchmarked against post Hartree-Fock (HF) \textit{ab initio} methods, QCISD and MP2, at very large basis sets, 6-311+G(2df,p) or 6-311++G(3df,3pd). From a relative energy perspective, our results showed that the accuracy to the benchmark calculations is strongly dependent on the functional used. More specifically, the HF coefficients systematically dictates the performance of the calculation, that is, B3LYP, B3PW91 and M06, all have between 20-27% HF and they perform very well at a high basis set. In contrast, M06HF (100% HF) and M062X (~56%) overestimate the relative energy values whereas M06L (0% HF) underestimates them at the same basis set. This highlights the importance of benchmarking the system to be studied. In terms of geometry, B3LYP always over
predicted bond lengths in virtually all scenarios whereas M06 and M062X were consistently the best performers.

In Chapter 4, we investigated a possible mechanism for the hydrolysis of DMP by Glycerophosphodiesterase (GpdQ). Specifically, we investigated the $\mu$–OH directed mechanism as previously proposed in our group with homonuclear compositions. Our results show that mechanism could occur in a single step with the concomitant nucleophilic attack and protonation of the leaving group. The ordering of the calculated barriers are in agreement with the experimental determined $k_{cat}$ for the phosphodiester bis(p-nitrophenyl) phosphate (bpNPP). How the different metal combinations achieved remarkably different rates was explored. Our investigation suggested that reactant destabilization of specifically $\angle_{phosO\alpha–P–phosO\beta}$ contributes mostly to the enzyme’s ability to speed up the reaction from the uncatalyzed hydrolysis, marked by a 39.9 kJ mol$^{-1}$. Collectively, we hypothesized that the fine-tuning due to the metals could be two fold. First, the stabilization of the leaving group was predominate in the Cd(II)/Cd(II) and Mn(II)/Mn(II) combinations indicated by a shortening of the His82-H…phosOR coupled with an elongation of phosOR…phosP over the Zn(II)/Zn(II) and Co(II)/Co(II) combinations. Second, the increased interactions between the phosphate and the metals systematically increase the $\angle (phosO\alpha–P–phosO\beta)$ contributing up to a 11 kJ mol$^{-1}$ difference in relative energy. When extrapolating these results into a larger Ca(II)/Ca(II) GpdQ derivative, we determined that reactant strain was satisfactory on its own to explain the difference in the calculated barriers. Collectively, this work begins to provide a foundation for understanding a $\mu$–OH mechanism in GpdQ that now must be compared the proposed terminal water mechanism.

The final computational investigation was a preliminary study that thermodynamically quantifies boron’s use as a means to inhibiting $\beta$-lactamase, Chapter 5. Although this work is not complete, we can extrapolate information from the current results. First, we
modeled a β-lactam antibiotic whose mechanism is well characterized through previous computational work.\textsuperscript{11-12} We show computationally that potent Lewis acid in boron does significantly increases the rate of enzyme acylation compared to a standard β-lactam by 58.1 kJ mol\textsuperscript{-1}. This decrease can be contributed to the fact that Ser70 does not need to be activated prior to nucleophilic attack, as in required in of the less reactive carbonyl in a β-lactam. However a major draw back, as previously been mentioned is the high reversibility of enzyme-substrate bonding. This has now been quantified with only 47.9 kJ mol\textsuperscript{-1} to return from the “acyl enzyme” back to the reactant complex in the boron substitute compared to the irreversibility of β-lactam of 127.6 kJ mol\textsuperscript{-1}. In light of the reversibility, it appears that designing a cyclic ring structure for boron may end up providing a good strategy to inhibit β-lactamase as shown in RPX7009\textsuperscript{13} because of the stability of the four-coordinate boron center in the active site and subsequent hydrolysis could not occur in that state. Further studies are required and highly encouraged on cyclic boron containing ring structures to fine-tune the reversibility of this mechanism.

6.2 References


Chapter 6: Conclusions

Vita Auctoris

<table>
<thead>
<tr>
<th>NAME:</th>
<th>Daniel James Simard</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLACE OF BIRTH:</td>
<td>Cottam, ON</td>
</tr>
<tr>
<td>YEAR OF BIRTH:</td>
<td>1991</td>
</tr>
<tr>
<td>EDUCATION:</td>
<td>Essex High School, Essex, ON, 2009</td>
</tr>
<tr>
<td></td>
<td>University of Windsor, B.Sc., Windsor, ON, 2013</td>
</tr>
<tr>
<td></td>
<td>University of Windsor, M.Sc., Windsor, ON, 2015</td>
</tr>
</tbody>
</table>