

2013

Computational Insights into Redox and Metal Ion Biochemistry

Eric Andre Bushnell
University of Windsor

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

Recommended Citation

Bushnell, Eric Andre, "Computational Insights into Redox and Metal Ion Biochemistry" (2013). *Electronic Theses and Dissertations*. Paper 4915.

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 Insights into Redox and Metal Ion Biochemistry

By

Eric A. C. Bushnell

A Dissertation

Submitted to the Faculty of Graduate Studies
through the Department of Chemistry and Biochemistry
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada

© 2013 E. Bushnell

Declaration of Co-Authorship/Previous publications

I. Co-Authorship Declaration

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

Chapter 3 was done in collaboration with Dr. WenJuan Huang and Prof. Jorge Llano under the supervision of Prof. James W. Gauld.

Chapter 4 was done in collaboration with Dr. Edvin Erdtman, Prof. Jorge Llano and Prof. Leif A. Eriksson under the supervision of Prof. James W. Gauld.

Chapter 5 was done in collaboration with Mr. Rami Gherib under the supervision of Prof. James W. Gauld.

Chapter 6 was done in collaboration with my supervisor Prof. James W. Gauld.

Chapter 7 was done in collaboration with Miss. Riam Jamil under the supervision of Prof. James W. Gauld.

Chapter 8 was done in collaboration with Mr. Grant B. Fortowsky under the supervision of Prof. James W. Gauld.

Chapter 9 was done in collaboration with Mr. Phil De Luna under the supervision of Prof. James W. Gauld.

Chapter 10 was done in collaboration with Mr. Joel N. Almasi under the supervision of Prof. 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 qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publications

This thesis includes eight original papers that have been previously published for publication in peer-reviewed journals as follows:

	Citation	Status
Chapter 3	Bushnell, E. A. C.; Huang, W.J.; Llano, J.; Gauld, J. W. <i>J. Phys. Chem. B</i> 2012 , <i>116</i> , 5205-5212.	Published
Chapter 4	Bushnell, E. A. C.; Erdtman, E.; Llano, J.; Eriksson, L. A.; Gauld, J. W. <i>J. Comput. Chem.</i> 2011 , <i>32</i> , 822-834.	Published
Chapter 5	Bushnell, E. A. C.; Gherib, R.; Gauld, J. W. <i>J. Phys. Chem. B</i> (Accepted, May 15 th 2013)	Published
Chapter 6	Bushnell, E. A. C.; Gauld, J. W. <i>J. Comput. Chem.</i> 2013 , <i>34</i> , 141-148.	Published
Chapter 7	Bushnell, E. A. C.; Jamil, R.; Gauld, J. W. <i>J. Biol. Inorg. Chem.</i> 2013 , <i>18</i> , 343-355.	Published
Chapter 8	Bushnell, E. A. C.; Fortowsky, G. B.; Gauld, J. W. <i>Inorg. Chem.</i> 2012 , <i>51</i> , 13351-13356.	Published
Chapter 9	De Luna, P.; Bushnell, E. A. C.; Gauld, J. W. <i>J. Phys. Chem. A</i> , 2013 , <i>117</i> (19), 4057-4065	Published
Chapter 10	Almasi, J. N.; Bushnell, E. A. C.; Gauld, J. W. <i>Molecules</i> 2011 , <i>16</i> , 8569-8589.	Published

I certify that I have obtained written permission from the copyright owners to include the above-published material in my thesis. I certify that the above material describes work completed while registered as a graduate student at the University of Windsor.

I declare that to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violates any proprietary rights and that the 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 reference 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 written permission from the copyright owners to include such material 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

Redox reactions are ubiquitous and essential in nature. The enzymes that catalyse these reactions enable the efficient flow of electrons within organisms as well synthesising essential biomolecules. Importantly, these enzymes ensure that the respective cellular chemical reactions occur at life-sustainable rates. Notably, many enzymes that utilize such reactions for catalysis require cofactors and/or coenzymes that may include metal ions. In this thesis, redox and metal-ion biochemistry are investigated using computational methods.

Dedication

This work is dedicated to my family.

Acknowledgements

Over the last several years I have had the opportunity to work with many talented and wonderful people. Importantly, with their help and support I have been able to overcome the many hurdles that hindered my goal of obtaining a Ph.D. Thus, I would like to take time to express my appreciation to these people. In particular, I would like to express great gratitude to my supervisor James W. Gauld for his mentoring and support during the past several years. Importantly, his valuable and constructive suggestions have helped me become a better scientist.

My sincere gratitude also goes to my Ph.D. committee for taking the time to read and evaluate this work. Furthermore, Prof. Samuel A. Johnson, Prof. Bulent Mutus, Prof. Joel E. Gagnon are thanked for their time and patient instructions and discussions involving my research during the past few years. Also Prof. Maria J. Ramos is thanked for agreeing to be my external examiner.

I would like to thank Prof. Leif Eriksson, Prof. Christopher Francklyn, Dr. Edwin Erdtman and Prof. Jorge Llano for their collaboration. With their help I was able to learn new and useful techniques used in the various projects over the years. Importantly, such knowledge will aid me in my future endeavors.

I am also grateful for all past and present members of the Gauld group. Dr. Haining Liu, Jesse Robinet, Dr. Wenjuan Huang, Bogdan Ion, Hisham Dokainish, Avtinder Basra, Riam Jamil, Joel Almasi, Grant Fortowsky, Natalia Mroz, Rami Gherib, Erum Kazim, Phil De Luna, Daniel Simard, Mohamed Aboelnga and Anthony Deschamps are all acknowledged for the discussions of quantum and computational chemistry. In addition, I wish to thank many of the above people in their involvement with my research. Without their help I would not have been so productive.

Thanks are also extended to Mr. Joe Lichaa for technical support. In particular, in maintaining the cluster at the University of Windsor. In addition I thank Mrs. Marlene Bezaire for all the help and direction for the last five years during my Ph.D. Without her help the past five years would have been much more confusing.

My deepest gratitude is extended to Patricia Bushnell for her support and encouragement throughout my study. In addition I would like to offer special appreciations to Ian, Lucienne, Anne, Paul, Jamielynne, Jillian, and Jackson Bushnell, John, Chris, Steve, George and Mary Merwin and Ange-Aimée Cook for their support and encouragement.

Lastly I thank the Natural Sciences and Engineering Research Council of Canada for a PGS-D Scholarship. In addition, I thank Canada Foundation for Innovation, the Ontario Innovation Trust and SHARCNET for additional computational resources and graduate scholarships.

Table of Contents

Declaration of Co-authorship/Previous Publication-----	iii
Abstract -----	vi
Dedication -----	vii
Acknowledgements-----	viii
List of Tables-----	xvii
List of Figures-----	xix
List of Schemes -----	xxvii
List of Appendices-----	xxix
List of Abbreviations-----	xxxi

Chapter 1

Introduction

1.1 Introduction -----	2
1.2 References-----	6

Chapter 2

Theoretical Methods

2.1 Introduction -----	10
2.2 Schrödinger Equation-----	10
2.3 Born-Oppenheimer (BO) Approximation-----	11
2.4 Orbital Approximation -----	12
2.5 Basis set Expansion -----	13
2.6 Basis Sets-----	14
<i>Split Valence</i> -----	14
<i>Polarization Functions</i> -----	14

<i>Diffuse Functions</i> -----	15
<i>Effective Core Potentials</i> -----	15
<i>Dispersion correcting potentials (DCP)</i> -----	15
2.7 Variational Theorem-----	16
2.8 Hartree-Fock (HF) Theory -----	16
2.9 Electron Correlation -----	17
<i>Møller-Plesset Perturbation Theory</i> -----	18
2.10 Density Functional Theory -----	18
2.11 Molecular Mechanics -----	21
2.12 Quantum Mechanics/Molecular Mechanics-----	22
2.13 Molecular Dynamics -----	24
2.14 Solvation-----	25
2.15 Computational Tools -----	26
<i>Geometry Optimization</i> -----	26
<i>Frequency Calculations</i> -----	26
<i>Single Point Calculations</i> -----	27
<i>Redox Potentials</i> -----	27
<i>Potential Energy Surfaces (PES)</i> -----	28
2.16 References -----	30

Chapter 3

A MD Investigation into Substrate Binding and Identity of the Catalytic Base in the Mechanism of ThrRS

3.1 Introduction -----	34
3.2 Computational Methods-----	38
<i>Design of Chemical Model</i> -----	38
<i>Solvation and Annealing</i> -----	38
<i>Molecular Dynamics Production Runs</i> -----	39

3.3 Results and Discussion -----	40
<i>Effects of protonation at His309-N^c on its direct hydrogen-bonding to Ado76-2'-OH</i> -----	42
<i>Effects of an additional active-site water on the hydrogen bonding between His309-N^c and Ado76-2'-OH.</i> -----	48
3.4 Conclusions -----	53
3.5 References-----	54

Chapter 4

The First Branching Point in Porphyrin Biosynthesis: A Systematic Docking, MD and QM/MM Study of Substrate Binding and Mechanism of UROD

4.1 Introduction -----	58
4.2 Computational Methods -----	62
<i>Molecular Docking</i> -----	63
<i>Molecular Dynamics Equilibration</i> -----	63
<i>QM/MM calculations</i> -----	64
4.3 Results and Discussion -----	66
<i>Binding of the URO-III substrate to the UROD active site</i> -----	66
<i>UROD–URO-III Complex I</i> -----	66
<i>UROD–URO-III Complex II</i> -----	68
<i>UROD–URO-III Complex III</i> -----	68
<i>Identifying possible mechanistic acids</i> -----	69
<i>QM/MM Investigation of the Catalytic Mechanism of hUROD</i> -----	74
<i>The Initial Proton Transfer</i> -----	75
<i>Acetate decarboxylation</i> -----	77
<i>Abstraction of a Proton by Arg37</i> -----	79
4.4 Conclusions -----	81

4.5 References-----	81
---------------------	----

Chapter 5

Insights into the Catalytic Mechanism of Coral Allene Oxide Synthase: A Dispersion Corrected Density Functional Theory Study

5.1 Introduction -----	87
5.2 Computational Methods -----	89
<i>Molecular Docking</i> -----	89
<i>QM calculations</i> -----	90
<i>Chemical Model</i> -----	92
5.3 Results and Discussion -----	93
<i>The catalytic mechanism of cAOS, without dispersion corrections</i> -----	93
<i>The Effects of Including Dispersion Corrections</i> -----	101
<i>The Effects of Choice of DFT functional</i> -----	104
5.4 Conclusions -----	107
5.5 References -----	110

Chapter 6

An Assessment of Standard, Hybrid, Meta and Hybrid-meta GGA Density Functional Theory Methods for Open-Shell Systems: The Case of the Non-Heme Iron Enzyme 8R-LOX

6.1 Introduction -----	115
6.2 Computational Methods -----	119
<i>Molecular Docking</i> -----	119
<i>Molecular Dynamics Equilibration</i> -----	119
<i>QM/MM Model</i> -----	120
<i>QM/MM calculations</i> -----	121

6.3 Results and Discussion -----	122
<i>The effect of DFT functional choice on optimized geometries</i> -----	122
<i>The effect of DFT method choice on calculated spin densities in ⁴II and ⁶II</i> -----	126
<i>The effect of DFT functional choice on the relative energies of ⁴II and ⁶II</i> -----	129
6.4 Conclusions -----	132
6.5 References -----	133

Chapter 7

Gaining insight into the chemistry of lipoxigenases: A computational investigation into the catalytic mechanism of (8R)-LOX

7.1 Introduction -----	139
7.2 Computational Methods -----	143
<i>Molecular Docking</i> -----	143
<i>Molecular Dynamics Relaxation</i> -----	143
<i>QM/MM calculations</i> -----	144
<i>QM/MM Chemical Model</i> -----	145
7.3 Results and Discussion -----	147
<i>Quartet, Sextet and Octet Reactive Complexes</i> -----	147
<i>Activation of the substrate: the initial PCET step</i> -----	150
<i>Formation of the peroxy ($-OO^{\bullet}$) radical</i> -----	153
<i>Conformational change in the active site</i> -----	155
<i>Regeneration of the Fe(III) center</i> -----	160
7.4 Conclusions -----	162
7.5 References -----	163

Chapter 8

Model Iron-Oxo Species and the Oxidation of Imidazole: Insights into the Mechanism of OvoA and EgtB?

8.1 Introduction -----	170
------------------------	-----

8.2 Computational Methods	173
8.3. Results and Discussion	175
8.4 Conclusions	183
8.5 References	183

Chapter 9

A Density Functional Theory Investigation into the Binding of the Antioxidants

Ergothioneine and Ovothiol to Copper

9.1 Introduction	190
9.2 Computational Methods	192
9.3 Results and Discussion	193
<i>Geometrical Assessment of $[Cu(ESH)_3]^+$</i>	194
<i>Geometrical Assessment of $[Cu(OSH)_3]^+$</i>	197
<i>Geometrical Assessment of $Cu(OS^-)_2$</i>	199
<i>Geometrical Assessment of $Cu(ES^-)_2$</i>	201
<i>Effects of OSH and ESH on Free Energies of Cu(II)/Cu(I) Redox Cycling</i>	204
9.4 Conclusions	208
9.5 References	209

Chapter 10

A QM/MM-Based Computational Investigation on the Catalytic Mechanism of

Saccharopine Reductase

10.1 Introduction	215
10.2 Computational Methods	218
10.3. Results and Discussion	221
<i>The pK_a of the Substrate Glutamate's α-Amine</i>	221
<i>Mechanism for Formation of Saccharopine</i>	223
<i>Formation of a Carbinolamine Intermediate</i>	226
<i>Rearrangement of the Carbinolamine Intermediate II</i>	227

<i>Formation of the Schiff Base and its Reduction</i> -----	229
<i>Obtaining More Accurate and Reliable Energies for the Mechanism of Saccharopine Reductase</i> -----	230
<i>The Inclusion of Electron Correlation Effects</i> -----	230
<i>The Effects of the Protein's Anisotropic Polar Environment</i> -----	233
<i>The Effects of Increasing the Basis Set Size</i> -----	236
10.4 Conclusions-----	237
10.5 References -----	240
Chapter 11	
Conclusions	
11.1 Conclusions-----	245
11.1 References -----	252
Vita Auctoris	
Vita Auctoris -----	253

List of Tables

Table 2.1. Chemical potentials of an electron and proton in various reference states. ³⁸ -----	28
Table 3.1. Summary of Michaelis complex models considered in this present study. -----	42
Table 5.1. Spin densities on select atoms obtained at the B3LYP*/6-311+G(2df,p) level of theory.-----	95
Table 6.1. Absolute differences ($ \Delta r $) between optimized ⁴II and ⁶II structures, obtained at the ONIOM(DFT _i /BS1:AMBER94) level of theory for selected interactions. All absolute differences are in Å. -----	124
Table 6.2. Selected spin densities in intermediate complexes ⁴II and ⁶II as calculated at the ONIOM(DFT _i /BS2//DFT _i /BS1:AMBER94)–EE level of theory (see text). -----	127
Table 6.3. The relative energies of the intermediate complexes ⁴II and ⁶II obtained at the different ONIOM(DFT _i /BS2//DFT _i /BS1:AMBER94)–EE levels of theory. -----	130
Table 6.4. The energies for the quartet and sextet systems at the ONIOM(M06–L/BS2//B3LYP*/BS1:AMBER94) and ONIOM(M06/BS2//B3LYP*/BS1:AMBER94) levels of theory in the electronic embedding formulism. -----	132
Table 8.1. Adiabatic free energies (kJ mol ⁻¹) for oxidation of the Im via ET and PCET. -----	176
Table 8.2. Adiabatic free energies (kJ mol ⁻¹) for reduction of the Fe(III)–O ₂ ^{•-} complexes (A) via ET or PCET. -----	177
Table 8.3. Adiabatic free energies (kJ mol ⁻¹) for reduction of the Fe(IV)=O complexes (B) via ET or PCET.-----	179

Table 8.4. Adiabatic free energies (kJ mol^{-1}) for reduction of Fe-OO-S complexes (C; see Figure 9.6) via PCET.-----	182
Table 9.1. The aqueous solution relative free energies (kJ mol^{-1}) of <i>cis</i> and <i>trans</i> - ϵ -Cu(ES ⁻) ₂ and <i>cis</i> and <i>trans</i> - δ -Cu(ES ⁻) ₂ calculated at the DFT _i /6-311+G(2df,p) + ΔG_{Corr} level of theory (DFT _i = B3LYP, BP86, M06 or M06L).-----	203
Table 9.2. Reduction potentials (V) for the reduction of Cu(OS ⁻) ₂ and Cu(ES ⁻) ₂ complexes in aqueous solution based on Eq. 9.1 and 9.2 calculated at the IEF-PCM-DFT _i /6-311+G(2df,p) + ΔG_{Corr} level of theory.-----	206
Table 9.3. Reduction potentials (V) for reduction of the disulfides OSSO, ESSE and MeSSMe in aqueous solution calculated at the IEF-PCM-DFT _i /6-311+G(2df,p) + ΔG_{Corr} level of theory.-----	207

List of Figures

Figure 1.1. Schematic illustrations of the (a) non-heme Fe(II) cofactor of 8R-Lipoxygenase and the coenzyme (b) heme and (c) NADH.-----	3
Figure 1.2. The key steps in the synthesis of Protoporphyrin IX. ²⁶ Intermediates formed are represented in black while the enzymes that catalyze the respective transformations are in blue. -----	4
Figure 1.3. A schematic representation of (a) Cpd I (X = cysteinyl, histidyl or tyrosyl depending on enzyme family) and (b) non-heme iron Cpd II analog (L's may be ligating substrate(s) or active site residue(s)). -----	6
Figure 2.1. A schematic representation of the increasingly complex DFT functionals in the attempt to reach the exact functional. ¹⁰ -----	20
Figure 2.2. A schematic representation of a QM/MM model in which the surrounding environment is modeled at the MM level of theory while the reacting center is at the QM level of theory. -----	23
Figure 2.3. A PES for a two step mechanism which consists of a reactive complex (RC), two transition states (TS), intermediate complex (IC) and product complex (PC). -----	29
Figure 3.1. Active site of ThrRS bound with: (a) Thr-AMP [PDB ID: 1EVL], ¹⁹ (b) tRNA ^{Thr} and AMP [PDB ID: 1QF6]. ²⁰ -----	37
Figure 3.2. RMSDs in the positions of the imidazole of His309, phenol of Tyr462, ribose sugar ring of Ado76 and the threonyl moiety during the 10 ns simulation of model I. ---	43
Figure 3.3. Overlay of the five representative average structures of model I obtained from cluster analysis of the RMSDs during the 10 ns MD production simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those	

used in the calculation of the RMSDs presented in **Figure 3.2**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white)]-----43

Figure 3.4. RMSDs in the positions of the imidazole of His309, phenol of Tyr462, ribose sugar ring of Ado76 and the threonyl moiety during the 10 ns simulation of model

I-H⁺. -----45

Figure 3.5. Overlay of the five representative average structures of model **I-H⁺** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.4**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white)] -----47

Figure 3.6. RMSDs in the positions of the His309 imidazole, phenol of Tyr462, ribose sugar ring of Ado76, threonyl moiety and added "bridging" water molecule during the 10 ns simulation of model **II**. -----49

Figure 3.7. Overlay of the five representative average structures of model **II** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.6**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white); added H₂O (green)]. -----50

Figure 3.8. RMSDs in the positions of the His309 imidazole, phenol of Tyr462, ribose sugar ring of Ado76, threonyl moiety and added "bridging" water molecule during the 10 ns simulation of model **II-H⁺**.-----51

Figure 3.9. Overlay of the five representative average structures of model **II-H⁺** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.8**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white); added H₂O (green)] -----52

Figure 4.1. Structural models of the active site: (a) Arrangement of the catalytically active residues of UROD in complex I. (b) QM/MM model with residues in QM (non-shaded region) and MM (shaded region) layers. Residues in red modeled as side chains with atoms fixed at the truncation position. Remainder of residues modeled as side chain and backbone with C α positions fixed. Residues in blue represent side chains modeled as hydrogen (S85 and L341). ----- 65

Figure 4.2. Schematic representation of the binding modes of URO-III in the active site of UROD. Hydrogen-bonding interactions are encoded as follows: green dashed line, bonding with the side chain R; blue dashed line, bonding with backbone amide; olive-green dashed line, bonding with water; magenta dashed line, salt bridge. ----- 67

Figure 4.3. Distance fluctuations between C2 of ring **D** (**D**-C2) and both the nearest water oxygen [$r(\mathbf{D}\text{-C2}\cdots\text{OH}_2)$] and the closest guanidinium proton on Arg37 [$r(\mathbf{D}\text{-C2}\cdots\text{H-Arg37})$] in in UROD-URO-III complexes **I**, **II** and **III** (Cx-**I**, Cx-**II**, Cx-**III**) over 100 ps of MD equilibration. ----- 69

Figure 4.4. The models of (a) 3-acetopyrrole and proposed intermediate models (b) 3-methylenepyrrole derivative (c) 3-methylpyrrole (d) N-ethylguanidino, (e) 3-acetopyrrole coordinated to an acetate and (e) 3-methylenepyrrole derivative coordinated to an acetate.----- 71

Figure 4.5. Effect of medium polarity on proton affinities of C2 in 3-acetopyrrole, with and without hydrogen bond to acetate, 3-methylenepyrrole and N-ethylguanidino.----- 72

Figure 4.6. Effect of medium polarity on the basicities of the N-ethylguanidino group and the pyrrolic C3' centre, with and without hydrogen bonding with acetate.----- 73

Figure 4.7. Free energy diagram for proton transfer from the guanidinium of Arg37 to the C2 centre of ring **D**. ----- 75

Figure 4.8. Free energy profile for the decarboxylation of ring **D**. ----- 77

- Figure 4.9.** Optimized structure of **TS2**, the transition structure for the decarboxylation of ring **D** with concomitant proton transfer from Arg50 to the C3' centre. For clarity, not all hydrogens are shown.----- 78
- Figure 4.10.** Gibbs energy profile for abstraction of the initially transferred proton to C2 by Arg37. ----- 80
- Figure 5.1.** The active site bound-substrate model used in the investigation of AOS. Carbon atoms labeled with an asterisk were fixed to their final MM minimized position. ----- 92
- Figure 5.2.** Free energy surfaces for the overall catalytic mechanism of cAOS obtained at the IEF-PCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory. Surface color code: red (doublet); black (quartet) and blue (sextet). -----94
- Figure 5.3.** The optimized structures of **RC** and **IC1** for the doublet and quartet states. Bond lengths are in Ångstroms (Å) where the values in red correspond to the former state, while those in black are for the latter state.----- 96
- Figure 5.4.** The optimized structures of **IC2** and **PC** for the doublet and quartet states. Bond lengths are in Ångstroms (Å) where the values in red correspond to the former state, while those in black are for the latter state.----- 98
- Figure 5.5.** Schematic illustration of the dispersion-corrected free energy surfaces obtained at the IEF-PCM-B3LYP*/6-311+G(2df,p). + ΔG_{Cor} + $\Delta \text{Disp}_{\text{Cor}}$ level of theory for the overall catalytic mechanism of cAOS. Surface color code: red (doublet); black (quartet) and blue (sextet).----- 102
- Figure 5.6.** Free energy surfaces for the catalytic mechanism of cAOS obtained at the IEF-PCM-BP86/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory. Surface color code: red (doublet; S=1/2); black (quartet; S=3/2).----- 104
- Figure 5.7.** Free energy surfaces for the catalytic mechanism of cAOS obtained at the IEF-PCM-DFT_i/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory (DFT_i = M06,

B3LYP, B3LYP* and BP86): **(a)** doublet ($S=1/2$) state and **(b)** quartet ($S=3/2$) state. The values given in parentheses indicate the %HF contribution in the functional.----- 107

Figure 6.1. Schematic representation of the QM/MM model used. Groups in the inner circle have been modeled at the QM level of theory, while the residues in the outer circle have been modeled at the MM level of theory. Colour code for residues: included in their entirety (black); modeled as Gly, i.e., only the backbone was included with R-groups replaced by a hydrogen (red); modeled as only their R-group, i.e., only their C_{α} and side chains included (blue). ----- 121

Figure 6.2. The QM-region model used for the active site of 8R-LOX with the pentadienyl intermediate (**II**) bound. ----- 123

Figure 7.1. The Fe coordination site as observed in an X-ray Crystal Structure of the holoenzyme form of 8R-LOX (PDB: 3FG1).⁶ ----- 140

Figure 7.2. The QM/MM model used in the investigation of 8R-LOX. **(a)** Those components included in the high layer (QM) are shown in tube and ball format while those of the low layer are shown in wire format. For clarity, hydrogen atoms have not been shown. **(b)** Schematic representation of the QM/MM model used: groups in the inner circle were included in the QM region while those in the outer circle were included in the MM region. Note, residues in black were included in their entirety, those in red only had their peptide backbone included, while those in blue only included their side chains. ----- 146

Figure 7.3. Schematic illustration of the PESs obtained, at the ONIOM(B3LYP*/BS2//B3LYP*/BS1:AMBER94)–EE level of theory, for the overall catalytic mechanism of LOX. Color code: red, black and blue surfaces represent the octet, sextet and quartet systems, respectively. ----- 147

Figure 7.4. The electronic configurations obtained from the spin densities at the ONIOM(B3LYP*/BS2//B3LYP*/BS1:AMBER94)–EE level of theory. While the orbitals of the Fe center are not expected to be those for an ideal octahedral complex they

have been drawn that way to simplify discussion. Note that the electronic configurations presented above only represent key structures during the mechanism of 8R–LOX. ---- 149

Figure 7.5. Schematic illustration of the optimized structures of the reactant complexes (**RC**) with key selected distances (Ångstroms) shown, as obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory. Color code: blue (quartet), black (sextet), red (octet).----- 149

Figure 7.6. Schematic illustration of the optimized structures, obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory, of **TS1** and the subsequent radical-intermediate **IC1** for the PCET process with selected distances (Ångstrom) shown. Color code: blue (quartet), black (sextet), red (octet).----- 151

Figure 7.7. Optimized structures, obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory, with key selected distances for the TS of the attack of O₂ (**TS2**) and the subsequent peroxy–radical intermediate (**IC2**). Color code: blue (quartet), black (sextet).----- 154

Figure 7.8. Schematic illustration of the dihedral angles ω_1 ($\angle C_{10}-C_9-C_8-C_7$) and ω_2 ($\angle C_9-C_8-O-O^\bullet$). ----- 156

Figure 7.9. Optimized structures obtained at the ONIOM(B3LYP/6-31g(d):AMBER)-ME level of theory with key selected dihedral angles (degrees) for the quartet (blue) and sextet (black) intermediate **IC3**. ----- 156

Figure 7.10. Optimized structures, obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory, with key selected distances (Ångstroms) and dihedral angles (degrees) for **TS4** and intermediate **IC4**. Color code: Color code: blue (quartet), black (sextet).----- 157

Figure 7.11. The 3D PES for the rotation of the peroxy group to an angle required for H• abstraction. The energies were obtained at the IEF-PCM–B3LYP/6-311G(2df,p)//B3LYP/6-31G(d) level of theory.----- 159

- Figure 7.12.** Optimized structures, obtained at the ONIOM(B3LYP/6-31g(d):AMBER)-ME level of theory, with key selected distances for the transition structure (**TS5**) of the reduction of the Fe–OH₂ center and the subsequent product (**PC**). Color code: blue (quartet), black (sextet).----- 161
- Figure 8.1.** The initial five- and six-coordinate Fe(III)–O₂^{•-} complexes considered herein. ----- 174
- Figure 8.2.** The initial five- and six-coordinate Fe(IV)=O complexes considered herein. ----- 178
- Figure 8.3.** The initial five- and six-coordinate ferryl-peroxy-sulfur complexes considered herein. ----- 180
- Figure 9.1.** Schematic illustration of possible complexes formed by ligation of 4-thiol-N¹-methyl-5-methylimidazole (OSH) or 2-thiol-4-methylimidazole (ESH) to Cu(I) and Cu(II). ----- 194
- Figure 9.2.** Plot of the optimized $r_{Avg}(\text{Cu(I)}\cdots\text{S})$ bond distances (in Ångstroms) for [Cu(ESH)₃]⁺.----- 195
- Figure 9.3.** Plot of the optimized $r_{Avg}(\text{Cu(I)}\cdots\text{S})$ bond distances (in Ångstroms) for [Cu(OSH)₃]⁺.----- 197
- Figure 9.4.** Plots of the optimized (a) $r_{Avg}(\text{Cu(II)}\cdots\text{S})$ and (b) $r_{Avg}(\text{Cu(II)}\cdots\text{N})$ distances (in Ångstroms) for *trans*-Cu(OS⁻)₂.----- 200
- Figure 9.5.** Plot of the optimized (a) $r_{Avg}(\text{Cu(II)}\cdots\text{S})$ and (b) $r_{Avg}(\text{Cu(II)}\cdots\text{N}_\delta)$ distances (in Ångstroms) for *trans*- δ -Cu(ES⁻)₂.----- 202
- Figure 10.1.** (a) The X-ray crystal structure (PDB ID: 1E5Q)–derived QM/MM model used to investigate the catalytic mechanism of saccharopine reductase. (b) Schematic representation of the QM/MM model: groups in the inner and outer circles have been modelled at the HF/6-31G(d) and AMBER94 levels of theory, respectively. Note, residues, waters and functional groups in black have been included in the model in their entirety, while residues in red have only had their peptide backbone included. ----- 220

- Figure 10.2.** The PAs for the α -amines of AASA and Glu with respect to the local environment. The horizontal dashed line represents the PA of H₂O in solution. ----- 222
- Figure 10.3.** Overall PES for the catalytic mechanism of saccharopine reductase obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory with inclusion of Gibbs corrections. ----- 224
- Figure 10.4.** Optimized structures obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory of the reactant complex (**RC**), transition structure (**TS1**) and the carbinolamine intermediate **I1** with selected distances shown (in Ångstroms). ----- 225
- Figure 10.5.** Optimized structures of **TS2**, **I2**, **TS3**, **I3**, **TS4** and **I4** obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory with selected distances shown (in Ångstroms).----- 225
- Figure 10.6.** Optimized structures of **TS5**, **I5**, **TS6** and **PC** obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory with selected distances shown (in Ångstroms).----- 228
- Figure 10.7.** Overall PES's obtained for the catalytic mechanism of saccharopine reductase at the (i) ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME + Gibb's Corrections (dashed blue line), (ii) ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–EE + Gibb's Corrections (dotted pink line), and (iii) ONIOM(MP2/6-311G(d,p)//HF/6-31G(d):AMBER94)–EE + Gibb's Corrections (solid black line) levels of theory. ----- 231

List of Schemes

Scheme 1.1. The main stages in protein synthesis. -----	2
Scheme 3.1. Aminoacyl transfer from aa-AMP to Ado76-3'-OH (second half-reaction) as catalyzed by a class-II aaRS. -----	35
Scheme 3.2. Proposed mechanism for aminoacyl transfer as catalyzed by ThrRS via deprotonation of 2'-OH-ribonucleoside by His309. ¹⁴ -----	36
Scheme 4.1. Overall sequential decarboxylation reaction catalyzed by UROD: the acetate of ring D is decarboxylated first followed by those of rings A , B and C ^{1,2,4,5} . -----	58
Scheme 4.2. Proposed 'blueprint' general acid–base mechanism for the UROD-catalyzed decarboxylation of the acetates of URO-III ^{11,12,20} . HA and HB represent general acids.-----	60
Scheme 5.1. The proposed catalytic mechanism of cAOS. ¹¹ For reasons of clarity not all hydrogens have been added while the porphyrin ring is shown as a porphine ring. -----	87
Scheme 6.1. The generally accepted mechanism of LOXs. -----	117
Scheme 6.2. The electronic configurations for the quartet (red) and sextet (black) ground states of the first intermediate complex (II) in the mechanism of LOXs.-----	118
Scheme 7.1. The overall peroxidation of arachidonic acid as catalyzed by the LOX family member 8R-LOX to give the product 8R-hydroperoxyeicosatetraenoic acid (8R-HPETE). -----	139
Scheme 7.2. The general mechanism of Lipoygenases (LOXs). -----	142
Scheme 8.1. Proposed mechanism for formation of a histidyl-sulfoxide via a radical mechanism with coupling between the sulfoxide and histidyl occurring at the latter C δ position (i.e. synthesis of 5-HisCysSO). ¹⁷ -----	171

Scheme 8.2. Proposed intermediates for the formation of the histidyl-sulfoxide based on thermodynamic stability. -----	183
Scheme 9.1. Schematic illustration of (a) Ovothiol (OSH_A : $n = 0$; OSH_B : $n = 1$; and OSH_C : $n = 2$) and (b) Ergothioneine, in their preferred state at biological pH. -----	191
Scheme 10.1. The catalytic mechanism of saccharopine reductase as proposed by Johansson <i>et al.</i> ⁴ -----	217
Scheme 10.2. The general acid/base catalytic mechanism as proposed by Vashishtha <i>et al.</i> ¹ -----	217

List of Appendices

Table A1. Optimized xyz-coordinates of the QM/MM structures for the proposed decarboxylation of ring D obtained at the ONIOM(B3LYP/6-31G(d):AMBER). Single point energies calculated at the B3LYP/6-311+G(2df,p) level of theory with corrections obtained at the ONIOM(B3LYP/6-31G(d):AMBER) level of theory.

Figure B1. Schematic illustration of the PESs obtained at the IEF-PCM-M06/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Coor} energies level of theory for the overall catalytic mechanism of cAOS. Color code: blue, black and red surfaces represent the S=5/2, 3/2 and 1/2 states, respectively.

Figure B2. Schematic illustration of the distances between centers of mass of the substrate (or intermediate) and porphine ring. Geometries obtained at the B3LYP/BS1 level of theory Color code: blue, black and red surfaces represent the S=5/2, 3/2 and 1/2 states, respectively.

Figure B3: The optimized structures, with selected bond lengths in Angstroms (Å), of **TS1**, **TS2** and **TS3** for the S=1/2 (doublet: red values) and S=3/2 (quartet: black values) states.

Table B1. XYZ's of the optimized geometries at the B3LYP/BS1 level of theory.

Table C1. Brief summary of the properties of ${}^6\text{RC}$.

Table C2. Optimized (see Computational Methods section) $L_n\text{-Fe-}L_m$ ($m \neq n$) angles in ${}^8\text{RC}$, ${}^6\text{RC}$ and ${}^8\text{RC}$.

Table C3. Calculated (see Computational Methods) spin densities on select atom centers within the various QM/MM optimized reactant, intermediate, product complexes and transition structures.

Table C4. Cartesian coordinates of the various QM/MM models. AMBER atom types and charges have been included.

Table D1. XYZ's for the optimized structures at the B3LYP/6-31G(d) level of theory

Table E1. Optimized Copper-Sulfur Bond Lengths^a for [Cu(ESH)₃]⁺

Table E2. Optimized Copper-Sulfur Bonds^a for [Cu(OSH)₃]⁺

Table E3. Optimized Nitrogen-Copper and Sulfur-Copper Bond Lengths^a for *trans*- and *cis*-[Cu(OS⁻)₂]

Table E4. Optimized Nitrogen-Copper and Sulfur-Copper Bond Lengths^a for *trans*- and *cis*-[Cu(ϵ -ES⁻)₂] and *trans*- and *cis*-[Cu(δ -ES⁻)₂] Complexes

Table E5. Optimized Carbon-Sulfur Bond Lengths^a for [Cu(ESH)₃]⁺

Table E6. Optimized Carbon-Sulfur Bond Lengths^a for [Cu(OSH)₃]⁺

Table E7. Optimized Carbon-Sulfur Bond Lengths^a for Cu(OS⁻)₂

Table E8. Optimized Carbon-Sulfur Bond Lengths^a for Cu(ES⁻)₂

Table E9. XYZ's of the optimized geometries at the DFTi/6-311+G(2df,p) level of theory

List of Abbreviations and Symbols

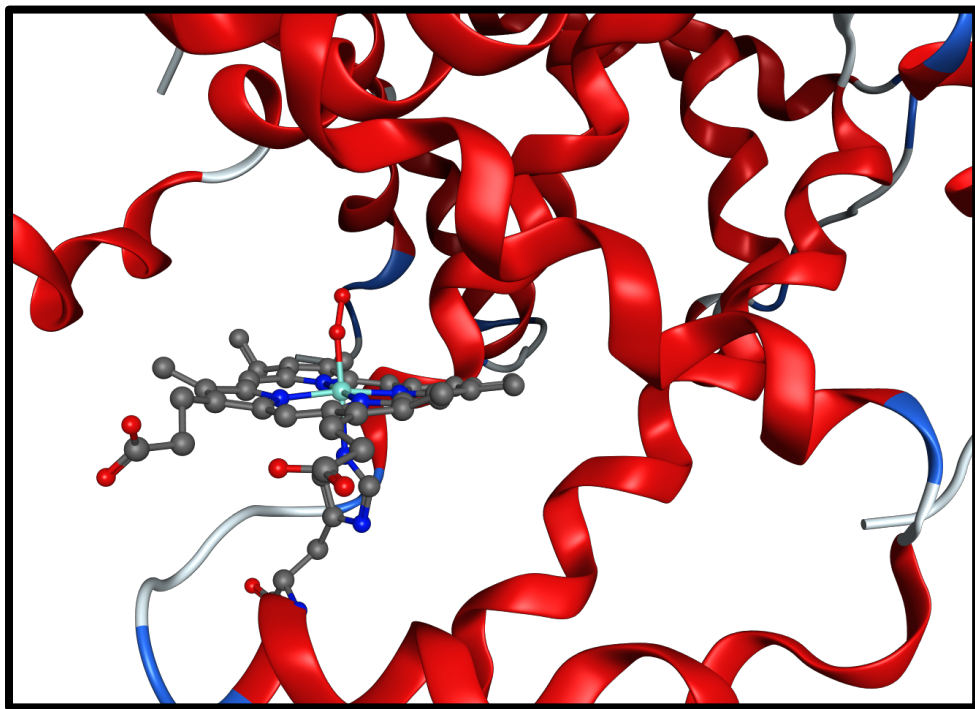
8(<i>R</i>)-HPETE	8(<i>R</i>)-hydroperoxyeicosatetraenoic
AA	Arachidonic acid
aaAMP	Aminoacyl-adenylate
aaRS	α Aminoacyl-tRNA synthetase
AASA	α -aminoadipate- δ -semialdehyde
aa-tRNA ^{aa}	Aminoacyl-tRNA ^{aa}
AOS	Allene oxide synthase
ATP	Adenosine triphosphate
Arg	Arginyl
Asp	Aspartyl
BO	Born-Oppenheimer
BS	Basis set
Cpd I	Compound I
Cpd II	Compound II
DCP	Dispersion correcting potentials
DFT	Density Functional Theory
Dmit	<i>N,N'</i> -dimethylimidazolethione
EC	Electron correlation
ECP	Effective Core Potentials
EE	Electronic embedding
EER	Exchange enhanced reactivity

EgtB	2-histidyl- γ -glutamyl cysteine sulfoxide synthase
ESH	Ergothioneine
ET	Electron transfer
FF	Force field
GGA	Generalized Gradient approximation
ΔG_{corr}	Gibbs Free energy correction
GSH	Glutathione
GTO	Gaussian type orbitals
HF	Hartree-Fock
His	Histidyl
IC	Intermediate complex
Im	Imidazole
IEF	Integral-equation formalism
KIE	Kinetic isotope effect
LDA	Local density approximation
LA	Linoleic acid
LOX	Lipoxygenases
MD	Molecular Dynamics
ME	Mechanical embedding
MM	Molecular mechanics
MOE	Molecular Operating Environment
MSH	Methylthiol
MSR	Multistate reactivity
NAD^+	Nicotinamide adenine dinucleotide
OSH	Ovothiol
OvoA	5-histidylcysteine sulfoxide synthase
PA	Proton Affinity

PC	Product complex
PCM	Polarizable continuum model
PES	Potential energy surfaces
PP _i	Pyrophosphate
PPIX	Protoporphyrin IX
Por	Porphine
PCET	Proton-coupled electron transfer
QM/MM	Quantum Mechanics/Molecular Mechanics
RC	Reactive complex
RMSD	Root mean square deviation
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
SATP	Standard ambient temperature and pressure
SD	Spin density
SHE	Standard hydrogen electrode
SI	Spin inversion
sLO	Soybean-LOX
SP	Single point
TS	Transition state
UROD	Uroporphyrinogen decarboxylase
vdW	van der Waal
vac	Vacuum
ZPVE	Zero point vibrational energies

Chapter 1

Introduction



1.1 Introduction

Enzymes, proteins that have catalytic properties, are crucial to all organisms; they ensure many essential cellular chemical reactions occur at life-sustainable rates.^{1,2} For instance, the half-life for the hydrolysis of a glycosidic bond under standard conditions in solution is approximately five-million years.³ Yet glycoside hydrolases are capable of catalyzing this reaction at rates of 1000 times per second.³ Another quintessential example is Uroporphyrinogen decarboxylase which is capable of performing the quadruple decarboxylation of its substrate URO-III in mere seconds.⁴ Astonishingly, the half-life of this reaction in the absence of this enzyme is 2.3 billion years in solution under standard conditions.⁵

The proper functioning of proteins is dependent upon their structure. The latter is fundamentally governed by the sequence of their constituent amino acids (*i.e.* their primary structure). Consequently, any mutation to this sequence can result in the malfunction or inactivation of the respective enzyme. Every protein is encoded in an organisms DNA as genes. The overall process by which these genes lead to the synthesis of protein can be thought of as occurring in two main stages: transcription and translation (**Scheme 1.1**). In the first stage the DNA is transcribed into corresponding strands of RNA.⁶ It is this RNA that is then read in the next stage by the cellular machinery and “translated” into proteins themselves.



Scheme 1.1. The main stages in protein synthesis.

The aminoacyl-tRNA synthetase (aaRS) family of enzymes play a central role in translation. Specifically, they catalyze the activation and coupling of free amino acids to their cognate tRNA molecule (tRNA^{aa}) to form the corresponding aminoacyl-tRNA^{aa}

moiety (aa-tRNA^{aa}).⁶ Impressively, aaRSs catalyze their respective reactions with less than a 0.001% chance of error⁹ and as a result are commonly thought of as models of enzymatic specificity.^{7,8} It is this incredibly small chance of misacylation that helps prevent mutations in the primary structure of a protein. The aa-tRNA^{aa} then binds to the ribosome allowing for elongation of the protein being formed one residue at a time.

With translation (*i.e.* protein synthesis) complete, many enzymes are fully capable of performing their catalytic role. However, many require added groups called cofactors or coenzymes in order to exhibit their catalytic behaviour.¹⁰ The former are generally inorganic metal ions while the latter represent organic or metalloorganic molecules.¹¹ For example, some common coenzymes are nicotinamide adenine dinucleotide (NAD⁺), heme and tetrahydrobiopterin (see **Figure 1.1**). Many of these species' are essential for enzymes that utilize redox chemistry. In fact, the reduced nicotinamide containing NADH and NADPH are often involved in two electron redox reactions and have been said to be among the most vital coenzymes in all living systems.¹²

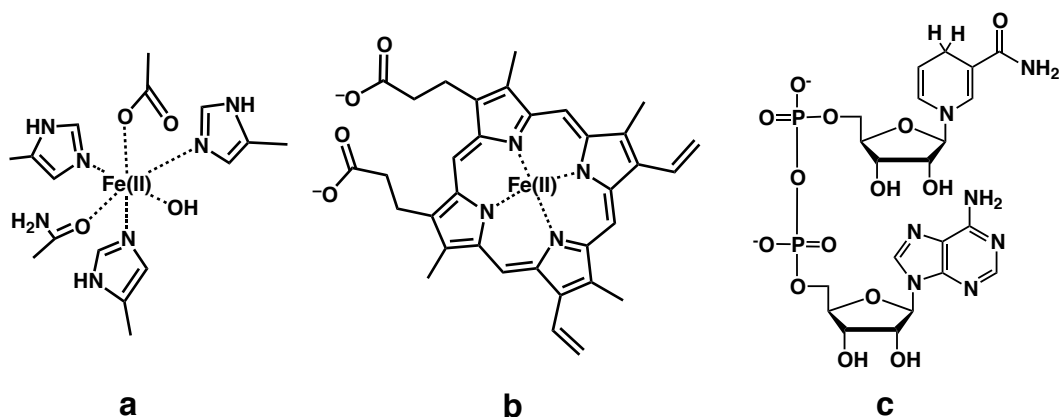


Figure 1.1. Schematic illustrations of the (a) non-heme Fe(II) cofactor of 8R-Lipoxygenase and the coenzymes (b) heme and (c) NADH.

Transition metals, due in large part to their redox properties, also have very important and unique roles as cofactors.^{13,14} Moreover, the possibility of multistate reactivity exists,

thus providing further means of enhancing chemical reactions.^{15,16} Indeed, such a phenomena may not only allow a particular reaction to happen faster but may in fact enable processes to occur that might otherwise not.¹⁷ Of the various metals available to biological systems it has been stated that iron is the most important, and in such cases it exists in two forms: heme and non-heme (**Figure 1.1**).¹⁸⁻²¹

Heme is essential for most organisms.^{26,29} Indeed, 75% of iron in the human body is contained within this cofactor.²⁹ Protoporphyrin IX (PPIX), the organic ring of heme (**Figure 1.1**), is the most complex and abundant macrocycle found in nature.²⁶ While its synthesis proceeds via several steps (**Figure 1.2**) the exact processes involved in its formation still remain uncertain.²⁶ PPIX is typically characterized as a strong field ligand that puts geometrical restrictions on the metal center.²² Furthermore, heme typically ligates to the protein via its iron center to a histidyl, cysteinyl or tyrosyl residue. The resulting heme-proteins have a diverse range of functions including O₂ transportation, signal transduction pathways, electron transport and redox catalysis.²³⁻²⁸

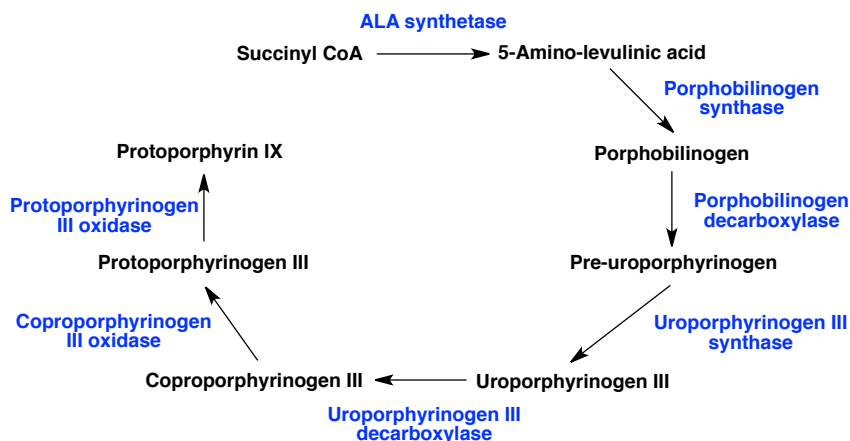


Figure 1.2. The key steps in the synthesis of Protoporphyrin IX.²⁶ Intermediates formed are represented in black while the enzymes that catalyze the respective transformations are in blue.

In contrast, in non-heme iron systems, weak field ligands generally coordinate the Fe. Additionally, such ligands are also usually less restrictive in terms of geometry, thus allowing for greater structural variation.²² Furthermore, the residues that coordinate the metal center show increased variability and may include, for example, aspartate, glutamate or inorganic sulfur. Non-heme enzymes have also been found to have diverse biological functions such as nitrogen fixation, photosynthesis and electron transfer.²¹

Regardless, of their differences, a very important reaction that both classes of enzymes catalyze is the activation of O₂ for mono- and di-oxygenation of substrates.¹⁰ This is due in part to the fact that the use of metals can have a dramatic impact on key oxo-intermediates by contributing to longer lifetimes and thus, enabling greater control in their respective reactivities.¹⁰ For instance, a common intermediate in O₂ activation by iron-containing metalloenzymes is a high-valent oxo-ferryl (FeIV=O) species (**Figure 1.3**). In heme systems, two such complexes may be formed and are referred to as compound I and II (Cpd I and II).³⁰ A key difference between them is that in the former PPIX exists as an oxidized cationic radical while in the latter it does not. While, Cpd I is generally the more reactive species, in non-heme iron systems only the Cpd II analogs exist.³⁰⁻³⁸ Regardless, this Fe(IV)=O group is generally considered to be the definitive biochemical oxidant.^{30,31,38-40} Many heme and non-heme iron enzymes simultaneously employ other coenzymes such as NAD⁺ and tetrahydrobiopterin to provide, for example, required additional electrons.¹⁰

Besides providing a means of catalyzing essential reactions, metals can also cause severe problems. In particular, free metal ions (*e.g.* Fe(II) and Cu(II)) in solution can create reactive oxygen species that may react with important biomolecules (*e.g.* DNA) resulting in their oxidative damage.⁴¹ Consequently, organisms have developed a number of processes for protecting against such damage and the agents that cause them. One common approach is the use of small biomolecules, antioxidants, that will complex

and/or react with these free metals and inhibit or prevent their undesirable effects. Thus, the importance of the roles of metal ions in biological systems cannot be overstated.

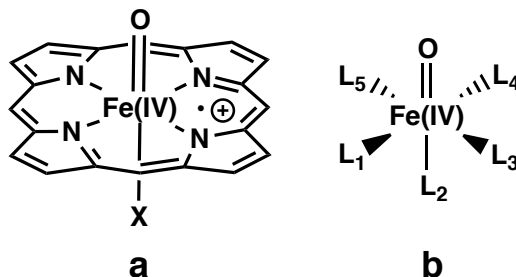


Figure 1.3. A schematic representation of (a) Cpd I (X = cysteinyl, histidyl or tyrosyl depending on enzyme family) and (b) non-heme iron Cpd II analog (L's may be ligating substrate(s) or active site residue(s)).

1.2 References

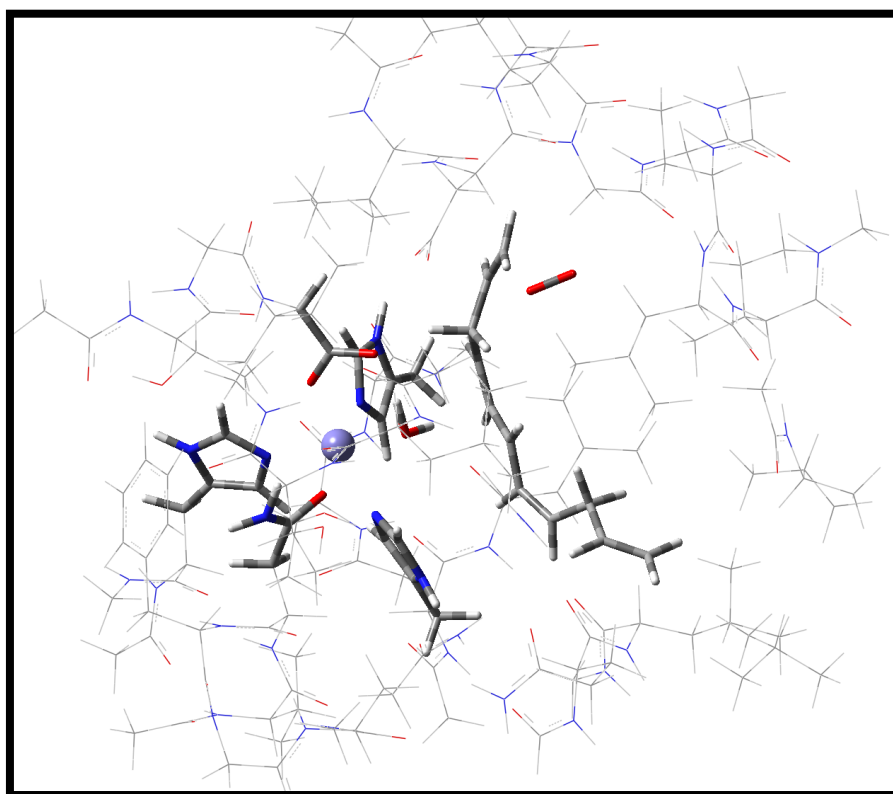
- (1) Liu, H. N.; Gault, J. W. *J. Phys. Chem. B* **2008**, *112*, 16874-16882.
- (2) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431-12441.
- (3) Huang, W. J.; Llano, J.; Gault, J. W. *J. Phys. Chem. B* **2010**, *114*, 11196-11206.
- (4) Bushnell, E. A. C.; Erdtman, E.; Llano, J.; Eriksson, L. A.; Gault, J. W. *J. Comput. Chem.* **2011**, *32*, 822-834.
- (5) Lewis, C. A., Wolfenden, R. *Proc. Nat. Acad. Sci. U.S.A.* **2008**, *105*, 17328-17333.
- (6) Creighton, T. E. *Proteins: Structures and Molecular Properties*; Second ed.; W.H. Freeman and Company: New York, 1993.
- (7) Cusack, S. *Curr. Opin. Struct. Biol.* **1997**, *7*, 881-889.
- (8) Hussain, T.; Kamarthapu, V.; Kruparani, S. P.; Deshmukh, M. V.; Sankaranarayanan, R. *Proc. Nat. Acad. Sci. U.S.A.* **2010**, *107*, 22117-22121.
- (9) Ibba, M.; Soll, D. *Annu. Rev. Biochem.* **2000**, *69*, 617-650.

- (10) Klinman, J. P. *Acc. Chem. Res.* **2007**, *40*, 325-333.
- (11) Nelson, D. L.; Cox, M. M. *Lehninger: Principles of Biochemistry*; 4th ed.; W.H. Freeman and Company: New York, 2005.
- (12) Carugo, O.; Argos, P. *Proteins* **1997**, *28*, 29-40.
- (13) Hughes, T. F.; Friesner, R. A. *J. Chem. Theory Comput.* **2011**, *7*, 19-32.
- (14) Hughes, T. F.; Friesner, R. A. *J. Chem. Theory Comput.* **2012**, *8*, 442-459.
- (15) Poli, R.; Harvey, J. N. *Chem. Soc. Rev.* **2003**, *32*, 1-8.
- (16) Shaik, S.; Hirao, H.; Kumar, D. *Acc. Chem. Res.* **2007**, *40*, 532-542.
- (17) Poli, R.; Harvey, J. N. *Chem. Soc. Rev.* **2003**, *32*, 1-8.
- (18) Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L. *Chem. Rev.* **2004**, *104*, 939-986.
- (19) de Visser, S. P. *Angew. Chem. Int. Ed.* **2006**, *45*, 1790-1793.
- (20) Neidig, M. L.; Solomon, E. I. *Chem. Commun.* **2005**, 5843-5863.
- (21) Greenwood, N. N., Earnshaw, A. *Chemistry of the Elements*; 2nd ed.; Elsevier: New York, 1997.
- (22) Liu, H. N.; Llano, J.; Gauld, J. W. *J. Phys. Chem. B* **2009**, *113*, 4887-4898.
- (23) Battistuzzi, G.; Bellei, M.; Bortolotti, C. A.; Sola, M. *Arch. Biochem. Biophys.* **2010**, *500*, 21-36.
- (24) Dawson, J. H. *Science* **1988**, *240*, 433-439.
- (25) Fita, I.; Rossmann, M. G. *J. Mol. Biol.* **1985**, *185*, 21-37.
- (26) Heinemann, I. U.; Jahn, M.; Jahn, D. *Arch. Biochem. Biophys.* **2008**, *474*, 238-251.
- (27) Meunier, B.; de Visser, S. P.; Shaik, S. *Chem. Rev.* **2004**, *104*, 3947-3980.
- (28) Rodgers, K. R. *Curr. Opin. Chem. Biol.* **1999**, *3*, 158-167.
- (29) Guallar, V.; Wallrapp, F. H. *Biophys. Chem.* **2010**, *149*, 1-11.
- (30) Bollinger, J. M.; Krebs, C. *Curr. Opin. Chem. Biol.* **2007**, *11*, 151-158.
- (31) Bruijninx, P. C. A.; van Koten, G.; Gebbink, R. *Chem. Soc. Rev.* **2008**, *37*, 2716-2744.

- (32) Chen, H.; Cho, K.-B.; Lai, W.; Nam, W.; Shaik, S. *J. Chem. Theory Comput.* **2012**, *8*, 915-926.
- (33) Galonić Fujimori, D.; Barr, E. W.; Matthews, M. L.; Koch, G. M.; Yonce, J. R.; Walsh, C. T.; Bollinger, J. M.; Krebs, C.; Riggs-Gelasco, P. J. *J. Am. Chem. Soc.* **2007**, *129*, 13408-13409.
- (34) Karlsson, A.; Parales, J. V.; Parales, R. E.; Cibson, D. T.; Eklund, H.; Ramaswamy, S. *Science* **2003**, *299*, 1039.
- (35) Krebs, C.; Galonić Fujimori, D.; Walsh, C. T.; Bollinger, J. M. *Acc. Chem. Res.* **2007**, *40*, 484-492.
- (36) Matthews, M. L.; Krest, C. M.; Barr, E. W.; Vaillancourt, F. H.; Walsh, C. T.; Green, M. T.; Krebs, C.; Bollinger, J. M. *Biochemistry* **2009**, *48*, 4331-4343.
- (37) Mbughuni, M. M.; Chakrabarti, M.; Hayden, J. A.; Bominaar, E. L.; Hendrich, M. P.; Munck, E.; Lipscomb, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16788-16793.
- (38) Nam, W. *Acc. Chem. Res.* **2007**, *40*, 465-465.
- (39) Cho, K.-B.; Chen, H.; Janardanan, D.; de Visser, S. P.; Shaik, S.; Nam, W. *Chem. Commun.* **2012**, *48*, 2189-2191.
- (40) van der Donk, W. A.; Krebs, C.; Bollinger, J. M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 673-683.
- (41) Bushnell, E. A. C.; Llano, J.; Eriksson, L. A.; Gauld, J. W. Mechanisms of Mutagenic DNA Nucleobase Damages and Their Chemical and Enzymatic Repairs Investigated by Quantum Chemical Methods, In *Selected Topics in DNA Repair*; Chen, P. C., Ed.; InTech: 2011.

Chapter 2

Theoretical Methods



2.1 Introduction

Computational chemistry is the application of the equations of quantum chemistry to the study of chemical systems via the use of computers. Ever increasing computational power coupled with algorithmic developments has enabled this approach to provide insights into most if not all fields of chemistry. In particular, there are now numerous computational methods currently available including density functional theory, quantum mechanics/molecular mechanics and molecular dynamics. Individually or in combination with each other, these allow highly accurate and reliable investigations to be performed on chemical systems ranging from an atom or molecule consisting of a just a few atoms to those composed of tens of thousands of atoms, *e.g.* solvated proteins. Importantly these methods allow for the investigation of stable and long-lived species as well as short-lived or highly reactive species that in some cases cannot be experimentally observed. This is particularly true of transition metal-containing species in which, for example, complex metal...ligand interactions can lead to geometrical and electronic structures that are markedly more highly varied compared to main-group compounds.¹

Within this thesis a variety of computational chemistry methods have been used to, for instance, gain new and detailed insights into the catalytic mechanisms of several metallo-enzymes and the properties of various metal-ligand complexes. Notably, the fundamental theories underpinning the methods used have been discussed in detail in numerous textbooks.²⁻⁵ Thus, in this chapter only a brief synopsis of the central theorems and methods utilized in the subsequent various chapters of this thesis are described.

2.2 Schrödinger Equation

A fundamental equation of quantum chemistry is the time-dependent Schrödinger equation (eq. **2.1**). This equation relates the change in a system defined by its wavefunction with respect to time. That it is defines how the system evolves in time.

$$-\frac{i}{\hbar} \frac{\partial \Psi(r,t)}{\partial t} = \hat{H}\Psi(r,t) \quad (\text{eq. 2.1})$$

\hat{H} is the Hamiltonian operator and $\Psi(r,t)$ is the state wavefunction that describes the particular system and depends on the positions of the particles of interest and time. It is postulated that the wavefunction contains all information about the system.

However, many of the systems of interest do not typically change with respect to time. Thus, through separation of variables we are able to express the above equation as the time-independent Schrödinger equation (eq. 2.2).³

$$\hat{H}\Psi(r) = E\Psi(r) \quad (\text{eq. 2.2})$$

In this form the wavefunction is now solely a function of the positions of the particles and the Hamiltonian operator also no longer depends on time. In particular, for molecules the field-free non-relativistic Hamiltonian \hat{H} can be written in atomic units as:

$$\hat{H} = -\frac{1}{2} \sum_i^N \nabla_i^2 - \frac{1}{2} \sum_a^N \frac{\nabla_a^2}{M_a} + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_a^N \sum_{b>a}^N \frac{Z_a Z_b}{r_{ab}} - \sum_i^N \sum_a^N \frac{Z_a}{r_{ia}} \quad (\text{eq. 2.3})$$

In equation 2.3 the first and second terms represent the kinetic energies of the electrons and nuclei, respectively. The third, fourth and fifth terms represent the electron-electron, nuclei-nuclei and electron-nuclei electrostatic interaction energy terms, respectively.

Unfortunately, however, a problem exists in that the above equation 2.2 is unsolvable for all but the simplest of systems.³ Thus, approximations must be made in order to be able to practically apply and solve this equation for most chemical systems of interest.

2.3 The Born-Oppenheimer Approximation

Due to their significantly larger masses, nuclei move considerably more slowly than electrons. Consequently, to a reasonable approximation, one can treat electrons as moving in a field of fixed nuclei, *i.e.* they are stationary.³ This is known as the Born-Oppenheimer (BO) Approximation and aids in simplifying both the molecular Hamiltonian and wavefunction. Specifically, the kinetic energy of the nuclei, the second

term in eq. 2.3, is now zero. Concomitantly, the nuclear-nuclear interaction term, the fourth term in eq. 2.3, is now a constant (V_{NN}) for any given nuclear configuration.² The resulting Hamiltonian can now be written as:

$$\hat{H} = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} - \sum_i^N \sum_a^N \frac{Z_a}{r_{ia}} + V_{NN} \quad (\text{eq. 2.4})$$

Notably, the first three terms in eq. 2.4 are commonly collectively referred to as the electronic Hamiltonian:

$$\hat{H}_{el} = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} - \sum_i^N \sum_a^N \frac{Z_a}{r_{ia}} \quad (\text{eq. 2.5})$$

Thus, the resulting Schrödinger equation can now be represented as:

$$(\hat{H}_{el} + V_{NN})\Psi_{el}(\mathbf{r}_i; \mathbf{r}_N) = (E_{el} + V_{NN})\Psi_{el}(\mathbf{r}_i; \mathbf{r}_N) \quad (\text{eq. 2.6})$$

where the $\Psi_{el}(\mathbf{r}_i; \mathbf{r}_N)$ is known as the electronic wavefunction and is parameterically dependent on the positions of the nuclei. That is, for each position of the nuclei, one will obtain a new set of electronic wavefunctions.

2.4 Orbital Approximation

However, even with the BO approximation the electronic Hamiltonian still remains unsolvable for all but one-electron systems, *e.g.* H and He⁺. This is due to the fact that while the kinetic energy term and the electron-nuclei potential energy term are both one-electron terms, the electron-electron repulsion term depends on two electrons. However, to a reasonable approximation it can be assumed that each electron in a N-electron system moves independent of each another (*i.e.* no interaction) and thus, each can be assigned its own one-electron orbital. That is, the two electron electron-electron potential energy term (second term in eq. 2.5) is effectively neglected. This enables the N-electron electronic Hamiltonian to be written as a linear combination of N one-electron electronic Hamiltonians:

$$\hat{H}_{el} = \left(-\frac{1}{2}\nabla_1^2 - \sum_a^N \frac{Z_a}{r_{1a}}\right) + \left(-\frac{1}{2}\nabla_2^2 - \sum_a^N \frac{Z_a}{r_{2a}}\right) + \dots + \left(-\frac{1}{2}\nabla_n^2 - \sum_a^N \frac{Z_a}{r_{na}}\right) \quad (\text{eq. 2.7})$$

Furthermore, this allows the N-electron wavefunction to be written as a product of N one-electron functions:

$$\Psi_{el} = \prod_i^N (\psi_{el})_i \quad (\text{eq. 2.8})$$

However, an issue exists in that the resulting product in eq. 2.8 is no longer a proper wavefunction given that it fails to obey the Pauli exclusion principle: an electronic wavefunction must be anti-symmetric with respect to interchange of two electrons.³ Fortunately, Slater³ showed that an anti-symmetric wavefunction is obtained if it is represented in the form of a determinant, as shown below:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(e_1) & \dots & \psi_N(e_1) \\ \vdots & \ddots & \vdots \\ \psi_1(e_N) & \dots & \psi_N(e_N) \end{vmatrix} \quad (\text{eq. 2.9})$$

where $\psi_i(e_j)$ represents the i^{th} MO for the j^{th} electron, and the factor $1/\sqrt{N!}$ ensures that the wavefunction is normalized.

2.5 Basis Set Expansion

While expressing the wavefunction as one-electron orbitals represents a significant step towards solving the Schrödinger equation, further approximations must be made. In particular, we do not know the exact form of the one-electron functions used to express the wavefunction. Roothaan, however, proposed that each such orbital be expressed as a linear combination of known functions (basis functions).³ Collectively, the set of these functions is referred to as a basis set. In theory an infinite number of functions, a complete set, should be used but in practice only a limited number of functions can be used.³ An example of a MO expressed as a sum of basis functions is given below in eq. 2.10:

$$\psi_i = \sum_{\mu=1}^K c_{\mu i} \phi_{\mu i} \quad (\text{eq. 2.10})$$

where $c_{\mu i}$ is the coefficient of the μ^{th} basis function (*i.e.* $\phi_{\mu i}$) in the linear expansion of the i^{th} MO (*i.e.* ψ_i). If the set of basis functions represents atomic orbitals then eq. 2.10 is referred to as a *linear combination of atomic orbitals* (LCAOs).

2.6 Basis Sets

Basis sets are commonly composed of Gaussian type orbitals (GTOs). This is because not only are the integrals involved in the various calculations computationally more tractable and cheaper but notably, the product of two GTOs is another centered between the two original functions. While there is a quite diverse variety of basis sets available, they in general share several common features which are briefly discussed below.

Split-Valence. In general, chemical reactions and interactions between atoms or molecules involve the valence electrons. In split-valence basis sets the valence orbitals are described by two or more basis functions which are then combined to describe the particular valence orbital. This approach allows one to better account for the fact that orbitals in molecules may expand or contract depending on their environment.² If two or three basis functions are used to describe each valence orbital the resulting basis set is referred to as double-zeta (*e.g.* 6-31G) or triple-zeta (*e.g.* 6-311G), respectively.

Polarization Functions. These are functions of higher angular momentum than the occupied valence orbitals that can be included in the basis set. For example, *p*- and *d*-functions are common polarization functions added onto hydrogens, while *d*- and *f*-functions are commonly added to heavy atoms such as carbon. Such functions enable one to better describe the polarization of electrons in an orbital by an environment.² Circumstances in which such functions are often required is for the reasonable predictions of the geometries of species' containing stretched bonds, *e.g.* transition states, or those involving second row elements and hypervalent bonding situations.² The 6-

31G(d,p) basis set is an example of a common double-zeta basis set that contains *d*-polarization functions on heavy atoms and *p*-polarization functions on hydrogens.

Diffuse Functions. In many chemical systems long-range interactions are important to their structure and properties, or they contain electrons that are more distant from the nuclei (*e.g.* excited states and anions).² In order to better describe the orbitals involved in such cases one can add diffuse functions; functions that contain a smaller exponential coefficient and hence are spatially much larger. The 6-31+G basis set is an example of a split-valence double-zeta basis set which includes diffuse functions on only heavy atoms (*i.e.* non-hydrogen atoms). Such functions correspond to the addition of orbitals that have the same angular momentum quantum number as those of the valence orbitals.

Effective Core Potentials (ECP). As noted above, many chemical properties of interest, *e.g.* ionization energies, are determined by the valence electrons. However, as the size of an atom increases so does the number of core electrons. Hence, an increasing fraction of the cost of a calculation is spent dealing with core electrons. However, it is possible to replace the core electrons with an effective core potential (ECP) and thus, substantially reduce the cost of calculations. In addition, one can more easily take into account the fact that for very heavy atoms relativistic effects become significant via the use of relativistic-corrected ECPs. Notably, ECP-containing basis sets, *e.g.* LANL2DZ, are often used in the study of transition metal-containing systems.

Dispersion correcting potentials (DCP). In some systems such as protein-ligand complexes, van der Waal (vdW) interactions can be important.⁶ These interactions can be challenging to describe using standard basis sets.⁷ However, it is possible to correct for this shortcoming via the use of dispersion correcting potentials. In particular, two functions are added to each carbon atom within the system being considered where one is attractive while the other is repulsive.⁸ Together they account for the dispersive type interactions that may exist between non-covalently bound molecules. Importantly, this

approach has been shown to greatly improve the reliability of describing weakly interacting systems.⁸ Further details can be found in the work of Mackie and DiLabio^{8,9}

2.7 Variational Theorem

Given an approximate way to represent the wavefunction (Slater determinant; Ψ_{Approx}) of a chemical system, the variational theorem provides one with a means of evaluating the accuracy of the energy obtained. As shown in equation 2.11, this theorem states that for an approximate wavefunction that obeys the same boundary conditions of the exact wavefunction, the energy obtained (E_{Approx}) will always be higher than the exact non-relativistic, time-independent energy (E_{Exact}).

$$E_{\text{Exact}} \leq E_{\text{Approx}} = \frac{\langle \Psi_{\text{Approx}}^* | \hat{H} | \Psi_{\text{Approx}} \rangle}{\langle \Psi_{\text{Approx}}^* | \Psi_{\text{Approx}} \rangle} \quad (\text{eq. 2.11})$$

2.8 Hartree-Fock (HF) Theory

In the 1920's Douglas Hartree proposed a means to solve the Schrödinger equation. Given that the instantaneous interaction between electrons cannot be calculated he suggested to treat them in an average way.³ In particular, the electron-electron interaction term (*i.e.* $\sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}}$) was written as:

$$\sum_i^N \sum_{j>i}^N \int \frac{\rho_j dv_j}{r_{ij}} \quad (\text{eq. 2.12})$$

where $\rho_j dv_j$ represents the averaged density for electron j. The assumption that an electron sees all of the other electrons in an average way speeds up the calculations but affects the accuracy of the calculations.² In the original proposal a product of one-electron orbitals was used as the trial wavefunction (eq. 2.8). However, as noted above, such an approach does not satisfy the requirement that an electronic wavefunction be antisymmetric. Thus, in the 1930's Fock proposed that a Slater determinant (eq. 2.9) be used instead, thus the birth of HF theory.

With the Hartree-Fock operator and initial trial wavefunction now chosen the next step is to obtain the energy of the system. A peculiarity of the HF operator is that it depends on its own eigenfunctions by way of eq. **2.12**, hence the energy must be obtained via an iterative process. For a system of N electrons it begins with the determination of an improved orbital for electron 1 within an averaged field of the remaining $N-1$ electrons generated from an initial guess of their orbitals. Then, an improved orbital for electron 2 is obtained within an averaged field generated from the improved orbital of electron 1 and the initial guessed orbitals of the remaining electrons 3 to N . This then continues for the remaining electrons. Once improved orbitals have been obtained for all N electrons one then returns to electron 1 and obtains a new improved orbital within an averaged field generated from the improved orbitals of the other $N-1$ electrons and so on. This iterative process continues until the improved orbitals obtained in one entire cycle are obtained in the subsequent cycle. As a result, this is also known as the self-consistent field method. One can now construct an optimized wavefunction and the energy of the system can then be calculated in a straightforward manner.

However, by treating the electron-electron interactions in an average way the instantaneous correlation between electrons has been neglected. The difference between the exact energy and the Hartree-Fock energy represents the correlation energy.

$$E_{Exact} - E_{HF} = E_{Corr} \quad (\text{eq. 2.13})$$

Importantly, the correlation energy is always negative, thus, it is a stabilizing effect.

2.9 Electron Correlation

When discussing electron correlation (EC) we generally describe it as one of two forms: dynamic and static. Dynamic EC represents the short-range effects caused by the repulsion felt by electrons.

Static EC is very system dependant and can have medium to long-range effects. It exists when a single Slater determinant is not sufficient to describe the chemical system. For example, ozone is best described by a series of resonance structures and thus its wavefunction is better represented by a linear combination of degenerate Slater determinants. Systems with significant static EC are typically called multi-reference systems.¹⁰

Møller-Plesset Perturbation Theory. A common conventional wavefunction-based method used to account for the missing dynamic EC is based on perturbation theory.^{2,3,11} In summary, the method works by promoting electrons from occupied to virtual MOs. That is, it includes excited states into the Hartree-Fock wavefunction. By effectively allowing electrons to move away from and thus avoid each other, it recovers at least in part the missing dynamic EC.

2.10 Density Functional Theory (DFT)

The central approach in DFT is that the energy of the system can be calculated from the electron density of the system and can be mathematically expressed as shown in eq. 2.14. Notably, unlike wavefunction-based methods the central quantity on which DFT methods are based, the density is experimentally observable.

$$E_0 = E[\rho(x, y, z)] \quad (\text{eq. 2.14})$$

E_0 is the ground state energy, $E[...]$ is the energy functional and $\rho(x, y, z)$ is the density function. Unlike the wavefunction which depends on the spin and three spatial coordinates of each electron, the electron density is a function of only three variables (x, y, z). The connection between the wavefunction and density function is:¹²

$$\rho(x, y, z) = \int \int \dots \int |\Psi|^2 ds_1 d\tau_2 \dots d\tau_n \quad (\text{eq. 2.15})$$

The first Hohenberg-Kohn theorem states that the ground-state electron density uniquely determines the Hamiltonian and wavefunction.¹² Consequently, it provides a basis for the calculation of the ground state energy and properties of the system from the density. Unfortunately, however, it makes no mention of the form of the energy functional.¹² Similarly, we do not know the exact form of the density function. However, the Kohn-Sham theorem in part assumes that the electrons are non-interacting. As a result, the total density can be approximated by a linear combination of one-electron densities constructed from Kohn-Sham orbitals:

$$\rho(x, y, z) = \sum_i^N |\Phi_i^{KS}|^2 \quad (\text{eq. 2.16})$$

where Φ_i^{KS} represents the i^{th} Kohn-Sham orbital. Importantly, this approximation allows the energy functional to be expressed as:

$$E_o[\rho] = T_{cl}[\rho] + V_{cl}[\rho] + E_{XC}[\rho] \quad (\text{eq. 2.17})$$

where $T_{cl}[\rho]$ represents the exact kinetic energy of the non-interacting system; $V_{cl}[\rho]$ corresponds to the potential energy for the nuclei-electron and electron-electron interactions; and $E_{XC}[\rho]$ represents the exchange-correlation energy. While the first two terms are known exactly the last term, the exchange-correlation functional, is unknown. Notably, it is this term that corrects for the difference between the kinetic energy of the system of non-interacting electrons and the real system. Furthermore, it also corrects for the self-interaction error in $V_{cl}[\rho]$ as well as the neglect of exchange between parallel spin electrons. It is noted that for all current DFT functionals it is their expression used for $E_{XC}[\dots]$ in which they differ. As described by Perdew and Schmidt,¹³ a *ladder* of approximations for the exchange-correlation energy as a functional of the electron density can be constructed (**Figure 2.1**). At the lowest rung of this *ladder* the exchange-correlation energy is only dependent on the local density at a particular point; the local density approximation (LDA). With the movement to higher *rungs* we see increasingly more complex components included into the exchange-correlation functional. For

instance, the exchange-correlation energy may also depend on the gradients (GGA) or second derivatives (Meta-GGA) of the electron density. The details of the various functional types can be found in several recent works.^{2,7,11-15}

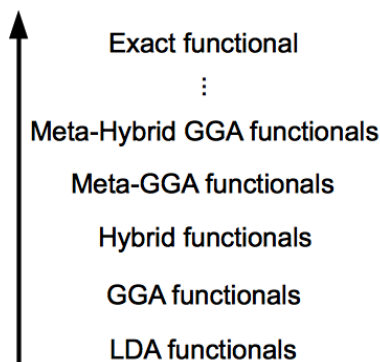


Figure 2.1. Schematic representation of a ladder of increasingly more complex DFT functionals.¹³

The $E_{XC}[\dots]$ component represents a key feature of DFT methods in that they inherently include electron correlation, even if it is only approximate because we do not know its exact form. In fact, a central difference between HF and DFT is that the former is a deliberately approximate theory whose development was in part motivated by an ability to solve the relevant equations exactly. In contrast, the latter is an exact theory but the relevant equations must be solved approximately because a key operator has an unknown form.¹²

Because of this, DFT methods are in general the most cost effective to use in order to achieve a given level of accuracy.¹² In fact, they are the methods of choice in the investigation of biochemical systems and more importantly in the study of those containing transition metals.^{12,16-18} In particular, the B3LYP functional (shown in eq. **2.18**) has proven to be one of the most dependable for such studies.¹⁷⁻²⁴ For instance, it has been shown to have average errors of 0.018 Å in metal-ligand (M–L) bond lengths

for complexes containing 3rd-row transition metals,²⁵ Meanwhile, for relative energies it has been shown to have errors of $\sim 21.0 \text{ kJ mol}^{-1}$.¹⁸

$$E_{XC}^{B3LYP} = aE_X^{B88} + bE_X^{LDA} + (1 - b)E_X^{HF} + cE_C^{LYP} + (1 - c)E_C^{LDA} \quad (\text{eq. 2.18})$$

It is a combination of LDA and GGA functionals as well as the exchange operator from HF theory. Due to the inclusion of the latter it is also known as a hybrid DFT method. The coefficients a , b and c were obtained via a best fit to experimental atomization energies, ionization energies and proton affinities.¹²

2.11 Molecular Mechanics

In molecular mechanics (MM), the electrons of the system are ignored; only the nuclei of the system are considered when calculating the energy of the system.²⁶ While, such methods cannot describe bond breaking and forming processes it does allow for the investigation of very large systems compared to those possible via the HF and DFT approaches discussed above.²⁷ Notably, the ability to define the energy of the system as a function of solely nuclei is actually a consequence of the BO approximation.²⁶

The only interactions considered in MM are those that exist between nuclei and these can be generally classified as bonded or non-bonded. In the case of bonded interactions we have terms to account for the bonds, angles and torsional angles while for non-bonded interactions we have terms to account for the vdW and electrostatic interactions. Consequently, the MM energy (E^{MM}) can mathematically be represented as in eq. 2.19:

$$E^{MM} = \sum E_B + \sum E_A + \sum E_{TA} + \sum E_{vdW} + \sum E_C \quad (\text{eq. 2.19})$$

The first two terms on the right, $\sum E_B$ and $\sum E_A$, represent the energy of the system with respect to the bonds and angles respectively, and are commonly written as a simple harmonic expression.²⁸ The third term represents the torsional energy ($\sum E_{TA}$) of the system.²⁸ Finally, the last two terms, $\sum E_{vdW}$ and $\sum E_C$, represent the non-bonded vdW

and electrostatic interactions, respectively. In general, the vdW interaction energies are expressed as a 6-12 Lennard-Jones potential while the electrostatic interactions are calculated via Coulomb's law.²⁸

It is noted that MM methods are empirical. Thus, each of the various methods available have been developed to study different chemical systems such as organic molecules, complex liquids or proteins.²⁸ What separates the various MM methods is the exact form of the energy expression and parameters used within it. Notably, the combination of the mathematical expression and all parameters together form the particular MM force field (FF). For instance, AMBER is a commonly used FF which was developed to model conformations and intermolecular interactions accurately for proteins, nucleic acids and other biomolecules.²⁸ For AMBER its mathematical form is:

$$\begin{aligned}
 E^{AMBER} = & \sum_{bonds} k_r (r - r_o)^2 + \sum_{angles} k_\theta (\theta - \theta_o)^2 \\
 & + \sum_{torsions} \frac{V_n}{2} [1 + \cos(n\phi - \omega)] \\
 & + \sum_i \sum_{j>i} \left[\frac{A_{ij}}{(r_{ij})^{12}} - \frac{B_{ij}}{(r_{ij})^6} \right] + \sum_i \sum_{j>i} \left[\frac{q_i q_j}{(r_{ij})} \right] \quad (\text{eq. 2.20})
 \end{aligned}$$

2.12 Quantum Mechanics/Molecular Mechanics (QM/MM)

In 1976 Warshel and Levitt introduced the concept of QM/MM.²⁹ This method aims to combine the advantages and strengths of the faster MM methods with those of the slower QM methods. In particular, the region in which bond making and breaking is treated at the high QM level of theory while the remaining part (*i.e.* the environment) is treated at the lower MM level of theory (**Figure 2.2**).²⁷

The advantage of such an approach is that with the use of only a DFT or HF approach, system sizes are typically limited to 200-400 atoms. However, in the case of QM/MM models, chemical model sizes of 20000-30000 atoms can be investigated.³⁰ It is noted that in general only a 1000 atoms or so are left free to move while the remainder are held

fixed.³⁰ Consequently, QM/MM methods have been increasingly applied to the study of, for example, enzymatic reactions and previously reviewed; see, for example, Senn and Thiel²⁹ Llano and Gauld³¹ and Sousa et al.³²

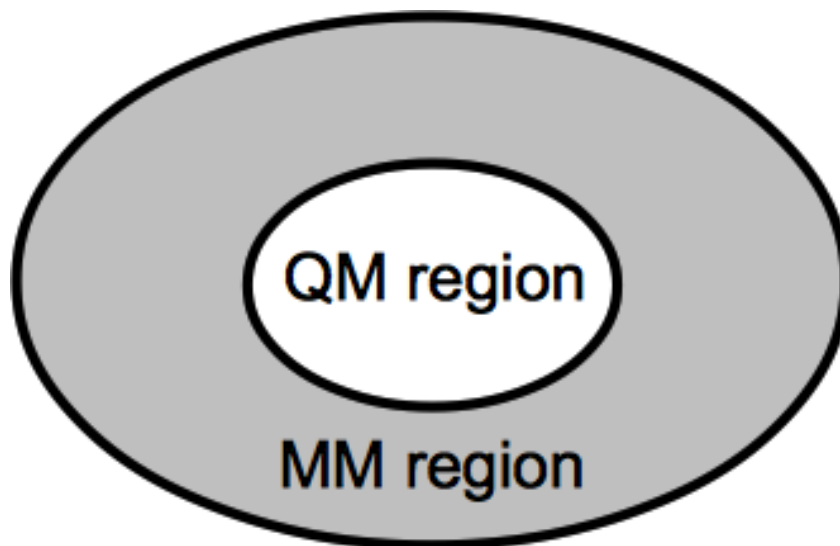


Figure 2.2. A schematic representation of a QM/MM model in which the surrounding environment is modeled at the MM level of theory while the reacting center is at the QM level of theory.

Of the various QM/MM formalisms available the one used within this thesis is the ONIOM approach. It is noted that in the calculation of the ONIOM energy a mechanical or electronic embedding scheme can be used. For the former, all interactions between layers are treated at the MM level of theory. In the latter, the electronic interactions are treated at the QM level of theory, while the remaining interactions remain modeled at the MM level of theory. Importantly, such a scheme allows the points charges within the outer MM layer to affect (*e.g.* polarize) the chemical environment of the inner QM layer.

It should be noted, however, that QM/MM-based approaches can suffer from the same limitations as either a solely DFT, HF, or MM approach; the dynamic behaviour of the enzyme is not fully taken into account.^{33,34}

2.13 Molecular Dynamics

In the methods discussed above we generally work with static systems. In contrast, in Molecular Dynamics (MD) simulations by integrating Newton's laws of motion one can generate sequential configurations of a system as it changes with respect to time.²⁶

Due to the incredibly large number of calculations needed, MD is an empirical based method like MM. Thus, by using a particular forcefield the forces acting on the system are calculated which are then used to determine the sequential configurations with respect to time.²⁶ However, as the equations of motion involved in MD calculations cannot be analytically solved, such methods are instead based on a finite difference model.²⁶ Hence, the equations are solved for finite time steps where the forces acting on the atoms at a particular time are assumed constant. Then, using the equations of motion, we calculate where the system will be after this finite time step. The forces are then recalculated given the new positions of the atoms.

Thus, with the use of MD methods we are able to investigate how chemical systems (ranging from a single small molecule consisting of a just a few atoms to those composed of tens of thousands of atoms) change with respect to time. However, in general, these methods are used for the simulation of very large molecular systems that have many degrees of freedom.²⁶ For instance, MD methods have been successfully applied to the study substrate/inhibitor–protein interactions, protein-protein interactions as well as many other aspects.³⁵⁻³⁹

It is noted that because MD methods are an empirical based method like MM we are unable to investigate chemical reactions that involve the breaking and forming of bonds. However, a QM/MM type approach can be used where MD methods are used to allow the environment to change with respect to time. Such a method allows us to determine a statistical average of possible alternative pathways of an enzymatic reaction.^{2,5,40,41} In

particular, from these simulations the equilibrium averages of the MD trajectories can then be used to calculate the free energies that accompany the chemical changes.

2.14 Solvation

When examining biochemical systems it is often important to also consider the protein environment. As noted above, we can use a QM/MM approach and explicitly model the steric and electronic effects of the surrounding often non-homogeneous environment. However, this can be time-consuming and challenging or even unnecessary. Alternatively, however, one can use an implicit solvation model such as a polarizable continuum model (PCM) approach.⁴²

The implicit solvation model used in this thesis is the integral-equation formalism polarizable continuum model (IEF-PCM).⁴³⁻⁴⁶ It puts interlocking cavities around each atom of the solute. The charge distribution of the solute polarizes the dielectric continuum which in turn polarizes the solute charge distribution.⁴⁷ While a PCM-based method is simpler than using a QM/MM method, only the general polarity effects of the surrounding solvent/environment are modeled. That is, explicit solvent-solute interactions (*e.g.* hydrogen bonds) are not modeled and the polarity of the surrounding environment is treated as being homogeneous. It is noted in some cases where it is essential to include explicit hydrogen bonds with the solvent we must include these necessary groups within the model and treat the remaining environment using a PCM approach.⁴²

In those cases where a PCM approach is used to model a protein environment it is common to use a dielectric constant of 4.0. This represents a compromise between the value of 3 for the core of a protein and 80 for water.⁴²

2.15 Computational Tools

As highlighted above, there are a variety of tools available to the computational chemist. Most, however, are applied in order to obtain optimized structures, harmonic vibrational frequencies, relative energies, and the generation of potential energy surfaces (PESs). The latter can give invaluable insight into, for example, enzymatic mechanisms such as barrier heights.

Geometry Optimizations. In chemistry we are often interested in the nature and properties of reactants, products, intermediates and transition states of a chemical reaction. Importantly, such points on a PES are stationary points. Mathematically this refers to the fact that their first derivatives of the energy with respect to the nuclear coordinates are zero as shown in eq. 2.20:

$$\frac{\partial E}{\partial r_1}, \frac{\partial E}{\partial r_2}, \frac{\partial E}{\partial r_3}, \dots, \frac{\partial E}{\partial r_N} = 0 \quad (\text{eq. 2.20})$$

By determining the conditions at which such derivatives are zero provides us with the various minima (*e.g.* reactants, products and intermediates) and maxima (*e.g.* transition states) along the PES. To differentiate between these two types of critical points, however, one requires the calculation of the second derivatives.

Frequency Calculations. The second derivatives of the energy with respect to the coordinates of the nuclei are obtained from harmonic vibrational frequency calculations. As noted above, these values are important for determining the nature of the various stationary points along a PES. In particular, if:

- (i) the second derivatives are all positive then we have a minimum;
- (ii) all but one is positive then we have a maximum along a single reaction coordinate (*i.e.* a first-order saddle point or transition state).

For the latter, the one negative value corresponds to an imaginary frequency and represents the mode of vibration that leads from one minimum to another.

In addition, frequency calculations also provide valuable thermochemical information such as zero point vibrational energies (ZPVE) and Gibb's Free Energy corrections. Such energies are ignored in the calculation of the electronic energy since the nuclei are assumed to be fixed in space. However, from QM we know that even if the vibrational quantum number is zero the system (if non-linear) will have a vibrational energy of:

$$ZPVE = \sum_i^{3N-6} \left(\frac{1}{2} + v_i\right) h\nu_i \quad (\text{eq. 2.21})$$

where N is the number of atoms within the system, v_i is the vibrational quantum number and ν_i is frequency of the vibration.

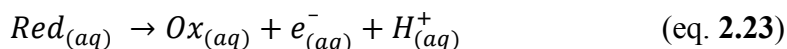
The inclusion of these corrections when calculating energies enables one to calculate, for example, the energy of a system at 0 K (the ZPVE correction) or their free energies at 298 K.

Single Point Calculations. It is generally accepted that reliable geometries can be obtained using a moderately sized basis set. However, this is generally not the case when calculating relative energies.^{2,11} Thus, a common approach to calculating such energies is to use the geometries obtained at a lower level of theory and perform single point calculations at a higher level of theory and is represented as:

$$\text{Method}_A/\text{Basis Set}_A // \text{Method}_B/\text{Basis Set}_B \quad (\text{eq. 2.22})$$

where $\text{Method}_A/\text{Basis Set}_A$ is the level of theory used for the single point energy calculation, and $\text{Method}_B/\text{Basis Set}_B$ is the level of theory used to obtain the optimized geometry.

Redox Potentials. In the calculation of redox potentials a common approach is to treat protons and electrons as independent ions as shown in the following half reaction:



Unfortunately, the energies of the electron and proton are not obtainable using standard computational tools. Instead one can use the chemical potential of an electron or

proton under various conditions obtained by means of a first principles quantum and statistical mechanics approach, the details of which can be found in the work of Llano and Eriksson.⁴⁸ These energies are summarized below in **Table 2.1**.

Table 2.1. Chemical potentials of an electron and proton in various reference states.⁴⁸

Quantity	eV	kJ mol ⁻¹
$\mu_{0K}^{\text{vac}}(e^-)$	0.0	0.0
$\mu_{298K, \text{latm}}^{\ominus, \text{SHE}}(e^-)$	-4.34 ± 0.02	-418.5 ± 2.1
$\mu_{298K, \text{latm}}^{\ominus, \text{aq}}(e^-)$	-1.6638 ± 0.04	-160.5 ± 2.1
$\mu_{298K, \text{latm}}^{\ominus, \text{aq}}(H^+)$	-11.6511 ± 0.02	-1124.2 ± 2.1

With the values given in **Table 2.1** all that remains is to calculate the absolute chemical potentials of $Red_{(aq)}$ and $Ox_{(aq)}$. These are obtained via single point calculations, as discussed above, with inclusion of the appropriate Gibbs corrections. Such methods can and have been used to provide insight into, for example the oxidative power of various non-heme iron-oxo complexes⁴¹ and the processes involved in the oxidative damage of DNA^{49,50}.

Potential Energy Surfaces (PES). In a recent article we discussed the applications of potential energy surfaces in the study of enzymatic reactions.⁵¹ Thus, rather than discuss them in detail we instead herein present a concise review.

As noted above, when investigating chemical reactions chemists and biochemists are usually only interested in key, mechanistically relevant structures; *e.g.* the reactants, products, intermediates and transition states.⁵ Alternatively, it can be said that they are interested in the passage of a chemical system through various energy minima and maxima, and the structural and thermochemical relationships between them.⁴ The reactants, intermediates and products all exist on a multi-dimensional-surface. However,

one can represent the lowest energy path that interconnects each of these mechanistically relevant points as a surface, a PES, that generally involves only two coordinates; energy and reaction coordinate (**Figure 2.3**).² Notably, such a concept would not be possible if it were not for the BO approximation.¹¹

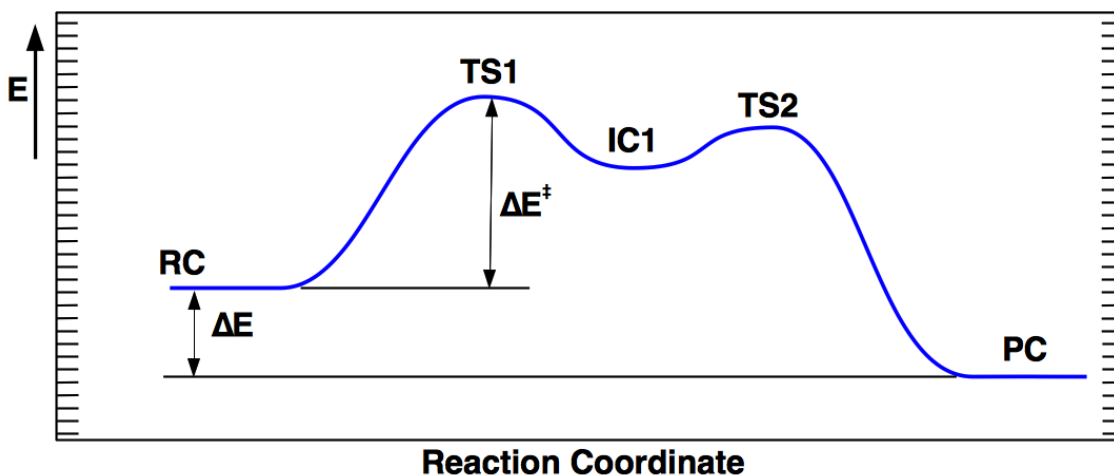


Figure 2.3. A PES for a two-step mechanism that consists of a reactive complex (RC), two transition states (TS1 and TS2), intermediate complex (IC1) and a product complex (PC).

A PES can provide a considerable amount of information and insight into a chemical system. For example, the sequence of steps that occur in the reaction can be determined and illustrated. In addition, the energetic differences between a transition state and its associated minima gives the activation barrier(s) (*i.e.* ΔE^\ddagger) for going from one intermediate complex to another. Importantly, from a PES one can determine if a reaction is feasible given the reaction conditions.

2.16 References

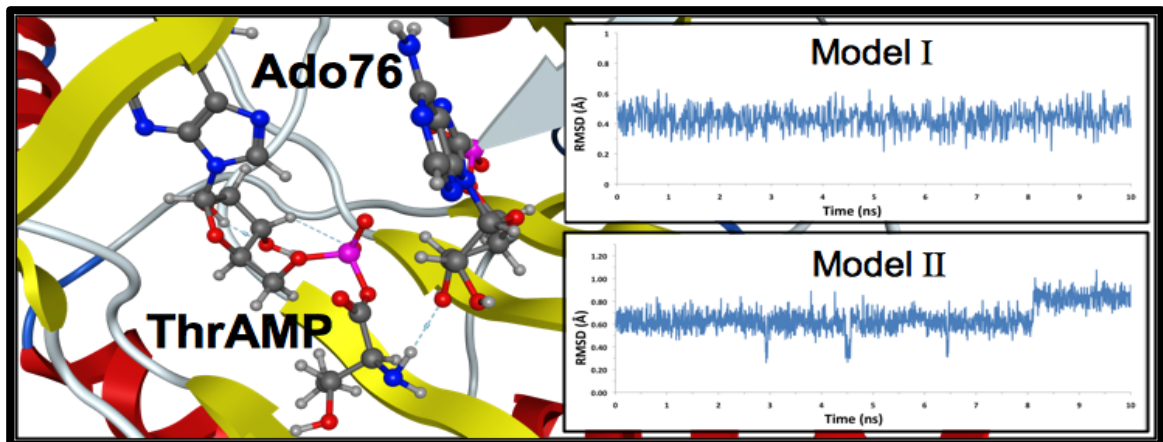
- (1) Harvey, J. N.; Bathelt, C. M.; Mulholland, A. J. *J. Comput. Chem.* **2006**, *27*, 1352-1362.
- (2) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons Ltd.: New York, **2002**.
- (3) Levine, I. N. *Quantum Chemistry*; Prentice Hall: New Jersey, **1991**.
- (4) Lewars, E. *Computational Chemistry: Introduction to the Theory and Applications on Molecular and Quantum Mechanics*; Springer: New York, **2003**.
- (5) Young, D. *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*; Wiley-Interscience: New York, **2001**.
- (6) Johnson, E. R.; Keinan, S.; Mori-Sanchez, P.; Contreras-Garcia, J.; Cohen, A. J.; Yang, W. T. *J. Am. Chem. Soc.* **2010**, *132*, 6498-6506.
- (7) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439-10452.
- (8) Mackie, I. D.; DiLabio, G. A. *J. Phys. Chem. A* **2008**, *112*, 10968-10976.
- (9) DiLabio, G. A. *Chem. Phys. Lett.* **2008**, *455*, 348-353.
- (10) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757-10816.
- (11) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: New York, **1999**.
- (12) Koch, W., Holthausen, M.C. *A Chemist's Guide to Density Functional Theory*; Second ed.; Wiley-VCH: New York, **2001**.
- (13) Perdew, J. P.; Schmidt, K. *AIP Conf. Proc.* **2000**, *577*, 1-20.
- (14) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
- (15) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.
- (16) Lill, S. O. N.; Siegbahn, P. E. M. *Biochemistry* **2009**, *48*, 1056-1066.
- (17) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695-701.

- (18) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729-738.
- (19) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (20) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- (21) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403-412.
- (22) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
- (23) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.
- (24) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.
- (25) Yang, Y.; Weaver, M. N.; Merz, K. M. *J. Phys. Chem. A* **2009**, *113*, 9843-9851.
- (26) Leach, A. R. *Molecular Modelling: Principles and Applications*; Second ed.; Pearson Education Limited: England, **2001**.
- (27) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.
- (28) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
- (29) Senn, H. M.; Thiel, W. *Angew. Chem.-Int. Edit.* **2009**, *48*, 1198-1229.
- (30) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949-1017.
- (31) Llano, J., Gauld, J. W. Mechanistics of Enzyme Catalysis: From Small to Large Active-Site Models, In *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta, C. F., Ed.; Wiley-VCH: Weinheim, **2010**; Vol. 2, p 920.
- (32) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431-12441.
- (33) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645-15650.
- (34) Zhang, Y. K.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459-4463.

- (35) Moreira, I. S.; Martins, J. M.; Ramos, R. M.; Fernandes, P. A.; Ramos, M. J. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 404-414.
- (36) Sousa, S. F.; Coimbra, J. T. S.; Paramos, D.; Pinto, R.; Guimares, R. S.; Teixeira, V.; Fernandes, P. A.; Ramos, M. J. *J. Mol. Model.* **2013**, *19*, 673-688.
- (37) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307-327.
- (38) Pyrkosz, A. B.; Eargle, J.; Sethi, A.; Luthey-Schulten, Z. *J. Mol. Biol.* **2010**, *397*, 1350-1371.
- (39) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6620-6625.
- (40) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142-151.
- (41) Hu, L. H.; Soderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2011**, *7*, 761-777.
- (42) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421-437.
- (43) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032-3041.
- (44) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506-10517.
- (45) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151-5158.
- (46) Tomasi, J.; Mennucci, B.; Cancès, E. *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211-226.
- (47) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999-3093.
- (48) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193-10206.
- (49) Bushnell, E. A. C.; Fortowski, G. B.; Gault, J. W. *Inorg. Chem.* **2012**, *51*, 13351-13356.
- (50) Bushnell, E. A. C.; Llano, J.; Eriksson, L. A.; Gault, J. W. Mechanisms of Mutagenic DNA Nucleobase Damages and Their Chemical and Enzymatic Repairs Investigated by Quantum Chemical Methods, In *Selected Topics in DNA Repair*; Chen, P. C., Ed.; InTech: **2011**.
- (51) Bushnell, E. A. C.; Huang, W.; Gault, J. W. *Adv. Phys. Chem.* **2012**, *2012*.

Chapter 3

A MD Investigation into Substrate Binding and Identity of the Catalytic Base in the Mechanism of ThrRS

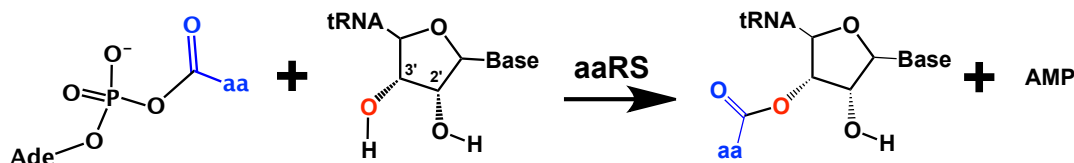


3.1 Introduction

Aminoacyl-tRNA synthetases (aaRS's) are ubiquitous in nature with central roles in a range of physiological processes including apoptosis, inflammation and porphyrin biosynthesis.^{1,2} They are perhaps most well-known, however, for their critical role in protein biosynthesis. More specifically, they catalyze the 'activation' of amino acids and attachment to their cognate tRNA.² The amino acid residues can then be polymerized by the cell's ribosomes to produce the genetically encoded proteins. For each of the twenty "standard" α -amino acids there exists a respective aaRS.³ While differing in structure and having a low degree of sequence similarity, they do exhibit a number of commonalities.²⁻¹⁰ In particular, the overall tRNA aminoacylation process as catalyzed by all aaRS's proceeds via two half-reactions. In the first, they catalyze the reaction of their specific target amino acid with adenosine triphosphate (ATP) to give the corresponding aminoacyl-adenylate (aaAMP) derivative with release of pyrophosphate (PP_i). In the second half, aaRS's catalyze the transfer of the aminoacyl (aa) group from the aaAMP to the 2'- or 3'-position of ribose of the cognate tRNA (tRNA^{aa}) at the Ado76 nucleotidyl residue (aa-tRNA^{aa}).^{2,11,12}

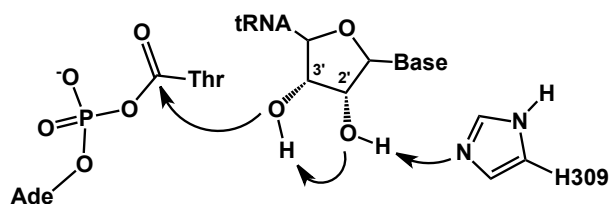
Typically, in class-I aaRS's, aminoacylation occurs at the Ado76-2'-oxygen, and in the class II, at the Ado76-3'-oxygen. For both classes, however, it has been proposed that this process occurs via similar mechanisms (**Scheme 3.1**).^{9,13} Specifically, a base within the active site is thought to deprotonate the target hydroxyl of the ribose sugar of the Ado76 residue. This enhances the nucleophilicity of the hydroxyl's oxygen and facilitates its attack at the carbonyl carbon of the amino acid moiety of the aaAMP substrate. In general, however, the exact identity of the Brønsted base is unclear. Indeed, the active sites of aaRS's typically lack any residue that may act as a proton acceptor in the esterification process.^{13,14} In some cases, the basic group appears to be a non-bridging phosphate oxygen of the aaAMP substrate itself.^{13,15} In a recent detailed computational

study on a histidyl-RNA synthetase (HisRS),¹⁶ we showed that it was thermodynamically feasible for the *pro-S* non-bridging phosphate oxygen to act as the required mechanistic base. In fact, it has been suggested¹⁰ that such a substrate-assisted catalytic process may be a common approach in aaRS's.^{10,13,17}



Scheme 3.1. Aminoacyl transfer from aa-AMP to Ado76-3'-OH (second half-reaction) as catalyzed by a class-II aaRS.

In the case of the class-II threonyl-tRNA synthetase (ThrRS), however, recent mutagenesis studies by Minajigi and Francklyn¹⁴ have suggested that the aminoacyl-adenylate substrate's phosphate is *not* essential to the mechanism. In fact, only small decreases in the rate of reaction by no more than 3-fold were observed upon substitution of either the *pro-R* or *-S* non-bridging phosphate oxygens. In contrast, substitution of an active-site histidine (His309) by alanine had a significantly larger effect, decreasing the reaction rate by ~240-fold.¹⁴ Hence, it was proposed that His309 may be the mechanistic base and furthermore, that it may directly or indirectly (via a water molecule) deprotonate the Ado76-2'-OH group of the cognate tRNA (tRNA^{Thr}), which then subsequently deprotonates the adjacent Ado76-3'-OH group (**Scheme 3.2**).¹⁴ Notably, for HisRS, substitution of the His-AMP substrate's *pro-S* oxygen resulted in a considerably more marked rate decrease of ~10⁴-fold.¹⁷



Scheme 3.2. Proposed mechanism for aminoacyl transfer as catalyzed by ThrRS via deprotonation of 2'-OH-ribonucleoside by His309.¹⁴

An important step in any enzymatic reaction is the formation of a fully reactive enzyme-substrate complex, that is, the binding and positioning of substrate(s) and active-site functional groups, e.g., residues, cofactors and water molecules. In addition, the structure of such complexes can provide invaluable insights into the catalytic pathway of that enzyme such as the identity of possible key active site functional groups and their potential mechanistic roles. To date, unfortunately, no experimentally derived NMR or X-ray structure for ThrRS has been reported in which both substrates for the second half-reaction (ThrAMP and tRNA^{Thr}) are bound within its active site.

Several X-ray crystallographic structures have been obtained, however, in which a substrate or substrate-analogues are bound within the active site of ThrRS.¹⁸⁻²¹ In particular, several structures have been obtained in which only the substrate-analogues ThrAMS¹⁹⁻²¹ (**Figure 3.1a**) or SerAMS¹⁸ are bound, while another has been obtained²⁰ with both AMP *and* Thr-tRNA^{Thr} simultaneously bound (**Figure 3.1b**). From these structures, it was concluded that the aminoacyl-adenylate substrates *pro-R* and *-S* non-bridging oxygens likely interact with an arginyl and asparagyl residue, respectively. In addition, its aminoacyl moieties α -amino and side-chain hydroxyl (γ -OH) groups are bound to an active site Zn(II) ion (**Figure 3.1a**). This bidentate coordination has been proposed to be an essential characteristic of ThrRS, allowing the enzyme to discriminate

against valine, which is the isosteric analog of threonyl.²¹ It should be noted that on binding Thr, a Zn(II)-bound water is displaced but may remain within the fully-bound active site, as a water was detected in several crystallographic structures.¹⁸⁻²¹ In addition, the Ado76-2'-OH group of the Thr-tRNA^{Thr} moiety is thought to interact with the side-chains of tyrosyl (Tyr462) and histidyl (His309), **Figure 3.1b**. The Tyr462...2'-OH interaction has been suggested to be important in binding of Thr-tRNA^{Thr} to ThrRS. In particular, a peptide 'loop' containing Tyr462 undergoes a conformational change forming a hydrogen-bond between the two groups. This resulting interaction is thought to help stabilize the active-site region by enabling the formation of additional interactions such as stacking of the highly conserved Phe461 and Asn312 residues.²¹

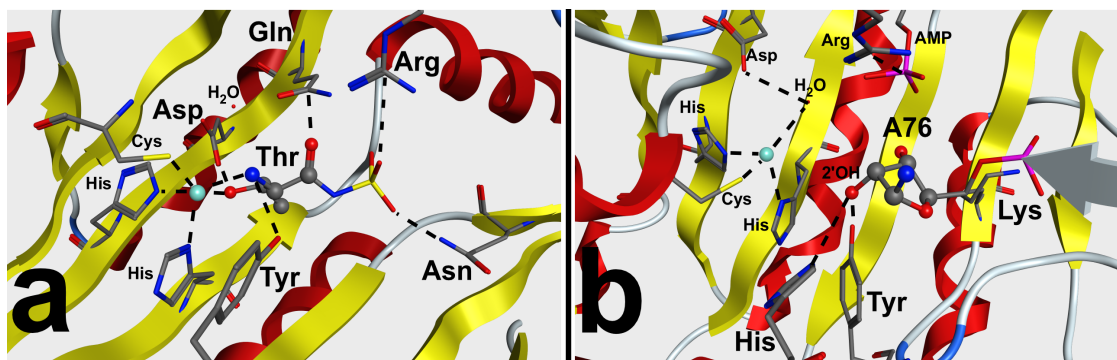


Figure 3.1. Active site of ThrRS bound with: (a) Thr-AMP [PDB ID: 1EVL],¹⁹ (b) tRNA^{Thr} and AMP [PDB ID: 1QF6].²⁰

Computational methods are now widely applied in the study of enzyme chemistry including, for example, the structure and properties of enzyme-substrate/intermediate complexes.²² In particular, it is noted that molecular dynamics (MD) methods have successfully been applied previously to the study of various aspects of aaRS chemistry.^{15,23-31}

In this present chapter, MD simulations have been used to investigate the structure of viable Michaelis complexes for the enzyme ThrRS with both tRNA^{Thr} and ThrAMP

bound within its active site, i.e., complexes that would lead to aminoacyl transfer in the second half-reaction. In particular, we have examined the effects of both a neutral and protonated His309-N^ε side-chain. Moreover, we considered the possible involvement of an active-site water in bridging between His309-N^ε and Ado76-2'-OH, thus acting as a proton shuttle in the deprotonation of the latter.¹⁴

3.2. Computational Methods

The Molecular Operating Environment (MOE)³² software package was used for all calculations.

Design of Chemical Model. It has been experimentally observed that conformational changes occur within the active site of ThrRS upon binding of the tRNA^{Thr} cosubstrate.^{29,33} Hence, an X-ray crystal structure of a ThrRS···tRNA^{Thr}+AMP complex (PDB ID: 1QF6)^{29,33} was used as the template structure. This structure was then manually modified using MOE in order to include the missing threonyl moiety. Specifically, threonyl was added to the AMP substrate and positioned in the active site in accordance with that observed in X-ray crystal structures of the enzyme-substrate analogue complexes ThrRS···ThrAMS (PDB ID: 1EVL and 1NYQ),^{14,18} ThrRS···SerAMS (PDB ID: 1FYF),^{18,19,21} and ThrRS···Thr (PDB ID: 1NYR).²¹ The α-NH₂ group of the Thr moiety (Thr-NH₂) and His309 residue were modeled as neutral.

Solvation and Annealing. Prior to data collection, an MD simulation was performed in order to obtain the corresponding solvated enzyme–substrates complex and to enable it to undergo thermal relaxation. Specifically, the enzyme–substrate complex was surrounded by a 7-Å spherical layer of water molecules. An ellipsoidal potential wall with a scaling constant of 2 was then placed around the resulting solvated complex in order to force the system to lie within the volume of space defined by the ellipsoid. In addition, the distance between the 3'-OH and C=O of the ThrAMP moiety was restrained by a force of 10 kcal

mol^{-1} to a distance of 2.6 Å, thus allowing the whole system to reach a conformation consistent with the mechanistically relevant reactive conformation. It should be noted that this restraint was removed in all subsequent production runs. The damping functional factor included in the electrostatic and van der Waals potentials were set to decay smoothly beyond 8 – 10 Å. The geometry of the solvated complex was then optimized using the AMBER99 force field until the root mean square gradient of the total energy fell below $0.05 \text{ kcal mol}^{-1} \text{ Å}^{-1}$.

The MD simulation was then performed under constrained pressure and temperature while the equations of motion were coupled with the Nosé–Poincaré thermostat³⁴ and the time step for numerical integration was set to 2 fs. The system was annealed by heating it from 150 to 300 K over a period of 25 ps, holding the temperature constant at 300 K for an additional period of 25 ps, then heating from 300 to 400 K over a period of 25 ps, and subsequently holding the temperature constant at 400 K for a further interval of 375 ps. At the 450-ps mark, the system was allowed to cool down by decreasing the temperature from 400 to 300 K over a period of 25 ps and then holding the temperature constant at 300 K for an interval of 25 ps.

Molecular Dynamics Production Runs. From the final structure of the above annealing simulation, four enzyme–substrate complexes were then obtained. The complexes differed in the protonation state of His309 and in the position of the water molecule in the vicinity of Ado76-2'-OH and His309 as a potential hydrogen bond bridge between the two. The structure of each of the complexes was then optimized using the AMBER99 force field until the root mean square gradient of the total energy fell below $0.05 \text{ kcal mol}^{-1} \text{ Å}^{-1}$. The damping functional factor included in the electrostatic and van der Waals potentials were set to decay smoothly beyond 8 to 10 Å. In addition, only those residues, nucleobases and waters within 15 Å of the ThrAMP and Ado76 moieties were free to move, leaving all the other atoms fixed at the positions that they end up after the MD annealing. As described above for the annealing process, the subsequent MD simulations

were performed under constrained pressure and temperature, the equations of motion were coupled with the Nosé–Poincaré thermostat,³⁴ and the time step for numerical integration was set to 2 fs. Each of the simulations was then run over 15 ns for equilibration, and the last 10-ns data of this equilibration phase were used for RSMD and cluster analyses.

3.3 Results and Discussion

As detailed in the Introduction, aaRSs are thought to use a common catalytic approach for the second half-reaction. In particular, a base deprotonates either the Ado76-2'-OH or Ado76-3'-OH hydroxyl group of the cognate tRNA cosubstrate. The resulting Ado76-2'-O⁻ or Ado76-3'-O⁻ oxyanion can then nucleophilically attack the carbonyl of the substrate aaAMP.^{9,16} In contrast to that proposed for other aaRSs, in the case of ThrRS, an active site histidyl residue (His309) has been suggested to act as the catalytic base.¹⁴ Specifically, its side-chain His309-N^ε centre deprotonates the Ado76-2'-OH group either directly or via a bridging H₂O molecule. This is then followed by proton transfer from Ado76-3'-OH to the resulting Ado76-2'-O⁻ oxyanion (see **Scheme 3.2**). Within an aqueous environment at SATP, the p*K*_a of histidine's imidazole is approximately 6.0.³⁵ However, this value can significantly fluctuate under the influence of the local protein environment. For example, in aqueous solution at SATP, the p*K*_a of the guanidinium side-chain of arginine is ~12.5.³⁵ Yet, in the case of the enzyme UROD, we have previously shown that the reduced local polarity of its active-site environment markedly lowers the p*K*_a's of the R-groups of two active-site arginyl residues such that they in fact may act as proton donors in the catalytic mechanism.³⁶

Hence, since the aminoacyl transfer can potentially either be acid- or base-catalyzed, we have considered the case of His309-N^ε being either protonated (His309-N^εH⁺) or neutral (His309-N^ε:) in the initial Michaelis complex. In the case of His309-N^εH⁺, the

possible catalytic mechanism could involve transfer of the proton to the carbonyl oxygen of the ThrAMP substrate. This would enhance the electrophilicity of the adjacent carbonyl carbon centre, and thus, its susceptibility to nucleophilic attack by the Ado-3'-OH oxygen.

Accordingly, four models for the Michaelis complex of ThrRS, summarized in **Table 3.1**, were generated. We took into account that the water is either absent or present and that the His-N^ε centre is either neutral (models **I** and **II**) or protonated (models **I-H⁺** and **II-H⁺**). We also considered complexes in which the α-amine of the threonyl moiety was protonated. However, these led to structures that differed significantly from those experimentally obtained. A similar approach has recently been applied to investigate the structural dynamics of the riboswitch in the active site of glucosamine-6-phosphate synthetase with changing protonation states.³⁷ In particular, Banas et al.³⁷ carried out MD simulations involving various protonation states of three crucial active-site moieties to probe the dominant protonation states of these key active-site residues. In our investigation of each model of the Michaelis complex, the root mean square deviations (RMSDs) in the positions of the His309 imidazole, Tyr462 phenol, Ado76 ribose ring and the threonyl moiety were calculated over the last 10 ns of the production run. These RMSDs were calculated with respect to the minimized starting structure to ensure that the conformational equilibrium was reached. In the case of Model **II** and **II-H⁺**, the position of the bridging water was also included in RMSD calculation. Specifically, the phenol ring of Tyr462 was included because it has been suggested to be important in binding of Thr-tRNA to ThrRS. As stated in the Introduction a peptide 'loop' containing Tyr462 undergoes a conformational change upon complete substrate binding such that the Tyr462-Ph-OH moiety forms a hydrogen-bond to the Ado76-2'-OH group, with Tyr462 acting as the hydrogen-bond donor. This conformational change and resulting interaction is proposed to help stabilize the active-site region.²¹

Table 3.1. Summary of Michaelis complex models considered in this present study.

Model	His309	Thr-NH ₂	Water
I	Neutral	Neutral	Absent
I-H⁺	Protonated	Neutral	Absent
II	Neutral	Neutral	Present
II-H⁺	Protonated	Neutral	Present

Effects of protonation at His309-N^ε on its direct hydrogen-bonding to Ado76-2'-OH.

We began by considering the effects of protonating His309, specifically at its imidazole N^ε centre (His309-N^ε), on its hydrogen bonding with Ado76-2'-OH. For model **I** (i.e., neutral His309-N^ε with no bridging H₂O) the RMSDs in the positions of the His309 imidazole, Tyr462 phenol, Ado76 ribose ring and the threonyl moiety relative to the optimized initial starting structure was calculated for each configuration obtained after equilibration during the production run (see Computational Methods). A plot of the calculated RMSDs is shown in **Figure 3.2**.

As can be seen, during the production run the positions of the His309 imidazole, Tyr462 phenol, Ado76 ribose ring, and the threonyl moiety maintain a quite consistent configuration. Indeed, with very few exceptions almost all RMSDs lie within a quite narrow range of 0.2 – 0.6 Å; indicating that there were no significant changes in their positions during the 10 ns production simulation.

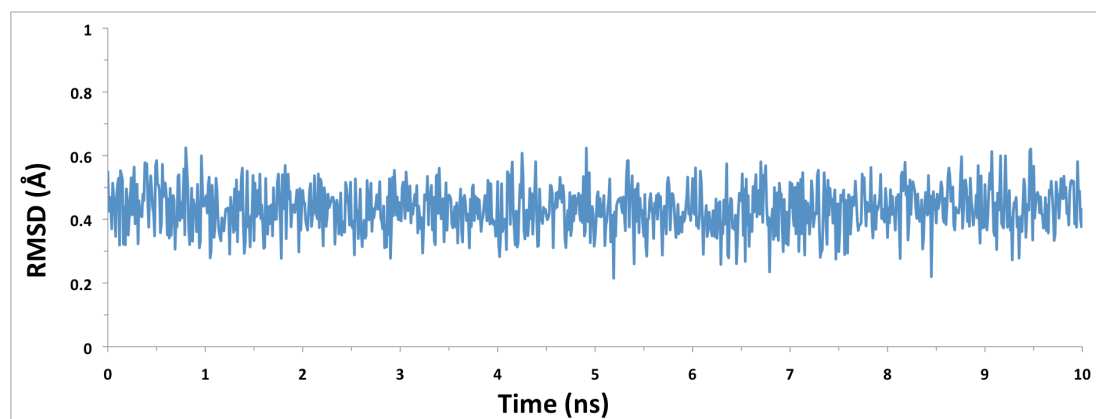


Figure 3.2. RMSDs in the positions of the imidazole of His309, phenol of Tyr462, ribose sugar ring of Ado76 and the threonyl moiety during the 10 ns simulation of model **I**.

The structures sampled during the simulation were investigated further using clustering analysis in order to group the RMSDs into five clusters. A representative average structure of each cluster was then selected and their bound active sites then overlaid with the others as shown in **Figure 3.3**.

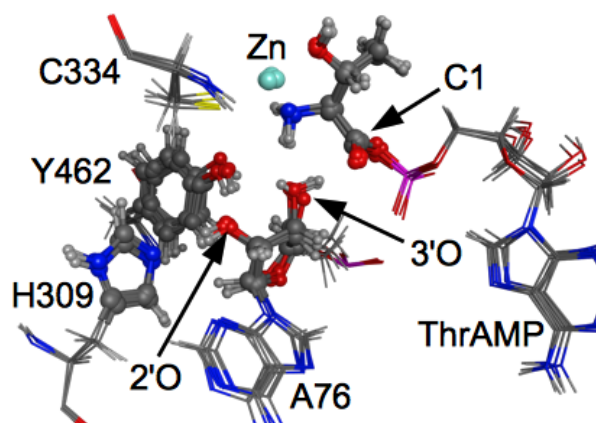


Figure 3.3. Overlay of the five representative average structures of model **I** obtained from cluster analysis of the RMSDs during the 10 ns MD production simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.2**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white)].

As can be seen in **Figure 3.3**, the overlaid average structures show that within the fully bound active site a number of proposed mechanistically important interactions are quite consistent while some others appear to be more variable. For instance, in order to act as the required base as detailed in the Introduction, the His309 residue must obtain a proton from the Ado76-2'-OH group.¹⁴ Over the 10 ns production run the 2'-OH group is consistently positioned such that its proton is directed towards the neutral imidazole ϵ -nitrogen (N^ϵ) centre of the His309 residue. Furthermore, it has an average His309- $N^\epsilon \cdots O-2'-Ado76$ distance of 3.02 Å which is only moderately shorter than observed experimentally (3.29 Å) in the available X-ray crystallographic structure (PDB ID: 1QF6).²⁰ Similarly, a consistent Tyr462-OH $\cdots O-2'-Ado76$ hydrogen bonding interaction is also observed. Furthermore, it has an average Tyr462- $O \cdots O-2'-Ado76$ distance of 2.88 Å; which is also in good agreement with that observed (2.81 Å) in the available X-ray crystallographic structure (PDB ID: 1QF6).²⁰

In contrast, considerably greater variability in the five average structures is observed for the orientation of the mechanistically key Ado76-3'-OH group. As detailed above (see **Scheme 3.2**), it has been proposed that the Ado76-2'-oxygen deprotonates the adjacent Ado76-3'-OH group.¹⁴ However, over the course of the simulation only approximately 10.7% of the conformers obtained were observed to have a suitable orientation for a 2'- $O \cdots H-O-3'$ hydrogen bond. Furthermore, during the 10 ns simulation, the 2'- $O \cdots HO-3'$ distance varies quite significantly from 1.77 to 3.88 Å with a markedly long average distance of 3.11 Å. This elongated hydrogen-bond distance suggests that in model I (i.e., a neutral His309 residue and no additional active site water), it is unlikely that the Ado76-2'-oxygen would be able to deprotonate the adjacent Ado76-3'-OH group.

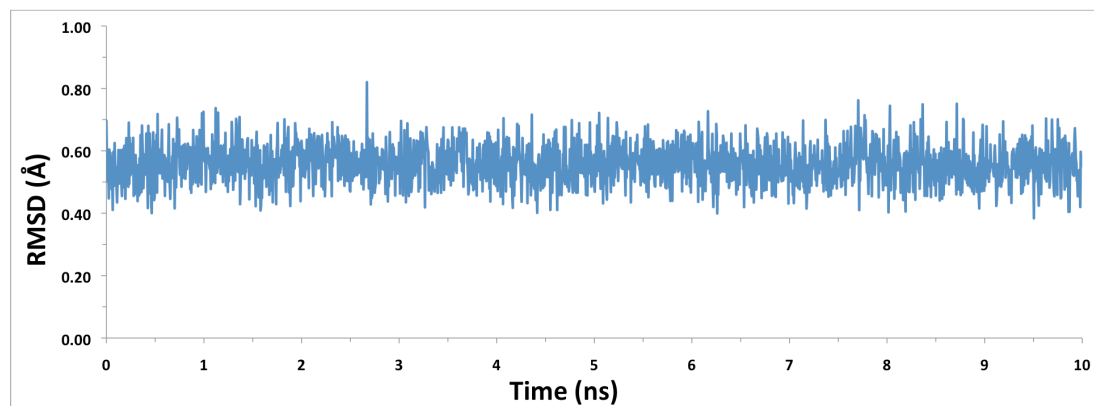


Figure 3.4. RMSDs in the positions of the imidazole of His309, phenol of Tyr462, ribose sugar ring of Ado76 and the threonyl moiety during the 10 ns simulation of model **I-H⁺**.

The effects of protonating His309 on the above fully bound active-site were then examined using model **I-H⁺**; i.e., His309-N^εH⁺ with no additional active site water. The RMSDs for the positions of the His309 imidazole, phenol of Tyr462, ribose ring of Ado76 and the threonyl moiety relative to the starting structure were calculated for each structure and plotted in **Figure 3.4**. Similar to that observed above for the corresponding unprotonated fully bound active-site model **I**, overall, the positions of the above four moieties maintain a reasonably consistent configuration throughout the 10-ns simulation. However, the average RMSD is now moderately larger by approximately 0.2 Å, ranging from slightly below 0.4 Å to just over 0.8 Å (*c.f.*, **Figure 3.2**). This may indicate an increased mobility of the groups within the active site.

A cluster analysis was then performed on the RMSD's in **Figure 3.4** to obtain five clusters. A representative average structure of each was then obtained and overlaid (**Figure 3.5**). As observed in the unprotonated fully bound active site (model **I**), all five structures have quite similar hydrogen bonding networks. However, they also exhibit some key differences between each other and, importantly, from that observed in model **I**. For example, due to the fact that it is now protonated, the His309 imidazole now acts as a hydrogen bond donor via its His309-N^εH⁺ moiety to the 2'-oxygen of Ado76. Indeed,

this interaction is consistent in all five average structures. Furthermore, the average His309-N^E:...O-2'-Ado76 distance during the simulation is 2.97 Å. This is slightly shorter by 0.05 Å than that observed in the corresponding 'neutral' bound active-site (model **I**: 3.02 Å, see **Figure 3.3**), and consequently, 0.34 Å shorter than in the X-ray crystallographic structure (PDB ID: 1QF6).²⁰ Thus, protonation of His309 results in a strengthening of its interaction with Ado76.

As a consequence of this now reversed and strengthened His309...Ado76 interaction, in contrast to that observed in model **I**, the Ado76-2'-OH group is now able to act as a hydrogen-bond donor towards the phenolic oxygen of Tyr462. Furthermore, the Tyr462-O...O-2'-Ado76 distance has increased by 0.21 Å compared to model **I** to 3.09 Å. This indicates that the hydrogen bond between these two groups has weakened. Indeed, the Tyr462-O...HO-2'-Ado76 was not consistently maintained throughout the simulation. In fact, in approximately 43% of the conformers obtained, the Ado76-2'-OH group formed an intramolecular hydrogen bond with the adjacent Ado76-3'-oxygen, i.e., Ado76-2'-OH...O-3'-Ado76. Thus, protonation of the imidazole of the His309 residue appears to weaken the interaction between the Tyr462-OH and Ado76-2'-OH groups, which has been previously proposed to aid in stabilizing the fully bound active-site.²¹ As well, in the fully bound active site model **I-H⁺** the Ado76-2'-oxygen is unlikely to act as a Brønsted base and deprotonate the Ado-3'-OH group.

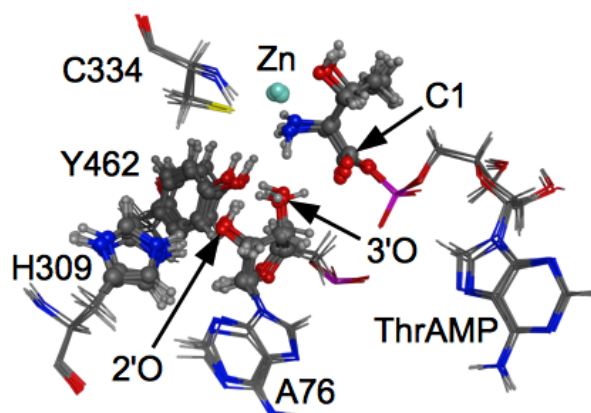


Figure 3.5. Overlay of the five representative average structures of model **I-H⁺** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.4**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white)].

In an acid catalyzed aminoacyl transfer process the His309-N^ε-H⁺ proton could potentially transfer via the Ado-2'-OH and Ado-3'-OH groups onto the aminoacyl's carbonyl oxygen. This would enhance the electrophilicity of the adjacent carbonyl carbon (C1), and in turn, its susceptibility to nucleophilic attack. However, from **Figure 3.5**, it can be seen that the Ado76-3'-OH does not consistently hydrogen bond to the carbonyl oxygen but rather to the Tyr462-OH group. Thus, it is unlikely that His309 would indirectly protonate the carbonyl oxygen. In addition, a mechanistically important geometric parameter is undoubtedly the distance between the Ado76-3'-oxygen and the C1 centre of the threonyl moiety of the ThrAMP substrate. For model **I-H⁺**, the average Ado76-3'O...C1-ThrAMP distance is 3.02 Å. Notably, this is 0.07 Å greater than that observed (2.95 Å) in the corresponding "unprotonated fully bound active site" model **I** (2.95 Å). Thus, protonation of the imidazole of the His309 residue appears to also

negatively affect suitable positioning of the Ado76-3'-OH for nucleophilic attack at the threonyl's C1 centre.

Effects of an additional active-site water on the hydrogen bonding between His309-N^ε and Ado76-2'-OH. It has been alternatively suggested that an additional active-site water may act as a hydrogen-bond bridge between His309-N^ε and the Ado76-2'-OH group.¹⁴ In order to examine this possibility, a water molecule was added to both models **I** and **I-H⁺** in the vicinity of His309-N^ε and Ado76-2'-OH, generating models **II** and **II-H⁺** respectively (see Computational Methods).

For the fully bound active site containing an extra water but a neutral His309 residue, i.e., model **II**, the RMSDs for each conformation during the 10-ns period was determined based on the position of the His309 imidazole, Tyr462 phenol ring group, Ado76 ribose sugar, threonyl moiety and the added "bridging" water. The resulting values are shown plotted in **Figure 3.6**.

It can be clearly seen upon comparison with the RMSDs obtained for the corresponding model **I** (**Figure 3.2**), that introduction of the additional water significantly increases fluctuations in the positioning of the above moieties in the bound active site. In particular, at approximately 3.0, 4.5 and 6.5 ns in the collection period, large, sudden but short-lived deviations lasting just fractions of a nanosecond are observed with RMSDs decreasing by ~ 0.3 Å. In addition, at approximately 8 ns a sudden, large increase in the RMSDs of ~ 0.2 Å is observed lasting ~ 2 ns. All these fluctuations in this RMSD profile reflect the intermittence of the hydrogen-bonding network of the bridging water molecule within the active site, and this contributes to disrupt the hydrogen-bonding network of Ado76.

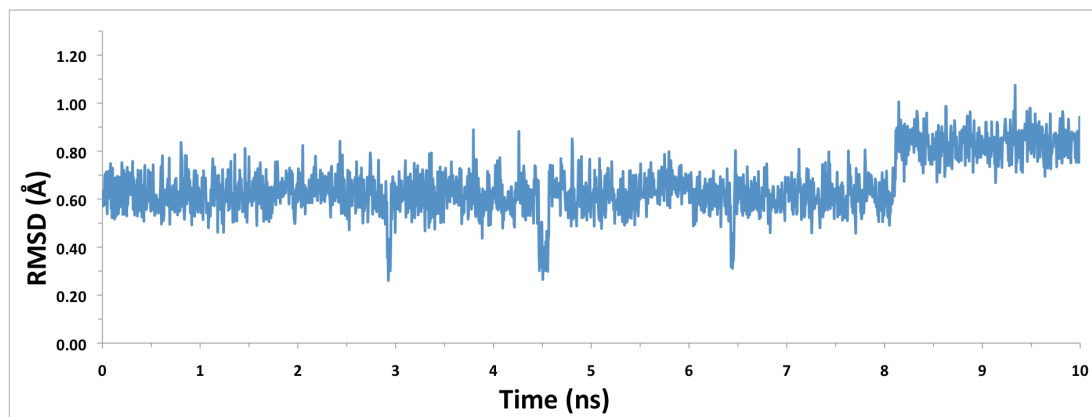


Figure 3.6. RMSDs in the positions of the His309 imidazole, phenol of Tyr462, ribose sugar ring of Ado76, threonyl moiety and added "bridging" water molecule during the 10 ns simulation of model **II**.

The RMSDs were then subjected to a clustering analysis as per models **I** and **I-H⁺**. Again, five clusters were produced and a representative average structure obtained for each. These structures are shown overlaid with each other in **Figure 3.7** and clearly show the greater variability in the relative positioning of the His309 imidazole, phenol of Tyr462, ribose sugar ring of Ado76, threonyl moiety and "bridging" water molecule. Importantly, despite the average His309-N^ε...O-2'-Ado76 distance being quite long at 3.44 Å, at no time does it appear that the water inserts itself between His309 and Ado76-2'-OH. Instead, it prefers to sit "to the side" of both groups. As a consequence, the average Ado76-2'O...OH₂ distance is also markedly long at 2.75 Å while that for His309-N^ε...OH₂ is significantly even longer at 3.68 Å. This further suggests that any hydrogen bonding interactions between the H₂O and Ado76 and His309 are moderate or quite weak, respectively.

In comparison to that observed in models **I** and **I-H⁺**, the Ado76-2'-OH group has markedly greater mobility. As a result, it alternates between acting as a hydrogen bond donor with either the added H₂O moiety or His309 imidazole N^ε centre. The inclusion of the water was also found to negatively impact the interaction between the phenolic

hydroxyl of Tyr462 and the Ado76-2'-OH groups. Indeed, while a Tyr462-OH \cdots O-2'-Ado76 is consistently observed in all five average structures (**Figure 3.7**), the average Tyr462-O \cdots O-2'-Ado76 distance has increased markedly by 0.38 Å (model **I**) to 3.26 Å. This suggests that while the interaction between these two groups is maintained, it has been notably weakened. In addition to resulting in less stabilization of the fully bound active site, this would also likely result in less stabilization of the “fully bound active site-conformation” of the peptide 'loop' containing Tyr462 (as noted in the introduction).²¹

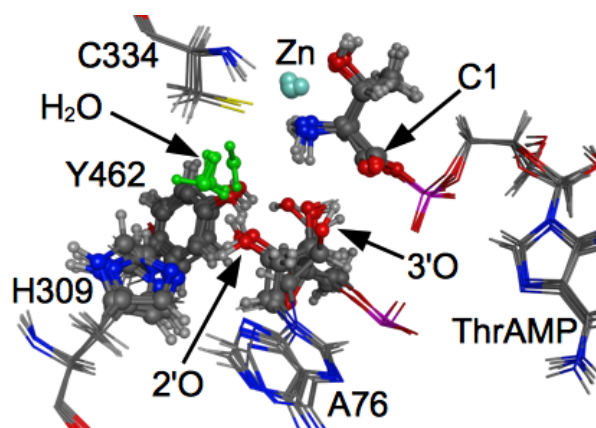


Figure 3.7. Overlay of the five representative average structures of model **II** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.6**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white); added H₂O (green)].

The above changes resulting from inclusion of the water when His309 is neutral, also negatively impact the relative positioning of the mechanistically important Ado-3'-OH oxygen with respect to the threonyl's carbonyl carbon centre (C1). Specifically, the average Ado-3'-O \cdots C1=O distance has increased by 0.09 Å from that observed in model **I** (2.95 Å) to 3.04 Å in the present model **II**.

The effect of protonating the imidazole of His309 in the bound active site containing an additional water was then examined using model **II-H⁺**. Alternatively, this can be thought of as considering the effects of adding a water to the His309-protonated bound active-site model **I-H⁺**. As for the above model **II**, the RMSDs for each conformation during the 10-ns simulation (**Figure 3.8**) was determined based on the positions of the His309 imidazole, Tyr462 phenol, Ado76 ribose sugar, threonyl moiety and the added "bridging" water. In contrast to that observed for model **II**, no sudden, short- or long-lived fluctuations occur in the RMSDs. Instead, apart from the slightly higher fluctuations during the first nanosecond, they remain fairly consistent throughout the 10 ns simulation with most values lying between 0.9 ± 0.1 Å.

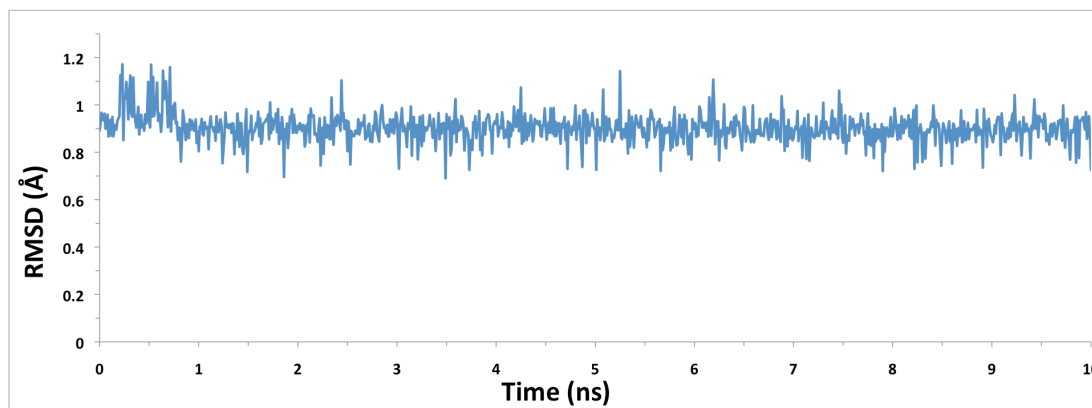


Figure 3.8. RMSDs in the positions of the His309 imidazole, phenol of Tyr462, ribose sugar ring of Ado76, threonyl moiety and added "bridging" water molecule during the 10 ns simulation of model **II-H⁺**.

Again, a cluster analysis of the RMSDs was performed and a representative average structure obtained of each of the five resulting clusters. These are shown overlapped with each other in **Figure 3.9**. Similar to that seen for model **II**, i.e., neutral His309 with an added H₂O, the extra water in model **II-H⁺** does not bridge between His309-N^ε and the Ado76-2'-OH. In fact, the water instead appears to form a hydrogen bonding bridge

between His309-N^ε and the Ado76-3'-OH group. Indeed, the average His309-N^ε⋯OH₂ and H₂O⋯O-3'-Ado76 distances are 2.84 and 3.24 Å, respectively. While the latter is 0.49 Å longer than the analogous H₂O⋯O-2'-Ado76 distance in model **II**, it is noted that the former (i.e., His309-N^ε⋯OH₂) is markedly shorter by 0.84 Å than in model **II**. As a result, the added water appears to be better positioned to act as a bridge between the His309 residue's imidazole and the Ado76 nucleotidyl residue.

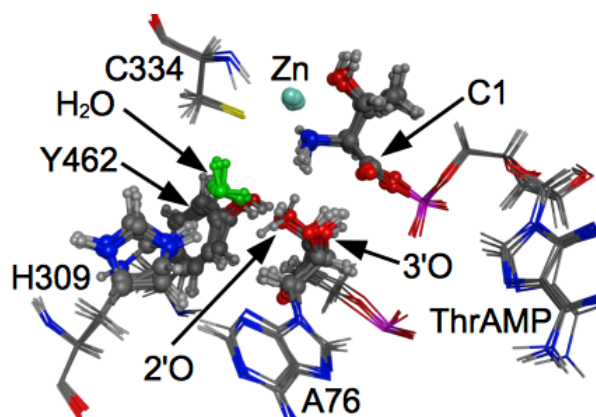


Figure 3.9. Overlay of the five representative average structures of model **II-H⁺** obtained from cluster analysis of the RMSDs during the 10 ns MD simulation. For clarity, not all hydrogens are shown. The enlarged atoms (except the Zn ion) are those used in the calculation of the RMSDs presented in **Figure 3.8**. [Color code: P (pink); C (grey); O (red); N (dark blue); S (yellow); Zn (light blue); H (white); added H₂O (green)].

Importantly, as can be seen in **Figure 3.9**, the ribose ring of Ado76 has dramatically altered its position. In fact, its Ado76-2'-OH hydroxyl now hydrogen bonds with the carbonyl oxygen of the ThrAMP substrate with an average Ado76-2'-O⋯O=C1 distance of 3.02 Å. Furthermore, the average distance between the mechanistically key Ado76-3'-OH oxygen and the threonyl's carbonyl carbon centre has increased significantly by 0.5 to 3.53 Å. Thus, the combination of a protonated His309 residue and addition of a water disfavours nucleophilic attack by the Ado76-3'-oxygen at the Thr-AMP's C1 centre.

In addition, it is noted that the proposed²¹ important stabilizing Tyr462...O-2'-Ado76 interaction has now been broken, with the Tyr462 phenolic hydroxyl instead hydrogen bonding with the Ado76-3'-OH group. The average distance for this latter interaction during simulations is 2.75 Å.

3.4 Conclusions

Molecular dynamics methods were used to investigate the structure of the fully bound active site of ThrRS for the second half-reaction, i.e., with both ThrAMP and tRNA^{Thr} bound. In particular, the ability of His309 to act as either a mechanistic base (i.e. neutral) as proposed by Minajigi and Francklyn,¹⁴ or acid (i.e. protonated), without or with the assistance of a "bridging" water was examined.

In the cases where His309 was protonated or unprotonated, but with no additional "bridging" water added, it was clearly seen that a stable interaction between His309 and Ado76 was formed.¹⁴ Moreover, the Ado76-3'O...C1-ThrAMP average distances of 2.95 Å and 3.02 Å in models **I** and **I-H⁺**, respectively, suggest that nucleophilic attack of C1 is possible for both protonation states. However, the +0.07 Å difference observed in **I-H⁺** does suggest that protonation of His309 negatively affects suitable positioning of the Ado76-3'-OH for nucleophilic attack. Interestingly though, regardless of protonation state of His309 it is unlikely that the Ado76-2'-oxygen would be able to deprotonate the adjacent Ado76-3'-OH group allowing for nucleophilic attack of the C1 carbon of ThrAMP. Moreover, the possibility of an acid catalyzed amino-acyl (i.e. the indirect protonation of the carbonyl oxygen by His309) was also found to be unlikely. It is noted that in a recent cluster-DFT investigation we considered the involvement of His309 in the catalytic mechanism of ThrRS and it was found that it did not act as the general base.³⁸

Upon addition of a water molecule to "bridge" the His309-N^e:...O-2'-Ado76 interaction, for both the neutral and protonated His309, significant disruptions to the

orientation of the active-site groups are observed. That is, the current results suggest that an additional water does not bridge by hydrogen bonding between His309 imidazole and Ado76-2'-OH moieties regardless of the protonation state of His309. In addition, for neutral His309, it does not markedly affect the mechanistically important Ado76-3'-O...C1=O distance (e.g., model **I** versus **II**). In contrast, for protonated His309, it causes significant distortions in the fully bound active-site conformation (e.g., model **II** versus **II-H⁺**). Therefore, a water is not necessary nor does it enhance positioning of the tRNA^{Thr} cosubstrate for nucleophilic attack at the C1 centre of the Thr-AMP substrate.

3.5 References

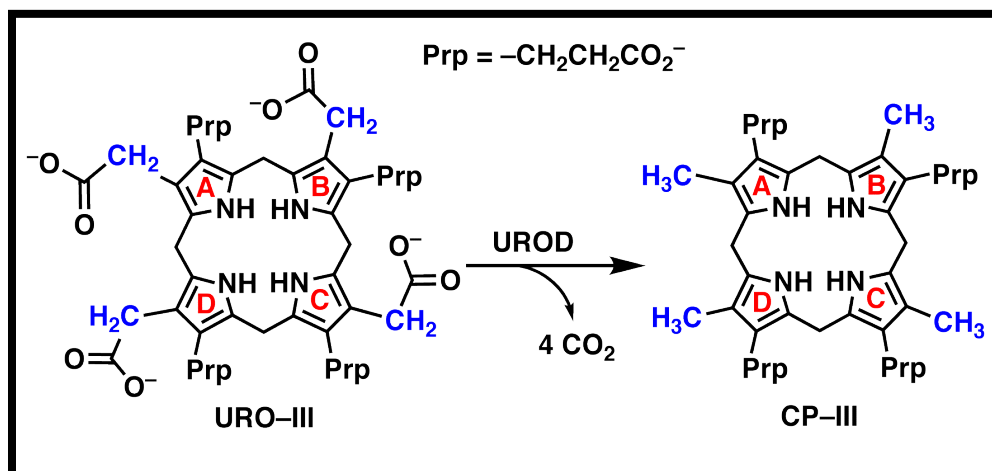
- (1) Heinemann, I. U.; Jahn, M.; Jahn, D. *Arch. Biochem. Biophys.* **2008**, *474*, 238-251.
- (2) Safro, M. G.; Moor, N. A. *Mol. Biol.* **2009**, *43*, 211-222.
- (3) Cusack, S.; Berthetcolominas, C.; Hartlein, M.; Nassar, N.; Leberman, R. *Nature* **1990**, *347*, 249-255.
- (4) Kisselev, L. L.; Favorova, O. O. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1974**, *40*, 141-238.
- (5) Mirande, M. *Prog. Nucl. Acid Res. Mol. Biol.* **1991**, *40*, 95-142.
- (6) Schimmel, P. *Annu. Rev. Biochem.* **1987**, *56*, 125-158.
- (7) Vasil'eva, I. A.; Moor, N. A. *Biochem.-Moscow* **2007**, *72*, 247-263.
- (8) Cusack, S. *Curr. Opin. Struct. Biol.* **1997**, *7*, 881-889.
- (9) Arnez, J. G.; Moras, D. *Trends Biochem.Sci.* **1997**, *22*, 211-216.
- (10) Uter, N. T.; Perona, J. J. *Biochemistry* **2006**, *45*, 6858-6865.
- (11) Ibba, M.; Soll, D. *Annu. Rev. Biochem.* **2000**, *69*, 617-650.
- (12) Malde, A. K.; Mark, A. E. *J. Am. Chem. Soc.* **2009**, *131*, 3848-3849.
- (13) Liu, H. N.; Gauld, J. W. *J. Phys. Chem. B* **2008**, *112*, 16874-16882.

- (14) Minajigi, A.; Francklyn, C. S. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17748-17753.
- (15) Pyrkosz, A. B.; Eargle, J.; Sethi, A.; Luthey-Schulten, Z. *J. Mol. Biol.* **2010**, *397*, 1350-1371.
- (16) Liu, H. N.; Gauld, J. W. *J. Phys. Chem. B* **2008**, *112*, 16874-16882.
- (17) Guth, E.; Connolly, S. H.; Bovee, M.; Francklyn, C. S. *Biochemistry* **2005**, *44*, 3785-3794.
- (18) Dock-Bregeon, A. C.; Sankaranarayanan, R.; Romby, P.; Caillet, J.; Springer, M.; Rees, B.; Francklyn, C. S.; Ehresmann, C.; Moras, D. *Cell* **2000**, *103*, 877-884.
- (19) Sankaranarayanan, R.; Dock-Bregeon, A. C.; Rees, B.; Bovee, M.; Caillet, J.; Romby, P.; Francklyn, C. S.; Moras, D. *Nat. Struct. Biol.* **2000**, *7*, 461-465.
- (20) Sankaranarayanan, R.; Dock-Bregeon, A. C.; Romby, P.; Caillet, J.; Springer, M.; Rees, B.; Ehresmann, C.; Ehresmann, B.; Moras, D. *Cell* **1999**, *97*, 371-381.
- (21) Torres-Larios, A.; Sankaranarayanan, R.; Rees, B.; Dock-Bregeon, A. C.; Moras, D. *J. Mol. Biol.* **2003**, *331*, 201-211.
- (22) Llano, J., Gauld, J. W. *Mechanistics of Enzyme Catalysis: From Small to Large Active-Site Models*, In *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta, C. F., Ed.; Wiley-VCH: Weinheim, 2010; Vol. 2, p 920.
- (23) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307-327.
- (24) Bharatham, N.; Bharatham, K.; Lee, Y.; Lee, K. W. *Biophys. Chem.* **2009**, *143*, 34-43.
- (25) Budiman, M. E.; Knaggs, M. H.; Fetrow, J. S.; Alexander, R. W. *Proteins* **2007**, *68*, 670-689.
- (26) Hansia, P.; Ghosh, A.; Vishveshwara, S. *Mol. Biosyst.* **2009**, *5*, 1860-1872.
- (27) Hughes, S. J.; Tanner, J. A.; Miller, A. D.; Gould, I. R. *Proteins* **2006**, *62*, 649-662.
- (28) Kapustina, M.; Carter, C. W. *J. Mol. Biol.* **2006**, *362*, 1159-1180.

- (29) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6620-6625.
- (30) Thompson, D.; Lazenec, C.; Plateau, P.; Simonson, T. *Proteins* **2008**, *71*, 1450-1460.
- (31) Yamasaki, S.; Nakamura, S.; Terada, T.; Shimizu, K. *Biophys. J.* **2007**, *92*, 192-200.
- (32) Molecular Operating Environment; 2009.10 ed.; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2009.
- (33) Sekine, S.; Nureki, O.; Dubois, D. Y.; Bernier, S.; Chenevert, R.; Lapointe, J.; Vassilyev, D. G.; Yokoyama, S. *Embo J.* **2003**, *22*, 676-688.
- (34) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. *J. Comput. Phys.* **1999**, *151*, 114-134.
- (35) Nelson, D. L.; Cox, M. M. *Lehninger: Principles of Biochemistry*; 4th ed.; W.H, Freeman and Company: New York, 2005.
- (36) Bushnell, E. A. C.; Erdtman, E.; Llano, J.; Eriksson, L. A.; Gault, J. W. *J. Comput. Chem.* **2011**, *32*, 822-834.
- (37) Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114*, 8701-8712.
- (38) Huang, W. J.; Bushnell, E. A. C.; Francklyn, C. S.; Gault, J. W. *J. Phys. Chem. A* **2011**, *115*, 13050-13060.

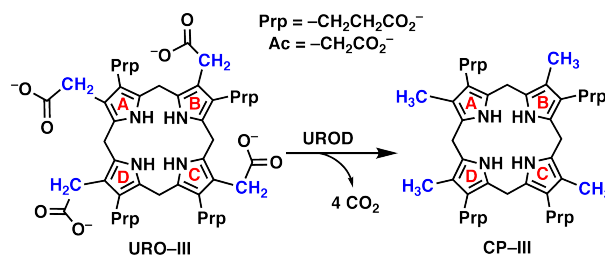
Chapter 4

The First Branching Point in Porphyrin Biosynthesis: A Systematic Docking, MD and QM/MM Study of Substrate Binding and Mechanism of UROD



4.1 Introduction

Porphyrin is a prosthetic group that enables many proteins to perform their roles in enzymatic catalysis, ligand transport, electron transfer and light harvesting¹⁻⁴. The porphyrin ring is synthesized in a multi-stage multi-enzymatic pathway. Specifically, the fifth stage of this pathway involves the enzyme uroporphyrinogen decarboxylase (UROD), which catalyzes the first branching point in the biosynthesis of tetrapyrrole molecules: the sequential non-symmetric decarboxylation of the four acetates of both uroporphyrinogen III (URO-III) and uroporphyrinogen I (URO-I) to give coproporphyrinogen III (CP-III) and coproporphyrinogen I (CP-I), respectively^{1,2,4,5}. It should be noted that only the asymmetric CP-III, and not the C_4 -symmetric CP-I isomer, is a viable precursor to the metabolically functional intermediate protoporphyrin IX. The overall chemical equation for the sequential decarboxylation of URO-III is given in **Scheme 4.1**.



Scheme 4.1. Overall sequential decarboxylation reaction catalyzed by UROD: the acetate of ring **D** is decarboxylated first followed by those of rings **A**, **B** and **C**^{1,2,4,5}.

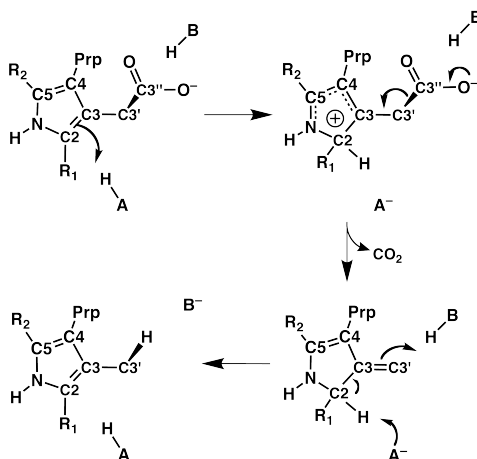
Human UROD (hUROD) exists as a homodimer^{1,2,4,6-8}. Due to the non-symmetric relationship between the acetates of URO-III, it was believed that the two active sites formed an extended cleft that enabled reaction intermediates to shuttle between them during the course of reaction^{2,4,8}. However, it is now more generally accepted that all four decarboxylations occur within a single active site^{7,9-15}. For instance, in a kinetic study on

bovine hepatic UROD, Straka and Kushner¹³ obtained a Hill coefficient of approximately 1, which indicates that its two active sites are non-cooperative. In addition, stereospecific labeling experiments^{11,12} showed that the stereochemical configuration of all four acetate methylenes is retained during reaction, and hence, it was concluded that all four decarboxylations occur within a single active site and via the same mechanism. This has been further supported by mutagenesis studies⁹. In particular, the mutagenesis study of Phillips *et al.*¹⁴ found that the dimeric form of UROD likely only helps create a cleft large enough to hold the tetrapyrrole substrate during catalysis. The study also sustained that a shuttling of intermediates between active sites is unlikely to happen.

Three electrostatic regions have been identified within the active site of each UROD monomer itself: a negative, a polar-positive and a non-charged region¹. The latter is believed to provide a hydrophobic area that aids binding of the substrate's relatively non-polar tetrapyrrole core^{7,13}. Moreover, the negatively charged region contains an invariant aspartyl (hUROD: Asp86) that is thought to help orientate the URO-III substrate by binding its pyrrolic –NH– groups^{1,7}. This is supported by the observation that the replacement of Asp86 by asparagine (Asn86) led to almost complete loss of the enzyme's activity⁷. Consequently, the negatively charged region has also been proposed to play an essential catalytic role, possibly by stabilizing various mechanistic intermediates^{1,7}.

In contrast, the polar-positive region contains a number of residues whose roles are not clearly understood. For instance, it contains tyrosyl, histidyl, lysyl and several argyl residues, which have been proposed to interact with the substrate's carboxylates and thus help in substrate binding and recognition^{7,16}. In addition, several of these residues have been shown to be highly conserved (hUROD: Arg37, Arg41, Tyr164 and His339), and as a result, have been suggested as possibly being involved in catalysis². However, based on the findings of their mutagenesis studies, Wyckoff *et al.*⁹ and Phillips *et al.*⁷ concluded that His339 and Tyr164 are evidently *not* essential for catalysis. In fact, mutation of His339 had little to no effect on the rate at which the initial decarboxylation occurred,

i.e., that of ring **D**. Nevertheless, it did result in accumulation of the first mechanistic intermediate, i.e., that with methyl on ring **D** and acetates on rings **A**, **B** and **C**⁹. In contrast, it has been found that one or more active-site arginyl residues (hUROD: Arg37, Arg41 and Arg50) are catalytically essential^{1,16-19}. However, their exact mechanistic role remains unclear.



Scheme 4.2. Proposed 'general blueprint' acid–base mechanism for the UROD-catalyzed decarboxylation of the acetates of URO-III^{11,12,20}. HA and HB represent general acids.

Barnard and Akhtar^{11,12,20} have proposed the mechanistic 'general blueprint' for UROD enzymes given in **Scheme 4.2**. Specifically, in the first step, a general acid (HA) protonates the pyrrole's C2 centre to generate a delocalized carbocation. The resulting electronic rearrangement weakens the C3'–C3'' bond within the acetate group, which leads to its decarboxylation with formation of a C3=C3' double bond. This is followed by facile addition of a proton at C3' by a second general acid (HB) in concert with deprotonation of the C2–H group by A[−] to yield the final decarboxylated pyrrole. In addition, due to retention of configuration at the newly formed methyl group, an acidic residue within the active site is likely to initially interact with the leaving carboxyl and to subsequently protonate the C3' centre^{2,11,12}. Furthermore, due to the similar pH ranges of

UROD enzymes obtained from very different organisms, Juárez *et al.*²¹ have proposed that they all exploit similar decarboxylation mechanisms.

Several specific mechanisms based on the 'blueprint' idea have since been proposed^{2,7,17,22}. They can be classed according to two broad criteria: (i) the identity of residues assumed to be acting as the general acids HA and HB, and (ii) the binding configuration of the URO-III substrate in the active site.

For instance, both Martins *et al.*² and Lewis *et al.*¹⁷ have proposed mechanisms in which Asp86 acts as the initial general acid HA. However, while Martins *et al.*² pointed out to the active site residue Tyr64 acting as the second general acid HB, Lewis *et al.*¹⁷ proposed that it is instead Arg37. Martins *et al.*² based their mechanistic proposal on an enzyme–substrate complex devised by manually docking the substrate to the free active site of the enzyme crystal structure [PDB accession code: 1URO]. But the crystal structure of an enzyme–product complex reported later⁷ shed light into a substrate's binding mode that would be incompatible with the mechanism proposed by Martins *et al.*². It should be noted that Lewis *et al.*¹⁷ investigated the CP-I–enzyme crystal structure instead of the CP-III–enzyme complex, and CP-III and not CP-I is the precursor to the metabolically relevant intermediate protoporphyrin IX. In addition, several experimental studies have concluded that Asp86 is most likely in its ionized form, i.e., Asp86–COO[−]^{1,7}.

Based on an X-ray crystallographic structure of a CP-III–UROD complex, Phillips *et al.*⁷ suggested that the first mechanistic proton is instead likely donated from the solvent. This was justified by the lack of a suitably placed general acid HA within the active site, which were able to protonate the 2-position of pyrrole. Furthermore, the authors were unable to conclusively target a suitable candidate for the acid HB responsible for protonation of C3' (**Scheme 4.2**).

In another attempt to explore the mechanistic blueprint, Silva and Ramos²² performed a computational density functional study of an small chemical model (ca. 70 atoms)

devised from the UROD–product complex crystal structure obtained by Phillips *et al.*⁷. In particular, their chemical model did not explicitly include all active-site residues or constraints. In addition, the model does not take into account the substrate's orientation within the active site. Their results indicate that a catalytic mechanism as the one outlined in **Scheme 4.2** would indeed be thermodynamically feasible. Specifically, the authors found that the active-site residue Arg37 is both suitably positioned and capable of acting as the general acid HA that protonates the substrate's C2 centre. Furthermore, they proposed that the second proton, that from HB, is donated by the solvent (**Scheme 4.2**). However, the experimental evidence given by Barnard and Akhtar^{11,12} shows that the second proton must be donated by an active-site residue, and not by the solvent, so that the stereochemical configuration at C3' may be retained.

Clearly, our current understanding of the catalytic mechanisms of the UROD class of enzymes is insufficient and incomplete. Hence, the aim of this work is to explore and find the catalytically relevant binding modes of the URO-III substrate within the active site of UROD by applying flexible docking in conjunction with a force field scoring function and a subsequent molecular dynamics annealing of the candidate enzyme–substrate complexes. Furthermore, the enzyme–substrate complex with the largest binding energy is chosen to build a large active-site chemical model of hUROD. Then, the enzymatic mechanism of the first decarboxylation of URO-III, i.e., that of the acetate on ring **D**, is investigated through combined quantum mechanical and molecular mechanical methods in the ONIOM formalism.

4.2 Computational Methods

The Molecular Operating Environment (MOE)²³ software package was used to perform molecular docking and molecular dynamics (MD) simulations of the URO-III

substrate within the active site of UROD. These calculations were done with the CHARMM22 force field²⁴.

Molecular Docking. The initial structure of UROD was extracted from the crystallographic structural model of hUROD complexed with CP-III (PDB accession code: 1R3Y) by removing all the coordinates from the crystallographic waters, counterions and CP-III. The coordinates of hydrogens were added using the MOE default method. All residues within 7 Å of the CP-III molecule were designated as belonging to the active site. Then, URO-III was placed in the active site using the proxy triangle method, which is designed to dock large polyatomic multi-conformational ligands. The binding free energy of each enzyme–substrate complex generated by this procedure was estimated with the London dG scoring function. Then, the geometries of the top one hundred complexes were optimized using the Forcefield refinement scheme in conjunction with the CHARMM22 force field. From this set, the final top thirty complexes with the lowest CHARMM22 energies were selected for further analyses, and their binding free energies were recalculated with the London dG scoring function.

Of these 30 structures, only a small subset was found to also have the mechanistically required interaction between the leaving carboxylate of the acetate on ring **D** and a potentially acidic active-site residue. This subset was retained and was itself divided into three groups based on the identity of the potential acidic residue coordinated to the acetate on ring **D**. Namely, those in which the residue is: (i) Arg50, (ii) Tyr164 or (iii) His339. For each of these, the complex with the largest free energy within its group was selected. Finally, these three complexes are hereafter referred to as enzyme–substrate complex **I**, **II** and **III**, respectively.

Molecular Dynamics Equilibration. MD simulations were performed to allow the solvated enzyme–substrate complexes undergo thermal relaxation using a protocol as follows. Each enzyme–substrate complex was surrounded with a 7-Å spherical layer of water molecules. An ellipsoidal potential wall with a scaling constant of 2 was placed

around the solvated enzyme–substrate complex, in order to force the system to lie within the volume of space defined by the ellipsoid. The damping functional factor included in the electrostatic and van der Waals potentials were set to decay smoothly beyond 8 to 10 Å. The geometry of each solvated complex was then optimized using the CHARMM22 force field until the root mean square gradient of the total energy fell below 0.05 kcal a.u.⁻¹. The MD simulations were performed under constrained pressure and temperature. The equations of motion were coupled with the Nosé–Poincaré thermostat²⁵ and the time step for numerical integration was set to 1 fs. Initially, the system was heated from 150 to 300 K for a period of 50 ps, followed by an equilibration period of 100 ps at the constant temperature of 300 K and pressure of 1 atm. A typical structure from the trajectory was then optimized with the CHARMM22 force field for the final analyses.

QM/MM calculations. Combined quantum mechanical and molecular mechanical (QM/MM) methods in the ONIOM formalism with mechanical embedding²⁶⁻³⁴ were applied as implemented in the *Gaussian03* program suite³⁵. The stationary points of the potential energy surface (PES) were located using a two-layer ONIOM model consisting of a QM layer that combined the density functional method B3LYP³⁶⁻³⁸ with the 6-31G(d) basis set. The MM layer was described with the AMBER94 force field³⁹. Frequency analyses of all stationary points were done at the same level of theory, i.e., ONIOM(B3LYP/6-31G(d):AMBER94), in order to characterize minima and transition states and to calculate zero-point vibrational energy and Gibbs energy corrections at SATP. Single-point energy calculations on the optimized structures were done at the ONIOM(B3LYP/6-311+G(2df,p):AMBER94) level of theory.

We used a large active-site chemical model to investigate the reaction mechanism, as illustrated in **Figure 4.1**. It included URO-III and all active site residues immediately surrounding it, i.e., first-shell residues. In addition, for those portions of the substrate exposed to solvent, the first solvation shell was retained. It should be noted that the α -carbons of each residue were held fixed at their final MM minimized positions in order to

ensure integrity of the active site during the calculations. Such an approach has been commonly used in the computational investigation of the catalytic mechanisms of enzymes, and its applicability and reliability has been discussed in detail elsewhere^{40,41}. A subset of the complete model, centered on the reactive region of the active site was then selected for the high-level QM treatment. Specifically, it consisted of that component of the substrate and those residues directly involved in the first decarboxylation reaction. That is, the model contained the pyrrole ring **D** and its acetate group of URO-III and the side chains of Arg37 and Asp86. In addition, it also included the side chains of the possible second mechanistic acid Arg50 (**Figure 4.1b**).

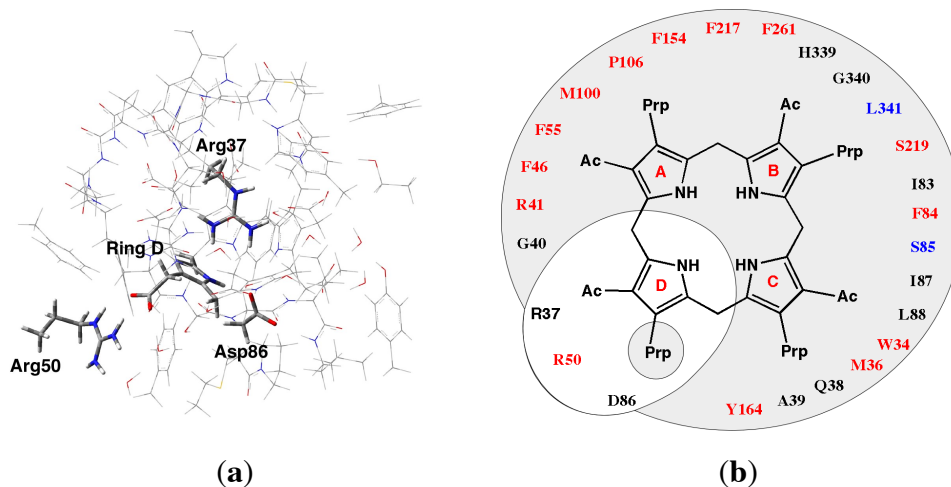


Figure 4.1. Structural models of the active site: **(a)** Arrangement of the catalytically active residues of UROD in complex I. **(b)** QM/MM model with residues in QM (non-shaded region) and MM (shaded region) layers. Residues in red modeled as side chains with atoms fixed at the truncation position. Remainder of residues modeled as side chain and backbone with C α positions fixed. Residues in blue represent side chains modeled as hydrogen (S85 and L341).

The general affects of the polarity of the surrounding environment were modeled using the IEF-PCM approach with dielectric constants (ϵ) of 4, 10 and 78.39. The two

former values have been commonly used previously to model the polarity found within a protein⁴⁰, while the latter is that of water.

4.3 Results and Discussion

Binding of the URO-III substrate to the UROD active site. In all three optimized enzyme–substrate complexes **I**, **II** and **III**, the URO-III substrate binds to the active site by its pyrrole –NH– groups that hydrogen-bond to the carboxylate of Asp86. In addition, the active-site residue Arg37 positions itself above the tetrapyrrole rings and forms several arene–cation interactions. Both of these active-site–substrate interactions are compatible with the X-ray crystallographic model of the UROD–CP-III complex previously obtained by Phillips *et al.*⁷.

The key hydrogen-bonding interactions of URO-III with the active-site residues and solvent molecules are schematically represented in **Figure 4.2**. It is clearly noted that URO-III displays three distinct binding modes in the active site of UROD, in which the mechanistically key acetate of ring **D**, the first to be decarboxylated by UROD, interacts with a different active-site residue.

UROD–URO-III Complex I. This complex shows the strongest binding of the URO-III substrate to the enzyme, which amounts to $-246.3 \text{ kcal mol}^{-1}$. In this case, Arg50 is the only active-site residue that interacts directly with the acetate on ring **D** (**D**–Ac[−]). Specifically, the guanidinium forms two relatively short hydrogen bonds with the acetate with lengths of 1.59 and 1.67 Å. In addition, the acetate also accepts a hydrogen bond from water. Furthermore, the adjacent propionate on ring **D** (**D**–Prp[−]) accepts hydrogen bonds from both a backbone –NH– and the Arg41 guanidinium ion with lengths of 1.78 and 1.77 Å respectively.

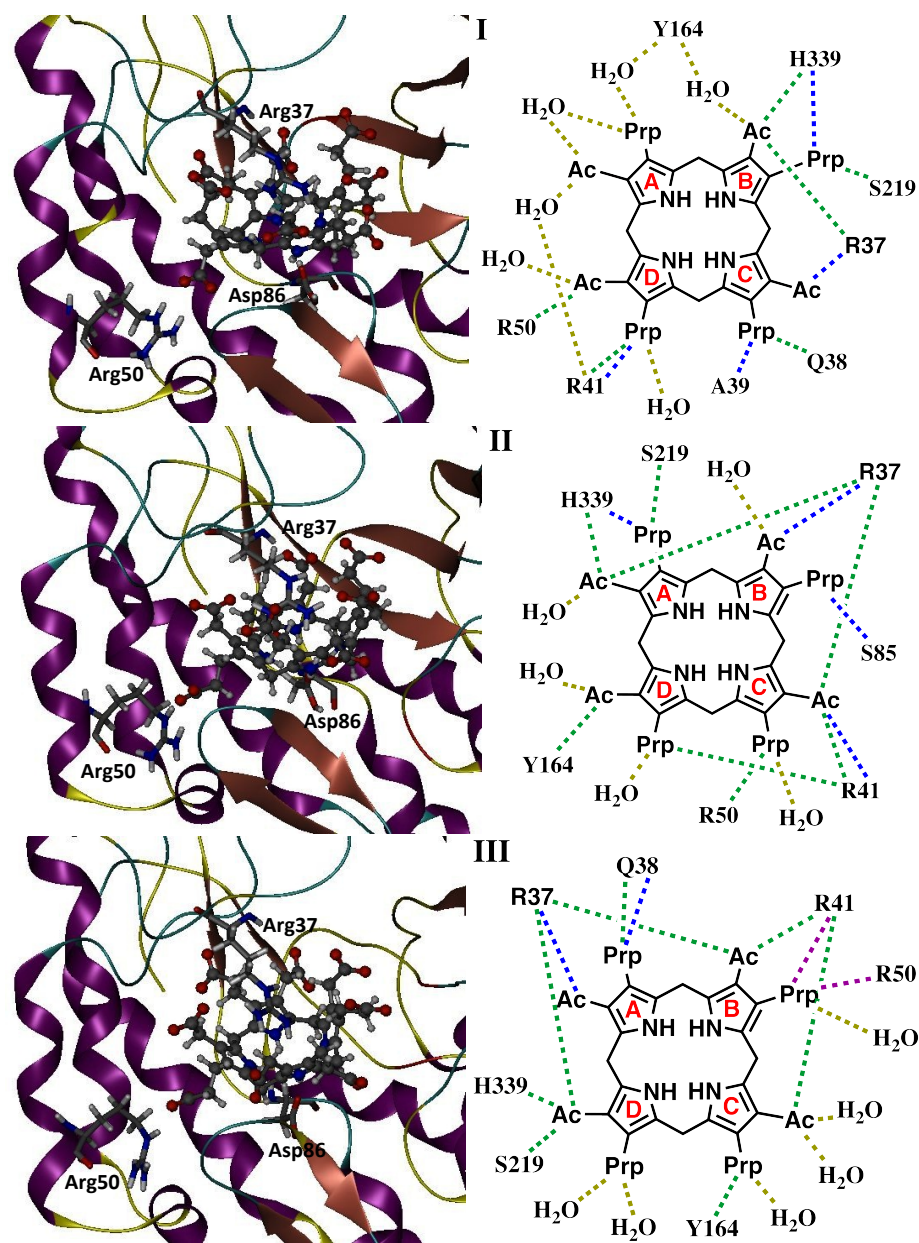


Figure 4.2. Schematic representation of the binding modes of URO-III in the active site of UROD. Hydrogen-bonding interactions are encoded as follows: green dashed line, bonding with the side chain R; blue dashed line, bonding with backbone amide; olive-green dashed line, bonding with water; magenta dashed line, salt bridge.

Unlike in the enzyme–substrate complexes **II** and **III**, the carboxylates on ring **A** are both entirely exposed to the solvent and form no direct hydrogen bonds with the active-

site residues. In contrast, the carboxylates on ring **C** only hydrogen-bond with active-site residues, namely Arg37, Ala39 and Gln38. Moreover, the carboxylates on ring **B** hydrogen-bond directly with several nearby residues (His339, Ser219 and Arg37) and via an H₂O with the hydroxyl of Tyr164.

UROD–URO-III Complex II. In this case, URO-III binds to the enzyme with a free energy of $-231.8 \text{ kcal mol}^{-1}$, which is $14.5 \text{ kcal mol}^{-1}$ higher than that for complex **I**. Interestingly, complex **II** shows a larger number of enzyme–substrate hydrogen-bonding interactions than complex **I**. In particular, the **D**–Ac[−] group is now hydrogen-bonded to the hydroxyl of Tyr164 at 1.63 \AA and to a water molecule (**Figure 4.1**). However, Arg50, which is the potential candidate for HB in complex **I**, simply assists the URO-III binding via hydrogen bonds to **C**–Prp[−]. Furthermore, the analogous potential candidate for HB in conformer **III**, i.e., His339, is involved in binding both carboxylates on ring **A** in conjunction with Ser219 and Arg37.

In complex **II**, unlike in **I** and **III**, every ring forms hydrogen bonds with at least one water molecule. However, complex **II** has only five hydrogen bonds of URO-III with water: two involving ring **D** and the remainder involving each of the other rings. Moreover, every carboxylate of URO-III forms at least one direct hydrogen bond with either a side chain or an amide backbone group or both in the active site.

UROD–URO-III Complex III. This complex is characterized by the weakest binding, with a free energy that amounts to $-223.5 \text{ kcal mol}^{-1}$. This may be due to a disruption of the arrangement of the active-site residues caused by a secondary-structure transition from β -strand to coil near the active site (**Figure 4.2**). In complex **III**, the key **D**–Ac[−] group is an acceptor of several hydrogen bonds from active-site residues. More specifically, it forms a short hydrogen bond (1.58 \AA) with the protonated imidazole of His339 and two markedly longer hydrogen bonds with the guanidinium of Arg37 and hydroxyl of Ser219 at distances of 1.78 and 1.77 \AA , respectively. The potential mechanistic acidic residues in complexes **I** (i.e., Arg50) and **II** (i.e., Tyr164) are now

involved in binding URO-III via the propionates of rings **B** and **C**, respectively. Moreover Arg50 does not hydrogen-bond to the propionate. Rather, the guanidinium groups of Arg50 and Arg41 stack on either side of the propionate, forming a positive electrostatic region in which the carboxylate sits.

Identifying possible mechanistic acids. According to the 'blueprint' mechanism outlined in **Scheme 4.2**^{11,12,20}, decarboxylation is initiated by protonation of the 2-position of pyrrole by some presumed and unidentified acid HA that should be available in the vicinity of ring **D**. Thereafter, a second acid HB should protonate the =C3' centre resulting from acetate decarboxylation.

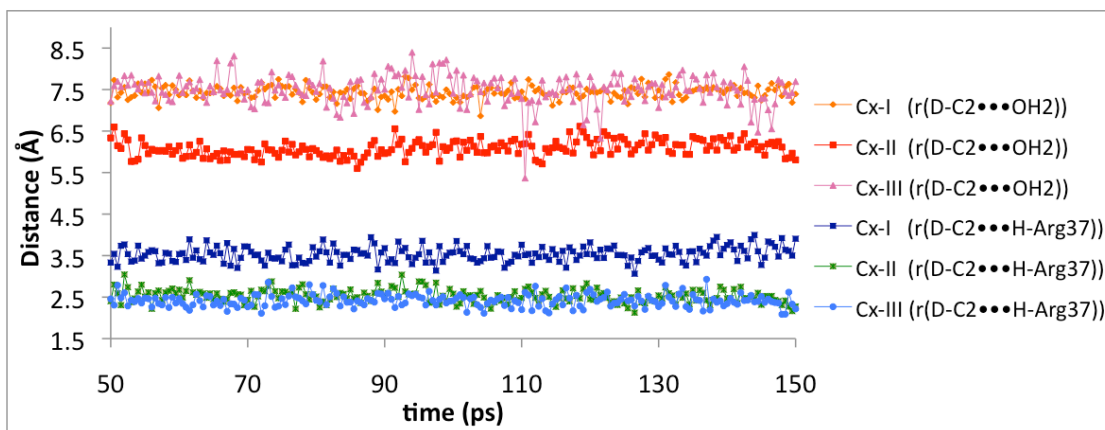


Figure 4.3. Distance fluctuations between C2 of ring **D** (**D**–C2) and both the nearest water oxygen [$r(\mathbf{D}\text{--C2}\cdots\text{OH}_2)$] and the closest guanidinium proton on Arg37 [$r(\mathbf{D}\text{--C2}\cdots\text{H}\text{--Arg37})$] in in UROD–URO-III complexes **I**, **II** and **III** (Cx-**I**, Cx-**II**, Cx-**III**) over 100 ps of MD equilibration.

It has been previously suggested that the initial proton may in fact originate from the solvent⁷. We investigated this possibility by tracking the distance fluctuations between C2 of ring **D** (**D**–C2) and the nearest water [$r(\mathbf{D}\text{--C2}\cdots\text{OH}_2)$] in each UROD–URO-III complex over 100 ps of MD equilibration. First of all, it is noted that $r(\mathbf{D}\text{--C2}\cdots\text{OH}_2)$ in all three complexes ranges from 5.60 to 8.40 Å (**Figure 4.3**). These distances are clearly

significantly longer than those typical O \cdots O and O \cdots N hydrogen bonds, i.e., 1.5 – 2.5 Å, and thus, it seems unlikely that the initial proton could be donated from the aqueous solvent.

Nonetheless, Arg37 may in fact act as the initial acid HA because of its positioning above the tetrapyrrole rings²². Hence, we have also examined the distance fluctuations between the closest guanidinium proton on Arg37 and D–C2 [$r(\text{D-C2}\cdots\text{H}^+\text{Arg37})$] in each UROD–URO-III complex over the same 100 ps of MD equilibration (**Figure 4.3**). It is found that these distances range from 2.0 to 4.0 Å, and thus, are all significantly shorter than those involving water. In fact, in each complex the Arg37 guanidinium group was found to be the closest source of protons available to D–C2. For instance, the shortest average distances between an Arg37 guanidinium proton and D–C2 are 3.54, 2.53 and 2.41 Å for complexes **I**, **II** and **III**, respectively.

The idea that arginine may act as an acid/base has been previously proposed for several other enzymes⁴²⁻⁴⁶. However, Lewis and Wolfenden¹⁷ have experimentally examined the effects of pH on the non-enzymatic aqueous decarboxylation of pyrrole-3-acetate. They estimated that the pK_a of C2 in pyrrole-3-acetate is –3.5. Since this pK_a is considerably smaller than the pK_a of guanidinium in aqueous arginine (~13.5), they concluded that it was unlikely that Arg37 would be able to act as the initial mechanistic acid in UROD. However, the local electrostatic environment surrounding the side chain of a residue can significantly change its pK_a ⁴⁷.

Hence, we investigated the influence of relevant local interactions of functional groups with arginine within the active site of UROD and their influence on basicity of the residue's side chain. In particular, we examined the effects of varying the polarity of the medium on the basicities of the guanidinium group of arginine (modeled as N-ethylguanidino), the C2 centres of the initial substrate and final product (modeled as 3-acetopyrrole and 3-methylpyrrole, respectively), and the C2 and C3' centres of the proposed decarboxylated intermediate (modeled as a 3-methylenepyrrole derivative).

Furthermore, we also included the influence of hydrogen bonding with the side-chain carboxyl of Asp86 (modeled as acetate) and the $-NH-$ group of pyrrole (**Figure 4.4**).

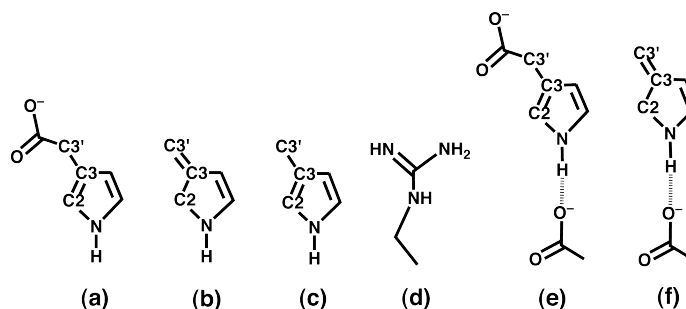


Figure 4.4. The models of (a) 3-acetopyrrole and proposed intermediate models (b) 3-methylenepyrrole derivative (c) 3-methylpyrrole (d) N-ethylguanidino, (e) 3-acetopyrrole coordinated to an acetate and (f) 3-methylenepyrrole derivative coordinated to an acetate.

The proton affinities (PA's) were estimated as the negative change in the internal energy of protonation of the side chain at 0 K with the proton coming from the vacuum state, i.e., $PA = -[E_0(BH^+) - E_0(B) - E_0(H^+)]$, where $E_0(H^+) = 0$. The basicities were estimated as the negative change in the internal energy of protonation of the side chain at 0 K with the proton coming from an aqueous ideal-dilute solution at pH 7. The standard energy of bulk solvation of the proton in aqueous solution was taken to be -262.40 kcal mol $^{-1}$ because this value is the only one consistent with the Born–Haber-type cycles of both hydrogen and the electron.^{48,49} That is, $Basicity = PA - 262.40$ kcal mol $^{-1}$.

We began by first examining the basicities relating to the first protonation reaction, namely the basicities of the pyrrole ring C2 and the arginine's guanidinium **Figure 4.4**. In a continuum medium with the dielectric constant of water ($\epsilon = 78.39$), the basicity of N-ethylguanidino is 42.4 kcal mol $^{-1}$, which is larger than the basicity of the C2 centre in either 3-acetopyrrole itself (19.8 kcal mol $^{-1}$), or when it is hydrogen-bonded with the acetate via the $-NH-$ group of pyrrole (31.8 kcal mol $^{-1}$). In addition, it is also larger than

that of the decarboxylated pyrrole derivative 3-methylpyrrole ($16.5 \text{ kcal mol}^{-1}$), the product model.

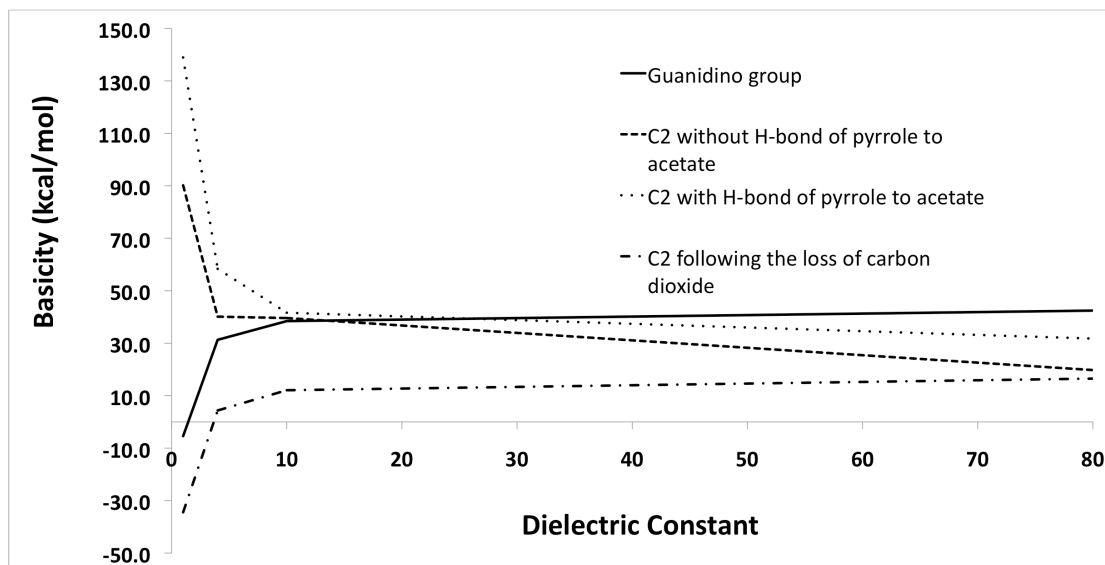


Figure 4.5. Effect of medium polarity on proton affinities of C2 in 3-acetopyrrole, with and without hydrogen bond to acetate, 3-methylenepyrrole and N-ethylguanidino.

However, as the medium polarity (ϵ) decreases from that of water to $\epsilon = 10$, the basicity of Arg decreases slightly to $38.4 \text{ kcal mol}^{-1}$. In contrast, the basicity of C2 in both 3-acetopyrrole without and with hydrogen bonding to acetate via its $-\text{NH}-$ group increases to 39.6 and $41.6 \text{ kcal mol}^{-1}$, respectively. Consequently, now they are both larger than that of arginine. Further reduction of the medium dielectric constant to 4 and subsequently to 1 (which represents the gas phase at SATP) causes the basicity of Arg to decrease even more markedly to 31.3 and $-5.4 \text{ kcal mol}^{-1}$, respectively. However, the basicity of C2 in both 3-acetopyrrole with and without hydrogen bonding with acetate shows the opposite trend, i.e., rising more rapidly as the dielectric constant decreases (**Figure 4.5**). Notably, it is generally held that the polarity of an enzyme active site typically lies in the range of $\epsilon = 4 - 10$ ^{40,47}. Thus, the above results suggest that when an arginine residue is placed within hUROD's active site, its basicity decreases such that it is

better able to act as an acid. Furthermore, the combined effect of low polarity of the active site and the hydrogen bonding via the pyrrolic –NH– group of ring **D** to Asp86 increases the basicity of C2 in ring **D** so that it may accept a proton from the guanidinium cation of Arg37. It also noted that the basicity of C2 in 3-methylpyrrole (i.e., our model of final product) is significantly lower (average of 27.1 kcal mol⁻¹) than that of the N-ethylguanidino group (**Figure 4.5**).

It has previously been suggested that the observed retention of stereochemistry at the C3' centre after decarboxylation, indicates that the second mechanistic acid must also be one that initially interacts with **D**-Ac⁻ ^{11,12}. Considering the structure of the complex **I**, which is the preferred substrate binding mode, then Arg50 is a plausible candidate to act as the acid in the first step of the catalytic mechanism. However, the ability of Arg50 to fulfill this role depends on the basicity of its guanidino group as well as that of the protonation site of the mechanistic intermediate modeled as 3-methylenepyrrole, i.e., its C3' centre (see **Figure 4.2**).

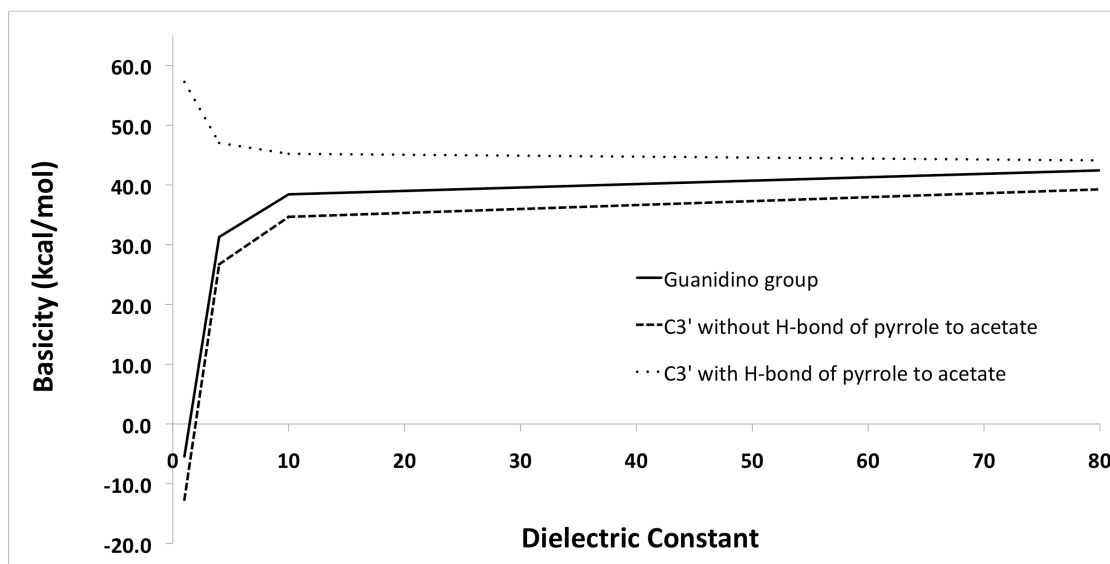


Figure 4.6. Effect of medium polarity on the basicities of the N-ethylguanidino group and the pyrrolic C3' centre, with and without hydrogen bonding with acetate.

As seen in **Figure 4.6**, without an acetate hydrogen-bonded to the pyrrolic $-\text{NH}-$ group, the basicity of C3' is slightly reduced by 3 – 7 kcal mol⁻¹ with respect to that of the N-ethylguanidino for the dielectric constants used. However, when the acetate is hydrogen-bonded to the pyrrolic $-\text{NH}-$ group, the basicity of C3' becomes larger than that of the N-ethylguanidino for any dielectric constant considered. In fact, the basicity of C3' increases as ϵ decreases. This suggests that Arg50 could in fact act as the second required mechanistic acid, and furthermore, that one role of Asp86 is to enhance the basicity of C3' in key mechanistic intermediates.

As noted above, we have also examined the basicity of C2 in 3-methylpyrrole (**Figure 4.3**) in order to determine the ability of the final proposed mechanistic intermediate formed after decarboxylation and protonation of the C3' centre to donate a proton back from its $-\text{C}2\text{H}_2-$ group to the guanidino group of Arg37. It is found that for any dielectric constant, the basicity of C2 in 3-methylpyrrole is markedly lower than that of the N-ethylguanidino by ~ 27 kcal mol⁻¹ (**Figure 4.5**). Hence, once the initial substrate URO-III has been decarboxylated and the C3' centre protonated by a nearby acid, the guanidino group of Arg37 would easily be able to deprotonate the $-\text{C}2\text{H}_2-$ group. It is expected that the basicity of C2 after the loss of CO₂ would be lower than that of N-ethylguanidino regardless of whether Asp86 was hydrogen-bonded to the pyrrole amine. This is because, on average, the basicity of C2 (prior to the loss of CO₂) and C3' increased by ~ 23 kcal mol⁻¹, which would not be a significantly enough change in the case of the basicity of C2 after the loss of CO₂.

QM/MM Investigation of the Catalytic Mechanism of hUROD. Following the above binding and proton affinity studies, we then examined the mechanism for the first decarboxylation as catalyzed by hUROD, i.e., that of the acetate on ring **D** in complex **I**. The chemical model (**Figure 4.1**) devised for this part of the study was based on our findings that: (i) the binding of the URO-III substrate to the active site of hUROD is the strongest in complex **I**, (ii) previous experimental studies have found that both Tyr164

and His339 are not catalytically essential^{7,9}, and (iii) the calculated basicities on the small models indicate that both Arg37 and Arg50 could act as the two mechanistically required general acids HA and HB (cf. **Scheme 4.2**).

The Initial Proton Transfer. The first step in the proposed mechanism is a proton transfer from an acid to the C2 centre of ring **D** (**Scheme 4.2**). Hence, we considered the ability of Arg37 to act as the initial acid. The calculated free energy profiles of reaction are shown in **Figure 4.7** and the Cartesian coordinates of the fully optimized structures are given in **Table A1** of the Appendix.

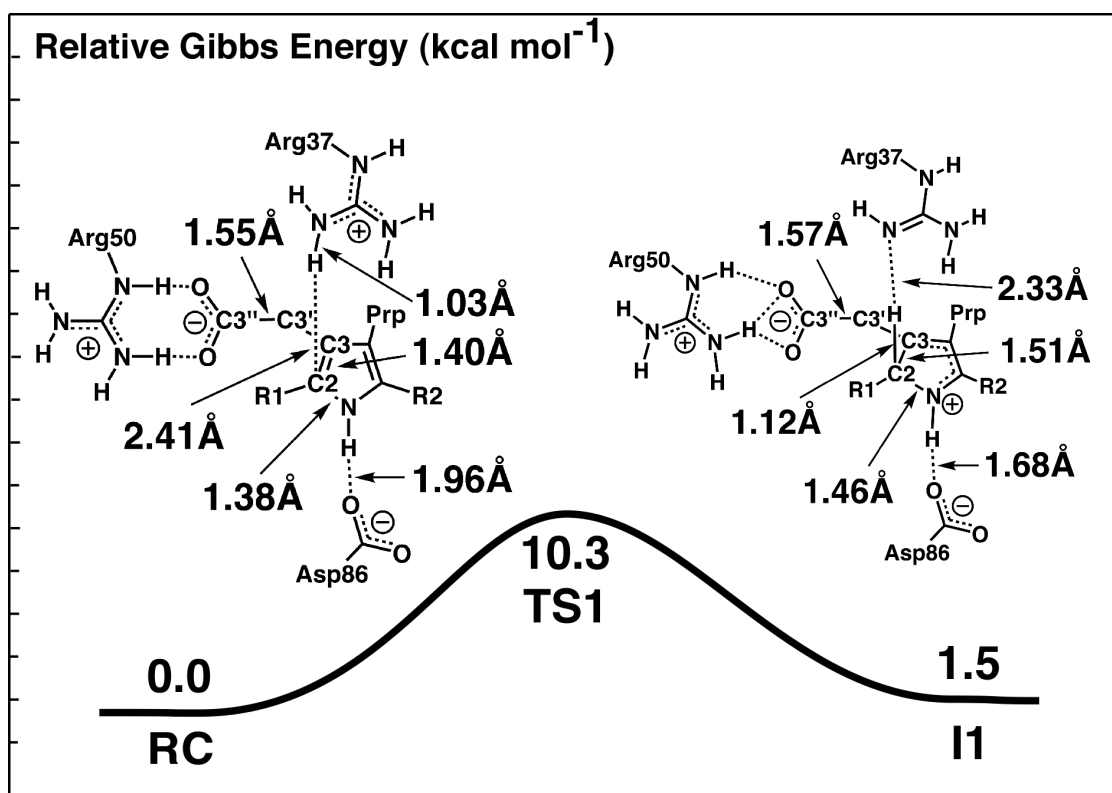


Figure 4.7. Free energy diagram for proton transfer from the guanidinium of Arg37 to the C2 centre of ring **D**.

In the initial reactant complex (**RC**), the nearest proton of Arg37 guanidinium cation is 2.41 Å apart from C2 in ring **D**. In addition, the pyrrolic NH group in ring **D** and

carboxylate of Asp86 form a moderately strong hydrogen bond, $\mathbf{D-NH}\cdots\text{OOC-Asp86}$, of length 1.96 Å with an almost linear $\text{N-H}\cdots\text{O}$ angle of 176.6°. It is noted that the optimized $\mathbf{D-NH}\cdots\text{O}$ distances in **RC** are in good agreement with the corresponding crystallographic distances⁷. Furthermore, the $\text{C3}'\text{-C3}''$ bond has a length of 1.55 Å, typical for a C-C single bond, while the C3-C2 and C2-N-D bonds have lengths of 1.40 Å and 1.38 Å, respectively.

Transfer of the nearest guanidinium proton of Arg37 to the substrate occurs via **TS1** at a cost of 10.3 kcal mol⁻¹ (**Figure 4.7**). The resulting intermediate **II** in which the Arg37 guanidinium is now neutral while the pyrrole is now formally protonated at the C2 position lies just marginally higher in energy than the initial **RC** by just 0.9 kcal mol⁻¹. A barrier of 10.3 kcal mol⁻¹ seems low in comparison to the barrier of 18.5 kcal mol⁻¹ obtained from the experimentally reported k_{cat} of 0.16 s⁻¹^{15,17,40}. However, as summarized by Juarez *et al.*²¹ decarboxylation of URO-III generating the 7-carboxylate intermediate is most likely not the rate-limiting step, and in fact, for several variants of UROD, the rate-limiting step appears to be the decarboxylation of the 7-carboxylate intermediate.

Previously, it has been suggested that a possible limitation of a QM/MM-based approach is that the dynamic behaviour of the enzyme is not fully taken into account.^{50,51} However, Zhang *et al.*⁵⁰ found that regardless of the variations in the protein environment, the role of the respective groups involved in the catalysis are likely to be very consistent. It is noted that for UROD from various species, the barriers to decarboxylation of ring **D** estimated experimentally range from 2.0 to 12.3 kcal mol⁻¹^{16,21}. Thus, our calculated barrier lies within the range of values reported from experiment.

In **II**, the resulting ring has considerably lost conjugation. This is illustrated by the fact that C2-N-D and C2-C3 bonds have markedly lengthened to 1.46 Å and 1.51 Å, respectively. That is, they now have significantly reduced double-bond character. Furthermore, the $\text{C3}'\text{-C3}''$ bond has lengthened slightly to 1.57 Å. Thus, in agreement with that previously proposed^{11,12}, protonation at the pyrrole C2 centre does appear to

weaken the C–C bond with the acetate group. In addition, the D–NH \cdots ⁻OOC–Asp86 hydrogen bond shortens dramatically to 1.68 Å. In addition, for the newly formed C2–H, the bond length was calculated to be 1.12 Å.

Acetate decarboxylation. In agreement with the mechanism proposed by Barnard and Akhtar,^{11,12,20} the next step is found to be decarboxylation of the acetate moiety (**Scheme 4.2**). However, in contrast to that previously suggested, this is found to effectively occur simultaneously with proton transfer from the guanidinium group of Arg50 to the C3' centre of the substrate. The free energy profile for this step is illustrated in **Figure 4.8**.

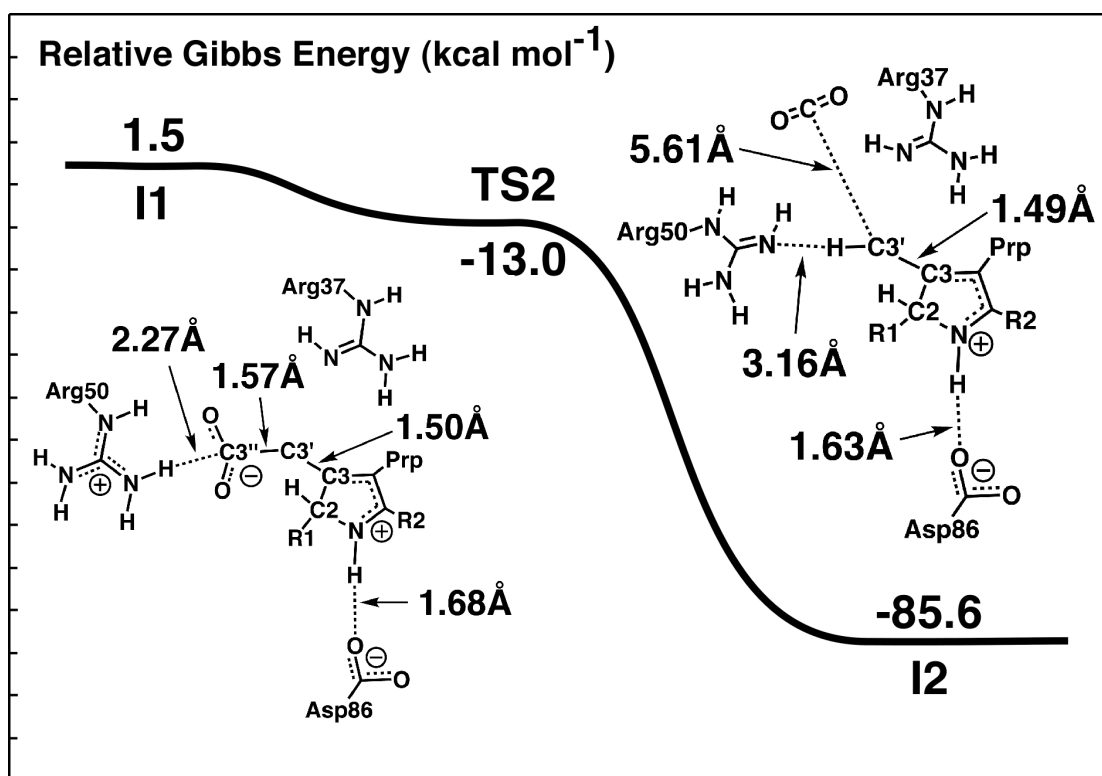


Figure 4.8. Free energy profile for the decarboxylation of ring **D**.

Specifically, this catalytic step occurs via **TS2** with a relative energy barrier *lower* than that of **I1**. This indicates that at SATP, the loss of CO₂ occurs without a barrier. The optimized structure of **TS2** is shown in **Figure 4.9**. As can be seen, the cleaving O₂C3'...C3' distance has increased markedly in length to 2.64 Å, while the angle between

oxygens of the leaving CO_2 has increased to 163.5° . Furthermore, the $\text{C3}'\text{-C3}$ bond gains marked double-bond character, as indicated by its now considerably reduced length of 1.37 \AA . Simultaneously, however, the $\text{Arg50-NH}\cdots\text{C3}''$ and $\text{Arg50-NH}\cdots\text{C3}'$ distances are 2.77 and 4.08 \AA , respectively. This also suggests that the Arg50 residue shifts such that there is greater room for the leaving CO_2 molecule. Thus, while decarboxylation and protonation of the $\text{C3}'$ centre do occur in one step, the optimized structure of **TS2** indicates that the $\text{C3}''\text{-C3}'$ bond is essentially cleaved prior to proton transfer from Arg50 to the $\text{C3}'$ centre.

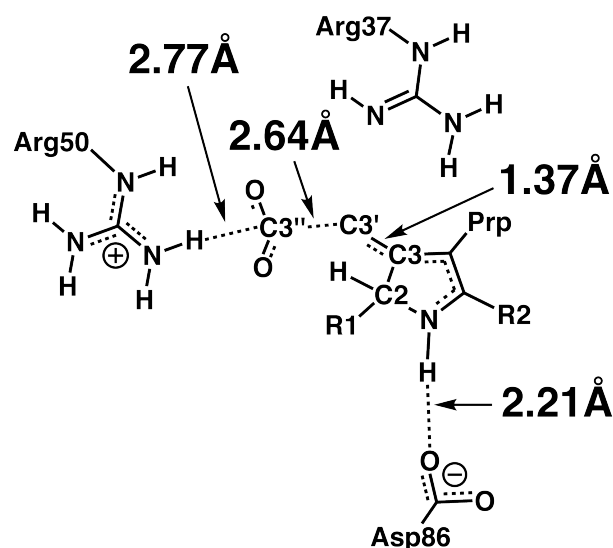


Figure 4.9. Optimized structure of **TS2**, the transition structure for the decarboxylation of ring **D** with concomitant proton transfer from Arg50 to the $\text{C3}'$ centre. For clarity, not all hydrogens are shown.

Applying DFT computational methods in combination with smaller chemical models, it was previously concluded that the presence of Asp86 and its hydrogen bond to the pyrrole's ring amine actually hindered the release of CO_2 .²² In contrast, we find that during the decarboxylation step the $\text{D-NH}\cdots\text{OOC-Asp86}$ hydrogen bond lengthens significantly by 0.53 \AA to 2.21 \AA in **TS2**. This reduced $\text{Asp86}\cdots\text{pyrrole}$ carbocation

interaction causes a destabilization of the carbocation, thus enhancing the feasibility of decarboxylation. Indeed, we note that the presently calculated barrier for decarboxylation, $\sim 13 \text{ kcal mol}^{-1}$ lower in energy than the reactive complex, is significantly lower than that previously estimated using the smaller chemical models, which was found to be $\sim 20 \text{ kcal mol}^{-1}$ higher in energy than the reactive complex²².

The resulting decarboxylated intermediate formed, **I2**, lies $85.6 \text{ kcal mol}^{-1}$ lower in energy than the initial substrate bound–active site complex **RC**. Importantly, in **I2**, the stereochemistry of the C3' centre has been retained, in agreement with experimental observations^{11,12,20}. In addition, it should also be noted that the C3'–C3 bond has now lengthened to 1.49 \AA , typical of a single carbon–carbon bond. Intriguingly, during this reaction step Arg50 undergoes rotation about the C χ –C δ bond. As a result, rather than donating its proton to the C3' centre, which was initially hydrogen bonded to the substrate's acetate group, it alternatively transfers the other proton of the same amino group. Furthermore, the **D**–NH \cdots ⁻OOC–Asp86 hydrogen bond has again shortened markedly to 1.63 \AA . Finally, the cleaved CO₂ moves further away at an equilibrium distance between C3' and C3'' of 5.61 \AA .

Abstraction of a Proton by Arg37. The final catalytic step in the overall mechanism is regeneration of the protonated Arg37 residue by transfer of a proton from the substrate's –C2H₂– group to its neutral guanidino group. The Gibbs free energy profile calculated for this step is shown in **Figure 4.10** along with the optimized structures of **I2** and the final active site-bound product complex (**PC**) with selected bond lengths given.

In **I2** the guanidino group of Arg37 via its terminal imine weakly interacts with a proton on –C2H₂–, i.e., Arg37–N γ \cdots H⁺C2–**D**, at a distance of 2.23 \AA . The proton transfer from the –C2H₂– group to the Arg37 imine proceeds via **TS3** at a cost of just $3.1 \text{ kcal mol}^{-1}$. This low barrier is likely explained by the higher proton affinity of the Arg37 with respect to that of the decarboxylated intermediate **I2**, as previously discussed. The resulting final active-site-bound product complex **PC** lies 10 kcal mol^{-1} lower in energy

than **I2**, or 96.3 kcal mol⁻¹ lower in energy than the initial active-site-bound substrate complex **RC**.

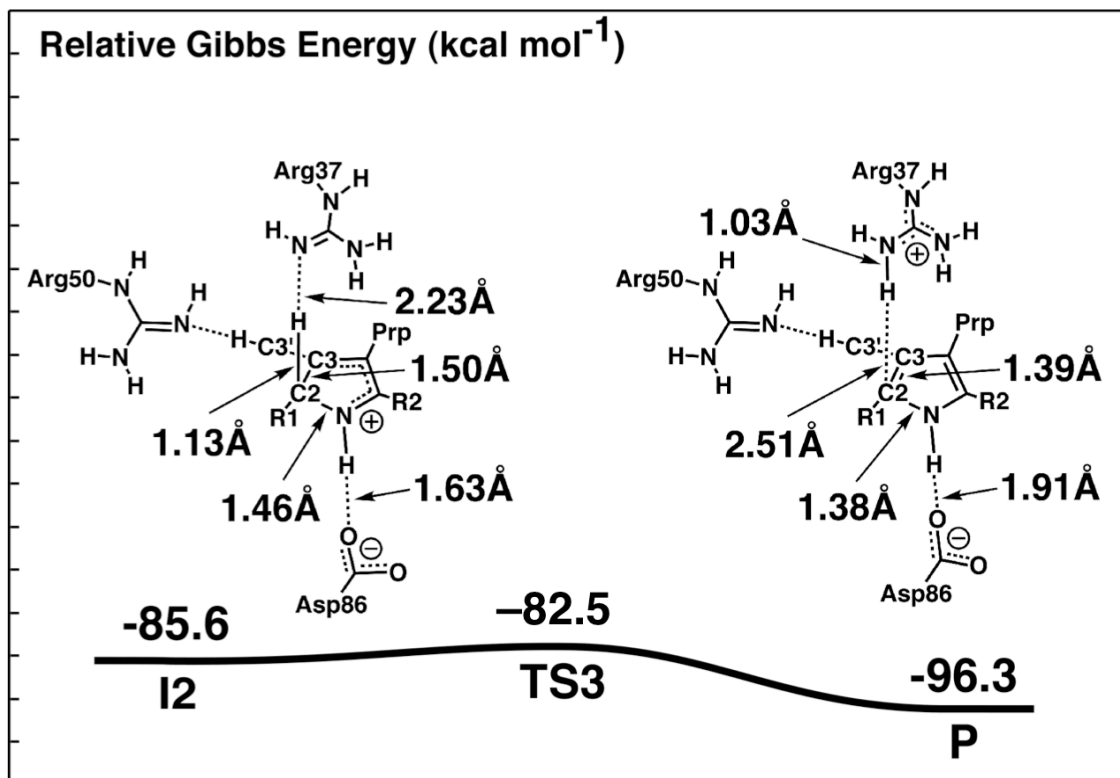


Figure 4.10. Gibbs energy profile for abstraction of the initially transferred proton to C2 by Arg37.

In the complex **PC** the C2–N–**D** and C2–C3 bonds have both shortened to 1.38 Å and 1.39 Å, respectively, which indicates that conjugation has been restored in the pyrrole ring. In addition, as the ring no longer formally has any carbocation character, the **D**–NH...⁻OOC–Asp86 hydrogen bond is lengthened to 1.91 Å, i.e., similar in length to that observed in the initial complex **RC**. Furthermore, the Arg37–N γ ...C2–**D** distance has now also increased to 2.51 Å.

4.4 Conclusion

Docking and MD simulations show that several complexing modes are possible between URO-III and the active site of hUROD. Specifically, in all complexes investigated, we found that the NH groups of the pyrrole rings are coordinated to Aps86, and that Arg37 is positioned above each tetrapyrrole ring forming arene–cation interactions. Thus, the combined use of docking and MD simulations has allowed us deeper understanding of the relevant interactions of URO-III with the active-site residues. In particular, it was observed that the complex in which URO-III binds the strongest to the active site of hUROD, several of carboxylate groups of URO-III hydrogen-bond to the backbone amide groups of the active site. Moreover, Ser219 and Arg41 along with Tyr164 and His339 interact with URO-III to properly orient the substrate within the active site.

The QM/MM calculations shows that Arg37 most likely acts as the initial acid that protonates C2, in agreement with previous small model DFT calculations²². We found that the rate-limiting step involved the proton transfer from Arg37 to C2 of URO-III, with an activation Gibbs energy of 10.3 kcal mol⁻¹, which is in good agreement with the experimentally determined range of 2.0 to 12.3 kcal mol⁻¹^{16,21}.

4.5 References

- (1) Fan, J., Liu, Q., Hao, Q., Teng, M., Niu, L. *J. Bacteriol.* **2007**, 3573-3580.
- (2) Martins, B. M., Grimm, B., Mock, H. -P., Huber, R., Messerschmidt, A. *J. Biol. Chem.* **2001**, 276, 44108-44116.
- (3) Mock, H. P., Trainotti, L., Kruse, E., Grimm, B. *Plant Mol.Biol.* **1995**, 28, 245-256.
- (4) Heinemann, I. U.; Jahn, M.; Jahn, D. *Arch. Biochem. Biophys.* **2008**, 474, 238-251.

- (5) Jackson, A. H., Sancovich, H. A., Ferramola, A. M., Evans, N., Games, D. E., Matlin, S. A., Elder, G. H., Smith, S. G. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.* **1976**, 273, 191-206.
- (6) Phillips, J. D., Parker, T.L., Schubert, H.L., Whitby, F. G., Hill C. P., Kushner, J. *P. Blood* **2001**, 98, 3179-3185.
- (7) Phillips, J. D., Whitby, F. G., Kushner, J. P., Hill C. P. *EMBO J.* **2003**, 22, 6225-6233.
- (8) Whitby, F. G., Phillips, J.D., Kushner, J.P., Hill, C.P. *EMBO J.* **1998**, 17, 2463-2471.
- (9) Wyckoff, E. E., Phillips, J.D., Sowa, A.M., Franklin, M.R., Kushner, J.P. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1996**, 1298, 294-304.
- (10) Akhtar, M. *New Comprehensive Biochemistry: Biosynthesis of Tetrapyrroles*; Elsevier: London, 1991; Vol. 19.
- (11) Barnard, G. F., Akhtar, M. *J. Chem. Soc., Chem. Commun.* **1975**, 13, 494-496.
- (12) Barnard, G. F., Akhtar, M. *J. Chem. Soc., Perkin Trans. 1* **1979**, 2354-2360.
- (13) Straka, J., Kushner, J.P. *Biochemistry* **1983**, 22, 4664-4672.
- (14) Phillips, J. D.; Warby, C. A.; Whitby, F. G.; Kushner, J. P.; Hill, C. P. *J. Mol. Biol.* **2009**, 389, 306-314.
- (15) de Verneuil, H., Sassa, S., Kappas, A. *J. Biol. Chem.* **1983**, 258, 2454-2460.
- (16) Chaufan, G.; de Molina, M. D. R.; de Viale, L. C. S. *Int. J. Biochem. Cell Biol.* **2001**, 33, 621-630.
- (17) Lewis, C. A., Wolfenden, R. *PNAS* **2008**, 105, 17328-17333.
- (18) Decataggi, S. B., Demolina, M. D. R., Deviale, L. *Int. J. Biochem.* **1991**, 23, 675-679.
- (19) Jones, R. M., Jordan, P. M. *Biochem. J.* **1993**, 293, 703-712.
- (20) Akhtar, M. *Ciba Found. Symp.* **1994**, 180, 131-152.

- (21) Juaréz, A. B., Aldonatti, C., Vigna, M. S., Ríos de Molina, M. C. *Can. J. Microbiol.* **2007**, *53*, 303-312.
- (22) Silva, P. J., Ramos, M. J. *J. Phys. Chem. B* **2005**, *109*, 18195-18200.
- (23) MOE; 2009.10 ed.; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2009.
- (24) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
- (25) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. *J. Comput. Phys.* **1999**, *151*, 114-134.
- (26) Bearpark, M. J.; Ogliaro, F.; Vreven, T.; Boggio-Pasqua, M.; Frisch, M. J.; Larkin, S. M.; Robb, M. A. In *Computation in Modern Science and Engineering Vol 2, Pts a and B*; Simos, T. E., Maroulis, G., Eds.; Amer Inst Physics: Melville, 2007; Vol. 2, p 583-585.
- (27) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Theochem-J. Mol. Struct.* **1999**, *461*, 1-21.
- (28) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959-1967.
- (29) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179.
- (30) Morokuma, K.; Musaev, D. G.; Vreven, T.; Basch, H.; Torrent, M.; Khoroshun, D. V. *IBM J. Res. Dev.* **2001**, *45*, 367-395.
- (31) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
- (32) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.
- (33) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432.

(34) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760-769.

(35) Frisch, M. J., Trucks, G.W., Schlegel,H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A.Jr., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A.; Revision D.02 ed. Wallingford CT, 2004.

(36) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

(37) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(38) Lee, C., Yang, W., Parr, R.G. *Phys. Rev. B: Condens. Matte* **1988**, *37*, 785-789.

(39) Case, D. A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J. *J. Comput. Chem.* **2005**, *26*, 1668-1688.

(40) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232-240.

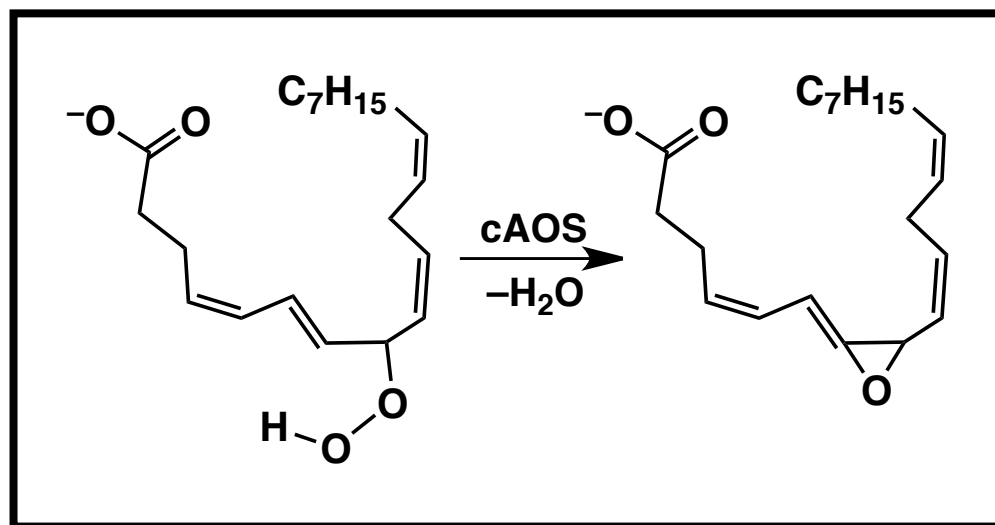
(41) Llano, J., Gault, J. W. In *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta, C. F., Ed.; Wiley-VCH: Weinheim, 2010; Vol. 2, p 920.

(42) Schlippe, Y. V. G., Hedstrom, L. *Arch. Biochem. Biophys.* **2005**, *433*, 266-278.

- (43) Gan, L., Seyedsayamdost, M.R., Shuto, S., Matsuda, A., Petsko, G.A., Hedstrom, L. *Biochemistry* **2003**, *42*, 857-863.
- (44) Bossi, R. T., Negri, A., Tedeschi, G., Mattevi, A. *Biochemistry* **2002**, *41*, 3018-3024.
- (45) Tedeschi, G., Ronchi, S., Simonic, T., Treu, C., Mattevi, A., Negri, A. *Biochemistry* **2001**, *40*, 4738-4744.
- (46) Lam, B. K., Penrose, J.F., Xu, K., Baldasaro, M.H., Austen, K.F. *J. Biol. Chem.* **1997**, *272*, 13923-13928.
- (47) Schutz, C. N.; Warshel, A. *Proteins* **2001**, *44*, 400-417.
- (48) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193-10206.
- (49) Llano, J.; Eriksson, L. A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2426-2433.
- (50) Zhang, Y. K.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459-4463.
- (51) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645-15650.

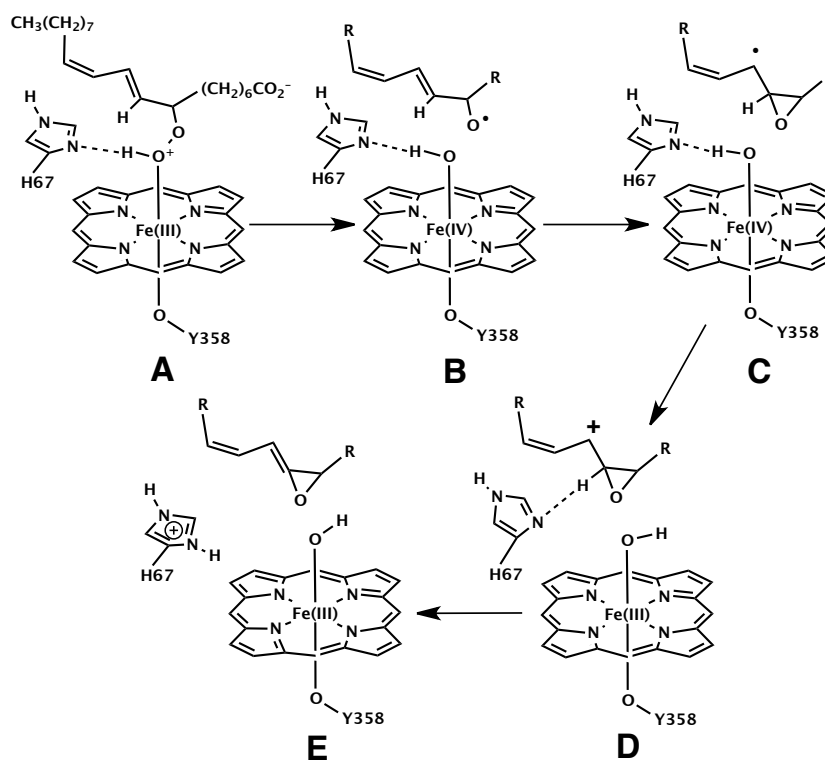
Chapter 5

Insights into the Catalytic Mechanism of Coral Allene Oxide Synthase: A Dispersion Corrected Density Functional Theory Study



5.1 Introduction

Oxylipins are involved in numerous signaling and development processes in almost all living organisms.¹⁻⁵ A key enzyme in their synthesis is allene oxide synthase (AOS).⁶ In plants this enzyme (pAOS) is a member of the CYP74A sub-family of the P450 family of hemoproteins with a cysteinyl residue as the heme's proximal ligand.⁷⁻¹⁰ In contrast, in coral AOS (cAOS) the active site structure shows a remarkable resemblance to Catalase as it has a tyrosyl proximal ligand.¹¹⁻¹³ Regardless of these differences it has been proposed that both enzymes catalyse their reactions *via* similar chemistry.⁷



Scheme 5.1. The proposed catalytic mechanism of cAOS.¹¹ For reasons of clarity not all hydrogens have been added while the porphyrin ring is shown as a porphine ring.

The proposed mechanism (**Scheme 5.1**) for cAOS begins with the binding of 8(*R*)-hydroperoxyeicosatetraenoic (8(*R*)-HPETE) acid (**A**).¹¹ Subsequently, homolytic

cleavage of the O–O bond occurs generating an alkoxy radical and compound II (Cpd II) intermediate complex with an Fe-bound hydroxyl group (**B**). It is noted that an alternative form of Cpd II exists where the oxygen is doubly bound to the Fe center (*i.e.* deprotonated). Cyclization of the alkoxy radical results in formation of an allylic epoxide radical (**C**) that is then subsequently oxidized *via* electron transfer (ET) onto Cpd II.¹¹ This results in the formation of a cationic center on the epoxide intermediate (**D**). The last step is proton transfer from the substrate's C9 center to an active site histidyl (H67) residue to generate the C=C double bond adjacent to the epoxide moiety (**E**).¹¹

Recently the mechanism of pAOS was investigated by Cho et al.¹⁴ using a QM/MM approach. More specifically, they used the density functional theory hybrid method B3LYP/LACVP (*i.e.* 6-31G on all atoms + LANL2DZ on Fe) and the CHARMM27 force field for the high- and low-layers, respectively. Notably, unlike that experimentally proposed,¹⁵ they did not obtain an Fe-peroxy bound reactive complex for either the high spin (*i.e.* sextet; $S=5/2$) or low spin (*i.e.* doublet; $S=1/2$) state with the latter being the most favoured.¹⁴ Instead, they concluded that O–O homolytic bond cleavage of the peroxide substrate occurred concomitantly with substrate binding. Moreover, a spin inversion from the sextet state to the doublet state occurred as well.¹⁴ This step was found to be rate limiting with the subsequent generation of the allylic epoxide radical occurring without a barrier. Due to the lack of an active site histidyl in pAOS, the last step, oxidation of the epoxide allylic radical proceeded in one step *via* hydrogen atom transfer from the substrate to the oxo-ferryl species rather than in two separate steps as proposed for cAOS.¹⁴ Unfortunately, as they only investigated the mechanism for the doublet state (with the exception of the sextet reactive complex), the possible role of multistate reactivity (MSR), a common feature in transition metal chemistry, was not considered.¹⁶

The interactions between substrates and proteins are commonly a combination of van der Waals (vdW), electrostatic and hydrogen bonding interactions.¹⁷ Indeed, it has been shown that inclusion of vdW effects is important for reliably computationally modeling

the binding of small molecules to model heme systems.¹⁸ However, the proper description of such dispersion interactions is a well-known limitation of commonly used DFT functionals such as B3LYP.¹⁹ For instance, a number of computational studies have examined the effects of dispersion on the reactions of P450-enzymes using both small DFT-cluster and QM/MM approaches.²⁰⁻²³ From the results obtained it was concluded that for the reactions considered the inclusion of dispersion effects lowered barriers considerably. Importantly, they were as a result in better agreement with experiment.^{21,22} Recently, Hirao²⁴ examined the performance of various DFT functionals with empirical dispersion corrections included to properly describe the Co–C bond dissociation energies of methyl cobalamin, a corrin ring containing molecule. It was found that good agreement with experiment²⁵ was only obtained upon inclusion of dispersion effects.²⁴ In addition, Hirao²⁶ examined O₂ binding by the non-heme iron enzyme *myo*-inositol oxygenase using a dispersion corrected QM/MM approach. Notably, they concluded that the reliable modeling of vdW interactions was important for properly describing the binding of O₂ within the active site.

In this current work the mechanism by which cAOS catalyses the synthesis of allene oxide from its hydroperoxy substrate has been computationally investigated using a DFT-chemical cluster approach. To the best of our knowledge it is the first such study on cAOS. In particular we have systematically examined the effect of dispersion interactions and multi-state reactivity along the mechanism, the choice of DFT functional and the role of the tyrosyl proximal ligand of cAOS compared to the cysteinyl found in pAOS.

5.2 Computational Methods

Molecular Docking. The Molecular Operating Environment (MOE)²⁷ software package was used to perform the molecular docking calculations. The crystal structure of the AOS-lipoxygenase protein from *P. homomalla* (PDB: 3DY5) was used as a template.

Prior to docking, all crystallographic waters and counter-ions were removed as well as the C-terminal lipoygenase domain. Experimentally it was shown that the catalytic activity of the N-terminal AOS domain was retained upon deletion of the C-terminal domain.⁷ The coordinates of hydrogens were then added using the MOE default method. The substrate, modeled as 8R-hydroperoxyeicosatetraenoic acid (8R-HPETE), was then docked within the active site. The binding free energy of each enzyme-substrate complex generated by this procedure was estimated with the London dG scoring function. The geometries of the top one hundred complexes were then optimized using the Forcefield refinement scheme in conjunction with the AMBER99 force field.²⁸ After minimization the binding free energies were recalculated with the London dG scoring function. The top scoring complex was then minimized using the AMBER99 force field until the root mean square gradient of the total energy fell below 0.05 kcal a.u.⁻¹. It is noted that in all top scoring conformers the substrates' peroxy moiety was ligated to the Fe center while its carboxy (R-COO⁻) head-group formed a salt bridge with K105.

QM calculations. The Gaussian09 program suite was used to perform all DFT-cluster calculations.²⁹ This approach has been successfully widely used on related systems and reviewed in detail elsewhere.^{30,31} The stationary points on the free energy surface (PES) were located at the B3LYP/BS1 level of theory.³²⁻³⁷ The combination of basis functions defined by BS1 are the 6-31G basis set on all atoms except Fe and the peroxy oxygens. For Fe a combination of the 6-311+G(2df) basis set (for valence orbitals) and LANL2DZ ECP (for core orbitals) was used while the peroxy oxygens were described by the 6-31+G(d) basis set. It has been previously shown that diffuse functions are essential for a proper description of the O-O homolytic cleavage process in the formation of Cpd I.³⁸ Frequency analyses of all stationary points was done at the same level of theory in order to characterize them as minima or transition states as well as to obtain the corresponding Gibbs free-energy corrections (ΔG_{Cor}). Notably, B3LYP has been successfully used to investigate the mechanism and properties of catalase enzymes.³⁹⁻⁴¹ Importantly, as noted

in the Introduction the active site structure of cAOS shows a remarkable resemblance to Catalase as it has a tyrosyl proximal ligand.¹¹⁻¹³

Dispersion interactions, which are not well described in B3LYP, were corrected for *via* the use of single point calculations involving dispersion correcting potentials (DCP).⁴² More specifically, single point calculations were performed at the B3LYP/6-31+G(d,p) level of theory on the above optimized geometries. These were then recalculated using a modified 6-31+G(d,p) basis set in which two basis functions have been added on each carbon, as detailed in the work of DiLabio.^{43,44} For each complex the difference in energies was taken as the dispersion correction ($\Delta\text{Disp}_{\text{Cor}}$). It is noted that the use of DCPs has been shown to provide more reliable reaction thermodynamics.⁴⁴

Reducing the amount of exact exchange in B3LYP to 15% (*i.e.* use of the B3LYP* method) has been shown to give relative energies in better agreement with experiment.¹⁸ Moreover, it has been stated that the use of B3LYP* has been shown to be better when describing the oxidation of transition metal containing compounds providing an improvement over the standard B3LYP functional.^{18,30} Furthermore, as shown in a recent investigation of 8R-LOX geometry optimizations are less sensitive to the amount of exact exchange in B3LYP, it was found that in general the energies are more sensitive.⁴⁵ Hence relative energies were determined *via* single point calculations at the IEFPCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 level of theory and corrected by inclusion of ΔG_{Cor} and ΔD_{Cor} , unless otherwise noted.⁴⁶ The IEFPCM approach with a dielectric constant of 4.0 was used to model the polarizing effect of the protein environment.^{31,47-50}

The M06 functional has been suggested to provide a better account of dispersion interactions.^{51,52} Hence, the above B3LYP/BS1 obtained structures were re-optimized at the M06/BS1 level of theory. However, this caused rearrangements of the alcohol radical intermediate complex such that it no longer was able to lead to formation of the allene oxide product. In addition, relative energies were calculated at the IEFPCM-M06/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory (**Figure B1**). Unfortunately, the

resulting relative energies were found to be enzymatically unrealistic and thus, are not discussed herein.

Chemical Model: The final MM minimized top scoring complex obtained above was used to generate the active site bound-substrate chemical model shown in **Figure 5.1**. In particular, it included the substrate 8R-HPETE modeled as 1-Hydroperoxy-n-hexane, and the side chains of R354, T66, H67 and Y358. These were included as the side chain of R354 directly interacts with the proximal Y358 and helps stabilize the negatively charged tyrosinate. Moreover, it has been suggested that they are important for the Catalase's catalytic function.^{12,53} H67 and T66 were retained as they hydrogen bond with the substrate. Furthermore, mutation of T66 by valine causes a reduction in the catalytic activity of AOS,¹³ while H67 has been proposed to accept a proton during the mechanism (see **Scheme 5.1**). The heme's porphyrin ring was modeled by a porphine ring.

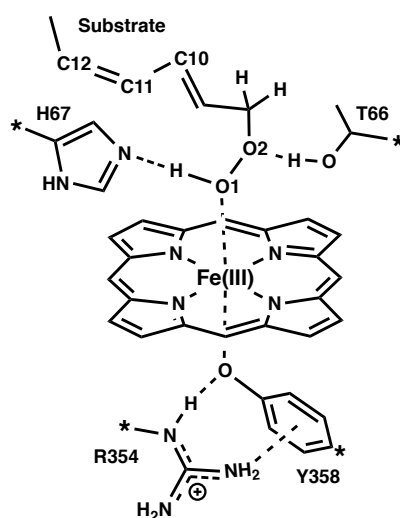


Figure 5.1: The active site bound-substrate model used in the investigation of AOS. Carbon atoms labeled with an asterisk were fixed to their final MM minimized position.

5.3 Results and Discussion

The free energy surfaces for the overall reaction of cAOS in the *overall* doublet ($S=1/2$), quartet ($S=3/2$) and sextet ($S=5/2$) states, obtained at the IEF-PCM-B3LYP*/6-311G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory, are shown in **Figure 5.2**. All energies are given relative to that of the quartet reactant complex (${}^4\text{RC}$) unless otherwise noted. The surface for the sextet state consistently lies higher in energy than those obtained for the doublet and quartet states. Hence, it is unlikely to contribute to the catalytic mechanism of cAOS. It is noted that this is in contrast to that obtained by Cho et al.¹⁴ in which their initial complex of enzyme with an unbound peroxide substrate (see Introduction) was found to prefer the sextet state. Thus, structures and energies corresponding to this state are not discussed hereafter.

The catalytic mechanism of cAOS, without dispersion corrections. For the initial reactant complex (**RC**) the quartet state (${}^4\text{RC}$) is preferred over ${}^2\text{RC}$ with the latter lying 11.2 kJ mol⁻¹ higher in energy (**Figure 5.2**). The calculated spin densities (SD) for ${}^2\text{RC}$ and ${}^4\text{RC}$ are 0.98 and 2.95, respectively (**Table 5.1**). These values indicate that in both complexes the unpaired electrons are effectively centered on the Fe. Structurally, in ${}^2\text{RC}$ and ${}^4\text{RC}$ the $r(\text{Fe}\cdots\text{O1O2})$ distances are 1.88 and 2.85 Å, respectively (**Figure 5.3**). That is, in the doublet state the Fe \cdots O1 interaction is quite strong while in the quartet state it is significantly reduced. In contrast, in a previous study on pAOS no stable initial active site-bound-substrate could be found.¹⁴ This may be due to the fact that in their investigation the basis set used (*i.e.* LACVP) included no polarization functions.

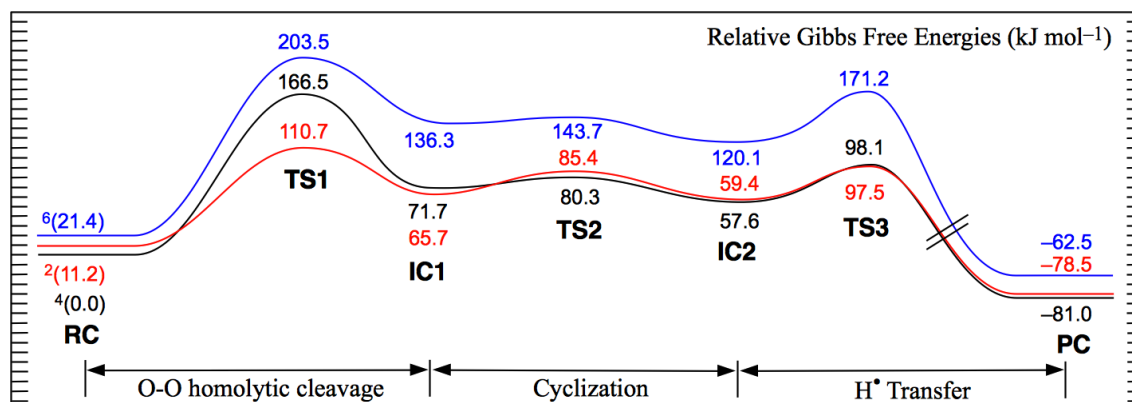


Figure 5.2: Free energy surfaces for the overall catalytic mechanism of cAOS obtained at the IEF-PCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory. Surface color code: red (doublet); black (quartet) and blue (sextet).

In ${}^2\text{RC}$, due to the strong $\text{Fe}\cdots\text{O1}$ interaction, the proton on the substrate's peroxide moiety has transferred onto the imidazole of H67 (**Figure 5.3**). In contrast, in ${}^4\text{RC}$ the peroxide moiety remains protonated. In either case, however, there exists a strong hydrogen bond between the peroxy moiety and H67; in ${}^2\text{RC}$ $r(\text{H67-N}_\epsilon\text{H}^+\cdots\text{O1}) = 1.57 \text{ \AA}$ while in ${}^4\text{RC}$ $r(\text{H67-N}_\epsilon\cdots\text{H-O1}) = 1.61 \text{ \AA}$. Notably, despite the differences in the strength of the $\text{Fe}\cdots\text{O1}$ interaction the O1-O2 and C8-O2 bond lengths in ${}^2\text{RC}$ and ${}^4\text{RC}$ are not significantly different (see **Figure 5.3**). Furthermore, to help position the substrate, in both reactant complexes T66 forms a hydrogen-bond to the peroxide moiety with $r(\text{T66-OH}\cdots\text{O2})$ distances of 1.97 (${}^2\text{RC}$) and 1.98 \AA (${}^4\text{RC}$).

The proximal tyrosyl ligand forms a reasonably strong interaction via its negatively charged phenolic oxygen with the Fe centre in both reactant complexes with $r(\text{Y358-O}\cdots\text{Fe}) = 1.95$ and 2.03 \AA for the doublet and quartet states, respectively. Simultaneously, it also forms a strong hydrogen bond with the side-chain guanidinium of the arginyl residue R354 with $r(\text{R354-N}_\epsilon\text{H}^+\cdots\text{O-Y358}) = 1.69$ and 1.63 \AA in ${}^2\text{RC}$ and ${}^4\text{RC}$, respectively. It has been suggested^{12,53} that R534 is important in the mechanism of cAOS due its involvement in the Catalase mechanism function. Indeed, removal of the

presence of R354 from the QM model significantly perturbed the Y358-O \cdots Fe interaction. For example, in $^2\mathbf{RC}$ it lead to significant lengthening in $r(\text{Y358-O}\cdots\text{Fe})$ of 0.55 Å.

Table 5.1. Spin densities on select atoms obtained at the B3LYP*/6-311+G(2df,p) level of theory.

Complex	Fe	Por	O1^a	O2^a	C10^a	C11^a	C12^a
$^2\mathbf{RC}$	0.98	-0.05	0.07	0.00	0.00	0.00	0.00
$^4\mathbf{RC}$	2.95	-0.10	0.00	0.00	0.00	0.00	0.00
$^2\mathbf{IC1}$	1.47	-0.11	0.52	-0.66	-0.06	0.02	-0.06
$^4\mathbf{IC1}$	1.48	-0.10	0.60	0.74	0.05	-0.02	0.05
$^2\mathbf{IC2}$	1.43	-0.10	0.64	-0.07	-0.63	0.27	-0.65
$^4\mathbf{IC2}$	1.43	-0.10	0.64	0.07	0.63	-0.27	0.65
$^2\mathbf{PC}$	1.00	-0.06	0.01	0.00	0.00	0.00	0.00
$^4\mathbf{PC}$	2.83	-0.05	0.04	0.00	0.00	0.00	0.00

^a Atom labels are defined in **Figure 5.1**.

The first step along the overall mechanism is cleavage of the peroxy bond. This occurs via $^2\mathbf{TS1}$ and $^4\mathbf{TS1}$ with free energy barriers of 110.7 and 166.5 kJ mol⁻¹, respectively (**Figure 5.2**). Both of these barriers are considerably higher than that previously obtained for pAOS in the doublet state of 75.7 kJ mol⁻¹.¹⁴ If only the doublet surface is considered for cAOS then the currently calculated barrier for cleavage is only 99.5 kJ mol⁻¹. Importantly, however, in cAOS $^4\mathbf{RC}$ is the lowest free energy reactant complex. Hence, it appears that cleavage of the peroxy bond preferably occurs with spin inversion (SI) from the quartet to doublet surface. Notably, the occurrence of SI has been suggested to be common in transition metal chemistry.¹⁶

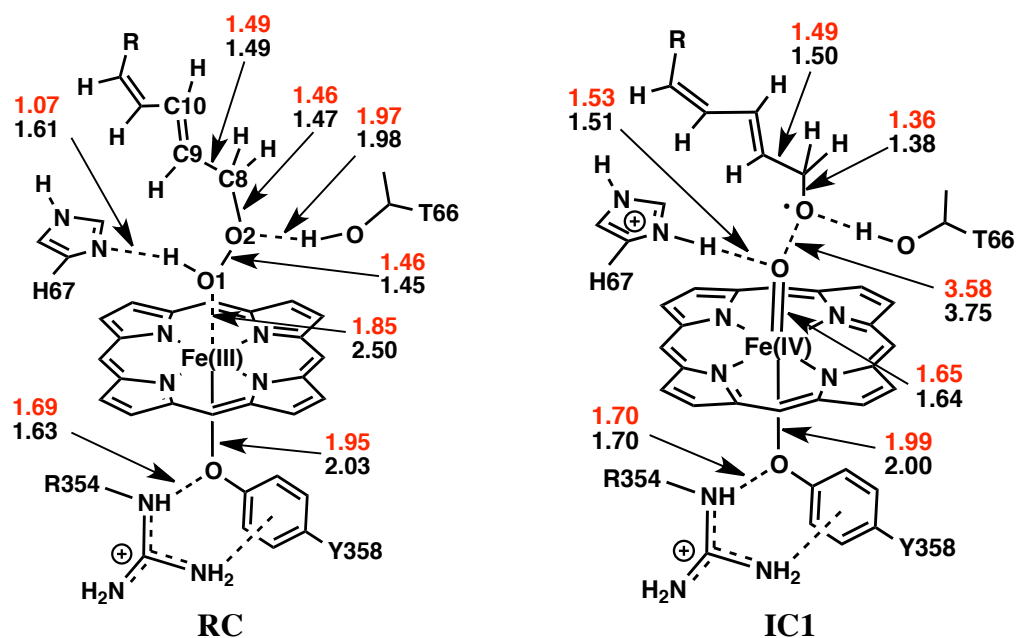


Figure 5.3: The optimized structures, with selected bond lengths in Ångströms (Å), of **RC** and **IC1** for the S=1/2 (doublet: red values) and S=3/2 (quartet: black values) states.

In ${}^2\text{TS1}$ and ${}^4\text{TS1}$ the O1 \cdots O2 peroxy bond has lengthened considerably by at least 0.24 Å to 1.87 and 1.71 Å, respectively (**Figure B3**). In the latter case (${}^4\text{TS1}$) the substrate's peroxide proton transfers to His67 concomitant with peroxide bond cleavage. That is, in both the doublet (in ${}^2\text{RC}$) and quartet (in ${}^4\text{TS1}$) state the substrate's hydroperoxy proton is transferred to the imidazole of His67 early in the overall mechanism to give a protonated His67 (His67-H $^+$). This suggests that the experimentally proposed¹¹ role of His67 as the base that abstracts a proton during the later reduction of the Cpd II intermediate is unlikely (**Scheme 5.1**). Notably, the transition structures have markedly different electronic configurations on their Fe centers compared to the corresponding reactive complexes. Consequently, the Fe-O1-O2 angles in ${}^2\text{TS1}$ and ${}^4\text{TS1}$ of 122.5° and 168.7° respectively, vary significantly by 6.2° and 54.9° respectively relative to their corresponding reactive complexes (**Table B1**).

For both the doublet and quartet states the resulting intermediate formed (**IC1**) is an oxo-ferryl Cpd II type intermediate (*i.e.* Pro(Fe(IV)=O)) with ${}^2\text{IC1}$ lying 6.0 kJ mol $^{-1}$

lower in energy than ${}^4\mathbf{IC1}$. For both states, however, formation of Cpd II is endergonic with ${}^2\mathbf{IC1}$ and ${}^4\mathbf{IC1}$ lying 65.7 and 71.7 kJ mol^{-1} higher in energy than ${}^4\mathbf{RC}$, respectively (**Figure 5.2**). The optimized Fe(IV)=O bond lengths in ${}^2\mathbf{IC1}$ and ${}^4\mathbf{IC1}$ are quite short at 1.65 and 1.64 Å respectively, indicating that they have considerable double bond character (**Figure 5.3**). Conversely, the now cleaved O1 \cdots O2 distance has lengthened considerably to 3.58 (${}^2\mathbf{IC1}$) and 3.75 (${}^4\mathbf{IC1}$) Å, respectively. Only quite minor changes in the Y358-O \cdots Fe interaction length are observed upon forming Cpd II. As can be seen in **Table 5.1**, there is very little or no spin density (SD) observed on either the tyrosyl or porphyrin ring. In contrast, the SDs on the Fe center of 1.47 (${}^2\mathbf{IC1}$) and 1.48 (${}^4\mathbf{IC1}$) suggest that the oxo-ferryl center likely has a pair of unpaired electrons (*i.e.* triplet) with parallel spin in both complexes. Furthermore, the Fe-bound oxygen has SDs of 0.52 and 0.60, respectively. For the alkoxy intermediate its total SDs are calculated to be -0.93 (${}^2\mathbf{IC1}$) and 0.99 (${}^4\mathbf{IC1}$) with the spin density predominantly localized on the oxygen radical itself which has SDs of -0.66 and 0.74, respectively. However, despite a shortening and strengthening of the T66-OH \cdots O2 interaction and hence stabilization of the intermediates oxygen radical centre (O2), spin delocalization is observed. Notably, the hydrogens geminal to O2 are calculated to now have a marked increase in SDs. This delocalization is also evidenced by the fact that its $r(\text{C8-O2}^\bullet)$ bond has shortened considerably by ~ 0.1 Å in both states.

The above Cpd II intermediate differs from that calculated for pAOS which, due in part to the lack of an active site histidyl, was protonated and consequently exhibits a considerably longer Fe–O bond length of 1.87 Å.¹⁴ Furthermore, it was a bi-radical with essentially no spin density on the Fe–O oxygen but significant SD on the proximal Fe-ligating sulfhydryl (-0.24) and porphyrin ring (-0.88).¹⁴

The next step is cyclization of the alkoxy radical to form an epoxide. This occurs via ${}^2\mathbf{TS2}$ and ${}^4\mathbf{TS2}$ (**Figure B3**) with a Gibbs barrier of 19.7 and 8.6 kJ mol^{-1} relative to ${}^2\mathbf{IC1}$ and ${}^4\mathbf{IC1}$, respectively. The resulting epoxide-containing Cpd II intermediates ($\mathbf{IC2}$) are

just 59.4 and 57.6 kJ mol⁻¹ higher than ⁴RC on the doublet and quartet surfaces, respectively (**Figure 5.2**). That is, epoxide formation is an exergonic process as they lie lower in free energy than their corresponding **IC1** complex by -6.3 and -14.1 kJ mol⁻¹ for the doublet and quartet states, respectively. Cyclization causes both the C8–O2 and C10–C9 bonds to lengthen considerably with the latter having lost its double bond character (**Figure 5.4**). Furthermore, the unpaired electron of the former alkoxy radical now exists as an allylic radical delocalized on carbons C10, C11 and C12 as indicated by the SDs given in **Table 5.1**. In the case of the Cpd II moiety itself, epoxide has an insignificant effect on its geometry.

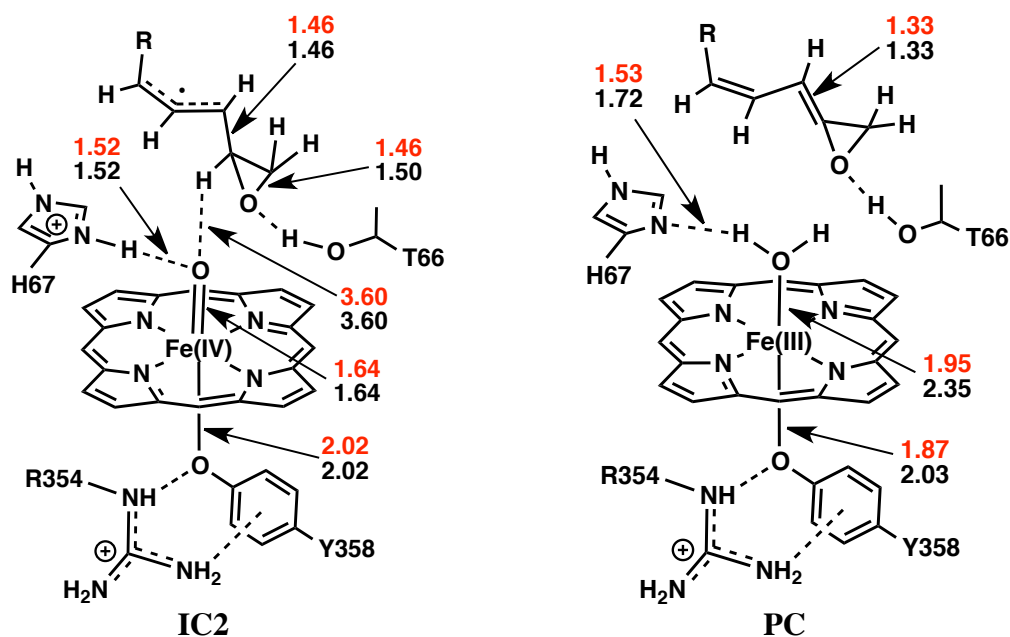


Figure 5.4: The optimized structures, with selected bond lengths in Ångströms (Å), of **IC2** and **PC** for the S=1/2 (doublet: red values) and S=3/2 (quartet: black values) states.

The final step is the generation of the allene oxide *via* the removal of a H[•] from the epoxides C9-H group. However, the Fe=O...H-C9 distance in **IC2** is approximately 3.60 Å for both the doublet and quartet states (**Figure 5.4**). Thus, in the transition structure for this process (**TS3**) the epoxide-allylic radical has shifted markedly closer to the Fe(IV)=O

moiety. However, it was found that the His67-H⁺ proton, which is hydrogen bonded to the Fe-bound oxygen in **IC2** (**Figure 5.4**), has now essentially transferred to the oxo-ferryl moiety as indicated by the relatively short $r(\text{Fe}-\text{O}\cdots\text{H}^+)$ distances of 1.00 and 1.01 Å in ²**TS3** and ⁴**TS3**, respectively (**Figure B3**). Thus, in either **TS3** the iron-oxygen moiety is perhaps better described as an Fe(IV)–OH that is now involved in H[•] abstraction from the epoxide radical. In ²**TS3** the FeO1 \cdots H[•] distance has shortened to 1.56 Å while the H[•] \cdots C9 bond has elongated to 1.17 Å. On the quartet free energy surface, ⁴**TS3** appears to occur notably later along the reaction coordinate as suggested by its markedly shorter FeO1 \cdots H[•] distance of 1.38 Å and concomitantly significantly more elongated H[•] \cdots C9 distance of 1.24 Å. Despite such structural changes the free energies for ²**TS3** and ⁴**TS3** are only 38.1 and 40.5 kJ mol⁻¹ relative to their corresponding **IC2** complexes (**Figure 5.2**).

Complete transfer of the hydrogen atom, i.e. formation of the product complex (**PC**), results in the C10–C9 bond shortening considerably due to restoration of its double bond character. Concomitantly, the C10–C9–C8–O2 torsional angle has increased to 176.9° and 178.0° in ²**PC** and ⁴**PC** respectively, indicating near planarity of the allene oxide functional group. Importantly, with the H[•] fully transferred onto the Fe-bound ⁻OH group an Fe-bound water has been formed (**Figure 5.4**). As a result, the Fe \cdots O1H₂ distance has increased to 1.95 and 2.35 Å in ²**PC** and ⁴**PC**, respectively. It is noted that, similar to that observed for the reactant complexes ²**RC** and ⁴**RC**, the calculated spin densities (SD) of ²**PC** and ⁴**PC** indicate that the unpaired electrons are essentially centered on the Fe(III) metal ion (**Table 5.1**). In particular, for ²**PC** and ⁴**PC** the calculated SDs on the Fe are 1.00 and 2.83, respectively. The relative free energies of ²**PC** and ⁴**PC** are markedly lower than that of ⁴**RC** by 78.5 and 81.0 kJ mol⁻¹, respectively (**Figure 5.2**). That is, the overall formation of allene oxide via a PCET mechanism as catalysed by cAOS is exergonic.

The above free energy surfaces suggest that the mechanism of cAOS may involve multi-state reactivity (MSR).³¹ More specifically, the reaction begins in the quartet state with ⁴RC. Then, a spin inversion occurs allowing for a more facile O–O cleavage on the doublet surface via ²TS1. More specifically, the barrier for this process on the doublet surface is 55.8 kJ mol⁻¹ lower in free energy than would otherwise be required on the quartet surface (**Figure 5.2**)! It should be noted that given that ⁴PC lies lower in free energy than ²PC a second SI likely occurs later in the mechanism and prior to final product formation. Thus, the use of a common "non-corrected" computational approach to investigating enzymatic processes suggests that the reaction of cAOS requires the use of MSR.³¹ It is noted that regardless of the possible SI processes, the overall mechanism of cAOS is calculated to occur with considerably higher Gibbs free energies than that obtained previously¹⁴ for the analogous pathway in pAOS. In particular, the free energies for the entire process (**Figure 5.2**) are on average 30.2 kJ mol⁻¹ higher than those for pAOS.¹⁴ The most significant difference occurring for the H[•] abstraction process (*i.e.* **TS3**). Potentially this difference may be due to the presence of a ligating tyrosyl residue in cAOS rather than the cysteinyl in pAOS resulting in a more reactive Cpd II intermediate. It is noted, that these energetic differences may also partly be due to the differences in the computational models used (*i.e.* B3LYP*-based QM-cluster versus use of B3LYP within a QM/MM framework¹⁴). However, as noted above the Cpd II intermediate of cAOS differs from that calculated for pAOS which was found to be bi-radical with essentially no spin density on the Fe–O oxygen but significant SD on the proximal Fe-ligating sulfhydryl (-0.24) and porphyrin ring (-0.88).¹⁴

As noted in the Introduction, however, non-covalent interactions often dominate in the interactions between ligands and proteins.¹⁷ Furthermore, previous investigations of P450-enzymes have suggested that the barriers for the reactions were considerably affected by the inclusion of corrections for dispersion interactions.²⁰⁻²³

The Effects of Including Dispersion Corrections. The effects of dispersion interactions were modeled by, as described in the Computational Methods, inclusion of a dispersion correction ($\Delta\text{Disp}_{\text{Cor}}$) to the above free energy surfaces. The resulting corrected free energy surfaces for the doublet, quartet and sextet states are presented in **Figure 5.5**. Again, the sextet surface is calculated to lie significantly higher in energy for all complexes and transition structures along the enzymes mechanistic pathway. In fact, inclusion of $\Delta\text{Disp}_{\text{Cor}}$ has effectively raised the relative energies of all but one of the sextet species' by 1.6 (**IC1**) – 22.8 (**PC**) kJ mol^{-1} with respect to $^4\text{RC}_{\text{dc}}$ (c.f. **Figure 5.2**). The only exception is **TS3** which decreases by 4.5 kJ mol^{-1} . This simply reflects that all sextet species except **TS3** are stabilized to a lesser extent by dispersion interactions than $^4\text{RC}_{\text{d}}$ (see below). Hence, unless otherwise noted discussion is herein limited to the doublet and quartet surfaces.

Upon correcting for dispersion interactions ^2RC is stabilized by 31.3 kJ mol^{-1} relative to ^4RC . In fact, in contrast to that seen for the uncorrected free energies (c.f. **Figure 5.2**) $^2\text{RC}_{\text{dc}}$ now lies *lower* in energy than $^4\text{RC}_{\text{dc}}$ by 20.1 kJ mol^{-1} (**Figure 5.5**)! This preferential stabilization of $^2\text{RC}_{\text{dc}}$ relative to $^4\text{RC}_{\text{dc}}$ is likely due to the fact that in the former the hydroperoxy substrate is significantly more tightly ligated to the heme Fe center (**Figure 5.3**). If we consider the distance between centers of mass for the hydroperoxy substrate and porphine ring (**Figure B2**) we find a difference of 0.362 Å. As a consequence, the dispersion interactions between the substrates carbon chain and the heme are greater. This enhanced stabilization is also observed, though to a lesser extent, in the transition structure for homolytic O–O bond cleavage **TS1_{dc}** (**Figure 5.2**). Specifically, the free energy barrier for this step is now lowered by 6.2 kJ mol^{-1} relative to $^2\text{RC}_{\text{dc}}$ (a stabilization of 26.3 kJ mol^{-1} relative to $^2\text{RC}_{\text{dc}}$). Regardless of this stabilization, O–O homolytic cleavage remains the rate-limiting step in the overall mechanism on both the doublet and quartet surfaces. Importantly, this reordering of states suggests that spin inversion (SI) is not necessary for O–O cleavage given this occurs

lowest on the doublet surface with a Gibbs barrier of $104.5 \text{ kJ mol}^{-1}$ (**Figure 5.5**). This process, however, remains the rate-limiting step in the overall mechanism of cAOS. For the resulting Cpd II intermediate **IC1**, opposite effects appear to be seen for the doublet and quartet states. More specifically, ${}^2\text{IC1}_{\text{dc}}$ appears to be destabilized by 15.4 kJ mol^{-1} relative to ${}^2\text{RC}_{\text{dc}}$ while ${}^4\text{IC1}_{\text{dc}}$ appears to be stabilized by 10.4 kJ mol^{-1} relative to the initial reactant complex ${}^4\text{RC}_{\text{dc}}$. However, this is due to the fact that ${}^2\text{RC}_{\text{dc}}$ is significantly stabilized by dispersion interactions due to tighter binding of the substrate to the heme. But, upon O—O bond cleavage the alkoxy radical formed is shifted away from the heme and hence ${}^2\text{IC1}_{\text{dc}}$ and later stationary points along the mechanism pathway are *less* stabilized by dispersion interactions (**Figure 5.3**). However, we do see slightly greater stabilization of ${}^2\text{IC1}$ with respect to ${}^4\text{IC1}$ when dispersion effects are included (*i.e.* ${}^2\text{IC1}_{\text{dc}}$ lies 11.5 kJ mol^{-1} lower in energy than ${}^4\text{IC1}_{\text{dc}}$ as opposed to only 6.0 kJ mol^{-1} seen in **Figure 5.2**). This can be better understood if we again consider the distance between centers of mass for the alkoxy radical and porphine ring. In particular, like the reactive complexes we find the intermediate is slightly closer to the porphine ring for the doublet system than the quartet system where a difference of 0.211 \AA was found (**Figure B2**).

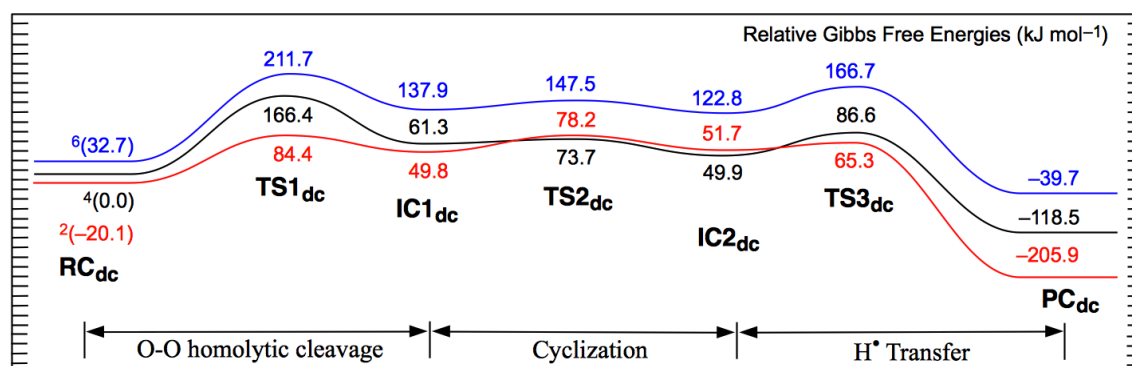


Figure 5.5: Schematic illustration of the dispersion-corrected free energy surfaces obtained at the IEF-PCM-B3LYP*/6-311+G(2df,p). + ΔG_{Cor} + $\Delta \text{Disp}_{\text{Cor}}$ level of theory for the overall catalytic mechanism of cAOS. Surface color code: red (doublet); black (quartet) and blue (sextet).

The barriers for conversion of IC1_{dc} into the alternate Cpd II species IC2_{dc} via TS2_{dc} are 28.4 and 12.4 kJ mol^{-1} for the doublet and quartet surfaces, respectively (**Figure 5.5**). Notably, ${}^4\text{TS2}_{\text{dc}}$ lies 3.5 kJ mol^{-1} lower in energy than ${}^2\text{TS2}_{\text{dc}}$. Furthermore, unlike that observed for the order of the states for IC1_{dc} , ${}^4\text{IC2}_{\text{dc}}$ lies lower in energy than ${}^2\text{IC2}_{\text{dc}}$ by 1.8 kJ mol^{-1} . Notably, as seen in **Figure 5.2**, ${}^4\text{IC2}$ also lies 1.8 kJ mol^{-1} lower in energy than ${}^2\text{IC2}$. If we again consider the distance between centers of mass for the allylic epoxide radical intermediate and porphine ring we find an insignificant difference of 0.001 Å (**Figure B2**). Thus, as seen in **Figure 5.2**, multi-state reactivity could potentially play a role in interconversion of the two mechanistic Cpd II species with the quartet surface providing a slightly lower energy pathway (**Figure 5.5**).

As observed for the non-dispersion corrected surfaces (**Figure 5.2**), the dispersion corrected barriers for the final step are again lower on the doublet surface than on the quartet surface. However, now the difference is much greater with ${}^2\text{TS3}_{\text{dc}}$ lying lower in energy than ${}^4\text{TS3}_{\text{dc}}$ by 21.3 kJ mol^{-1} . That is, the final reaction step once again preferentially proceeds via a possible spin inversion from the quartet to doublet surface. Notably, in contrast to that observed for the uncorrected free energy surfaces (c.f. **Figure 5.2**), the inclusion of dispersion effects leads to the suggestion that the overall mechanism of cAOS could occur *without* the necessity of a spin inversion from the doublet to quartet surfaces. For the product complexes, and in contrast without correcting for dispersion interactions (**Figure 5.2**), ${}^2\text{PC}_{\text{dc}}$ now lies lower in energy than ${}^4\text{PC}_{\text{dc}}$ by 87.4 kJ mol^{-1} ! That is, similar to the reactive complexes, accounting for dispersion interactions results in a reordering of the states for the product complex. If we consider the distance between centers of mass for the allene oxide product and porphine ring (**Figure B2**) we find a difference of 0.713 Å! Hence, ${}^2\text{PC}_{\text{dc}}$ is strongly stabilized by such effects due to the allene oxide itself being considerably closer to the heme ring in ${}^2\text{PC}_{\text{dc}}$ compared to ${}^4\text{PC}_{\text{dc}}$. Like the previous situation the overall mechanism of cAOS with dispersion corrections (with the exception of ${}^2\text{PC}$) is calculated to occur with higher Gibbs free energies than

that obtained previously¹⁴ for the analogous pathway in pAOS. In particular, the free energies for the entire process (excluding ²PC) are on average 37.3 kJ mol⁻¹ higher than those for pAOS.¹⁴

The Effects of Choice of DFT functional. As recently discussed by Ye and Neese⁵⁴ the ordering of spin states can be sensitive to the functional chosen. More specifically, they investigated several inorganic complexes that have been shown to be problem cases for DFT methods.⁵⁴ They concluded that in general the B2PLYP functional in conjunction with large and flexible basis sets gave the best qualitative agreement with experiment.⁵⁴ However, for complexes containing strong-field ligands all DFT methods employed predicted the correct ground-state multiplicity.⁵⁴ Reiher et al.⁴⁶ have shown that the relative ordering of states is sensitive to the amount of the HF contribution. In particular, they concluded that the most reliable description of transition-metal complexes with sulfur-rich first coordination spheres was obtained when the HF contribution in B3LYP was reduced to 15%.⁴⁶ Notably, in heme systems the ligands that bind the Fe centre are typically characterized as strong field ligands.⁵⁵

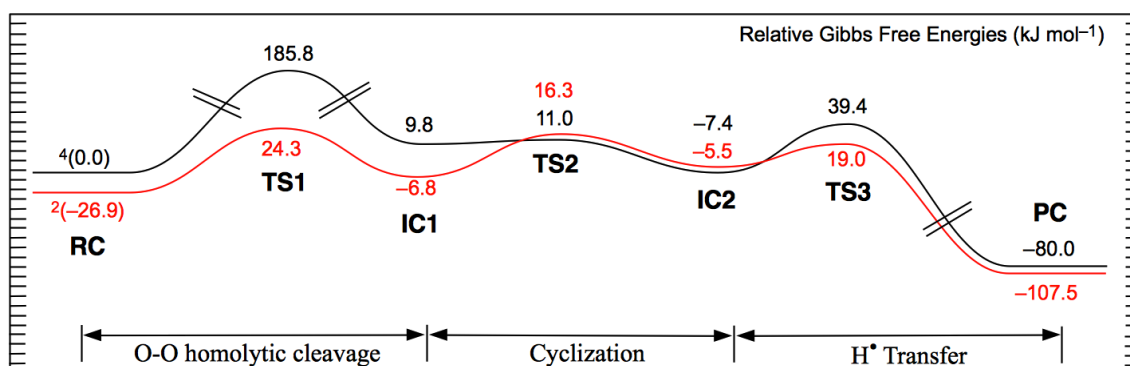


Figure 5.6. Free energy surfaces for the catalytic mechanism of cAOS obtained at the IEF-PCM-BP86/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory. Surface color code: red (doublet; S=1/2); black (quartet; S=3/2).

Thus, we also investigated the mechanism of cAOS using the GGA functional (i.e. 0% HF contribution), BP86. It is noted that DCPs were not developed for use with BP86 hence the relative free energies were obtained at the IEF-PCM-BP86/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory and are given in **Figure 5.6**. However, the corresponding free energy surface in which $\Delta \text{Disp}_{\text{Cor}}$ has been included is provided in **Figure B2**.

As can be seen in **Figure 5.6**, the use of the BP86 functional results in ^4RC no longer being the most favoured state. Instead, it now lies 26.9 kJ mol^{-1} higher in energy than ^2RC . Furthermore, the relative free energy of **PC** is also now lowest for the doublet state. Interestingly, however, while the relative ordering of the states of **RC** and **PC** are reversed compared to that obtained using B3LYP* (**Figure 5.2**), those of the remaining stationary points along the pathway are not. For example, in the case of **TS2** and **IC2** the quartet state is still lower in energy compared to the doublet state. Importantly, using the BP86 functional the barrier for O-O homolytic cleavage in the $S=1/2$ state is significantly lower at only 51.2 kJ mol^{-1} . In contrast, using the B3LYP* functional this same reaction has a calculated barrier of $110.7 \text{ kJ mol}^{-1}$ (**Figure 5.2**). Furthermore, the use of the BP86 functional suggests that in order to be enzymatically feasible the mechanism of cAOS does not appear to require the involvement of MSR.

In order to better compare the effect of changing functional on each of the free energy surfaces for both the doublet and quartet states, the overall mechanisms obtained at the IEF-PCM-DFT_i/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory (DFT_i = M06, B3LYP, B3LYP* and BP86) for each state are shown in **Figure 5.7**. As can be seen, the results suggest that for both states there is a correlation between the amount of %HF contribution in the functional and the reduction in the relative energies of the stationary points along the pathway. For both states when we change the functional from M06 to B3LYP we see an average reduction in relative free energies of 6.2 and 26.9 kJ mol^{-1} for the doublet ($S=1/2$) and quartet ($S=3/2$) states, respectively. Reducing the %HF

contribution in B3LYP from 20 to 15% (i.e., on changing from B3LYP to B3LYP*) further average reductions are observed of 12.0 (doublet) and 31.4 (quartet) kJ mol^{-1} . Lastly, on changing the functional from B3LYP* to BP86 further average reductions in the relative energies are observed of 28.1 and 39.1 kJ mol^{-1} for the doublet ($S=1/2$) and quartet ($S=3/2$) states, respectively. That is, for both states as the %HF contribution is reduced, so are in general the relative free energies of the various stationary points along the surface (relative to **RC**). In fact, overall, the average reduction in relative energies obtained upon going from the M06 to BP86 functional are 46.3 and 97.3 kJ mol^{-1} for the doublet and quartet states, respectively. The largest change on both surfaces occurs for **TS2** which sees a lowering in its relative energy of 74.6 ($^2\text{TS2}$) and 146.7 ($^4\text{TS2}$) kJ mol^{-1} . Thus, comparison of the results obtained using the B3LYP* (**Figure 5.2**), BP86 (**Figure 5.6**) and M06 (**Figure B1**) functionals suggests that the mechanism of cAOS is also sensitive to the choice of functional, with the quartet state being most affected.

Interestingly, from **Figure 5.7** it can be seen that the M06 functional predicts a higher barrier for O–O cleavage (i.e. **TS1**) in comparison to the other three functionals. However, for the same functional the relative energy of **TS3** with respect to **IC2** is considerably less than that obtained with the other functionals. In particular, $^4\text{TS3}$ actually lies lower in energy than $^4\text{IC2}$. This indicates that using the M06 functional, and in contrast to that obtained using any of the other functionals, the final step of the mechanism is predicted to occur essentially without a barrier at 298 K on the quartet surface. That is, after cyclization of the alkoxy radical, H^\bullet abstraction from the epoxide intermediate by Cpd II is predicted to readily occur. On the doublet surface (**Figure 5.7a**) $^2\text{TS3}$ lies slightly higher in energy than $^2\text{IC2}$ by only 7.2 kJ mol^{-1} . Thus, for either state use of the M06 functional leads to the prediction of a low barrier for H^\bullet abstraction by Cpd II. Overall, however, the most thermodynamically and kinetically favourable free energy surfaces, but still qualitatively correct, are obtained using the BP86 functional (i.e. IEF-PCM-BP86/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory). That is, the

functional with the lowest %HF contribution (0%) in general gives the lowest relative energies (with respect to **RC**) of the intermediates and transition structures. In contrast, M06 has the highest %HF contribution (27%) of the four functionals considered herein and in general gives the highest relative free energies of the intermediates and transition structures with respect to **RC**.

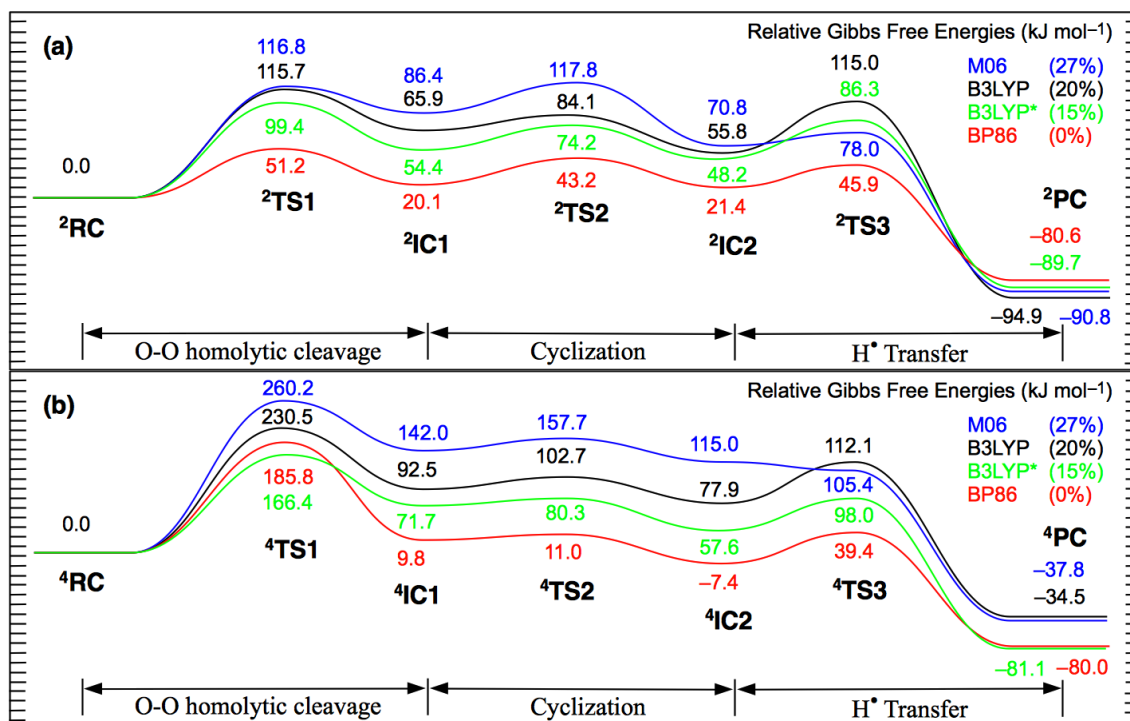


Figure 5.7. Free energy surfaces for the catalytic mechanism of cAOS obtained at the IEF-PCM-DFT_i/6-311G(2df,p)//B3LYP*/BS1 + ΔG_{cor} level of theory (DFT_i = M06, B3LYP, B3LYP* and BP86): (a) doublet (S=1/2) state and (b) quartet (S=3/2) state. The values given in parentheses indicate the %HF contribution in the functional.

5.4 Conclusions

In this present work the mechanism by which coral Allene Oxide Synthase (cAOS) catalyses the formation of allene oxide from its hydroperoxy substrate has been computationally investigated using a DFT-chemical cluster approach. Specifically, we

have examined the effect of dispersion interactions and multi-state reactivity along the mechanism, and the effect of the tyrosyl proximal ligand of cAOS compared to the cysteinyl found in pAOS.

In the reactant complex (**RC**) the hydroperoxy substrate forms a strong Fe-O—O-C crosslink in the overall S=1/2 (doublet) state with $r(\text{Fe—O}) = 1.85 \text{ \AA}$. In contrast, in the overall S=3/2 (quartet) state this link is much weaker at $r(\text{Fe—O}) = 2.50 \text{ \AA}$. Regardless, the overall mechanism begins with cleavage of the peroxy O—O bond to give a Cpd II-type intermediate with concomitant formation of an alkoxy radical. Subsequently, the latter species undergoes a rearrangement to give an epoxide with a delocalized allylic radical. The final step is hydrogen abstraction from the epoxide to give an Fe-bound H₂O and an epoxide. Thus, cAOS utilizes a mechanism that is similar to that for pAOS.

The mechanism of cAOS, however, appears to differ from that of pAOS in several key features. In particular, the initial Cpd II intermediate formed has a markedly different overall electronic configuration to that calculated for pAOS. This is likely due to both the presence of a histidyl active site residue in cAOS, which is lacking in pAOS, and a ligating tyrosyl residue. Furthermore, the mechanism occurs with considerably higher Gibbs free energies of reaction than that for the analogous pathway in pAOS. However, it is noted that these energetic differences may be partly due to differences in the computational models used to previously¹⁴ study pAOS versus that used herein.

From the results obtained at the IEF-PCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory the inclusion of dispersion effects results in considerable changes to the free energy surface for the mechanism. For instance, without dispersion effects the homolytic O—O bond cleavage likely occurs with SI from the quartet to doublet surface. However, with dispersion corrections the energy ordering of the various states of RC is altered such that SI is not needed for the initial step as the overall S=1/2 (doublet-state) reactive complex (*i.e.* **²RC_{dc}**) is now most favoured. Similarly, the occurrence of SI in product formation is also now unlikely to occur when dispersion effects are included due

in part to reordering of the relative free energies of the product complexes; ${}^2\mathbf{RC}_{\text{dc}}$ is now significantly more favoured. The contribution of dispersion effects directly correlates with the changes observed along the mechanisms pathway with regards to the distance between the center of mass of the substrate and heme.

Thus, in contrast to that observed for the uncorrected free energy surfaces (i.e. at the IEF-PCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory) the inclusion of dispersion effects leads to the suggestion that the overall mechanism of cAOS could occur *without* the need for spin inversion. Importantly, the present investigation infers that while energetic differences may exist due to the various electronic configurations of the Fe center subtle effects such as the vdW distances between substrate and enzyme can also significantly affect the energetics. Importantly, such effects may result in very different qualitative and quantitative results as shown here.

In addition we investigated the effect of changing functional (i.e. at the IEF-PCM-DFT_i/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory; DFT_i = M06, B3LYP, B3LYP*, BP86) on the free energy surfaces for both the doublet and quartet states. For both states there is in general a correlation between the amount of %HF contribution in the functional and the reduction in the relative energies of the stationary points along the pathway. That is, the functional with the lowest %HF contribution (0%) in general gives the lowest relative free energies (with respect to \mathbf{RC}) of the intermediates and transition structures. In contrast, M06, which has the highest %HF contribution (27%) of the four functionals considered herein, in general gives the highest relative free energies of the intermediates and transition structures with respect to \mathbf{RC} . In fact, the average reduction in relative free energies obtained upon going from the M06 to BP86 functional is 46.3 and 97.3 kJ mol⁻¹ for the doublet and quartet states, respectively.

5.5 References

- (1) Caldwell, G. S. *Mar. Drugs* **2009**, *7*, 367-400.
- (2) Sebolai, O. M.; Pohl, C. H.; Kock, L. J. F.; Chaturvedi, V.; del Poeta, M. *Prostaglandins Other Lipid Mediat.* **2012**, *97*, 17-21.
- (3) Vera, J.; Castro, J.; Gonzalez, A.; Moenne, A. *Mar. Drugs* **2011**, *9*, 2514-2525.
- (4) Brodhun, F.; Feussner, I. *Febs J.* **2011**, *278*, 1047-1063.
- (5) Tsitsigiannis, D. I.; Keller, N. P. *Trends Microbiol.* **2007**, *15*, 109-118.
- (6) Andreou, A.; Brodhun, F.; Feussner, I. *Prog. Lipid Res.* **2009**, *48*, 148-170.
- (7) Boutaud, O.; Brash, A. R. *J. Biol. Chem.* **1999**, *274*, 33764-33770.
- (8) Feussner, I.; Wasternack, C. *Annu. Rev. Plant Biol.* **2002**, *53*, 275-297.
- (9) Grechkin, A. *Prog. Lipid. Res.* **1998**, *37*, 317-352.
- (10) Gao, B.; Boeglin, W. E.; Brash, A. R. *Arch. Biochem. Biophys.* **2008**, *477*, 285-290.
- (11) Oldham, M. L.; Brash, A. R.; Newcomer, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 297-302.
- (12) Abraham, B. D.; Sono, M.; Boutaud, O.; Shriner, A.; Dawson, J. H.; Brash, A. R.; Gaffney, B. J. *Biochemistry* **2001**, *40*, 2251-2259.
- (13) Tosha, T.; Uchida, T.; Brash, A. R.; Kitagawa, T. *J. Biol. Chem.* **2006**, *281*, 12610-12617.
- (14) Cho, K. B.; Lai, W. Z.; Hamberg, M.; Raman, C. S.; Shaik, S. *Arch. Biochem. Biophys.* **2011**, *507*, 14-25.
- (15) Lee, D. S.; Nioche, P.; Hamberg, M.; Raman, C. S. *Nature* **2008**, *455*, 363-U327.
- (16) Poli, R.; Harvey, J. N. *Chem. Soc. Rev.* **2003**, *32*, 1-8.
- (17) Johnson, E. R.; Keinan, S.; Mori-Sanchez, P.; Contreras-Garcia, J.; Cohen, A. J.; Yang, W. T. *J. Am. Chem. Soc.* **2010**, *132*, 6498-6506.

- (18) Siegbahn, P. E. M.; Blomberg, M. R. A.; Chen, S. L. *J. Chem. Theory Comput.* **2010**, *6*, 2040-2044.
- (19) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439-10452.
- (20) Lai, W.; Chen, H.; Cohen, S.; Shaik, S. *J. Phys. Chem. Lett.* **2011**, *2*, 2229-2235.
- (21) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3232-3237.
- (22) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Chem. Theory Comput.* **2012**, *8*, 4637-4645.
- (23) Schyman, P.; Lai, W. Z.; Chen, H.; Wang, Y.; Shaik, S. *J. Am. Chem. Soc.* **2011**, *133*, 7977-7984.
- (24) Hirao, H. *J. Phys. Chem. A* **2011**, *115*, 9308-9313.
- (25) Hung, R. R.; Grabowski, J. J. *J. Am. Chem. Soc.* **1999**, *121*, 1359-1364.
- (26) Hirao, H. *J. Phys. Chem. B* **2011**, *115*, 11278-11285.
- (27) Molecular Operating Environment; 2010.10 ed.; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2010.
- (28) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049-1074.
- (29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.;

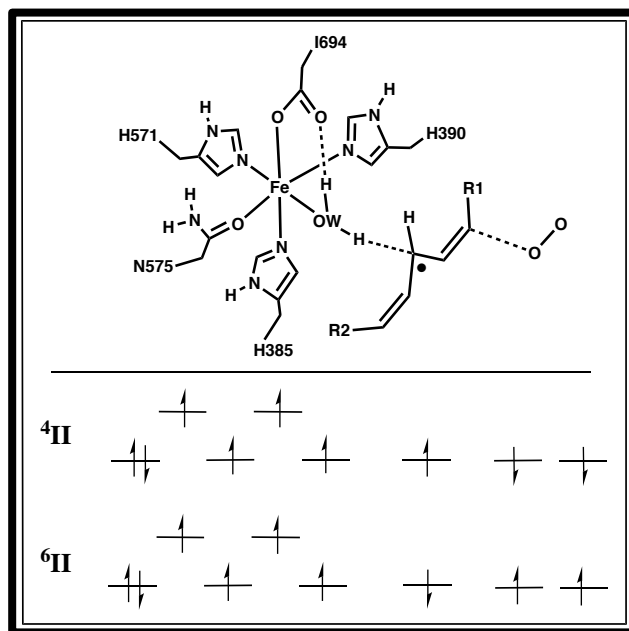
Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision B.01, Wallingford CT, 2010.

- (30) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643-651.
- (31) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232-240.
- (32) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (33) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- (34) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403-412.
- (35) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
- (36) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.
- (37) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.
- (38) Chen, H.; Hirao, H.; Derat, E.; Schlichting, I.; Shaik, S. *J. Phys. Chem. B* **2008**, *112*, 9490-9500.
- (39) Alfonso-Prieto, M.; Vidossich, P.; Rovira, C. *Arch. Biochem. Biophys.* **2012**, *525*, 121-130.
- (40) Vidossich, P.; Alfonso-Prieto, M.; Rovira, C. *J. Inorg. Biochem.* **2012**, *117*, 292-297.
- (41) de Visser, S. P. *Inorg. Chem.* **2006**, *45*, 9551-9557.
- (42) Lill, S. O. N. *J. Phys. Chem. A* **2009**, *113*, 10321-10326.
- (43) DiLabio, G. A. *Chem. Phys. Lett.* **2008**, *455*, 348-353.
- (44) Mackie, I. D.; DiLabio, G. A. *J. Phys. Chem. A* **2008**, *112*, 10968-10976.
- (45) Bushnell, E. A. C.; Gauld, J. W. *J. Comput. Chem.* **2013**, *34*, 141-148.
- (46) Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *107*, 48-55.
- (47) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032-3041.
- (48) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506-10517.
- (49) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151-5158.

- (50) Tomasi, J.; Mennucci, B.; Cancès, E. *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211-226.
- (51) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
- (52) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.
- (53) Fita, I.; Rossmann, M. G. *J. Mol. Biol.* **1985**, *185*, 21-37.
- (54) Ye, S.; Neese, F. *Inorg. Chem.* **2010**, *49*, 772-774.
- (55) Lundberg, M.; Morokuma, K. *J. Phys. Chem. B* **2007**, *111*, 9380-9389.

Chapter 6

An Assessment of Standard, Hybrid, Meta and Hybrid-meta GGA Density Functional Theory Methods for Open-Shell Systems: The Case of the Non-Heme Iron Enzyme 8R-LOX



6.1 Introduction

Metalloenzymes, in particular those containing iron, catalyse a broad range of metabolically important reactions within organisms including C—H bond activation, hydrolysis and DNA repair.¹⁻⁵ A detailed understanding of their chemistry and catalytic mechanism provides invaluable insights into their biochemical function as well as those of related species. In addition, it can also provide fundamental chemical insights into, for instance, catalysis, as well as enable the development of new more effective therapeutic treatments. Consequently, they have been the subject of numerous experimental and computational investigations.

For computational investigations on iron-containing metalloenzymes hybrid density functional theory (DFT) methods, in particular B3LYP, have become the standard tools of choice.⁶⁻⁹ It has been widely applied to a variety of such systems and shown to be able to provide useful chemical and mechanistic insights.⁴⁻⁶ This is perhaps surprising given that the B3LYP functional was parameterized based on reference molecules that do not include metals.¹⁰ Indeed, for some metal-containing systems such approaches have been shown not to give reasonable agreement with experiment or to in fact fail.^{11,12} In particular, for Fe-containing systems with near-degenerate states such methods are often unable to give a qualitatively or quantitatively correct ordering of states.¹⁰ For example, in a study by Ghosh et al.¹² on an FeIII(Por)Cl complex the B3LYP method predicted that a quartet and not the experimentally observed sextet was the ground state. Unfortunately, accurate determination of the relative energies of near-degenerate states can be essential in elucidating a given systems chemistry.^{10,13}

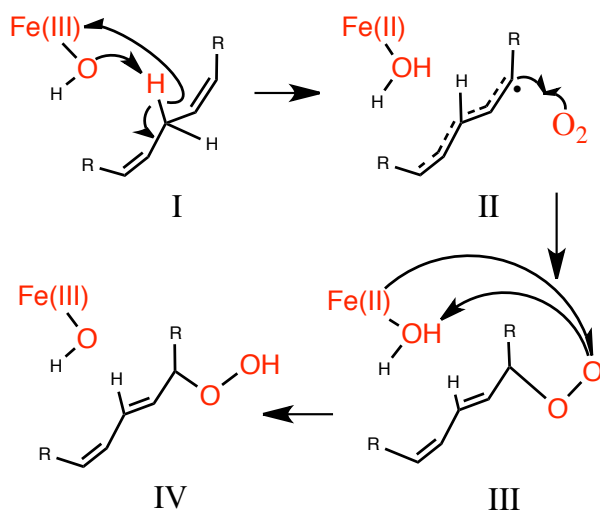
The reliability of such methods in describing the relative energies can be sensitive to the amount of exact Hartree-Fock (HF) exchange included.¹⁰ In particular, it has been proposed that for metal-containing systems with unpaired electrons a HF contribution of 15% provides better accuracy.^{10,14-17} Indeed, Reiher et al.¹⁰ showed that for B3LYP the

reduction from the widely used value of 20% to 15% HF contribution (denoted as B3LYP*) was able to reproduce the energetics of all Fe(II)–S complexes they considered as part of their investigation. In addition, the optimized structural parameters were also in better agreement with experimental X-ray structures. Importantly, Solomon et al.¹⁸ later showed that in the case of the G2 test set B3LYP* preserved the reliability of B3LYP. Thus, it was concluded that for many compounds the reliability and accuracy of the B3LYP functional is retained, while providing a much better description for TM-containing compounds. However, more recently, Hughes and Friesner¹⁹ constructed a database of experimental spectra of 57 octahedral first-row transition metal complexes that included V, Ni, Mn, Cr, Fe and Co. In addition, they also varied the ligands to include examples of M–C, –N, –O, –S, –N, –F and –Cl bonds. They then examined the effects of reducing the HF contribution in B3LYP from 20 to 15% on its ability to reproduce spin-splitting in such complexes. They concluded that reducing the percent HF contribution did not lead to *general* satisfactory agreement with experiment.¹⁰ Notably, in all of these systems any unpaired electrons were formally located on the sole metal-ion centre. That is, the effect of the percent HF contribution on systems containing multiple centers of unpaired electrons has not, to the best of our knowledge, been examined.

Lipoxygenases (LOXs) are a ubiquitous family of non-heme iron enzymes involved in the stereo- and regiospecific peroxidation of fatty-unsaturated acid substrates, typically arachidonic acid (AA) or linoleic acid (LA), by molecular oxygen.²⁰⁻²⁴ The active enzyme contains a ferric (Fe(III))–iron that exists as a high spin hexa-coordinated center. More specifically, it is ligated by a hydroxyl (OH) and the R-groups or backbones of five active site residues: one being the C-terminal isoleucyl's carboxylate while three others are histidyl imidazoles.²⁵⁻³² In animals the sixth ligand is another histidyl imidazole while in plants it is the R-group amide of an asparagyl.

While these differences exist, it is generally accepted that the mechanism for peroxidation is consistent for *all* LOXs and is as shown in **Scheme 6.1**. The first key step

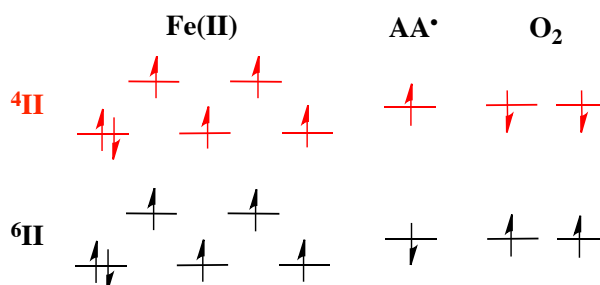
of the process involves a proton-coupled electron transfer (PCET) from the substrate to the Fe(III)-OH moiety to form a pentadienyl-type radical (**I**) while the high-spin Fe(III) is reduced to a high-spin Fe(II) centre, as shown by experimental magnetic susceptibility studies, but now ligated to a H₂O.²⁹ Subsequently, the pentadienyl radical undergoes an attack by O₂ from the *opposite* side of the substrate to that of the Fe centre to give the corresponding peroxy-radical (**II**). Experimental investigations have suggested that this activation of the substrates C—H bond by the Fe(III)-OH group occurs prior to any kinetically productive interaction between the enzyme–substrate complex and O₂.³³ Furthermore, the ferrous and ferric centers in **I** and **II** respectively *do not* bind O₂, if at all, before the formation of the pentadienyl radical.^{33,34}



Scheme 6.1. The generally accepted mechanism of LOXs.

Indeed, experimental EPR studies have shown that there exists an equilibrium between the pentadienyl and peroxy-radicals,^{33,35,36} i.e., the O₂ attack is fully reversible. Thus, in the case of LOXs the enzyme activates the substrate by generating the pentadienyl radical intermediate making it vulnerable to O₂ attack. This contrasts with that observed in most heme-containing enzymes or the family of non-heme α -ketoglutarate dependent dioxygenases that utilize O₂ to oxidize organic substrates:^{2,3} the Fe ion binds the O₂ and

facilitates its activation prior to reaction with the substrate. Thus, in the case of LOXs the experimental evidence suggests that the key pentadienyl + O₂ intermediate (**II**) involves three open-shell centers: the triplet O₂, the substrate-derived radical and the Fe center. Given the high-spin state of the Fe center and the triplet O₂ (i.e., seven unpaired electrons) there are only two possible electronic configurations that allow for spin conservation following the attack of O₂ (i.e., **II** → **III**), illustrated in **Scheme 6.2**. Specifically, it involves the arrangement of electrons such that total spins of 3/2 or 5/2 are obtained generating a quartet or sextet state represented as ⁴**II** and ⁶**II**, respectively.



Scheme 6.2. The electronic configurations for the quartet (red) and sextet (black) ground states of the first intermediate complex (**II**) in the mechanism of LOXs.

This key intermediate (**Scheme 6.1: II**) in the mechanism of the biochemically important LOX family of enzymes presents a clear example of a multi-centered open-shell system. The resulting complex likely involves three open shell centers with a total number of seven unpaired electrons. These open shell centers being the Fe(II), the pentadienyl radical intermediate (hereafter referred to as AA•) and the molecular oxygen. The catalytically relevant arrangement of these electrons leads to the possibility of having the mechanistic intermediate **II** with total spins of 3/2 or 5/2; that is, a quartet (⁴**II**) or sextet (⁶**II**) state (**Scheme 6.2**). It is noted that, as for a singlet bi-radical, a single determinant of KS orbitals (or MOs in the case of HF theory) is incapable of properly describing such multi-centered open-shell systems. Instead, an MC-SCF approach must

be used. Unfortunately, however, such methods are computationally impractical for the study of most biochemical systems due to the often necessary requirement for large chemical models.

In order to assess the applicability of commonly used DFT methods for the study of multi-centered open-shell systems, and as a first step towards obtaining an understanding of the chemistry of LOXs, we have examined the performance of a range of hybrid, meta and hybrid-meta GGA density functionals to reliably determine the structures and energetics of the mechanistic intermediates ^4II and ^6II of 8R-LOX. In addition, we have also considered the effect of varying the %HF contribution in these methods within an ONIOM-type QM/MM model.

6.2 Computational Methods

The Molecular Operating Environment (MOE)³⁷ software package was used to perform all docking and molecular dynamics (MD) annealing and relaxation of the system. These calculations were done with the AMBER99 force field.³⁸

Molecular Docking. The initial crystal structure of 8R-LOX (PDB: 3FG1) was used as a template for docking. All crystallographic waters and counter-ions were removed. The coordinates of hydrogens were then added using the MOE default method. The substrate, modeled as (4Z,7Z)-1,4,7,10-Undecatetraene, was added in the active site manually and oriented such that the H \bullet to be abstracted was within hydrogen distance and interacting with the Fe-OH group (**Figure 6.2**).

Molecular Dynamics Equilibration. With the substrate docked within the active site MD simulations were then performed to allow the solvated enzyme-substrate complex to undergo thermal relaxation. In particular, the initial enzyme-substrate complex was solvated with a 7 Å spherical layer of water molecules. In order to force the system to lie within a volume of space an ellipsoidal potential wall with a scaling constant of 2 was

placed around the solvated enzyme–substrate complex. The damping functional factor included in the electrostatic and van der Waals potentials were set to decay smoothly beyond 8 to 10 Å. Prior to running the simulations the geometry of the solvated complex was optimized using the AMBER99 force field until the root mean square gradient of the total energy fell below 0.05 kcal a.u.⁻¹. The MD simulations were performed under constrained pressure and temperature. The equations of motion were coupled with the Nosé–Poincaré thermostat³⁹ and the time step for numerical integration was set to 1 fs. Initially, the system was heated from 150 to 300 K for a period of 50 ps, followed by an equilibration period of 100 ps at the constant temperature of 300 K and pressure of 1 atm. A typical structure from the trajectory was then optimized with the AMBER99 force field for the subsequent QM/MM analyses (see below).

QM/MM Model. A large active–site chemical model was then obtained from the above final optimized structure for all QM/MM calculations. More specifically, it was chosen such that it included the truncated form of the substrate and all active site residues immediately surrounding it, i.e., first–shell residues (**Figure 6.1**). In addition, all second–shell residues surrounding the Fe center were retained. It is noted that for added consistency the same initial structure was used in all optimizations. As a result, any differences observed in the final optimized structures are due solely to the methods and multiplicity chosen.

The α –carbons of each residue were held fixed at their final MM minimized positions (see above) in order to ensure integrity of the active site during the calculations. Such an approach has been commonly used in the computational investigation of the catalytic mechanisms of enzymes, and its applicability and reliability has been discussed in detail elsewhere.^{40,41} A subset of the complete model centered on the reactive region of the active site was then selected for the high–level QM treatment (**Figure 6.1**: inner circle). Specifically, it consisted of the substrate and those groups directly involved in the reaction. That is, it contained the truncated model substrate the side chains of His385,

His390, His571 and Asn575. In addition it contained the carboxylate of the terminal Ile694, the attacking O₂ and the Fe–OH center. The O₂ was manually added after the simulation in a cavity consisting of the Gly428 residue proposed to be essential in the controlling the O₂ for attack of C8.⁴² To form the first intermediate complex (**II**: **Scheme 6.1**) a hydrogen was manually transferred from the substrate to Fe(III)–OH. This complex was the starting point for all further QM/MM calculations.

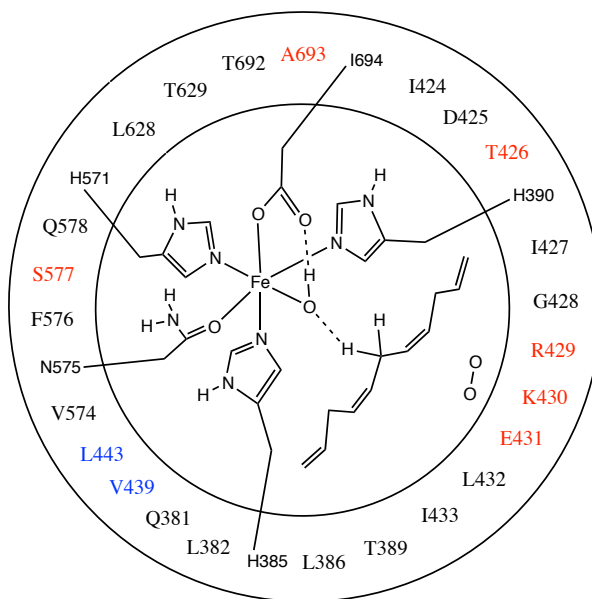


Figure 6.1. Schematic representation of the QM/MM model used. Groups in the inner circle have been modeled at the QM level of theory, while the residues in the outer circle have been modeled at the MM level of theory. Colour code for residues: included in their entirety (black); modeled as Gly, i.e., only the backbone was included with R-groups replaced by a hydrogen (red); modeled as only their R-group, i.e., only their C_α and side chains included (blue).

QM/MM calculations. Combined quantum mechanical and molecular mechanical (QM/MM) methods in the ONIOM formalism⁴³⁻⁵¹ were performed as implemented in the *Gaussian* program suite.⁵² The optimized structures for **4II** and **6II** (**Scheme 6.2**) were

obtained using the ONIOM(DFT_i/BS1:AMBER) level of theory in the mechanical embedding formalism.^{7,8,53,54} DFT_i represents the BP86, B3LYP[±] (10% HF contribution), B3LYP*, B3LYP, M06 and M06-L functionals.¹⁰ The combination of basis functions defined by BS1 was the 6–31G(d) basis set on all atoms but Fe, where the LANL2DZ+ECPs basis set was used. Due to the fixing of atoms within the chemical models the energies reported are the solely electronic energies. Frequency calculations were performed to validate the nature of the stationary point. Single–point energy calculations on the optimized structures were done at the respective ONIOM(DFT_i/BS2//DFT_i/BS1:AMBER) level of theory in the electronic embedding formalism. BS2 is defined as the combination of the 6–311G(2df,p) functions plus the LANL2DZ ECPs for the iron. Diffuse functions were not used because, as discussed by Martin et al.,⁵⁵ their inclusion on Fe is a poor match when used with the underlying 6–31G or 6–311G basis sets.

6.3 Results and Discussion

The effect of DFT functional choice on optimized geometries. As noted above, the optimized geometries of the quartet and sextet states of the intermediate complex **II** (i.e. ⁴**II** and ⁶**II**) in the mechanism of LOXs (see **Scheme 6.1**) were examined using a variety of density functional methods to describe the high (QM) region of the QM/MM model, the MM method (AMBER) being kept constant. Thus, for simplicity only the key structural changes observed in the QM chemical model region, which is schematically shown in **Figure 6.2**, are discussed herein. It is noted, however, that the most significant changes upon changing the DFT method did occur in the QM region. As can be seen in **Figure 6.2** the metal ion forms a single interaction with each of five amino acid residues; the R-group imidazole of 3 histidyl's, the C-terminus of an isoleucyl and the R-group amide oxygen of an asparaginyll. It also forms a sixth metal-ligand (M-L) interaction with

an hydroxyl ion oxygen. Given the distinct nature of the three open-shell centres, a chosen DFT method should give the same or very similar M-L distances for the quartet and sextet states.

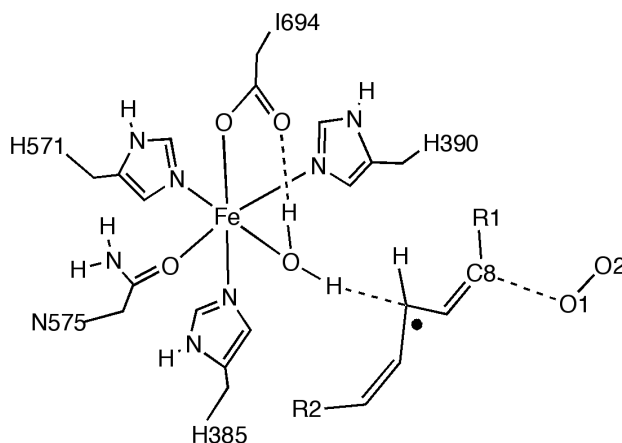


Figure 6.2. The QM-region model used for the active site of 8R-LOX with the pentadienyl intermediate (**II**) bound.

As seen in **Table 6.1**, however, the M–L bond lengths do differ between the structures of the quartet ($^4\mathbf{II}$) and sextet ($^6\mathbf{II}$) states. However, the degree of the differences (reported as the absolute value of Δr between complexes) depends on the functional used. For instance, the averages of the absolute differences (**Table 6.1**; $\text{Average}^{|\Delta r|}$) between the optimized bond lengths for $^4\mathbf{II}$ and $^6\mathbf{II}$ as obtained using the BP86, B3LYP $^\pm$, B3LYP*, B3LYP, M06 and M06-L functionals is 0.076, 0.045, 0.009, 0.039, 0.038 and 0.054 Å, respectively. That is, the B3LYP* functional gives a markedly smaller $\text{Average}^{|\Delta r|}$ value than the other functionals considered. In contrast, the BP86 and M06-L functionals, i.e., those with a 0% Hartree-Fock (HF) contribution, have the greatest differences in M–L bond lengths between states. It is noted, however, that M06-L does perform better than BP86 as indicated by it having an $\text{Average}^{|\Delta r|}$ value that is ~ 0.2 Å smaller. This is likely expected given that unlike the latter functional, M06-L was trained with compounds that contain transition metals.^{56,57}

Table 6.1. Absolute differences ($|\Delta r|$) between optimized ^4II and ^6II structures, obtained at the ONIOM(DFT_i/BS1:AMBER94) level of theory for selected interactions. All absolute differences are in Å.

Interaction	DFT _i					
	BP86	B3LYP [±]	B3LYP [*]	B3LYP	M06-L	M06
Fe-H385	0.109	0.061	0.012	0.047	0.042	0.068
Fe-H390	0.091	0.035	0.002	0.036	0.015	0.064
Fe-H571	0.034	0.021*	0.005*	0.013*	0.032	0.041
Fe-N575	0.050*	0.053*	0.004*	0.056*	0.160	0.008
Fe-I694	0.100*	0.064*	0.021*	0.051*	0.028	0.026*
Fe-OW	0.072	0.035	0.010	0.031	0.050*	0.021*
O-O	0.068*	0.007	0.000	0.000	0.057*	0.047*
^a Average ^{\Delta r}	0.076	0.045	0.009	0.039	0.054	0.038

* complexes in which the quartet bond length is longer. ^a Averages were calculated only for the M–L interactions, that is the O–O bond length was ignored.

For B3LYP and M06 (both of which have 20% HF contribution) similar values of Average^{|\Delta r|} were measured of 0.039 and 0.038 Å respectively. It is interesting to note that while the averages are similar there are significant differences when comparing individual M–L interactions. With respect to individual interaction distances all of the B3LYP-based functionals (B3LYP[±], B3LYP^{*} and B3LYP) predicted the H571···, N575··· and I694···Fe interactions to be shorter, and the H385···, H390··· and HO···Fe interactions to be longer in the sextet state compared to that obtained for the quartet state. Similarly, the BP86 functional predicts the N575···Fe and I694···Fe distances to be shorter in the sextet state. In contrast, both the M06 and M06-L functionals predict the sextet state to have longer M–L interactions for all ligands, the only exception occurring for the optimized I684-COO⁻···Fe distance obtained using the M06-L method.

Similar trends were also observed for the differences in the L_n-M-L_m (where $m \neq n$) bond angles (not shown). Specifically, the BP86 method again gave the largest value (2.8°) for the difference between $^4\mathbf{II}$ and $^6\mathbf{II}$, while the M06 and M06-L functionals gave smaller average values of 1.8° and 2.4° , respectively. The smallest differences were again obtained using the B3LYP-based functionals, B3LYP $^\pm$, B3LYP* and B3LYP, which gave averages of 1.7° , 0.4° and 1.4° , respectively; again B3LYP* gave the smallest error of all functionals considered.

For the triplet O_2 , one of the three "open-shell centres" in intermediate \mathbf{II} , it was also found that the optimized O–O bond lengths obtained within $^4\mathbf{II}$ and $^6\mathbf{II}$ and the size of the differences between them, depended on the choice of functional (**Table 6.1**). For instance, for the quartet and sextet complexes the BP86 functional gave $r(O-O) = 1.306$ and 1.238 \AA respectively, i.e., the O_2 moiety having a longer bond length within the $^4\mathbf{II}$ complex (**Table 6.1**). This was also the largest difference (0.068 \AA) observed in lengths between these two states of any functional considered. For the M06 and M06-L functionals the O–O bond lengths were again predicted to differ significantly though now by 0.047 and 0.057 \AA respectively, with $^4\mathbf{II}$ again having the longer length. In contrast, all of the B3LYP-based functionals had only minor or negligible differences in O_2 bond lengths between the $^4\mathbf{II}$ and $^6\mathbf{II}$ complexes. Indeed, the observed differences for the B3LYP $^\pm$, B3LYP* and B3LYP functionals were just 0.007 , 0.000 and 0.000 \AA , respectively. It is also noted that of these three latter methods, only B3LYP $^\pm$ predicted O_2 to have a slightly longer bond length within the sextet state.

For the substrate radical itself (AA^\bullet), modeled as a pentadienyl radical, all DFT methods considered herein gave C–C bond lengths for the quartet and sextet states that were in very close agreement with each other differing by 0.01 \AA or less. Similar to that noted above for the L_n-M-L_m (where $m \neq n$) bond angles, the M06 and M06-L functionals gave slightly larger differences in bond angles for AA^\bullet in $^4\mathbf{II}$ and $^6\mathbf{II}$ or the bond angles than obtained using the B3LYP-based functionals (B3LYP $^\pm$, B3LYP* and B3LYP).

More specifically, the use of M06 and M06-L gave differences in the carbon-backbone bond angles of $\sim 2.9^\circ$ and $\sim 2.8^\circ$ respectively, while B3LYP $^\pm$, B3LYP* and B3LYP all had negligible differences (0.0°). Much more significant differences amongst the DFT methods were observed for the carbon-backbone dihedral angles. For instance, the BP86, B3LYP $^\pm$, B3LYP* and B3LYP methods all gave differences in backbone dihedral angles between $^4\mathbf{II}$ and $^6\mathbf{II}$ that were $< 2.0^\circ$. In contrast, for some of these same backbone dihedral angles the M06 and M06-L methods gave differences of 14.1° and 12.3° , respectively. This may be due in part to the fact that in the QM/MM optimized structures it was observed that the active sites of $^4\mathbf{II}$ and $^6\mathbf{II}$ remained fairly consistent, except when using the M06 and M06-L functionals. For the latter two methods AA $^\bullet$ shifted considerably in the $^4\mathbf{II}$ complex. As a result, the substrate is able to adopt a markedly more planar geometry than in the corresponding $^6\mathbf{II}$ complexes.

Overall, the above results appear to suggest that within a QM/MM framework the B3LYP* method is the preferred DFT functional for obtaining consistent optimized structures of a QM region containing multi open-shell centres.

The effect of DFT method choice on calculated spin densities in $^4\mathbf{II}$ and $^6\mathbf{II}$. The spin densities of various key sites within $^4\mathbf{II}$ and $^6\mathbf{II}$ were then examined. In particular, the spin densities on the Fe(II) centre, the oxygen atoms of the O₂ moiety (O1 and O2), and the carbon centre (C8) of AA $^\bullet$ that is attacked by the O2. In addition, the sum of the spin densities on carbons on the pentadienyl radical intermediate itself (AA $^\bullet$) were also considered. All of these values obtained are provided in **Table 6.2**.

It can be seen that for both possible multiplicities the spin density (SD) on the Fe center is fairly consistent regardless of the choice of DFT method. More specifically, for $^4\mathbf{II}$ and $^6\mathbf{II}$ it is calculated to lie in the ranges of 3.94 — 3.80 and 3.87 — 3.82, respectively. Such SD values for an Fe(II) have been previously concluded to be indicative of a high spin arrangement.² The BP86 and M06-based methods generally gave values near or at the upper ends of these ranges. For example, use of the BP86 functional

gives values of 3.92 and 3.87 for the Fe(II) centre in ^4II and ^6II respectively (**Table 6.2**). In contrast, the B3LYP* and B3LYP methods generally gave values towards the lower ends of these ranges.

Table 6.2. Selected spin densities in intermediate complexes ^4II and ^6II as calculated at the ONIOM(DFT_i/BS2//DFT_i/BS1:AMBER94)–EE level of theory (see text).

DFT _i		Fe	O1	O2	AA•	C8
BP86	^4II	3.92	−0.93	−0.90	0.55	0.20
	^6II	3.87	0.92	0.90	−0.97	−0.35
B3LYP [±]	^4II	3.84	−0.99	−0.96	0.87	0.30
	^6II	3.85	0.95	0.93	−0.97	−0.34
B3LYP*	^4II	3.80	−1.01	−0.98	0.99	0.36
	^6II	3.82	0.99	0.97	−0.98	−0.36
B3LYP	^4II	3.82	−1.01	−0.98	1.00	0.37
	^6II	3.82	1.01	0.98	−0.98	−0.37
M06	^4II	3.84	−1.01	−0.97	0.71	0.47
	^6II	3.86	1.02	0.97	−0.95	−0.35
M06-L	^4II	3.94	−0.97	−0.91	0.98	0.34
	^6II	3.85	0.95	1.00	−0.98	−0.35

The calculated SDs on O1 and O2 indicate the presence of the triplet species for all methods used with ranges of 0.92–1.02 and 0.90–1.00, respectively. Notably, the BP86 method gives the lowest SD values for O1 and O2 in both states while the B3LYP[±] and M06-L methods give a low SD value for O2 in the sextet and quartet states respectively (see **Table 6.2**). Meanwhile the B3LYP*, B3LYP and M06 methods give SDs on O1 and O2 that are closer to 1 for both ^4II and ^6II .

The most significant variations in calculated spin densities, however, were observed in AA• itself. As can be seen in **Table 6.2**, the B3LYP*, B3LYP and M06-L methods all give absolute SD values for AA• (the sum of the spin densities on carbons on the pentadienyl radical) for both ⁴II and ⁶II that are within 0.02 of 1.0. Furthermore, they also give consistent absolute SD values for the C8 carbon in both states that are approximately 1/3, ranging from 0.34 – 0.37. In contrast, the BP86 method significantly underestimates the absolute SD on both AA• and C8 for the quartet state with values of just 0.55 and 0.20 respectively (**Table 6.2**). Meanwhile, the corresponding absolute values in the sextet state using the same functional are 0.97 and 0.35. This difference in calculated SD values is due to the fact that the use of the BP86 method to describe the QM region in the QM/MM model results in greater delocalization of the SD over the amino acid residues ligated to the Fe(II) centre in the quartet system (⁴II) than in the sextet (⁶II). This is expected given that DFT tends to suffer from delocalization errors.⁵⁸⁻⁶⁰ In fact from **Table 6.2** it can be seen that as the %HF is increased, the spin density AA• in the quartet system approaches a value of 1.00. That is, as the %HF contribution increases a greater localization of density is observed, which is expected given that HF theory tends to cause over-localization.⁵⁹ Consistency between ⁴II and ⁶II is reached when the percent HF contribution included is at 15%; i.e. for the B3LYP* method. However, it is noted that even with the use of the B3LYP[±] method (i.e., 10% HF contribution) only a modest underestimation of the SD values of AA• and C8 in the quartet state compared to the sextet state values is observed (**Table 6.2**). More specifically, the absolute SD values of AA• and C8 in the quartet state are 0.10 and 0.04 lower respectively than obtained for the sextet state using the same functional.

For ⁴II the use of the M06 functional provides an unexpected result. With the use of M06 (which has 20% HF contribution) a marked underestimation of the absolute SDs on AA• (0.71) with a concomitant overestimation of the value on C8 (0.47) is observed. This was again found to be due to greater delocalization of SD over the amino acid residues

ligated to the Fe(II) center in ${}^4\mathbf{II}$ compared to that in ${}^6\mathbf{II}$. However, as stated above, M06-L (which has 0% HF contribution) gives an absolute SD value for AA• for both the quartet and sextet complexes very close to 1.0.

Thus, by examining the performance of a range of hybrid, meta and hybrid-meta GGA density functionals the above results suggest that at least for ONIOM QM/MM calculations, a reliable description of the spin density distribution can be obtained when the B3LYP*, B3LYP or M06-L functionals are used to describe the QM-region. Thus, it appears that while significant differences in geometry existed between ${}^4\mathbf{II}$ and ${}^6\mathbf{II}$ for the B3LYP and M06-L functionals (in comparison to B3LYP*) the proper description of the electronic distribution appears to be less sensitive to geometrical differences, but rather is sensitive to the functional used.

The effect of DFT functional choice on the relative energies of ${}^4\mathbf{II}$ and ${}^6\mathbf{II}$. The reliable and accurate calculation of the thermochemistry of a chemical system is important not only in the elucidation of enzymatic mechanisms but arguably, is a common goal of computational studies. Hence, the ability of the various DFT methods to reliably calculate the relative energy difference between the ${}^4\mathbf{II}$ and ${}^6\mathbf{II}$ complexes, i.e., between the quartet and sextet states of the 8R-LOX mechanistic intermediate \mathbf{II} , was also examined. The results obtained at the ONIOM(DFT_i/BS2//DFT_i/BS1:AMBER) level of theory within the electronic embedding formalism (EE) are presented in **Table 6.3**. That is, relative energies were calculated by performing single-point calculations in which the same DFT functional was used to describe the QM-region as for the optimized geometry being used, but now in combination with a significantly larger basis set (BS2; see Computational Methods).

Table 6.3. The relative energies of the intermediate complexes ${}^4\text{II}$ and ${}^6\text{II}$ obtained at the different ONIOM(DFT_i/BS2//DFT_i/BS1:AMBER94)–EE levels of theory.

Functional		$\Delta E(\text{kcal/mol})$
BP86	${}^4\text{II}$	10.9
	${}^6\text{II}$	0.0
B3LYP $^\pm$	${}^4\text{II}$	-0.4
	${}^6\text{II}$	0.0
B3LYP*	${}^4\text{II}$	0.0
	${}^6\text{II}$	0.0
B3LYP	${}^4\text{II}$	-0.6
	${}^6\text{II}$	0.0
M06	${}^4\text{II}$	11.0
	${}^6\text{II}$	0.0
M06-L	${}^4\text{II}$	10.8
	${}^6\text{II S}$	0.0

From **Table 6.3** it can be clearly seen that when the B3LYP $^\pm$, B3LYP* and B3LYP functionals are used to describe the QM region, the energy differences between the quartet and sextet states are within 1 kcal mol $^{-1}$ with the largest difference occurring for B3LYP (0.6 kcal mol $^{-1}$). It is noted that both the B3LYP $^\pm$ and B3LYP functionals predict the quartet (${}^4\text{II}$) to lie marginally lower in energy than the sextet (${}^6\text{II}$). However, as noted in the introduction Hughes and Friesner¹⁹ found that for several transition metal complexes reduction from 20% to 15% HF contribution in B3LYP appeared to cause larger errors in the electronic transitions when compared to experimental values. However, upon closer inspection of their data the B3LYP* method in fact appears to give improved results for the Fe containing compounds. In particular smaller errors in the $e_g \rightarrow t_{2g}$ and $t_{2g} \rightarrow e_g$ spin-forbidden transitions for Fe(TRENCAM) $^{-3}$ and Fe(CN) $_6^{+3}$

respectively were obtained. In contrast, for the Ni, Cr and Mn complexes that they considered a reduction to 15% HF contribution did result in larger errors in the electronic transitions when compared to experimental values.

With the use of the BP86 functional the quartet is predicted to be 10.9 kcal mol⁻¹ higher in energy than the sextet. Interestingly, with both the M06 and M06-L functionals energy differences very similar to those obtained with the BP86 functional are observed (see **Table 6.3**). In particular, the M06 functional predicts the quartet to be 11.0 kcal mol⁻¹ higher in energy than the sextet while the M06-L functional similarly calculates it to be 10.8 kcal mol⁻¹ higher in energy. Vancoille et al.⁶¹ have previously shown that for several heme models, the M06 and M06-L functionals overstabilized the high-spin state with respect to the low-spin state. However, as discussed above, the optimized geometries obtained for **4II** and **6II** using the M06 and M06-L functionals to describe the QM-region show significant differences. Thus, it is unclear if the differences in energy between the quartet and sextet states observed for these two functionals are due to the functionals themselves or differences in geometries.

Hence, relative energies were then re-determined via single point calculations on the optimized geometries of the quartet and sextet obtained at the ONIOM(B3LYP*/BS1:AMBER)-ME level of theory; the level that gave the most consistent agreement between the quartet and sextet geometries. That is, relative energies were determined at the ONIOM(DFT_γ/BS2//B3LYP*/BS1:AMBER)-EE level of theory and are presented in **Table 6.4**.

As can be seen the energy difference between the two multiplicities has now been reduced significantly to just 0.9 and 1.1 kcal mol⁻¹ for the M06 and M06-L functionals, respectively! It is interesting to note that both functionals predict the quartet state to be slightly higher in energy. This is in contrast to that observed for B3LYP[±] and B3LYP (**Table 6.3**). Regardless, it appears that the failure of M06 and M06-L lies predominantly in the determination of the geometry of the quartet system and not preferential

stabilization of high- or low-spin. This is further illustrated by the fact that calculation of relative energies at the ONIOM(B3LYP*/BS2//M06/BS1:AMBER)-EE and ONIOM(B3LYP*/BS2//M06-L/BS1:AMBER)-EE levels of theory gives energy differences between the quartet and sextet of 5.2 and 8.2 kcal mol⁻¹, respectively.

Table 6.4. The energies for the quartet and sextet systems at the ONIOM(M06-L/BS2//B3LYP*/BS1:AMBER94) and ONIOM(M06/BS2//B3LYP*/BS1:AMBER94) levels of theory in the electronic embedding formalism.

Functional	Mult.	$\Delta E(\text{kcal/mol})$
M06-L	⁴ II	1.1
	⁶ II	0.0
M06	⁴ II	0.9
	⁶ II	0.0

6.4 Conclusions

The performance of a range of hybrid, meta and hybrid-meta GGA density functionals to reliably provide geometries, spin densities and relative energies of multi-centered open-shell complexes within an ONIOM QM/MM methodology has been examined. More specifically, the ability of the BP86, B3LYP[±], B3LYP*, B3LYP, M06 and M06-L functionals to provide reliable structures, spin densities and relative energies of a multi-centered open-shell mechanistic intermediate complex **II** in the mechanism of the non-heme iron metalloenzyme 8R-LOX was considered. The latter complex **II** contains three open shell centers; an Fe(II), a substrate-derived pentadienyl radical and a molecular oxygen (O₂).

From the results obtained the B3LYP* functional, i.e., the B3LYP functional but now with only a 15% rather than a 20% contribution from Hartree-Fock, appears to provide the most reliable geometries. In particular, the use of the B3LYP* method gave the smallest average differences between the catalytically relevant quartet (^4II) and sextet (^6II) complexes.

In contrast, reliable descriptions of the spin density distribution appeared to be obtained using the B3LYP*, B3LYP or M06-L functionals. Thus, while the B3LYP and M06-L functionals appeared less suited for the proper description of the geometries of ^4II and ^6II , they are capable of properly describing their electronic distribution.

In the case of the relative energies between ^4II and ^6II the use of the B3LYP* functional at the ONIOM(DFT_i/BS2//DFT_i/BS1:AMBER) level of theory within the electronic embedding formalism provided a difference of 0.0 kcal mol⁻¹ between the two states. However, B3LYP[±] and B3LYP also predicted differences in energies of less than 1 kcal mol⁻¹. The erroneously large differences in relative energies predicted using the M06 and M06-L functionals, when using geometries obtained using the same functionals, was found to be due to errors in their underlying optimized geometries. The use of more reliable structures of ^4II and ^6II lead to M06 and M06-L predicting only small relative energy differences between the two complexes.

Overall, the results obtained suggest that for systems with multiple centers having unpaired electrons the B3LYP* appears most well rounded to provide reliable geometries, electronic structures and relative energies.

6.5 References

- (1) Robinet, J. J.; Gault, J. W. *J. Phys. Chem. B* **2008**, *112*, 3462-3469.
- (2) Liu, H. N.; Llano, J.; Gault, J. W. *J. Phys. Chem. B* **2009**, *113*, 4887-4898.

- (3) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949-1017.
- (4) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421-437.
- (5) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729-738.
- (6) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695-701.
- (7) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- (9) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
- (10) Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *107*, 48-55.
- (11) Kepenekian, M.; Calborean, A.; Vetere, V.; Le Guennic, B.; Robert, V.; Maldivi, P. *J. Chem. Theory Comput.* **2011** (Web).
- (12) Ghosh, A.; Persson, B. J.; Taylor, P. R. *J. Biol. Inorg. Chem.* **2003**, *8*, 507-511.
- (13) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643-651.
- (14) Lundberg, M.; Morokuma, K. *J. Phys. Chem. B* **2007**, *111*, 9380-9389.
- (15) Reiher, M. *Inorganic Chemistry* **2002**, *41*, 6928-6935.
- (16) Schenk, G.; Pau, M. Y. M.; Solomon, E. I. *J. Am. Chem. Soc.* **2004**, *126*, 505-515.
- (17) Siegbahn, P. E. M.; Blomberg, M. R. A.; Chen, S. L. *J. Chem. Theory Comput.* **2010**, *6*, 2040-2044.
- (18) Salomon, O.; Reiher, M.; Hess, B. A. *Journal of Chemical Physics* **2002**, *117*, 4729-4737.
- (19) Hughes, T. F.; Friesner, R. A. *J. Chem. Theory Comput.* **2011**, *7*, 19-32.
- (20) Neau, D. B.; Gilbert, N. C.; Bartlett, S. G.; Boeglin, W.; Brash, A. R.; Newcomer, M. E. *Biochemistry* **2009**, *48*, 7906-7915.
- (21) Borowski, T.; Broclawik, E. *J. Phys. Chem. B* **2003**, *107*, 4639-4646.
- (22) Zhang, Y.; Gan, Q. F.; Pavel, E. G.; Sigal, E.; Solomon, E. I. *J. Am. Chem. Soc.* **1995**, *117*, 7422-7427.

- (23) Ivanov, I.; Heydeck, D.; Hofheinz, K.; Roffeis, J.; O'Donnell, V. B.; Kuhn, H.; Walther, M. *Arch. Biochem. Biophys.* **2010**, *503*, 161-174.
- (24) Minor, W.; Steczko, J.; Bolin, J. T.; Otwinowski, Z.; Axelrod, B. *Biochemistry* **1993**, *32*, 6320-6323.
- (25) Glickman, M. H.; Klinman, J. P. *Biochemistry* **1995**, *34*, 14077-14092.
- (26) Dunham, W. R.; Carroll, R. T.; Thompson, J. F.; Sands, R. H.; Funk, M. O. *Eur. J. Biochem.* **1990**, *190*, 611-617.
- (27) Nelson, M. J. *J. Am. Chem. Soc.* **1988**, *110*, 2985-2986.
- (28) Zhang, Y.; Gebhard, M. S.; Solomon, E. I. *J. Am. Chem. Soc.* **1991**, *113*, 5162-5175.
- (29) Gaffney, B. J. *Annu. Rev. Biophys. Biomolec. Struct.* **1996**, *25*, 431-459.
- (30) Cheesbrough, T. M.; Axelrod, B. *Biochemistry* **1983**, *22*, 3837-3840.
- (31) Petersson, L.; Slappendel, S.; Vliegthart, J. F. G. *Biochim. Biophys. Acta* **1985**, *828*, 81-85.
- (32) Vanderheijdt, L. M.; Feiters, M. C.; Navaratnam, S.; Nolting, H. F.; Hermes, C.; Veldink, G. A.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1992**, *207*, 793-802.
- (33) Glickman, M. H.; Klinman, J. P. *Biochemistry* **1996**, *35*, 12882-12892.
- (34) Chamulitrat, W.; Mason, R. P. *J. Biol. Chem.* **1989**, *264*, 20968-20973.
- (35) Nelson, M. J.; Cowling, R. A.; Seitz, S. P. *Biochemistry* **1994**, *33*, 4966-4973.
- (36) Nelson, M. J.; Seitz, S. P.; Cowling, R. A. *Biochemistry* **1990**, *29*, 6897-6903.
- (37) Molecular Operating Environment; 2010.10 ed.; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2010.
- (38) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049-1074.
- (39) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. *J. Comput. Phys.* **1999**, *151*, 114-134.
- (40) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232-240.
- (41) Llano, J., Gault, J. W. In *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta, C. F., Ed.; Wiley-VCH: Weinheim, 2010; Vol. 2, p 920.

- (42) Coffa, G.; Brash, A. R. *Proc. Nat. Acad. Sci. U.S.A.* **2004**, *101*, 15579-15584.
- (43) Bearpark, M. J.; Ogliaro, F.; Vreven, T.; Boggio-Pasqua, M.; Frisch, M. J.; Larkin, S. M.; Robb, M. A. In *Computation in Modern Science and Engineering Vol 2, Pts a and B*; Simos, T. E., Maroulis, G., Eds.; Amer Inst Physics: Melville, 2007; Vol. 2, p 583-585.
- (44) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Theochem-J. Mol. Struct.* **1999**, *461*, 1-21.
- (45) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959-1967.
- (46) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179.
- (47) Morokuma, K.; Musaev, D. G.; Vreven, T.; Basch, H.; Torrent, M.; Khoroshun, D. V. *IBM J. Res. Dev.* **2001**, *45*, 367-395.
- (48) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
- (49) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.
- (50) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432.
- (51) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760-769.
- (52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachar, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann,

R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian 09, Revision B.01, Gaussian, Inc.: Wallingford CT, 2010.

(53) Lee, C., Yang, W., Parr, R.G. *Phys. Rev. B: Condens. Matte* **1988**, *37*, 785-789.

(54) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(55) Martin, J.; Baker, J.; Pulay, P. *J. Comput. Chem.* **2009**, *30*, 881-883.

(56) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(57) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.

(58) Boguslawski, K.; Marti, K. H.; Legeza, Ö.; Reiher, M. *J. Chem. Theory Comput.* **2012**, *8*, 1970-1982.

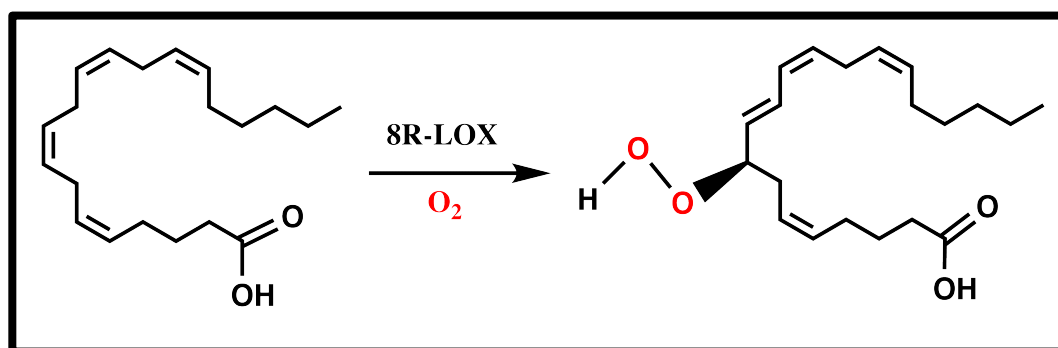
(59) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289-320.

(60) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Science* **2008**, *321*, 792-794.

(61) Vancoillie, S.; Zhao, H. L.; Radon, M.; Pierlout, K. *J. Chem. Theory Comput.* **2010**, *6*, 576-582.

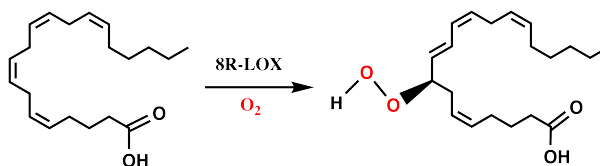
Chapter 7

Gaining insight into the chemistry of lipoxygenases: A computational investigation into the catalytic mechanism of (8*R*)-LOX



7.1 Introduction

Lipoxygenases (LOXs) are a ubiquitous family of non-heme iron enzymes found in bacteria, plants and animals.¹⁻⁵ Importantly, they catalyze the stereo- and regio-specific peroxidation by molecular oxygen (O_2) of fatty-unsaturated acids that contain one or more (1Z,4Z)-pentadiene systems such as linoleic (LA) or arachidonic (AA) acid as shown in **Scheme 7.1**.^{1,6-9} Little is known about the roles of the products formed by such enzymes in plants. In contrast, in the case of mammals the products of LOXs, known as eicosanoids, act as potent cell effector molecules and are critical for normal cell function. Indeed, they have been linked to several physiological disorders including atherosclerosis and cancer.^{10,11} They have also been shown to be important in pathogenic bacteria;^{5,11-13} for example, acting as anaphylactic and inflammatory agents.^{1,5,12,14-16} Thus, it has been suggested that a greater understanding of the mechanism by which LOXs oxidize their substrates may also provide valuable insights into the design of novel pharmaceuticals.⁵



Scheme 7.1. The overall peroxidation of arachidonic acid as catalyzed by the LOX family member 8R-LOX to give the product 8R-hydroperoxyeicosatetraenoic acid (8R-HPETE).

Experimentally it has been shown that in the native resting LOX enzymes the iron exists in its ferrous form, i.e., Fe(II). However, upon activation it is oxidized to its ferric form, i.e., Fe(III).^{17,18} *In vitro* activation, oxidation of the Fe center, has been shown to occur only after the addition of the peroxide product (8R-HPETE) itself.^{14,19} Based on magnetic susceptibility studies and crystallographic data it has been concluded that both the Fe(II) and Fe(III) centers exist as high spin hexa-coordinated centers that are ligated

to an hydroxyl (OH⁻) and five active site residues (**Figure 7.1**).^{11,20-26} While the LOX enzymes have been found to possess a high degree of sequence similarity, even between highly evolutionary diverged species,²⁷⁻³⁰ differences are observed in the latter ligating residues. In the case of plant LOXs they are three histidyl and an asparagyl residue, and the C-terminus carboxylate.^{1,13} In contrast, in mammalian LOXs a histidyl replaces the asparagyl residue.^{11,13} It has been noted that despite these differing coordination environments the ferric form of 15-hLOX (human-LOX) and soybean-LOX-1 (sLO-1) give very similar ESR spectra.¹¹

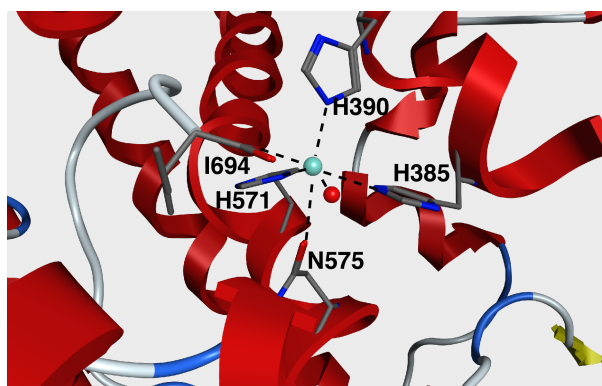


Figure 7.1. The Fe coordination site as observed in an X-ray Crystal Structure of the holoenzyme form of 8R-LOX (PDB: 3FG1).⁶

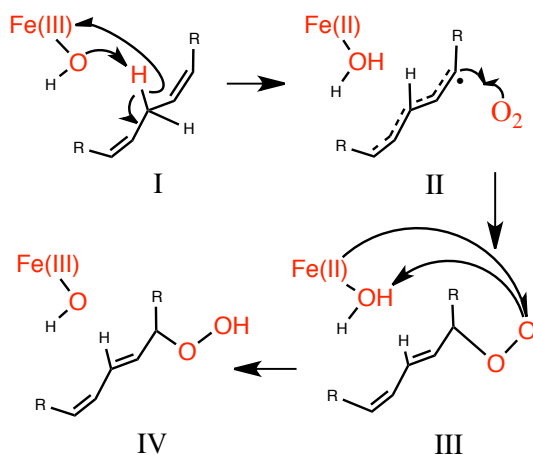
Due to the commonalities observed amongst the LOX family members it is thought that they share similar reactivities and catalytic mechanisms. It is noted that as a result sLO-1 is often used as a prototypical LOX in investigations on their mechanism and reactivity even from different species.^{7,14,31} The generally accepted mechanism for LOXs is shown in **Scheme 7.2**.

The overall peroxidation process is thought to begin with abstraction of a hydrogen atom from the -CH₂- lying between the two -HC=CH- groups of the substrate by the Fe(III)-OH moiety (**Scheme 7.2-I**). Based on the fact that experimentally a large primary kinetic isotope effect (KIE) of ~80 has been measured, it has been concluded that this

step proceeds via a tunneling mechanism.^{9,32} Furthermore, evidence for such a tunneling has been seen in both plant and mammalian LOXs.³³ This step has been investigated both experimentally and computationally in detail.^{7,14,20,29,31,33-38} In particular, Hatcher et al.³⁸ applied multistate continuum theory to the H[•] abstraction step catalyzed by sLO.³⁹⁻⁴¹ Using such an approach they were able to reach agreement with the experimental temperature dependence of the calculated rate as well as correctly predict the primary kinetic isotope effect. Moreover, they showed that the H[•] transfer step in fact occurred via a proton-coupled electron transfer (PCET) process. Olsson et al.³⁵ also investigated the initial H[•] transfer process. First, using a DFT-based empirical valence bond method, they obtained a PES for a chemical model consisting of the reacting system and its surrounding protein+water environment. Then, using a quantum classical path version of the centroid path integral approach they were able to obtain the quantum correction to the classical activation free energy.^{42,43} They also obtained primary KIE effects in good agreement with those obtained experimentally. In addition, again in agreement with experiment, they also showed that Fe(III)-bound OH is involved in abstraction of the substrate's H[•].^{7,44} Furthermore, however, the H⁺ and e⁻ of the PCET step were shown to each have their own unique acceptor: the electron going to the Fe center, while the H⁺ goes to the -OH (**Scheme 7.2-I**).^{33,38}

This step is followed by attack of a molecular oxygen (O₂) at the substrate, antarafacial to the Fe center as shown in **Scheme 7.2-II**. Experimental EPR spectroscopic studies have suggested that this attack of O₂ is fully reversible. Indeed, further experimental evidence has been obtained indicating that both the organic substrate and peroxy radical exist in equilibrium.^{28,31,45} Computationally, the steps following hydrogen transfer are considerably less well studied. Borowski et al.⁷ used a DFT-cluster approach consisting of a model of the intermediate radical and O₂ within a gas-phase environment, in the absence of the Fe center, to investigate this reaction step. They concluded that O₂

attack occurs with a small barrier of approximately 2 kcal mol⁻¹ and that the peroxy-radical derivative (**7.2-III**) lies decidedly lower in energy than **7.2-II** (see **Scheme 7.2**).



Scheme 7.2. The general mechanism of Lipoyxygenases (LOXs).

The last step in the overall mechanism is reduction of **7.2-III** via transfer of a H[•] from the Fe(II)–OH₂ moiety to give the final desired hydroperoxy product (**Scheme 7.2-IV**). Unfortunately, how such a transfer is achieved between two moieties that appear to at least initially be on opposite sides of the fatty acid chain is unknown.¹⁴ However, Borowski et al.⁷ used a DFT-cluster approach to also computationally investigate this step in the mechanism of sLO-1. Notably, the active-site intermediate chemical cluster utilized consisted of models of only the peroxy-intermediate derivative and the non-heme iron and its ligands. Based on their results they concluded that the most likely pathway involved the formation of a seven coordinate Fe–peroxy intermediate known as the purple intermediate.^{31,46} An experimental X-ray crystal structure has been obtained of a LOX with the peroxy derivative bound to its Fe-center in such an arrangement.¹⁵ However, it has also been suggested that due to the steric restrictions imposed on the substrate by the active site residues this purple intermediate is not catalytically relevant.^{46,47} Moreover, it is only experimentally observed upon addition of excess

product, needed to oxidize the Fe(II) center (as discussed above).^{8,46} Thus, a number of important questions unfortunately still remain about the reactivity and mechanism of LOXs.^{6,9,14} This is further complicated by the fact that there are currently no X-ray crystal structures available of the initial enzyme–substrate complex.^{6,9}

In this present investigation we have used a number of modern computational chemistry methods to collectively gain a better understanding of the mechanism of LOXs after the hydrogen transfer has occurred. More specifically, Docking and Molecular Dynamics (MD) simulations involving the entire enzyme have been utilized to obtain a representative bound enzyme-substrate complex. In addition, ONIOM QM/MM methods in combination with large extensive active site and environment models derived from the MD studies have been utilized to investigate the overall catalytic mechanism of 8R-LOX.

7.2 Computational Methods

The Molecular Operating Environment (MOE)⁴⁸ software package was used to perform the docking and molecular dynamics (MD) simulations, all of which utilized the AMBER99 force field.⁴⁹

Molecular Docking. The X-ray crystal structure of 8R-LOX (PDB ID: 3FG1)⁶ was used as a template for docking. All crystallographic waters and counter-ions were removed and the coordinates of hydrogens were then added using the MOE default method. The substrate (AA), modeled as (4Z,7Z)-1,4,7,10-Undecatetraene, was added in the active site manually and oriented such that the leaving H• is within H-bonding distance to the Fe-OH center.

Molecular Dynamics Relaxation. The enzyme–substrate complex was solvated with a 7-Å spherical layer of water molecules. In order to force the system to lie within the volume of space, an ellipsoidal potential wall with a scaling constant of 2 was placed around the solvated complex. The damping functional factor included in the electrostatic

and van der Waals potentials was set to decay smoothly from 8 to 10 Å. The simulations were performed under constrained standard pressure and temperature. The equations of motion were coupled with the Nosé–Poincaré thermostat⁵⁰ and the time step for numerical integration was set to 1 fs. Initially, the system was heated from 150 to 300 K for a period of 50 ps, followed by an equilibration period of 100 ps at the constant temperature of 300 K and pressure of 1 atm. It should be noted that prior to the simulation the geometry of the solvated complex was optimized using the AMBER99 force field until the root mean square gradient of the total energy fell below 0.05 kcal a.u.⁻¹. Following the MD simulation a typical structure from the trajectory was then optimized with the AMBER99 force field for the final analyses.

QM/MM calculations. Combined quantum mechanical and molecular mechanical (QM/MM) methods in the ONIOM formalism⁵¹⁻⁵⁹ were applied as implemented in the *Gaussian03* and *Gaussian09* program suites.^{60,61} Previously, it has been shown that B3LYP* (B3LYP in which the % Hartree-Fock contribution has been reduced to 15%) optimized structural parameters of metal coordination sites are often in good agreement with experimental X-ray structures.⁶² Importantly, in a previous investigation by us it was found that greatest agreement in structures, energetics and electronic distributions for intermediate II (cf. **Scheme 7.2**) was obtained with reduction of the %HF contribution to 15%.⁶³ Hence, optimized geometries of stationary points along the potential energy surface (PES) were obtained at the ONIOM(B3LYP*/BS1:AMBER94) level of theory in the mechanical embedding (ME) formalism.^{62,64-67} The combination of basis functions defined by BS1 is the 6-31G(d) basis set on all atoms but Fe, where LANL2DZ+ECP's was used (i.e. LACVP* basis set). Harmonic vibrational frequencies of all stationary points were obtained at the same level of theory in order to characterize them as either minima or transition states. Relative energies were obtained by performing single-point energy calculations on the above optimized structures at the ONIOM(B3LYP*/BS2:AMBER94) level of theory in the electronic embedding (EE)

formalism. The combination of basis functions defined by BS2 was the 6-311G(2df,p) basis set on all atoms where the ECPs contained in LANL2DZ were used for the core electrons of Fe. Due to the fixing of atoms within the chemical model only ONIOM energies are used herein unless otherwise noted. The AMBER94 charges and parameters for the QM layer, which include the Fe binding site and truncated substrate, were taken from MOE.

It should be noted that we also obtained optimized structures of **RC**, **TS1** and **IC1**, in their quartet and octet configurations, in order to assess the importance of adding polarization functions on Fe, specifically *f*-functions. In particular, for the set of *f*-functions an exponent value of 2.462 and coefficient value of 1.000 were used; taken from the LANLDZ(*f*) basis set contained within the EMSL Basis Set Exchange.^{68,69} Importantly, it was found that for these structures very little changes were observed in the coordination geometry of the Fe center upon increasing the basis set on Fe from LANL2DZ to LANL2DZ(plus additional *f*-functions on Fe from LANLDZ(*f*)). In fact, the measured RMSDs for the change in the Fe–L bond lengths for the six complexes investigated was quite minor at only 0.030 Å, while those for the L_n–Fe–L_m (m≠n) angles for the six structures was just 1.4° (see Appendix, **Table C4**). Thus, it was concluded that the present choice of the LACVP* basis set for obtaining optimized geometries represented an acceptable compromise.

QM/MM Chemical Model: The large active-site chemical model illustrated in **Figure 7.2** was used to investigate the catalytic mechanism. It included the truncated form of AA (modeled as (4Z,7Z)-1,4,7,10-Undecatetraene) and all active site residues immediately surrounding it, i.e., first-shell residues. In addition, all second-shell residues surrounding the Fe center were retained. It is noted that no water molecules were found to exist within this environmental shell as seen in **Figure 7.2**.

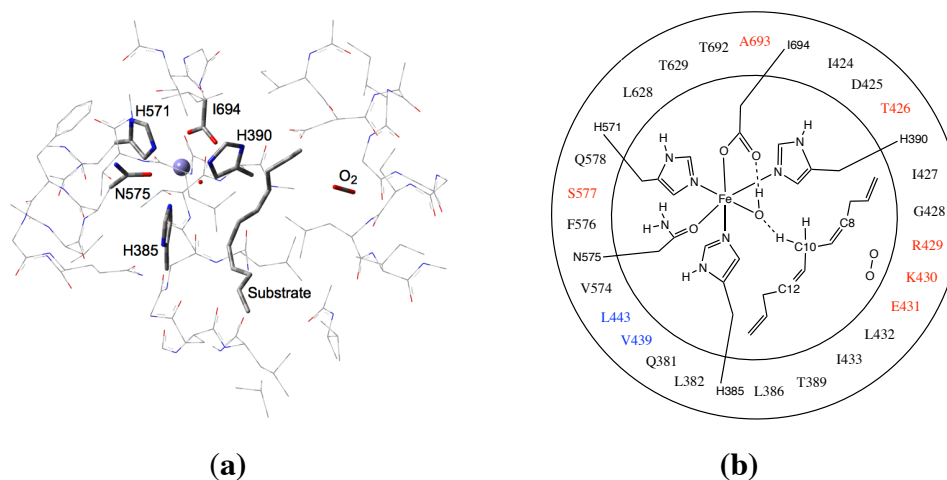


Figure 7.2. The QM/MM model used in the investigation of 8R-LOX. **(a)** Those components included in the high layer (QM) are shown in tube and ball format while those of the low layer are shown in wire format. For clarity, hydrogen atoms have not been shown. **(b)** Schematic representation of the QM/MM model used: groups in the inner circle were included in the QM region while those in the outer circle were included in the MM region. Note, residues in black were included in their entirety, those in red only had their peptide backbone included, while those in blue only included their side chains.

It should be noted that the α -carbons of each residue were held fixed at their final MM minimized positions. Such an approach is common in computational investigations of the catalytic mechanisms of enzymes, and its applicability and reliability has been discussed in detail elsewhere.^{70,71} A subset of the complete model, centered on the reactive region of the active site was then selected for the high-level QM treatment. Specifically, it consisted of the substrate and those groups directly involved in the reaction. That is, the model contained the truncated substrate and the side chains of His385, His390, His571 and Asn575. In addition, it contained the carboxylate of Ile694, the attacking O_2 and the Fe-OH center. It should be noted that, as described in the Introduction, kinetic experiments have shown that O_2 attacks antarafacial to the leaving H^\bullet ,^{11,72} and that it does

not bind to the ferric or ferrous center prior to the formation of the peroxy radical.^{31,73} Hence, in the present models the O₂ moiety was manually added after the MD simulations into the cavity formed by Gly428 which has previously been proposed to control the selectivity of the O₂ attack at C8 of the substrate.⁷⁴

7.3 Results and Discussion

For LOXs, several multiplicities are possible with varying electronic configurations on the Fe center. As previously noted, experimental magnetic susceptibility studies have shown that throughout the reaction the Fe center exists in a high-spin state.¹¹ In the present study it was found that in the case of a doublet multiplicity, similar to the octet surface (see below), a favourable pairing of electrons enabling C—O bond formation (**Figure 7.3: TS2**) is not possible. Moreover, the Fe-center would be in an intermediate spin configuration and thus, would not likely represent a viable species in the catalytic mechanism of 8R-LOX. Thus, only systems with total multiplicities of four, six and eight were investigated herein. The resulting potential energy surfaces (PESs) generated are presented in **Figure 7.3** while the respective optimized geometries of the species involved, with selected bond lengths, are compiled in **Figures 7.5, 7.6, 7.9** and **7.10**.

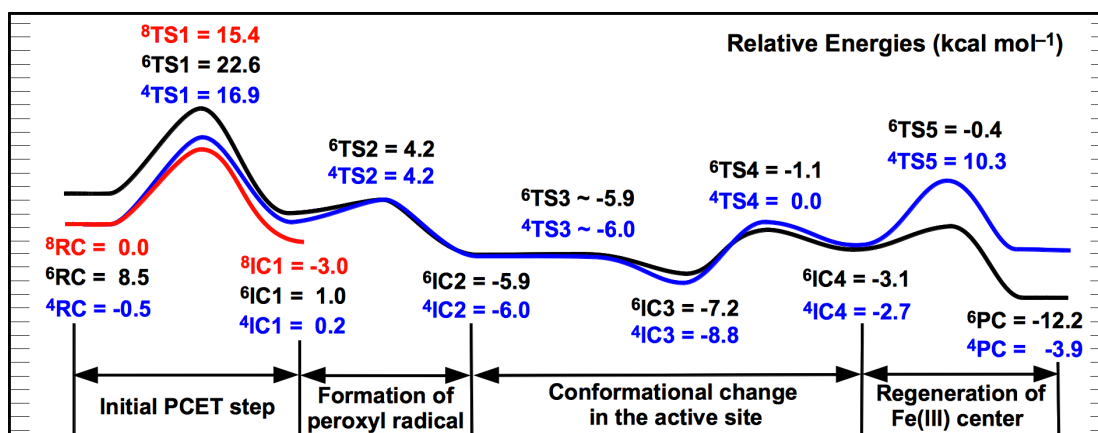


Figure 7.3. Schematic illustration of the PESs obtained, at the ONIOM(B3LYP*/BS2//B3LYP*/BS1:AMBER94)-EE level of theory, for the overall

catalytic mechanism of LOX. Color code: red, black and blue surfaces represent the octet, sextet and quartet systems, respectively.

Quartet, Sextet and Octet Reactive Complexes. We began by comparing the energies and structures of the fully-bound initial reactive complexes; that is, when the substrate and O₂ are bound within the active site.

It should be noted that for comparing relative energies along the PESs, the reactive complex with octet multiplicity (⁸RC) was chosen as the reference point. Importantly, however, ⁴RC lies only slightly lower in energy than ⁸RC by 0.5 kcal mol⁻¹, i.e., they are essentially thermo-neutral (**Figure 7.3**). The only difference between ⁴RC and ⁸RC is the direction of spin for the electrons on O₂ as shown in **Figure 7.4**. Thus, the presence of O₂ appears to have little effect on the energy of the high-spin Fe-center. In contrast, ⁶RC is 8.5 kcal mol⁻¹ higher in energy than ⁸RC due to the differing electronic configuration on the Fe center (**Figure 7.4**). In particular, the Fe-center now exists in an intermediate-spin state and not a high-spin state as expected in an ideal octahedral (**Figure 7.4**).

The optimized geometries of the three reactive complexes obtained (⁸RC, ⁶RC and ⁴RC) are schematically shown in **Figure 7.5**. Due to the complexities inherent in the large QM/MM model, it is not possible to show the entire structures, but nor is it necessary. The largest structural differences between the reactive complexes was observed to occur in the iron···ligand (i.e., Fe-L) bonds.

In ⁸RC, it is noted that the residues all have similar Fe-L interaction lengths. For instance, for the three histidyl residues the distances between their imidazole nitrogen and the Fe of 2.17 (His571), 2.20 (His390) and 2.30 (His385) Å. Meanwhile, the interaction between the R-group amide oxygen of Asn575 and Fe has a distance of 2.30 Å (see **Figure 7.5**) while the shortest Fe-L length is observed for Fe-OH, $r(\text{Fe-OH}) = 1.83$ Å. Very similar distances were obtained for ⁴RC. Indeed, for the six Fe-L bond lengths an RMSD of 0.02 Å is obtained for differences between ⁴RC and ⁸RC (**Figure 7.5**). The

largest change of 0.04 Å occurs for Fe–His390. These small differences are expected given that the only difference in the octet and quartet multiplicities arises from flipping of the unpaired spins on the O₂, which is removed from the Fe-center.

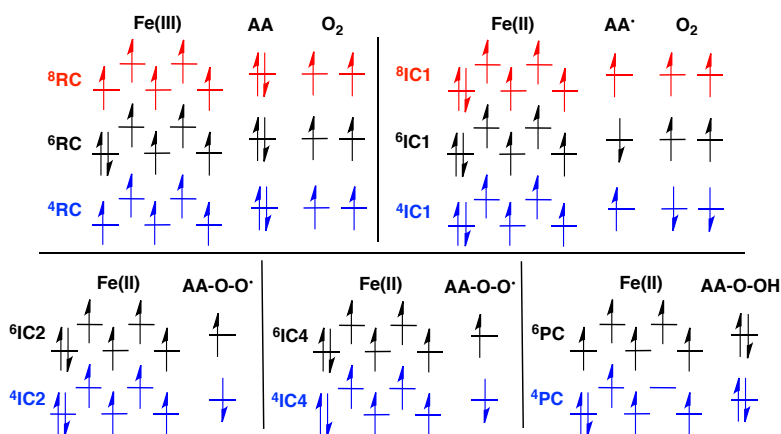


Figure 7.4. The electronic configurations obtained from the spin densities at the ONIOM(B3LYP*/BS2//B3LYP*/BS1:AMBER94)–EE level of theory. While the orbitals of the Fe center are not expected to be those for an ideal octahedral complex they have been drawn that way to simplify discussion. Note that the electronic configurations presented above only represent key structures during the mechanism of 8R–LOX.

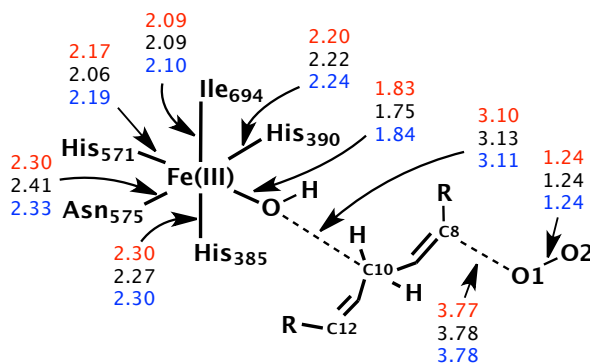


Figure 7.5. Schematic illustration of the optimized structures of the reactant complexes (RC) with key selected distances (Ångstroms) shown, as obtained at the ONIOM(B3LYP/6-31G(d):AMBER)–ME level of theory. Color code: blue (quartet), black (sextet), red (octet).

Activation of the substrate: the initial PCET step. The overall catalytic mechanism of LOXs begins with abstraction of a H• from the substrate via a PCET. As noted in the introduction this process has been studied in great detail for these enzymes, in particular via extensive experimental and computational investigations on sLO-1.^{7,14,20,29,31,33-38} Thus, in this present study, we simply compare the current results obtained for 8R-LOX to those previously obtained for sLO-1. To the best of our knowledge the present investigation is the first to systematically consider the abstraction (PCET) process for several multiplicities of the fully-bound active site that includes the O₂ moiety.

The presently calculated barriers for the PCET process shown in **Figure 7.3** are strictly classical, i.e., no contribution from QM tunneling. For 8R-LOX the relative energies of ⁸TS1 and ⁴TS1 are close to each other at 15.4 and 16.9 kcal mol⁻¹ respectively, while ⁶TS1 lies decidedly higher in energy at 22.6 kcal mol⁻¹ (**Figure 7.3**). It should be noted that Glickman and Klinman³¹ have shown that the rate of H• abstraction is independent of the presence of O₂ and thus, the quartet and octet should lie equal in energy. It is well known that hybrid DFT methods tend to overestimate the stability of high spin systems.⁷⁵ Thus, the slight over-stabilization of ⁸TS is likely due to the greater number of parallel-spin electrons. However, the significantly higher relative energy of ⁶TS1 suggests that it is unlikely to be an important contributor to the PCET.

In comparison, previous computational investigations on sLO-1 have obtained a similar range of 12.1 – 20.8 kcal mol⁻¹ for the classical barrier of the PCET process.^{7,34,35,37} In addition, in sLO-1 the distance between the substrate carbon being oxidized (C11) and the Fe-OH oxygen, $r(\text{C11}\cdots\text{OFe})$, was 2.60-2.70 Å.^{7,34,35,37} In comparison in each current transition structure for the PCET process the key $r(\text{C10}\cdots\text{OFe})$ distance is 2.63 (⁴TS1 and ⁶TS1) and 2.64 (⁸TS1) Å (**Figure 7.6**). Thus, it appears that the PCET processes in 8R-LOX and sLO-1 are very similar in terms of classical barriers and geometries.

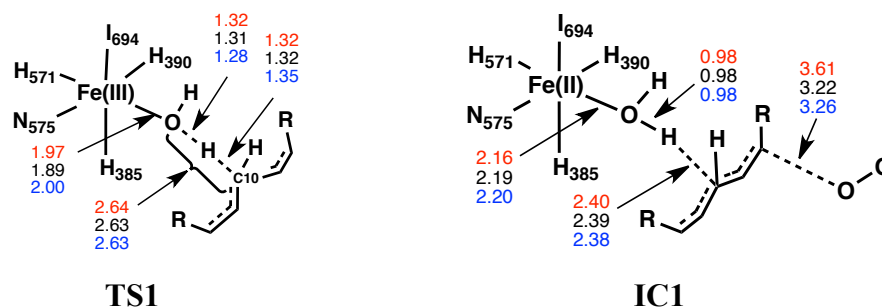


Figure 7.6. Schematic illustration of the optimized structures, obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory, of **TS1** and the subsequent radical-intermediate **IC1** for the PCET process with selected distances (Ångstrom) shown. Color code: blue (quartet), black (sextet), red (octet).

With regards to the possible impact of inclusion of QM tunneling effects it is noted that Kamerlin and Warshel⁷⁶ have stated that tunneling is influenced by the distance between the two reactants; forcing them together inhibits tunneling. Meanwhile, Knapp et al.⁷⁶ stated that the process of tunneling is very sensitive to the shape of the barrier. Thus, it can be expected that reaction barriers of similar height and involving similar distances between the reactants would likely result in similar contributions of tunneling to the PCET. Hence, given that LOXs are thought to exhibit commonalities of mechanism and reactivity, the present results suggest that the contribution of tunneling to the initial PCET step may be similar to all LOXs.^{7,14}

It is noted that of the previous theoretical investigations on LOXs some were performed in the absence of an active site environment while all involved a high-spin Fe complex in the absence of O₂.^{7,34,35,37} The present results for 8R-LOX thus suggest that the active sites of LOXs may not sterically force the substrate and Fe center closer than seen in the isolated reaction. Furthermore, the polar environment of the active site appears to not be significant for lowering the classical barrier, where tunneling is the major factor in reducing the barrier for PCET.

Upon PCET, the differences between multiplicities for the Fe–L bond lengths in the resulting pentadienyl radical intermediate complex (**IC1**) is reduced. This is expected given that the spin densities obtained for $^8\mathbf{IC1}$, $^6\mathbf{IC1}$ and $^4\mathbf{IC1}$ the absolute values for the Fe center, H₂O and pentadienyl radical intermediate are essentially the same. In particular, for all three multiplicities the iron now exists as a high-spin Fe(II) center (**Figure 7.4**). Interestingly, for $^6\mathbf{IC1}$ and $^4\mathbf{IC1}$ $r(\text{O}_2\cdots\text{C8})$ is ~ 0.4 Å shorter than in $^8\mathbf{IC1}$. This is likely caused by the favourable pairing of electrons between O₂ and the pentadienyl radical in $^4\mathbf{IC1}$ and $^6\mathbf{IC1}$. Regardless, however, $^8\mathbf{IC1}$ lies 3.2 kcal mol⁻¹ lower in energy than $^4\mathbf{IC1}$. This is again likely due to over-stabilization of the former with its larger number of parallel-spin electrons. Importantly, however, $^6\mathbf{IC1}$ now only lies 0.8 kcal mol⁻¹ higher in energy than $^4\mathbf{IC1}$. Thus, unlike $^6\mathbf{RC}$ and $^6\mathbf{TS1}$, the sextet may contribute to the ground state of **IC1**. It is noted that the difference between $^6\mathbf{IC1}$ and $^8\mathbf{IC1}$ is the direction of the unpaired electron's spin on the pentadienyl intermediate (**Figure 7.4**). Given that spin inversion is common in TM complexes it is conceivable that rather than generate $^6\mathbf{IC1}$ from $^6\mathbf{RC}$ via the high energy $^6\mathbf{TS1}$, it could instead be formed via spin inversion in $^8\mathbf{IC1}$.⁷⁷

Formation of the peroxy (–OO•) radical. Once the pentadienyl radical intermediate is formed, the subsequent step is attack at its C8 position by the O₂ moiety to give the corresponding peroxy intermediate. In agreement with experiment, at no time in our present investigations did the O₂ bind to the ferric or ferrous center prior to the formation of the peroxy radical. Moreover, the O₂ attacked the pentadienyl radical antarafacial to the leaving H•.^{11,31,72,73} Attempts to locate a TS for this process for any multiplicity considered within the current computational model was unsuccessful. Thus, the O₂ attack was examined via detailed relaxed PES scans.

On both the quartet and sextet PESs a relaxed scan along the O₂⋯C8 interaction coordinate gave a maximum in energy at $r(\text{O}_2\cdots\text{C8}) = 2.68$ Å. These transition structures, $^4\mathbf{TS2}$ and $^6\mathbf{TS2}$ respectively, are shown in **Figure 7.7** along with key selected bond

lengths. It is noted that in both TSs the O–O bond of the O₂ moiety itself has lengthened to 1.26 Å. Also, with respect to their corresponding fully-bound pentadienyl radical intermediate complexes (⁴IC1 and ⁶IC1) they give upper estimates to the energetic barriers for the attack of O₂ of 4.0 (⁴TS2) and 3.2 (⁶TS2) kcal mol⁻¹, respectively (**Figure 7.3**).

In contrast, in the case of the octet PES no maximum was found; the energy of the system simply increased with decreasing $r(\text{O}_2 \cdots \text{C}8)$ distances. In particular, the energy was found to increase above 40 kcal mol⁻¹, well beyond the limits of a typical enzymatic reaction.⁷⁸ This is due to the fact that no favourable pairing of electrons occurs between the reacting molecules (AA• and O₂) on the octet surface to enable C8–O₂ bond formation (see **Figure 7.4**). However, similar to that noted above, spin inversion could occur on the octet PES as O₂ approaches the carbon backbone. Specifically, a spin flip of the lone electron on the pentadienyl backbone would allow a hop from the octet to sextet PES, thus allowing for C8–O₂ bond formation.

As noted above, Borowski and Broclawik⁷ have previously computationally investigated this reaction step, attack of O₂, using DFT methods and a chemical model consisting of only the pentadienyl radical and O₂. That is, without the Fe center or active site residues being present. They obtained a barrier for this step of 2 kcal mol⁻¹ and with $r(\text{O}_2 \cdots \text{C})$ in the reaction TS of 2.35 Å.⁷ In their study the pentadienyl radical was planar and they observed that the barrier for attack increased when its planarity was perturbed.⁷ This suggests that the protein environment of 8R–LOX has a destabilizing effect on this process by distorting the geometry of the radical intermediate and thus, increasing the barrier by ~2 kcal mol⁻¹. Moreover, in the presently optimized structures of **TS2** the $r(\text{O}_2 \cdots \text{C})$ distance is ~0.30 Å longer than obtained by Borowski and Broclawik.⁷ It is noted that given the nature of the TS entropy is likely to have an important contribution. However, because the PESs in **Figure 7.3** represent changes in electronic energy this entropic effect is ignored. Such approximations may lead to highly inaccurate relative

density was ~ 1.0 with ~ 0.30 on the oxygen adjacent to the carbon backbone and ~ 0.70 on the distal C8–O₂ oxygen. Thus, the unpaired electron on AA–OO• is essentially wholly localized on the peroxy (–OO•) moiety (see **Figure 7.4**).

As noted in the introduction EPR evidence indicates that the attack of O₂ is fully reversible.^{28,31,45} From the PESs obtained (**Figure 7.3**) it can be seen that for both ⁴IC2 and ⁶IC2 the reverse barrier (dissociation of the peroxy radical to O₂ + pentadienyl radical) is only 10.2 and 10.1 kcal mol⁻¹, respectively. Thus, while IC2 is thermodynamically more favoured for both the quartet and sextet, the low barriers obtained for the forward and importantly reverse reactions support the conclusion that this reaction step is fully reversible.

Conformational change in the active site. With formation of the peroxy intermediate the proposed next step is oxidation of the Fe-center and thus regeneration of the catalytic center. Potentially, this may occur via the direct formation of an Fe-peroxy intermediate, referred to as the purple intermediate. This requires rotation of the peroxy intermediate from an antarafacial to suprafacial arrangement within the active site, i.e., so that the peroxy (–OO•) moiety is now directly adjacent or bound to the Fe(II) center. Alternatively, it could occur via a PCET in which a hydrogen is formally transferred from the Fe(II)–OH₂ moiety onto the –OO• moiety.^{31,46} However, regardless of the pathway, steric restrictions imposed by active site residues on the substrate may inhibit either process.^{46,47} Within the current QM/MM model the most direct pathway for a conformational change that would enable oxidation of the Fe center (via inner or outer shell electron transfer) was examined. More specifically, it involves counter-clockwise rotation about the C9–C8 bond (ω_1 ; \angle C10–C9–C8–C7) and C8–O bond (ω_2 ; \angle C9–C8–O–O•) as illustrated in **Figure 7.8**. It is noted that there are several possible pathways to reach the necessary conformer for oxidation of the Fe center, however, rotation of ω_1 allows for the change from antarafacial to suprafacial arrangement of the –O₂• moiety with respect to the iron center.

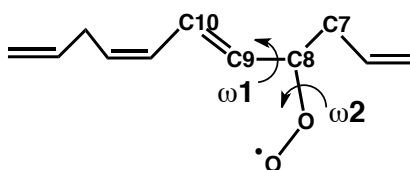


Figure 7.8. Schematic illustration of the dihedal angles ω_1 (\angle C10-C9-C8-C7) and ω_2 (\angle C9-C8-O-O *).

It is noted that for both ${}^4\text{IC2}$ and ${}^6\text{IC2}$, in which the peroxy group is antarafacial to the Fe center, ω_1 and ω_2 are approximately -64.0° and -13.0° , respectively. Rotation about ω_1 (with ω_2 free to rotate) the energy minimums ${}^4\text{IC3}$ and ${}^6\text{IC3}$ were obtained and are shown schematically in **Figure 7.3**. These complexes lie 8.8 and 7.2 kcal mol $^{-1}$ lower in energy than ${}^8\text{RC}$, or 2.8 and 1.3 kcal mol $^{-1}$ lower in energy than ${}^4\text{IC2}$ and ${}^6\text{IC2}$, respectively. While, a barrier likely exists (${}^4\text{TS3}$ and ${}^6\text{TS3}$) between IC2 and IC3 it was found in the scan no barriers for interconversion (within one decimal point) of either the quartet or sextet (**Figure 7.3**) were obtained at the level of theory used.

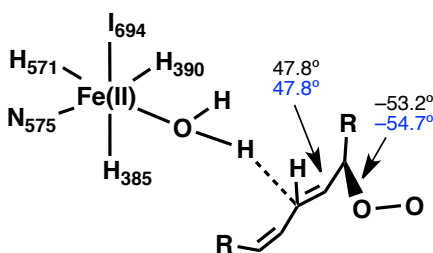


Figure 7.9. Optimized structures obtained at the ONIOM(B3LYP/6-31g(d):AMBER)-ME level of theory with key selected dihedral angles (degrees) for the quartet (blue) and sextet (black) intermediate IC3 .

The likely reason being that the intrinsic barrier is low where at the level of theory used and choice of 20° increments this barrier could not be found. Regardless, this suggests that within the enzyme proper, attack of O_2 at the C8 position of the pentadienyl

radical intermediate more or less leads directly to formation of **IC3**. Now, structurally, in ${}^4\mathbf{IC3}$ ω_1 and ω_2 are 47.8° and -55.0° while in ${}^6\mathbf{IC3}$ ω_1 and ω_2 are quite similar at 47.8° and -53.2° (**Figure 7.9**).

However, in both ${}^4\mathbf{IC3}$ and ${}^6\mathbf{IC3}$ the peroxy moiety is still antarafacial to the metal center. Thus, detailed further PES scans of rotation about ω_1 and ω_2 were performed to investigate pathways to formation of a suprafacial peroxy complex. Specifically, an energy maximum was obtained on the quartet and sextet surfaces when $\omega_1 = 115.5^\circ$ for both, and $\omega_2 = 19.9^\circ$ and 19.8° respectively. These resulting complexes designated as ${}^4\mathbf{TS4}$ and ${}^6\mathbf{TS4}$ (**Figure 7.10**) are 8.8 and 6.2 kcal mol $^{-1}$ higher in energy than ${}^4\mathbf{IC3}$ and ${}^6\mathbf{IC3}$, respectively (**Figure 7.3**). It should be noted that **TS4** should be taken as upper limits to rotation from an antarafacial to suprafacial arrangement of the peroxy moiety.

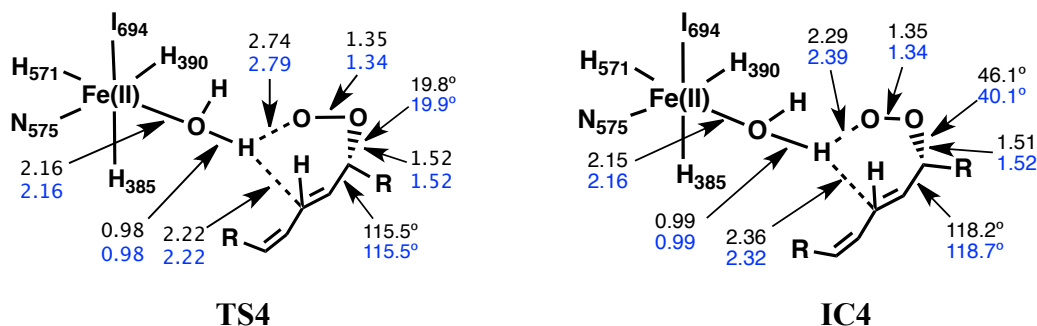


Figure 7.10. Optimized structures, obtained at the ONIOM(B3LYP/6-31G(d):AMBER)-ME level of theory, with key selected distances (Ångstroms) and dihedral angles (degrees) for **TS4** and intermediate **IC4**. Color code: Color code: blue (quartet), black (sextet).

In the suprafacial peroxy intermediate complexes formed, **IC4** (**Figure 7.10**) ω_1 and ω_2 are 118.7° and 40.1° for the quartet (${}^4\mathbf{IC4}$) and 118.2° and 46.1° respectively for the sextet (${}^6\mathbf{IC4}$). In addition, in both complexes the water bound to the Fe has re-oriented such that it now forms a hydrogen bond with the $-\text{OO}^\bullet$ moiety with $r(\text{OO}^\bullet \cdots \text{H}_2\text{O})$

distances of 2.39Å. In general, only quite minor changes are observed in key bond lengths compared to their corresponding "pre-rotated" peroxy complexes **IC2** (cf. **Figure 7.7**). It is also noted that the change in orientation of the peroxy moiety and formation of the H-bond had little effect on the measured spin densities for **IC4** which very similar to those seen in **IC2** multiplicities. Thus, as shown in **Figure 7.4**, similar electronic configurations exist. Formation of **IC4** is slightly endothermic with respect to **IC2**, with ${}^4\text{IC4}$ and ${}^6\text{IC4}$ modestly higher in energy than ${}^4\text{IC2}$ and ${}^6\text{IC2}$ by 2.8 and 3.3 kcal mol⁻¹, respectively.

In order to gain a greater understanding of the role and effect of the protein environment on the rotation of ω_1 and ω_2 , we have also considered the analogous process in aqueous solution. More specifically, the model shown in **Figure 7.8** and at the IEF-PCM-B3LYP/6-311G(2df,p)//B3LYP/6-31G(d) level of theory (water as solvent) was used. Then, ω_1 was scanned over a full 360° rotation in increments of 20° with ω_2 fixed. For each of these values of ω_1 , ω_2 was scanned over a full 360° in increments of 20°. All other degrees of freedom were not fixed. The resulting 3D PES is shown in **Figure 7.11**. As can be seen, overall six minima are found on the PES. The complex **A**, with ω_1 and ω_2 values of -64.0° and -13.0° respectively, is similar to complex ${}^4\text{IC2}$ obtained using the above QM/MM approach which had essentially the same values for ω_1 and ω_2 (see above). Thus, all relative energies in **Figure 7.11** are determined using **A** as the energy zero-point.

However, the minima labeled **B** and **C** in **Figure 7.11** are thermodynamically slightly favoured being -1.7 and -1.5 kcal mol⁻¹ lower in energy than **A**, respectively. Structure **C** is of interest as it is the aqueous solution minimum with angles most similar to that obtained for both ${}^4\text{IC4}$ and ${}^6\text{IC4}$ with ω_1 and ω_2 values of 127.4° and 76.4° respectively (${}^4\text{IC4}$: $\omega_1 = 118.2^\circ$, $\omega_2 = 40.1^\circ$; ${}^6\text{IC4}$: $\omega_1 = 118.2^\circ$, $\omega_2 = 46.1^\circ$). Meanwhile **B** ($\omega_1 = 127.4^\circ$, and $\omega_2 = -63.6^\circ$) does not correspond to an energy minimum on the QM/MM PES.

The most direct pathway connecting **C** and **A** involves initial rotation of ω_1 to first form **B**, followed by rotation of ω_2 to give **C**. It is noted that this is similar to the pathway examined in the above QM/MM investigation. As can be seen in **Figure 7.11**, the largest barrier encountered along this sequential rotation route is only $2.5 \text{ kcal mol}^{-1}$.

Thus, the steric interactions and polarizing effects provided by the active site do have a modest affect upon the rotation of the peroxy intermediate. In particular, **IC4** (analogous to **C**, **Figure 7.11**) is no longer the thermodynamically favoured conformer. Instead, it lies 3.3 (quartet) and 2.8 (sextet) kcal mol^{-1} higher in energy than **IC2** (active site complex analogous to the aqueous structure **I**). Meanwhile the complex corresponding to the angles seen in **B** is no longer a minimum. Furthermore, within the active site (**Figure 7.3**) the highest energetic cost for any rotational barrier is $8.8 \text{ kcal mol}^{-1}$. This represents an approximate difference of only 6 kcal mol^{-1} in comparison to the energies obtained in aqueous solution (**Figure 7.11**). Thus, while the active site has perturbed the rotational profile for the peroxy intermediate it has not made its rotation unfeasible.

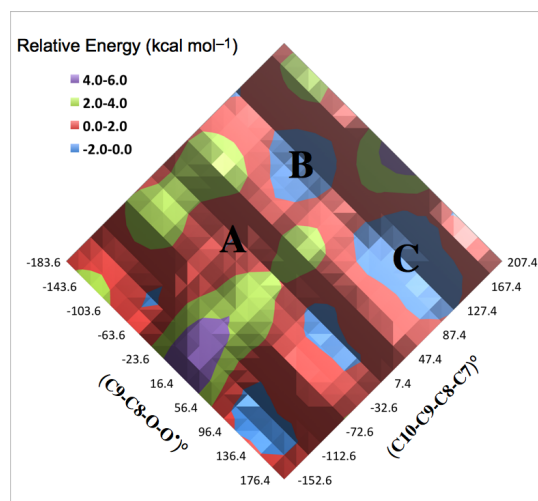


Figure 7.11. The 3D PES for the rotation of the peroxy group to an angle required for H• abstraction. The energies were obtained at the IEF-PCM-B3LYP/6-311G(2df,p)//B3LYP/6-31G(d) level of theory.

Regeneration of the Fe(III) center. The next and final mechanistic step after formation of the suprafacial peroxy intermediates ${}^4\text{IC4}$ and ${}^6\text{IC4}$ is oxidation of the Fe center and formation of the final peroxide (C8–O₂H) product.

Previously, Borowski and Broclawik⁷ investigated two mechanistic possibilities using modest sized QM models: direct oxidation of the Fe-center or via a PCET. The former proceeds via the formation of a seven-coordinate complex (the "purple intermediate"). Within their models, while this process was exothermic by $\sim 6 \text{ kcal mol}^{-1}$, they were unable to obtain the reaction barrier. Furthermore, the subsequent heterolytic Fe \cdots OOH cleavage was found to occur in concert with transfer of a proton from the Fe(II)–OH₂ moiety to the C8–OO \cdot group (i.e., giving the final desired products) with a barrier of $10.1 \text{ kcal mol}^{-1}$. For the alternative pathway via PCET, without complexation of the peroxy group to the Fe(II) center, a transition structure and hence barrier leading to the desired product complex could not be obtained.⁷

Thus we initially considered the PCET pathway within the present larger computational model. For quartet multiplicity (i.e., ${}^4\text{IC4}$) the process occurs via ${}^4\text{TS5}$ with a barrier of $13.0 \text{ kcal mol}^{-1}$ relative to ${}^4\text{IC4}$ ($10.3 \text{ kcal mol}^{-1}$ relative to ${}^8\text{RC}$). In contrast, for sextet multiplicity (i.e., ${}^6\text{IC4}$) the PCET process occurs via ${}^6\text{TS5}$ with a markedly lower barrier of $2.7 \text{ kcal mol}^{-1}$ relative to ${}^6\text{IC4}$ ($-0.4 \text{ kcal mol}^{-1}$ relative to ${}^8\text{RC}$)! It is noted that ${}^4\text{IC4}$ and ${}^6\text{IC4}$ have very similar relative energies, differing by just $0.4 \text{ kcal mol}^{-1}$, with the sextet lying slightly lower in energy (see **Figure 7.3**). Spin inversion within ${}^4\text{IC4}$ could give rise to ${}^6\text{IC4}$ and thus enable the higher barrier for PCET along the quartet surface to be avoided.

We have also examined the alternative mechanism for formation of the final product complexes via a "purple intermediate"; that is, via a Fe(II) \cdots OO–C8 peroxy-crosslinked intermediate. Specifically, detailed scans of the PES were conducted starting from ${}^6\text{IC4}$ (**Figure 7.10**) in which the Fe \cdots O distance was systematically shortened in increments of 0.1 \AA . However, such a shortening was observed to result in a rapid increase in relative

energy for both multiplicities: at a $\text{Fe}\cdots\text{OO}^-$ distance of 3.1 Å the relative energy had already increased by $\sim 14 \text{ kcal mol}^{-1}$! That is, no peroxy crosslink has been formed and the energy is already above the highest barrier of $13.0 \text{ kcal mol}^{-1}$ obtained above for the alternative PCET process. Thus, in agreement with previous experimental investigations^{46,47} it appears unlikely that the "purple intermediate" is catalytically relevant. Comparison with the results from the previous gas-phase computational study of Borowski and Broclawik⁷ the current investigation suggests that within the active site formation of a "purple intermediate" is disfavoured.

The resulting product complexes ^4PC and ^6PC are 3.9 and $12.2 \text{ kcal mol}^{-1}$ respectively lower in energy than ^8RC (Figure 7.3). Interestingly, from Figure 7.4 it can be seen that upon oxidation of the Fe center ^4PC exists in an intermediate spin state, while for ^6PC the Fe center *now* lies in a high-spin state. Magnetic susceptibility studies have suggested that throughout the reaction the iron center exists in a high-spin state.¹¹ Thus, in agreement with the present results ^4PC appears unlikely to be a relevant product complex.

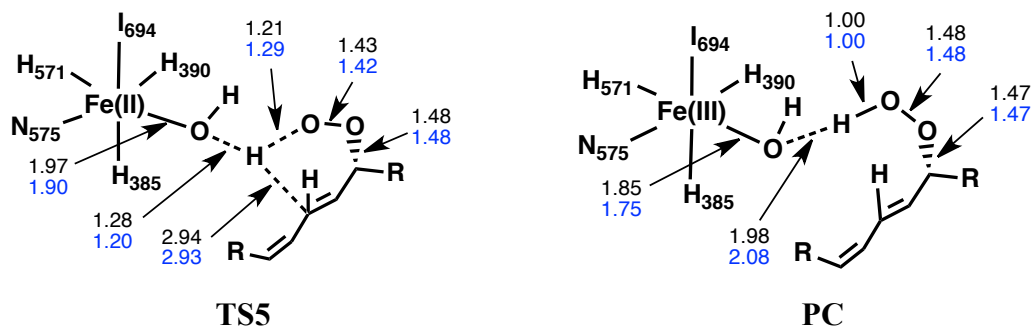


Figure 7.12. Optimized structures, obtained at the ONIOM(B3LYP/6-31g(d):AMBER)-ME level of theory, with key selected distances for the transition structure (TS5) of the reduction of the Fe-OH₂ center and the subsequent product (PC). Color code: blue (quartet), black (sextet).

7.4 Conclusions

A detailed systematic computational investigation has been performed on the catalytic mechanism of 8R-LOX using an ONIOM QM/MM-based approach.

The results obtained for the initial PCET of 8R-LOX were compared to results obtained for sLO-1. From the present QM/MM results of 8R-LOX the classical barriers for the initial PCET show good agreement with those obtained for sLO-1. Moreover, the distances between the Fe-center (being reduced) and carbon-center (being oxidized) were very similar. Potentially, this suggests that the contribution of tunneling for the initial PCET may be similar among all LOXs.

For the AA[•] complex **IC1**, the quartet, sextet and octet were found to all lie close in energy (i.e. within 4 kcal mol⁻¹ of each other). Indeed, while the octet is energetically preferred the subsequent required C8-OO bond formation cannot occur due to unfavourable electron spin pairing. Instead, spin inversion to the quartet or even sextet state, which lie within 0.8 kcal mol⁻¹ of each other, must occur. As a result, the barrier for C8-OO bond formation proceeds via a barrier of 4.2 kcal mol⁻¹ with respect to ⁸RC for either multiplicity.

The peroxy -OO[•] moiety in ⁴IC2 and ⁶IC2 is antarafacial to the Fe center. The rotation of the substrate about both its C9-C8 and C8-OO bonds enables the required conformational change to occur that brings the -OO[•] moiety suprafacial to the Fe centre. Importantly, throughout this process the quartet and sextet state stationary points (i.e., minima and transition structures) are calculated to lie within 1.6 kcal mol⁻¹ of each other with the highest rotational step barrier (⁴TS4) being only 8.8 kcal mol⁻¹. For the current model, contrary to previous suggestions that steric interactions between substrate and active site might inhibit rotation, the impact of the active site was found to not be that significant. However, the active site did inhibit the formation of the seven-coordinate intermediate (see below). In addition, the presently determined energetics for formation

of the initial peroxy intermediate **IC2** and its subsequent conformational change within the active site suggest that 8R-LOX may utilize multistate reactivity.

The next and final step in the overall mechanism is formation of the peroxide product AA-OOH. This process is determined to preferentially occur via a PCET process in which the Fe(II) is oxidized and thus regenerates a Fe(III) center while concomitantly a proton of the Fe-bound H₂O transfers onto the peroxy moiety of AA-OO• to give the desired AA-OOH product. On the *now* quartet PES this step occurs with a barrier of 13.0 kcal mol⁻¹. Importantly, however, on the *now* sextet PES the barrier is significantly lower at just 2.3 kcal mol⁻¹. Furthermore, in agreement with experimental observations the resulting product ⁶PC is preferred, being markedly lower in energy than either ⁴PC or any of the initial fully bound active site complexes. The alternate pathway for product formation via a seven-coordinate Fe-complex, the "purple intermediate", was calculated to be at least energetically uncompetitive with the above PCET process.

7.5 References

- (1) Minor, W.; Steczko, J.; Bolin, J. T.; Otwinowski, Z.; Axelrod, B. *Biochemistry* **1993**, *32*, 6320-6323.
- (2) Andreou, A.; Brodhun, F.; Feussner, I. *Prog. Lipid Res.* **2009**, *48*, 148-170.
- (3) Kuhn, H. *Prostaglandins Other Lipid Mediat.* **2000**, *62*, 255-270.
- (4) Kuhn, H.; Walther, M.; Kuban, R. J. *Prostaglandins Other Lipid Mediat.* **2002**, *68-9*, 263-290.
- (5) Toledo, L.; Masgrau, L.; Marechal, J. D.; Lluch, J. M.; Gonzalez-Lafont, A. *J. Phys. Chem. B* **2010**, *114*, 7037-7046.
- (6) Neau, D. B.; Gilbert, N. C.; Bartlett, S. G.; Boeglin, W.; Brash, A. R.; Newcomer, M. E. *Biochemistry* **2009**, *48*, 7906-7915.
- (7) Borowski, T.; Broclawik, E. *J. Phys. Chem. B* **2003**, *107*, 4639-4646.

- (8) Zhang, Y.; Gan, Q. F.; Pavel, E. G.; Sigal, E.; Solomon, E. I. *J. Am. Chem. Soc.* **1995**, *117*, 7422-7427.
- (9) Ivanov, I.; Heydeck, D.; Hofheinz, K.; Roffeis, J.; O'Donnell, V. B.; Kuhn, H.; Walther, M. *Arch. Biochem. Biophys.* **2010**, *503*, 161-174.
- (10) Clapp, C. H.; Strulson, M.; Rodriguez, P. C.; Lo, R.; Novak, M. J. *Biochemistry* **2006**, *45*, 15884-15892.
- (11) Gaffney, B. J. *Annu. Rev. Biophys. Biomolec. Struct.* **1996**, *25*, 431-459.
- (12) Kuhn, H.; O'Donnell, V. B. *Prog. Lipid Res.* **2006**, *45*, 334-356.
- (13) Andreou, A.; Feussner, I. *Phytochem. Rev.* **2009**, *70*, 1504-1510.
- (14) Schenk, G.; Neidig, M. L.; Zhou, J.; Holman, T. R.; Solomon, E. I. *Biochemistry* **2003**, *42*, 7294-7302.
- (15) Skrzypczak-Jankun, E.; Bross, R. A.; Carroll, R. T.; Dunham, W. R.; Funk, M. O. *J. Am. Chem. Soc.* **2001**, *123*, 10814-10820.
- (16) Samuelsson, B.; Ramwell, P.; Paletti, R.; Folco, G.; Granstrom, E.; et-al; *Advances in Prostaglandin, Thromboxane, and Leukotriene Research*; Raven: New York, 1995; Vol. 23.
- (17) Degroot, J. J. M.; Veldink, G. A.; Vliegthart, J. F. G.; Boldingh, J.; Wever, R.; Vangelder, B. F. *Biochim. Biophys. Acta* **1975**, *377*, 71-79.
- (18) Schilstra, M. J.; Veldink, G. A.; Vliegthart, J. F. G. *Biochemistry* **1994**, *33*, 3974-3979.
- (19) Veldink, G. A.; Vliegthart, J. F. *Adv. Inorg. Biochem.* **1984**, *6*, 139-161.
- (20) Glickman, M. H.; Klinman, J. P. *Biochemistry* **1995**, *34*, 14077-14092.
- (21) Dunham, W. R.; Carroll, R. T.; Thompson, J. F.; Sands, R. H.; Funk, M. O. *Eur. J. Biochem.* **1990**, *190*, 611-617.
- (22) Nelson, M. J. *J. Am. Chem. Soc.* **1988**, *110*, 2985-2986.
- (23) Zhang, Y.; Gebhard, M. S.; Solomon, E. I. *J. Am. Chem. Soc.* **1991**, *113*, 5162-5175.

- (24) Cheesbrough, T. M.; Axelrod, B. *Biochemistry* **1983**, *22*, 3837-3840.
- (25) Petersson, L.; Slappendel, S.; Vliegthart, J. F. G. *Biochim. Biophys. Acta* **1985**, *828*, 81-85.
- (26) Vanderheijdt, L. M.; Feiters, M. C.; Navaratnam, S.; Nolting, H. F.; Hermes, C.; Veldink, G. A.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1992**, *207*, 793-802.
- (27) Gardner, H. W. *Biochim. Biophys. Acta* **1991**, *1084*, 221-239.
- (28) Nelson, M. J.; Cowling, R. A.; Seitz, S. P. *Biochemistry* **1994**, *33*, 4966-4973.
- (29) Rickert, K. W.; Klinman, J. P. *Biochemistry* **1999**, *38*, 12218-12228.
- (30) Siedow, J. N. *Annu. Rev. Plant Physiol. Plant Molec. Biol.* **1991**, *42*, 145-188.
- (31) Glickman, M. H.; Klinman, J. P. *Biochemistry* **1996**, *35*, 12882-12892.
- (32) Knapp, M. J.; Rickert, K.; Klinman, J. P. *J. Am. Chem. Soc.* **2002**, *124*, 3865-3874.
- (33) Iyengar, S. S.; Sumner, I.; Jakowski, J. *J. Phys. Chem. B* **2008**, *112*, 7601-7613.
- (34) Lehnert, N.; Solomon, E. I. *J. Biol. Inorg. Chem.* **2003**, *8*, 294-305.
- (35) Olsson, M. H. M.; Siegbahn, P. E. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126*, 2820-2828.
- (36) Sumner, I., Iyengar, S.S., *J. Chem. Theory Comput.* **2010**, *6*, 1698-1710.
- (37) Tejero, I.; Eriksson, L. A.; Gonzalez-Lafont, A.; Marquet, J.; Lluch, J. M. *J. Phys. Chem. B* **2004**, *108*, 13831-13838.
- (38) Hatcher, E.; Soudackov, A. V.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2004**, *126*, 5763-5775.
- (39) Hammes-Schiffer, S. *Acc. Chem. Res.* **2001**, *34*, 273-281.
- (40) Soudackov, A.; Hammes-Schiffer, S. *J. Chem. Phys.* **1999**, *111*, 4672-4687.
- (41) Soudackov, A.; Hammes-Schiffer, S. *J. Chem. Phys.* **2000**, *113*, 2385-2396.
- (42) Hwang, J. K.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 10053-10058.
- (43) Hwang, J. K.; Warshel, A. *J. Am. Chem. Soc.* **1996**, *118*, 11745-11751.

- (44) Scarrow, R. C.; Trimitsis, M. G.; Buck, C. P.; Grove, G. N.; Cowling, R. A.; Nelson, M. J. *Biochemistry* **1994**, *33*, 15023-15035.
- (45) Nelson, M. J.; Seitz, S. P.; Cowling, R. A. *Biochemistry* **1990**, *29*, 6897-6903.
- (46) Nelson, M. J.; Chase, D. B.; Seitz, S. P. *Biochemistry* **1995**, *34*, 6159-6163.
- (47) Knapp, M. J.; Seebeck, F. P.; Klinman, J. P. *J. Am. Chem. Soc.* **2001**, *123*, 2931-2932.
- (48) Molecular Operating Environment; 2010.10 ed.; Chemical Computing Group Inc.: Montreal, Quebec, Canada, 2010.
- (49) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049-1074.
- (50) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. *J. Comput. Phys.* **1999**, *151*, 114-134.
- (51) Bearpark, M. J.; Ogliaro, F.; Vreven, T.; Boggio-Pasqua, M.; Frisch, M. J.; Larkin, S. M.; Robb, M. A. In *Computation in Modern Science and Engineering Vol 2, Pts a and B*; Simos, T. E., Maroulis, G., Eds.; Amer Inst Physics: Melville, 2007; Vol. 2, p 583-585.
- (52) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Theochem-J. Mol. Struct.* **1999**, *461*, 1-21.
- (53) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959-1967.
- (54) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179.
- (55) Morokuma, K.; Musaev, D. G.; Vreven, T.; Basch, H.; Torrent, M.; Khoroshun, D. V. *IBM J. Res. Dev.* **2001**, *45*, 367-395.
- (56) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
- (57) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.
- (58) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432.
- (59) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760-769.

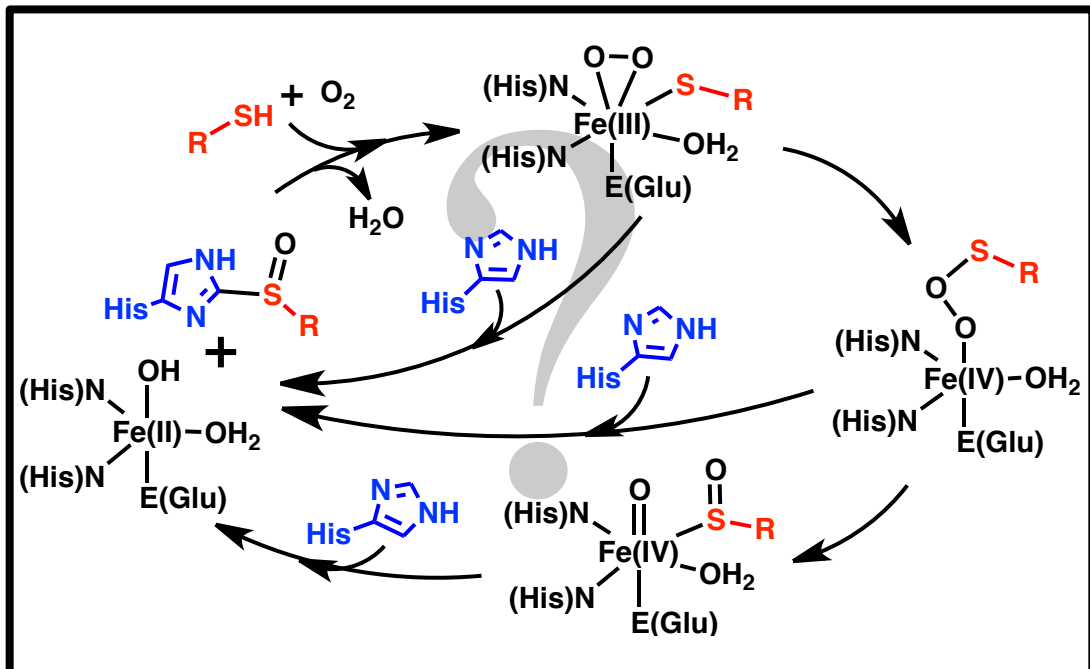
(60) Frisch, M. J., Trucks, G.W., Schlegel,H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A.Jr., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A.; Revision D.02 ed. Wallingford CT, 2004.

(61) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Revision B.01 ed. Wallingford CT, 2010.

- (62) Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2001**, *107*, 48-55.
- (63) Bushnell, E. A. C.; Gault, J. W. *J. Comput. Chem.* **2013**, *34*, 141-148.
- (64) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- (65) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (66) Lee, C., Yang, W., Parr, R.G. *Phys. Rev. B: Condens. Matter* **1988**, *37*, 785-789.
- (67) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
- (68) Feller, D. *J. Comp. Chem.* **1996**, *17*, 1571-1586.
- (69) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model* **2007**, *47*, 1045-1052.
- (70) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232-240.
- (71) Llano, J., Gault, J. W. In *Quantum Biochemistry: Electronic Structure and Biological Activity*; Matta, C. F., Ed.; Wiley-VCH: Weinheim, 2010; Vol. 2, p 920.
- (72) Wu, F.; Gaffney, B. J. *Biochemistry* **2006**, *45*, 12510-12518.
- (73) Chamulitrat, W.; Mason, R. P. *J. Biol. Chem.* **1989**, *264*, 20968-20973.
- (74) Coffa, G.; Brash, A. R. *Proc. Nat. Acad. Sci. U.S.A.* **2004**, *101*, 15579-15584.
- (75) Hughes, T. F.; Friesner, R. A. *J. Chem. Theory Comput.* **2011**, *7*, 19-32.
- (76) Kamerlin, S. C. L.; Warshel, A. *J. Phys. Org. Chem.* **2010**, *23*, 677-684.
- (77) Poli, R.; Harvey, J. N. *Chem. Soc. Rev.* **2003**, *32*, 1-8.
- (78) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729-738.

Chapter 8

Model Iron-Oxo Species and the Oxidation of Imidazole: Insights into the Mechanism of OvoA and EgtB?



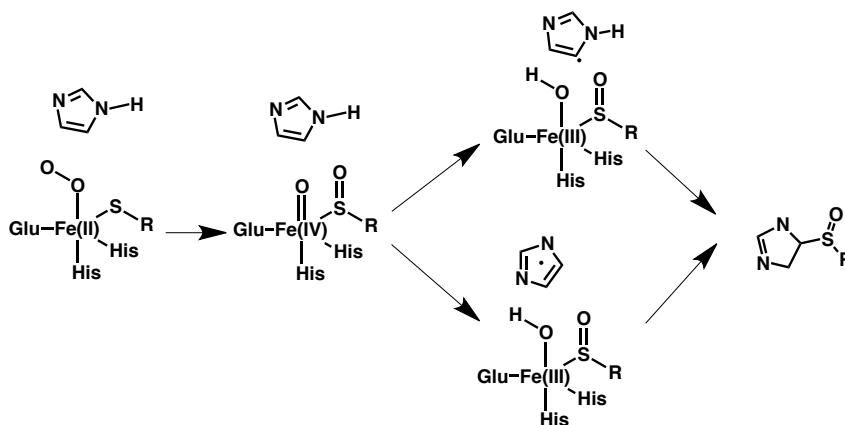
8.1 Introduction

Activation of O₂ is well-established as a key physiological approach to, for instance, activation of metabolites and incorporation of oxygen into biomolecules.^{1,2} Within cells this process is typically catalyzed by enzymes that depend on metal cofactors, e.g., copper and iron.^{3,4} Commonly formed intermediates in O₂ activation by iron-containing metalloenzymes include the high-valent oxo-ferryl Fe(IV)=O containing species' Compounds I and II (Cpd I and II, respectively) in heme-enzymes,⁵ and Cpd II analogs in non-heme enzymes.^{4,12} These moieties are generally considered to be the 'ultimate biochemical oxidants'^{5,6,12-14} with their reactivity tuned by such factors as redox potential and spin state of the metal-oxo moiety.¹⁵ Indeed, as a result they are able to transform relatively stable bonds, e.g., C–H, via oxygen insertion.¹⁵ Due to their preference for a high-spin ground state (GS) which thus allows for exchange enhanced reactivity (EER), it has been stated that non-heme Fe(IV)=O species exhibit a greater reactivity than their heme analogues which prefer a low-spin GS.¹⁶

For instance, 5-histidylcysteine sulfoxide synthase (*OvoA*) and 2-histidyl- γ -glutamyl cysteine sulfoxide synthase (*EgtB*) are two non-heme iron-containing enzymes that activate O₂ as part of their catalytic mechanism.¹⁷ More specifically, they use O₂, histidine and a cysteine-derivative as cosubstrates to synthesize their respective sulfoxides.^{17,18} The active sites of both enzymes utilize a conserved Fe binding motif, that is also common among O₂ activating enzymes,¹⁹ involving two histidyl imidazoles (Im's) and a carboxylate side chain in a facial ligation.^{17,20} Such an arrangement leaves several possible sites for Fe-substrate binding. Due to these similarities in reactants, product and active site, it has been proposed that *OvoA* and *EgtB* share similar chemistry.²⁰ However, their exact catalytic mechanisms are unknown.

OvoA is found in marine organisms^{21,22} such as sea urchins, scallops, starfish and the annelid *Platynereis dumerilii*,^{20,23} as well as human pathogenic parasites of the *Trypano-*

soma genus.^{20,24} In the case of EgtB it is found in non-yeast fungi, mycobacteria and cyanobacteria.²⁵ Importantly, the products of OvoA and EgtB (hereafter referred to as 5- and 2-HisCysSO, respectively) are later converted to OSH and ESH (both of which are mercaptohistidine derivatives) to provide essential protection against oxidative damage. In particular, ESH has been shown to scavenge reactive oxygen species and radicals such as singlet oxygens, hydroxyl radicals, hypochlorous acid, and peroxy radicals.²⁶⁻³¹ Importantly ESH's medical potential is very promising where aerosols have been developed to treat chronic inflammatory diseases such as asthma.³²⁻³⁴ Furthermore, it has been stated that ESH is an important chemoprotector present in humans.³⁵ It is noted that OSH, while an important antioxidant, has been suggested to have additional physiological roles.²⁰ For example, it has been proposed to act as a male pheromone in the *P. dumerilii*.^{20,23} Consequently, because of its wide presence in many organisms the antioxidant and scavenging abilities of OSH has been studied extensively.^{17,21,24,36-43}



Scheme 8.1. Proposed mechanism for formation of a histidyl-sulfoxide via a radical mechanism with coupling between the sulfoxide and histidyl occurring at the latter's C δ position (i.e. synthesis of 5-HisCysSO).¹⁷

The proposed mechanism of OvoA (**Scheme 8.1**), i.e., synthesis of 5-HisCysSO, begins with O₂ and cysteine binding to the ferrous (Fe(II)) center to form an Fe(III)–O₂^{•−}

containing complex.¹⁷ The latter species then attacks the Fe-bound cysteine, homolytically cleaving its O–O bond with formation of a sulfoxide···Fe(IV)=O complex. Three possible pathways have been proposed for the next step.¹⁷ In one, the histidine nucleophilically attacks at the sulfoxide sulfur center with concomitant deprotonation of N δ to directly give the product (not shown).¹⁷ In the alternate two pathways, however, the histidine is oxidized via a proton-coupled electron (PCET) onto the Fe(IV)=O moiety. Specifically, a H \cdot is abstracted from either its C δ -H or N δ -H groups to form an sp² C-centered or a π -delocalized radical, respectively (**Scheme 8.1**).¹⁷ The resulting histidyl radical is suggested to then attack the Fe-bound sulfoxide to give the final product. However, a number of key central questions remain including the nature of the oxidizing non-heme Fe species and its coordination environment; the most likely resulting histidyl radical resulting from oxidation; and the apparent need for formation of a sulfoxide-containing intermediate when it is not present in the final product (i.e. OSH or ESH).

Thus, using a DFT-cluster approach in combination with a first principles quantum and statistical mechanics⁴⁴ approach we have computationally investigated the half-reactions for the oxidation of Im and the reduction of several possible iron-oxygen complexes via an electron (ET) or proton-coupled-electron (PCET) transfer. It is noted that because there are currently no available X-ray crystal structures for *OvoA* or *EgtB* the use of small model Fe-complexes herein does not provide a conclusive answer to the mechanisms of *OvoA* or *EgtB*. In particular, it does not explicitly account for the environmental effects provided by the secondary shell of active site residues within the respective enzymes. However, such a model approach can provide fundamental insights into the changes in oxidative power of the Fe center as the coordination around the center is changed. Specifically, we investigated twenty-one possible Fe-cluster models (i.e. the oxidized and reduced forms). For each of these we considered several possible multiplicities to give in total one-hundred and ninety-six different Fe-complexes.

8.2 Computational Methods

As noted above there are currently no available X-ray crystal structures for OvoA or EgtB, thus, we have chosen to investigate seven possible models differing in their Fe-coordination arrangements (**Figure 8.1**). In all complexes the ligating glutamate was modeled as formic acid. The ligating histidines were modeled as imidazoles. In all models these residues ligate the Fe-center in a facial arrangement. For A₁-A₅ complexes methylthiol was used to model the ligating cysteine. In the case of A₁ and A₂ we have modeled a five-coordinate Fe center. For these complexes the thiol was either *cis* or *trans* to the formate ligand, respectively. For A₃-A₅ we have modeled a six-coordinate Fe center where water was added to fill the sixth coordination site of the metal. For these complexes we have generated initial complexes such that the MeS⁻, H₂O or O₂ were *trans* to the carboxylate ligand, respectively. In the case of A₆ and A₇ 2-amine-methylthiol was used to model a bidentately ligating cysteine. The difference between these two was whether the thiol was *trans* or *cis* to the carboxylate ligand. It is noted that for EgtB which uses γ -glutamyl-cysteine as a substrate the coordination modes for A₆ and A₇ are unlikely.

For the calculation of the half-reaction free energies a DFT-cluster model was used in combination with a first principles quantum and statistical mechanics approach.⁴⁴ For the reactions considered herein the protons and electrons were treated as independent ions. Thus, their chemical potentials have been taken to be that of a solvated free electron with respect to a SHE reference state ($-418.5 \text{ kJ mol}^{-1}$) and for a proton in a dilute aqueous environment ($-1124.2 \text{ kJ mol}^{-1}$) as previously obtained by means of a first principles quantum and statistical mechanics approach.⁴⁴

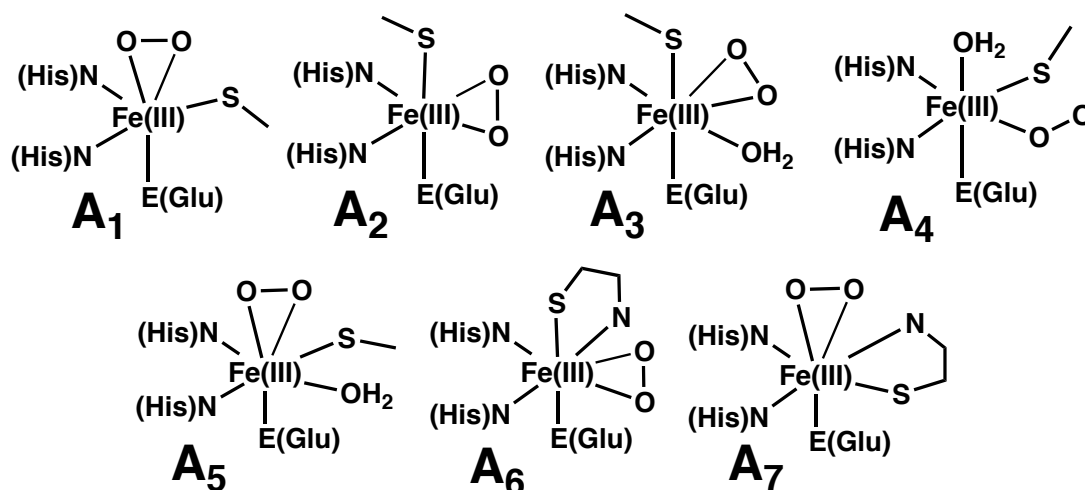


Figure 8.1. The initial five- and six-coordinate Fe(III)-O₂⁻ complexes considered herein.

For the remaining species (i.e. imidazole and Fe-complexes) the Gaussian 09⁴⁵ suite of software was used. In the present investigation we have chosen to use the 6-31G(d) basis set on all atoms including Fe (i.e., an effective core potential basis set for Fe was not used). As a test we ran the septet **A**-complexes in **Figure 8.1** (the reason being that they were the favoured starting complexes; see below) and found that the RMSD in Fe-O₂ and Fe-S bond lengths to be 0.04 Å when the SDD effective core potential basis set for Fe while the 6-31G(d) basis set was used for all other atoms was used. This combination of basis sets has been shown to be reliable in both mono- and bi-nuclear Fe-containing enzymes.^{46,47} Furthermore, the bonding of the O₂ (i.e. whether end-on or side-on was preferred) was identical between basis sets used. Thus, it is believed that the differences in using the 6-31G(d) basis set solely would not significantly affect the key results obtained herein. Thus, the optimized structures and Gibbs corrections (ΔG_{corr}) were obtained at the B3LYP/6-31G(d) level of theory (See Appendix **Table D1**).⁴⁸⁻⁵³ Relative energies were obtained via single-point calculations at the IEF-PCM-B3LYP/6-311G(2df,p)//B3LYP/6-31G(d) + ΔG_{corr} level of theory.⁵⁴⁻⁵⁷ In addition single point energies were also calculated using the IEF-PCM-M06/6-311G(2df,p)//B3LYP/6-31G(d) + ΔG_{corr} method.^{58,59} However, the results obtained were qualitatively similar to those obtained with the B3LYP func-

tional and thus, while provided in the following tables (in parentheses) are not discussed. Diffuse functions were not used because, as discussed by Martin et al.,⁶⁰ their inclusion on Fe is a poor match when used with the underlying 6–31G or 6–311G basis sets. Water was chosen as the solvent because the electron and proton reference energies were defined in an aqueous environment. We did perform additional calculation with a dielectric constant to better model an active site environment. However, while the absolute energies changed the key results obtained remained consistent. Using these calculated free energies as well as the chemical potentials of a solvated free electron (with respect to a SHE reference state) and proton (in a dilute aqueous environment) the half-reaction free energies were obtained as per the approach outlined by Llano and Eriksson.⁴⁴

8.3 Results and Discussion

In order to evaluate the mechanistic feasibility of the various possible iron-oxygen oxidants we began by first examining the inherent free energy cost of oxidizing the R-group Im of histidine (modeled as imidazole) via either an ET or PCET process. In the SHE reference state the loss of an electron from Im to give the radical cation $\text{Im}^{\bullet+}$ is endothermic by $186.0 \text{ kJ mol}^{-1}$ (**Table 8.1**). However, the coupling of ET with proton loss from either the $\text{C}_\delta\text{-H}$ or $\text{C}_\epsilon\text{-H}$ moieties of the Im, i.e. PCET to give a deprotonated neutral Im-derived radical ($\text{Im}(-\text{H})^\bullet$), is markedly even more endothermic with a free energy cost of 250.1 and $251.5 \text{ kJ mol}^{-1}$, respectively.

As previously noted,¹⁷ in the resulting radical species' (i.e. $\text{Im-C}_\delta(-\text{H})^\bullet$ and $\text{Im-C}_\epsilon(-\text{H})^\bullet$) the unpaired electron is localized on the respective carbon as an sp^2 -radical.¹⁷ However, while the alternate PCET process involving proton loss from the Im's $\text{N}_\delta\text{-H}$ group is still endothermic, the free energy cost is significantly lower at $171.0 \text{ kJ mol}^{-1}$. In fact, it is now less than that of the ET process alone (see **Table 8.1**)! Again, as noted previously,¹⁷ in the resulting $\text{Im-N}_\delta(-\text{H})^\bullet$ radical the unpaired electron is delocalized over the π -

system of the Im itself.¹⁷ As noted above, the formation of 2- and 5-HisCysSO is believed to follow similar chemistries. Thus, it seems that of the above three possible processes, formation of Im-N δ (-H) \cdot is thermodynamically most favoured.

Table 8.1. Adiabatic free energies (kJ mol⁻¹) for oxidation of the Im via ET and PCET.

ET	Species Formed	PCET	Species Formed
		171.0 (163.3)	Im-N δ (-H) \cdot
186.0(189.5)	Im \cdot ⁺	250.1 (238.6)	Im-C δ (-H) \cdot
		251.5 (241.2)	Im-C ϵ (-H) \cdot

While Fe(IV)=O is generally considered the stronger oxidant, we first investigated the free energies of reducing the various possible ferrous-O₂ complexes (**A**) shown in **Figure 8.1** via ET or PCET processes. For each complex we calculated the energies and geometries of the singlet (bi-radical), triplet, quintet and septet multiplicities. Notably, each **A** complex was found to prefer a septet GS with the iron-center in its ferric state (Fe(III)) with the bound O₂ moiety better represented as a superoxide radical, O₂ \cdot ⁻. Furthermore, in each complex (except for **A**₄) the O₂ was bound side-on. These results agree with previous computational investigations.^{4,61} In particular, Chen et al.⁴ found that for [(TMC)O₂Fe(II)]²⁺ the septet GS was preferred with O₂ bound side-on. In addition, Chung et al.⁶¹ found that in several non-heme complexes the septet GS was preferred. However, the side- or end-on binding of O₂ depended on the steric-crowding about the Fe.⁶¹ For the reduction of the **A**-complexes via ET we calculated the energies and geometries of the doublet, quartet and sextet complexes. The reduction of each **A** complex via a single ET is endothermic by at least 60.4 kJ mol⁻¹ (**Table 8.2**). In the case of **A**₅, reduction via ET caused cleavage of an Fe \cdots Im ligation and thus, the resulting complex was ignored in further studies. With reduction it was found that the preferred state was no longer solely the high-spin state. While for **A**₁, **A**₆, **A**₇ the sextet state was preferred, for

\mathbf{A}_2 , \mathbf{A}_3 and \mathbf{A}_4 the quartet state was favoured. The resulting spin and charge densities imply that the added electron goes onto the iron center thus resulting in a ferrous-superoxo ($\text{Fe(II)-O}_2^{\bullet-}$) complex.

Table 8.2. Adiabatic free energies (kJ mol^{-1}) for reduction of the $\text{Fe(III)-O}_2^{\bullet-}$ complexes (**A**) via ET or PCET.

Complex	ET ⁶²	PCET ⁶²
\mathbf{A}_1	89.1	-57.7 (-55.2)
\mathbf{A}_2	76.6	-55.6 (-42.6)
\mathbf{A}_3	73.0	-59.3 (-62.0)
\mathbf{A}_4	60.4	-84.7 (-87.4)
\mathbf{A}_5	NA	-63.1 (-64.1)
\mathbf{A}_6	80.2	-76.5 (-78.3)
\mathbf{A}_7	66.4	-77.6 (-86.5)

In contrast, if ET is coupled with PT (i.e. a PCET) to the distal oxygen of the bound O_2 , reduction of the **A** complexes becomes exothermic (**Table 8.2**). Like the reduction via ET we investigated the formation of the doublet, quartet and sextet multiplicities. It is noted that the free energies for reduction via PCET whereby the proton was localized on the proximal oxygen was also calculated. However, the energies obtained were 50 kJ mol^{-1} less exothermic than those provided in **Table 8.2**, thus will not be discussed hereafter. In all cases it was found that upon reduction via PCET the resulting complex \mathbf{A}_1 - \mathbf{A}_7 preferred the quartet multiplicity. The resulting spin and charge densities suggest the formation of ferric-peroxide (Fe(III)-OOH) complexes. Chung et al.⁶¹ have suggested that the oxidizing power of Fe(III)-O_2^- is related to the energy of the $\pi^*(\text{O}_2)$ orbital (i.e. the orbital which the added electron populates). More specifically, the lower its energy the greater the free energy change for reduction.⁶¹ Thus, it is perhaps not surprising

that ET is endothermic while PCET is significantly exothermic given the different moieties into which the added electron goes. Of the **A** complexes considered, the most powerful oxidant (**A**₄; **Figure 8.1**) is the only one in which the O₂ moiety is bound to the iron end-on. Furthermore, a H₂O is ligated to the Fe trans to the carboxylate while the cysteine is monodentately bound via its sulfur. However, the free energy of reduction of **A**₄ via a PCET process is only $-84.7 \text{ kJ mol}^{-1}$ (**Table 8.2**). This is not in itself sufficient to overcome the free energy required to oxidize the Im ($171.0 \text{ kJ mol}^{-1}$). Thus, it appears unlikely that any of the Fe(III)–O₂^{•-} complexes are suitable mechanistic oxidants.

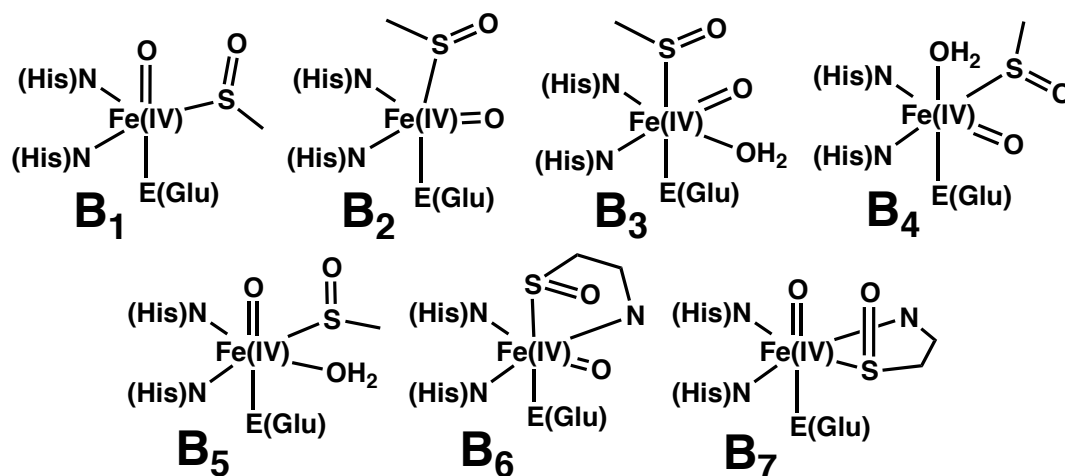


Figure 8.2. The initial five- and six-coordinate Fe(IV)=O complexes considered herein.

Next, we considered the free energy associated with reduction of the possible Fe(IV)=O containing complexes (**B**) shown in **Figure 8.2**. Using the optimized structures of the **A** complexes we manually cleaved the O–O bond with concomitant S=O bond formation. Again, we considered all possible relevant multiplicities. For each complex we calculated the energies and geometries of the singlet, triplet and quintet multiplicities. Unlike **A** (which were found to all exist in the septet) it was found that **B**₂, **B**₄, **B**₅ and **B**₇ were found to have a triplet GS while a quintet GS was favoured in **B**₁, **B**₃ and **B**₆. Interestingly, in the latter three complexes the cysteine sulfur is ligated to the iron trans to the

carboxylate. Previous investigations of Fe(IV)=O complexes have noted that a triplet GS is generally favoured.^{16,61,63-82} All complexes preferred a quintet GS at the M06/6-311G(2df,p)//B3LYP/6-31G(d) level.

Table 8.3. Adiabatic free energies (kJ mol⁻¹) for reduction of the Fe(IV)=O complexes (**B**) via ET or PCET.

Complex	ET ⁶²	PCET ⁶²
B ₁	90.0	-126.6 (-130.2)
B ₂	65.7	-114.1 (-128.8)
B ₃	38.7	-143.0 (-143.4)
B ₄	57.9	-143.9 (-136.1)
B ₅	44.4	-146.5 (-126.5)
B ₆	69.1	-139.6 (-156.3)
B ₇	67.5	-116.2 (-139.6)

Upon reduction via ET all of the resulting anionic complexes had a sextet GS (in comparison to quartet and doublet systems) except those arising from **B**₁ and **B**₅ which instead had a quartet GS. For the high-spin anions the calculated spin and charge densities indicated the formation of an Fe(III)-O⁻ complex. For the majority of the reduced complexes the iron's coordination environment was disrupted. Specifically, in **B**₃, **B**₄ and **B**₅ the H₂O was no longer ligated to the Fe-center but instead hydrogen bonded to the Fe(III)-O⁻ moiety and either the sulfoxide or carboxylate oxygen. For **B**₂ and **B**₇ an Fe···Im ligation was broken with the Im instead hydrogen bonding to the Fe(III)-O⁻ and either the sulfoxide's α-amine or S=O oxygen. Only for those anions arising from **B**₁ and **B**₆ did the Fe-center retain its coordination. As observed for the above Fe(III)-O₂⁻ complexes, reduction of all Fe(IV)=O species' via ET is endothermic (**Table 8.3**). Now, however, the process is on average less endothermic. For example, for **B**₃ reduction via ET costs just 38.7 kJ mol⁻¹.

This differs from that observed by Chung et al.⁶¹ where the Fe(III)–O₂^{•−} complexes had the higher electron affinities.

In contrast, reduction of the Fe(IV)=O complexes (**B**) via PCET did not disrupt the iron's six-coordinate ligation. However, there was again variation in the preferred GS multiplicity of the resulting complexes. While in general the quartet state was favored, **B**₂, **B**₃ and **B**₆ instead had a sextet GS. For the high-spin complexes the calculated spin densities indicated the formation of an Fe(III)-OH complex. Energetically, reduction via PCET was again found to be exothermic. Now, however, it is considerably more favored by at least 31 kJ mol^{−1} than seen for the **A** (Fe(III)–O₂^{•−}) complexes. Yet still, their exothermicity, or oxidant power, is insufficient to overcome the inherent cost (171.0 kJ mol^{−1}) associated with oxidizing the Im.

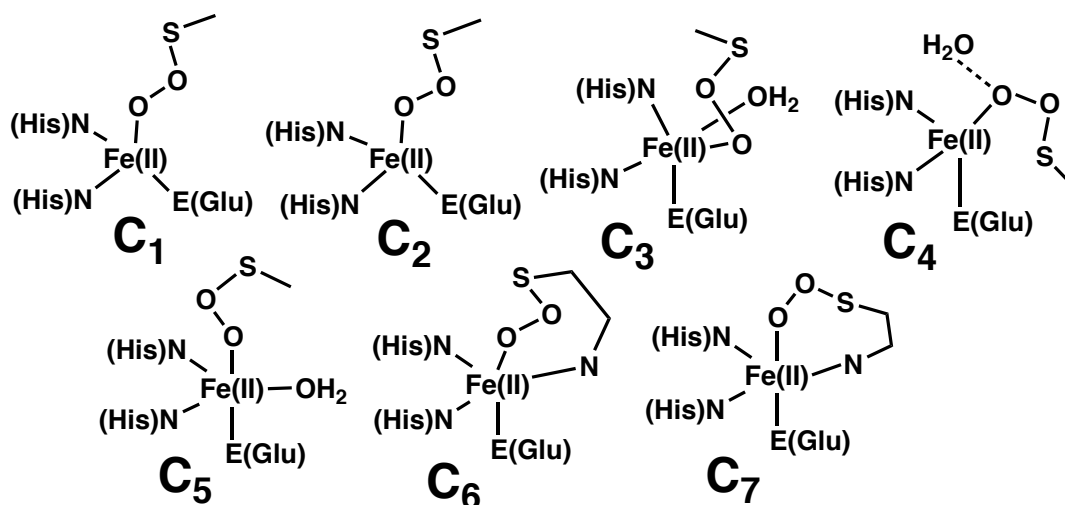


Figure 8.3. The initial five- and six-coordinate ferryl-peroxy-sulfur complexes considered herein.

Previously, de Visser and Straganz⁸³ have investigated computationally the enzyme cysteine dioxygenase (CDO), which catalytically dioxygenates a cysteine. While CDO has three ligating imidazole a key intermediate in the mechanism is a Fe(IV)=O···sulfoxide complex (as proposed in the *OvoA* and *EgtB*). Importantly, they

observed that a mechanistic intermediate containing an Fe-OO-S linkage forms prior to O–O homolytic bond cleavage. Thus, we considered the redox abilities of seven such intermediate complexes (**C**) for our present models (**Figure 8.3**). Using the optimized structures of the **A** complexes we formed an S–O bond. For each new **C** complex we calculated the geometries and energies (see Computational Methods) for the singlet, triplet and quintet multiplicities. It is noted that **C**₆ and **C**₇ are analogous to those obtained by de Visser and Straganz (i.e. bidentate ligation of cysteine).⁸³ All **C** complexes were found to prefer a quintet GS with the calculated spin densities suggesting that the Fe is in a +2 oxidation state. In agreement with previous studies,⁸³ in the optimized structures of each of the **C**-type complexes no Fe···S interaction was observed. It is noted that as a result, the Fe-center's coordination geometry in **C**₁ and **C**₂ are quite similar with a distorted trigonal-bipyramidal geometry. In particular, the peroxy moiety was essentially trans to the carboxylate ligand. In the case of **C**₄ the water also dissociated from the Fe center and instead hydrogen bonded to the proximal oxygen of the ferrous-peroxy-sulfur moiety.

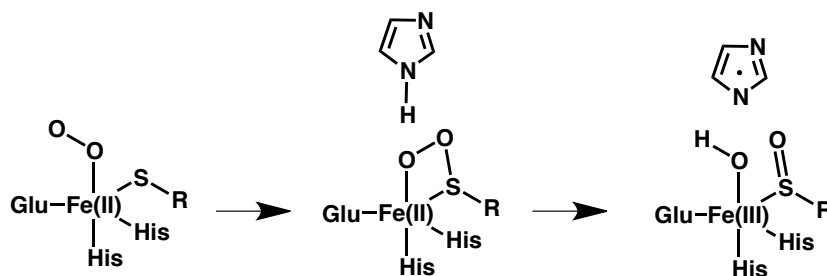
Unlike the previous systems, free energies of reduction via ET could not be obtained as the addition of an electron to each **C** complex resulted in their collapse to mechanistically infeasible complexes. As for the Fe(III)–O₂^{•-} and Fe(IV)=O complexes, reduction of the Fe-OO-S complexes via a PCET was thermodynamically favorable (**Table 8.4**). However, the exothermicity of the process is now significantly greater. Indeed, all Fe-OO-S complexes have potentials around or above 200 kJ mol⁻¹. Thus, they are all now capable of oxidizing Im to give an Im-N δ (-H)[•] radical. In such cases, subsequent C–S bond formation would require a formal proton shuttle from C δ -H or C ϵ -H to give the 2- or 5-HisCysSO product. Interestingly, however, both computational methods indicate that several of the complexes are in fact sufficiently strong enough to oxidize histidine via a PCET to directly form HisC δ (-H)[•] or HisC ϵ (-H)[•]. This would allow for C-S bond formation without the need for an intramolecular proton shuttle.

Table 8.4. Adiabatic free energies (kJ mol^{-1}) for reduction of Fe-OO-S complexes (**C**; see **Figure 8.3**) via PCET.

Complex	PCET ⁶²
C ₁	-211.1 (-274.1)
C ₂	-205.3 (-195.1)
C ₃	-203.9 (-214.6)
C ₄	-215.2 (-245.4)
C ₅	-265.1 (-284.9)
C ₆	-208.7 (-238.8)
C ₇	-268.3 (-308.4)

Upon reducing each of the Fe-OO-S complexes in **Figure 8.3** the peroxy O–O bond was cleaved, resulting in formation of complexes containing an iron-oxygen species with a Fe-bound sulfoxide! Specifically, for **C**₂, **C**₃, **C**₄ and **C**₅ the complex formed contained an Fe(II)-OH moiety and a weakly interacting sulfoxide radical while for **C**₁, **C**₆ and **C**₇ it contained an Fe(III)-OH moiety with a bound sulfoxide. It is noted that the need for sulfoxidation by *OvoA* and *EgtB* in the C–S bond formation has previously puzzled experimentalists.¹⁷ This is due partly to the fact that formation of a sulfoxide moiety appears unnecessary for subsequent steps and furthermore, it is not present in the final product.^{17,18,20,84} The results obtained herein using small model Fe-complexes suggest that formation of a sulfoxide could be a consequence of the reduction of a powerful mechanistic Fe-OO-S oxidant in order to oxidize the histidine cosubstrate. As noted in the Introduction, the use of such model biomimetic Fe-complexes does not conclusively elucidate the mechanisms of enzymes such as *OvoA* or *EgtB*. However, they can provide fundamental insights into the oxidative power of the Fe center in such systems and its dependence upon its coordination. Given the current model systems a possible pathway for enzymes such as *OvoA* and *EgtB* that involves the most thermodynamically favoured in-

intermediates is given in **Scheme 8.2**. Once further experimental data in particular structures are obtained for relevant enzyme complexes, their mechanisms can be elucidated in detail.



Scheme 8.2. Proposed intermediates for the formation of the histidyl-sulfoxide based on thermodynamic stability.

8.4 Conclusion

In summary, oxidation of histidine was thermodynamically most favorable for formation of a $\text{HisN}\delta(-\text{H})^{\bullet}$ radical via a PCET process. Of the small model iron-oxygen oxidants considered, only the ferrous-peroxy-sulfur (i.e. Fe-O-O-S) complexes were found to be inherently capable of performing this oxidation. Furthermore, several such complexes were also able to oxidize histidine to generate the higher energy radicals $\text{HisC}\delta(-\text{H})^{\bullet}$ and $\text{HisC}\epsilon(-\text{H})^{\bullet}$. Importantly, from the results the need to form the sulfoxide is rather a consequence in the formation of a more powerful oxidant in the model Fe-OO-S complexes providing insight into the puzzling need for sulfoxidation.

8.5 References

- (1) Kovaleva, E. G.; Lipscomb, J. D. *Nat. Chem. Biol.* **2008**, *4*, 186-193.
- (2) Yi, C. Q.; Jia, G. F.; Hou, G. H.; Dai, Q.; Zhang, W.; Zheng, G. Q.; Jian, X.; Yang, C. G.; Cui, Q. A.; He, C. A. *Nature* **2010**, *468*, 330-U223.
- (3) Klinman, J. P. *Acc. Chem. Res.* **2007**, *40*, 325-333.

- (4) Chen, H.; Cho, K.-B.; Lai, W.; Nam, W.; Shaik, S. *J. Chem. Theory Comput.* **2012**, *8*, 915-926.
- (5) Bollinger, J. M.; Krebs, C. *Curr. Opin. Chem. Biol.* **2007**, *11*, 151-158.
- (6) Bruijninx, P. C. A.; van Koten, G.; Gebbink, R. *Chem. Soc. Rev.* **2008**, *37*, 2716-2744.
- (7) Galonić Fujimori, D.; Barr, E. W.; Matthews, M. L.; Koch, G. M.; Yonce, J. R.; Walsh, C. T.; Bollinger, J. M.; Krebs, C.; Riggs-Gelasco, P. J. *J. Am. Chem. Soc.* **2007**, *129*, 13408-13409.
- (8) Karlsson, A.; Parales, J. V.; Parales, R. E.; Cibson, D. T.; Eklund, H.; Ramaswamy, S. *Science* **2003**, *299*, 1039.
- (9) Krebs, C.; Galonić Fujimori, D.; Walsh, C. T.; Bollinger, J. M. *Acc. Chem. Res.* **2007**, *40*, 484-492.
- (10) Matthews, M. L.; Krest, C. M.; Barr, E. W.; Vaillancourt, F. d. r. H.; Walsh, C. T.; Green, M. T.; Krebs, C.; Bollinger, J. M. *Biochemistry* **2009**, *48*, 4331-4343.
- (11) Mbughuni, M. M.; Chakrabarti, M.; Hayden, J. A.; Bominaar, E. L.; Hendrich, M. P.; Munck, E.; Lipscomb, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16788-16793.
- (12) Nam, W. *Acc. Chem. Res.* **2007**, *40*, 465-465.
- (13) Cho, K. B.; Chen, H.; Janardanan, D.; de Visser, S. P.; Shaik, S.; Nam, W. *Chem. Commun.* **2012**, *48*, 2189-2191.
- (14) van der Donk, W. A.; Krebs, C.; Bollinger, J. M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 673-683.
- (15) Gupta, R.; Lacy, D. C.; Bominaar, E. L.; Borovik, A. S.; Hendrich, M. P. *J. Am. Chem. Soc.* **2012**.
- (16) de Visser, S. P. *Angew. Chem. Int. Ed.* **2006**, *45*, 1790-1793.
- (17) Braunshausen, A.; Seebeck, F. P. *J. Am. Chem. Soc.* **2011**, *133*, 1757-1759.
- (18) Seebeck, F. P. *J. Am. Chem. Soc.* **2010**, *132*, 6632-6633.
- (19) Que, L. *Nat. Struct. Biol.* **2000**, *7*, 182-184.

- (20) Vogt, R. N.; Spies, H. S. C.; Steenkamp, D. J. *Eur. J. Biochem.* **2001**, *268*, 5229-5241.
- (21) Turner, E.; Klevit, R.; Hager, L. J.; Shapiro, B. M. *Biochemistry* **1987**, *26*, 4028-4036.
- (22) Shapiro, B. M.; Hopkins, P. B. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1991**, *64*, 291-316.
- (23) Rohl, I.; Schneider, B.; Schmidt, B.; Zeeck, E. *Z. Naturforsch. C* **1999**, *54*, 1145-1147.
- (24) Ariyanayagam, M. R.; Fairlamb, A. H. *Mol. Biochem. Parasitol.* **2001**, *115*, 189-198.
- (25) Cheah, I. K.; Halliwell, B. *Biochim. Biophys. Acta-Mol. Basis Dis.* **2012**, *1822*, 784-793.
- (26) Aruoma, O. I.; Spencer, J. P. E.; Mahmood, N. *Food Chem. Toxicol.* **1999**, *37*, 1043-1053.
- (27) Ey, J.; Schomig, E.; Taubert, D. *J. Agric. Food Chem.* **2007**, *55*, 6466-6474.
- (28) Cheah, I. K.; Halliwell, B. *Biochim. Biophys. Acta-Mol. Basis Dis.* **2012**, *1822*, 784-793.
- (29) Hartman, P. E. *Method Enzymol.* **1990**, *186*, 310-318.
- (30) Akanmu, D.; Cecchini, R.; Aruoma, O. I.; Halliwell, B. *Arch. Biochem. Biophys.* **1991**, *288*, 10-16.
- (31) Zhu, B.-Z.; Mao, L.; Fan, R.-M.; Zhu, J.-G.; Zhang, Y.-N.; Wang, J.; Kalyanaraman, B.; Frei, B. *Chem. Res. Toxicol.* **2010**, *24*, 30-34.
- (32) Hand, C. E.; Taylor, N. J.; Honek, J. F. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1357-1360.
- (33) Rahman, I.; Gilmour, P. S.; Jimenez, L. A.; Biswas, S. K.; Antonicelli, F.; Aruoma, O. I. *Biochem. Biophys. Res. Commun.* **2003**, *302*, 860-864.
- (34) Kirkham, P.; Rahman, I. *Pharmacol. Ther.* **2006**, *111*, 476-494.

- (35) Asmus, K. D.; Bensasson, R. V.; Bernier, J. L.; Houssin, R.; Land, E. J. *Biochem. J.* **1996**, *315*, 625-629.
- (36) Shapiro, B. M. *Science* **1991**, *252*, 533-536.
- (37) Holler, T. P.; Hopkins, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 4837-4838.
- (38) Zoete, V.; Bailly, F.; Catteau, J. P.; Bernier, J. L. *J. Chem. Soc., Perkin Trans. 1* **1997**, 2983-2988.
- (39) Zoete, V.; Bailly, F.; Vezin, H.; Teissier, E.; Duriez, P.; Fruchart, J. C.; Catteau, J. P.; Bernier, J. L. *Free Radic. Res.* **2000**, *32*, 515-+.
- (40) Zoete, V.; Vezin, H.; Bailly, F.; Vergoten, G.; Catteau, J. P.; Bernier, J. L. *Free Radic. Res.* **2000**, *32*, 525-+.
- (41) Jacob, C. *Nat. Prod. Rep.* **2006**, *23*, 851-863.
- (42) Krauth-Siegel, R. L.; Bauer, H.; Schirmer, H. *Angew. Chem. Int. Ed.* **2005**, *44*, 690-715.
- (43) Spies, H. S. C.; Steenkamp, D. J. *Eur. J. Biochem.* **1994**, *224*, 203-213.
- (44) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193-10206.
- (45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J.A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachar, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.;

Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian 09, Revision B.01, Gaussian, Inc.: Wallingford CT, 2010.

(46) Hirao, H. *J. Phys. Chem. B* **2011**, *115*, 11278-11285.

(47) Hirao, H.; Morokuma, K. *J. Phys. Chem. Lett.* **2010**, *1*, 901-906.

(48) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(49) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

(50) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403-412.

(51) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.

(52) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.

(53) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.

(54) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032-3041.

(55) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506-10517.

(56) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151-5158.

(57) Tomasi, J.; Mennucci, B.; Cancès, E. *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211-226.

(58) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

(59) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.

(60) Martin, J.; Baker, J.; Pulay, P. *J. Comput. Chem.* **2009**, *30*, 881-883.

(61) Chung, L. W.; Li, X.; Hirao, H.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 20076-20079.

(62) Free energies for ET were calculated via the equation $\text{Ox}_{(\text{aq})} + \text{e}^-_{(\text{aq})(\text{SHE})} = \text{Red}_{(\text{aq})}$. Free energies for PCET were calculated via the equation $\text{Ox}_{(\text{aq})} + \text{H}^+_{(\text{aq})} + \text{e}^-_{(\text{aq})(\text{SHE})} = \text{Red}_{(\text{aq})}$.

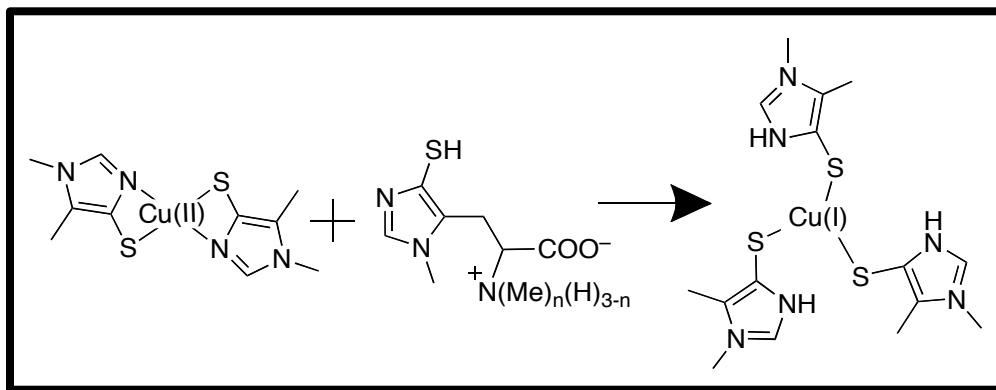
(63) Bassan, A.; Borowski, T.; Siegbahn, P. E. M. *Dalton Trans.* **2004**, 3153-3162.

(64) Blomberg, L. M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2004**, *9*, 923-935.

- (65) Borowski, T.; Blgomberg, M. R. A.; Siegbahn, P. E. M. *Chem. Eur. J.* **2008**, *14*, 2264-2276.
- (66) de Visser, S. P. *J. Am. Chem. Soc.* **2006**, *128*, 9813-9824.
- (67) de Visser, S. P. *J. Am. Chem. Soc.* **2010**, *132*, 1087-1097.
- (68) Decker, A.; Rohde, J. U.; Klinker, E. J.; Wong, S. D.; Que, L.; Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 15983-15996.
- (69) Geng, C. Y.; Ye, S. F.; Neese, F. *Angew. Chem. Int. Ed.* **2010**, *49*, 5717-5720.
- (70) Hirao, H.; Kumar, D.; Que, L.; Shaik, S. *J. Am. Chem. Soc.* **2006**, *128*, 8590-8606.
- (71) Hirao, H.; Morokuma, K. *J. Am. Chem. Soc.* **2009**, *131*, 17206-17214.
- (72) Hirao, H.; Morokuma, K. *J. Am. Chem. Soc.* **2010**, *132*, 17901-17909.
- (73) Hirao, H.; Que, L.; Nam, W.; Shaik, S. *Chem. Eur. J.* **2008**, *14*, 1740-1756.
- (74) Johansson, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Chem. Phys.* **2008**, *129*.
- (75) Lai, W. Z.; Shaik, S. *J. Am. Chem. Soc.* **2011**, *133*, 5444-5452.
- (76) Latifi, R.; Bagherzadeh, M.; de Visser, S. P. *Chem. Eur. J.* **2009**, *15*, 6651-6662.
- (77) Lundberg, M.; Kawatsu, T.; Vreven, T.; Frisch, M. J.; Morokuma, K. *J. Chem. Theory Comput.* **2009**, *5*, 222-234.
- (78) Michel, C.; Baerends, E. J. *Inorg. Chem.* **2009**, *48*, 3628-3638.
- (79) Neidig, M. L.; Decker, A.; Choroba, O. W.; Huang, F.; Kavana, M.; Moran, G. R.; Spencer, J. B.; Solomon, E. I. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12966-12973.
- (80) Salomon, O.; Reiher, M.; Hess, B. A. *J. Chem. Phys.* **2002**, *117*, 4729-4737.
- (81) Shaik, S.; Chen, H.; Janardanan, D. *Nat. Chem.* **2011**, *3*, 19-27.
- (82) Ye, S.; Neese, F. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 1228-1233.
- (83) de Visser, S. P.; Straganz, G. D. *J. Phys. Chem. A* **2009**, *113*, 1835-1846.
- (84) Ishikawa, Y.; Israel, S. E.; Melville, D. B. *J. Biol. Chem.* **1974**, *249*, 4420-4427.

Chapter 9

A Density Functional Theory Investigation into the Binding of the Antioxidants Ergothioneine and Ovothiol to Copper



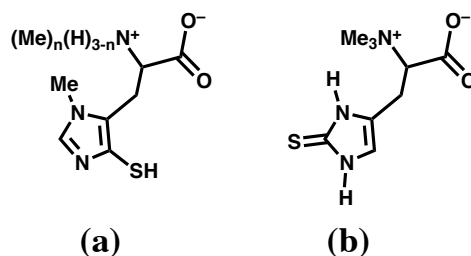
9.1 Introduction

Reactive oxygen (ROS) and nitrogen (RNS) species' are produced within cells, for instance, on exposure to ionizing radiation, as metabolic byproducts, or the action of free metal ions (e.g., Fe(II), Cu(I)).¹ Their generation can cause oxidative and nitrosative stress of important biomolecules including nucleic acids, proteins and lipids, resulting in their damage, malfunctioning or degradation.²⁻⁵ For example, their action upon DNA is known to lead to a range of alterations including nucleobase and sugar modifications, phosphodiester backbone cleavage, base-pair mismatching.¹ Such stress has been linked to aging and a variety of pathological disorders including cancer, Alzheimer's and Parkinson's disease.⁶⁻¹⁰

In response cells have developed a range of approaches to repair or mediate against such damages. In particular, they often use antioxidants to scavenge ROS and RNS or to bind free metal ions (e.g., copper) to form redox-inactive complexes.^{5,11-14} Many of these species contain sulfur and are increasingly being exploited as antioxidant therapeutics.¹⁴ In humans a major antioxidant is glutathione (GSH), an alkyl thiol derivative of cysteine. However, several other powerful antioxidants have been found that are mercaptohistidine derivatives; that is, sulfur-containing derivatives of histidine. Two such examples that have been attracting increasing attention are ovoidiol (OSH)¹⁵⁻²⁰ and ergothioneine (ESH)^{5,11,12,21-24} and are illustrated in **Scheme 9.1**. The former has been found to occur in three possible forms (A, B and C) that differ in the degree of methylation of the α -amine nitrogen.

Experimentally, OSH has been proposed to be one of the more powerful natural antioxidants.²⁵ It has been found in numerous organisms including the eggs of sea urchins, scallops, and starfish where it is believed to help protect against an increase in the concentration of H₂O₂ upon fertilization.^{16,15} In others, however, it has been suggested to have quite different roles. For example, in *Platynereis dumerilii* it has been proposed to act as a male pheromone.^{26,27} Meanwhile, ESH is found in some fungi, mycobacteria

and cyanobacteria.²⁴ While it can not be synthesized within humans, it is obtained via the diet and is thought to possibly act as an important chemoprotector.²⁸ In fact, aerosols containing it have been developed to treat chronic inflammatory diseases such as asthma.²⁹⁻³¹ In contrast to GSH, OSH and many other sulfur-containing antioxidants, ESH exists predominantly as its thione tautomer at biological pH (**Scheme 9.1**).²⁴ This is thought to contribute to its high stability against degradation, disulfide formation, and auto-oxidation.²⁴ While the exact role of ESH in humans is unknown, it has been shown to inhibit peroxynitrite-dependent nitration of tyrosine and formation of xanthines and hypoxanthines.²⁴ In addition, both OSH and ESH have been found to chelate copper and inhibit copper-induced oxidative damage of, for example, low-density lipoproteins,¹⁴ DNA and proteins.^{5,32,33}



Scheme 9.1. Schematic illustration of (a) Ovothiol (OSH_A: n = 0; OSH_B: n = 1; and OSH_C: n = 2) and (b) Ergothioneine, in their preferred state at biological pH.

OSH and ESH themselves have each been studied both computationally²⁹ and experimentally.^{17-20,22,25} For instance, Hand et al.²⁹ used a density functional theory (DFT)-based approach to examine the properties of ESH and found that the B3LYP/6-311G+(d,p) level of theory provided best agreement with X-ray crystallographic structures. Unfortunately, however, there has been only limited studies on their interactions with biologically important metal ions such as Cu(II) and Cu(I).^{13,14,34} Based on results obtained from mass spectrometry and EPR it has been suggested that the OSH-Cu(II) complex exists in a 2:1 ratio of OSH to copper, with the OSH ligated to the metal

ion via both its imidazole's N_{ϵ} and thiol's sulfur centre in a square planar geometry.¹⁴ Furthermore, using the relatively low level computational method used, PM3(TM), they also concluded that *trans*-Cu(OSH)₂ was preferred by 12.1 kJ mol⁻¹.¹⁴ In contrast, there currently exists no experimental crystal structure of, nor has there been any computational study on, ESH-Cu(II) complexes. Kimani et al.³⁴ have examined the binding of the ergothioneine analogue *N,N'*-dimethylimidazolethione (dmit) with Cu(I). In particular, in the crystal structure obtained it was shown that the Cu(I) binds three dmit ligands in a distorted trigonal planar geometry with an average Cu...S ligation length of 2.245 Å.³⁴ However, it is unclear how the di-methylation of the imidazole nitrogens in dmit affects its binding and chemistry with Cu(I). To date, there have been no computational investigations on or crystal structures obtained of OSH-Cu(I) complexes.

Herein we examine complexes formed upon ligation of the antioxidants ESH and OSH with the biologically important metal ions Cu(I) and Cu(II) using density functional theory (DFT) based methods. In particular, we have assessed the abilities of a range of DFT methods to reliably describe the structures of such complexes and their redox properties.

9.2 Computational Methods

All calculations were performed with the Gaussian 09 suite of programs.³⁵ In order to assess the ability of DFT methods to provide reliable optimized structures and energies, a series of DFT functionals and basis sets were used. Specifically, the GGA BP86,^{36,37} hybrid-GGA B3LYP,³⁸⁻⁴³ meta-GGA M06L^{44,45} and meta-hybrid-GGA M06^{44,45} functionals were selected and applied in combination with the 6-31G(d), 6-31G(d,p), 6-311G(d,p), 6-311G(2d,p), 6-311G(2df,p), 6-311+G(2df,p) and 6-311G++(3df,3pd) basis sets. The BP86 and B3LYP functionals were chosen as they have been shown to provide reliable structural parameters as well as reaction energies for transition metal (TM)-

containing complexes.^{46,47} In the case of M06 and M06L, both were parameterized with training sets containing TM complexes and importantly, have been shown to provide reliable results for TM containing systems.^{44,48} Except for 6-311++G(3df,3pd), the basis sets chosen represent a set that systematically differs by a single change, e.g., 6-31G(d,p) differs from 6-31G(d) by the inclusion of *p*-functions on hydrogens.

Aqueous solution-phase free energies were obtained via single point calculations on the above optimized geometries using the Integral Equation Formalism of the Polarizable Continuum Model (IEF-PCM),⁴⁹⁻⁵² with water as the solvent. Free energy corrections (ΔG_{corr}) were determined from the calculation of harmonic vibrational frequencies at the same level of theory as that used to obtain the optimized geometry.

Previous experimental studies into the one-electron oxidation of ESH (via pulse radiolysis) found that the deletion of the amino acid portion, i.e., the α -amino and carboxylate groups, had little effect.^{28,29} Thus, in this present study OSH and ESH were modeled by 4-thiol-N¹-methyl-5-methylimidazole and 2-thiol-4-methylimidazole, respectively. However, for simplicity we refer to the truncated molecules as OSH and ESH. Initial structures were modeled based on the results of previous experimental studies^{14,34} on ESH and/or OSH binding to Cu(I/II) (see above). In particular, trigonal-planar and square-planar geometries were assumed for the Cu(I) and Cu(II) complexes, respectively.^{14,34} For both the OSH- and ESH-Cu(II) complexes the *cis*- and *trans*-conformations were investigated, while for the latter the possible binding of ESH via either the δ - or ϵ -nitrogen to the copper was also investigated.

9.3 Results and Discussion

For the complexes investigated herein the bond lengths of main interest, and that also showed greatest dependence on level of theory, involved the copper ion (i.e., Cu \cdots X). For instance, sequentially increasing the basis set from 6-31G(d) to 6-311++G(3df,3pd)

caused changes in the C—S bonds of up to approximately 0.01 Å in total, while, in contrast, the Cu⋯X interactions varied by more than 0.1 Å! Thus, for the ESH-/OSH-Cu(I) and ESH-/OSH-Cu(II) complexes only the Cu⋯S and Cu⋯S/N bond lengths are discussed herein, unless otherwise noted. Complete optimized structures of all complexes considered are provided in the Appendix, **Tables E1-E4**.

Geometrical Assessment of [Cu(ESH)₃]⁺. In [Cu(ESH)₃]⁺ each mercaptohistidine ligates to the metal ion via *only* their sulfur centre (**Figure 9.1**). Notably, upon binding the C—S bond lengthens markedly by approximately 0.04 Å from 1.67 to 1.71 Å. This indicates that the thione sulfur now has considerable thiolate character. In agreement with that previously³⁴ observed for the analogous [Cu(dmit)₃]⁺ complex (*cf.* Introduction), the optimized geometry of [Cu(ESH)₃]⁺ adopts a distorted trigonal planar geometry about the Cu(I) ion in which each of the mercaptohistidine rings is rotated about their Cu(I)⋯S bond, thus giving the complex an overall propeller-like appearance.

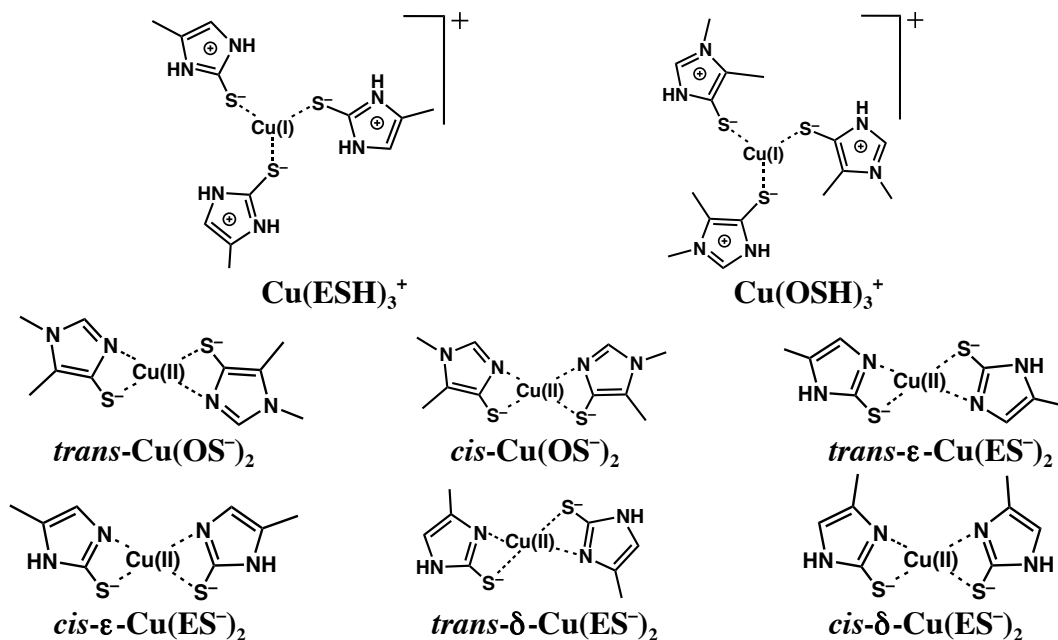


Figure 9.1. Schematic illustration of possible complexes formed by ligation of 4-thiol-N¹-methyl-5-methylimidazole (OSH) or 2-thiol-4-methylimidazole (ESH) to Cu(I) and Cu(II).

With regards to the optimized structures themselves it should be noted that at each level of theory only relatively minor differences were observed between the three Cu \cdots S bond lengths with the maximum deviation being no more than 0.03 Å (Table E1). Furthermore, as can be seen in Figure 9.2, the trends and size of changes observed in these distances upon sequentially increasing the basis set size was quite similar amongst the different functionals.

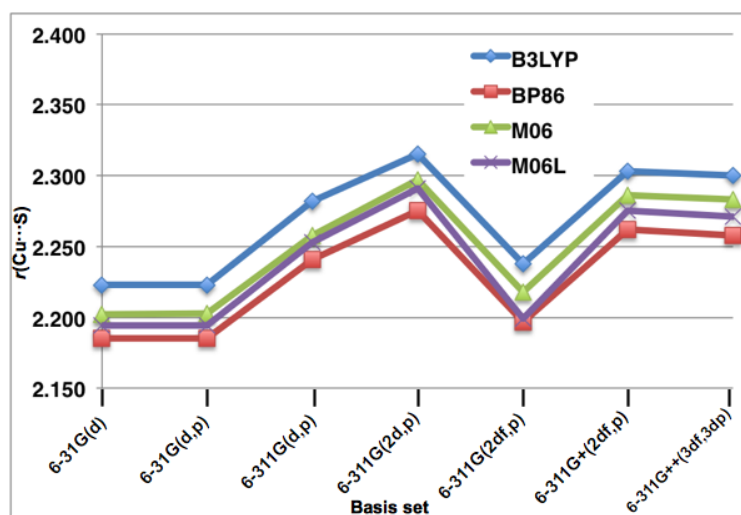


Figure 9.2. Plot of the optimized $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ bond distances (in Ångstroms) for $[\text{Cu}(\text{ESH})_3]^+$.

With the 6-31G(d) basis set the average Cu(I) \cdots S distance, $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$, obtained for $[\text{Cu}(\text{ESH})_3]^+$ with the B3LYP, BP86, M06 and M06L functionals is 2.223, 2.185, 2.202 and 2.194 Å respectively (Table E1). As might be expected, the inclusion of *p*-functions on the hydrogens (6-31G(d) \rightarrow 6-31G(d,p)) has negligible effect on the Cu \cdots S bonds which change by 0.001 Å or less. In contrast, improving the valence description from double- to triple- ζ (6-31G(d,p) \rightarrow 6-311G(d,p)) causes $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ to significantly lengthen by 0.055 – 0.059 Å. A further lengthening of 0.033 – 0.039 Å is

observed upon inclusion of a second set of d -functions on the heavy atoms (6-311G(d,p) \rightarrow 6-311G(2d,p)).

Interestingly, upon addition of f -functions on heavy atoms (6-311G(2d,p) \rightarrow 6-311G(2df,p)) the $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ distance significantly shortens by 0.077 – 0.092 Å. As a result, the Cu(I) \cdots S bond distances are now in reasonable agreement with those obtained using the considerably smaller double- ζ 6-31G(d) and 6-31G(d,p) basis sets. For example, using the BP86 functional, the $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ values obtained using the 6-31G(d,p) and 6-311G(2df,p) basis sets are just 2.185 and 2.197 respectively, a difference of just 0.012 Å (**Figure 9.2**). However, the inclusion of a set of diffuse functions on heavy atoms (6-311G(2df,p) \rightarrow 6-311+G(2df,p)) causes a considerable and almost equally opposite lengthening in $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ of 0.065 – 0.76 Å. For all DFT functionals considered herein, improving the basis set further results in quite minor decreases of 0.004 Å or less (**Figure 9.2**). As a result, the 6-311G(2d,p) and 6-311+G(2df,p) basis sets give $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ values within 0.020 Å of those obtained using the considerably larger 6-311++G(3df,3pd) basis set.

As discussed in the Introduction Kimani et al.³⁴ obtained an X-ray crystal structure of [Cu(dmit)₃]⁺ and in which the average Cu(I) \cdots S bond length was 2.245 Å.³⁴ Given the similar binding modes between dmit and ESH we have chosen to use this experimental geometrical parameter as a benchmark value. The level of theory that gives closest agreement with this value is BP86/6-311G(d,p); $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S}) = 2.241$ Å (**Table E1**). However, the BP86/6-311+G(2df,p) and BP86/6-311++G(3df,3pd) levels of theory give values that are only 0.017 and 0.013 Å larger respectively, corresponding to overestimations of 0.8% or less. In contrast, the B3LYP, M06 and M06L functionals in combination with the 6-311++G(3df,3pd) basis set all predict decidedly longer $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ distances of 2.271 – 2.300 Å (**Table E1**). Notably, for all basis sets the B3LYP method consistently gives longer $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ distances than any of the other methods considered (**Figure 9.2**). Previous computational studies of 3d-metal ion

containing compounds also found that BP86 was in general the most reliable for predicting structural parameters.^{46,47}

Geometrical Assessment of $[\text{Cu}(\text{OSH})_3]^+$. Much like the $[\text{Cu}(\text{I})(\text{ESH})_3]^+$ complex it was found that in $[\text{Cu}(\text{OSH})_3]^+$ the OSH ligands preferentially bind through their sulfurs. Furthermore, the complex adopts a propeller-like trigonal planar geometry, though the "twisting" of the mercaptohistidine ligands about their $\text{Cu}(\text{I})\cdots\text{S}$ bonds is now more pronounced. However, unlike for $[\text{Cu}(\text{ESH})_3]^+$, the three $\text{Cu}\cdots\text{S}$ distances have greater variability, differing from each other by as much as 0.2 Å depending on the level of theory (**Table E2**). Given the equivalency of the ligands and the likely fluxional nature of the complexes, we will continue to discuss changes with respect to $r_{\text{Avg}}(\text{Cu}(\text{I})\cdots\text{S})$.

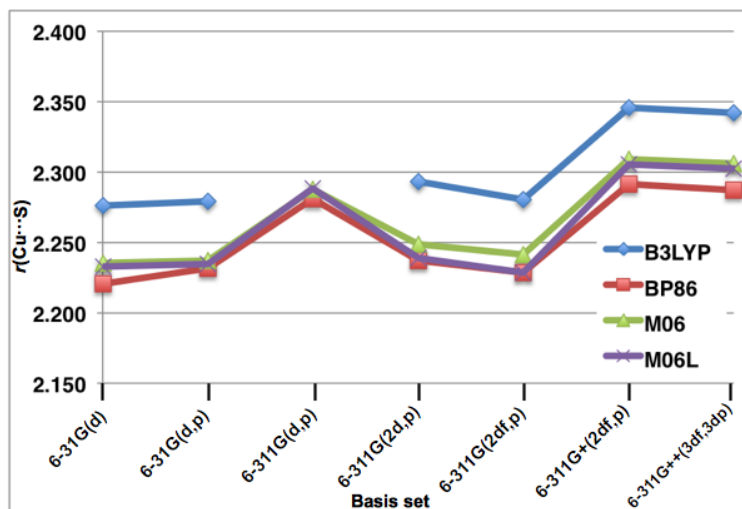


Figure 9.3. Plot of the optimized $r_{\text{Avg}}(\text{Cu}(\text{I})\cdots\text{S})$ bond distances (in Ångstroms) for $[\text{Cu}(\text{OSH})_3]^+$.

Overall, a similar trend is observed to that described for $[\text{Cu}(\text{ESH})_3]^+$; as the basis set is increased from 6-31G(d) to 6-311++G(3df,3pd) the $r_{\text{Avg}}(\text{Cu}(\text{I})\cdots\text{S})$ value is increased. However, there are a number of distinct differences. With the 6-31G(d) basis set the average $\text{Cu}(\text{I})\cdots\text{S}$ distance [$r_{\text{Avg}}(\text{Cu}(\text{I})\cdots\text{S})$] obtained for $[\text{Cu}(\text{OSH})_3]^+$ with the B3LYP,

BP86, M06 and M06L functionals is 2.276, 2.220, 2.235 and 2.233 Å respectively (**Table E2**). The inclusion of *p*-functions on the hydrogens (6-31G(d) → 6-31G(d,p)) has negligible effect on the Cu⋯S bonds which on average change by 0.005 Å. The greatest lengthening occurred for the BP86 functional where $\Delta r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S}) = 0.012$ Å. Meanwhile, improving the valence description from double- to triple- ζ (6-31G(d,p) → 6-311G(d,p)) again causes $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ to significantly lengthen by 0.049 – 0.500 Å (**Figure 9.3**). However, at the B3LYP/6-311G(d,p) level one OSH ligands in fact dissociated from the Cu center!

In contrast to that observed for $[\text{Cu}(\text{ESH})_3]^+$, the inclusion of a second set of *d*-functions on the heavy atoms (6-311G(d,p) → 6-311G(2d,p)) leads to a significant shortening in $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ of 0.039 – 0.049 Å. Furthermore, the B3LYP method again predicts a trigonal planar complex. The addition of *f*-functions again causes $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ to decrease by 0.007 – 0.013 Å. Interestingly, this change is considerably less than that seen for the ESH analogue. However, like $[\text{Cu}(\text{ESH})_3]^+$ addition of diffuse functions (6-311G(2df,p) → 6-311+G(2df,p)) caused a significant increase in $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ of 0.063 – 0.077 Å. For all DFT functionals considered, minor decreases of 0.004 Å or less were seen when the basis set was significantly improved further to 6-311++G(3df,3pd) (**Figure 9.3**).

While no crystal structures for $[\text{Cu}(\text{OSH})_3]^+$ or suitable analogue exist, it may be assumed that as the BP86/6-311+G(2df,p) level of theory provided best agreement between $[\text{Cu}(\text{ESH})_3]^+$ and $[\text{Cu}(\text{dmit})_3]^+$, it might also provide the most reasonable Cu–S bond lengths. With this basis set the BP86 method gives an $r_{\text{Avg}}(\text{Cu(I)}\cdots\text{S})$ distance of 2.291 Å, approximately 0.03 Å longer than that observed at the same level of theory for $[\text{Cu}(\text{ESH})_3]^+$. This likely reflects in part steric effects due to the methylated imidazole in OSH. In addition, there are likely to be differences in the sulfur's electron donating abilities as a result of differences in stabilization of the thiolate's negative charge via resonance in OSH compared to ESH.

We also considered the binding of OSH to Cu(I) via its N_{ϵ} centre. However, these complexes were found to be at least 64.0 kJ mol^{-1} higher in energy than the thiolate bound analogues (at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory). A preference for Cu \cdots S versus Cu \cdots N ligation has also been seen in the binding of S-nitrosothiols to Cu(I).⁵³ To better understand the process we chose to calculate the energy difference between the thiol and its zwitterionic tautomer. Interestingly, the imidazolium-thiolate form is in fact preferred in solution lying 11.8, 13.3, 4.4 and 6.2 kJ mol^{-1} lower in energy at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory (DFT_i = B3LYP, BP86, M06 or M06L, respectively). In comparison the energy difference between the ESH tautomers was 44 kJ mol^{-1} with the thione being preferred.

Geometrical Assessment of Cu(OS⁻)₂. In the optimized structures of *cis*- and *trans*-Cu(OS⁻)₂ each deprotonated OSH ligand (OS⁻) bidentately binds to the Cu(II) center via their N_{ϵ} and the thiolate sulfur centres. This is in agreement with previous results obtained from mass spectrometry and EPR studies.¹⁴ However, only when the 6-311+G(2df,p) basis set or larger was used, for all functionals, was the geometry predicted to be square planar in agreement with previous EPR studies.¹⁴ It is noted that the changes in the average Cu(II) \cdots N/S bond lengths are very similar between the *cis*- and *trans*-complexes (**Table E3**). Thus, only the *trans*-Cu(OS⁻)₂ complex will be discussed herein.

The average Cu(II) \cdots S distance obtained for *trans*-Cu(OS⁻)₂ with the B3LYP, BP86, M06 and M06L functionals in combination with the 6-31G(d) basis set is 2.414, 2.434, 2.363 and 2.382 Å respectively. Meanwhile, the average Cu(II) \cdots N distance obtained with the B3LYP, BP86, M06 and M06L functionals is 1.942, 1.899, 1.940 and 1.936 Å respectively. Notably, unlike the Cu(I) complexes, for all basis sets the BP86 method consistently gives the longest average Cu(II) \cdots S distances. Simultaneously, it also gives decidedly shorter $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ lengths than the other functionals (**Figure 9.4**).

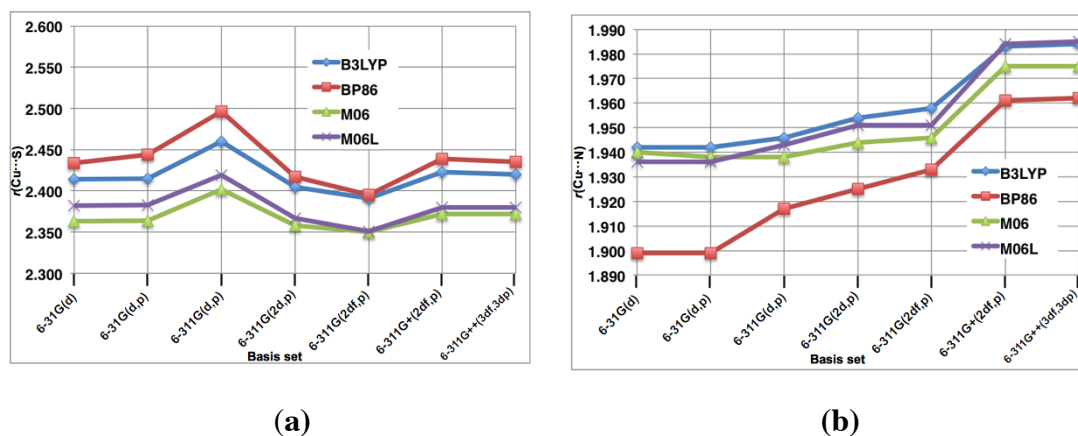


Figure 9.4. Plots of the optimized (a) $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ and (b) $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ distances (in Ångstroms) for *trans*-Cu(OS⁻)₂.

For the $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ values in [Cu(OS⁻)₂], apart from that noted above, the overall trends observed with changes in the basis set are similar to that described for [Cu(OSH)₃]⁺. However, there are some notable differences. For instance, the B3LYP method consistently maintained both Cu(II)⋯S bonds for all basis sets considered. Furthermore, the decrease in $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ length upon inclusion of a second set of *d*-functions on heavy atoms (6-311G(d,p) → 6-311G(2d,p)) is almost double that observed in [Cu(OSH)₃]⁺. Regardless, for any given functional, the $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ distance obtained using the small 6-31G(d) basis set is in reasonable agreement with that obtained using the considerably more extensive 6-311++G(3df,3pd) basis set (**Figure 9.4a**).

As seen for $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$, the $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ distances are generally unaffected by the addition of *p*-functions on hydrogen atoms. Greater effects are observed upon improving the valence description from double- to triple- ζ with $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ increasing by 0.000 – 0.018 Å. In contrast, the inclusion of a second set of *d*-functions on the heavy atoms (6-311G(d,p) → 6-311G(2d,p)), lengthens the Cu(II)⋯N bonds (**Figure 9.4b**). Furthermore, the addition of *f*-functions on heavy atoms (6-311G(2d,p) → 6-311G(2df,p)) in general causes similar increases in $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ of 0.000 – 0.008 Å as observed for the addition of a second set of *d*-functions (**Table E3**). The largest change in

$r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$, for all functionals, is in fact seen upon addition of diffuse functions on heavy atoms (6-311G(2df,p) \rightarrow 6-311+G(2df,p)) which lengthen by 0.025 – 0.033 Å. Importantly, as noted above, at the 6-311+G(2df,p) basis set the geometry found to be completely planar, thus, agreeing with the conclusions of previous experimental EPR studies.¹⁴ Negligible changes are seen upon increasing the basis set further to 6-311++G(3df,3pd) basis set. Thus, a reliable description of the $\text{Cu}(\text{OS}^-)_2$ complexes appears to only be obtained using a basis set of 6-311+G(2df,p) or larger.

A previous computational study at the PM3(TM) level of theory predicted *trans*- $\text{Cu}(\text{OS}^-)_2$ to be 12.1 kJ mol⁻¹ lower in energy than *cis*- $\text{Cu}(\text{OS}^-)_2$.¹⁴ At the IEF-PCM-M06/6-311+G(2df,p) + ΔG_{Corr} or IEF-PCM-M06L/6-311+G(2df,p) + ΔG_{Corr} levels of theory, i.e., those involving a functional of the M06 family, *trans*- $\text{Cu}(\text{OS}^-)_2$ lies marginally lower in energy by 2.3 and 1.2 kJ mol⁻¹, respectively. However, at the IEF-PCM-B3LYP/6-311+G(2df,p) + ΔG_{Corr} and IEF-PCM-BP86/6-311+G(2df,p) + ΔG_{Corr} levels of theory *trans*- $\text{Cu}(\text{OS}^-)_2$ lies higher in energy by 7.2 and 3.5 kJ mol⁻¹, respectively.

Geometrical Assessment of $\text{Cu}(\text{ES}^-)_2$. The binding of deprotonated ESH (ES^-) to Cu(II) involves the thione sulfur and an imidazole nitrogen. However, unlike OSH two nitrogens are available for binding; N_ϵ or N_δ of the imidazole ring. Thus, a total of four possible complexes were investigated and are shown in **Figure 9.1**. Similar to $\text{Cu}(\text{OS}^-)_2$ the optimized structures of $\text{Cu}(\text{ES}^-)_2$ were found to have a distorted square planar geometry with the use of smaller basis sets. Only with the use of the 6-311+G(2df,p) basis set or larger were the complexes predicted to be square planar. It is noted that the trends in $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N})$ and $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ were generally quite similar in the four $\text{Cu}(\text{ES}^-)_2$ complexes (**Table E4**). Thus, herein only the *trans*- δ - $\text{Cu}(\text{ES}^-)_2$ complex will be discussed (**Figure 9.5**).

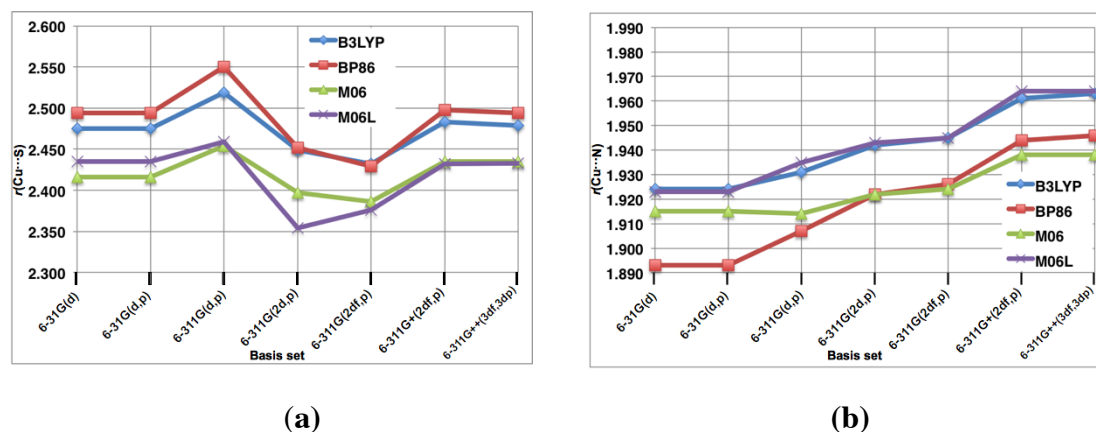


Figure 9.5. Plot of the optimized (a) $r_{Avg}(Cu(II)\cdots S)$ and (b) $r_{Avg}(Cu(II)\cdots N_\delta)$ distances (in Ångströms) for $trans-\delta-Cu(ES^-)_2$.

With the 6-31G(d) basis set the average Cu(II)···S distance obtained for $trans-Cu(ES^-)_2$ with the B3LYP, BP86, M06 and M06L functionals is 2.475, 2.494, 2.416 and 2.435 Å respectively. The average Cu(II)···N_δ distance obtained with the B3LYP, BP86, M06 and M06L functionals is 1.924, 1.893, 1.915 and 1.923 Å respectively. The overall trends observed for all functionals for both $r_{Avg}(Cu(II)\cdots S)$ and $r_{Avg}(Cu(II)\cdots N_\delta)$ are similar to that described above for $trans-Cu(OS^-)_2$. For example, as can be seen in **Figure 9.5a**, $r_{Avg}(Cu(II)\cdots S)$ in general increases upon systematically improving the basis set from 6-31G(d) to 6-311G(d,p) before then decreasing upon inclusion of a second set of *d*- or a set of *f*-functions on heavy atoms (i.e., 6-311G(d,p) → 6-311G(2d,p) → 6-311G(2df,p)). Further improving the basis set by inclusion of diffuse functions (6-311G(2df,p) → 6-311G+(2df,p)) causes $r_{Avg}(Cu(II)\cdots S)$ to lengthen such that it once again is in close agreement with that obtained using the much smaller 6-31G(d) or 6-31G(d,p) basis set. Meanwhile, in general, $r_{Avg}(Cu(II)\cdots N_\delta)$ again lengthens as the basis set is systematically improved from 6-31G(d,p) to 6-311G+(2df,p).

However, there are a number of notable differences observed in the trends of each method compared to that observed for the other complexes. In particular, for all other complexes the M06 and M06L functionals gave $r_{Avg}(Cu(II)\cdots S)$ and $r_{Avg}(Cu(II)\cdots N)$

values that were in close agreement with each other for all basis sets used herein. In *trans*- δ -Cu(ES⁻)₂, however, with the 6-311G(2d,p) basis set the M06L functional gives an $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{S})$ length that is markedly shorter by almost 0.05 Å than that obtained using the M06 method (**Figure 9.5a**). It is noted that the M06 and M06L functionals differ primarily by the amount of Hartree-Fock (HF) exchange they contain. Specifically, the former contains 27% HF exchange while the latter has none. For the $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N}_\delta)$ distance the M06L functional consistently gives values that are in close agreement with those obtained using the B3LYP method. In contrast, the M06 functional gives $r_{\text{Avg}}(\text{Cu(II)}\cdots\text{N}_\delta)$ distances that are 0.01 – 0.02 Å shorter. In fact, for the 6-311G(2d,p) basis sets and larger it gives the shortest values of all functionals considered.

Table 9.1. The aqueous solution relative free energies (kJ mol⁻¹) of *cis* and *trans*- ϵ -Cu(ES⁻)₂ and *cis* and *trans*- δ -Cu(ES⁻)₂ calculated at the DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory (DFT_i = B3LYP, BP86, M06 or M06L).

DFT _i	<i>trans</i> - δ -Cu(ES ⁻) ₂	<i>trans</i> - ϵ -Cu(ES ⁻) ₂	<i>cis</i> - δ -Cu(ES ⁻) ₂	<i>cis</i> - ϵ -Cu(ES ⁻) ₂
B3LYP	0.0	1.5	13.9	5.0
BP86	0.0	10.1	6.7	1.8
M06	0.0	3.0	5.9	1.8
M06L	0.0	1.3	4.3	6.2

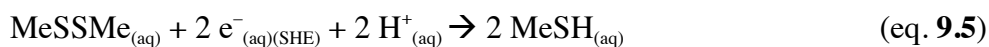
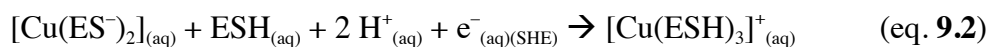
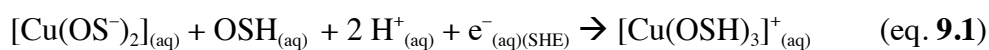
Based on the above results, the aqueous relative free energies of the four *cis* and *trans* ϵ - and δ -Cu(ES⁻)₂ complexes were calculated at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory (DFT_i = B3LYP, BP86, M06 or M06L) and are given in **Table 9.1**. The relative ordering of the geometrical isomers differs between each method. Importantly, however, all four functionals consistently predict them to lie close in energy

(<14 kJ mol⁻¹) to each other with *trans*- δ -Cu(ES⁻)₂ always being the most preferred. Hence, it is likely that all four isomers exist in aqueous solution.

From the above geometrical assessments it can be concluded that converged optimized structures for any of the given DFT functionals is obtained with the 6-311+G(2df,p) basis set, i.e., at the DFT_i/6-311+G(2df,p) level of theory. Hence, using these structures we then investigated the Cu(II)/Cu(I) redox cycling to gain a better understanding of the possible role of these two mercaptohistidine antioxidants in preventing copper induced oxidative damage.

Effects of OSH and ESH on Free Energies of Cu(II)/Cu(I) Redox Cycling. The reduction potentials for the reversible-half-reactions defined by equations **9.1** – **9.5** were calculated. It is noted that Cu(II) binds two OS⁻ or ES⁻ moieties while Cu(I) binds three OSH or ESH molecules. It is noted that Cu(II) binds two OS⁻ or ES⁻ moieties while Cu(I) binds three OSH or ESH molecules. It is noted that the proton affinities (PA) of the Cu(ESH)₃⁺ and Cu(OSH)₃⁺ complexes were calculated at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{corr} level of theory, with water as the solvent. The structures for the Cu(ESH)₂(ES⁻) and Cu(OSH)₂(OS⁻) are provided in **Table E9**. However, in the optimization of Cu(OSH)₂(OS⁻) at the B3LYP/6-311+G(2df,p) level of theory an OSH dissociated from the Cu center. However, the calculated PAs for the remaining DFT functionals were found to be at least 200 kJ mol⁻¹ greater than that of water. Thus, in aqueous environment it is expected that the Cu(ESH)₃⁺ and Cu(OSH)₃⁺ complexes would be the most populated. Hence, in modeling the reduction of the Cu(II) complexes two protons have been added as well as a third OSH or ESH for binding to the resulting Cu(I) ion (Eq. **9.1** and **9.2**). Thus, the calculated potentials include the binding of the third mercaptohistidine to the Cu center as well as the protonation of the two bound OS⁻ moieties. For comparative reasons the reduction potentials for the reversible-half-reaction of methylthiol (MeSH), a model for glutathione, is included.

For calculating reduction potentials we have treated both the electrons and protons as independent ions (Eq. **9.1** – **9.5**). Hence, their chemical potentials with respect to a particular reference state are needed. For a free electron in a dilute aqueous solution with respect to a SHE reference state (i.e. $e^-_{(aq)(SHE)}$) a chemical potential of $-418.5 \text{ kJ mol}^{-1}$ was used. Meanwhile, for a proton in a dilute aqueous environment (i.e. $H^+_{(aq)}$) a chemical potential of $-1124.2 \text{ kJ mol}^{-1}$ was used. These energies were obtained via a first principles quantum and statistical mechanics approach as detailed by Llano and Eriksson.⁵⁵ The resulting calculated reduction potentials for both isomers of $[Cu(OS^-)_2]$ and the four isomers of $[Cu(ES^-)_2]$ are given in **Table 9.2**.



As noted above the free energies of the two isomers of $[Cu(OS^-)_2]$ lie in close agreement with each other, as do those of the four isomers of $[Cu(ES^-)_2]$ (**Table 9.1**). Hence, it is likely that for each complex all possible isomers may exist simultaneously. Thus, we have calculated a weighted average according to Boltzmann statistics however, for simplicity we herein discuss in terms of an average reduction potential for each complex. For $Cu(ES^-)_2$ the B3LYP, BP86, M06 and M06L functionals give average values of 1.43, 1.39, 1.59 and 1.40 V, respectively. In contrast, the average values for $Cu(OS^-)_2$, with the exception of that obtained using the M06 method, are all notably lower at 1.32, 1.32, 1.58 and 1.21 V, respectively. The M06 functional on average predicts the largest reduction potentials for both complexes while M06L predicts the lowest (**Table 9.2**). This may only be due in small part to the degree of inclusion of HF

exchange, M06 contains 27% HF exchange while the latter has 0%. However, the potentials calculated using B3LYP (20% HF exchange) and BP86 (0% HF exchange) are identical values within two decimal places.

Table 9.2. Reduction potentials (V) for the reduction of $\text{Cu}(\text{OS}^-)_2$ and $\text{Cu}(\text{ES}^-)_2$ complexes in aqueous solution based on Eq. 9.1 and 9.2 calculated at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory.

DFT _i	Cu(OS ⁻) ₂			Cu(ES ⁻) ₂				
	<i>trans</i>	<i>cis</i>	Avg.*	<i>trans-ε</i>	<i>cis-ε</i>	<i>trans-δ</i>	<i>cis-δ</i>	Avg.*
B3LYP	1.39	1.32	1.32	1.44	1.48	1.42	1.57	1.43
BP86	1.35	1.32	1.32	1.48	1.40	1.38	1.45	1.39
M06	1.58	1.60	1.58	1.61	1.64	1.58	1.64	1.59
M06L	1.21	1.22	1.21	1.41	1.44	1.39	1.44	1.40

* Calculated using a weighted average according to Boltzmann statistics.

The standard reduction potential for reduction of $\text{Cu}(\text{II})_{(\text{aq})}$ to $\text{Cu}(\text{I})_{(\text{aq})}$, is 0.16 V.⁵⁶ Hence, it appears that (regardless of functional used) in the presence of either OSH or ESH, Cu(II) becomes considerably more oxidizing with the latter ligand having the greatest effect. Moreover, this increased potential for reduction of Cu(II) to Cu(I) suggests that the reversible cycling between the two ions is less likely to occur, That is, ligation by OSH or ESH appears to inhibit the ability of Cu(II)/Cu(I) to cause oxidative damage of important biomolecules.

In the reduction of Cu(II) to Cu(I) involving OSH or ESH the reduction potential of the relevant reactions involving the antioxidants must also be considered. In these reactions, shown in equations 9.3 – 9.5, the disulfide is reversibly reduced to give two antioxidant molecules. The resulting calculated potentials are given in **Table 9.3**. Interestingly, unlike that seen in **Table 9.2** M06 now predicts the smallest values while

B3LYP predicts the greatest for all complexes investigated. Moreover, M06L only predicts reduction potentials slightly greater than that of M06. Regardless though for both OSH and ESH the reduction potentials obtained using any DFT functional is positive. Furthermore, all suggest that OSH is markedly more reducing and thus, more reactive to oxidizing species in agreement with that proposed by Marjanovic et al.²⁵ In addition, and in agreement with a previous computational study by Hand et al.,²⁹ it is noted that both OSH and ESH are calculated to have more positive potentials compared to MeSH. More specifically, they are stronger oxidizing agents by on average 0.07 and 0.46 V, respectively.

Table 9.3. Reduction potentials (V) for reduction of the disulfides OSSO, ESSE and MeSSMe in aqueous solution calculated at the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory.

DFT _i	OSSO _(aq) → 2 OSH _(aq) (Eq. 3)	ESSE _(aq) → 2 ESH _(aq) (Eq. 4)	MeSSMe _(aq) → 2 MeSH _(aq) (Eq. 5)
B3LYP	0.39	0.83	0.33
BP86	0.31	0.70	0.25
M06	0.22	0.58	0.14
M06L	0.27	0.63	0.21
Avg.	0.30	0.69	0.23

It is noted that the value of 0.07 V obtained for the reduction potential of OSH relative to MeSH is in reasonable agreement with that obtained experimentally by Weaver and Rabensteine⁵⁷ for that of OSH relative to GSH; 0.17 V. Furthermore, Asmus et al.²⁸ have suggested that the reduction potential of ESH has an upper limit of ~1 V and with which the present results are in good agreement. In contrast, the reduction potential of GSH has been experimentally predicted to be -0.26 V vs. the SHE at biological pH

while the average value calculated herein for MeSH is +0.23 V vs. the SHE.^{57,58} However, as stated by Llano and Erikson,⁵⁴ while the measurement of reduction potentials in aqueous solutions is relatively straightforward, many factors limit their interpretation including whether the reactions are reversible or if multiple redox couples are present. Notably, such factors have been observed in the redox chemistry of OSH.¹³ Fortunately, while experimental measurements are seldom directly comparable to calculated values,⁵⁴ as evident above the latter are useful for examining relative changes in a system.

By combining the calculated potentials for Eq. 9.1 and 9.2 with those of Eq. 9.3 and 9.4 respectively, the relative abilities of the two systems in protecting against oxidative damage can be examined. At the IEF-PCM-DFT_i/6-311+G(2df,p) + ΔG_{Corr} level of theory (DFT_i = B3LYP, BP86, M06 or M06L) the calculated potentials for reduction of Cu(II) with formation of the OSSO disulfide are 0.94, 1.02, 1.36 and 0.94 V respectively. In contrast, for the analogous reaction with ESH the calculated reduction potentials were 0.61, 0.69, 1.02 and 0.77 V respectively. Hence, in agreement with experiment the reduction of Cu(II) to Cu(I) with concomitant formation of the disulfide is favoured for either OSH and ESH with the former antioxidant likely preferred given its greater reduction potential.

9.4 Conclusions

The ability of a range of density functional theory (DFT) methods to provide reliable optimized structures and thermochemical properties of Cu(I) and Cu(II) complexes with OSH and ESH has been examined. More specifically, the GGA functional BP86,^{36,37} hybrid GGA functional B3LYP,³⁸⁻⁴³ meta-GGA functional M06L^{44,45} and meta-hybrid GGA functional M06,^{44,45} in combination with 6-31G(d), 6-31G(d,p), 6-311G(d,p), 6-311G(2d,p), 6-311G(2df,p), 6-311+G(2df,p) and 6-311G++(3df,3dp) were used.

For all complexes studied, the largest impact of change in the basis set occurred in optimized lengths of both the (Cu(II)···N) and (Cu(II)···S) bonds. In general, convergence in their lengths occurred with the 6-311+G(2df,p) basis set. Indeed, for the Cu(II) complexes only use of the 6-311+G(2df,p) basis set or larger correctly predicted the complexes to be square planar. Comparison of optimized structures with the experimental data available for [Cu(ESH)₃]⁺ suggests that the BP86/6-311+G(2df,p) level of theory gives closest agreement with experiment.

Both the relative energies of the various complexes as well as reduction potentials associated with reduction of Cu(II)···ESH/OSH complexes to the corresponding Cu(I)-containing complexes were investigated. It is noted that M06 on average predicts the largest reduction potentials 0.25 V in comparison to the other functionals used. Regardless, the reduction potentials calculated for reduction of Cu(II) to Cu(I) with formation of the respective OSH/ESH disulfides (i.e., OSSO and ESSE) suggest that such processes are thermodynamically favourable in the presence of either OSH and ESH, but are preferred for OSH. However, regardless of which antioxidant OSH or ESH is used, the increased reduction potential for reduction of Cu(II) to Cu(I) suggest that the oxidation of Cu(I) to Cu(II) is less likely to occur. As a result, the redox cycling of Cu(I)/Cu(II) is inhibited. Thus, OSH and ESH at least in part inhibit the oxidative damaging abilities of copper ions in biochemical systems by altering their reduction potentials and inhibiting the recycling of Cu(I) to Cu(II).

9.5 References

- (1) Bushnell, E. A. C.; Llano, J.; Eriksson, L. A.; Gault, J. W. Mechanisms of Mutagenic DNA Nucleobase Damages and Their Chemical and Enzymatic Repairs Investigated by Quantum Chemical Methods, . In *Selected Topics in DNA Repair*; Chen, P. C., Ed.; InTech, 2011.

- (2) Gaetke, L. M.; Chow, C. K. *Toxicol.* **2003**, *189*, 147.
- (3) Aruoma, O. I.; Halliwell, B.; Gajewski, E.; Dizdaroglu, M. *Biochem. J.* **1991**, *273*, 601.
- (4) Burrows, C. J.; Muller, J. G. *Chem. Rev.* **1998**, *98*, 1109.
- (5) Zhu, B. Z.; Mao, L.; Fan, R. M.; Zhu, J. G.; Zhang, Y. N.; Wang, J.; Kalyanaraman, B.; Frei, B. *Chem. Res. Toxicol.* **2011**, *24*, 30.
- (6) Valko, M.; Rhodes, C. J.; Moncol, J.; Izakovic, M.; Mazur, M. *Chem.-Biol. Interact.* **2006**, *160*, 1.
- (7) Valko, M.; Izakovic, M.; Mazur, M.; Rhodes, C. J.; Telser, J. *Mol. Cell. Biochem.* **2004**, *266*, 37.
- (8) Valko, M.; Leibfritz, D.; Moncol, J.; Cronin, M. T. D.; Mazur, M.; Telser, J. *Int. J. Biochem. Cell Biol.* **2007**, *39*, 44.
- (9) Valko, M.; Morris, H.; Cronin, M. T. D. *Curr. Med. Chem.* **2005**, *12*, 1161.
- (10) Alvarez-Idaboy, J. R.; Galano, A. *J. Phys. Chem. B* **2012**, *116*, 9316.
- (11) Hartman, P. E. *Methods Enzymol.* **1990**, *186*, 310.
- (12) Akanmu, D.; Cecchini, R.; Aruoma, O. I.; Halliwell, B. *Arch. Biochem. Biophys.* **1991**, *288*, 10.
- (13) Zoete, V.; Vezin, H.; Bailly, F.; Vergoten, G.; Catteau, J. P.; Bernier, J. L. *Free Radic. Res.* **2000**, *32*, 525.
- (14) Zoete, V.; Bailly, F.; Vezin, H.; Teissier, E.; Duriez, P.; Fruchart, J. C.; Catteau, J. P.; Bernier, J. L. *Free Radic. Res.* **2000**, *32*, 515.
- (15) Shapiro, B. M.; Hopkins, P. B. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1991**, *64*, 291.
- (16) Turner, E.; Klevit, R.; Hager, L. J.; Shapiro, B. M. *Biochemistry* **1987**, *26*, 4028.
- (17) Ariyanayagam, M. R.; Fairlamb, A. H. *Mol. Biochem. Parasitol.* **2001**, *115*, 189.
- (18) Holler, T. P.; Hopkins, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 4837.

- (19) Zoete, V.; Bailly, F.; Catteau, J. P.; Bernier, J. L. *J. Chem. Soc., Perkin Trans. 1* **1997**, 2983.
- (20) Bailly, F.; Zoete, V.; Vamecq, J.; Catteau, J. P.; Bernier, J. L. *FEBS Lett.* **2000**, 486, 19.
- (21) Aruoma, O. I.; Whiteman, M.; England, T. G.; Halliwell, B. *Biochem. Biophys. Res. Commun.* **1997**, 231, 389.
- (22) Aruoma, O. I.; Spencer, J. P. E.; Mahmood, N. *Food Chem. Toxicol.* **1999**, 37, 1043.
- (23) Ey, J.; Schomig, E.; Taubert, D. *J. Agric. Food. Chem.* **2007**, 55, 6466.
- (24) Cheah, I. K.; Halliwell, B. *Biochim. Biophys. Acta-Mol. Basis Dis.* **2012**, 1822, 784.
- (25) Marjanovic, B.; Simic, M. G.; Jovanovic, S. V. *Free Radical Biol. Med.* **1995**, 18, 679.
- (26) Rohl, I.; Schneider, B.; Schmidt, B.; Zeeck, E. *Z. Naturforsch., C: Biosci.* **1999**, 54, 1145.
- (27) Vogt, R. N.; Spies, H. S. C.; Steenkamp, D. J. *Eur. J. Biochem.* **2001**, 268, 5229.
- (28) Asmus, K. D.; Bensasson, R. V.; Bernier, J. L.; Houssin, R.; Land, E. J. *Biochem. J.* **1996**, 315, 625.
- (29) Hand, C. E.; Taylor, N. J.; Honek, J. F. *Bioorg. Med. Chem. Lett.* **2005**, 15, 1357.
- (30) Rahman, I.; Gilmour, P. S.; Jimenez, L. A.; Biswas, S. K.; Antonicelli, F.; Aruoma, O. I. *Biochem. Biophys. Res. Commun.* **2003**, 302, 860.
- (31) Kirkham, P.; Rahman, I. *Pharmacol. Ther.* **2006**, 111, 476.
- (32) Motohash.N; Mori, I.; Sugiura, Y.; Tanaka, H. *Chem. Pharm. Bull.* **1974**, 22, 654.
- (33) Hanlon, D. P. *J. Med. Chem.* **1971**, 14, 1084.
- (34) Kimani, M. M.; Wang, H. C.; Brumaghim, J. L. *Dalton Trans.* **2012**, 41, 5248.
- (35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji,

H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision B.01. Wallingford CT, 2010.

(36) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.

(37) Perdew, J. P.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800.

(38) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(40) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.

(41) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(42) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(43) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(44) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(45) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(46) Buhl, M.; Kabrede, H. *J. Chem. Theory Comput.* **2006**, *2*, 1282.

(47) Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 27.

(48) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

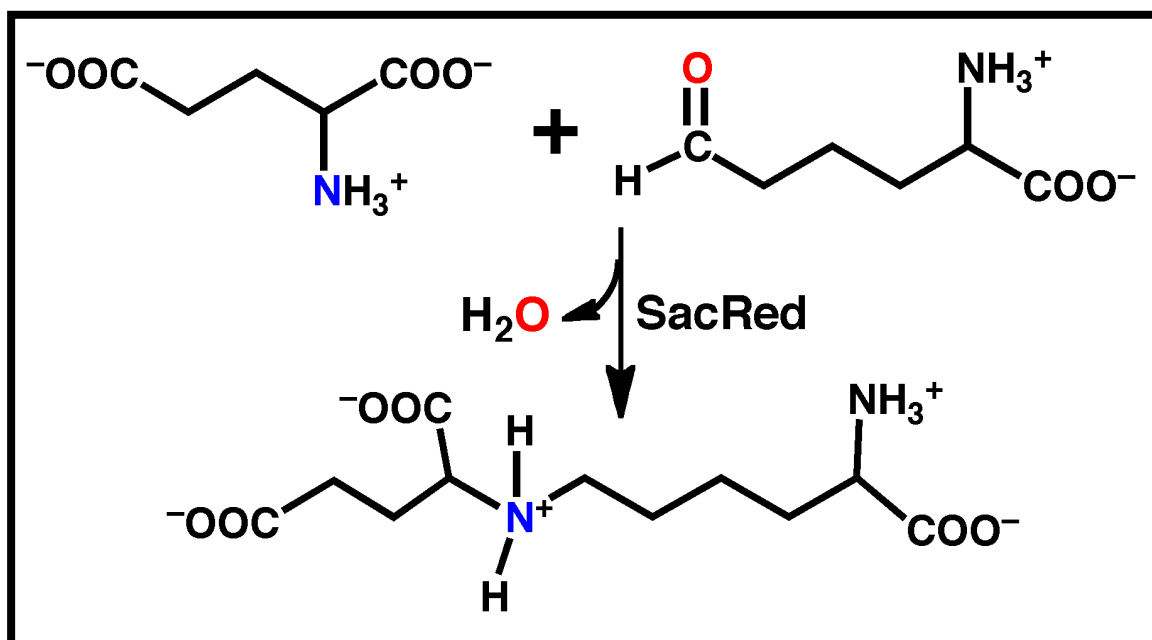
(49) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032.

(50) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506.

- (51) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151.
- (52) Tomasi, J.; Mennucci, B.; Cancès, E. *J. Mol. Struct. THEOCHEM* **1999**, *464*, 211.
- (53) Baciu, C.; Cho, K. B.; Gauld, J. W. *J. Phys. Chem. B* **2005**, *109*, 1334.
- (54) Baciu, C.; Gauld, J. W. *J. Phys. Chem. A* **2003**, *107*, 9946.
- (55) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193.
- (56) Lide, D. *CRC Handbook of Chemistry and Physics, 90th Edition (CRC Handbook of Chemistry and Physics)*; CRC, 2009.
- (57) Weaver, K. H.; Rabenstein, D. L. *J. Org. Chem.* **1995**, *60*, 1904.
- (58) Millis, K. K.; Weaver, K. H.; Rabenstein, D. L. *J. Org. Chem.* **1993**, *58*, 4144.
- (59) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193.

Chapter 10

A QM/MM-Based Computational Investigation on the Catalytic Mechanism of Saccharopine Reductase



10.1 Introduction

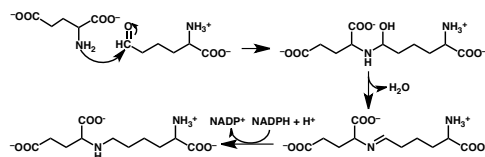
The genetic material of organisms contains the codons for twenty “standard” α -amino acids. Despite their central importance for the construction of cellular proteins and enzymes, however, not all cells are able to synthesize all twenty *de novo*. For instance, the ability to biosynthesize the essential amino acid L-lysine is limited to some green plants, bacteria, fungi, and cyanobacteria.^{1,2} In addition, it has been observed to occur via just two distinct routes: (i) the diaminopimelate (green plants, bacteria, and lower fungi), and (ii) the α -aminoadipate (cyanobacteria and higher fungi) pathways.³ For example, the fungal species, *Candida albicans*, *Cryptococcus neoformans*, *Aspergillus fumigatus*, *Saccharomyces cerevisiae* and *Magnaporthe grisea* have all been shown to utilise the α -aminoadipate pathway.¹ The former three are all human fungal agents and pose a risk to those with compromised immune systems such as AIDS, cancer, and transplant patients.⁴ *Magnaporthe grisea*, on the other hand, affects many grass and crop species’ and is perhaps best known for causing rice blast disease.⁵ Thus, this pathway represents an attractive target for the development of new fungicides.⁶

Saccharopine reductase is a key enzyme in the α -aminoadipate pathway. Specifically, it catalyses the condensation of α -aminoadipate- δ -semialdehyde (AASA) with glutamic acid and subsequent reduction by NADPH of the resulting Schiff base to give the L-lysine precursor saccharopine.¹ It has been found that the substrates of saccharopine reductase bind in the order of NADPH, AASA and lastly glutamate.¹ However, after the binding of the substrates, two possible catalytic mechanisms have been proposed for saccharopine reductase.^{1,4} Johansson *et al.*⁴ obtained X-ray crystal structures of the apo-enzyme and an enzyme···saccharopine/NADPH complex. Based on these structures they suggested that there are no suitable active-site acid/base residues able to facilitate the mechanistically required proton transfers. Hence, they concluded that the observed catalytic rate enhancement of saccharopine reductase is instead due to favourable

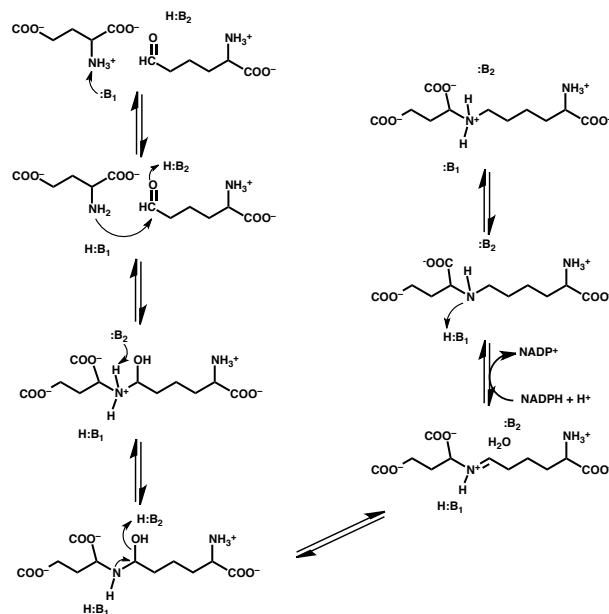
positioning of the substrates with respect to each other within the active-site.⁴ Consequently, they proposed the catalytic mechanism outlined in **Scheme 10.1**. Notably, the α -amino of the glutamic acid is initially neutral, while the α -carboxylate groups are both ionised. Thus, the first step in the overall mechanism is nucleophilic attack of the Glu-NH₂ nitrogen (^{Glu}N) at the R-group carbonyl carbon (^{AASA}C) of AASA. This occurs with concomitant transfer of a proton to AASA's R-group carbonyl oxygen (^{AASA}O) and loss of a proton from the bridging amine to give a carbinolamine intermediate. The latter then formally undergoes a 1,3-intramolecular proton transfer from its ^{Glu}NH- moiety to the newly formed hydroxyl (^{AASA}OH) group, resulting in loss of water and formation of an unprotonated-Schiff base intermediate. In the third and final step the Schiff base is reduced via hydride transfer from the NADPH cofactor with a concomitant protonation by an unknown moiety to give saccharopine.

More recently, Vashishtha *et al.*¹ experimentally examined pH-rate profiles and solvent deuterium kinetic isotope effects of saccharopine reductase from *S. cerevisiae*. Based on their observations, they concluded that it utilises an acid/base mechanism involving two active-site residues and proposed the catalytic mechanism outlined in **Scheme 10.2**. Specifically, an active-site base (**B1**), estimated¹ to have a pK_a in the range of 5.6–5.7, initially deprotonates the glutamates protonated α -amino group. However, as noted by Johansson *et al.*⁴ it is not clear what group may act as this general base. As a result, the ^{Glu}N centre is then able to nucleophilically attack at the ^{AASA}C centre. However, in contrast to that proposed by Johansson *et al.*⁴ this occurs with concomitant protonation of the ^{AASA}O centre by an acidic active site residue (**H:B2**) with an estimated pK_a of 7.8–8. Based on their X-ray crystal structure of the enzyme-product complex, Johansson *et al.*⁴ concluded that no obvious active site residue exists to facilitate this protonation. However, Vashishtha *et al.*¹ have alternatively suggested that an aspartyl (Asp126) may be able to protonate the oxyanion formed during nucleophilic attack of the glutamates amine. This results in formation of a ^{Glu}N-protonated carbinolamine intermediate.

Subsequently, **B2** abstracts a proton from the intermediate's $-\text{GluNH}_2^+-$ moiety before transferring it onto the newly formed nearby $^{\text{AASA}}\text{OH}$ group, resulting in formation of a protonated-Schiff base intermediate with loss of water (see **Scheme 10.2**). The Schiff-base is then reduced by hydride transfer from the NADPH cofactor onto the imines carbon centre, thus giving saccharopine. In the last two steps the initial active-site is regenerated with assistance of the product itself; **H:B1** transfers its proton to **B2** via the saccharopine's $-\text{GluNH}-$ group to reform **B1** and **H:B2**, *i.e.*, their initial states. Unfortunately, they were unable to identify the exact active-site acid/base residues involved.



Scheme 10.1. The catalytic mechanism of saccharopine reductase as proposed by Johansson *et al.*⁴



Scheme 10.2. The general acid/base catalytic mechanism as proposed by Vashishtha *et al.*

At present, there have been no computational investigations on the catalytic mechanism of saccharopine reductase. However, Schiff base formation has been extensively studied both experimentally and computationally due in part to their common occurrence as reaction intermediates in biochemistry and chemistry.^{4,7-9} From these studies it has been shown that Schiff base formation depends on several factors including the solvent, pH, and the chemical nature of the reactants.^{7,10-15} Mechanistically, it can be thought of occurring in two stages: (i) initial formation of a carbinolamine-type intermediate via nucleophilic attack of an amino group at a carbonyl carbon, followed by (ii) loss of its carbinolamine hydroxyl as water to give the corresponding imine. Overall, Schiff base formation is favoured at neutral pH. However, markedly lower reaction barriers are obtained if a water or some other suitable moiety facilitates the required proton transfers.⁷ In contrast, under acidic conditions, *i.e.*, those in which the attacking amino group is initially protonated, the first stage is slow as it requires deprotonation of the amino group while the subsequent stage, loss of water, is quite rapid.

Elucidation of an enzymes catalytic mechanism is central to a complete understanding of its biochemical role and the development of effective inhibitors. In this present study we have used ONIOM quantum mechanics/molecular mechanics (QM/MM) computational methods to investigate the overall mechanism of saccharopine reductase. In particular, we have examined the initial protonation states of key substrate functional groups and possible mechanistic roles of the substrates own acid/base groups.

10.2. Computational Methods

For all calculations the combined quantum mechanical and molecular mechanical (QM/MM) method in the ONIOM¹⁶⁻²⁴ formalism was applied as implemented within the Gaussian 09²⁵ program suite.

Density functional theory (DFT) is a common tool for investigating biochemical reactions.²⁶ However, it has a tendency to underestimate barriers, in particular, those corresponding to proton transfers.²⁷ In contrast, Hartree–Fock (HF) tends to overestimate barriers for proton transfer.²⁷ However, in a related computational investigation Williams²⁸ studied the condensation reaction between ammonia and formaldehyde at the HF/3–21G level of theory. They found that for nucleophilic attack of the amino at the carbonyl carbon, the lowest barrier to formation of the carbinolamine intermediate was obtained when two water molecules were involved in the reaction, in agreement with experimental predictions.²⁸ Later, as part of a computational study on Schiff base formation in the same chemical system, Hall and Smith²⁹ re–examined the reaction steps leading to formation of the carbinolamine intermediate at the considerably higher G2(MP2,SVP) level of theory. Importantly, they obtained the same series of reaction steps leading to formation of the carbinolamine intermediate as previously found by Williams.²⁸ Although the relative energies differed.

Thus, all geometry optimizations were performed at the ONIOM(HF/6-31G(d):AMBER94) level of theory within the mechanical embedding (ME) formalism.³⁰ Harmonic vibrational frequency calculations of stationary points along the potential energy surface (PES) were performed at the same level of theory in order to characterize them as minima or transition structures and to calculate Gibbs free energy corrections at standard ambient temperature and pressure (SATP).

Relative energies were then obtained by performing single point (SP) calculations at higher levels of theory based on the above–optimized geometries, with inclusion of the appropriate free energy correction. Specifically, SP calculations were performed at the; (i) ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME; (ii) ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94) within the electronic embedding (EE) formalism; and (iii) ONIOM(MP2/6-311G(d,p)//HF/6-31G(d):AMBER94)–EE levels of theory. These were chosen in order to enable systematic consideration of the effects of (i) incorporation

saccharopine moiety in the X-ray crystal structure was replaced by the two cosubstrates accordingly. Specifically, the C–N bond was cleaved and hydrogens added to the nitrogen to regenerate the initial glutamate substrate while an oxygen was added to the carbon centre thus reforming the initial α -aminoadipate- δ -semialdehyde cosubstrate. Hydrogens were added to the active-site model with all ionisable functional groups being modelled in their most likely protonation state at pH = 7. To ensure the integrity of the model during calculations the α -carbon of each residue was held fixed at its crystal structure position. It should be noted that residue Tyr100 was included in the QM layer for the examination of the initial protonation state of the Glu-NH₂ moiety and in the MM layer for the mechanism studies (see below).

10.3. Results and Discussion

The pK_a of the Substrate Glutamate's α -Amine. It has been suggested¹ that for favourable binding, the amines of both the glutamate and AASA substrates must initially be protonated. Then, once bound, the glutamate's α -NH₃⁺ group deprotonates thus enabling it to act as a nucleophile.¹ Hence, prior to an investigation of the catalytic mechanism the likely initial protonation state of the glutamate's α -amine (Glu-NH₂) was examined.

In particular we have considered the proton affinities (PAs) of the α -amine of glutamate and AASA in aqueous solution and when bound in the active site, and that of H₂O_(aq), *i.e.*, a water in the bulk aqueous environment. It is noted that the PA of an acidic group is simply the difference in electronic energy between a base (X⁻) and its conjugate acid (HX) as shown in Equation 10.1:

$$\text{PA} = E(\text{X}^-) - E(\text{HX}) \quad (\text{eq. 10.1})$$

More specifically, the PAs of AASA-NH_{2(aq)}, Glu-NH_{2(aq)} and H₂O_(aq) were obtained at the IEF-PCM(ϵ =78.3553)/MP2/6-311+G(2df,p)//HF/6-31G(d) level of theory. The

PAs of AASA–NH₂ and Glu–NH₂ when bound within the active site were obtained at the ONIOM(MP2/6-311+G(2df,p)//HF/6-31G(d):AMBER94) level of theory within the electronic embedding formalism. It is noted that in each of these “bound systems” the substrates carboxylates were modelled in their ionised forms (*i.e.*, –COO[–]). The PAs obtained are illustrated in **Figure 10.2**.

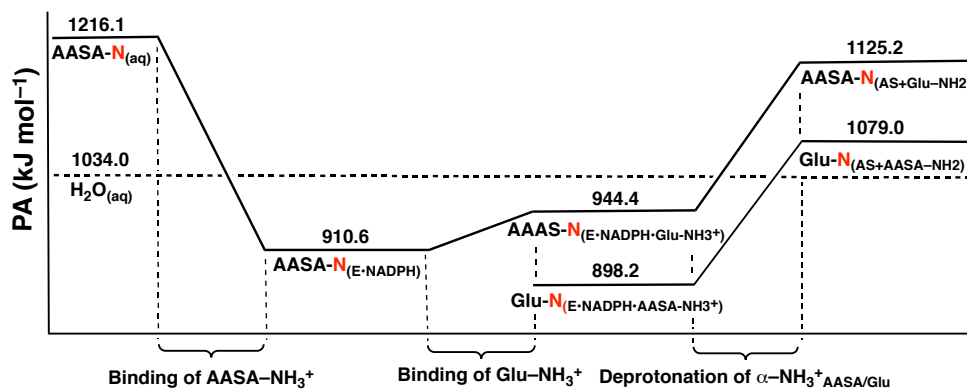


Figure 10.2. The PAs for the α -amines of AASA and Glu with respect to the local environment. The horizontal dashed line represents the PA of H₂O in solution.

In aqueous solution the PA of AASA–NH₂ is calculated to be 1,216.1 kJ mol⁻¹, which is significantly greater than that calculated for H₂O_(aq) (1,034.0 kJ mol⁻¹). Given the similarities in the α -amines of Glu and AASA, their PAs are expected to be in close agreement. Thus, in aqueous solution both AASA–NH₂ and Glu–NH₂ are likely protonated.

As noted by Vashishtha *et al.*¹ the NADPH binds within the active site first to give an E•NADPH complex, followed by AASA and then Glu. From **Figure 10.2** it can be seen that upon binding of AASA to E•NADPH the PA of its α -amine (AASA–N_(E•NADPH)) drops significantly to 910.6 kJ mol⁻¹. In fact, it is now lower than that of H₂O_(aq), thus suggesting that once bound the AASA–NH₃⁺ group could readily donate a proton to the bulk solution.

The next step is binding of Glu to the E•NADPH•AASA complex. As noted above, it has been suggested¹ that for binding the α -amine of Glu must be protonated. From **Figure 10.2** it can be seen that upon binding of Glu-NH₃⁺ the PA of AASA-NH₂ (AASA-N_(E•NADPH•Glu-NH₃⁺)) increases to 944.4 kJ mol⁻¹. In order for the α -amine of Glu to act as a nucleophile it must be neutral, *i.e.*, Glu-NH₃⁺ must lose a proton. As noted by Johannson *et al.*⁴ there appears to be no suitable base within the active site to deprotonate Glu-NH₃⁺. However, from **Figure 10.2** it can be seen that deprotonation of the glutamate α -amine results in a significant increase in the PA of AASA-NH₂ to 1,125.2 kJ mol⁻¹. In contrast, the PA of the resulting Glu-NH₂ moiety is lower at 1,079.0 kJ mol⁻¹. Importantly, both of these PAs are now larger than that of H₂O_(aq). This suggests that it is unlikely for the α -amine of both substrates (AASA and Glu) to be at least simultaneously neutral; in such a case they both can potentially accept a proton from the bulk aqueous solution. However, if the more basic AASA-NH₂ group does take up a proton, *i.e.*, becomes AASA-NH₃⁺, the PA of the Glu-NH₂ decreases considerably to 898.2 kJ mol⁻¹, and is now in fact lower than that of H₂O_(aq). Furthermore, it has the lowest PA of all possible both-substrates-bound configurations considered herein. That is, the preferred configuration of the fully bound active site has a neutral Glu-NH₂ and protonated AASA-NH₃⁺.

Mechanism for Formation of Saccharopine. The overall potential energy surface (PES) obtained for the catalytic mechanism of saccharopine reductase at the ONIOM(HF/6-31G(d):AMBER94)-ME level of theory with inclusion of Gibbs free energy corrections is presented in **Figure 10.3**. The optimized geometries of the corresponding reactant, product, intermediate complexes and transition structures, with selected distances, are presented in **Figures 10.4–10.6**.

In the optimized structure of the reactant complex (RC) both cosubstrates form intramolecular hydrogen bonds. Specifically, in the glutamate moiety its α -amino and –carboxylate groups form a reasonably strong hydrogen bond with an NH \cdots O distance of

2.12 Å (**Figure 10.4**). Meanwhile, in the α -aminoadipate- δ -semialdehyde (AASA) cosubstrate its R-group carbonyl oxygen weakly hydrogen bonds with its protonated α -amino group with an ${}^{\text{AASA}}\text{CO}\cdots\text{H}_3\text{N}^+\text{-AASA}$ distance of 3.05 Å. More importantly, however, the distance between the nitrogen centre of Glu-NH₂ and the R-group carbonyl carbon of AASA (${}^{\text{AASA}}\text{C}$), *i.e.*, $r(\text{Glu-N}(\text{H}_2)\cdots(\text{O})\text{C}^{\text{AASA}})$, is 3.58 Å. While this distance is quite long, it is shorter than that observed between these same two centres in gas-phase calculations on the complexed isolated substrates (*i.e.* in the absence of active-site and NADPH); 3.83 Å (not shown). It is also noted that the distance from the hydrogen on NADPH to be transferred later in the mechanism as a hydride to ${}^{\text{AASA}}\text{C}$ is quite long at 4.38 Å (see **Figure 10.4**). Thus, at least initially the two co-substrates and cofactor appear to interact only weakly within the active-site. However, they all appear to be suitably positioned to react further.

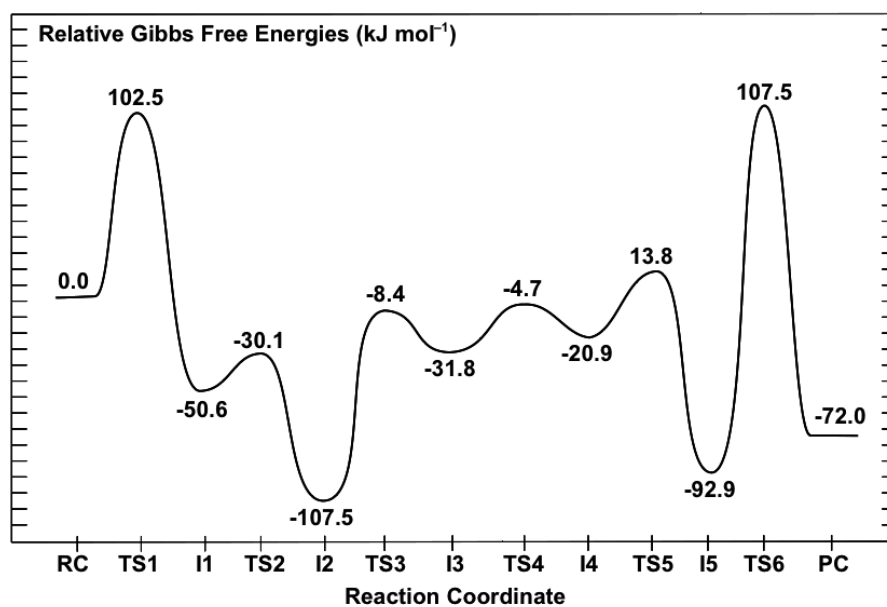


Figure 10.3. Overall PES for the catalytic mechanism of saccharopine reductase obtained at the ONIOM(HF/6-31G(d):AMBER94)-ME level of theory with inclusion of Gibbs corrections.

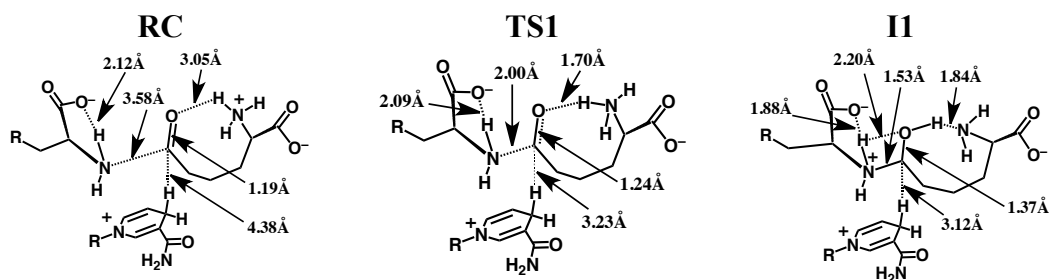


Figure 10.4. Optimized structures obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory of the reactant complex (**RC**), transition structure (**TS1**) and the carbinolamine intermediate **I1** with selected distances shown (in Ångstroms).

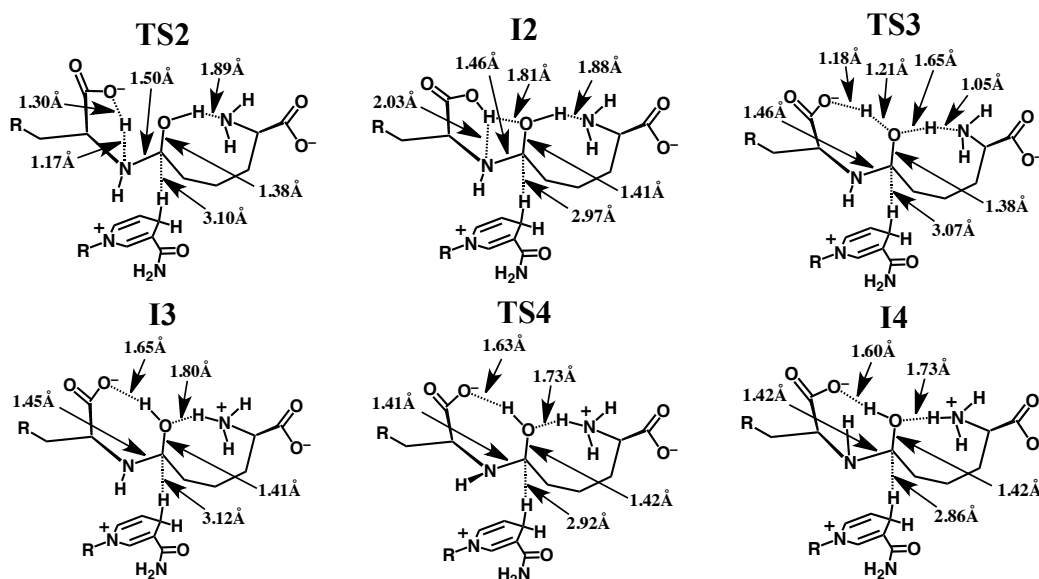


Figure 10.5. Optimized structures of **TS2**, **I2**, **TS3**, **I3**, **TS4** and **I4** obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory with selected distances shown (in Ångstroms).

Formation of a Carbinolamine Intermediate. The first step in the overall pathway is nucleophilic attack of the glutamates α -amino nitrogen at the R-group carbonyl carbon centre of the cosubstrate AASA. This occurs via transition structure **TS1** at a cost of $102.5 \text{ kJ mol}^{-1}$ at the ONIOM(HF/6-31G(d):AMBER94)–ME + Gibbs corrections level of theory (**Figure 10.3**). In **TS1** the $\text{GluN}\cdots\text{C}^{\text{AASA}}$ distance has shortened markedly to 2.00

Å while concomitantly the O–C^{AASA} bond has lengthened to 1.24 Å, *i.e.*, has reduced double-bond character (**Figure 10.4**). The nucleophilicity of the attacking amine nitrogen is likely slightly enhanced by the modest decrease in the length of the glutamates intramolecular α -NH \cdots -OOC–Glu hydrogen bond to 2.09 Å. In addition, however, the electrophilicity of the ^{AASA}C centre is enhanced by the significant decrease in the intramolecular ^{AASA}CO \cdots H₃N⁺–AASA hydrogen bond to just 1.70 Å.

In the resulting carbinolamine intermediate **II**, lying 50.6 kJ mol⁻¹ lower in energy than **RC**, the newly formed ^{AASA}C–N^{Glu} bond has a length of 1.53 Å, slightly longer than a typical C–N single bond (HF/6-31G(d): $r(\text{CH}_3\text{--NH}_2) = 1.46$ Å). Concomitantly, the ^{AASA}C–O bond has lengthened to 1.37 Å and a proton has now been transferred onto its oxygen centre from the ⁺H₃N–AASA group. While this is a substantial increase from that observed in **RC**, it is still shorter than for a typical C–O single bond (HF/6-31G(d): $r(\text{CH}_3\text{--OH}) = 1.40$ Å). This is likely due to the fact that the newly formed ^{AASA}COH group maintains a short, strong hydrogen bond with the AASA–NH₂ nitrogen centre (**Figure 10.4**). It should be noted that the ^{Glu}NH \cdots -OOC–Glu hydrogen bond has also shortened to 1.88 Å. Furthermore, the distance between the mechanistically important NADPH hydrogen and the intermediates ^{AASA}C centre has decreased markedly by 1.26 Å to 3.12 Å in **II**.

Vashishtha *et al.*¹ have suggested that an acidic residue within the active-site with a pK_a of 7.8–8.0, possibly an aspartate (Asp126), protonates the oxyanion formed during nucleophilic attack of the glutamates amine. In the optimized structure of **RC**, the side chain of Asp126 hydrogen bonds with that of Arg243 and thus, it would seem unlikely to be able to act as a general acid. Furthermore, the pK_a of an aspartate R-group carboxylate in aqueous solution is 3.8. Hence, the protein environment would have to significantly perturb its pK_a upwards by approximately 4 or more units. In addition, in the optimized structure of **RC** the side chains of Asp126 and Arg243 form a hydrogen bonded ion pair and thus, Asp126 is unlikely to be able to act as a general acid. In contrast, in aqueous

solution primary amines such as that of the cosubstrate AASA (AASA-NH_3^+) typically have $\text{p}K_a$'s in the range of 9–10 and these values can be lowered when placed within the less polar environment of a protein's active site. Indeed, the $\text{p}K_a$ measured by Vashishtha *et al.*¹ is only slightly lower than what one would anticipate for AASA-NH_3^+ , the acidic group that protonates the $^{\text{AASA}}\text{O}$ centre in our present mechanism, in aqueous solution.

Rearrangement of the Carbinolamine Intermediate II. Before Schiff base formation the carbinolamine intermediate **II** must undergo a rearrangement to allow for loss of H_2O ; specifically, deprotonation of the bridging $-\text{NH}_2^+$ moiety and inversion of the resulting $-\text{NH}-$ group.⁷

In saccharopine reductase this proceeds in a stepwise manner with the first being transfer of a proton from the bridging $-\text{Glu}^{\text{NH}_2^+}$ group onto what was initially the glutamate's carboxylate group. This occurs via **TS2** with a quite low barrier of only 20.5 kJ mol^{-1} relative to **II** (**Figure 10.3**) The resulting “neutral” carbinolamine intermediate **I2** lies significantly lower in energy than **II** by 56.9 kJ mol^{-1} , most likely due to the neutralisation of charges. As can be seen in **Figure 10.5**, in **I2** the $\text{Glu}^{\text{N}}-\text{C}^{\text{AASA}}$ bond has shortened considerably by 0.07 Å to 1.46 Å; a typical C–N single bond length (see above). Concomitantly, the C–OH bond has lengthened by 0.04 Å to 1.41 Å which similarly, is a length more typical of a C–O single bond (see above). The mechanistically important NADPH hydrogen is also now significantly closer by 0.15 Å to the intermediates $^{\text{AASA}}\text{C}$ centre. It should be noted, that the proton transferred onto the glutamate's carboxylate now forms a bifurcated hydrogen bond with both the bridging $-\text{Glu}^{\text{NH}}-$ nitrogen and the oxygen of the $^{\text{AASA}}\text{C}-\text{OH}$ moiety (**Figure 10.5**). Furthermore, the $\text{Glu}-\text{COOH}\cdots\text{N}^{\text{Glu}}$ interaction inhibits the bridging $-\text{Glu}^{\text{NH}}-$ group from inverting. As noted in the Introduction, Vashishtha *et al.*¹ have suggested that a general base with a $\text{p}K_a$ in the range of 5.6–5.7 deprotonates the α -amine of glutamate prior to nucleophilic attack. However, as detailed above, upon binding the $\text{Glu}-\text{NH}_3^+$ group appears able to readily lose a proton to the bulk aqueous solution. The above results suggest that the

carboxylate originating from the substrate glutamate may be the acid/base group measured. Indeed, while the pK_a of the $\alpha\text{-COO}^-$ of glutamate in aqueous solution is 2.1, that of acetic acid is markedly higher at 4.8. Furthermore, as noted above, the low polarity of the protein environment can also induce a shift in measured pK_a 's.

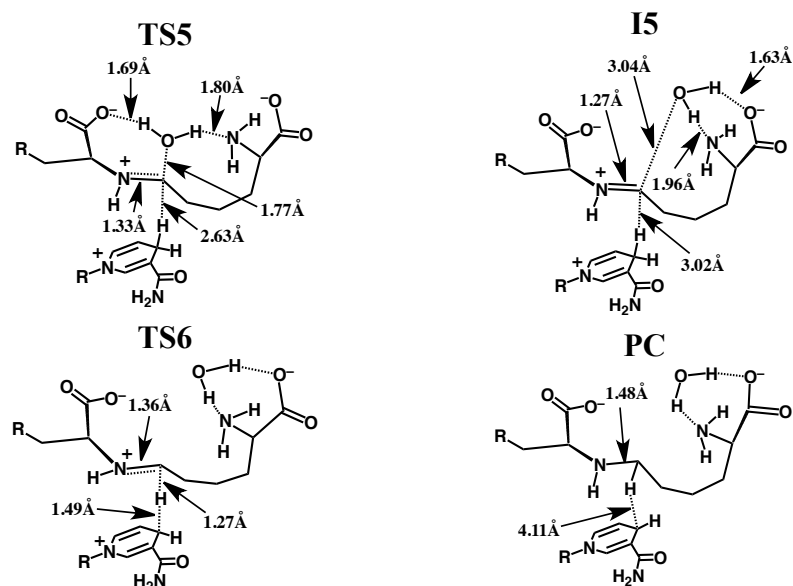


Figure 10.6. Optimized structures of **TS5**, **I5**, **TS6** and **PC** obtained at the ONIOM(HF/6-31G(d):AMBER94)–ME level of theory with selected distances shown (in Ångstroms).

Thus, the next step is transfer of the Glu–COOH proton onto the carbinolamine's $^{\text{AASA}}\text{C}\text{-OH}$ hydroxyl group, which itself simultaneously donates its proton to the nearby $\text{H}_2\text{N}\text{-AASA}$ amino moiety. This double proton transfer proceeds via **TS3** at a cost of 99.1 kJ mol^{-1} with respect to **I2**. It is noted that in **TS3** the –OH proton is essentially almost wholly transferred onto the $\text{H}_2\text{N}\text{-AASA}$ moiety while that of the Glu–COOH lies almost equidistant between both the carboxylate and carbinolamine oxygen's (see **Figure 10.5**). The resulting “charged” carbinolamine intermediate **I3** lies 75.7 kJ mol^{-1} higher in energy than **I2**, which is still 31.8 kJ mol^{-1} lower than that of the initial reactant complex

RC (Figure 10.3). Importantly, this reaction has now removed the $\text{Glu-COOH}\cdots\text{N}^{\text{Glu}}$ interaction. It is noted that we were unable to obtain any carbinolamine intermediate that contained a neutral Glu-COOH group that did not have a $\text{Glu-COOH}\cdots\text{N}^{\text{Glu}}$ hydrogen bond.

The bridging $-\text{GluNH}-$ moiety is then able to undergo an inversion as is required to enable loss of the carbinolamine $-\text{OH}$ group as H_2O to form the Schiff base.⁷ This inversion allows for the overlap in the non-bonding MO containing the nitrogen's lone pair and the anti-bonding MO of the C-O bond. Such an overlap weakens the C-O leading to a more facile bond cleavage process. The process of inversion occurs via **TS4** with a barrier of 12.6 kJ mol^{-1} with respect to **I3**, to give the alternate carbinolamine intermediate **I4** lying 10.9 kJ mol^{-1} higher in energy than **I3**. As can be seen in **Figure 10.5**, in **I4** a marginal shortening and lengthening of the $^{\text{AASA}}\text{C-N}$ and $^{\text{AASA}}\text{C-O}$ bonds, respectively is observed. In addition, the distance between the mechanistically important NADPH hydrogen and $^{\text{AASA}}\text{C}$ centre has decreased further to 2.86 \AA (**Figure 10.5**).

Formation of the Schiff Base and its Reduction. Once the bridging $-\text{GluNH}-$ has inverted, formation of the corresponding “N-protonated” Schiff base **I5** can then occur via loss of the carbinolamine $-\text{OH}$ as water. This is achieved in one step by transfer of a proton from the AASA-NH_3^+ group onto the $^{\text{AASA}}\text{C-OH}$ oxygen centre and occurs via **TS5** with a barrier of only 34.7 kJ mol^{-1} with respect to **I4** (**Figure 10.3**). The resulting imine intermediate **I5** lies markedly lower in energy than **I4** by 72.0 kJ mol^{-1} . It is noted that in **I5** the bridging $^{\text{Glu}}\text{N}-^{\text{AASA}}\text{C}$ bond has shortened significantly to 1.27 \AA while the water that was released remains hydrogen bonded to both the α -amino and $-\text{carboxylate}$ of the initial AASA cosubstrate. In addition, while the distance between the NADPH hydrogen and the $^{\text{AASA}}\text{C}$ centre has increased by 0.16 \AA to 3.02 \AA , it is still significantly closer than observed in **RC** (cf. **Figure 10.4**).

The final step in the overall catalytic pathway is formation of the saccharopine product via reduction of the Schiff base by a hydride transfer from NADPH onto the $^{\text{AASA}}\text{C}$ centre

of **I5**. It is noted that experimentally it has been found that nucleophilic attack of the N=C double bond only occurs when the Schiff base is protonated (*i.e.* when in its iminium form) as is the case for **I5**.^{29,31} This H⁻ transfer step proceeds via **TS6** with a barrier of 107.5 kJ mol⁻¹ with respect to **RC** at the ONIOM(HF/6-31G(d):AMBER94)-ME + Gibbs corrections level of theory. This barrier is lower than the generally accepted upper thermodynamic limit for enzymatic reactions of approximately 120 kJ mol⁻¹.³² However, it corresponds to a reaction barrier of 200.4 kJ mol⁻¹ relative to **I5** and thus, at the level of theory above is predicted to at least be kinetically unfavourable (**Figure 10.3**).

The final active-site bound-saccharopine complex (**PC**) is -72.0 kJ mol⁻¹ lower in energy than the initial active-site bound-reactant complex **RC**. Thus, overall, the pathway is calculated to be exothermic and thus thermodynamically favoured at the ONIOM(HF/6-31G(d):AMBER94)-ME + Gibbs corrections level of theory (**Figure 10.3**).

Obtaining More Accurate and Reliable Energies for the Mechanism of Saccharopine Reductase. As noted in the Computational Methods, previous studies have shown that the Hartree-Fock level of theory can provide a reliable mechanistic pathway for Schiff base formation, although the associated relative energies may be less accurate. However, by careful choice of higher levels of theory one can systematically consider the effects of, for example, electron correlation and the polarity of the protein environment surrounding the reactants and enzyme active-site. This is usually done by performing single-point (SP) calculations at higher levels of theory that are based on the geometries optimized at a lower level of theory, in this case ONIOM(HF/6-31G(d):AMBER94)-ME. These provide more accurate relative energies and hence, potential energy surfaces. Thus PES's were then obtained at several systematically higher levels of theory and which are presented in **Figure 10.7**. In order to facilitate comparison with the PES in **Figure 10.3**, the relative energy of RC at all levels of theory have been set to zero.

The Inclusion of Electron Correlation Effects. In the ONIOM(HF/6-31G(d):AMBER94)-ME approach the key reactive region, the QM-region, is described

by the Hartree–Fock method. This method, however, lacks inclusion of electron correlation effects which can be important in describing bond making and breaking processes. Thus, relative energies were obtained at the higher ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME level of theory with inclusion of Gibbs corrections. That is, single–points were performed in which the key QM–region is now described using the conventional electron correlation approach MP2/6-31G(d). The resulting PES obtained is shown in **Figure 10.7**; dashed blue line.

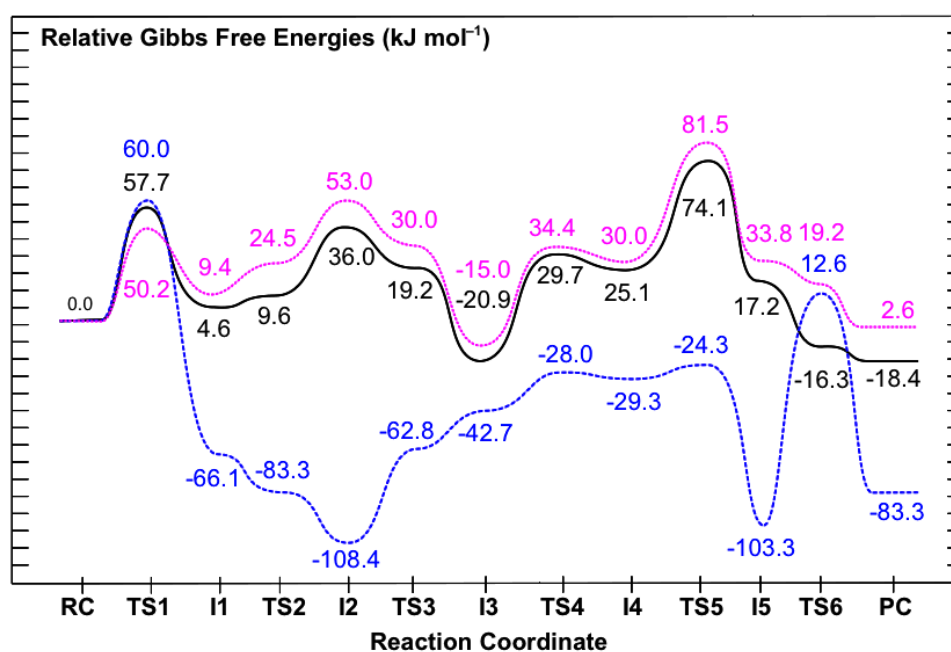


Figure 10.7. Overall PES's obtained for the catalytic mechanism of saccharopine reductase at the (i) ONIOM(MP2/6-31G(d))//HF/6-31G(d):AMBER94)–ME + Gibb's Corrections (dashed blue line), (ii) ONIOM(MP2/6-31G(d))//HF/6-31G(d):AMBER94)–EE + Gibb's Corrections (dotted pink line), and (iii) ONIOM(MP2/6-311G(d,p))//HF/6-31G(d):AMBER94)–EE + Gibb's Corrections (solid black line) levels of theory.

It can be seen clearly that a number of significant changes in relative energies occur along the catalytic pathway. In particular, those of the intermediates and product complex

all decrease by 0.9–15.5 kJ mol⁻¹ with respect to **RC**. However, the largest effects are observed for the transition structures (TS's) which all decrease considerably by 23.3–94.9 kJ mol⁻¹ with respect to **RC**. For example, the reaction barrier for the initial nucleophilic attack of the Glu–NH₂ nitrogen at the R–group carbonyl carbon of AASA is significantly reduced by 42.5 kJ mol⁻¹ to 60.0 kJ mol⁻¹; namely is, the relative energy of **TS1** for formation of the ^{Glu}N–C^{AASA} bond decreases. In fact, at this level of theory this reaction step now represents the rate-limiting step of the overall catalytic mechanism.

The resulting “N–protonated” carbinolamine **I1** lies lower in energy than **RC** by –66.1 kJ mol⁻¹, a modest lowering of 15.5 kJ mol⁻¹ (cf. **Figure 10.3**). The subsequent proton transfer from the bridging –^{Glu}NH₂⁺– moiety to the Glu–COO⁻ group now occurs essentially without a barrier via **TS2** to give the “neutralized” carbinolamine intermediate **I2**. Therefore, **I1** has become kinetically and thermodynamically unstable with respect to rearrangement to **I2**. Indeed, re-optimization of **I1** at the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME level of theory gave **I2** directly (not shown). Notably, **I2** is again the lowest energy intermediate along the catalytic pathway with respect to **RC** and now has a relative energy of –108.4 kJ mol⁻¹. This corresponds to a marginal decrease of just 0.9 kJ mol⁻¹, the smallest observed of any intermediate upon inclusion of electron correlation effects.

A large reduction in the calculated barrier for the subsequent double-proton transfer via **TS2** to give the alternate carbinolamine intermediate **I3** is also observed. Specifically, it has been reduced by 54.4 kJ mol⁻¹. Consequently, similar to that observed for **I1** at the same level of theory, **I3** which now has a relative energy of –47.2 kJ mol⁻¹ with regards to **RC**, is kinetically and thermodynamically unstable with respect to rearrangement back to **I2**. Indeed, as for **I1**, re-optimization of **I3** at the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME level of theory directly gave **I2** (not shown).

The reaction barrier for inversion of the bridging –^{Glu}NH–; namely, rearrangement of **I3** to give the alternate carbinolamine intermediate **I4** via **TS4**, is calculated to be just

14.7 kJ mol⁻¹ with respect to **I3** (**Figure 10.7**). Notably, the resulting “inverted carbinolamine” intermediate **I4** is calculated to be only slightly stable with respect to rearrangement back to **I3** by 1.3 kJ mol⁻¹. The subsequent loss of water via **TS5** is found to occur at a very low cost of 5.0 kJ mol⁻¹ with respect to **I4** (**Figure 10.7**). Thus, while the energy of **TS4** relative to **RC** has decreased by 38.1 kJ mol⁻¹ upon inclusion of electron correlation effects, this corresponds to a decrease in the actual reaction barrier height of just 2.1 kJ mol⁻¹ (cf. **Figure 10.3**). Similar to that obtained at the lower ONIOM(HF/6-31G(d):AMBER94)–ME + Gibbs corrections level of theory, the resulting Schiff base intermediate **I5** is calculated to lie very low in energy relative to **RC**. Indeed, it has only modestly decreased by 10.4 kJ mol⁻¹ to -103.3 kJ mol⁻¹.

The largest impact of including the effects of electron correlation, however, is observed in the final step of the overall pathway; reduction of the Schiff base **I5** via hydride transfer from NADPH onto the intermediates ^{AASA}C centre to give the desired saccharopine product. Specifically, as can be seen in **Figure 10.7**, the relative energy of **TS6** with respect to **RC** decreases by 94.9 kJ mol⁻¹. As a result, the barrier for this final step is markedly reduced to 115.9 kJ mol⁻¹ with respect to **I5** and in fact, is now predicted to be enzymatically and kinetically feasible.³² The overall mechanism is again predicted to be exothermic with the final saccharopine-bound active-site complex **PC** lying lower in relative energy by 83.3 kJ mol⁻¹ than the initial reactant-bound active-site complex **RC** (**Figure 10.7**).

The Effects of the Protein’s Anisotropic Polar Environment. In the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME approach above the surrounding protein environment and its effects on the reactive QM region are only treated at the molecular mechanics (MM) level of theory. In contrast, in an electronic embedding (EE) formalism the point charges of the MM protein environment are included in the self-consistent optimization of the wave function. Consequently, it enables one to examine the effects of polarization on the reactive region (QM layer) by the anisotropic protein environment.

Thus, relative energies were then also obtained at the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–EE + Gibbs corrections level of theory. The resulting PES obtained is shown in **Figure 10.7**; dotted pink line. Comparison with the PES obtained at the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME + Gibbs corrections level of theory (**Figure 10.7**, dashed blue line) provides insight into the protein environment's influence on the catalytic mechanism.

As can be seen in **Figure 10.7**, inclusion of the polarizing effects of the protein environment has a tremendous influence on the overall pathway. More specifically, the relative energy of almost all intermediates, transition structures and the product complex are now significantly raised with respect to **RC** by 27.7–161.4 kJ mol⁻¹. The only exception occurs for the initial nucleophilic attack of Glu–NH₂ at the ^{ASA}C centre via **TS1** for which the barrier is instead reduced by 9.8 kJ mol⁻¹ to 50.2 kJ mol⁻¹. Importantly, as a result, this step is no longer rate-limiting in the overall pathway (see below).

Significant changes are also observed for the carbinolamine intermediates **I1**, **I2** and **I3** and the proton transfer reactions via **TS2** and **TS3** through which they interconvert. In particular, the initial “N-protonated” carbinolamine intermediate **I1** formed is now in fact slightly higher in energy than **RC** by 9.4 kJ mol⁻¹. Furthermore, it is stable with respect to rearrangement to the subsequent “neutralised” carbinolamine **I2** (see below). In contrast, **I2** now has the highest relative energy with respect to **RC**, 53.0 kJ mol⁻¹, of the three carbinolamine intermediates **I1**, **I2** and **I3**. In addition, it is thermodynamically and kinetically unstable with respect to rearrangement back to **I1** or to the subsequent “charged” carbinolamine **I3**. This is indicated by the fact that both **TS2** and **TS3** now have lower relative energies than **I2** of 24.5 and 30.0 kJ mol⁻¹, respectively (**Figure 10.7**). The complex **I3** continues to lie lower in energy than **RC**, but by a lesser margin of 15.0 kJ mol⁻¹. However, as a result it is now the lowest energy carbinolamine intermediate of all three and in fact, is *the* lowest energy intermediate obtained along the entire catalytic pathway.

A possible explanation for these observed changes may be found by considering the substrate glutamate's carboxylate and the active-site residues with which it interacts. In particular, in **I2** the Glu-COO⁻ group hydrogen bonds to the guanidinium of an arginyl (Arg224) and phenolic R-group of a tyrosyl (Tyr78). At the previous level of theory considered, ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)-ME, the Glu-COO⁻...Arg224/Tyr78 interactions were modelled at the MM level and thus, in effect, were modelled as a steric interaction. However, by now considering the polarizing effects of the protein environment these interactions are preferred when the Glu-COO⁻ group is anionic as in **I2** and **I3**. Furthermore, the lower relative energy for **I3** may reflect that there is also preference for having the positive charge on the intermediate further removed from the carboxylate and the positively charged Arg224 residue; in **I1** the Glu-COO⁻ hydrogen bonds directly with the bridging ^{-Glu}NH₂⁺ group while in **I3** it indirectly hydrogen bonds with the AASA-NH₃⁺ group via the carbinolamine -OH moiety. In addition, the predicted instability of **I2** suggests that it may resemble a transition structure for Glu-COO⁻-assisted proton transfer from ^{-Glu}NH₂⁺ to the carbinolamine hydroxyl oxygen, which would otherwise require an inherently high energy four-membered ring transition structure.^{7,29} This is analogous to previous studies that have found lower barriers in related systems for a water-assisted proton transfer from the bridging -NH₂⁺ to the carbinolamine oxygen.^{28,29,33}

Inversion of the bridging ^{-Glu}NH- moiety via **TS4** is calculated to have now a decidedly higher barrier of 49.4 kJ mol⁻¹ with respect to **I3**. This corresponds to an increase of 34.7 kJ mol⁻¹ compared to that obtained within the mechanical embedding formalism at the same level of theory (see above). Thus, the protein environment has a greater effect on this reaction step than does the inclusion of electron correlation which resulted in a comparatively slight increase of 2.1 kJ mol⁻¹.

Considerable changes upon inclusion of the polarizing effects of the protein environment are also observed for the subsequent loss of water via **TS5** to give the Schiff

base intermediate **I5**. In particular, the barrier for this step is now 51.5 kJ mol^{-1} with respect to **I4**, a ten-fold increase compared to that obtained using the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)–ME approach. In fact, this process now represents the overall rate-limiting step along the catalytic pathway. Furthermore, **I5** lies higher in energy than **RC** by 33.8 kJ mol^{-1} and importantly, is thermodynamically and kinetically unstable with respect to further reaction via **TS6** to give the final product complex **PC** (**Figure 10.7**). Therefore, hydride transfer from the NADPH cofactor to the ^{AASA}C centre of the imine now essentially occurs without a barrier. The complex **PC** is calculated to be marginally endothermic compared to **RC** by 2.6 kJ mol^{-1} .

The Effects of Increasing the Basis Set Size. In any computational study it is important to use a basis set that adequately describes the chemical system being studied. This is particularly true when examining bond making and breaking process or those systems that involve weak, long-range or charged interactions. Thus, we also chose to examine the effects of increasing the basis set size for the reactive region, the QM layer. Specifically, the PES for the overall catalytic mechanism was obtained at the ONIOM(MP2/6-311G(d,p)//HF/6-31G(d): AMBER94)–EE + Gibbs corrections level of theory and is shown in **Figure 10.7** (solid black line). This approach also represents the best, or benchmark, level of theory used in this present study.

One can clearly see that improving the basis set used for the QM layer from 6-31G(d) to 6-311G(d,p) does not change the overall reaction pathway. Indeed, for almost all of the intermediates, transition structures and product complex, their relative energy with respect to **RC** decreases by just -4.7 – $-37.5 \text{ kJ mol}^{-1}$. The only exception is again observed for **TS1**, *i.e.*, nucleophilic attack of Glu–NH₂ at the ^{AASA}C centre, which instead increases by 7.5 kJ mol^{-1} to 57.7 kJ mol^{-1} .

For example, the carbinolamine **I1** is again predicted to lie just slightly higher in energy than **RC** by 4.6 kJ mol^{-1} . Similarly, **I3** is again calculated to lie lower in energy than **RC**, though now by 20.9 kJ mol^{-1} , a 5.9 kJ mol^{-1} decrease (**Figure 10.7**).

Furthermore, it is still the lowest energy intermediate along the overall pathway. In addition, despite a decrease in its relative energy of 17.0 kJ mol^{-1} to 36.0 kJ mol^{-1} , **I2** is still predicted to be thermodynamically and kinetically unstable with respect to rearrangement back to **I1** or on to **I3**.

A marginal increase in the barrier of just 1.2 kJ mol^{-1} to 50.6 kJ mol^{-1} is observed for inversion of the bridging $-\text{GluNH}-$ moiety via **TS4**, *i.e.*, interconversion of **I3** and **I4**. In contrast, there is a modest reduction in the subsequent barrier height for loss of the carbinolamine $-\text{OH}$ group as water via **TS5**. Specifically, the barrier is reduced by 7.4 kJ mol^{-1} to 74.1 kJ mol^{-1} with respect to **RC**. However, this reaction process again remains the overall rate-limiting step in the catalytic mechanism of saccharopine reductase. It is noted that despite this relative decrease in the barrier height of **TS5**, the energy difference between **I3** and **TS5**, the lowest energy intermediate and highest barrier respectively, remains fairly constant upon increasing the basis sets size. Indeed, this difference decreases by just 1.5 kJ mol^{-1} to 95.0 kJ mol^{-1} (**Figure 10.7**).

The resulting Schiff base intermediate **I5** now lies 17.2 kJ mol^{-1} higher in energy than **RC**, but again can essentially undergo a barrierless hydride transfer from the NADPH cofactor to the $^{\text{AASA}}\text{C}$ centre to give the final saccharopine-bound active-site product complex (**PC**). It is noted that upon increasing the basis set size the overall mechanism has once again become slightly exothermic with **PC** lying 18.4 kJ mol^{-1} lower in energy than the initial reactant-bound active-site complex **RC**.

It is interesting to note that we also used the first principles quantum and statistical mechanics approach outlined by Llano and Eriksson³⁴, in combination with small chemical models consisting of only the Schiff base itself in **I5**, *i.e.*, no active-site, and the NADPH ring from which the hydride is donated. It was predicted that the inherent free energy change for hydride transfer favoured the reduced imine product by 33.3 kJ mol^{-1} . Within the active-site, at our presently used highest level of theory ONIOM(MP2/6-311G(d,p)//HF/6-31G(d):AMBER94)-EE + Gibbs corrections, **PC** lies 35.6 kJ mol^{-1}

lower in free energy than **I5**. This suggests that for this final mechanistic step the enzyme does not aim to target distinctly or favour binding of the product over the preceding Schiff base. Rather, it simply utilises the inherent favourable free energy change for the hydride transfer.

10.4. Conclusions

A series of systematically higher-level ONIOM-based computational methods have been used in order to examine the overall catalytic mechanism of saccharopine reductase and the effects of electron correlation and the anisotropic polar protein environment on the mechanism.

The enzymes overall mechanism was elucidated using the quantum mechanics/molecular mechanics (QM/MM) ONIOM(HF/6-31G(d):AMBER94) method within the mechanical embedding (ME) formalism. The present results suggest that the catalytic mechanism does not require the direct involvement of active-site residues in the required proton transfer processes. For example, the protonated α -amine (AASA-NH₃⁺) of the cosubstrate α -aminoadipate- δ -semialdehyde (AASA) is able to act as an acid. Specifically, during nucleophilic attack of the Glu-NH₂ group at the R-group aldehyde carbon of AASA it protonates the forming oxyanion centre. In addition, the glutamate's carboxylate (Glu-COO⁻) is able to assist the proton transfer from the bridging $-\text{GluNH}_2^+$ moiety in the formed initial carbinolamine intermediate (**I1**) to the nearby carbinolamine hydroxyl oxygen. Notably, at the ONIOM(HF/6-31G(d):AMBER94)-ME + Gibb's corrections level of theory the lowest energy intermediate along the overall pathway is calculated to be the "neutralised" carbinolamine intermediate **I2** in which both the α -carboxylate and α -amino of the initial glutamate and AASA cosubstrates respectively are neutral. The overall rate-limiting step was calculated to be hydride transfer via **TS6** from the NADPH cofactor onto the bridging $-\text{NH}^+=\text{C}-$ carbon centre of the Schiff base

intermediate **I5** to give the final product bound active-site complex **PC**. In particular, it lies $107.5 \text{ kJ mol}^{-1}$ higher in energy with regards to the initial reactant complex **RC**, but is $200.4 \text{ kJ mol}^{-1}$ higher in energy than the preceding Schiff base intermediate **I5** (**Figure 10.3**).

The inclusion of electron correlation effects, on the key reactive region (QM layer) by increasing the level of theory to ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94)-ME + Gibb's corrections, leads to considerable changes along the catalytic pathway. In particular, with respect to the initial reactant complex **RC** the relative energies of the mechanistic intermediates and product complex all decreased by $0.9\text{--}15.5 \text{ kJ mol}^{-1}$ while those of the transition structures (TS's) all decreased by $23.3\text{--}94.9 \text{ kJ mol}^{-1}$. As a result, while the "neutralised" carbinolamine **I2** remains the lowest energy intermediate along the pathway ($-108.4 \text{ kJ mol}^{-1}$), the thermodynamic rate-limiting step is now nucleophilic attack of Glu-NH₂ at the R-group aldehyde carbon (^{AASA}C) of the AASA cosubstrate via **TS1** at a cost of 60.0 kJ mol^{-1} (**Figure 10.7**). The largest single reaction step barrier again occurs for reduction of the Schiff base intermediate via **TS6**, though now greatly reduced at $115.9 \text{ kJ mol}^{-1}$.

Re-examination of the PES at the ONIOM(MP2/6-31G(d)//HF/6-31G(d):AMBER94) within the electronic embedding (EE) formalism with inclusion of Gibb's corrections enabled the effects of the polarizing protein environment on the reactive region (QM layer) to be investigated. Importantly, it was found that relative to the initial reactant complex **RC** almost all intermediates, transition structures and product complex were destabilized by $27.7\text{--}161.4 \text{ kJ mol}^{-1}$; namely, their relative energy was raised. The only exception occurred for the initial nucleophilic attack of Glu-NH₂ on the AASA cosubstrate via **TS1** for which the barrier decreased by 9.8 kJ mol^{-1} . Consequently, the carbinolamine **I3** lying 15.0 kJ mol^{-1} lower in energy than **RC** was now found to be the lowest energy intermediate along the overall pathway. Furthermore, the rate limiting step is now loss of water from the "inverted" carbinolamine **I4** via **TS5**, at a cost of 81.5 kJ

mol^{-1} with respect to **RC**, to give the Schiff base **I5**. In addition, the subsequent and final step in the overall pathway, reduction of the Schiff base, is found to occur now essentially without a barrier.

Increasing the size of the basis set used to describe the key QM layer from 6-31G(d) to 6-311G(d,p) was also considered by the use of the ONIOM(MP2/6-311G(d,p)//HF/6-31G(d):AMBER94)-EE method with inclusion of Gibb's corrections. In general, only comparatively modest decreases of 4.7–37.5 kJ mol^{-1} in the relative energy of all intermediates, transition structures and product complex with respect to **RC** were observed. The only exception being for **TS1** whose relative energy increased by 7.5 kJ mol^{-1} . This was also the highest level of theory used in this present study. The carbinoalamine **I3** is again the the lowest energy intermediate along the catalytic pathway being 20.9 kJ mol^{-1} lower in energy than **RC**. The overall rate-limiting step is the loss of water to give the Schiff base intermediate **I5** which occurs via **TS6** at a cost of 74.1 kJ mol^{-1} with respect to **RC**. A subsequent barrierless hydride transfer reduces **I5** to the final saccharopine product.

Experimentally, it has been suggested that two active site acid/base residues with $\text{p}K_a$'s of 5.6–5.7 and 7.8–8.0 facilitate the mechanistically required proton transfers.¹ However, it has also been experimentally suggested that there are no active site acid/base residues available and thus, the enzyme functions by orientating and positioning the substrates for reaction.⁴ The present results suggest that acid/base functional groups within the substrates themselves, specifically the α -amine of α -aminoadipate- δ -semialdehyde and α -carboxylate of glutamate, are able to catalyse the mechanistically required proton transfer reactions, in support of previous suggestions by Johansson *et al.*⁴ In addition, it is also suggested that the two $\text{p}K_a$ values experimentally measured by Vashishtha *et al.*¹ may in fact correspond to these two substrate functional groups. That is, based on the extensive and high-level computational models used herein, the present results suggest that saccharopine reductase catalyses the overall reaction by binding the

three required reactant molecules glutamate, α -aminoadipate- δ -semialdehyde and NADPH in an orientation and polar environment conducive to reaction.

10.5 References

- (1) Vashishtha, A. K.; West, A. H.; Cook, P. F. *Biochemistry* **2009**, *48*, 5899-5907.
- (2) Zabriskie, T. M.; Jackson, M. D. *Nat. Prod. Rep.* **2000**, *17*, 85-97.
- (3) Xu, H. Y.; Andi, B.; Qian, J. H.; West, A. H.; Cook, P. F. *Cell Biochem. Biophys.* **2006**, *46*, 43-64.
- (4) Johansson, E.; Steffens, J. J.; Lindqvist, Y.; Schneider, G. *Structure* **2000**, *8*, 1037-1047.
- (5) Talbot, N. J. *Trends in Microbiology* **1995**, *3*, 9-16.
- (6) Ribot, C.; Hirsch, J.; Batzergue, S.; Tharreau, D.; Notteghem, J. L.; Lebrun, M. H.; Morel, J. B. *J. Plant Physiol.* **2008**, *165*, 114-124.
- (7) Erdtman, E.; Bushnell, E. A. C.; Gauld, J. W.; Eriksson, L. A. *Comput. Theor. Chem.* **2011**, *963*, 479-489.
- (8) Kaya, I.; Dogan, F.; Bilici, A. *Polym. Int.* **2009**, *58*, 570-578.
- (9) Lochee, Y.; Jhurry, D.; Bhaw-Luximon, A.; Kalangos, A. *Polym. Int.* **2010**, *59*, 1310-1318.
- (10) Godoy-Alcantar, C.; Yatsimirsky, A. K.; Lehn, J. M. *J. Phys. Org. Chem.* **2005**, *18*, 979-985.
- (11) Ilieva, S.; Cheshmedzhieva, D.; Tasheva, D. *Tetrahedron* **2010**, *66*, 5168-5172.
- (12) Lo Presti, L.; Soave, R.; Longhi, M.; Ortoleva, E. *Acta Crystallogr. Sect. B-Struct. Sci.* **2010**, *66*, 527-543.
- (13) Platas-Iglesias, C.; Esteban, D.; Ojea, V.; Avecilla, F.; de Blas, A.; Rodriguez-Blas, T. *Inorg. Chem.* **2003**, *42*, 4299-4307.
- (14) Praveen, P. L.; Ajeetha, N.; Ojha, D. P. *Russ. J. Gen. Chem.* **2009**, *79*, 2267-2271.

- (15) Sato, Y.; Hata, M.; Neya, S.; Hoshino, T. *J. Phys. Chem. B* **2006**, *110*, 22804-22812.
- (16) Bearpark, M. J.; Ogliaro, F.; Vreven, T.; Boggio-Pasqua, M.; Frisch, M. J.; Larkin, S. M.; Robb, M. A. In *Computation in Modern Science and Engineering Vol 2, Pts a and B*; Simos, T. E., Maroulis, G., Eds.; Amer Inst Physics: Melville, 2007; Vol. 2, p 583-585.
- (17) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *Theochem-J. Mol. Struct.* **1999**, *461*, 1-21.
- (18) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959-1967.
- (19) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179.
- (20) Morokuma, K.; Musaev, D. G.; Vreven, T.; Basch, H.; Torrent, M.; Khoroshun, D. V. *IBM J. Res. Dev.* **2001**, *45*, 367-395.
- (21) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357-19363.
- (22) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.
- (23) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419-1432.
- (24) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760-769.
- (25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J.A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachar, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.;

Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian 09, Revision B.01, Gaussian, Inc.: Wallingford CT, 2010.

(26) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232-240.

(27) Koch, W., Holthausen, M.C. *A Chemist's Guide to Density Functional Theory*; Second ed.; Wiley-VCH: New York, 2001.

(28) Williams, I. H. *J. Am. Chem. Soc.* **1987**, *109*, 6299-6307.

(29) Hall, N. E.; Smith, B. J. *J. Phys. Chem. A* **1998**, *102*, 4930-4938.

(30) Case, D. A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J. *J. Comput. Chem.* **2005**, *26*, 1668-1688.

(31) Brault, M.; Pollack, R. M.; Bevins, C. L. *J. Org. Chem.* **1976**, *41*, 346-350.

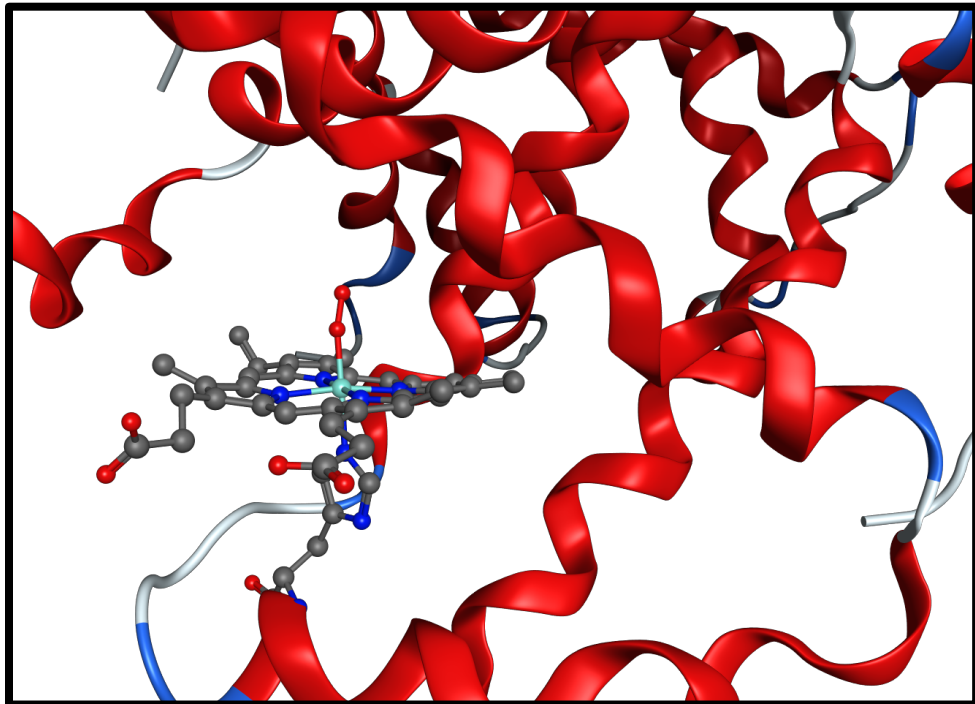
(32) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729-738.

(33) Rosenberg, S.; Silver, S. M.; Sayer, J. M.; Jencks, W. P. *J. Am. Chem. Soc.* **1974**, *96*, 7986-7998.

(34) Llano, J.; Eriksson, L. A. *J. Chem. Phys.* **2002**, *117*, 10193-10206.

Chapter 11

Conclusions



11.1 Conclusions

Using a broad range of computational tools and theoretical methods we have investigated several important biochemical systems. Importantly, such studies provide greater insight into redox and metal ion biochemistry.

In Chapter 3 MD methods allowed the investigation into the structure of the fully bound active site of ThrRS for the second half-reaction. In particular, the ability of His309 to act as either a mechanistic base or acid, without or with the assistance of a "bridging" water was examined. Upon addition of a water molecule to "bridge" the His309-N^E:...O-2'-Ado76 interaction, for both the neutral and protonated His309, significant disruptions to the orientation of the active-site groups was observed. With no additional "bridging" water added regardless of the protonation state of His309 the results obtained suggest that the active site histidyl is unlikely to be catalytically involved. That is, His309 is unlikely to act as a catalytic base or acid in the transfer of the threonyl moiety to ^{Thr}RNA. From these results a subsequent DFT-cluster investigation was performed that considered the involvement of His309 in the catalytic mechanism of ThrRS.¹ Importantly, it was found that it did not act as the general base or acid but rather as shown in the work of Huang et al.¹ a Zn within the active site ensures that the α -amine of the substrate remained neutral allowing it to act as the catalytic base.

In Chapter 4 docking and MD simulations allowed for the examination of several bound hUROD...URO-III complexes. Importantly, the results showed that several complexing modes are possible between URO-III and the active site of hUROD. In all complexes investigated, it was found that the -NH- groups of the pyrrole rings are coordinated to Aps86, and that Arg37 is positioned above each tetrapyrrole ring forming arene-cation interactions. However, of the various complexes it was observed that URO-III binds the strongest to the active site of hUROD in complex **I**, in which, several of carboxylate groups of URO-III hydrogen-bond to the backbone amide groups of the

active site. Moreover, for this complex Ser219 and Arg41 along with Tyr164 and His339 interact with URO-III to properly orient the substrate within the active site.

Using complex I the subsequent QM/MM calculations found that Arg37 most likely acts as the initial acid that protonates C2. Importantly, the rate-limiting step involved this proton transfer from Arg37 to C2 of URO-III, with an activation Gibbs energy of 10.3 kcal mol⁻¹, which is in good agreement with previously obtained experimental values that range from 2.0 to 12.3 kcal mol⁻¹.^{2,3} The process of decarboxylation on the other hand occurred without barrier. Notably, concomitant to the loss of the CO₂ the active site Arg50 donated a proton to the methylene resulting in the formation of a methyl group. The last step, deprotonation of C2 via Arg37 was found to occur with a very low barrier. Overall, the decarboxylation is thermodynamically favourable where the product complex lied considerably lower in energy than the reactant complex.

In Chapter 5 the mechanism by which coral allene oxide synthase (cAOS) catalyses the formation of allene oxide from its hydroperoxy substrate was computationally investigated using a DFT-cluster approach. Specifically, the effect of dispersion interactions and multi-state reactivity along the mechanism, and the effect of the tyrosyl proximal ligand of cAOS compared to the cysteinyl found in pAOS was examined.

In the reactant complex (**RC**) the hydroperoxy substrate was found to form a strong Fe···O interaction in the doublet state while for the quartet state this interaction was much weaker. However, for both states the overall mechanism begins with cleavage of the peroxy O–O bond to give a Cpd II-type intermediate with concomitant formation of an alkoxy radical. This latter species then undergoes a rearrangement resulting in an epoxide with a delocalized allylic radical. The Cpd II intermediate then abstracts a hydrogen atom from the epoxide to give an Fe(III)-bound H₂O and allene oxide. Thus, cAOS utilizes a mechanism that is similar to that for pAOS.

However, the mechanism of cAOS appears to differ from that of pAOS in several key features. The first being that the Cpd II intermediate formed has a markedly different

overall electronic configuration to that calculated for pAOS. This is likely due to both the presence of a histidyl active site residue in cAOS, which is lacking in pAOS, and a ligating tyrosyl residue. Second, the mechanism occurs with considerably higher Gibbs free energies of reaction than that for the analogous pathway in pAOS. However, it is noted that these energetic differences may be partly due to differences in the computational models used to previously⁴ study pAOS versus that used herein.

From the results obtained the inclusion of dispersion effects results in considerable changes to the free energy surface for the mechanism. For instance, at the IEF-PCM-B3LYP*/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory without dispersion effects the homolytic O–O bond cleavage likely occurs with SI from the quartet to doublet surface. However, with dispersion corrections the energy ordering of the various states of RC is altered such that SI is not needed for the initial step as the overall S=1/2 (doublet-state) reactive complex (*i.e.* ${}^2\mathbf{RC}_{\text{dc}}$) is now most favoured. Similarly, the occurrence of SI in product formation is also now unlikely to occur when dispersion effects are included due in part to reordering of the relative free energies of the product complexes; ${}^2\mathbf{RC}_{\text{dc}}$ is now significantly more favoured. The contribution of dispersion effects directly correlates with the changes observed along the mechanisms pathway with regards to the distance between the center of mass of the substrate and heme.

However, it is noted we also investigated the effect of changing functional (*i.e.* at the IEF-PCM-DFT_i/6-311+G(2df,p)//B3LYP/BS1 + ΔG_{Cor} level of theory; DFT_i = M06, B3LYP, B3LYP*, BP86) on the free energy surfaces for both the doublet and quartet states. For both states there is in general a correlation between the amount of %HF contribution in the functional and the reduction in the relative energies of the stationary points along the pathway. That is, the functional with the lowest %HF contribution (*i.e.* BP86) in general gives the lowest relative free energies (with respect to \mathbf{RC}) of the intermediates and transition structures. In contrast, M06, which has the highest %HF contribution (27%) of the four functionals considered herein, in general gives the highest

relative free energies of the intermediates and transition structures with respect to **RC**. In fact, the average reduction in relative free energies obtained upon going from the M06 to BP86 functional is 46.3 and 97.3 kJ mol⁻¹ for the doublet and quartet states, respectively.

In Chapter 6 a range of hybrid, meta and hybrid-meta GGA density functionals were investigated to see which were able to reliably provide geometries, spin densities and relative energies of multi-centered open-shell complexes within an ONIOM QM/MM methodology. Specifically, the BP86, B3LYP[±] (10% HF), B3LYP* (15% HF), B3LYP, M06 and M06-L functionals were used to provide structures, spin densities and relative energies of a multi-centered open-shell mechanistic intermediate complex that occurs during the mechanism of the non-heme iron metalloenzyme 8R-LOX. Specifically, this intermediate complex contains three open shell centers; an Fe(II) in the HS state, a substrate-derived pentadienyl radical and a triplet molecular oxygen (O₂). Overall, the results obtained suggest that for systems with multiple centers having unpaired electrons the B3LYP* appears most well rounded to provide reliable geometries, electronic structures and relative energies.

In Chapter 7 a detailed systematic computational investigation has been performed on the catalytic mechanism of 8R-LOX using an ONIOM QM/MM-based approach. From the results obtained it appears that the contribution of tunnelling to the initial PCET may be similar among all LOXs. In particular, comparison to the results obtained for sLO-1 good agreement in both energetics and geometries was observed.

Following the PCET process the AA[•] complex (i.e. **IC1**), the quartet, sextet and octet were found to all lie close in energy (i.e. within 4 kcal mol⁻¹ of each other). However, while the octet is energetically most preferred the subsequent required C8-OO bond formation cannot occur due to unfavourable electron spin pairing. Instead, spin inversion to the quartet or even sextet state must occur. For the resulting intermediate the peroxy -OO[•] moiety lies antarafacial to the Fe center. Thus, the rotation of the substrate about both its C9-C8 and C8-OO bonds is required to bring the -OO[•] moiety suprafacial to the

Fe centre. Throughout this process the quartet and sextet state stationary points (i.e., minima and transition structures) were calculated to lie within 1.6 kcal mol⁻¹ of each other with the highest rotational step barrier (⁴TS4) being only 8.8 kcal mol⁻¹. Hence, the active site appears not to significantly inhibit rotation.

The last step in the overall mechanism is formation of the peroxide product AA–OOH. This process occurred via a PCET process in which the Fe(II) is oxidized while concomitantly a proton of the Fe-bound H₂O transfers onto the peroxy moiety of AA–OO•. On the quartet PES this step occurs with a barrier of 13.0 kcal mol⁻¹. On the sextet PES this process occurs with a barrier significantly lower in energy at just 2.3 kcal mol⁻¹. In agreement with experimental observations the resulting product ⁶PC is preferred, being markedly lower in energy than either ⁴PC or any of the initial fully bound active site complexes.

In Chapter 8 small model iron-oxygen complexes were investigated to determine their ability to oxidize the imidazole ring of histidine. Of those considered, only the ferrous-peroxy-sulfur (i.e. Fe-O-O-S) complexes were inherently capable of performing this oxidation. Notably, these Fe-O-O-S species were found to be more oxidizing than the oxo-ferryl species considered herein. Importantly, the latter are considered to be the ultimate bio-oxidant. Thus, from the results the need to form the sulfoxide is rather a consequence in the formation of a more powerful oxidant in the model Fe-OO-S complexes providing insight into the puzzling need for sulfoxidation.

In Chapter 9 a range of density functional theory (DFT) methods were examined to see which provided reliable optimized structures of Cu(I) and Cu(II) complexes with OSH and ESH. More specifically, the GGA functional BP86,^{5,6} hybrid GGA functional B3LYP,⁷⁻¹² meta-GGA functional M06L^{13,14} and meta-hybrid GGA functional M06,^{13,14} in combination with 6-31G(d), 6-31G(d,p), 6-311G(d,p), 6-311G(2d,p), 6-311G(2df,p), 6-311+G(2df,p) and 6-311G++(3df,3dp) were used.

For all complexes studied, the largest impact of change in the basis set occurred in optimized lengths of both the (Cu(II)···N) and (Cu(II)···S) bonds. Notably, only when the 6-311+G(2df,p) basis set was used was convergence in their lengths observed. Furthermore, in the case of the Cu(II) complexes only with the 6-311+G(2df,p) basis set or larger was the correct square planar geometry predicted. Comparison of optimized structures with the experimental data available for [Cu(ESH)₃]⁺ suggests that the BP86/6-311+G(2df,p) level of theory gives closest agreement with experiment.

In addition to investigating the geometries both the relative energies of the various complexes as well as the potentials associated with reduction of Cu(II)···ESH/OSH complexes to the corresponding Cu(I)-containing complexes investigated. The calculated reduction potentials of the Cu(II)/Cu(I) couple with formation of the respective OSH/ESH disulfides (i.e., OSSO and ESSE) suggest that such processes are thermodynamically favourable in the presence of either OSH and ESH. Importantly, the increased reduction potential for reduction of Cu(II) to Cu(I) suggest that the oxidation of Cu(I) to Cu(II) is less likely to occur. As a result, the redox cycling of Cu(I)/Cu(II) is inhibited. Thus, OSH and ESH at least in part inhibit the oxidative damaging abilities of copper ions in biochemical systems by altering their reduction potentials and inhibiting the recycling of Cu(I) to Cu(II).

In Chapter 10 a series of systematically higher-level ONIOM-based computational methods were used to examine the catalytic mechanism of saccharopine reductase as well as the effects of electron correlation and the anisotropic polar protein environment on the mechanism. Based on the extensive and high-level computational models used herein, the present results suggest that saccharopine reductase catalyses the overall reaction by binding the three required reactant molecules glutamate, α -aminoadipate- δ -semialdehyde and NADPH in an orientation and polar environment conducive to reaction. Importantly, the enzymes overall mechanism does not require the direct involvement of active-site residues in the required proton transfer processes. Specifically, the protonated

α -amine (AASA-NH₃⁺) of the cosubstrate α -aminoadipate- δ -semialdehyde (AASA) was found to act as an acid protonating the forming oxyanion centre during nucleophilic attack of the Glu-NH₂ group at the R-group aldehyde carbon of AASA. In addition, the glutamate's carboxylate (Glu-COO⁻) assists the proton transfer from the bridging -^{Glu}NH₂⁺ moiety in the formed initial carbinolamine intermediate (**II**) to the nearby carbinolamine hydroxyl oxygen.

At the highest level of theory used (i.e. ONIOM(MP2/6-311G(d,p)//HF/6-31G(d):AMBER94)-EE with inclusion of Gibb's corrections) the catalytic mechanism obtained was found to be both kinetically feasible where the overall rate-limiting step was the loss of water to give the Schiff base intermediate **I5** which occurs via **TS6** at a cost of 74.1 kJ mol⁻¹ with respect to **RC**. A subsequent barrierless hydride transfer reduces **I5** to the final saccharopine product. Furthermore, the reaction was found to be thermodynamically favourable where the overall Gibbs free energy change was -18.4 kJ mol⁻¹.

While the material in this thesis does not embody a complete picture of redox and metal ion biochemistry, the results we have obtained do provide valuable insight into such chemistry. Importantly, such insight adds to the growing wealth of knowledge and ultimately provides a more complete understanding into the features of redox and metal ion biochemistry.

Every day new metallo-enzymes are being discovered with each having the potential to provide new insights into redox and metal ion biochemistry. Computational chemistry is an important tool that can be used to investigate these new systems. Indeed, the results generated from such investigations such as those presented herein provide direction into the design of new biomimetic catalysts. It is hoped that from nature inspiration selective and highly efficient biomimetic catalysts can be formed. Such catalysts may then be used in industry to enhance productivity and reduce costs while at the same time help prevent the production of unwanted pollution. Furthermore, as new biochemical systems are

explored, new opportunities to design better and more efficient computational methods emerge. Lastly, the best inhibitors are generally held to be transition state analogues; thus, by elucidating an enzymes mechanism via the use of computational methods it allows for new pharmaceuticals to be developed.

11.2 References

- (1) Huang, W. J.; Bushnell, E. A. C.; Francklyn, C. S.; Gault, J. W. *J. Phys. Chem. A* **2011**, *115*, 13050-13060.
- (2) Chaufan, G.; de Molina, M. D. R.; de Viale, L. C. S. *Int. J. Biochem. Cell Biol.* **2001**, *33*, 621-630.
- (3) Juárez, A. B., Aldonatti, C., Vigna, M. S., Ríos de Molina, M. C. *Can. J. Microbiol.* **2007**, *53*, 303-312.
- (4) Cho, K. B.; Lai, W. Z.; Hamberg, M.; Raman, C. S.; Shaik, S. *Arch. Biochem. Biophys.* **2011**, *507*, 14-25.
- (5) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406-7406.
- (6) Perdew, J. P.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800-8802.
- (7) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
- (9) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403-412.
- (10) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
- (11) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623-11627.
- (12) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.
- (13) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157-167.
- (14) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

Vita Auctoris

Name: Eric André Charles Bushnell
Place of Birth: Windsor, Ontario, Canada
Year of Birth: 1977

Education

Ph.D Chemistry September 2008 – June 2013
University of Windsor Windsor, Ontario

B.Sc Chemistry [H] September 2004 – May 2008
University of Windsor Windsor, Ontario

B.Sc Computer Science [G] September 1996 - June 2002
University of Windsor Windsor, Ontario

Scholarships

1. Killam Postdoctoral Fellowship (KPDF) awarded in February 2013.
2. NSERC Postgraduate Three Year Scholarship (PGS-3D) awarded in March 2010.
3. Ontario Graduate Scholarship (OGS) awarded in March 2010.
4. Alexander Graham Bell Canada Graduate Scholarships (CGS-M) awarded in March 2009.

5. Ontario Graduate Scholarship (OGS) awarded in March 2009.
-

Awards

5. HyperCube award. Awarded in May 2008.
-

Publications

Total papers (14): As first author (8 + 2 reviews/chapters), as a co-author (4).

14. **Bushnell, E.A.C.**, Gherib, R., Gault, J.W. Insights into the Catalytic Mechanism of Allene Oxide Synthase: A Dispersion Corrected Density Functional Theory Study. *Journal of Physical Chemistry B* (Accepted, May 15th 2013).
13. De Luna, P.*, **Bushnell, E.A.C.***, Gault, J.W. A Density Functional Theory Investigation into the Binding of the Antioxidants Ergothioneine and Ovothiol to Copper. *Journal of Physical Chemistry A* **2013**, *117*(19), 4057-4065. *Equal Contributions.
12. **Bushnell, E.A.C.**, Jamil, R., Gault, J.W. Insights into the Chemistry of Lipoygenases (LOXs): A Computational Investigation Into The Catalytic Mechanism of 8R-LOX. *Journal of Biological Inorganic Chemistry*. **2013**, *18*, 343-355.
11. **Bushnell, E.A.C.**, Gault, J.W. An assessment of pure, hybrid, meta, and hybrid-meta GGA density functional theory methods for open-shell systems: The case of the nonheme iron enzyme 8R-LOX. *Journal of Computational Chemistry* **2013**, *24*, 141-148.
10. **Bushnell, E.A.C.**, Fortowski, G.B., Gault, J.W. Iron-Oxo Species and the Oxidation of Imidazole: Insights into the Mechanism of OvoA and EgtB. *Inorganic Chemistry* **2012**, *51*, 13351-13356.

9. **Bushnell, E.A.C.**, Huang, W.J., Llano, J., Gauld, J.W. Molecular Dynamics Investigation into Substrate Binding and Identity of the Catalytic Base in the Mechanism of Threonyl-tRNA Synthetase. *Journal of Physical Chemistry B* **2012**, *116*, 5205–5212.
8. **Bushnell, E.A.C.**, Huang, W.J., Gauld, J.W. Applications of Potential Energy Surfaces in the Study of Enzymatic Reactions. *Advances in Physical Chemistry* **2012**, *2012*: 15 pages.
7. Ion, B.F., **Bushnell, E.A.C.**, De Luna P., Gauld, J.W. An MD and QM/MM Study on Ornithine Cyclodeaminase (OCD): A Tale of Two Iminiums. *International Journal of Molecular Sciences* (Invited to special issue: Enzyme Optimization and Immobilization). **2012**, *13*, 12994-13011.
6. Almasi, J.N.*, **Bushnell, E.A.C.***, Gauld, J.W. A QM/MM–Based Computational Investigation on the Catalytic Mechanism of Saccharopine Reductase. *Molecules* (Invited to special issue: Enzyme-Catalyzed Reactions). **2011**, *16*, 8569-8589. * Equal Contributions.
5. **Bushnell, E.A.C.**, Erdtman, E., Llano, J., Eriksson, L.A., Gauld, J.W. The first branching point in porphyrin biosynthesis: A systematic docking, molecular dynamics and quantum mechanical/molecular mechanical study of substrate binding and mechanism of uroporphyrinogen-III decarboxylase. *Journal of Computational Chemistry* **2011**, *32*, 822–834.
4. **Bushnell, E.A.C.**, Llano, J., Eriksson, L.A., Gauld, J.W. Mechanisms of Mutagenic DNA Nucleobase Damages and Their Chemical and Enzymatic Repairs Investigated by Quantum Chemical Methods. In: *Selected Topics in DNA Repair*. Clark C. Chen (Ed.), **2011**, 389-414, InTech, Croatia.
3. Erdtman, E., **Bushnell, E.A.C.**, Gauld, J.W., Eriksson, L.A. Computational studies on Schiff-base formation: Implications for the catalytic mechanism of

- porphobilinogen synthase. *Computational and Theoretical Chemistry* **2011**, *963*, 479-489.
2. Huang, W.J., **Bushnell, E.A.C.**, Francklyn, C.S., Gaud, J.W. The α -Amino Group of the Threonine Substrate as The General Base During tRNA Aminoacylation: A New Version of Substrate-Assisted Catalysis Predicted by Hybrid DFT. *Journal of Physical Chemistry A* (Invited to Richard F. W. Bader Festschrift special issue). **2011**, *115*: 13050–13060.
 1. Erdtman, E., **Bushnell, E.A.C.**, Gaud, J.W., Eriksson, L.A. Computational Insights into the Mechanism of Porphobilinogen Synthase. *Journal of Physical Chemistry B* **2010**, *114*: 16860–16870.
-

Conference Presentations

9. **Bushnell E.A.C.,*** Fortowsky G.B., Gaud J.W. (2012) The Oxidative Power of Key Intermediates in the Formation of Ovoidiol. 25th CSTCC (International Conference) in Guelph, ON, Can. Poster presentation. (PhD work)
8. **Bushnell, E.A.C.,*** Gaud, J.W., (2012) Effects of Computational Corrections on the Mechanism of Allene Oxide Synthase, CSC (International Conference) in Calgary, AB, Can. Oral presentation. (PhD work)
7. **Bushnell, E.A.C.,*** Gaud, J.W., (2011) Insights into LOX chemistry: The catalytic mechanism of 8R-LOX, CSC (International Conference) in Montreal, AB, Can. Oral presentation. (PhD work)
6. **Bushnell, E.A.C.,*** Gaud, J.W., (2011) Insights into LOX chemistry: The catalytic mechanism of 8R-LOX, ICBIC (International Conference) in Vancouver, BC, Can. Poster presentation. (PhD work)

5. **Bushnell, E.A.C.,*** Gauld, J.W., (2010) QM/MM Investigation of the Mechanism of 8R-Lipoxygenase, CSC (International Conference) in Toronto, AB, Can. Poster presentation. (PhD work)
4. **Bushnell, E.A.C.,*** Llano, J., Gauld, J.W., (2010) An Investigation into the Binding of Arachidonic Acid Within the Active Site of 8R-LOX, CSC (International Conference) in Calgary, AB, Can. Poster presentation. (PhD work)
3. **Bushnell, E.A.C.,*** Erdtman, E., Llano, J., Eriksson, L.A., Gauld, J.W., (2009) A Computational Investigation on the Mechanism of UROD, CSC (International Conference) in Calgary, AB, Can. Poster presentation. (PhD work)
2. **Bushnell, E.A.C.,*** Erdtman, E., Llano, J., Eriksson, L.A., Gauld, J.W., (2009) A Computational Investigation on the Mechanism of UROD, CCCC 7 (International Conference) in Halifax, NS, Can. Poster presentation. (PhD work)
1. **Bushnell, E.A.C.,*** Gauld, J.W., and Thadani, A.N., (2008) An Investigation of the Reactivity of Aldimines with Boron Derivatives, CSC (International Conference) in Edmonton, AB, Can. Poster presentation. (BSc work)