

2014

Computational Insights into Sulfur Redox Chemistry in Enzymology

Hisham Mohammed Mohammed Dokainish
University of Windsor

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

 Part of the [Chemistry Commons](#)

Recommended Citation

Dokainish, Hisham Mohammed Mohammed, "Computational Insights into Sulfur Redox Chemistry in Enzymology" (2014). *Electronic Theses and Dissertations*. Paper 5206.

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 Sulfur Redox Chemistry in Enzymology

By

Hisham M. Dokainish

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

2014

© 2014 Hisham M. Dokainish

Computational Insights Into Sulfur Redox Chemistry In Enzymology

by

Hisham Dokainish

APPROVED BY:

J. Honek, External Examiner
University of Waterloo

E. Maeva
Department of Physics

B. Mutus
Department of Chemistry & Biochemistry

J. Wang
Department of Chemistry & Biochemistry

J. Gauld, Advisor
Department of Chemistry & Biochemistry

September 12, 2014

Declaration of Co-Authorship

I. Co-Authorship Declaration

I hereby declare that this Dissertation incorporates material that is result of joint research, as follows:

Chapter 7: was done in collaboration with Mr. Daniel J. Simard and under the supervision of Prof. James W. Gauld.

Chapter 9: was done in collaboration with Mr. Bogdan F. Ion and 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 Dissertation, and have obtained written permission from each of the co-author(s) to include the above material(s) in my Dissertation.

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

II. Declaration of Previous Publication

This Dissertation includes 3 original papers in whole or in part that have been previously published/submitted for publication in peer reviewed journals, as follows:

	Citation	Status
Chapter 2 and 3	Gherib, R.; Dokainish, H. M.; Gauld, J. W. <i>Int. J. Mol. Sci.</i> 2014 , <i>15</i> , 401-422.	Published
Chapter 3	Dokainish, H. M.; Gauld, J. W. <i>Biochemistry</i> 2013 , <i>52</i> , 1814-1827.	Published
Chapter 9	Dokainish, H. M.; Ion, B, F.; Gauld, J. W. <i>Phys. Chem.</i>	Published

	<i>Chem. Phys.</i> , 2014 , <i>16</i> (24), 12462-12474.	
--	---	--

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

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

Abstract

In biochemistry, sulfur-containing biomolecules enrich the chemical diversity in cells. This occurs via their participation in several reactions including disulfide formation, metal-binding and redox catalysis. Since sulfur occurs in various oxidation states, it exhibits interesting chemistry and reactivity. In this Dissertation computational modeling techniques have been used to investigate several aspects of sulfur's unique chemistry.

Chapter 1 provides a brief introduction summarizing the importance of sulfur in biochemistry. In particular it discuss the chemical diversity and redox chemistry of sulfur in biology. Chapter 2 provides an overview of the computational methods commonly used in studying enzymatic reactions.

Chapter 3 discusses the reduction mechanism of S-methionine sulfoxide (Met-O) to methionine via a ubiquitous antioxidant enzyme known as methionine sulfoxide reductase A (MsrA). We used a wide range of computational methods including docking, molecular dynamics simulations and QM/MM calculations. Our results have proposed new roles for active site residues such as Asp87. In addition, the activation mechanism of the catalytic cysteine was revealed. Furthermore, the effect of active site tyrosyl residues after mutation were also considered and found in agreement with experimental results. The formation and reduction mechanism of sulfenic acid intermediate was computationally elucidated. It showed that sulfenic acid is formed via the activation of a water molecule in accordance with a new experimental study. Notably, our results support that sulfenic acid is readily reduced upon proper positioning of the second cysteinyl residue solving the controversy in experimental studies regarding this step.

In Chapter 4, we discuss our computational investigation on the last step in the reductase step of the catalytic mechanism of methionine sulfoxide reductase B (MsrB), the formation of disulfide directly or via sulfenic acid intermediate. Our previous work on MsrB showed that, unlike MsrA, sulfenic acid is not formed in the catalytic mechanism. Due to a disagreement with experimentalists, we considered investigating the effect of level of theory, the effect of model choice QM-cluster vs. QM/MM and the effect of choosing a starting structure from X-ray vs. MD. Our QM/MM results based on MD show that sulfenic acid occurs in the mechanism. We also considered running MD simulations throughout the reaction coordinates to test the effect of substrate binding on the distance between the two catalytic cysteinyl residues.

In Chapter 5, we discuss our synergistic use of docking, MD simulations and virtual screening to investigate the mechanism of MsrA activation via small ligands. The possibility of their direct molecular interaction with MsrA was tested using docking techniques identifying a novel allosteric site. Furthermore, the effect of ligand-MsrA binding was elucidated using MD simulations. Our MD results suggest that their binding facilitates the unfolding of two-stranded antiparallel β -sheet revealing the active site for subsequent reduction by thioredoxin (Trx). Finally, virtual screening was employed to identify other potential ligands to act as MsrA activators.

Chapter 6 discusses our computational investigation on the second step (formation sulfenyl-amide intermediate from sulfenic acid) in protein tyrosine phosphatase 1B (PTP1B). We propose an alternative pathway to the previously suggested high-energy barrier mechanism proposing new roles for active site residues. Furthermore, the role of non-covalent interactions was highlighted and supported using QTAIM and NBO analysis.

In Chapter 7 the reduction mechanism of peroxide via the archaeal peroxiredoxin (ApTPx) antioxidant enzyme was discussed. In addition, the formation of the previously proposed hypervalent intermediate from sulfenic acid was tested. The nature of

interactions in this highly coordinated species was elucidated using QTAIM and NBO analysis. Furthermore, the transferability of this unique sulfenic acid protection mechanism to other enzymes was also considered. Moreover, the overoxidation mechanism of sulfenic acid to sulfinic acid was examined for the first time.

Chapter 8 describes our MD and ONIOM investigation on the unique reduction mechanism of the 2-Cys peroxiredoxins (Prxs) sulfinic acid via sulfiredoxin (Srx). The previously suggested sulfinic acid phosphoryl ester and thiosulfinate intermediate were characterized in the active site. Understanding the generally irreversible sulfinic acid intermediate reduction mechanism can potentially enable the development of new antioxidant drugs.

In Chapter 9 we discuss our investigation of the cis-trans isomerization mechanism of maleate to fumarate via maleate isomerase (MI) enzyme. Small DFT models, QM-cluster approach, QM/MM and MD were employed to investigate previously proposed mechanisms. Previous X-ray and mass spectrometry studies have suggested a novel mechanism in the racemase family in which an active site cysteinyl residue covalently bonds to the substrate forming either a succinyl-cysteinyl or enediolate intermediate. Our calculations show that the enediolate intermediate is unstable and succinyl cysteine intermediate occurs in the mechanism. Also, proton affinities calculations of active site cysteinyl residues and substrate in protein environment indicated different roles for active site cysteinyl residues from the proposed ones. Notably, our QM/MM calculations of both pathways (experimental and proton affinities-based) have confirmed our conclusion. These results highlights that the mutation of active site residue might mislead the conclusion of the enzymatic mechanism.

Dedication

I dedicate this work to my loving family. Particularly, I dedicate this to my loving wife Marwa for her endless support and devotion. Also, I want to dedicate this to my uncle Dr. Hossam Kishawy for his guidance and assistance since the beginning of my academic journey.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Dr. James Gauld. His sound advice and careful guidance were invaluable over the course of my PhD study. His support during my studies as well as in finding a future job is genuinely appreciated. Indeed, working with Dr. Gauld helped me on different level to grow as scientist and as independent researcher.

In addition, I wish to thank my committee members, Dr. Maeva, Dr. Mutus and Dr. Wang who were more than generous with their expertise and precious time during the last four years. Also I would like to thank them for taking the time to read my Dissertation. Furthermore I would like to thank Dr. Mutus for writing me several reference letters helping me to find a future job. Finally, I would like to thank Dr. Honek for agreeing on being my external examiner.

Also, I would like to acknowledge the wonderful people I had the chance to work with in the Gauld group, including past and current members. In particular I want to thank Dr. Wenjuan Huang and Dr. Eric Bushnell for their help especially at the beginning of my studies. Similarly, I would like to thank Bogdan Ion, Daniel Simard and Rami Gherib for collaborating with me to study some of the presented work in this Dissertation. Also, I would like to thank Grant Fortowsky and Mohammed Aboelnga for the entertaining environment as well as many undergraduate students in the group. Without their valuable discussions and suggestions, I would not be able to complete many of my studies. Finally, a special thanks to Dan Simard for proof reading many of my work and job applications.

Table of Contents

Declaration of Co-authorship.....	III
Abstract.....	V
Dedication.....	VIII
Acknowledgments.....	IX
List of Figures.....	XV
List of Schemes.....	XX
List of Appendices.....	XXII
List of Abbreviations.....	XXIV
Chapter 1: Introduction.....	1
1.1 Introduction.....	2
1.2 References.....	6
Chapter 2: Overview of Computational Methods.....	8
2.1 Introduction.....	9
2.2 Computational Enzymology.....	9
2.3 Protein Dynamics and MD Simulations.....	10
2.4 Molecular Mechanical Force Fields.....	12
2.5 Quantum Mechanical (QM) Methods.....	14
2.6 Choosing a Starting Structure.....	16
2.7 Quantum Mechanical (QM)-Cluster Approach.....	17
2.8 Quantum Mechanical/Molecular Mechanical (QM/MM) Approach.....	19
2.8.1 QM/MM Additive Scheme.....	20

2.8.2 QM/MM Subtractive Scheme	21
2.8.3 QM/MM Electrostatic Interactions	22
2.8.4 QM/MM Boundary	24
2.8.5 Setting-up QM/MM Model	26
2.9 References	26

Chapter 3: An MD and QM/MM study on the Catalytic Reductase Mechanism of Methionine Sulfoxide Reductase A (MsrA): Formation and Reduction of a Sulfenic Acid.....

31	
3.1 Introduction.....	32
3.2 Computational Methods.....	36
3.3 Results and Discussion	40
3.3.1 Substrate Binding and Activation of the Catalytic Cysteine.....	40
3.3.2 Reduction of the Substrate with Formation of a Sulfenic Acid Intermediate	44
3.3.3 Reduction of the Sulfenic Acid with Formation of an Intramolecular Disulfide Bond	52
3.4 Conclusions.....	59
3.5 References.....	60

Chapter 4: Computational Approach Choice in Modeling Flexible Enzyme Active Sites.....

66	
4.1 Introduction.....	67
4.2 Computational Methods.....	70
4.2.1 DFT Calculations	70
4.2.2 QM/MM Models.....	71
4.2.3 Molecular Dynamics simulations.....	72

4.3 Results and Discussions	74
4.3.1 QM-Cluster Approach.....	74
4.3.2 QM/MM Results	77
4.3.3 Molecular Dynamics Simulations Results	81
4.4 Conclusions.....	86
4.5 References.....	88
Chapter 5: Small Molecules Activating an Antioxidant Enzyme.....	93
5.1 Introduction.....	94
5.2 Computational Methods.....	96
5.3 Results and Discussion	97
5.4 Conclusions.....	103
5.5 References.....	103
Chapter 6: Formation of Stable Iminol Intermediate in the Redox Regulation	
Mechanism of PTP1B	106
6.1 Introduction.....	107
6.2 Computational Methods.....	109
6.2.1 Protein Preparation and Molecular Dynamics Simulations	109
6.2.2 QM/MM Models and Calculations	112
6.3 Results and Discussions	113
6.3.1 Insights into Reactive Complex Structures.....	113
6.3.2 Formation of Sulfenyl-Amide via Direct Mechanism	114
6.3.3 Formation of Sulfenyl-Amide via Stepwise Mechanism.....	115
6.3.4 Formation of Sulfenyl-Amide using Neutral His214.....	119
6.4 Conclusions.....	120
6.5 References.....	122

Chapter 7: Pseudo-hypervalent Sulfur Intermediate as a Protective Mechanism in Peroxiredoxin Enzymes	126
7.1 Introduction.....	127
7.2 Computational Methods.....	130
7.2.1 Structural Preparation and Molecular Dynamic Simulations	130
7.2.2 QM/MM Models and Calculations	132
7.3 Results and Discussions	133
7.3.1 Sulfurane Intermediate Formation in ApTPx.....	133
7.3.2 Sulfenic Acid Oxidation Mechanism in ApTPx.	139
7.3.3 Sulfurane Formation in Human hORF6.....	140
7.4 Conclusions.....	143
7.5 References.....	145

Chapter 8: QM/MM Investigation of the Reduction Mechanism of Cysteine Sulfinic Acid in Peroxiredoxin via Sulfiredoxin	150
8.1 Introduction.....	151
8.2 Computational Methods.....	155
8.2.1 Protein Model Preparation	155
8.2.2 QM/MM Models and Calculations	157
8.3 Results and Discussions	159
8.3.1 Formation of Sulfinic Acid Phosphoryl Ester.....	159
8.3.2 Formation of Thiosulfinate Intermediate.	162
8.4 Conclusions.....	164
8.5 References.....	166

Chapter 9: Computational Investigations on the Catalytic Mechanism of Maleate Isomerase: the Role of the Active Site Cysteine Residues	170
9.1 Introduction.....	171
9.2 Computational Methods.....	175
9.2.1 DFT-Small Model Studies	175
9.2.2 QM-Cluster Studies.....	176
9.2.3 QM/MM Studies	177
9.3 Results and Discussion	179
9.3.1 DFT-Small Model Studies on Isomerization	179
9.3.2 The Ionization States of Cys76, Cys194 and the Substrate.	182
9.3.3 Cys76-Pathway: Mechanism with Cys76 as the Nucleophile	184
9.3.4 Cys194-Pathway: Mechanism with Cys194 as the Nucleophile. ...	192
9.4 Conclusions.....	195
9.5 References.....	199
Chapter 10: Conclusions	204
Appendix C	211
Vita Auctoris	217

List of Figures

Figure 1.1. Schematic illustration of some thiol-derived species obtained via redox in biochemical systems..	4
Figure 1.2. Schematic illustration of sulfenic acid proposed reactions in biochemistry.	5
Figure 2.1. Graph indicating the results of a search on ISI Web of Science using the criteria: Topic = computational and Topic = enzyme sorted by publication year.	11
Figure 2.2. Schematic representation of QM/MM partitioning model.	20
Figure 3.1. The QM/MM models used to examine the catalytic mechanism of MsrA for (A) formation of an enzyme-derived sulfenic acid intermediate from the initial enzyme-substrate complex and (B) reaction of the enzyme-derived sulfenic acid intermediate to give the final products. The inner circles represent the QM layer while the outer represent the MM layer. Color key for MM residues: included in their entirety (black); R-group replaced by hydrogen (red); R-group included in the MM layer while backbone is in the QM layer (blue).	40
Figure 3.2. Schematic illustration of the optimized structures of the pre-reactive complex (PRC), activated reactive complex (RC) and the transition structure (TS_{Act}) for their interconversion. For simplicity, only selected residues, functional groups and bond lengths (Angstroms) are shown.	42
Figure 3.3. Free energy surface (kJ mol^{-1}) obtained (see Computational Methods) for activation of the catalytic cysteine (Cys13) by proton transfer via a bridging H_2O to the active site glutamate (Glu52).	43

Figure 3.4. Free energy surface (kJ mol^{-1}) obtained (see Computational Methods) for formation of a sulfenic acid intermediate (Cys13SOH) involving the proton donors (A) Glu52COOH and Tyr92OH and (B) Glu52COOH and Tyr44OH.	44
Figure 3.5. Schematic illustration of the optimized structures with selected bond lengths (in Angstroms) of the minima and transition structures obtained for MsrA catalysed reduction of Met-O to Met with concomitant formation of an enzyme-derived sulfenic acid intermediate.	46
Figure 3.6. Free energy surface (kJ mol^{-1}) obtained (see Computational Methods) for the reduction of sulfenic acid-containing complex to give an intramolecular disulfide bond with formation of H_2O	55
Figure 3.7. Schematic illustration of the optimized structures for sulfenic acid reduction and formation of disulfide bond.	57
Figure 3.8. Schematic illustration of our computationally suggested mechanism.	58
Figure 4.1. The second QM/MM model for MsrB sulfonium cation intermediate optimized structure. The QM layer atoms are highlighted in sticks to the right of the figure.	73
Figure 4.2. Figure 4.2. Potential energy surface (PES) for direct disulfide bond formation pathway from optimization and single point calculations at several level of theory. Relative energies in kJ mol^{-1}	75
Figure 4.3. Potential energy surface (PES) for sulfenic acid formation pathway from optimization and single point calculations at several level of theory. Relative energies in kJ mol^{-1}	76
Figure 4.4. Figure 4.4. Optimized structures for the direct disulfide bond formation from the QM/MM calculations in which the sulfonium intermediate were obtained from MsrB crystal structure.	77

Figure 4.5. Figure 4.5. Optimized structures for the sulfenic acid formation pathway from the QM/MM calculations in which 3large were obtained from sulfonium intermediate MD simulation.	80
Figure 4.6. RMSD's of the QM layer vs time for the four models used to represent the catalytic cycle (A) Apoenzyme before substrate binding, (B) substrate bound active site, (C) sulfonium cation intermediate, and (D) sulfenic acid intermediate.....	81
Figure 4.7. Distance change between C α of Cys495 and Cys440 vs time the four models used to represent the catalytic cycle (A) Apoenzyme before substrate binding, (B) substrate bound active site, (C) sulfonium cation intermediate, and (D) sulfenic acid intermediate. Distances are represented in Angstrom.....	83
Figure 4.8. Superposition of the 4 MD's active site average structures in apoenzyme, Michaelis complex, sulfonium and sulfenic acid intermediates.	85
Figure 5.1. Top) docking predicted binding mode of Pergolide, bottom) schematic representation of Pergolide interactions in the MsrA allosteric site.	98
Figure 5.2. Molecular surface representation of the disulfide surrounding residues extracted from MD simulations. The surface exposure of the disulfide bond is represented by yellow colour.	99
Figure 5.3. Time evaluation of root-mean-square deviation (Rmsd) of Gln122-Thr132 segment and Gly196-Gly199 residues with and without Pergolide binding.	101
Figure 5.4. Superimposition of the MsrAOx NMR solution structure with that of MsrAOx MD generated structure upon Pergolide binding and removal.	102
Figure 6.1. Illustration of the QM/MM model used (Model I) in this study obtained from the average structure. The QM layer atoms are highlighted and shown in sticks representation.	111
Figure 6.2. Plot of the hydrogen-bond distance between His214 and Cys215 with respect to time in the MD simulation.	114

Figure 6.3. Potential energy surface obtained for the formation of sulfenyl-amide from sulfenic acid via iminol intermediate, see computational methods.	117
Figure 6.4. Stationary points obtained for model I for the sulfenyl-amide formation mechanism from sulfenic acid. All atoms in the high QM layer are included in the representation. However, only the highlighted residues are the ones involved in the reaction.	118
Figure 7.1. Illustration of the ApTPx QM/MM model used obtained from MD average structure. The QM layer atoms are highlighted and shown in sticks representation. ...	132
Figure 7.2. Illustration of the QM/MM optimized reactive complex of ApTPx. Only the QM layer is shown.	134
Figure 7.3. Potential energy surface obtained for the formation of pseudo-sulfurane intermediate in ApTPx.	136
Figure 7.4. Stationary points obtained for the pseud-sulfurane formation mechanism. Shown atoms represent the QM layer. However, only the highlighted are directly involved in the reaction.	138
Figure 7.5. Illustration of the optimized stationary points for sulfenic acid overoxidation to sulfinic acid including their relative energies. Shown atoms represent the QM layer. However, only the highlighted are the ones involved in the reaction.	140
Figure 7.6. Illustration of the optimized stationary points for pseudo-sulfurane formation mechanism in hORF6, including their relative energies. Shown atoms represent the QM layer. However, only the highlighted are the ones involved in the reaction.	141
Figure 8.1. The QM/MM model for the Srx-Prx Michaelis complex used to investigate the first step of the mechanism. The QM layer atoms are highlighted in sticks to the right of the figure.	156
Figure 8.2. The QM layer residues included in model II to study the second step of the reduction mechanism converting sulfinic acid phosphoryl ester into thiosulfinate.	158

Figure 8.3. Potential energy surface obtained for the formation of sulfinic acid phosphoryl intermediate and ADP from Prx-sulfinic acid and ATP via a concerted mechanism.	159
Figure 8.4. Schematic representation of the optimized stationary points for the overall reaction, from sulfinic acid to thiosulfinate product formation. All atoms in the QM layer are included however, we only highlighted the atoms included in the reaction as balls and sticks.	161
Figure 8.5. Potential energy surface obtained for the formation of thiosulfinate product complex via the interaction of Prx-sulfinic acid phosphoryl ester and Srx-Cys99 via a concerted mechanism.	163
Figure 9.1. The (a) HOMO and (b) LUMO of dianionic maleate in the gas-phase ($\epsilon = 1$).	181
Figure 9.2. The calculated proton affinities (see Computational Methods) of C76S ⁻ and C194S ⁻ before and after substrate binding, and the PA of the substrate in the active site.	182
Figure 9.3. PES obtained using a QM-cluster approach (see Computational Methods) of the catalytic mechanism of MI in which Cys76 acts as a nucleophile.	184
Figure 9.4. PES obtained using an ONIOM(QM/MM) approach (see Computational Methods) of the catalytic mechanism of MI in which Cys76 acts as a nucleophile.	188
Figure 9.5. PES obtained using an ONIOM(QM/MM) approach (see Computational Methods) of the catalytic mechanism of MI in which Cys194 acts as a nucleophile. ..	192

List of Schemes

Scheme 3.1. Proposed reductase mechanisms for MsrA via (A) a sulfenic acid intermediate, or (B) via direct attack.	34
Scheme 3.2. Proposed ¹ catalytic reductase mechanism for MsrA involving formation of a sulfenic acid intermediate via attack of a solvent water at the catalytic cysteine's sulfur centre.....	35
Scheme 6.1. The proposed dephosphorylation mechanism of phosphotyrosine in PTP as well as the oxidative regulation mechanism in PTP1B via formation of sulfenyl-amide and the subsequent reactivation.	110
Scheme 8.1. The proposed regulation mechanism of typical 2-Cys Prx via Srx upon the hyperoxidation of Prx sulfenic acid by hydrogen peroxide.	154
Scheme 9.1. Illustration of the overall isomerization reaction catalyzed by Maleate cis-trans Isomerase (MI).	173
Scheme 9.2. Proposed mechanisms for the maleate/fumarate isomerization reaction catalyzed by MI via an (A) succinyl-Cys or (B) enediolate intermediate.	174
Scheme 9.3. The active site-bound substrate chemical model of NfMI used for the QM-cluster studies (atoms fixed at their MM optimized coordinates are highlighted in red and atom numbering used for Maleate carbons is also shown).	177
Scheme 9.4. Schematic illustration of the chemical model, derived from the X-ray crystal structure PDB ID: 2XED, used in the ONIOM QM/MM calculations. The inner circle represents the high (QM)-layer while the outer represents the low (MM)-layer.	178
Scheme 9.5. Schematic illustration of the gas-phase optimized structures of the 3 possible substrate states ($\angle C1-C2-C3-C4 = 0.0^\circ$ in each).	180

Scheme 9.6. Schematic illustration of the optimized structures obtained using a QM-cluster approach (see Computational Methods) for the mechanism in which Cys76 acts as nucleophile.....	185
Scheme 9.7. Schematic illustration of optimized structures obtained using an ONIOM(QM/MM) approach (see Computational Methods) for the mechanism in which Cys76 acts as nucleophile.	190
Scheme 9.8. Schematic illustration of optimized structures obtained using an ONIOM(QM/MM) approach (see Computational Methods) for the mechanism in which Cys194 acts as nucleophile.	197

List of Appendixes

For access to appendixes A and B contact Dr. James Gauld at gauld@uwindsor.ca.

Figure A1. Overlay of the QM residues of the ten average structures obtained from cluster analysis of QM-RMSD of the reactive complex active site.

Figure A2. Overlay of the QM residues of the ten average structures obtained from cluster analysis of QM-RMSD of the sulfenic acid intermediate.

Figure A3. Structure obtained from MD simulations of the sulfenic acid intermediate IC3'. Black lines represent hydrogen bond interactions.

Figure A4. Potential energy surface (kJ mol^{-1}) obtained (see Computational Methods) for activation of the catalytic cysteine (Cys13) by proton transfer via a bridging H_2O to the active site glutamate (Glu52). Relative energies are obtained at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96) single point calculations using electronic embedding formalism.

Figure A5. Potential energy surface (kJ mol^{-1}) obtained for formation of a sulfenic acid intermediate (Cys13SOH) involving the proton donors (A) Glu52COOH and Tyr92OH and (B) Glu52COOH and Tyr44OH. Relative energies are obtained at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96) single point calculations using electronic embedding formalism. **Figure A6.** Potential energy surface (kJ mol^{-1}) obtained for the reduction of sulfenic acid-containing complex to give an intramolecular disulfide bond with formation of H_2O . Relative energies are obtained at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96) single point calculations using electronic embedding formalism.

Figure A7. Potential energy surface obtained for the formation of pseudo-sulfurane intermediate in ApTPx using several basis sets as well as ONIOM(MP2/6-31G(d,p):Amber96).

Figure A8. QM layer of the QM/MM model used to investigate the overoxidation mechanism of ApTPx.

Figure A9. QM layer of the QM/MM model used to investigate the pseudo-sulfurane intermediate formation in hORF6.

Table B1. XYZ-coordinates of the optimized structures of all species considered in this study.

Table B2. A full list of the top 292 potential MsrA activator ligands and their docking score.

Appendix C. Copyright permission

List of Abbreviated Symbols

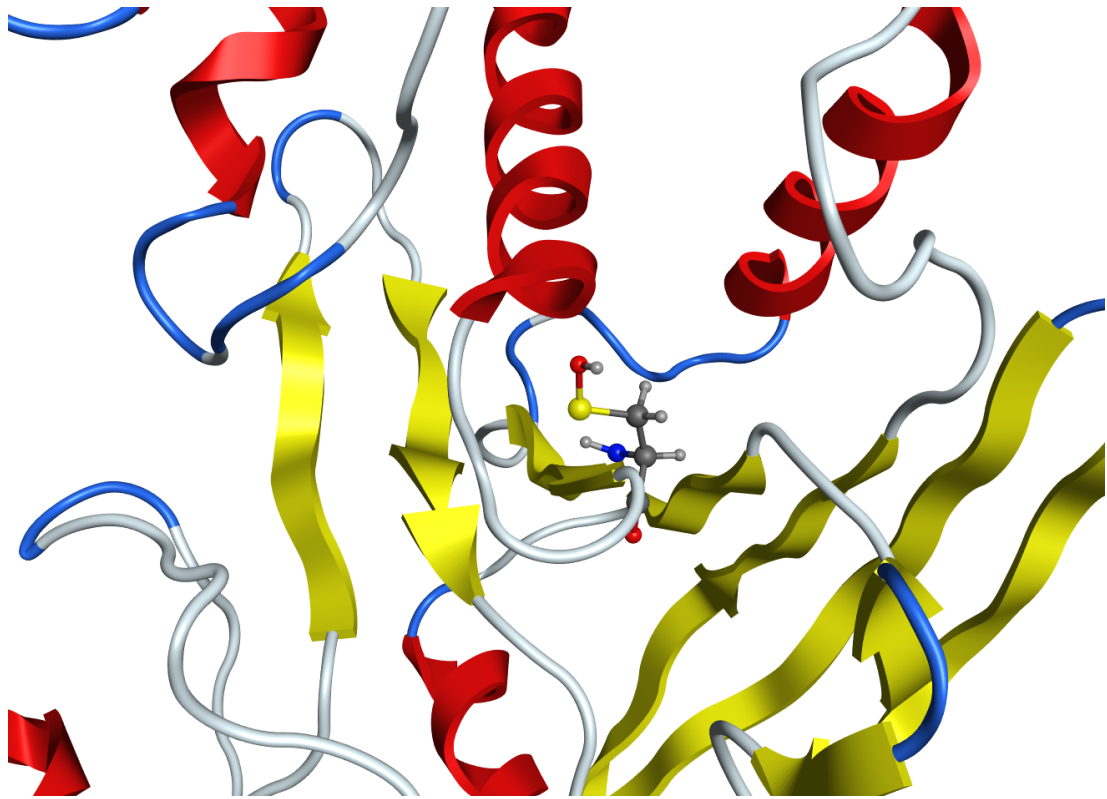
AMBER	Assisted Model Building and Energy Refinement
AMD	Arylmalonate decarboxylase
ApTPx	Thioredoxin peroxidase of <i>A. pernix</i> .
ATP	Adenosine triphosphate
CASSCF	Complete active space SCF
CC	Coupled cluster
CG	Coarse grain
CI	Configuration interaction
CTI	Cis-trans isomerization
Cys	Cysteine
DFT	Density functional theory
DFTB	Self-consistent charge density functional tight-binding method
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EE	Electrostatic embedding
EVB	Empirical valence bond
FAD	Flavin adenine dinucleotide
GHO	Generalized hybrid orbital
GluR	Glutamate racemase
GR	Glutathione reductase

GSH	Glutathione
HF	Hartree Fock
LDA	Local density approximation
MD	Molecular dynamics
ME	Mechanical embedding
Met	Methionine
Met-O	Methionine sulfoxide
MI	Maleate isomerase
MNDO	Modified neglect of diatomic differential overlap approximation
MM	Molecular mechanics
MOE	Molecular Operating Environment program
MP2	Second-order Møller-Plesset perturbation theory
Msr	Methionine sulfoxide reductase
NAD	Nicotinamide adenine dinucleotide
NBO	Natural bond orbitals
Nox	NADH oxidase
Npx	NADH peroxidase
PA	Proton affinity
PC	Product complex
PCM	Polarized continuum model
Perg	Pergolide
PerSO	Pergolide sulfoxide
PRC	Pre-reactive complex
ProR	Proline racemase
Prx	Peroxiredoxin
PTP	Protein tyrosine phosphatase
QCISD	Quadratic configuration interaction

QM	Quantum mechanics
QM/MM	Quantum mechanics/molecular mechanics
QTAIM	Quantum theory of atoms in molecule
RC	Reactive complex
RMSD	Root mean square deviation
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
SAM	S-adenosyl-L-methionine
Srx	Sulfiredoxin
Trx	Thioredoxin
TS	Transition state

Chapter 1

Introduction



1.1 Introduction:

The importance of sulfur in biochemistry has been recognized as early as the first half of the 20th-century with the discovery of insulin.¹ Subsequently it has been found to play a key role in all major classes of biomolecules such as proteins, nucleic acids, and carbohydrates, as well as cofactors and many metabolites.² In fact, sulfur is the only element other than N, C, and H that occurs in the 20 genetically encoded amino acids, being found in both cysteine (Cys) and methionine (Met).³ Sulfur-containing biomolecules greatly enrich the chemical diversity in cells in part as they can undergo a diverse variety of chemical reactions including nucleophilic substitution (e.g. thioredoxin), electron transfer (e.g. glutathione reductase (GR)), proton, hydrogen or hydride transfer (e.g. NADH peroxidase (Npx)), and oxygen atom transfer (e.g. NADH oxidase (Nox)).^{4,5}

This versatile chemistry of sulfur containing-biomolecules partially originates from being in the chalcogens group of the periodic table wherein sulfur and oxygen share similar chemical reactivity.^{6,7} However, being in lower position in the periodic table, sulfur has a lower electronegativity.⁶ As a result, sulfur-containing biomolecules are often better nucleophiles than oxygen-containing biomolecules. Additionally, disulfide bonds (RS–SR) are generally more stable than the peroxide bonds (RO–OR).^{6,7}

Another important contributing reason for sulfur's distinct reactivity is its electronic configuration of $[\text{Ne}]3s^23p^43d^0$, which contains an energetically accessible d-orbital that can be involved in bonding.⁸ Indeed, this electronic configuration enables sulfur to exhibit a range of oxidation states.⁸ Notably, in biochemistry, the oxidation states of sulfur have been found to vary from -2 to +6.^{5,8,9} Interestingly, it has been noted that there appears to be a correlation between the complexity of a cell and its sulfur content. For instance, in archaea the average occurrence of Cys in proteins is 0.4%-0.5% while in mammalian cells it is 2.26%.⁶

Thiol (-SH) is a common sulfur containing functional group within biochemical systems. Indeed, the unique R-group of Cys contains a thiol. At physiological pH it predominantly exists in its neutral form (i.e., RSH). For example, as the pK_a of the R-group thiol of Cys is around 9.⁸ However, the reactivity of a thiol can be increased by its deprotonation to form a thiolate anion (RS⁻). The latter is a much better nucleophile, and can react with soft or hard electrophiles such as carbonyl and phosphoryl groups.⁸ Notably, the high reactivity of RS⁻ plays key roles in a range of cellular functions such as catalysis, post-translational modifications, protein structure, metal binding, and detoxification of xenobiotics.^{4,5,10}

The R-group of Met contains a divalent sulfur atom (RSCH₃). Notably, Met has been found to play a number of crucial cellular roles, especially in redox signaling.^{11,12} This is due in part to the fact that it is one of the most readily oxidized amino acids.¹³ Indeed, due to its ability to undergo reversible oxidation under some conditions it has been suggested to play a role as a protein 'oxidation sink' or antioxidant.^{12,14} It is also noted that Met is the initiating amino acid in eukaryotic protein synthesis.¹⁵ In addition, upon its metabolism, Met can act as a methyl group donor, and form other sulfur containing amino acids such as homocysteine and cysteine.¹⁶

Unlike other redox reactions in biochemistry, Cys mediated redox systems are unique as they do not require metal ions or cofactors (e.g. NAD⁺ or FAD).^{4,17} In Cys the sulfur is in its fully reduced oxidation state of -2 and can readily undergo several redox reactions.^{4,18} This is due to the low oxidation/reduction potential of the Cys side chain; E°' -0.27 to -0.125 V.⁸ In general, higher oxidation states of Cys, specifically from -1 to +4, are less stable and a larger amount of energy is required for their further oxidation.⁶ Regardless, a diverse array of thiol-derived species' can be obtained via redox reactions in biochemical systems including sulfenic, sulfinic and sulfonic acids, thiosulfinate, thiosulfonate, disulfide trioxide, disulfone and disulfide (Figure 1.1).^{4,6,19} It is noted that the formation of a disulfide is the most common thiol redox modification.^{5,17}

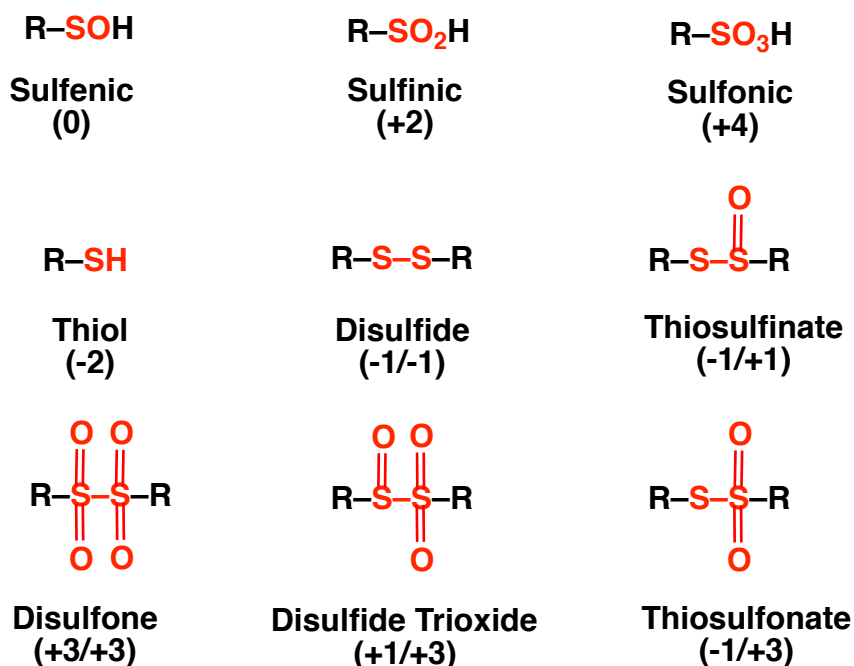


Figure 1.1. Schematic illustration of some thiol-derived species obtained via redox in biochemical systems.

With a sulfur oxidation state of 0, sulfenic acid (RSOH) has both nucleophilic and electrophilic character. As such, it is a chief intermediate in thiol redox chemistry with broad biological roles in signal transduction, non-enzymatic protein folding, protection against reactive oxygen species' (ROS) and modulating gene transcription.^{6,20,21} It is also now proposed to be an essential precursor intermediate for disulfide bond formation.²² RSOH can be formed via the direct oxidation of Cys by reactive oxygen or nitrogen species' (ROS/RNS), peroxyntirite, or hypochlorous acid. Alternatively, it can be formed indirectly either during some enzymatic mechanisms, or xenobiotic metabolism.^{21,23} Sulfenic acid is highly reactive and unstable,^{6,19} and can be irreversibly oxidized to form sulfinic or sulfonic acid.²⁰ Hence, it generally occurs as an intermediate along a reaction pathway.

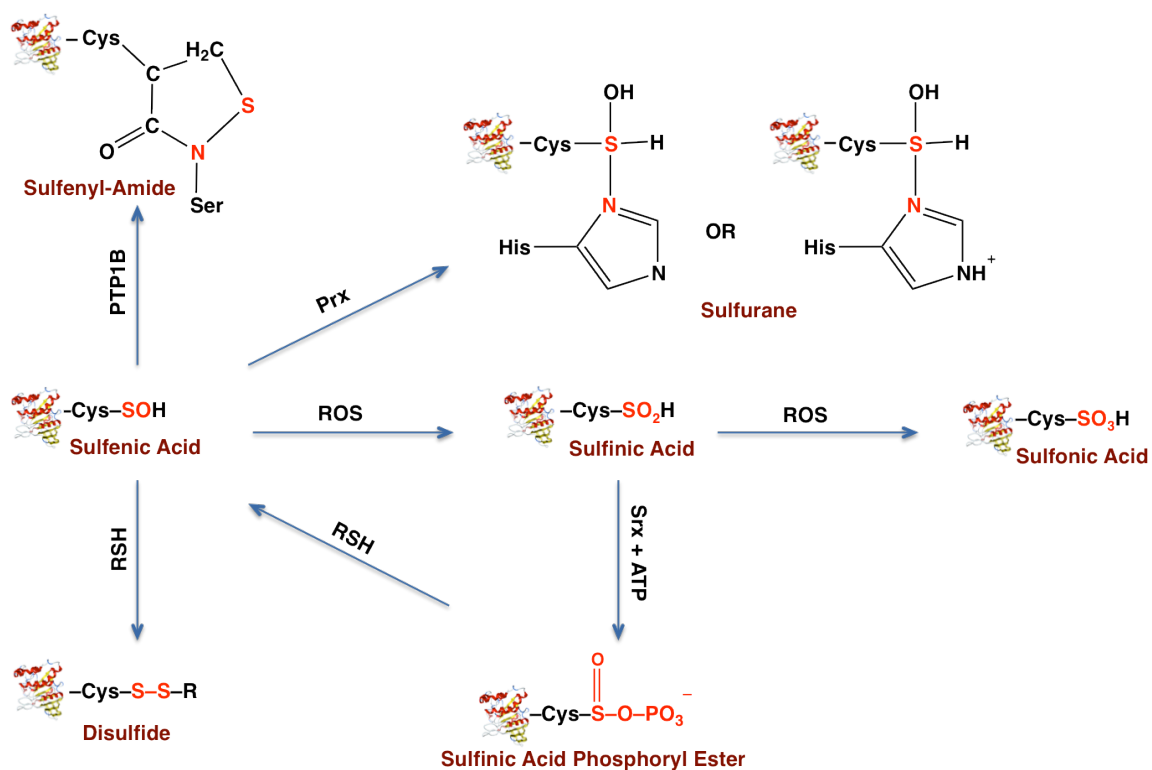


Figure 1.2. Schematic illustration of sulfenic acid proposed reactions in biochemistry.

The protection of sulfenic acid against overoxidation is crucial to maintaining cellular health.^{20,21} In fact, the presence of high oxidation state intermediates are hallmarks for many diseases.⁶ Consequently, many reactions in biochemical systems have been identified as protective mechanisms for sulfenic acid. The most common mechanism is its reduction to disulfide. However, some mechanisms involve the formation of unusual and unique species.^{6,19,24} For instance, using X-ray crystallography a sulfenyl-amide intermediate was identified in the enzyme protein tyrosine phosphatase 1B.²⁵ In archaeal peroxiredoxin (ApTPx) a unique electron density was identified in the crystal structure that was suggested to be a hypervalent sulfurane intermediate.²⁶ The oxidation of sulfenic acid to a sulfinic intermediate in some Prx's was found to be reversible via interactions with sulfiredoxin and an ATP cofactor.²⁷⁻²⁹

Despite its importance to biochemical systems, much of sulfur's redox chemistry, for example that of sulfenic acid, remains unknown or unclear. Notably, some of the above

proposed intermediates have not been experimentally or computationally characterized. In this thesis several computational modeling approaches have been applied, at times synergistically, to investigate sulfur-related biochemistry, in particular the redox chemistry of Cys and Met including the formation and chemistry of sulfenic acid.

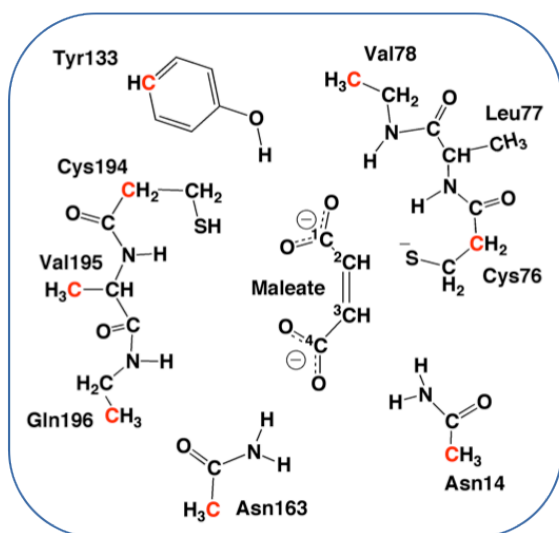
1.2 References

- (1) http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1955/press.html
[Accessed June 20, 2014]
- (2) Komarnisky, L. A.; Christopherson, R. J.; Basu, T. K. *Nutrition* **2003**, *19*, 54.
- (3) Cooper, A. J. L. *Annu. Rev. Biochem.* **1983**, *52*, 187.
- (4) Paulsen, C. E.; Carroll, K. S. *Chem. Rev.* **2013**, *113*, 4633.
- (5) Giles, N. M.; Giles, G. I.; Jacob, C. *Biochem. Biophys. Res. Commun.* **2003**, *300*, 1.
- (6) Gupta, V.; Carroll, K. S. *Biochim. Biophys. Acta-Gen. Subj.* **2014**, *1840*, 847.
- (7) Schaumann, E. *Sulfur-Mediated Rearrangements I*; Springer-Verlag Berlin: Berlin, Germany, 2007; Vol. 274, p 1.
- (8) Reddie, K. G.; Carroll, K. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 746.
- (9) Jacob, C.; Giles, G. L.; Giles, N. M.; Sies, H. *Angew. Chem. Int. Ed.* **2003**, *42*, 4742.
- (10) Hall, A.; Nelson, K.; Poole, L. B.; Karplus, P. A. *Antioxid. Redox Signal.* **2011**, *15*, 795.
- (11) Schöneich, C. *Biochim. et Biophys. Acta (BBA)-Proteins Proteom* **2005**, *1703*, 111.
- (12) Kim, G.; Weiss, S. J.; Levine, R. L. *Biochim. Biophys. Acta-Gen. Subj.* **2014**, *1840*, 901.

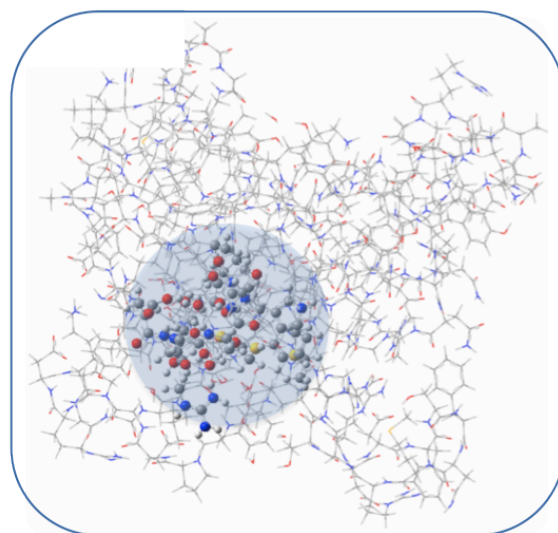
- (13) Dokainish, H. M.; Gault, J. W. *Biochemistry* **2013**, *52*, 1814.
- (14) Levine, R. L.; Mosoni, L.; Berlett, B. S.; Stadtman, E. R. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15036.
- (15) Laursen, B. S.; Sorensen, H. P.; Mortensen, K. K.; Sperling-Petersen, H. *U. Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101.
- (16) Brosnan, J. T.; Brosnan, M. E. *J. Nutr.* **2006**, *136*, 1636S.
- (17) Brandes, N.; Schmitt, S.; Jakob, U. *Antioxid. Redox Signal.* **2009**, *11*, 997.
- (18) Poole, L. B.; Nelson, K. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 18.
- (19) McGrath, A. J.; Garrett, G. E.; Valgimigli, L.; Pratt, D. A. *J. Am. Chem. Soc.* **2010**, *132*, 16759.
- (20) Roos, G.; Messens, J. *Free Radic. Biol. Med.* **2011**, *51*, 314.
- (21) Kettenhofen, N. J.; Wood, M. J. *Chem. Res. Toxicol.* **2010**, *23*, 1633.
- (22) Rehder, D. S.; Borges, C. R. *Biochemistry* **2010**, *49*, 7748.
- (23) Mansuy, D.; Dansette, P. M. *Arch. Biochem. Biophys.* **2011**, *507*, 174.
- (24) Jeong, J. H.; Jung, Y. S.; Na, S. J.; Jeong, J. H.; Lee, E. S.; Kim, M. S.; Choi, S.; Shin, D. H.; Paek, E.; Lee, H. Y.; Lee, K. J. *Mol. Cell. Proteomics* **2011**, *10*.
- (25) van Montfort, R. L. M.; Congreve, M.; Tisi, D.; Carr, R.; Jhoti, H. *Nature* **2003**, *423*, 773.
- (26) Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6238.
- (27) Jonsson, T. J.; Tsang, A. W.; Lowther, W. T.; Furdui, C. M. *J. Biol. Chem.* **2008**, *283*, 22890.
- (28) Jeong, W.; Park, S. J.; Chang, T. S.; Lee, D. Y.; Rhee, S. G. *J. Biol. Chem.* **2006**, *281*, 14400.
- (29) Lowther, W. T.; Haynes, A. C. *Antioxid. Redox Signal.* **2011**, *15*, 99.

Chapter 2

Overview of Computational Methods



QM-Cluster



QM/MM

2.1 Introduction

In 1888 J. L. Gay-Lussac wrote, “We are perhaps not far removed from the time when we shall be able to submit the bulk of chemical phenomena to calculation”.¹ Almost 40 years after Gay-Lussac's prophecy, Schrodinger and Dirac independently established the foundation of quantum chemistry conveying his dream to reality. However, the application of quantum chemical calculations remained limited; for decades only the investigation of simple systems involving one or a few atoms were practical. Later, the advent of computers together with the work of many scientists such as John Pople and Walter Kohn (Nobel Prize winners in chemistry 1998)² have pushed the limits of chemical calculations to deal with more complex systems. More recently, Karplus, Levitt and Warshel were awarded the Nobel Prize in chemistry 2013³ “for the development of multi-scale models for complex chemical systems”. As a result, the application of quantum chemical calculations has been extended to deal with complex systems not only in chemistry but also biology, enzymology, genetics, and engineering.

2.2 Computational Enzymology

Enzymes are remarkable natural biological catalysts that speed up chemical reactions dramatically to maintain cellular life. Elucidating the origin of enzymes' catalytic power as well as their mechanism of action is central to gaining a more complete understanding of biochemistry, as well as fundamental chemical principles. Indeed, such information could be used for several applications such as designing new catalysts as well as new therapeutic drugs (e.g. enzyme inhibitors).⁴ Enzymes are very versatile and utilize numerous mechanisms of action, making their characterization challenging.^{5,6} In addition, due to the nature of transition states and the fact that many reaction intermediates are highly reactive and transient,⁶ traditional experimental methods are generally unable to determine all atomistic-level details involved in their chemical reactions.^{7,8}

In the last two decades, computational modeling techniques have been established as a powerful tool to dissect in detail enzymatic chemical reactions describing the nature of intermediates and transition states.^{7,9,10} More importantly, several fundamental questions could be answered such as the origin of the enzyme catalytic power, the important catalytic residues and their mutations, the mode of ligand binding in the protein as well as the dynamics of its structure and energetics of the reaction.^{7,8} Thus, currently, the synergistic use of chemical modeling and laboratory experimentation is essential to study these powerful catalysts.^{6,8}

Computational enzymology (Figure 2.1) is a growing field of study with potential contributions in pharmaceutical, chemical and biotechnology industries.^{11,12} Several computational chemistry methods have been developed to help answer questions in enzymology. These methods include quantum mechanics (QM)-cluster, hybrid quantum mechanical/molecular mechanical (QM/MM), empirical valence bond (EVB), molecular dynamics simulations (MD) and Car-Parrinello molecular dynamics (CPMD).^{9,13-15} Notably, the aforementioned methods can achieve high accuracy when studying chemical systems. For example high-level QM-based methods can accomplish chemical accuracy near 1 kcal mol⁻¹,¹⁶ while parameterized-empirical based methods have also been shown to be able to provide good agreement with experiment.¹⁷

2.3 Protein Dynamics and MD Simulations

MD simulations are one of the fundamental tools in molecular modeling.¹⁸ They enable one to simulate the time dependent behavior of a molecular system by integrating Newton's laws of motion, providing insights into for example protein dynamics, protein-ligand and protein-protein interactions as well as conformational fluctuation. In computational enzymology, MD simulations can also be used to generate a number of plausible conformers of the initial reactive complex (RC).¹⁹ Subsequently the generated conformers can be analyzed to select either a single average structure to represent the

reactive complex or in some cases, several starting structures could be used to examine the effect of conformer variation on the mechanism.^{19,20} Notably, choosing an accurate RC is important to obtaining meaningful results in subsequent mechanistic calculations.^{10,21}

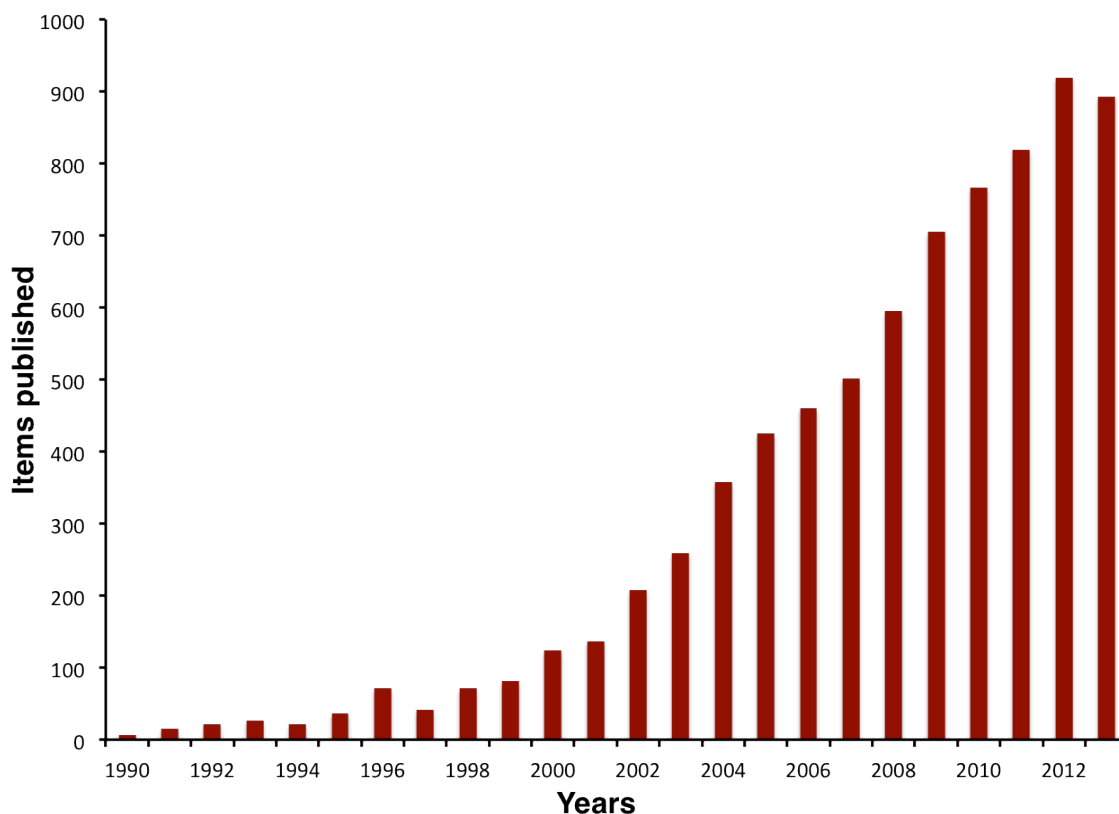


Figure 2.1. Graph indicating the results of a search on ISI Web of Science using the criteria: Topic = computational and Topic = enzyme sorted by publication year.

A large number of conformers are needed to calculate the free energy of an enzymatic system.²² This can be achieved by calculating molecular trajectories using several simulation methods, such as potential of mean force (PMF).²³ In general, understanding protein dynamics is important as many enzymes undergo large conformational changes during their catalytic cycle.¹⁹ Protein dynamics is involved in both substrate binding and product release, which could be the rate-limiting step of the mechanism.⁷

In biological systems, the size of the required chemical model is often quite large and complex. Similarly, the time scale required for their simulation is prolonged. Therefore, a cheaper level of theory is usually used to run the simulation. Typically, a molecular mechanical (MM) force field is used in simulating the trajectories of these systems. There are several computational programs to perform MD simulations such as AMBER, CHARMM, GROMOS and NAMD.²⁴⁻²⁷ The simulation starts by calculating the forces at particular time for a finite time period via solving the Newton equation of motion numerically.

2.4 Molecular Mechanics (MM) Force Fields

These methods use classical mechanics to predict the total energy of molecules as a function of their conformation. As a consequence, in MM force fields electrons are ignored and the system is expressed using empirical parameters derived from experiment or at times high-level conventional *ab initio* results. Hence, they are most applicable to ground state and covalent structures. Thus, processes that includes bond breaking or formation can not be described using MM.²⁸ However, MM force field is very useful in investigating energy changes upon structural rearrangements as in MD simulations.

Additionally, the total energy is calculated using simple functions,²⁸ an example of which is shown in equation 2.1.

$$E^{MM} = \sum E_{str} + \sum E_{bend} + \sum E_{tors} + \sum E_{vdw} + \sum E_{el} \quad (2.1)$$

In the above example the total MM energy of the system (E^{MM}) is written as the sum of simple harmonic terms of bond stretching (E_{str}) and valence angle bending (E_{bend}), a periodic term to represent the torsional angles (E_{tors}), while Van der Waals (E_{vdw}) and electrostatic (E_{el}) interaction contributions are calculated using Lennard-Jones potentials and Coulomb's law, respectively (equation 2.2).

$$E_{vdw} = \left(\frac{C_{12}}{R_{ij}}\right)^{12} - \left(\frac{C_6}{R_{ij}}\right)^6$$

$$E_{el} = \frac{e^2 Q_i Q_j}{4\pi\epsilon_0 R_{ij}} \quad (2.2)$$

The terms C_{12} and C_6 are the repulsion and attraction parameters respectively, Q_i is the atomic partial charge, R_{ij} is the interatomic distance, and e^2 and ϵ_0 are the unit charge and dielectric constant, respectively.

Many force fields have been developed for biomolecules (e.g. protein and DNA). This includes all atom force fields that represent all atoms in the system such as AMBER, CHARMM22-27 and OPLS/AA.^{24,26,29} Other force fields only deal with heavy atoms and polar hydrogen. In this case nonpolar hydrogens are represented as part of the carbon atom to which they are bonded. It is important to mention that atomic partial charges in most force fields are kept constant and do not change in response to environmental and/or conformational changes. Thus, electronic polarization is only included in an average way. Recently, there has been great interest and research into generating new polarizable force fields which would include electron polarization.^{30,31}

Although MD simulations based on MM methods are a well-established tool to simulate protein motion and dynamics, the size and the complexity of the biological systems generally allows for only relatively short simulations in the nanosecond (ns) time scale to be performed.³² Many biological problems would need much longer MD simulations (in micro to millisecond time scale) to be executed allowing for better sampling of the system.³³ In atomistic classical MD simulations the high barriers between minima often cannot be crossed. Several approaches have been developed to overcome these limitations. For instance, steered molecular dynamics allows the system to cross barriers by applying a force to the system forcing it to deviate from its initial conformational state.³⁴ This approach is very useful especially in studying the (un)folding of proteins and transportation of molecules through membrane proteins.^{35,36} Another

approach that has been fast growing is the generation of coarse-grained (CG) force fields.³² These force fields reduce computational cost by representing a small group of atoms as a single particle and thus allow for much longer simulation times. CG-MD has been successfully applied to study lipids and membrane proteins.^{37,38} Several CG force fields have been developed such as the all atom MARTINI force field.³⁹ Finally, other approaches are also been used such as targeted MD simulations and Monte Carlo MD simulations.^{40,41}

2.5 Quantum Mechanical (QM) Methods

Unlike MM methods, QM methods allow one to describe bond(s) formation/breaking processes in chemical systems.^{13,16} Thus in order to study the catalytic reactions in enzymes, generally, QM is the method of choice to obtain all detail including the geometries and energies of the unstable transition states and intermediates. To date, several different QM methods have been developed which can be categorized into three main types: 1) wavefunction-based (*ab initio*), 2) density functional theory (DFT) and 3) semiempirical methods. Each of these aim to find an approximate solution to the time-independent Schrödinger equation as it cannot be solved exactly for system containing more than one electron. The Hamiltonian of the non-relativistic time-independent Schrödinger equation of a molecular system is represented in equation 2.3.

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla^2 - \sum_{a=1}^n \frac{1}{2M_a} \nabla^2 - \sum_{i=1}^N \sum_{a=1}^n \frac{Z_a}{r_{ia}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{a=1}^n \sum_{\beta>a}^n \frac{Z_a Z_\beta}{r_{a\beta}} \quad (2.3)$$

The first two terms represent the kinetic energies of the electrons and nuclei, respectively. The following three terms describe the potential energy of electron-nuclei attraction, electron-electron repulsion and nuclei-nuclei repulsion, respectively. In addition, α and β represent the nuclei, i and j represent the electrons, and N and n represent the total number of electrons and nuclei in the system, respectively.

In order to address the complexity of the time-independent Schrödinger equation, several approximations must be made. Two of the most widely applied and common approximations are the (i) Born-Oppenheimer approximation in which the nuclei are assumed from the perspective of the electrons to be stationary, and (ii) orbital approximation in which the motions of the electrons are assumed to be independent of each other and hence each electron can be assigned its own spin orbital. In fact, QM methods differ based on the approximations applied. For example, the simplest *ab initio* method is based on Hartree Fock (HF) theory, which assumes that each electron can be considered as moving through an average field generated by the other electrons. As a result, the HF method is said to be an uncorrelated method. Unfortunately, the error in HF energies due to the neglect of electron correlation can lead to significant errors in describing molecular and electronic properties of the system. Therefore, several post-HF methods have been developed to account for electron correlation. These methods include Møller-Plesset perturbation theory (e.g. MP2), coupled cluster theory (CC), complete active space SCF (CASSCF) and Configuration interaction techniques (CI). Although all previously mentioned methods significantly improve the accuracy of the calculation, they cannot be applied to large systems due to their high computational cost.

DFT provides an alternative method that retains chemical accuracy at lower computational cost. It can calculate electronic and molecular properties using the molecule's electron density. This reduces the computational cost dramatically as the electron density is a function of three variables instead of the wavefunction's dependence on $3N$ variables. The main basis of DFT theory is that the energy of the system is a functional of the electron density.⁴² Unfortunately, the exact functional correlating the two is not known. Thus, numerous approximate functionals have been proposed and parameterized based on experimental results or high level *ab initio* calculations. These functionals can be categorized into: pure DFT (e.g. Local Density Approximate (LDA) methods),⁴³ generalized gradient approximation (GGA) (e.g. PBE),⁴⁴ hybrid-GGA

functionals (e.g. B3LYP)^{45,46} in which part of the exchange energy is calculated using HF, meta-GGA functionals (e.g. M06L) and hybrid-meta-GGA (e.g. M06-2X).⁴⁷ B3LYP was the most commonly employed functional in enzymology.⁶ However in the recent years other functionals have started to gain increasing use such as the M06 series.⁴⁷

Semiempirical methods are the least computationally demanding as they involve a number of approximations that aim to simplify the calculation, for instance, entire types of integrals. Therefore, they are also on average the least accurate among QM methods. Despite such drawbacks, they allow for the study of larger systems and are widely used in studying protein dynamics (e.g. QM/MM MD).⁴⁸ Therefore, several semiempirical methods have been developed based on different approximations. For instance, the modified neglect of diatomic differential overlap (MNDO) approximation includes some of the most widely used semi-empirical methods, AM1 and PM3.⁴⁹ Furthermore, a semiempirical DFT based method, self-consistent charge density functional tight-binding (DFTB) has been recently developed and has been successfully applied to enzymatic systems.⁵⁰

2.6 Choosing a Starting Structure

For most computational studies the choice of the starting chemical structure is very important. Typically, the primary model of the enzyme is obtained from a high-resolution X-ray structure of the protein in complex with the substrate, substrate analogue or inhibitors.^{6,9} Alternatively, docking techniques can be used to model the binding of the substrate in the active site. Although abovementioned approach is very useful, it suffers from limitations such as lacking protein dynamics and conformational variability. Therefore, another alternative is to use NMR structures as they provide information on protein dynamics. However, in general, NMR structures suffer from low resolution on atomic positions especially with respect to the flexible parts of the protein.¹⁰ In addition,

in some cases if there is no X-ray or NMR structure, homology modeling could be used to generate the starting structure.

Notably, in order to address the limitations of the generated structures, preparation steps need to be used to generate a more reliable representation of the *in vivo* structure of the enzyme-substrate complex via MD simulations. Once a representative structure is chosen, the mechanistic details of the reaction can be elucidated using several approaches. In enzymology, two main approaches are usually used to study chemical reactions: QM-cluster and quantum mechanical/molecular mechanical (QM/MM). It is noted that other methods can also be used such as empirical valence bond (EVB).^{9,13,15,28}

2.7 Quantum Mechanical (QM)-Cluster Approach

QM-cluster is one of the main approaches to study chemical reactions. It has been successfully used for many years to study enzymatic reactions with good accuracy, especially for metalloenzymes.^{51,52,53} In part, this is due to the fact that most enzymatic reactions occur within a small region (e.g. enzymes active site). In fact, using small models of the active site allows one to apply highly accurate QM methods, which cannot be used with larger models due to prohibitive computational cost. Furthermore, it provides an efficient tool to test the accuracy of several levels of theories against highly accurate methods and subsequently to select a suitable method for further calculations using larger models.

The initial model is usually derived from a high-resolution X-ray structure of an enzyme complex with substrate/substrate analogue or from the generated conformers of an enzyme-substrate complex using MD simulation. The enzyme-substrate complex is truncated to include a minimal number of atoms in the QM model that adequately describe or mimic the active site chemistry. In the case of using DFT, which is commonly used to study enzymes, a chemical model consisting of between 150-200 atoms is chosen.^{9,10} The selection of the active site atoms is quite crucial and might affect the

obtained results. Therefore it is important to include certain residues and interactions in the active site including:^{9,10} 1) the catalytic amino acids involved in the reactions, 2) residues that are responsible for substrate binding and conformation in the active site, 3) residues that can alter catalysis via their interactions (e.g. hydrogen bond, charge stabilization) with the catalytic residues or the substrate and 4) minimal part of the substrate without affecting its binding or properties. Generally, only models of the selected residues are included and only those parts of them that are necessary (e.g., R-groups). For instance, histidine and aspartate are often modeled by imidazole and acetate, respectively. Furthermore, a minimal number of peripheral atoms are usually fixed during the optimization to keep the active site conformation similar to the native enzyme structure.

If all important interactions and residues are maintained in the model, the QM-cluster can offer very useful and accurate insights in the reaction including stationary points on the potential energy surface (PES) (intermediates and transition states) and energetics. In general, a PES describes the relation between the system coordinates and its energy, which is important for a complete understanding of their chemical reactions. In fact, QM-cluster can be used as an efficient tool to compare between alternative catalytic pathways by obtaining their PESs ruling out those pathways whose activation energies are too high. Furthermore, free energies can be also estimated using frequency calculations, which are commonly done on the optimized structure. Moreover, since the QM-cluster model does not include the surrounding residues, the effect of enzymatic environment is modeled using a solvation model method and an appropriately chosen dielectric constant. Several implicit solvation approaches are usually used including IEF-PCM, COSMO and CPCM.⁵⁴ The values of the dielectric constant to mimic protein environments vary between 2-20 with a value of 4 being commonly used.⁹

Although a QM-cluster approach is very useful in revealing reaction chemical insights, in some cases this approach might fail.⁹ For instance, if the long-range

interactions are important for the reaction, truncating the surrounding residues can affect the calculated results dramatically. Furthermore, the dynamics of the enzyme are neglected and the investigated reaction is elucidated using a single model, which might not represent the reaction effectively.⁵⁵

2.8 Quantum Mechanical/Molecular Mechanical (QM/MM) Approach

QM/MM is a hybrid approach that allows for high accuracy calculations using QM methods, combined with the simplicity of cheap MM calculations.^{13,28} This combination enables the study of large enzymatic systems in which the core region (active site) where the bond(s) breaking/formation occur is described using QM calculations. The remainder is described using MM force field allowing for environmental effects and interactions to be accounted for in the calculations (see Figure 2.2). As in the QM-cluster approach, the QM layer should include catalytic residues, reactive region of the substrate and important interactions that can directly affect the catalytic mechanism. The MM layer can either include the rest of the enzyme/solvent atoms or part of the enzyme, e.g., all those residues within 15 Å from the center of the QM layer.

The beginning of the QM/MM era started with the pioneering work of Warshel and Levitt (Nobel prize winners in chemistry 2013) in 1976.⁵⁶ However, this method started to become widespread only in 1990 after the work of Field, Bash and Karplus who described in detail the coupling between the QM (using semiempirical method) and the MM layers (CHARMM forcefield) and compared the accuracy of the method against *ab initio* calculations and experimental data.⁵⁷ Since then QM/MM approaches have become established as an important tool for the study of bimolecular systems.⁵⁸ In general, transition states, intermediate, activation energies, zero point corrections and tunneling all can be elucidated using this approach.²⁸ There are several QM/MM implementations available in various computational chemistry programs.

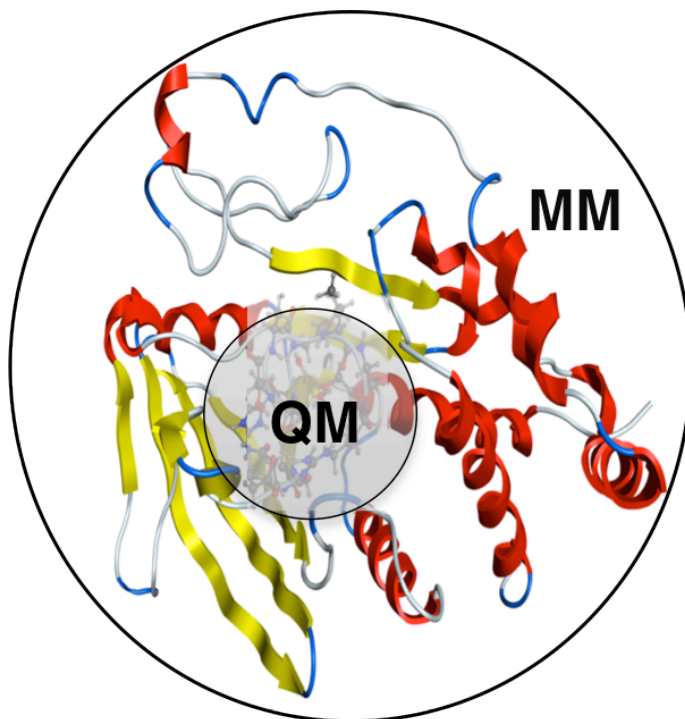


Figure 2.2. Schematic representation of QM/MM partitioning model.

The use of QM/MM has expanded and has been applied to study of inorganic/organometallic and solid-state systems.^{28,59} In addition it is currently also being used in free energies simulations allowing for better determination of ligand-protein binding.⁶⁰ A more recently developed extension of QM/MM is QM/MM-MD. It can, for example, be performed using a cheap QM method (e.g. DFTB) allowing for simultaneous modeling of the effects of protein dynamics.⁴⁸

2.8.1 QM/MM Additive Scheme

One of the main differences in QM/MM schemes is how the total energy of the system is calculated. In the additive scheme, the total energy is calculated by adding the individual energies of the QM layer, the MM layer and their coupling energy.⁵⁸ The full effective Hamiltonian (\hat{H}_{eff}) of the QM/MM system is:

$$\hat{H}_{eff} = \hat{H}_{QM,inner\ layer} + \hat{H}_{MM,outer\ layer} + \hat{H}_{QM/MM/coupling} \quad (2.4)$$

In which the first term is the quantum Hamiltonian of the inner layer (see equation 2.3) and the second term is the molecular mechanical Hamiltonian representing the MM layer (see equation 2.1). Finally, the third term in the equation represents the coupling between the two regions including electrostatic and van der Waals interaction. In this term, the MM atoms are described using point charges and van der Waals parameters (see equation 2.5).

$$\hat{H}_{QM/MM/coupling} = -\sum_{NM} \frac{q_M}{r_{NM}} + \sum_{nM} \frac{Z_n q_M}{R_{nM}} + \sum_{nM} \left(\frac{A_{nM}}{R_{nM}^{12}} - \frac{B_{nM}}{R_{nM}^6} \right) \quad (2.5)$$

In equation 2.5, M , N and n represent the MM atoms, the QM layer electrons and the QM layer nuclei, respectively. Furthermore, the first term represents the electrostatic interactions between the MM atoms and the QM layer electrons while the second term accounts for MM and QM layer nuclei electrostatic interactions. The last term represents the van der Waals interactions between the two layers. In general, the calculation of the QM/MM coupling term is problematic especially upon using link atoms and/or upon electrostatic embedding as will be discussed later in this chapter. The total energy of the system can be written as:

$$E_{QM/MM} = E_{QM,inner\ layer} + E_{MM,outer\ layer} + E_{QM/MM/coupling} \quad (2.6)$$

The additive scheme is a widely used approach in QM/MM and has been implemented in many software packages.

2.8.2 QM/MM subtractive scheme

Unlike the additive scheme, in QM/MM subtractive schemes the total energy of the system is obtained via addition and subtraction of component energy terms. More specifically, the total energy is calculated through three individual calculations: 1) the QM energy of the inner layer, 2) The MM energy of the entire system and 3) the MM

energy of the isolated inner layer. By subtracting the MM energy of the inner layer from the entire system, the coupling between the two layers is mainly calculated at the MM level of theory. The total energy of the system is the result of the subtraction of the MM inner system energy from the MM energy of the whole system, which is subsequently added to the QM energy of the inner layer as follow:

$$E_{QM/MM} = E_{QM,inner\ layer} + E_{MM,entire\ system} - E_{MM,inner\ layer} \quad (2.7)$$

In principle, this scheme can be used to extrapolate the energy of a system composed of n layers. Three different methods based on this scheme have been developed by Maseras and Morokuma including:^{61,62} 1) IMOMM, 2) IMOMO and 3) ONIOM. The latter allows for several layers to be used as well as numerous combinations of level of theories. It is important to mention that ONIOM is the method of choice in our studies. In addition, the following sections mainly discuss the interactions in ONIOM.

Although the subtractive QM/MM scheme does not require any parameterization of the coupling region, it still needs parameterization of the inner layer, which might be problematic to obtain accurately in certain cases (e.g. metal ions). However, all systematic errors are cancelled in the calculations. In addition, its implementation is straightforward as no coupling term to represent the interaction between the MM and the QM calculations is required.

2.8.3 QM/MM Electrostatic Interactions

One of the main challenges in QM/MM is to accurately describe the coupling between the two layers, including bonded and non-bonded interactions. More specifically electrostatic interactions are one of the effects that one may need to describe accurately. There are four different formalisms to evaluate QM/MM: A) mechanical embedding in which there is no polarization between the two layers, B) electrostatic embedding where

QM is polarized by MM, C) polarized embedding wherein MM is polarized by QM and D) Iterative polarization between QM and MM layers.

Mechanical embedding is the most computationally efficient as the QM layer is isolated during the calculations, that is, it is in a sense in the gas-phase. The coupling between the QM and MM layers is calculated at the MM level of theory that could be included in the force field. Since the QM inner system is evaluated in the absence of the MM layer, the MM layer only provides a mechanical/steric effect in the optimization of the QM layer. Furthermore, the interaction between the two regions is calculated based on fixed charges, which might change during the reaction during bond(s) formation/breaking in the inner layer. Thus in certain cases the partial charges of the QM inner layer should be calculated at each step along the reaction pathway. Similarly, Lennard-Jones parameters are usually kept constant in the calculations. All these limitations in the mechanical embedding formalism might affect the accuracy of some calculations.

Another alternative to improve the way electrostatic interactions are calculated is to use electrostatic embedding.⁶³ The affects of the electrostatic environment of the protein, surrounding the QM-region, can be modeled via the use of electrostatic embedding (EE). In this formalism, the charged MM atoms polarize the electron density of the QM layer. This is introduced into the calculation by adding a one electron operator to the QM Hamiltonian (equation 2.8).

$$\hat{H}_{QM,polarized} = \hat{H}_{QM} - \sum_J^M \frac{e^2 Q_J}{4\pi\epsilon_0 |r_i - R_J|} \quad (2.8)$$

The MM point charges are represented as Q_J . R_J and r_i represent the MM atom (J) and the electron (i) position, respectively.

Since the electronic Hamiltonian has an extra term, electrostatic embedding calculations are more computationally demanding.^{13,58} Therefore, it is commonly implemented via single point calculations on the structures optimized within the

mechanical embedding formalism. Indeed, the use of electrostatic embedding can provide more experimentally reliable results as the QM/MM coupling is now evaluated at the QM level of theory. Subsequently, no parameterization of the inner system is required. Despite all these advantages, electrostatic embedding might not be the best method to be used in all cases as the presence of the MM point charges can lead to the overpolarization of the QM inner layer, especially near the boundary region, in some cases. This mainly occurs upon using large basis set to describe the system leading to electrons spilling into the MM region.

In model C, the QM density polarizes the MM atoms but the polarized MM does not then polarize the QM layer.³¹ This requires the development of new force fields with flexible charges. Several approaches have been developed to model the polarization of MM atoms including the induced dipole model and the fluctuating charge models. Finally in model D both layers polarize each other via an iterative process. This requires the calculation of the MM polarization at every step of the self-consistent field iteration. Such calculations are very demanding and impractical to use with large QM/MM models. Unfortunately, such polarizable forcefield for bimolecular system is not available yet.

2.8.4 QM/MM Boundary

Since the QM and the MM layer are connected by covalent bonds, the partitioning of the two layers creates unpaired electrons in the QM subsystem, which is an artifact of the method. Several approaches have been developed to remedy this problem and saturate the valence of the frontier atoms.

The most straightforward approach and one that is commonly used, especially in ONIOM calculations, is the addition of monovalent link atoms to the QM subsystem. Hydrogen is the most commonly used link atom, however, in principle other atoms or even a group of atoms (e.g. methyl group) could be employed.⁵⁸ These atoms are only added during the calculation of the inner QM layer. Despite the simplicity of this

approach, it has several disadvantages:⁵⁸ 1) it complicates the energy expression, 2) it generates three extra degree of freedom for each link atom, 3) the generated bond is artificial and different from the original bond, and 4) it might lead to the overpolarization of the QM layer. Some of these problems are already addressed in chemical software packages. For instance, the extra degree of freedom is canceled by fixing the link atom along the partitioned bond between the QM and the MM atoms. Furthermore, in order to avoid overpolarization of the subsystem, many schemes have been developed including:^{31,58} 1) deleting MM charges atoms (e.g. Z1 scheme), 2) redistribution of the point charges around the link atom, and 3) deletion of one electron integrals of the link atom.

An alternative approach is the use of localized orbitals in which a doubly-occupied molecular orbital is used. Two commonly used such approaches are the: 1) hybrid orbital method, which introduces a molecular orbital at the boundary QM atom, and 2) generalized hybrid orbital approach (GHO)⁶⁴ which instead places a molecular orbital on the MM atom. In the hybrid orbital method (e.g. LSCF)⁶⁵ a hybrid occupied orbital is added at the QM atom and pointed towards the MM atom of the broken bond and kept frozen during the optimization. To apply this method the coefficient of the frozen orbital has to be calculated, which is usually done using small models assuming its transferability. Although this method does not cause overpolarization, it is more complicated than using link atoms. In GHO,⁶⁴ Gao and co-workers suggested the addition of the hybrid orbitals to the MM atom. This allows for the hybrid orbital linking the broken bond to be optimized while other auxiliary orbitals on the MM atom are fixed. Nevertheless, extensive parameterization is still needed. In general localized orbital approaches are more complicated. Comparative studies have shown that they can give slightly better results compared to the use of link atoms.

2.8.5 Setting-up QM/MM Model

In order to set-up a QM/MM model, special care is always required in choosing the boundary between the two layers. For instance, all atoms involved in bond breaking and formation must be included in the inner layer. Aromatic or conjugated systems cannot be divided between the two layers. Atoms that change hybridization during the reaction should also be included in the high layer. A general practice is to start the calculations with a minimum QM layer, and then to systematically increase it in order to examine the selection of model on the obtained energies.⁶⁶

As mentioned earlier the accurate description of the boundary and coupling between the two layers is important for obtaining reliable and accurate results. Therefore, it is always important to partition the layer by cutting non-polar bonds. Replacing C-C bonds by C-H when using link atoms also helps minimize any errors due to the need to partition. Each MM atom of the broken bonds has to be replaced by only one single link atom. It is also important to always cut single bonds to minimize the error. Furthermore, in the subtractive approach, the broken bonds have to be at least three bonds away from bond formation/breaking in the QM subsystem. This maintains the same MM parameters at the boundary region to be used during the reactions.

2.9 References

- (1) Thatcher, Oliver J. *The Ideas That Have Influenced Civilization*, in the Original Documents. 1907; Vol. 8. Reprint. London: Forgotten Books, 2013. Print.
- (2) http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/.
[Accessed July 2, 2014]
- (3) http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/.
[Accessed July 2, 2014]
- (4) Zheng, M.; Liu, X.; Xu, Y.; Li, H.; Luo, C.; Jiang, H. *Trends Pharmacol. Sci.* **2013**, *34*, 549.

- (5) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; Dover Publications: New York, USA. 1987.
- (6) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*, 689.
- (7) Gherib, R.; Dokainish, H. M.; Gauld, J. W. *Int. J. Mol. Sci.* **2014**, *15*, 401.
- (8) Karabencheva, T.; Christov, C. *Structural and Mechanistic Enzymology: Bringing Together Experiments and Computing*; Elsevier Academic Press Inc: San Diego, USA. 2012; Vol. 87, p 1.
- (9) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.
- (10) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *Chem. Soc. Rev.* **2012**, *41*, 3025.
- (11) Damborsky, J.; Brezovsky, J. *Curr. Opin. Chem. Biol.* **2014**, *19*, 8.
- (12) Ou-Yang, S. S.; Lu, J. Y.; Kong, X. Q.; Liang, Z. J.; Luo, C.; Jiang, H. L. *Acta Pharmacol. Sin.* **2012**, *33*, 1131.
- (13) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708.
- (14) Hutter, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 604.
- (15) Warshel, A. *Annu. Rev. Biophys. Biomolec. Struct.* **2003**, *32*, 425.
- (16) Mulholland, A. J. *Chem. Cent. J.* **2007**, *1*.
- (17) Braun-Sand, S.; Olsson, M. H. M.; Warshel, A. Computer modeling of enzyme catalysis and its relationship to concepts in physical organic chemistry. In *Advances in Physical Organic Chemistry*; Richard, J. P., Ed.; Elsevier Academic Press Inc: San Diego, USA. 2005; Vol. 40, p 201.
- (18) Karplus, M.; Gao, Y. Q.; Ma, J. P.; van der Vaart, A.; Yang, W. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **2005**, *363*, 331.
- (19) Zhang, Y. K.; Kua, J.; McCammon, J. A. *J. Phys. Chem. B* **2003**, *107*, 4459.
- (20) Doshi, U.; Hamelberg, D. *Adv. Exp. Med. Biol.* **2014**, *805*, 221.

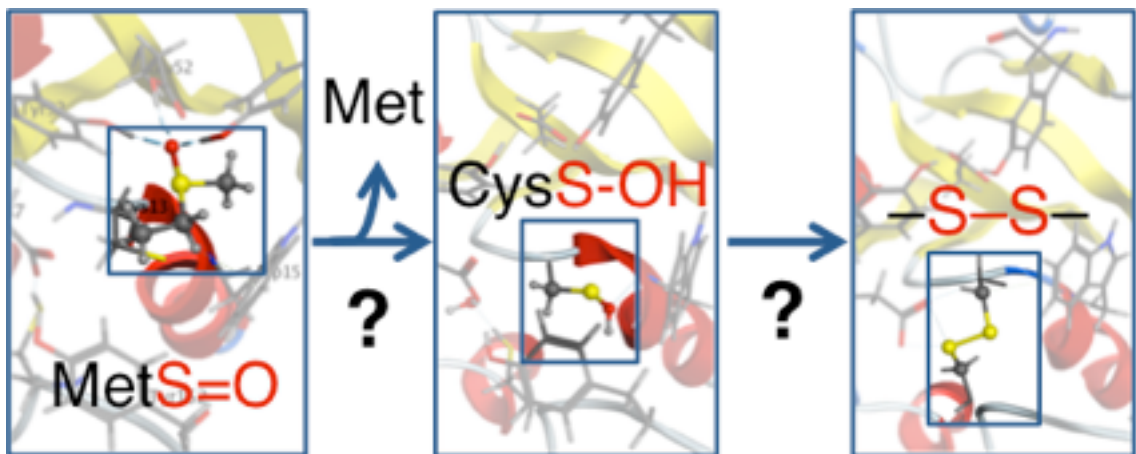
- (21) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645.
- (22) Woods, C. J.; Manby, F. R.; Mulholland, A. J. *J. Chem. Phys.* **2008**, *128*.
- (23) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
- (24) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671.
- (25) Van Gunsteren W.F., Daura X., Mark A.E. Encyclopedia of Computational Chemistry. John Wiley & Sons, Ltd; Chichester, UK. 2002; GROMOS Force Field.
- (26) 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.
- (27) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (28) Groenhof, G. Introduction to QM/MM Simulations. In *Biomolecular Simulations Methods and Protocols*; Monticelli, L., Salonen, Emppu Ed.; Humana Press; New York, USA. 2013; p 43.
- (29) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (30) Kim, B. C.; Young, T.; Harder, E.; Friesner, R. A.; Berne, B. J. *J. Phys. Chem. B* **2005**, *109*, 16529.
- (31) Thellamurege, N. M.; Hirao, H. *J. Phys. Chem. B* **2014**, *118*, 2084.
- (32) Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. *J. Struct. Biol.* **2007**, *157*, 593.
- (33) Lindahl, E.; Sansom, M. S. P. *Curr. Opin. Struct. Biol.* **2008**, *18*, 425.

- (34) Xu, Y. C.; Shen, J. H.; Luo, X. M.; Silman, I.; Sussman, J. L.; Chen, K. X.; Jiang, H. L. *J. Am. Chem. Soc.* **2003**, *125*, 11340.
- (35) Zhang, J.; Li, W. F.; Wang, J.; Qin, M.; Wu, L.; Yan, Z. Q.; Xu, W. X.; Zuo, G. H.; Wang, W. *IUBMB Life* **2009**, *61*, 627.
- (36) Khalili-Araghi, F.; Gumbart, J.; Wen, P. C.; Sotomayor, M.; Tajkhorshid, E.; Schulten, K. *Curr. Opin. Struct. Biol.* **2009**, *19*, 128.
- (37) Weingarth, M.; Prokofyev, A.; van der Crujisen, E. A. W.; Nand, D.; Bonvin, A.; Pongs, O.; Baldus, M. *J. Am. Chem. Soc.* **2013**, *135*, 3983.
- (38) Scott, K. A.; Bond, P. J.; Ivetac, A.; Chetwynd, A. P.; Khalid, S.; Sansom, M. S. P. *Structure* **2008**, *16*, 621.
- (39) Monticelli, L.; Kandasamy, S. K.; Periolo, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4*, 819.
- (40) Yu, S.; Larson, R. G. *Soft matter* **2014**, *10*, 5325.
- (41) Pedretti, A.; Labozzetta, A.; Lo Monte, M.; Beccari, A. R.; Moriconi, A.; Vistoli, G. *Biochem. Biophys. Res. Commun.* **2011**, *414*, 14.
- (42) Parr, R. G. *Annu. Rev. Phys. Chem.* **1983**, *34*, 631.
- (43) Seifert, G.; Porezag, D.; Frauenheim, T. *Int. J. Quantum Chem.* **1996**, *58*, 185.
- (44) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
- (45) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (46) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (47) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (48) Zhang, X.; Lei, M. *J. Theor. Comput. Chem.* **2013**, *12*.
- (49) Peshlherbe, G. H.; Hase, W. L. *J. Chem. Phys.* **1996**, *104*, 7882.
- (50) Elstner, M.; Frauenheim, T.; Kaxiras, E.; Seifert, G.; Suhai, S. *Phys. Status Solidi B-Basic Res.* **2000**, *217*, 357.
- (51) Himo, F.; Siegbahn, P. E. M. *Chem. Rev.* **2003**, *103*, 2421.

- (52) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643.
- (53) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R. Z.; Siegbahn, P. E. *M. Chem. Rev.* **2014**, *114*, 3601.
- (54) Guerard, J. J.; Arey, J. S. *J. Chem. Theory Comput.* **2013**, *9*, 5046.
- (55) Doshi, U.; McGowan, L. C.; Ladani, S. T.; Hamelberg, D. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5699.
- (56) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
- (57) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
- (58) Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
- (59) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185.
- (60) Menikarachchi, L. C.; Gascon, J. A. *Curr. Top. Med. Chem.* **2010**, *10*, 46.
- (61) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (62) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170.
- (63) Chui, C. O.; Shin, K. S.; Kina, J.; Shih, K. H.; Narayanan, P.; Moritz, C. A. Heterogeneous integration of epitaxial nanostructures: strategies and application drivers. In *Nanoepitaxy: Materials and Devices Iv*; Kobayashi, N. P., Talin, A. A., Islam, M. S., Eds.; Spie-Int Soc Optical Engineering; Bellingham, USA. 2012; Vol. 8467.
- (64) Pu, J. Z.; Gao, J. L.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 632.
- (65) Fornili, A.; Moreau, Y.; Sironi, M.; Assfeld, X. *J. Comput. Chem.* **2006**, *27*, 515.
- (66) Clemente, F. R. V., T.; Frisch, M. J. Getting the Most out of ONIOM: Guidelines and Pitfalls. In *Quantum Biochemistry: Electronic Structure and Biological Activity* Matta, C. F., Ed.; Wiley-VCH: Weinheim, 2010; Vol. 2, p 61.

Chapter 3

**An MD and QM/MM study on the Catalytic
Reductase Mechanism of Methionine Sulfoxide
Reductase A (MsrA): Formation and Reduction of
a Sulfenic Acid**



3.1 Introduction

Of the standard 20 genetically encoded amino acids methionine (Met) is one of the most easily oxidized.¹ In particular, oxidation at its R-group sulfur can lead to formation of methionine sulfoxide (Met-O).¹ Remarkably, Met-O can be reduced to give Met.² This uncommon amino acid redox chemistry has increasingly been shown to be critical for many biological processes including protein regulation, the calcium-induced signal transduction pathway and immune responses.^{1,3-6} However, such post-translational modifications induced by oxidative stress are also known to be involved in aging and age related diseases including cancer and Alzheimer's disease.⁷⁻⁹

Methionine sulfoxide reductases (Msr's) are a family of ubiquitous enzymes that reduce Met-O to Met.^{10,11} These enzymes have been shown to be important, for example, in cellular responses to oxidative stress,¹²⁻¹⁴ bacterial virulence,^{15,16} and against amyloid *b*-protein toxicity.¹⁷ There are two main classes of Msr's, A and B, which are stereospecific for the S- and R-Met-O epimers respectively.^{18,19} Both classes are thought to utilize the same overall reaction despite having quite distinct active site compositions and being structurally unrelated.²⁰ The few shared similarities include an active site cysteinyl and tryptophanyl that are thought to act as a nucleophile and help orient the substrate, respectively. In addition, both have several hydrogen bond donors that interact with the Met-O sulfoxide oxygen. However, the nature of these donors differs: MsrB has multiple ionizable and polar residues²¹ while MsrA contains two tyrosyl and a glutamyl residue.²² In vitro studies² of recombinant MsrA have shown it to be 10-fold more active than MsrB. Furthermore, for several species it has been found that knocking it out enhances their susceptibility to oxidative stress²³⁻²⁵ while its overexpression results in longer lifespans.^{26,27} Meanwhile, its down regulation in human breast cancer cells increased the disease's aggressiveness both in vivo and in vitro.²⁸

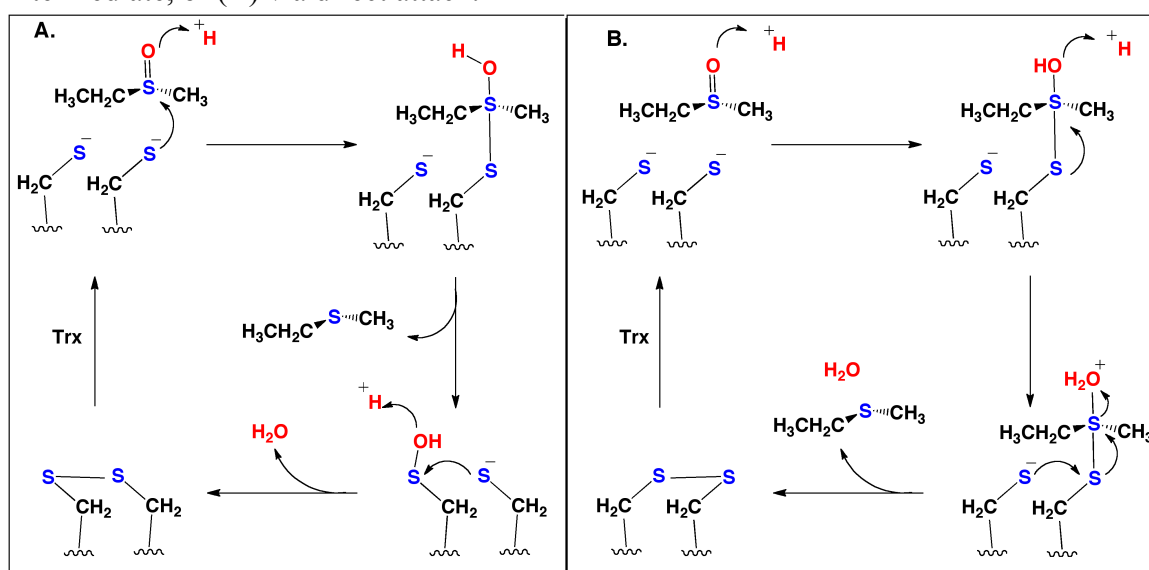
Several experimental kinetic studies on MsrA have examined the possible roles of various active site amino acid residues. For instance, substitution of the Cys51 in

Escherichia coli MsrA by serine deactivated the enzyme, thus confirming its essential catalytic role.²⁹ Similarly, substitution of the active site glutamyl (Glu94 in *Neisseria meningitidis*), thought to be important both in activation of Cys51 and as a proton donor to the substrate, by alanine drastically reduced the rate of catalysis 36500-fold.²² In contrast, mutation of either active site tyrosyl (Tyr82 or Tyr134 in *Neisseria meningitidis*) by Phe appeared to have little effect on the rate of reaction.²² However, simultaneous substitution of both by Phe decreased the catalytic rate significantly by ~10000-fold.²² In addition, using chemical probes and mass spectrometry a catalytic cysteine-derived sulfenic acid was detected in a wild type MsrA from *Escherichia coli*.²⁹ The ionization states of Cys51 (the catalytic cysteine) and a second active site cysteine Cys198, known as the recycling cysteine, have also been experimentally measured. In particular, it was observed that upon substrate binding the pK_a of Cys51 was reduced significantly from 9.5 to 5.7 in *Neisseria meningitidis*.²² However, it was only decreased to 8.0 if Glu94 was replaced by Gln or if both Tyr82 and Tyr134 were replaced by Phe. This was suggested to indicate that these three residues are important for polarizing the sulfoxide S=O bond which also leads to a decrease in the pK_a of Cys51.²²

Based on these and related experimental studies, several possible catalytic mechanisms for MsrA and by extension Msr's in general were proposed and are summarized in Scheme 3.1.^{18,30,31} Specifically, upon activation of Cys51 by some as yet unclear process to give the Cys51S⁻ thiolate, the proposed mechanisms **A** and **B** in Scheme 3.1 commonly begin with formation of a sulfurane via nucleophilic attack of Cys51S⁻ at the Met-O sulfur.³² However, two alternate mechanisms were suggested for the subsequent reactions of the sulfurane: (**A**) it undergoes a 1,2-shift of the sulfurane's –OH group from the substrate sulfur to the adjacent Cys51 sulfur to give the sulfenic acid Cys51SOH and the desired Met product or, (**B**) it is protonated at its oxygen to give a sulfonium. In the former the resulting catalytic cysteine-derived sulfenic acid intermediate readily reacts with the recycling cysteine, which itself has been activated and

is in its thiolate form, to give an intra-molecular disulfide bond. In contrast, in the latter proposed mechanism **B** the recycling cysteine thiolate directly attacks the Cys51S center of the sulfonium to give an intra-molecular disulfide bond and the reduced Met product, i.e., without formation of the sulfenic acid.¹⁴ In both cases, the active site is then ultimately regenerated via reduction of the disulfide bond by Thioredoxin (Trx).³³

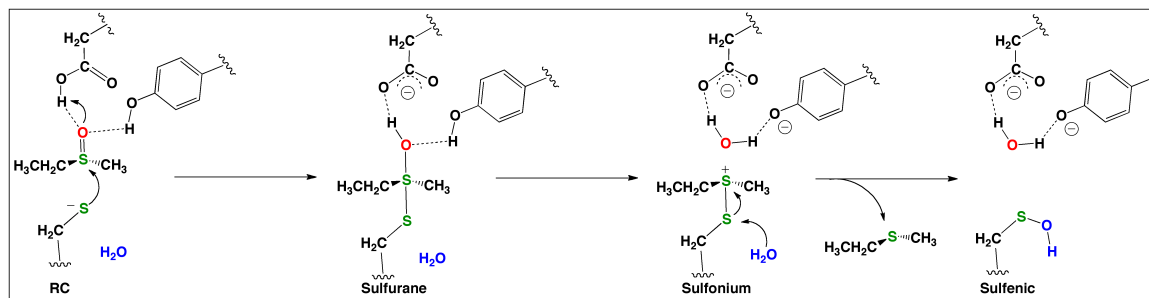
Scheme 3.1. Proposed^{30,31} reductase mechanisms for MsrA via (A) a sulfenic acid intermediate, or (B) via direct attack.



However, more recently Lim et al.¹⁸ using mass spectrometry and isotope labeling identified peptide modifications formed during the catalytic mechanism of mouse MsrA. To prevent other modifications such as disulfide bond formation they mutated Cys107, Cys218 and Cys227 to serine. They observed that the catalytic cysteine was converted to a sulfenic acid (CysSOH) during the course of the mechanism. Importantly, and in contrast to that previously proposed, the CysSOH oxygen was not derived from the initial Met-O substrate but instead from the aqueous solvent. Consequently, they proposed a modified mechanism for MsrA involving sulfonium and sulfenic acid intermediates highlighted in Scheme 3.2.

Computationally, there have been very few studies on the catalytic mechanism of Msr's. Recently, Thiriot et al.³⁴ used a DFT-cluster approach to gain insights into the reductase mechanism of MsrA. Specifically, the substrate was modeled by dimethyl sulfoxide (DMSO) and the R-groups of the active site residues Tyr134 and Tyr82, Glu94 and Cys51S⁻ by phenol and water, CH₃S⁻ and acetic acid respectively. In addition, general environment effects were included via use of a PCM solvation method with a dielectric constant (ϵ) of 2. Importantly, they showed that formation of a sulfonium intermediate may occur via nucleophilic attack of Cys51S⁻ at the substrate's sulfur with stepwise protonation of its sulfoxide oxygen by Glu94 and Tyr134. Furthermore, a sulfenic acid intermediate was calculated to lie quite low in energy. Unfortunately, a reaction pathway between the sulfonium ion and sulfenic acid was not elucidated. However, they suggested it may involve dissociation of the sulfonium with the H₂O moiety released then attacking the catalytic cysteinyl's sulfur to give a sulfenic acid.

Scheme 3.2. Proposed¹⁸ catalytic reductase mechanism for MsrA involving formation of a sulfenic acid intermediate via attack of a solvent water at the catalytic cysteine's sulfur centre.



Previously, we performed a computational investigation on the complete catalytic reductase mechanism of MsrB using a large DFT-cluster model and with the general protein environment modeled by a PCM solvation method with $\epsilon = 4$ as common in such approaches.³⁵ In particular, we showed that the mechanism may be initiated by

nucleophilic attack of the catalytic Cys at the substrate's sulfur with concomitant proton transfer from a nearby histidyl residue to form a sulfurane. A sulfonium cation was then formed via transfer of a proton from a second histidyl residue onto the sulfurane's oxygen. However, within the computational model used, two possible enzymatically feasible mechanisms were then obtained by which the Met product could be formed: (i) direct attack of the recycling cysteine thiolate at the sulfonium's catalytic cysteinyl sulfur, or (ii) via a sulfenic acid intermediate. The latter was found to not be feasible via intramolecular rearrangement of the sulfurane as previously suggested. Instead, it required a solvent water to attack the sulfonium's catalytic cysteinyl sulfur as experimentally observed in MsrA.¹⁸

A clearer elucidation of the catalytic mechanism of Msr's is central to a greater understanding of their role in aging, oxidative stress and other important physiological processes. In addition, due to their use of a diverse variety of sulfur compounds and chemistry, they represent a tremendous opportunity to obtain greater insights into the important area of sulfur biochemistry, in particular that of highly reactive species such as sulfenic acids. We have thus performed a detailed systematic computational investigation into the catalytic mechanism of MsrA. More specifically, we have complementarily applied Docking, Molecular dynamics (MD) simulations and an extensive ONIOM QM/MM approach to investigate activation of the catalytic cysteinyl and the subsequent reductase mechanism that leads to formation of Met and the intra-molecular disulfide bond.

3.2 Computational Methods

Docking and MD Studies. These calculations were performed using the Molecular Operating Environment (MOE) program.³⁶ The X-ray crystal structure (PDB ID: 1NWA) of MsrA from *Mycobacterium tuberculosis* (MsrA_{Mtb}) complexed with protein-bound methionine was used as a template for docking.³⁷ The substrate and all crystal structure

waters were removed then the substrate methionine sulfoxide (Met-O) docked into its active site where all residues within the first interaction shell of the catalytic cysteine (Cys13) were considered as in the active site. The London dG scoring function was used in conjunction with a force field refinement method in order to obtain the top 30 scoring structures. These were visually examined to choose the most suitable starting structure for further calculations. A molecular dynamics (MD) simulation was then performed on the chosen structure to allow for thermal relaxation. More specifically, the enzyme-substrate complex was spherically solvated up to 15 Å from the substrate. Then MD simulations were performed for 1 ns using a time step of 2 fs in a same manner as has been previously³⁸ described. The structures generated during the simulation were analysed and clustered according to the distance between the sulfur's of the catalytic cysteine and substrate into 10 clusters. The overlay of the 10 average structures of the generated clusters show similar interactions (See Appendix Figure A1); furthermore, the RMSD of the QM-region moieties (QM-RMSD) of the 10 structures was 0.12 Å, indicating a very consistent structure. The average structure of the most populated cluster was then optimized using the AMBER99 force field,³⁹ and the resulting structure (structure A') being used to construct a QM/MM chemical model to examine the catalytic mechanism of MsrA up to and including formation of a sulfenic acid intermediate.

A second MD simulation was performed to model possible rearrangements in the active site upon sulfenic acid formation. Similar to the protocol detailed above the X-ray crystal structure PDB ID: 1NWA was modified to remove the substrate and all crystallographic waters. The catalytic cysteine was then modified to form a sulfenic acid (CysSOH) and the resulting intermediate spherically solvated up to 15Å from the sulfur of the CysSOH moiety. The complex was equilibrated for 0.5 ns followed by a 1 ns production run as previously described.⁴⁰ The S-O bond of the CysSOH moiety was constrained to its optimized length as obtained at the B3LYP/6-31G(d) level of theory. Partial charges for the sulfenic acid were also obtained via a Mulliken population analysis

of **IC3** performed at the below QM/MM level of theory. The resulting structures obtained were cluster-analyzed based on the sulfenic acid hydrogen-bonding network into 10 clusters (see above). Similarly, the 10 average structures were overlaid ((See Appendix Figure A2) and analyzed and found to have a QM-RMSD of 0.24 Å. Again, the average structure of the most populated cluster was chosen and optimized using the AMBER99 force field.³⁹ The resulting complex, structure **B'**, was used to construct a QM/MM chemical model to examine the catalytic mechanism of MsrA for formation of the final product from the sulfenic acid intermediate.

QM/MM models and calculations. All QM/MM calculations were performed within the ONIOM formalism⁴¹ using the Gaussian 09⁴² suite of programs. The density functional theory method B3LYP; a combination of Becke's three parameter exchange functional⁴³ and Lee, Yang and Parr's correlation functional⁴⁴ was used for the high-layer (QM) while the AMBER96⁴⁵ force field was used for the low-layer (MM) parameterization. Optimized geometries (the default convergence criteria of Gaussian 09⁴² was used throughout) were obtained using the 6-31G(d) basis set for the high-layer and within the mechanical embedding (ME) formalism, i.e., at the ONIOM(B3LYP/6-31G(d):AMBER96)-ME level of theory. Frequency calculations were also obtained at this same level of theory in order to characterize the nature of the stationary points obtained (i.e., minima or transition structures) and to calculate Gibbs energy corrections (ΔG_{corr}). The single imaginary and first real frequencies for the optimized TSs are reported (See Appendix Table B1). Relative energies were then obtained via single point energy calculations at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96) level of theory within the electrostatic embedding (EE) formalism on the above optimized structures, and with inclusion of the appropriate ΔG_{corr} . All atoms in the QM layer were unfixed while for the MM layer only the C_{α} centres were kept fixed at their final MD positions. This and similar QM/MM approaches have been successfully used to explore a range of enzymatic systems.⁴⁶⁻⁵²

The chemical models used were obtained by truncating structures **A'** and **B'** to only include those residues in the first- and second-shell centered on the sulfur of the catalytic cysteine and where appropriate the substrate. The resulting two active sites models, **A** and **B**, enabled us to more effectively consider the different stages in the overall mechanism and are shown in Figure 3.1. More specifically, models **A** and **B** were used to investigate possible catalytic mechanisms of MsrA resulting in (i) formation of a enzyme-derived sulfenic acid intermediate and (ii) conversion of a sulfenic acid intermediate to the final products, respectively. In model **A** the Met-O substrate was modeled by ethylmethylsulfoxide. It should be noted that in model **A** (Figure 3.1A) the catalytic and recycling cysteinyls (Cys13 and 154), three tyrosyl (Tyr 92, 44 and 152), three charged residues (Glu52, Asp87, His155), three H₂O and the backbone of Trp15, Phe14 and Cys154 were included in the QM layer. The R-groups of Phe14 and Trp15 were kept in the low (MM) layer as DFT does not take into account Van der Waals interactions which are important for their interactions with the Met-O substrate.⁵³ It is worth noting that several QM/MM models were also tested that contained a smaller QM-region, e.g., lacking residues near the catalytic and recycling cysteinyl residues. However, these were found to not adequately describe the chemistry of the reactions that would likely occur as part of the mechanism. For model **B** (Figure 3.1B) the QM-layer included the catalytic and recycling cysteinyls (Cys13 and 154), two charged residues (Asp87 and His155), three H₂O and the backbone of Trp15 and Phe14.

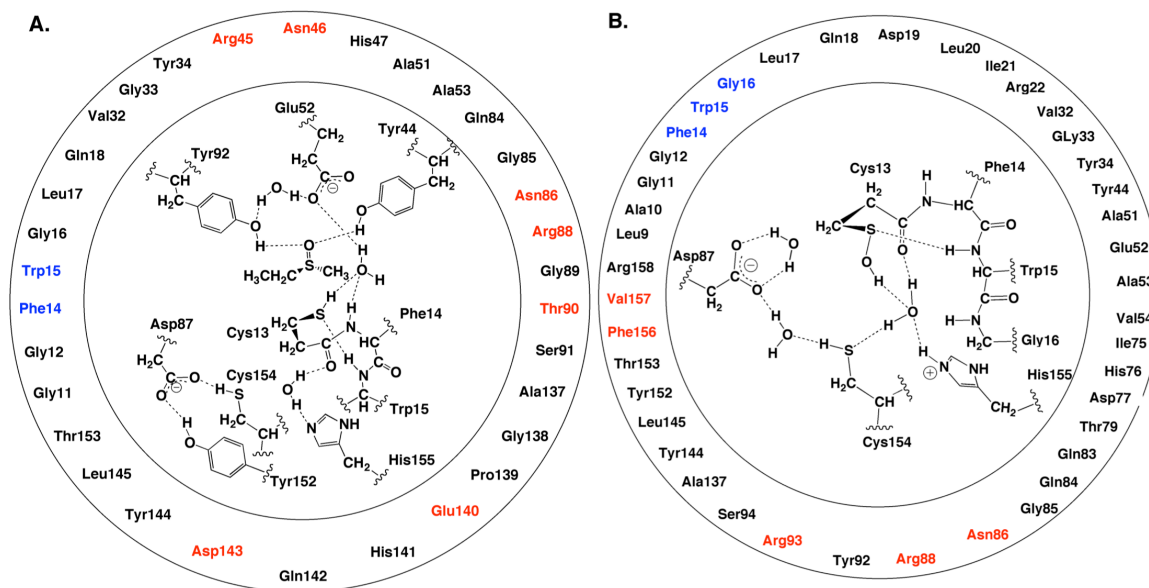


Figure 3.1. The QM/MM models used to examine the catalytic mechanism of MsrA for (A) formation of an enzyme-derived sulfenic acid intermediate from the initial enzyme-substrate complex and (B) reaction of the enzyme-derived sulfenic acid intermediate to give the final products. The inner circles represent the QM layer while the outer represent the MM layer. Color key for MM residues: included in their entirety (black); R-group replaced by hydrogen (red); R-group included in the MM layer while backbone is in the QM layer (blue).

3.3 Results and Discussion

3.3.1 Substrate binding and activation of the catalytic cysteine.

As noted in the introduction, all previously proposed mechanisms involve nucleophilic attack of the sulfur of the thiolate form of the catalytic cysteine (Cys13 in MsrA_{Mtb}) at the sulfur center (S_{sub}) of the Met-O substrate.^{18,31,33} In order for this to occur, however, Cys13 must be activated by deprotonation of its thiol. Indeed, as noted previously it has been shown experimentally²² that in the absence of the substrate the pK_a of the thiol of the catalytic cysteine is 9.5, but decreases to 5.7 upon substrate binding. X-ray crystal structures of MsrA's from different species have revealed that there is no basic

residue in close proximity to the catalytic cysteine,²² the closest being a glutamyl (Glu52 in MsrA_{Mtb}) ~ 5 Å away.³⁷ Nevertheless, Glu52 has been suggested to play a role in activating Cys13 and lowering the pK_a of its thiol.²²

In the optimized structure of the pre-reactive complex (**PRC**) in which the substrate is bound within the active site and Cys13 is neutral a water molecule forms a hydrogen bond bridge between the Cys13 thiol and the R-group carboxyl of Glu52 with distances of 2.55 and 1.67 Å, respectively (Figure 3.2). In the MD simulations on the PRC complex, a water is also observed to consistently be similarly positioned and hydrogen bonded (See Appendix Figure A3). Meanwhile, the backbone amide -NH- moiety of W15 forms a weak hydrogen bond with the Cys13SH sulfur with a length of 3.07 Å.⁵⁴ Meanwhile, the substrates sulfoxide oxygen (O_{sub}) is observed to form shorter, stronger hydrogen bonds with the phenol hydroxyls of Tyr92 and Tyr44 with lengths of 1.76 and 1.78 Å, respectively. As a result, the substrates $S_{\text{sub}}=O_{\text{sub}}$ bond lengthens markedly upon binding by 0.08 Å to 1.55 Å (See Appendix Table B1). It is noted that the distance between the sulfurs of Cys13 and the substrate, $r(\text{Cys13S}\cdots\text{S}_{\text{sub}})$, is quite long in **PRC** at 3.79 Å.

The structure and in particular hydrogen bond network within **PRC** suggests that Cys13SH may be able to transfer a proton via a bridging H_2O onto the R-group carboxylate of Glu52. At the B3LYP/6-31G(d) level of theory the proton affinities (PAs) of methylthiolate and acetate, models of the ionized R-groups of cysteine and glutamate, are calculated to lie just 34.3 kJ mol^{-1} apart, with the former being the more basic (See Appendix table B1). Indeed, Cys13SH is able to formally transfer a proton through a H_2O onto the carboxylate of the Glu residue. This step proceeds via **TS_{Act}** at a cost of just 38.5 kJ mol^{-1} (Figure 3.3). The slightly lower relative energy of **TS_{Act}** (10.8 kJ mol^{-1}) with respect to **RC** is an artifact of the use of free energy corrections and indicates that at 298 K **RC** can effectively convert back to **PRC** without a barrier. It indicates that the reverse reaction, i.e., proton transfer from Glu51COOH via a water to Cys13S⁻, effectively occurs without a barrier. As can be seen in the structure of **TS_{Act}** (Figure 3.2) the thiol's

proton lies approximately midway between the sulfur of Cys13 and the water's oxygen. Meanwhile, the water has almost wholly transferred its proton onto the glutamyl's carboxylate, $r(\text{Glu52COO}^- \cdots \text{H-OH}) = 1.08 \text{ \AA}$. This process may be facilitated by the moderately strong hydrogen bond between the H_2O and the backbone $-\text{NH}-$ of Phe14 in **PRC** and **TS_{Act}** which would help stabilize negative charge build-up on the water's oxygen and thus enhance the water's acidity. It should be noted that possible alternate mechanisms were also examined involving proton transfer from Cys13SH via a H_2O (i) directly onto the sulfoxide oxygen of the Met-O substrate and, (ii) to the glutamyl with concomitant nucleophilic attack of Cys13S^- at the substrate's sulfur centre. However, both mechanisms were found to be enzymatically unfeasible.

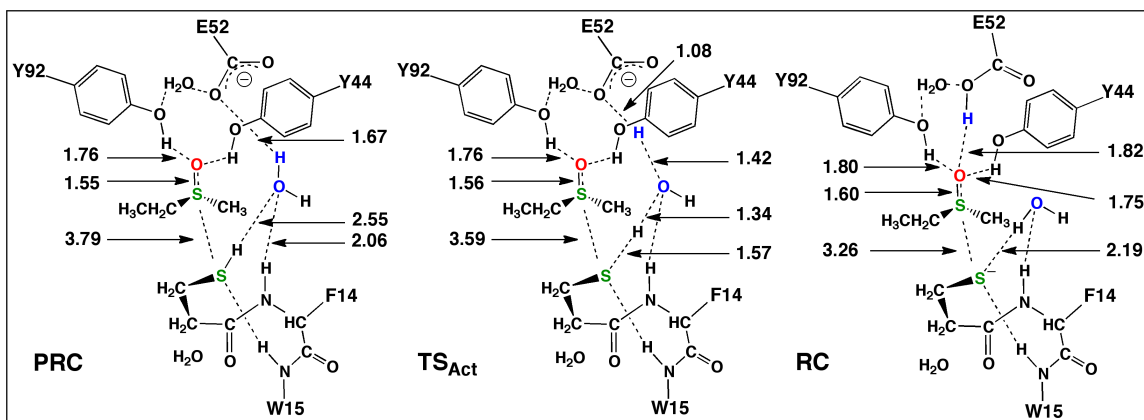


Figure 3.2. Schematic illustration of the optimized structures of the pre-reactive complex (**PRC**), activated reactive complex (**RC**) and the transition structure (**TS_{Act}**) for their interconversion. For simplicity, only selected residues, functional groups and bond lengths (Angstroms) are shown.

Two mechanistically relevant conformers of the reactive complex (**RC**) were optimized that differed only in the nature of the hydrogen bond formed by Glu52COOH. In one (**RC'**, not shown), it is hydrogen bonded via the H_2O with the now deprotonated thiol of Cys13. However, a low barrier rotation of the glutamyl's R-group about its $\text{C}_\beta-\text{C}_\gamma$ bond, i.e., a change in its $\angle\text{C}_\alpha\text{C}_\beta\text{C}_\gamma\text{O}$ angle, from -86.6° to -62.4° gives rise to the

alternate conformer **RC** shown in Figure 3.3 lying slightly higher in energy than **RC'** by 3.7 kJ mol⁻¹ and just 49.3 kJ mol⁻¹ higher in energy than **PRC**. In **RC**, the Glu52COOH group is now directly hydrogen bonded to the substrate's sulfoxide oxygen with $r(\text{Glu52COOH}\cdots\text{O}_{\text{subS}}) = 1.82 \text{ \AA}$ (Figure 3.2). The two phenol hydroxyl's (Tyr92OH and Tyr44OH) also remain strongly hydrogen bonded to the substrate's sulfoxide oxygen with distances of 1.80 and 1.75 Å, respectively. As a result, the sulfoxide's S_{sub}=O_{sub} bond has now elongated further to 1.60 Å, increasing the negative charge on O_{sub} and enhancing the susceptibility of S_{sub} to nucleophilic attack. In addition, in **RC** the distance between the sulfur's of Cys13 and Met-O has decreased significantly by 0.53 to 3.26 Å, such that they are now more suitably positioned for reaction.

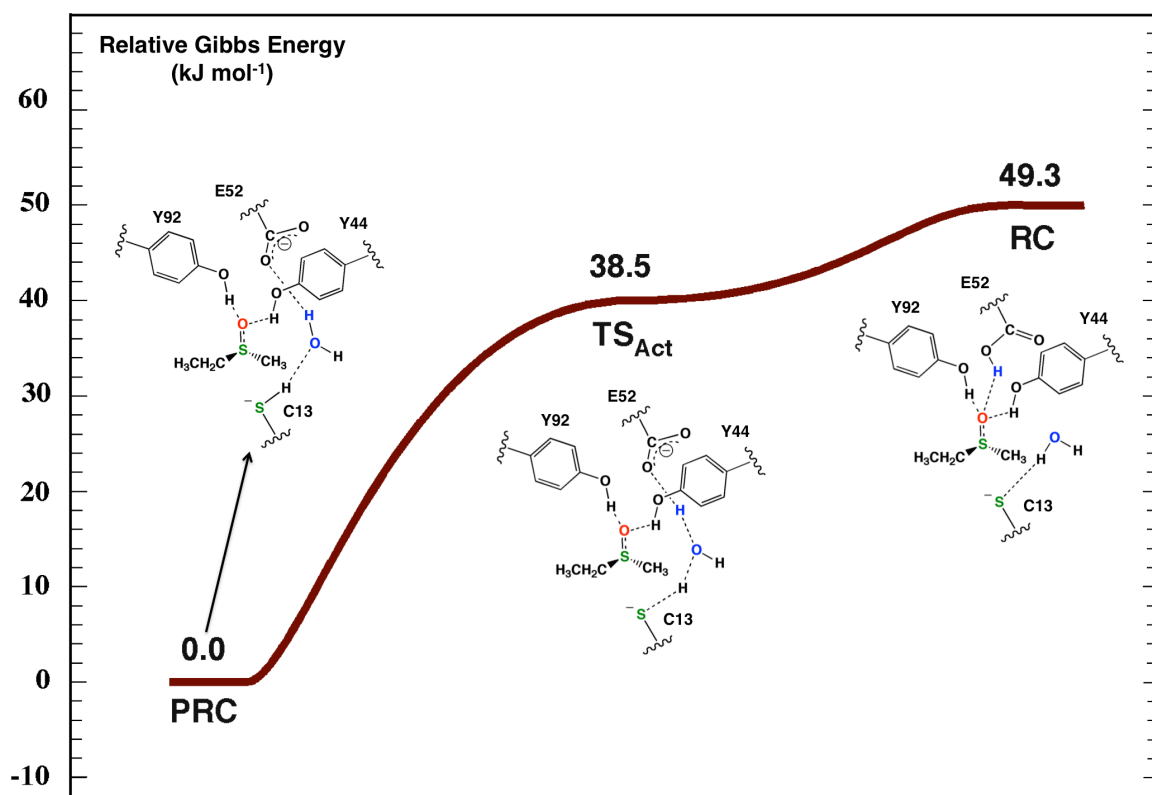


Figure 3.3. Free energy surface (kJ mol⁻¹) obtained (see Computational Methods) for activation of the catalytic cysteine (Cys13) by proton transfer via a bridging H₂O to the active site glutamate (Glu52).

3.3.2 Reduction of the substrate with formation of a sulfenic acid intermediate.

Following activation of the catalytic cysteine (Cys13) via proton transfer to the carboxylate of the glutamate (Glu52) residue, the latter is then able to transfer the proton onto the substrate's sulfoxide oxygen. Concomitantly, the now anionic sulfur of Cys13 nucleophilically attacks the substrate's sulfoxide sulfur center thus generating a sulfurane. This step proceeds via the transition structure **TS1** with a quite low activation energy of only 11.8 kJ mol^{-1} while the sulfurane intermediate formed, **IC1**, lies decidedly lower in energy than **RC** by 47.0 kJ mol^{-1} (Figure 3.4).

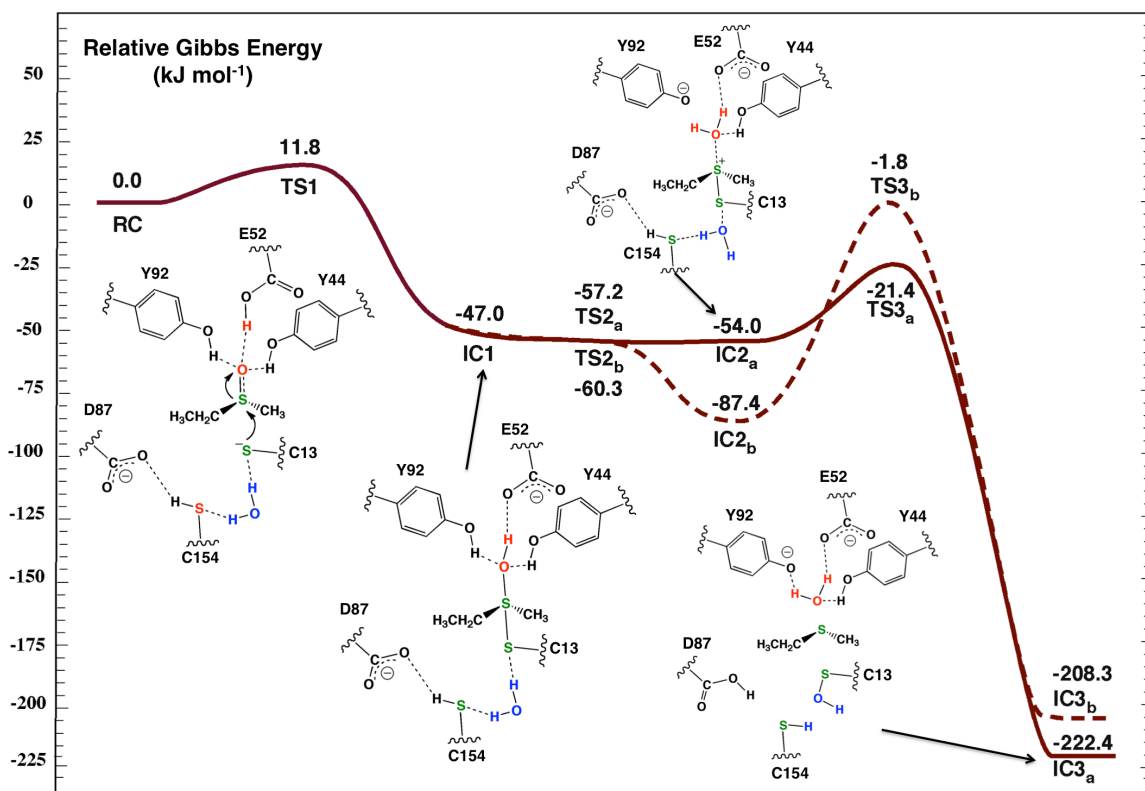


Figure 3.4. Free energy surface (kJ mol^{-1}) obtained (see Computational Methods) for formation of a sulfenic acid intermediate (Cys13SOH) involving the proton donors (A) Glu52COOH and Tyr92OH and (B) Glu52COOH and Tyr44OH.

The optimized structures of the reactant and product complex, intermediates and transition structures obtained, with selected bond and interaction distances in angstroms (Å), are schematically illustrated in Figure 3.5. As can be seen, in **TS1** the proton transferring from the Glu52COOH group is roughly equidistant to the substrates sulfoxide oxygen (O_{sub}) with $r(\text{Glu52COO}\cdots\text{H}^+)$ and $r(O_{\text{sub}}\cdots\text{H}^+) = 1.23$ and 1.22 Å respectively. More importantly, however, the substrates S—O bond has lengthened markedly to 1.74 Å while the key Cys13S \cdots S $_{\text{sub}}$ distance has shortened significantly by 0.51 Å to 2.75 Å. In the sulfurane intermediate **IC1** the S—O bond has lengthened even further to 2.31 Å while the Cys13S—S $_{\text{sub}}$ disulfide bond has now formed as indicated by its length of 2.23 Å. It is noted that these bond lengths are in reasonable agreement with a sulfurane intermediate we obtained previously^{35a} as part of a computational study on the mechanism of an MsrB using a DFT-large active site chemical model approach. In contrast, the S—OH and S—S lengths are considerably longer and shorter respectively than obtained by Thiriot et al.³⁴ in a computational study using more modest active site models. This suggests that the structure of the active site and its environment in MsrA and Msr's in general may act to help to stabilize a more polarized sulfurane intermediate.

In **IC1** the sulfurane's oxygen now forms even shorter and stronger hydrogen bonds with the phenol hydroxyls of both Tyr44 and Tyr92 compared to that observed in **RC** with distances of 1.60 and 1.56 Å, respectively (Figure 3.5). As noted previously, experimental²² mutation studies involving the two active site tyrosyl's suggest that substitution of either alone reduces the rate of the reaction only slightly while mutation of both has a severe negative effect on the rate. Previous computational studies^{34,35a} have suggested that the sulfurane hydroxyl (S $_{\text{sub}}$ -O $_{\text{sub}}$ H) can accept a proton from an acidic residue. This is likely facilitated in part by the sulfurane's oxygen O_{sub} greater basicity^{35a} than that of the sulfoxide oxygen in Met-O, which is likely further enhanced by the sulfurane's polarized nature within the active site. Thiriot et al.³⁴ suggested that one of the active site tyrosyl's may be able to act as such an acid and protonate the sulfurane's –

O_{sub}H oxygen. Given the similar distances for the hydrogen bond interactions between O_{sub} and Tyr44OH and Tyr92OH (Figure 3.5), we considered the feasibility of either tyrosyl acting as the second mechanistic acid.

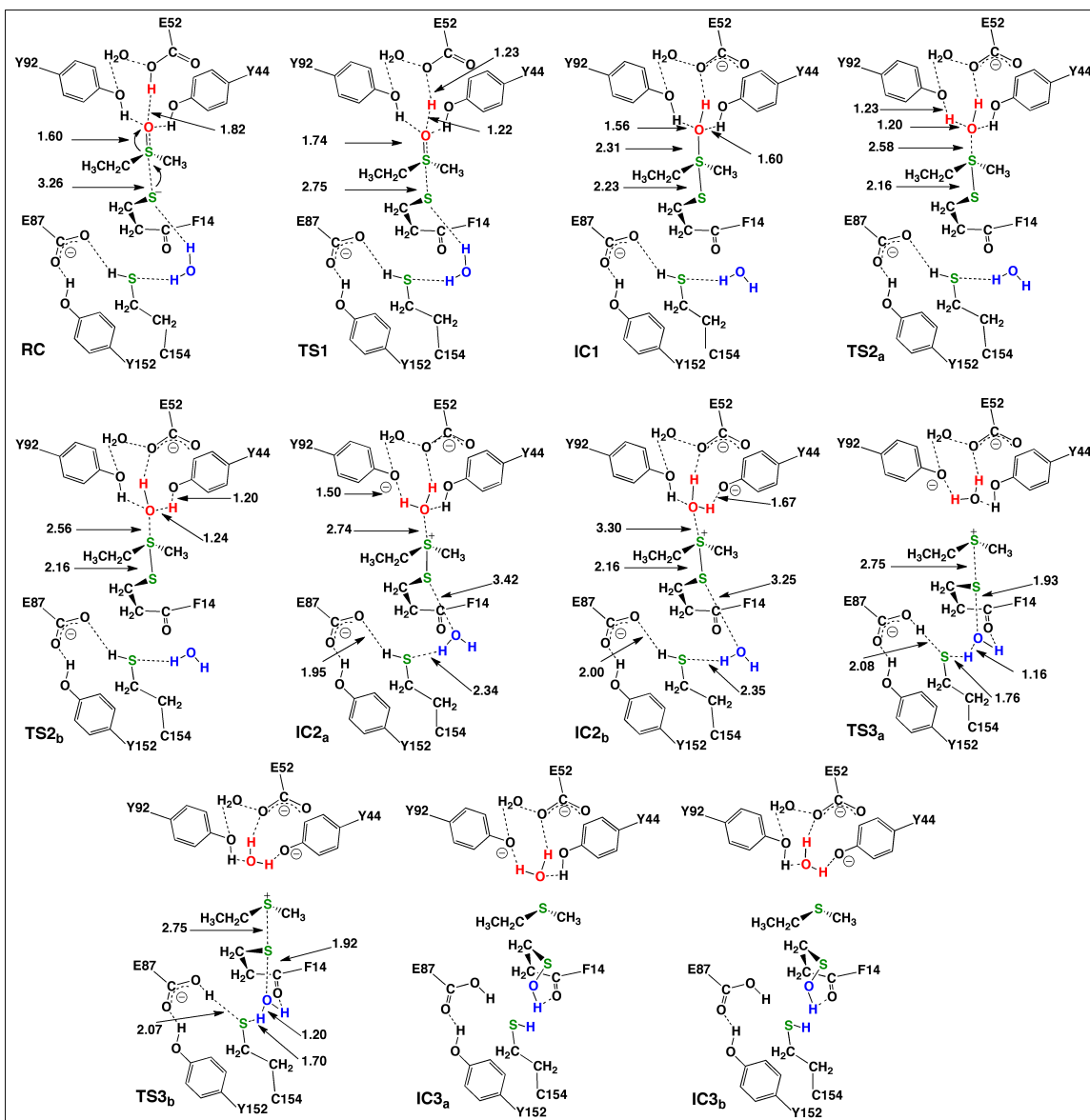


Figure 3.5. Schematic illustration of the optimized structures with selected bond lengths (in Angstroms) of the minima and transition structures obtained for MsrA catalysed reduction of Met-O to Met with concomitant formation of an enzyme-derived sulfenic acid intermediate.

The phenol hydroxyl of Tyr92 is able to transfer its proton essentially without a barrier onto the sulfurane's oxygen O_{sub} via **TS2_a** to give the sulfonium cation intermediate **IC2_a** (Figure 3.4). The latter lies slightly lower in energy than **IC1** by 7.0 kJ mol⁻¹ (i.e., -54.0 kJ mol⁻¹ with respect to **RC**). The fact that **TS2_a** (upon adding free energy corrections) is lower in energy than **IC1** and **IC2_a** suggests that these intermediates should be able to readily interconvert. In **TS2_a** the proton is almost equidistant between the sulfurane and tyrosyl oxygen's as indicated by the fact that the $r(\text{Tyr92O}\cdots\text{H})$ and $r(\text{Tyr92OH}\cdots\text{O}_{\text{sub}})$ distances are 1.23 and 1.20 Å, respectively (Figure 3.5). In the sulfonium complex **IC2_a** the $S_{\text{sub}}\cdots\text{OH}_2$ interaction has been effectively cleaved as indicated by its distance of 2.74 Å, while the Cys13S— S_{sub} disulfide bond has shortened to 2.13 Å (Figure 3.5).

Alternatively, Tyr44 can act as the second mechanistic acid. More specifically, it is able to transfer its phenolic proton onto the sulfurane's oxygen O_{sub} via **TS2_b** without a barrier (Figure 3.4). In contrast to that observed when Tyr92 acts as an acid, the resulting sulfonium complex formed, **IC2_b**, is calculated to lie decidedly lower in energy than **IC1** by 40.4 kJ mol⁻¹ (i.e., -87.4 kJ mol⁻¹ with respect to **RC**). This greater stabilization of the sulfonium may reflect in part enhanced stabilization of the positive charge on the Cys13S— S_{sub} disulfide by the now anionic phenol oxygen of Tyr44 at a distance of 2.80 Å. In contrast, the $S_{\text{sub}}\cdots\text{OH}_2$ distance in **IC2_b** is markedly longer than observed in **IC2_a** with a length of 3.30 Å (Figure 3.5).

Based in part on a previous computational study by Balta et al.,⁵⁵ Thiriot et al.³⁴ suggested that the sulfurane, upon protonation, may dissociate to a sulfonium cation and water molecule. The water moiety may then attack at the Cys13 sulfur center of the cationic disulfide to generate methionine (Met) and an enzyme-derived sulfenic acid (Cys13SOH). Indeed, for MsrA a catalytic cysteine-derived sulfenic acid has been detected in mutant and wild-type enzymes using chemical probes and mass spectrometry.²⁹ Unfortunately, they were unable to locate a pathway and barrier for this

process. Also, it has previously been suggested that the sulfurane intermediate could undergo an analogous 1,2-sigmatropic shift of its $-\text{O}_{\text{sub}}\text{H}$ group to generate Met and the sulfenic acid derivative.¹⁴ However, we previously showed that the barrier for such a process was enzymatically unfeasible.^{35a} Recently, experimental labeling studies have suggested that the oxygen of the sulfenic acid is in fact derived from the aqueous solvent and not the Met-O substrate.¹⁸ Notably, in the optimized structures of both **IC2_a** and **IC2_b** (see Figure 3.5) and in an X-ray crystal structure of the wild-type enzyme, a H_2O is positioned within the active site and oriented such that its oxygen is directed towards the Cys13S center of the sulfonium disulfide moiety. Specifically, in **IC2_a** and **IC2_b** the distance between the water oxygen (O_w) and sulfonium's Cys13 sulfur center is 3.42 and 3.25 Å respectively. Simultaneously, the water acts as a hydrogen bond donor to the sulfur of the recycling cysteine Cys154 with $r(\text{OH}_2 \cdots \text{SCys154}) = 2.34$ and 2.35 Å in **IC2_a** and **IC2_b**, respectively (Figure 3.5).

However, for nucleophilic attack of the water at Cys13S to occur, the H_2O needs to be either be first deprotonated or be able to donate a proton to a suitable base. In our present model Cys154 has been modeled as neutral based on its experimentally measure $\text{p}K_a$ value of 9.5.²² Thus, in order to be able to act as a suitable base its thiol must itself be able to transfer its proton to a suitable base. Importantly, during the initial MD simulations on the initial active site-bound substrate complex (see Computational Methods) the thiol of Cys154 was observed to consistently be directly hydrogen bonded to the R-group carboxylate of a spatially adjacent aspartyl (Asp87). In the QM/MM optimized structures of both **IC2_a** and **IC2_b** the Cys154SH group forms a short and strong hydrogen bond to Asp87; $r(\text{Cys154SH} \cdots \text{OOC-Asp87}) = 1.95$ Å. Consequently, the Cys154S—H bond itself is elongated now having a length of 1.37 Å in both complexes compared to 1.33 Å (at the B3LYP/6-31G(d) level of theory) in an isolated methylthiol. Thus, it appears that Cys154 may indeed be able to act as a suitable base to facilitate sulfenic acid formation. It is noted that within the active site region the only other residue

that may be able to act as base is His155. However, the distance between the nearest nitrogen of its R-group imidazole and H₂O is quite large (~4.85 Å). Hence, it would seem less likely to be a suitable base to facilitate sulfenic acid formation.

In the present study, for both situations in which Tyr92 or Tyr44 acts as the second mechanistic acid, the H₂O is able directly attack at the Cys13 sulfur center of the sulfonium's disulfide bond. Importantly, this occurs with concomitant proton transfer from the attacking H₂O onto the thiol of Cys154 which itself has transferred its proton to the adjacent R-group carboxylate of Asp87. For the case in which Tyr92 acted as an acid, this step proceeds via **TS3_a** at a cost of 32.6 kJ mol⁻¹ with respect to **IC2_a**. In the resulting intermediate complex obtained, **IC3_a**, the desired methionine (modeled by ethylmethylsulfide) product is bound within the active site. In particular, it is positioned such that its methyl is directed towards the R-group of the active site tryptophanyl residue (Trp15). Furthermore, however, the catalytic cysteine (Cys13) has now been oxidized to a sulfenic acid (Cys13SOH). Thermodynamically, the complex **IC3_a** lies significantly lower in energy than **IC2_a** by 168.4 kJ mol⁻¹ which is -222.4 kJ mol⁻¹ with respect to **RC**. Structurally, in **TS3_a** it can be seen that the Cys154SH proton is essentially wholly transferred onto the carboxylate of Asp87, $r(\text{Cys154S}\cdots\text{H}^+) = 2.08 \text{ \AA}$ (Figure 3.5). In contrast, the proton being abstracted from the attacking H₂O lies between the Cys154S and O_w centers with distances of 1.76 and 1.16 Å respectively (Figure 3.5). In contrast, for the pathway in which Tyr44 acted as an acid this process occurs via **TS3_b** with a markedly higher cost of 85.6 kJ mol⁻¹ with respect to **IC2_b** (Figure 3.4). In fact, the relative energy of **TS3_b** is higher than that of **TS3_a** by 19.6 kJ mol⁻¹. As obtained for **IC3_a**, in the resulting sulfenic acid-containing active site complex **IC3_b**, the desired ethylmethylsulfide product is again positioned within the active site such that its methyl is oriented towards the R-group of Trp15. In addition, **IC3_b** also lies considerably lower in energy than **IC2_b** and **RC**, though now by 120.9 and 208.3 kJ mol⁻¹, respectively. Notably, while **IC3_b** lies slightly higher in energy than **IC3_a** by 14.1 kJ mol⁻¹ (see Figure

3.5), they can interconvert with each other. This could occur by a proton transfer via the H₂O that is simultaneously hydrogen bonded to the carboxylate of Glu52 and the phenolic oxygen's of both Tyr92 and Tyr44 (see Figure 3.5). Structurally, **TS3_b** is very similar to **TS3_a** and hence is not discussed in detail herein.

We have also considered the catalytic pathway from **RC** to **IC2** in which either Tyr92 or 44 is substituted by a phenylalanyl (i.e., for the mutated enzymes Y92F or Y44F). Structurally, mutating either tyrosyl leads to a strengthening of the Met-O S—O bond (i.e., decreased polarization of the S—O bond via hydrogen bonds with the Tyr-OH groups) and concomitantly longer —S—S— bond in the sulfurane (**IC1**). Furthermore, a slight reduction is observed in the barrier for formation of **IC1**; specifically, it decreases from 14.6 kJ mol⁻¹ in the wild-type to 12.8 and 4.6 kJ mol⁻¹ in the Y44F and Y92F mutated enzymes, respectively (See Appendix Figure A4). It is noted that the magnitude of these energy changes are within the expected error margins of our calculations. Furthermore, these mutations lead to similar destabilizations of the resulting sulfurane (**IC1**) from -52.1 kJ mol⁻¹ (with respect to **RC**) to -21.0 and 22.4 kJ mol⁻¹, respectively. The subsequent sulfonium intermediates (**IC2**) are also destabilized, although now with a marked difference in magnitude between the two mutations, as indicated by the increase in its relative energy from -53.7 kJ mol⁻¹ (wild-type) to -51.8 (Y92F) and -28.7 (Y44F) kJ mol⁻¹. It is noted that this change in the relative energies of the subsequent sulfonium intermediates has little effect on the barrier for proton transfer from the remaining tyrosyl to the sulfurane's oxygen. For instance in the wild-type system the barrier for reaction via **TS2a** is 0.9 kJ mol⁻¹ while in Y92F and Y44F it is 0.8 and 3.0 kJ mol⁻¹, respectively. Importantly, however, assuming that the tyrosyl's have no effect on the activation of Cys13 (proton transfer from Cys13-SH via a H₂O to Glu52COO⁻), the rate-limiting step in the reductase stage remains formation of the sulfurane from **RC** (See Appendix Figure A5).

The above results may also give insights into observations from experimental²² mutagenesis studies in which the substitution of Tyr92 or Tyr44 (Tyr134 or Tyr82 in *Neisseria meningitidis*) by Phe was shown to have little effect on the catalytic rate. In particular, the experimentally measured difference in rates between the two mutations corresponded to a difference in rate-limiting barrier heights of $\sim 6 \text{ kJ mol}^{-1}$. The results of the present study emphasize the ability of either active site tyrosyl to act as a mechanistic acid. Furthermore, they also suggest only minor changes within the error margins of the present calculations in the thermodynamics of formation of the enzyme-sulfenic acid intermediate upon mutation of either tyrosyl.

It should be noted that alternative mechanistic pathways were also considered. In particular, the direct nucleophilic attack of the sulfur of the recycling cysteine (Cys154) at the catalytic cysteine's (Cys13) sulfur center was investigated. That is, the direct reaction of the sulfonium intermediate to give the final product complex (**PC**), bound Met with formation of an intramolecular Cys13S—SCys154 disulfide bond (see Scheme 3.1B), without formation of a sulfenic acid intermediate (see below). Previous computational studies have shown that if a thiolate is suitably positioned it can readily react with a sulfonium to give a disulfide, e.g., CysS—CysS, and an alkylated sulfide, e.g., Met.⁵⁶ However, in the sulfonium complexes **IC2_a** and **IC2_b**, the sulfurs of the two active site cysteinyl residues are at least 5.1 Å apart and with a Cys13C_α—C_αCys154 distance of 7.14 Å. Thus, direct nucleophilic attack of the thiol of Cys154 at the Cys13S center in either sulfonium complex appears unlikely to occur. It is noted that larger distances between the catalytic and recycling cysteines has been observed in the X-ray crystal structures of MsrA's from several species. For example, for an MsrA from *E. coli* the distance between the C_α carbons of the catalytic and recycling cysteinyl residues is approximately 11.0 Å!³⁷ However, if the Cys13S—S_{sub} bond was first cleaved, Cys13 would then be free to undergo a conformational change that reorients its sulfur closer to that of the recycling cysteine.

3.3.3 Reduction of the sulfenic acid with formation of an intramolecular disulfide bond

As noted for both of the above **IC3** complexes (**IC3_a** and **IC3_b**), the initial Met-O substrate has now been reduced to Met. Importantly, however, with cleavage of the sulfonium's Cys13S—S_{sub} bond and formation of the sulfenic acid Cys13SOH, the Met moiety is now free to leave the active site. This would likely enable water(s) to enter the active site, the anionic tyrosyl residue (Tyr44 or Tyr92) to regain a proton and may result in some changes in the hydrogen bonding network and possibly some repositioning of the active site residues. Hence, as detailed in the Computational Methods, in order to better model such changes an MD simulation was performed on a "resolvated Met-free" Cys13SOH-containing active site complex. The QM/MM optimized structure obtained at the same level of theory as above and hereafter referred to as **IC3'**, is shown in Figure 3.7.

In **IC3'**, both Tyr92 and Tyr44 are modeled as neutral while the R-group carboxylate of Asp87 is modeled as anionic. That is, the two former residues have regained or kept their proton while the latter has been deprotonated. It is noted that in both the average structure obtained from MD simulations and in the QM/MM optimized structure, a chain of waters was observed to interconnect the R-groups of Asp87 and the two tyrosyl residues. It thus appears possible that regeneration of the neutral tyrosyl residues could occur via proton transfer along a water chain from the Asp87COOH moiety generated during or possibly after sulfenic acid formation (See Appendix Figure A6).

Structurally, in **IC3'** the hydroxyl of the Cys13SOH sulfenic acid acts as a hydrogen bond donor to an active site H₂O that is itself simultaneously hydrogen bonded to the backbone carbonyl oxygen of Cys13, the R-group imidazole of His155 and the sulfur of Cys154. The thiol of Cys154 is also hydrogen bonded via another water molecule to the R-group carboxylate of Asp87. Due in part to this hydrogen bond network arrangement the hydroxyl of the Cys13SOH moiety sits approximately between the sulfur's of Cys13

and Cys154. Consequently, these two mechanistically important sulfur centers are not suitably positioned for the required nucleophilic attack of the Cys154S center at the sulfenic's acid sulfur to form an intramolecular Cys154S—SCys13 disulfide bond. Therefore, a rotation about the C—SOH bond of Cys13SOH (i.e., a change in its dihedral angle $\angle C_{\alpha}C_{\beta}SO$) with a slight alteration to the active site hydrogen bonding network is required in order to suitably position the sulfurs of Cys13S and Cys154 for further reaction. More specifically, a change in $\angle C_{\alpha}C_{\beta}SO$ from 243.9° to 110.0° leads to formation of the alternate conformer **IC3''** lying just slightly higher in energy than **IC3'** by 14.0 kJ mol^{-1} (Figure 3.6). As can be seen in Figure 3.7, in **IC3''** the Cys13SOH hydroxyl now acts as a hydrogen bond donor with a water that is itself hydrogen bonded to the R-group carboxylate of Asp87. That is, the sulfenic acid is no longer hydrogen bonded via a water molecule with the thiol of Cys154. Importantly, as a result of this rotation, the sulfurs of the catalytic and recycling cysteine are now better positioned with respect to each other for reaction and are closer than in **IC3'** by 0.39 \AA with a Cys154S \cdots SCys13 distance of 4.54 \AA . Furthermore, the sulfenic acids S—OH bond has lengthened slightly from 1.69 \AA (**IC3'**) to 1.72 \AA . Unfortunately, within the present computational model we were unable to optimize a transition structure (**TS4**) for interconversion of **IC3'** and **IC3''**. However, for isolated ethanesulfenic acid ($\text{CH}_3\text{CH}_2\text{SOH}$) at the B3LYP/6-31G(d) level of theory the analogous rotational barrier is very low at around 4.0 kJ mol^{-1} . Thus, while the active site sterics and the hydrogen bond network may increase the barrier for this step, it is still unlikely to be significant or rate-limiting. It is also important to note that sulfenic acid orientation in **IC3''** is in agreement with a sulfenic acid-containing X-ray crystal structure of MsrA from *N. meningitidis* (PDB ID: 3BQG).³³ Furthermore, two water molecules were also found to be positioned in the crystal structure such that they may be able to form a hydrogen bond bridge between the sulfenic acid's —SOH oxygen and a carboxylate oxygen of Asp87. In addition, a third water molecule was positioned such that it may form a hydrogen bond

bridge between the sulfenic acid's sulfur and the R-group of His155. However, the X-ray crystal structure as well as our MD structure also shows a chain of waters connecting the sulfenic acid oxygen to Glu52, Tyr44 and Tyr92, suggesting the possibility that one or more of these residues might also be able to play a role as proton donor's in the disulfide formation step.

The nucleophilicity of the Cys154S sulfur center, and hence its ability to attack at the sulfur of Cys13, would be facilitated by deprotonation of the Cys154 thiol. In the present computational model this can be achieved by proton transfer from Cys154SH via a water molecule onto the R-group carboxylate of Asp87. This step proceeds via **TS5** to give the Cys154S⁻ thiolate-containing complex **IC4** just 37.8 kJ mol⁻¹ higher in energy than **IC3**". In **IC4** the Asp87COOH group remains hydrogen bonded via a water with the Cys154S⁻ thiolate while the Cys154S...SCys13 distance has decreased further by 0.13 Å to 4.41 Å (Figure 3.6). The slightly lower relative energy of **TS5** (17.8 kJ mol⁻¹) with respect to **IC4** is an artifact of the use of single-point energy calculations and free energy corrections. It indicates that the reverse reaction, i.e., proton transfer from Asp87COOH via a water to Cys154S⁻, occurs without a barrier.

In order for the hydroxyl of the sulfenic acid to be a better leaving group for disulfide bond formation, it needs to gain a proton at some point and become H₂O. The now neutralized Asp87 residue could act as a suitable proton donor. In particular, it needs to act as a hydrogen bond donor either directly or indirectly (e.g., via water) with the Cys13SOH oxygen center. This can be achieved by rotation of its carboxylate's -OH moiety, i.e., rotation about the aspartyl's C_β-OH bond, such that it no longer hydrogen bonds indirectly with Cys154S⁻. This step proceeds via **TS6** with a barrier of 39.4 kJ mol⁻¹ with respect to **IC4**. This is only slightly higher by 28.0 kJ mol⁻¹ than that obtained for the analogous rotation in isolated propanoic acid at the B3LYP/6-31G(d) level of theory. This likely reflects in part general effects of the active site and the cleaving of the negatively charged Asp87COOH...OH₂...⁻SCys154 hydrogen bonding interaction. The

resulting alternate conformer formed, complex **IC4'**, lies 29.8 kJ mol⁻¹ lower in energy than **IC4**. The Cys154S...SCys13 distance is only marginally affected by this rotation being 4.43 Å in **IC4'**, while the sulfenic acids S—OH bond has lengthened marginally to 1.74 Å (Figure 3.7).

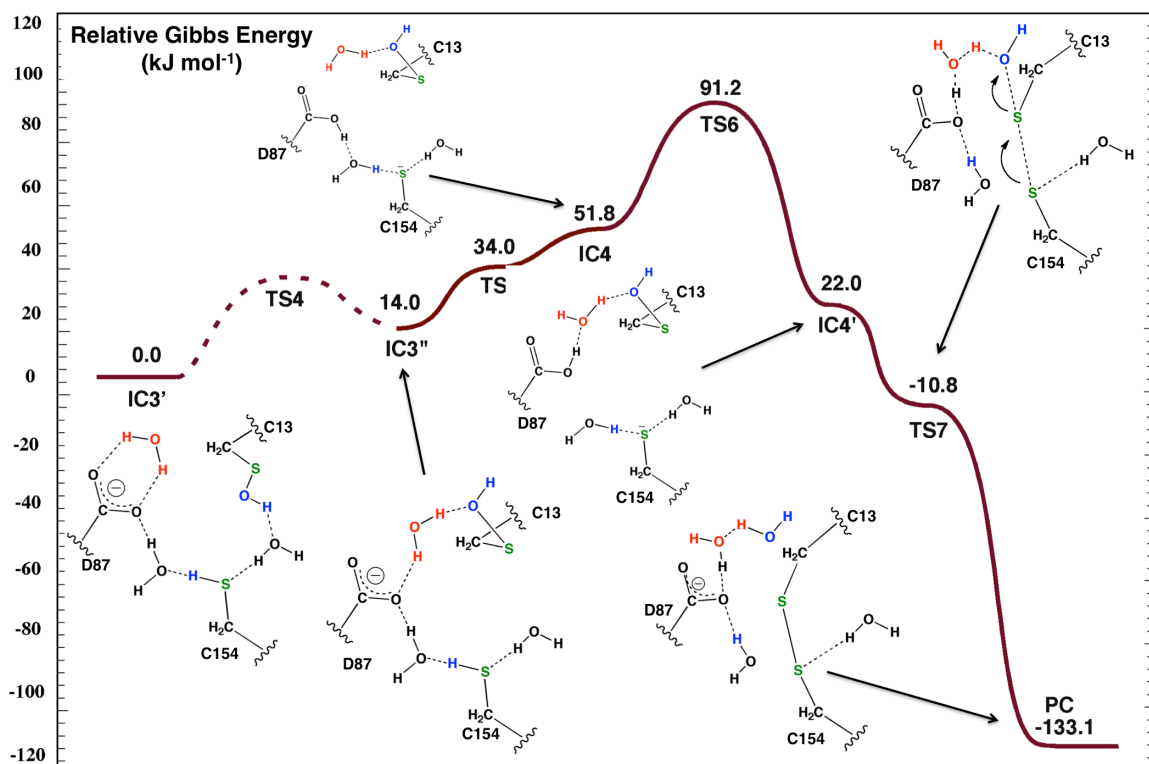


Figure 3.6. Free energy surface (kJ mol⁻¹) obtained (see Computational Methods) for the reduction of sulfenic acid-containing complex to give an intramolecular disulfide bond with formation of H₂O.

The final step of the overall mechanism is nucleophilic attack of the anionic sulfur of Cys154 at the sulfur center of the Cys13SOH sulfenic acid to give a Cys154S—SCys13 disulfide bond, and proceeds via **TS7**. As illustrated in Figure 3.7, formation of the disulfide bond occurs with concomitant transfer of the proton from Asp87COOH via a bridging water onto the leaving —OH group of the sulfenic acid to give a H₂O molecule. It is noted that this reaction has been computationally investigated previously using density

functional theory in combination with small chemical models.⁵⁶ Similar to the present results it was also found that disulfide bond formation involved an S_N2 mechanism in which S—S bond formation occurred with concomitant proton transfer onto the leaving –OH group.⁵⁶ Within the present extensive QM/MM approach, however, the relative energy of **TS7** is lower than that of **IC4'** (Figure. 3.6). This suggests that upon rearrangement of the Asp87COOH···OH₂···HOSCys13 hydrogen bond network within the active site environment of MsrA, the reduction of the sulfenic acid to give a disulfide bond can effectively occur without a barrier. The resulting final product complex (**PC**) formed lies considerably lower in energy than **IC3'** by 133.1 kJ mol⁻¹. The overall highly exothermic mechanism for sulfenic acid formation and its subsequent reduction supports the common experimental observation of the high reactivity of sulfenic acid and challenges in detecting its occurrence in the presence of the recycling cysteine.

It is noted that within the present computational model the rate-limiting step of the overall mechanism appears to occur after formation of the sulfenic acid and corresponds to rotation within the Asp87COOH moiety.⁵⁷ However, this approach is necessarily 'static' and does not include free energy corrections. In contrast, the hydrogen bond network within the active site is likely quite dynamic. Thus, the presently calculated results for this stage of the overall mechanism likely represent an upper thermodynamic value. Indeed, the approximate rate-limiting barrier obtained from experimental kinetic studies on the wild-type enzyme in the absence of Trx is 56.6 kJ mol⁻¹.³⁰ This is in good agreement with the rate-limiting barriers presently calculated for the two possible pathways for sulfenic acid (**IC3**) formation from **PRC**: 61.1 and 85.6 kJ mol⁻¹ in the case of Tyr92 and Tyr44 acting as the second mechanistic acid, respectively.

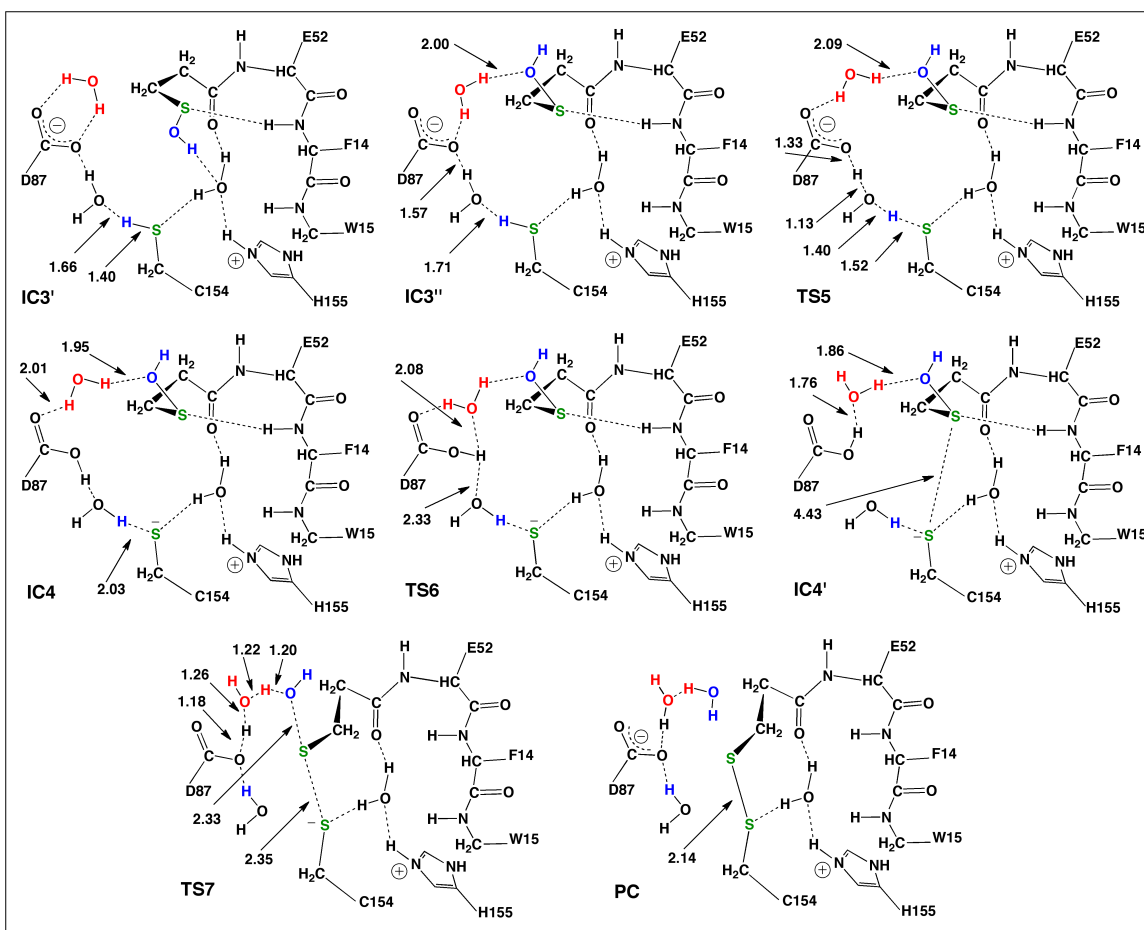


Figure 3.7. Schematic illustration of the optimized structures for sulfenic acid reduction and formation of disulfide bond

In the mechanism described above for sulfenic acid reduction after methionine formation, the active site model used began with an initially anionic Asp87 (i.e., Asp87COO⁻) and protonated His155 (i.e., His155-H⁺). That is, these residues were in their usual charge states at pH 7. Experimentally, Gand et al.⁵⁸ have examined the kinetic effects of mutating either Asp87 or His155 and concluded that a main role of Asp87 is in substrate binding. Our above MD simulations on the pre-reactive complex are in agreement as suggested by the observation of a consistent direct interaction between Asp87 and the substrate. However, the results of Gand et al. also found that mutating either Asp87 or His155 had similar kinetic effects.⁵⁸ We examined the effect of mutating

Asp87 to Alanine. More specifically, we replaced Asp87 by Ala in the pre-reactive complex ($\text{PRC}_{\text{Asp87Ala}}$), then resolved the complex using the same procedure as detailed in the Computational Methods. This was followed by a 1 ns MD simulation and the results obtained analyzed via cluster analysis. Notably, it was observed that in the $\text{PRC}_{\text{Asp87Ala}}$ the thiol of Cys154 consistently hydrogen bonds to the imidazole of His155 either directly or via a water molecule (See Appendix Figure A6) This suggests the possibility of a Cys154-activation pathway involving His155 that is analogous to that we have previously described involving Asp87. Hence, for completeness and given that the pK_a of the R-group of His lies near 7 and can be modified by the environment, we also considered an alternate mechanism in which the imidazole of His155 is neutral and thus able to act as a base. For this mechanism the R-group of Asp87 was modeled as neutral as it could still have participated in sulfenic acid formation by accepting a proton from Cys154SH. It should be noted that in the case of His155 facilitating Cys154 activation, no rotation of the hydroxyl in Asp87COOH would be necessary (i.e., $\text{IC4} \rightarrow \text{IC4}'$). Importantly, His155 was indeed able to act as a base and accept a proton from the thiol of Cys154 in the activation of the latter residue (i.e., $\text{IC3}' \rightarrow \text{IC4}$) with a calculated reaction barrier of 65.1 kJ mol^{-1} (See Appendix Table B1). While this is enzymatically feasible, it is higher than obtained for the above mechanism via **TS5**. This suggests that mutation of Asp87 may not significantly affect sulfenic acid reduction as either basic residue near the thiol of the recycling cysteine could potentially facilitate its activation.

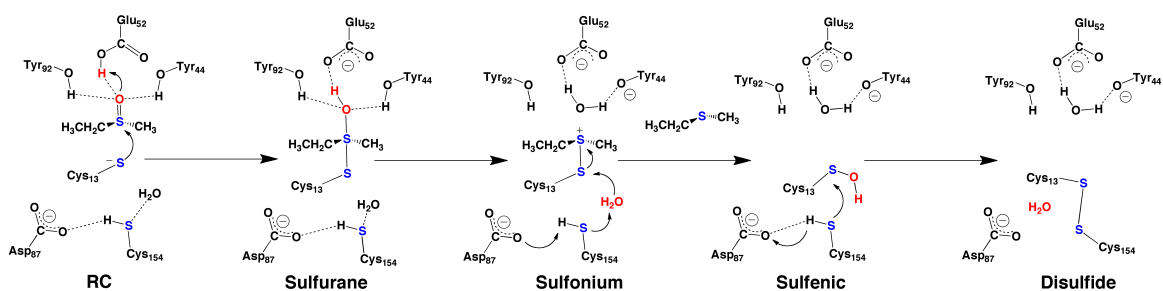


Figure 3.8. Schematic illustration of our computationally suggested mechanism.

3.4 Conclusion

The overall mechanism(s) by which methionine sulfoxide reductase A (MsrA) from *Mycobacterium tuberculosis* catalyzes the reduction of S-methionine sulfoxide (Met-O) to methionine (Met), i.e., the reductase stage, have been investigated via the complementary application of docking, molecular dynamics (MD) simulations and ONIOM (QM/MM) calculations. More specifically, docking and MD simulation were used to obtain solvated structures for the initial active site-bound substrate complex in which the catalytic cysteine (Cys13SH) is not yet activated (**PRC**), i.e., neutral. An ONIOM QM/MM approach in combination with a large active site model has been used to examine the mechanism of Cys13 activation and subsequent pathway(s) leading to formation of a sulfenic acid intermediate. In particular, Cys13SH is able to transfer its proton via a bridging water molecule onto the R-group carboxylate of the active site glutamate (Glu52) with a modest energy cost of 49.3 kJ mol^{-1} , to give the activated active site-bound substrate complex, **RC**. The now anionic Cys13S⁻ sulfur then nucleophilically attacks the sulfur of the Met-O substrate. This occurs with concomitant proton transfer from the now neutral Glu52COOH group onto the substrate sulfoxide oxygen. The resulting sulfurane intermediate is "polarized" by the active site environment. Consequently, the sulfurane's oxygen can readily accept a proton from the phenol hydroxyl of either of the active site tyrosyl residues, Tyr44 and Tyr92, with a negligible barrier or without a barrier, to give a sulfonium cation and water. The water formed is hydrogen bonded to the R-groups of both active site Tyr44 and Tyr92, and Glu52.

It was found that a water moiety was then able to directly and readily attack the sulfonium cation at its Cys13S center, i.e., the sulfur of the catalytic cysteine, to give a sulfenic acid Cys13SOH and methionine in one step. In the active site chemical model used, the recycling cysteine (Cys154) was neutral, i.e., Cys154SH. It is found that a neutral Cys154SH group is able to facilitate sulfenic acid formation by accepting a proton from the attacking H₂O while simultaneously transferring its proton via a water to the

nearby R-group carboxylate of Asp87. For the case in which Tyr92 acted as the second mechanistic acid this step occurs with a barrier of 32.6 while for the alternate case in which Tyr44 acted as an acid the barrier is higher but still enzymatically feasible at 85.6 kJ mol⁻¹. The resulting sulfenic acid intermediate complexes formed, **IC3_a** and **IC3_b** respectively, are calculated to lie significantly lower in energy than RC by -222.4 and -208.3 kJ mol⁻¹.

Reduction of the sulfenic acid to give an intramolecular Cys154S—SCys13 disulfide bond, with formation of H₂O, was found to occur via series of low barrier steps. Primarily, these steps involve the rearrangement of the active site hydrogen bond network and suitable positioning of the sulfur centers of Cys154 and Cys13 for reaction. It is found that nucleophilic attack by the sulfur of the neutral thiol of Cys154SH is facilitated by the R-group carboxylate of Asp87. In particular, the latter is able to accept the thiol proton and then transfer it via a H₂O onto the leaving -OH group of sulfenic acid (Cys13SOH) concomitant with formation of the disulfide bond. Further, our results suggest an S_N2 mechanism for disulfide bond formation, and that it would happen spontaneously after activation of Cys154 and suitable positioning of the Cys13SOH sulfur for nucleophilic attack by the thiolate sulfur of Cys154.

Thus, the present results also suggest that both active site cysteines can at least initially be neutral. During the course of the reaction they can be activated for formation of the sulfurane or reduction of the sulfenic acid via direct or indirect proton transfers to, for instance, the R-groups of some active site residues including Glu52 and Asp87.

3.5 References

- (1) Stadtman, E. R.; Moskovitz, J.; Levine, R. L. *Antioxid. Redox Signal.* **2003**, *5*, 577.
- (2) Koc, A.; Gladyshev, V. N. You have full text access to this contentMethionine Sulfoxide Reduction and the Aging Process. In *Biogerontology*:

Mechanisms and Interventions; Rattan, S. I. S., Akman, S., Eds.; Blackwell Publishing: Oxford, UK. 2007; Vol. 1100, p 383.

- (3) Levine, R. L.; Moskovitz, J.; Stadtman, E. R. *IUBMB Life* **2000**, *50*, 301.
- (4) Levine, R. L.; Mosoni, L.; Berlett, B. S.; Stadtman, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 15036.
- (5) Levine, R. L.; Berlett, B. S.; Moskovitz, J.; Mosoni, L.; Stadtman, E. R. *Mech. Ageing Dev.* **1999**, *107*, 323.
- (6) Agbas, A.; Moskovitz, J. *Curr. Signal. Transd. T.* **2009**, *4*, 46.
- (7) Haenold, R.; Wassef, D. M.; Heinemann, S. H.; Hoshi, T. *Age* **2005**, *27*, 183.
- (8) Stadtman, E. R. *Curr. Med. Chem.* **2004**, *11*, 1105.
- (9) Schöneich, C. *Biochim. Biophys. Acta* **2005**, *1703*, 111.
- (10) Ezraty, B.; Aussel, L.; Barras, F. *Biochim. Biophys. Acta* **2005**, *1703*, 221.
- (11) Weissbach, H.; Resnick, L.; Brot, N. *Biochim. Biophys. Acta* **2005**, *1703*, 203.
- (12) Trachootham, D.; Lu, W. Q.; Ogasawara, M. A.; Valle, N. R. D.; Huang, P. *Antioxid. Redox Signal.* **2008**, *10*, 1343.
- (13) Moskovitz, J. *Biochim. Biophys. Acta, Proteins Proteomics* **2005**, *1703*, 213.
- (14) Lowther, W. T.; Brot, N.; Weissbach, H.; Matthews, B. W. *Biochemistry* **2000**, *39*, 13307.
- (15) Chen, B. W.; Markillie, L. M.; Xiong, Y. J.; Mayer, M. U.; Squier, T. C. *Biochemistry* **2007**, *46*, 14153.
- (16) Kim, Y. K.; Shin, Y. J.; Lee, W. H.; Kim, H. Y.; Hwang, K. Y. *Mol. Microbiol.* **2009**, *72*, 699.
- (17) Moskovitz, J.; Maiti, P.; Lopes, D. H. J.; Oien, D. B.; Attar, A.; Liu, T. Y.; Mittal, S.; Hayes, J.; Bitan, G. *Biochemistry* **2011**, *50*, 10687.

- (18) Lim, J. C.; You, Z.; Kim, G.; Levine, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10472.
- (19) Boschi-Muller, S.; Gand, A.; Branlant, G. *Arch. Biochem. Biophys.* **2008**, *474*, 266.
- (20) Lowther, W. T.; Weissbach, H.; Etienne, F.; Brot, N.; Matthews, B. W. *Nat. Struct. Biol.* **2002**, *9*, 348.
- (21) Neiers, F.; Sonkaria, S.; Olry, A.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2007**, *282*, 32397.
- (22) Antoine, M.; Gand, A.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2006**, *281*, 39062.
- (23) Singh, V. K.; Moskovitz, J. *Microbiology-(UK)* **2003**, *149*, 2739.
- (24) Douglas, T.; Daniel, D. S.; Parida, B. K.; Jagannath, C.; Dhandayuthapani, S. *J. Bacteriol.* **2004**, *186*, 3590.
- (25) Nan, C. L.; Li, Y. J.; Jean-Charles, P. Y.; Chen, G. Z.; Kreymerman, A.; Prentice, H.; Weissbach, H.; Huang, X. P. *Biochem. Biophys. Res. Commun.* **2010**, *402*, 608.
- (26) Ruan, H.; Tang, X. D.; Chen, M. L.; Joiner, M. A.; Sun, G.; Brot, N.; Weissbach, H.; Heinemann, S. H.; Iverson, L.; Wu, C. F.; Hoshi, T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2748.
- (27) Moskovitz, J.; Bar-Noy, S.; Williams, W. M.; Berlett, B. S.; Stadtman, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12920.
- (28) De Luca, A.; Sanna, F.; Sallese, M.; Ruggiero, C.; Grossi, M.; Sacchetta, P.; Rossi, C.; De Laurenzi, V.; Di Ilio, C.; Favaloro, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 18628.
- (29) Boschi-Muller, S.; Azza, S.; Sanglier-Cianferani, S.; Talfournier, F.; Van Dorsselear, A.; Branlant, G. *J. Biol. Chem.* **2000**, *275*, 35908.

- (30) Antione, M.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2003**, *278*, 45352.
- (31) Lowther, W. T.; Brot, N.; Weissbach, H.; Honek, J. F.; Matthews, B. W. *Proc. Nat. Acad. Sci. U.S.A.* **2000**, *97*, 6463.
- (32) Boschi-Muller, S.; Olry, A.; Antoine, M.; Branlant, G. *Biochim. Biophys. Acta.* **2005**, *1703*, 231.
- (33) Ranaivoson, F. M.; Antoine, M.; Kauffmann, B.; Bosch-Muller, S.; Aubry, A.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2008**, *377*, 268.
- (34) Thiriote, E.; Monard, G.; Boschi-Muller, S.; Branlant, G.; Ruiz-Lopez, M. F. *Theor. Chem. Acc.* **2011**, *129*, 93.
- (35) Robinet, J. J.; Dokainish, H. M.; Paterson, D. J.; Gauld, J. W. *J. Phys. Chem. B* **2010**, *115*, 9202.
- (36) MOE, version 2009.10 (2009) Chemical Computing Group Inc., Montreal.
- (37) Taylor, A. B.; Benglis, D. M.; Dhandayuthapani, S.; Hart, P. J. *J. Bacteriol.* **2003**, *185*, 4119.
- (38) Bushnell, E. A. C.; Erdtman, E.; Llano, J.; Eriksson, L. A.; Gauld, J. W. *J. Comput. Chem.* **2010**, *32*, 822.
- (39) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- (40) Bushnell, E. A. C.; Huang, W.; Llano, J.; Gauld, J. W. *J. Phys. Chem. B* **2012**, *116*, 5205.
- (41) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (42) Gaussian 09, Revision C.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.

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

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

(45) 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.* **1996**, *118*, 2309.

(46) Lodola, A.; Sirirak, J.; Fey, N.; Rivara, S.; Mor, M.; Mulholland, A. J. *J. Chem. Theory. Comput.* **2010**, *6*, 2948.

(47) Xu, Q.; Li, L. Y.; Guo, H. *J. Phys. Chem. B* **2010**, *114*, 10594.

(48) Xiao, C.; Zhang, Y. *J. Phys. Chem. B* **2007**, *111*, 6229.

(49) 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, USA. 2010; Vol. 2, p 643.

(50) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.

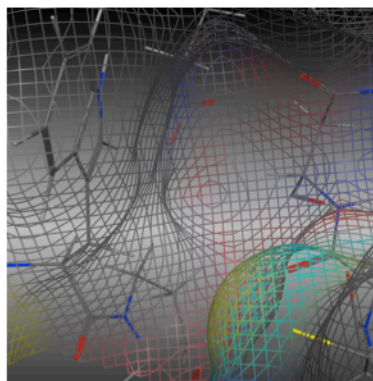
(51) Przybylski, J. L.; Wetmore, S. D. *Biochemistry* **2011**, *50*, 4218.

- (52) Tian, B. X.; Erdtman, E.; Eriksson, L. A. *J. Phys. Chem. B* **2012**, *116*, 12105.
- (53) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. T. *Chem. Rev.* **2012**, *112*, 289.
- (54) Zhou, P.; Tian, F. F.; Lv, F. L.; Shang, Z. C. *Proteins* **2009**, *76*, 151.
- (55) Balta, B.; Monard, G.; Ruiz-Lopez, M. F.; Antoine, M.; Gand, A.; Boschi-Muller, S.; Branlant, G. *J. Phys. Chem. A* **2006**, *110*, 7628.
- (56) Bayse, C. A. *Org. Biomol. Chem.* **2011**, *9*, 4748.
- (57) Laidler, K. J. *Chemical kinetics*; HarperCollinPublishers: New York, USA. 1987; 3rd ed.

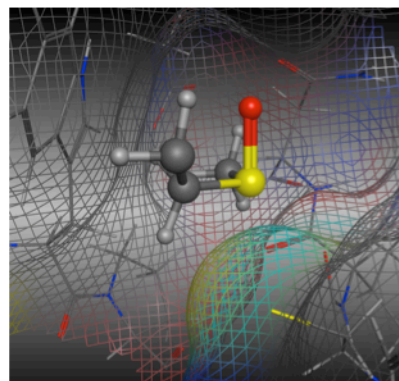
Chapter 4

Computational Approach Choice in Modeling

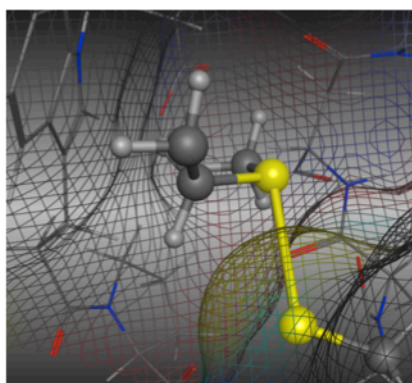
Flexible Enzyme Active Sites



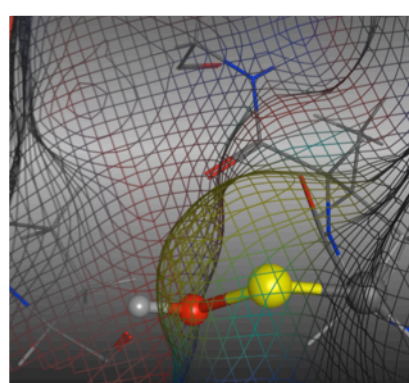
Apoenzyme



Michaelis



Sulfonium



Sulfenic

4.1 Introduction

In the last two decades, the applications of quantum chemistry have been extended from small molecules to protein and DNA.¹⁻⁴ More specifically, the introduction of density functional theory (DFT) methods as well as the vast growing of computational resources allowed for further applications in biochemistry.^{2,5} Enzymatic reactions are one of the main field where quantum chemistry have been applied successfully providing detailed information for numerous catalytic mechanisms.⁶⁻¹⁰ In order to model these proteins, there are two main approaches:^{1,3} 1) quantum cluster approach in which the active site is basically extracted from the crystal structure of the substrate analogue bound proteins; the structure is truncated including only models of the catalytic residues; mostly, the quantum model is less than 200 atoms. 2) Quantum mechanics/molecular mechanics (QM/MM) approach in which docking and molecular dynamic (MD) simulations are used to prepare the model; then the MM minimized structure is truncated up to 10-20Å from the active site and partitioned into two layers where the catalytic residues are only included in the quantum layer and the rest of the model is calculated using molecular mechanics force field. Both approaches have been successfully used to investigate several enzymatic mechanisms,¹¹⁻¹⁴ however they also have some downsides.^{1,9,15}

Earlier, we investigated in details the catalytic mechanism of Methionine sulfoxide reductase B (MsrB) using a quantum mechanical (QM)-cluster approach.¹⁶ The two previously experimentally proposed mechanisms were considered: (1) release of Met via sulfenic acid intermediate followed by nucleophilic attack of the recycling cysteine thiolate at sulfenic acid (-SOH) sulfur forming disulfide, (2) direct attack of the recycling cysteine thiolate at the sulfonium catalytic sulfur. Both mechanisms were found to be feasible! However, the direct sulfur attack was found to be energetically more favorable. More interestingly, it occurs via low barrier transition state of 23.8 kJ mol⁻¹ which is almost half the barrier for sulfenic acid formation, 54.6 kJ mol⁻¹. However, this energy difference is still small compared to computational error raising an important question, as

it is unclear if this difference is an artifact of the employed computational method or model. It is also important to mention that, there was a debate with experimentalists as they accept that a sulfenic acid intermediate would occur in the mechanism.^{17,18}

MsrB is subclasses of ubiquitous group of enzymes that play crucial roles in cellular protection mechanism against oxidative stress, Methionine sulfoxide reductase (Msr's), via reducing methionine sulfoxide (MetSO) back into methionine.^{19,20} MsrA, another subclass of the same group of enzymes, and MsrB are stereospecific to MetSO S- and R-epimers, respectively.²¹ Initially, the same reduction mechanism was proposed for both subclasses in which sulfenic acid was proposed to occur as a precursor intermediate to disulfide bond formation.^{22,23} In contrast, a direct sulfur attack on sulfonium cation intermediate mechanism was also suggested.²⁴ More importantly, sulfenic acid intermediate was only detected in wild type of MsrA.²⁵ However for MsrB, it was only detected upon mutating the second Cys.²⁶

Sulfenic acid is a fundamental key intermediate involved in redox chemistry of proteins.^{27,28} Its role in biochemistry diverge from signal transduction, non-enzymatic protein folding, protection against ROS and modulating gene transcription.²⁹⁻³⁴ It occurs either via direct oxidation of Cys by ROS, peroxynitrite and hypochlorous acid or indirectly during several enzymes catalytic reactions and xenobiotic metabolism.^{27,29,35-37} Furthermore, it has been also proposed to occur as an essential precursor intermediate for disulfide bond formation in several enzymes.^{25,31,38-40} Since sulfenic acid has both nucleophilic and electrophilic characters, it is highly reactive and unstable species.^{27,28} Therefore, once formed several fate reactions have been proposed to occur, such as overoxidation to sulfinic and sulfonic acids.^{29,41} Both reactions considered to be irreversible as their reduction require a very low pH.²⁹ Furthermore, sulfiredoxin (Srx) is the only enzyme that has been shown to reduce sulfinic acid in Prx.⁴²

In contrast, several reversible protection mechanisms have also been also proposed. For instance, Nakamura et al. suggested the formation of a hypervalent sulfur

intermediate in archaeal peroxiredoxin via forming S–N covalent bond between –SOH and neighboring His.⁴³ Likewise, it is protected in PTPs and OhrR via formation of sulfenyl-amide intermediate by forming a covalent bond with backbone nitrogen of the adjacent residue.^{44,45} More importantly, it could be stabilized and recycled via S-thiolation forming a mixed disulfide bond with Trx, Grx and GSH or using an intramolecular Cys as in Msr's.^{46,47}

Lately, we also investigated the formation and reduction mechanisms of sulfenic acid during the reduction mechanism of MetSO via MsrA in which sulfenic acid was found to form through a concerted TS via a low barrier of 30.7 kJ mol⁻¹. Furthermore, the reduction to disulfide was found to occur via several low barrier steps suggesting spontaneous reduction via an S_N2 mechanism upon the activation of second Cys and rotation of sulfenic acid to be in a proper position for nucleophilic attack explaining its high reactivity. In MsrA, no direct sulfur attack mechanism was obtained due to the large distance between the two sulfurs. More importantly, the C_s...C_s distance of the two cysteine was found to be around 11 Å in most MsrA crystal structures.⁴⁸

In contrast, in the MsrB crystal structure of *N.gonorrhoeae pilB* (PDB: 1L1D,⁴⁹ the one used in our previous DFT study), the distance between the two sulfurs is 3.29 Å. More importantly, the C_s...C_s distance is around 7.41 Å. Similar distances were also occur in other crystal structures such as in streptococcus pneumonia (PDB: 3E0M) and *Neisseria meningitidis* PilB (PDB: 3HCG) with a C_s...C_s distance of 8.29 Å and 8.47 Å, respectively.^{50,51} Earlier, Ranaivoson et al. suggested the presence of high degree of flexibility in the MsrB's second Cys containing loop, β2-β3.⁵¹ Since a distance of 13 Å were found to occur between the two C_s in *X. campestris* (PDB: 3HCI).⁵¹ This implies the need for conformational change prior to disulfide bond formation. More importantly, Trp65 lies in between the two Cys preventing the formation of disulfide bond. However, kinetically, these expected drastic conformational changes did not change the rate-limiting step. Furthermore, the C_s...C_s distance in the NMR structure of *Bacillus subtilis*

MsrB (PDB: 2KZN) is found to be 9.25 Å as well the distance between the two sulfurs is 7.46 Å.⁵² Hence, all solved crystal structures of MsrB suggest the high flexibility in the free, complex-like and sulfenic acid forms. However, the driving force for these changes are still unclear.

In this study, we reinvestigate the last step of the reductase step in MsrB via studying the effect of method selection, functional and basis sets on the catalytic mechanism. Further, we also investigate the effect of model preparation comparing quantum cluster approach to QM/MM. Furthermore, several MD studies have been used to understand the driving force for any of the suggested pathways. More importantly, investigating a series of conformational changes in the active site along the catalytic mechanism.

4.2 Computational Methods

4.2.1 DFT Calculations

All calculations were performed using Gaussian 09 suite of program.⁵³ The hybrid density functional theory method B3LYP; a combination of Becke's three parameter exchange functional and Lee, Yang and Parr's correlation,⁵⁴⁻⁵⁶ was used. Two basis sets including 6-31G(d) 5d, 6-31G(d,p) 5d were chosen for structures optimization. In addition, the HM-GGA M06-2X functional⁵⁷ was also used in accordance with the 6-31G(d) 5d basis set. Relative energies were obtained via single point energy calculations on the optimized structures using the 6-311G(d,p) 5d and the 6-311+G(2df,p) 5d basis sets. Frequency calculations were only used to characterize the optimized structures as minima or transition structures. The same active site model as in our previous study,¹⁶ extracted from the MsrB crystal structure of *N.gonorrhoeae* pilB (PDB: 1L1D)⁴⁹, was chosen to test the effect of functional and basis set choice. This model as previously described includes main active site residues.¹⁶

4.2.2 QM/MM Models

All QM/MM calculations were performed using Gaussian 09 suite of program using the ONIOM formalism.^{53,58} Truhlar's HDFT M06-2x was chosen for the QM layer calculations for better description of the hydrophobic interactions between the substrate and Trp442.⁵⁷ Furthermore, the AMBER96 force field was used for the MM layer as implemented in Gaussian.⁵⁹ The 6-31G(d) 5d basis set was chosen for structures optimization. However relative energies were obtained via single point energy calculations on the optimized structures at the ONIOM (M062X/6-311+G(2df,p):AMBER96) using the electronic embedding formalism to account for the polarization of the environment. One monomer of the same MsrB crystal structure of *N.gonorrhoeae* pilB (PDB: 1L1D)⁴⁹ was used to generate QM/MM models, in order to test the effect of model selection. Furthermore, two QM/MM models were used during calculations.

To mimic our previous DFT model we used the same crystal structure without running molecular dynamic simulations to keep the same starting distance between the two sulfur atoms. We isolated one of the monomers that include a cacodylate molecule in the active site. Then, the cacodylate was manually modified into Met. Likewise, two selenomethionine 509 and 464 were mutated back to Met. A PFROSST force field was used to parameterize and optimize the overall structure as implemented in MOE suite of program. Then, the QM/MM model was built using the whole monomer. No atoms were kept fixed during the calculations. A large QM model around the second Cys was chosen including Asp437, Ser438, Ala439, Cys440, Ala441, Trp442, Pro443, Ser444, Arg493, Asp412, Cys495, 3 water molecules and the substrate Met. In the QM/MM starting structure a bond was formed between Cys495 sulfur and the substrate forming a sulfonium cation intermediate as our starting structure.

The second QM/MM model was built using the same monomer but after minimizing and running an MD simulation for the sulfonium cation for 5 ns, simulation details are

explained in details in the following MD section. First, similar to previous model the cacodylate and MSE were mutated to Met. Then the catalytic Cys sulfur was bonded to the substrate sulfur forming a sulfonium cation intermediate. Afterwards, the whole system was solvated via adding 3434 water molecule to the system. Next, the system was minimized using the PFROSST force field until the root mean square gradient of the total energy become less than $0.1 \text{ Kcal mol}^{-1} \text{ \AA}^{-1}$. Subsequently, we performed a 5 ns simulation after performing a 100 ps equilibration step. The distance between the two sulfur of the catalytic and the second Cys, was used to cluster the resultant structures into 5 Clusters. Then an average structure was obtained from the highest populated cluster. Similar to previous QM/MM model, a large model around the second Cys including the same residues and 4 water molecules. Likewise, the whole monomer was used with no fixed atoms.

4.2.3 Molecular Dynamics Simulations

The MOE suite of program was used to prepare, minimize and analyze all structures. However, the NAMD program was used to run the MD simulations. An all atom force field PFROSST were used to parameterize all structures. It uses AMBER10 for macromolecules and for small molecules it uses parm@forsst parameters and AM1-BCC for charges. Four different stages in the catalytic mechanism were considered for simulations including the substrate free active site, the Michaelis complex with methionine sulfoxide, the sulfonium cation intermediate and sulfenic acid intermediate. All investigated structures were manually generated via modifying the MsrB *N.gonorrhoeae* pilB (PDB: 1L1D) crystal structure. Prior to simulations, all residues were solvated using the whole monomer as solute adding a layer of water up to 6 Å from the residues. The S–O bond in sulfenic acid were restrained to previous DFT measured value.

Solvated structures were minimized using PFROSST force field until the total energies become less than $0.1 \text{ Kcal mol}^{-1} \text{ \AA}^{-1}$. Next, 500 ps simulating annealing

simulations with time step of 2 fs were performed on the minimized structures allowing for thermal relaxation. All simulations were done under constant pressure however the temp varied during the annealing process. Starting by heating the system from 150° to 300° for 25ps followed by an equilibration step for another 25 ps. Then, it was heated again from 300° to 400° for 25 ps followed by a longer time equilibration step for 350 ps. Afterwards, it allowed to cool down to 300° for 25ps followed by a final equilibration for 50 ps. Subsequently, a production run for 15.5 ns, with time step of 2 fs were performed at constant temp (300°), in the case of wild, Michaelis complex, sulfonium cation and sulfenic acid intermediates.

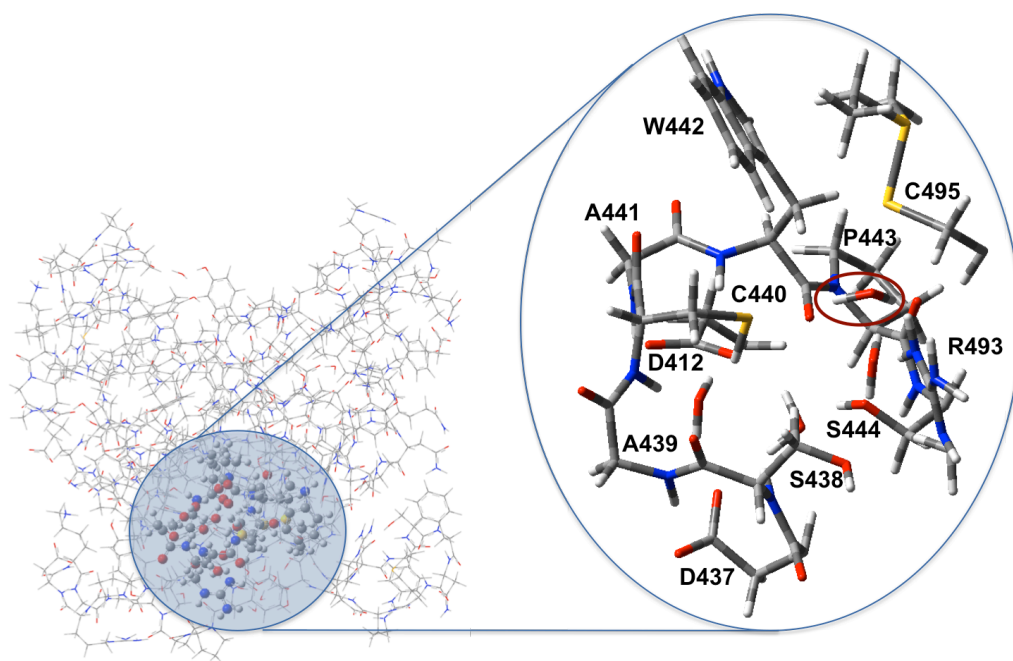


Figure 4.1. The second QM/MM model for MsrB sulfonium cation intermediate optimized structure. The QM layer atoms are highlighted in sticks to the right of the Figure.

Generated trajectories were then analyzed. Three main criteria were used to analyse the generated structures, the change in the $C_x \cdots C_x$ distance of the catalytic and second Cys

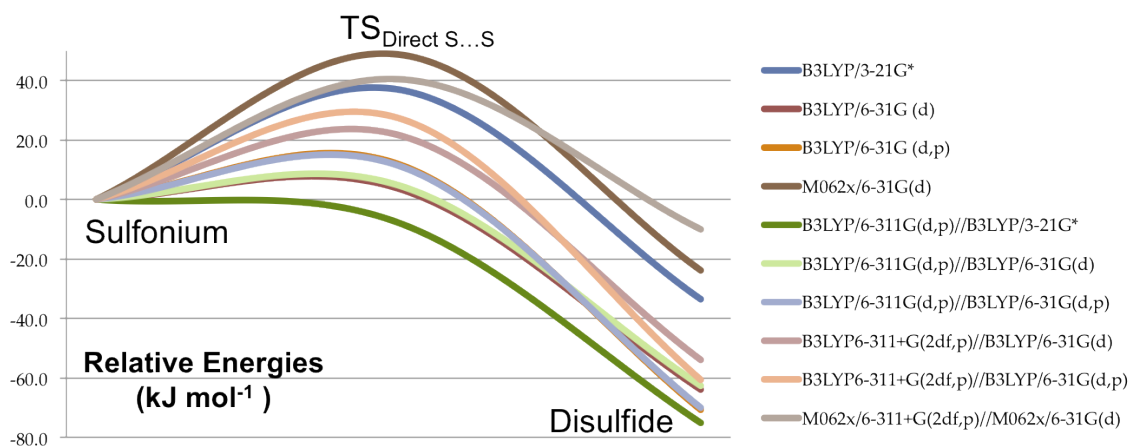
as well the change in the S-S distance. Furthermore, the RMSD of the active site QM residues were also obtained. Then, generated structures were clustered into five clusters in which the average structure was further investigated. The first 1.5 ns of the simulations were excluded from our analysis as it considered as equilibration period.

4.3 Results and Discussions

4.3.1 QM-cluster Approach

In order to determine the effect of the level theory on the energy difference between the two pathways, sulfenic acid and direct disulfide formation, several combinations of functionals and basis sets have been used, see Figure 4.2 and 4.3. Relative energies were compared at optimization level as well as single points calculations in gas phase. It is important to mention that all following discussed optimized TSs are enzymatically feasible. In our previous study, optimized structures were obtained at quite small basis set using the B3LYP/3-21G* 5d level of theory and the direct disulfide bond formation, before corrections, was slightly favoured by 2.0 kJ mol⁻¹. However this difference was significantly changed upon using a larger single point calculations at the B3LYP/6-311G(d,p) level of theory to 55.0 kJ mol⁻¹. Upon adding solvation and dispersion corrections, this difference was reduced to 30.2 kJ mol⁻¹. Therefore, we investigated the effect of using a larger basis set for optimization using polarization functions as well as adding f and diffuse function in the single points calculations.

Using the same model as in our previous paper, we optimized the two pathways starting from 3Large, the sulfonium cation intermediate, using B3LYP/6-31G (d) level of theory. Surprisingly, the energy difference between the two TSs increased to 60.5 kJ mol⁻¹ which is in agreement with our previous single point results emphasizing that a direct sulfur attack on the sulfonium sulfur is more favourable than forming sulfenic acid intermediate. This difference was slightly decreased upon adding p functions to the hydrogen atoms, 56.8 kJ mol⁻¹.



Level of theory	TS _{Direct S...S}	Disulfide
B3LYP/3-21G*	37.0	-33.5
B3LYP/6-31G (d)	3.7	-63.9
B3LYP/6-31G (d,p)	11.9	-70.7
M062x/6-31G(d)	48.8	-23.8
B3LYP/6-311G(d,p)//B3LYP/3-21G*	-7.8	-75.1
B3LYP/6-311G(d,p)//B3LYP/6-31G(d)	4.8	-62.6
B3LYP/6-311G(d,p)//B3LYP/6-31G(d,p)	11.4	-70.0
B3LYP6-311+G(2df,p)//B3LYP/6-31G(d)	21.7	-53.9
B3LYP6-311+G(2df,p)//B3LYP/6-31G(d,p)	27.4	-60.6
M062x/6-311+G(2df,p)//M062x/6-31G(d)	40.5	-10.0

Figure 4.2. Potential energy surface (PES) for direct disulfide bond formation pathway from optimization and single point calculations at several level of theory. Relative energies in kJ mol^{-1}

Single points on previous optimized structure using the same level of theory as in previous paper B3LYP/6-311G(d,p) have not changed the difference between the barriers, 60.5 and 55.2 kJ mol^{-1} for the B3LYP/6-31G(d) and B3LYP/6-31G(d,p), respectively. However, using a larger level of theory B3LYP/6-311+G (2df,p) adding f and diffuse functions slightly decreased the difference to 53.0 and 48.7 kJ mol^{-1} . These differences are still in agreement with our previous results at the B3LYP/6-311G(d,p)//B3LYP/3-21G*. Therefore, our new results with larger basis sets emphasized our previous conclusion in which the direct formation of disulfide bond is favoured over sulfenic acid formation.

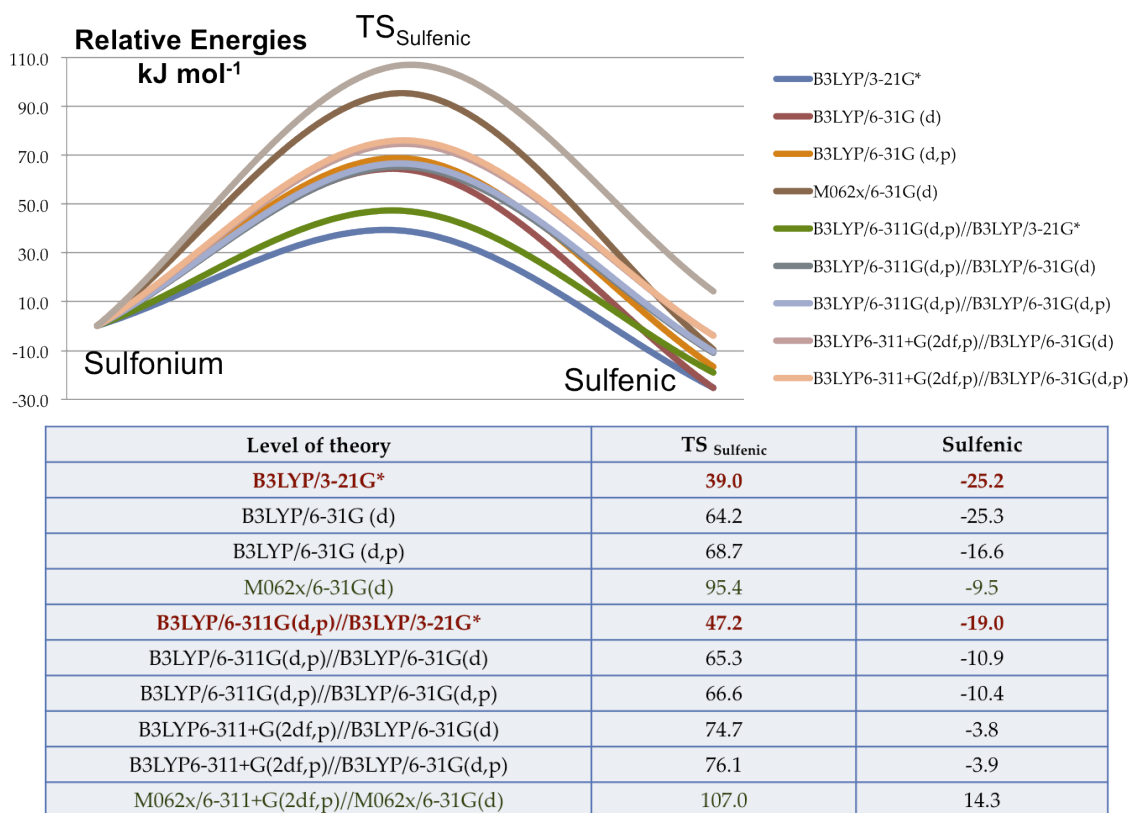


Figure 4.3. Potential energy surface (PES) for sulfenic acid formation pathway from optimization and single point calculations at several level of theory. Relative energies in kJ mol^{-1}

Since all previous results were obtained using the hybrid DFT B3LYP functional, we also considered using different functional to test if this difference is an artefact of the chosen method. Therefore, we also investigated the two pathways using meta hybrid DFT functional M062X. From previous calculations, the 6-31G(d) seems to be a reasonable basis set, so it has been employed. The energy difference between the two pathways at optimization level (M062X/6-31G(d)) gave a similar difference of $-46.6 \text{ kJ mol}^{-1}$, in favour of direct disulfide bond formation. More importantly, the sulfenic acid and disulfide TSs increased to 95.4 and 48.8 kJ mol^{-1} , respectively. Furthermore, the same results were obtained at the M062X/6-311G(2df,p) single point calculations showing an energy difference between the two pathways of 66.5 kJ mol^{-1} . Therefore, using different

basis sets and functionals does not change our previous conclusion; the direct disulfide bond formation is approximately 55 kJ mol^{-1} more favoured than the sulfenic acid intermediate formation, However both pathways are enzymatically feasible.

4.3.2 QM/MM Results

Since basis sets and functional change confirmed previous conclusions, we also considered investigating the effect of model choice. Therefore, we used a QM/MM model as described in the method section to explore both pathways starting from the sulfonium intermediate structure. A large QM layer around the recycling Cys440 has been chosen. In all our previous models we considered anionic recycling cysteine, however a neutral cysteine might affect the proposed mechanism. It is worth mentioning that, the pKa of the recycling cysteine has been experimentally suggested to be higher than 9.⁶⁰ Therefore, we attempted to calculate the proton affinity of Cys440. However, such calculations could not be obtained as in all our anionic models the optimization leads to the formation of the disulfide bond emphasizing that a direct nucleophilic attack on Cys495 sulfur forming a disulfide bond upon activation of Cys440 without an energy barrier.

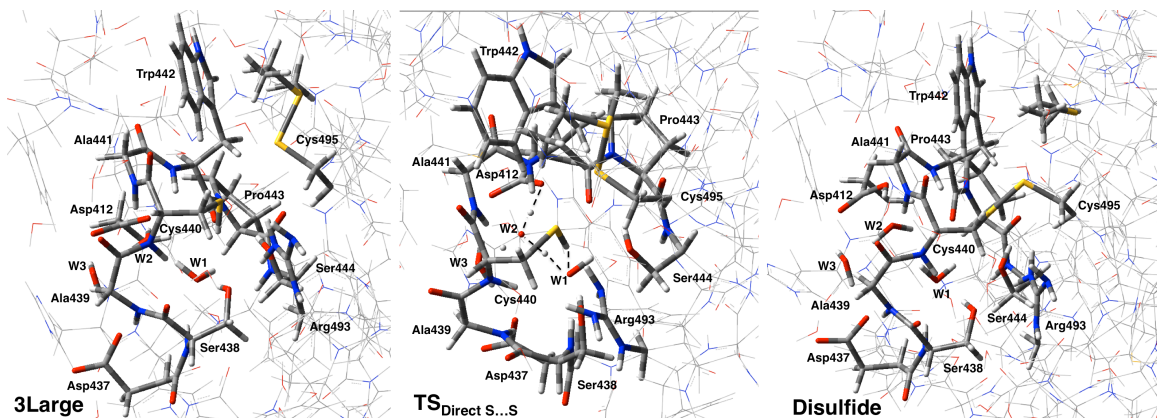


Figure 4.4. Optimized structures for the direct disulfide bond formation from the QM/MM calculations in which the sulfonium intermediate were obtained from MsrB crystal structure.

The optimized reactive complex of the neutral model shows that, the hydrogen of Cys440 is hydrogen bonded to Asp412 and Asp437 via several water molecules suggesting a new role for either of the aspartate residues to activate Cys440. It also shows that W1 (Figure 4.4) could act as a base as it has very strong hydrogen bonds with Ser438, NH backbone of Asp437 and W2 with distances of 1.68 Å, 1.99 Å and 1.51 Å, respectively. In addition it has a weaker hydrogen bonds to Ser444 and the backbone carbonyl of Pro443 with distances of 2.13 Å and 2.60 Å, respectively. Interestingly, we found that the activation of catalytic Cys495 occurs via a low barrier, 19.2 kJ mol⁻¹, concerted TS in which proton shuttle from Cys440 to Asp412 via two water molecules. The TS structure (Figure 4.4) show that, a proton is shared between W1 and W2 with distances of 1.26 Å and 1.17 Å, respectively, In turn, the O-H distances in W2 is elongated to 1.07 Å in which the proton is strongly interacting with Asp412 oxygen, 1.44 Å. Similarly, the S-H bond in Cys440 is elongated to 1.36 Å and is hydrogen bonded to W1 oxygen at a distance of 2.14 Å. IRC calculations shows that this TS leads to the activation of Cys440, however no stable intermediate was obtained as Cys440 concomitantly attack the sulfur centre of the sulfonium cation forming disulfide bond. The formed disulfide intermediate is 246.6 kJ mol⁻¹ lower than RC.

The sulfenic acid pathway is also investigated, however to activate a water molecule to attack the sulfonium sulfur the Cys440 needs to be activated first and as noted before once Cys440 is activated it forms the disulfide bond. Therefore, we investigated a concomitant activation of Cys440 and the sulfenic acid water, however no such TS was found to occur. Furthermore, the closest water is not in a proper location to attack; as including the whole monomers limited the ability of having that water in a proper position. These results actually question the validity of the QM cluster approach in which the surrounding of the active site are not included allowing for water addition in several position, which might not be realistic as in this case. In addition the nature of the TS differs based on chosen computational approach, QM-cluster or QM/MM, as in the

former one the TS represent the S \cdots S \cdots S bonds formation and dissociation however, the QM/MM study show that this TS would occur spontaneously upon Cys440 activation.

All previous results emphasize our original proposed mechanism were the direct disulfide bond formation is more favoured. As previously described we tested the choice of basis set, the functional and the computational approach, however all these result are based on the original crystal structure of MsrB in which the distance is 3.29 Å. Therefore, the obtained PES might be an artifact of the short distance between the two sulfurs in the crystal structure. Specially as discussed in the introduction that the recycling Cys440 is located in a flexible loop and the S \cdots S distance in the only solved NMR structure is over 7Å. Furthermore, in several species this distance has been found to be longer. Therefore, we ran an MD simulation on the sulfonium cation intermediate for 5ns and the generated structures were clustered and an average structure was obtained from the highest populated cluster. Interestingly, the S \cdots S distance was elongated during the simulation to approximately 5Å. Similar to previous QM/MM the whole monomer was used and no atoms were kept fixed allowing for full relaxation of the system.

The optimized structure of the average structure of the sulfonium cation shows that (Figure 4.5), the S \cdots S distance is 4.88 Å. More importantly, a water molecule is now located below the Cys495 sulfur and the S_{Cys495} \cdots O_{W1} distance is 2.44 Å. In turn, the W1 proton is hydrogen bonded to the S_{Cys495} sulfur with a distance of 2.29 Å. Furthermore, It also has several hydrogen bonds with carbonyl of Pro443 and W3 (to Arg493) with distances of 1.78 Å and 1.72 Å, respectively. The Cys440-H is strongly hydrogen bonded to W2 with a distance of 1.43Å, which in turn strongly polarized by two hydrogen bonds with Asp412 and Asp437 via a distance of 1.48Å and 1.66Å, respectively. This hydrogen bonding network emphasis our previous proposed mechanism in our previous MsrB and MsrA studies in which the recycling cysteine activates a water molecule forming sulfenic acid (Cys440 activates W1, in MsrB). Interestingly, this step was found to occur via a low barrier of 16.7 kJ mol⁻¹. In this TS (Figure 4.5), Cys440 hydrogen is shared with W2 with

distances of 1.50 Å and 1.43 Å, respectively. In turn the W2 O-H bond is elongated to 1.05 Å strongly interacting with Asp412, 1.78 Å. As previous no stable intermediate of the anionic sulfur was optimized, however the activation of Cys440 occur with concomitant proton transfer from W1 to Cys440 activating W1 and forming sulfenic acid. The generated sulfenic acid intermediate is 120.6 kJ mol⁻¹ lower in energy than the sulfonium intermediate, 3Large.

These results raise many questions about the quantum cluster approach especially in dealing with proteins containing flexible loops showing that misleading results might be obtained upon using models based on crystal structures without proper preparation. It also emphasizes the importance of the starting structure in determining enzymatic mechanisms. Such proper preparation is missing in the Quantum cluster approach as the models are commonly based on the crystal structure with the substrate analogue. Our QM/MM MD based calculations showed that a sulfenic acid intermediate is formed in the MsrB catalytic mechanism which is contradicting our DFT cluster approach as well as QM/MM crystal structure based calculations.

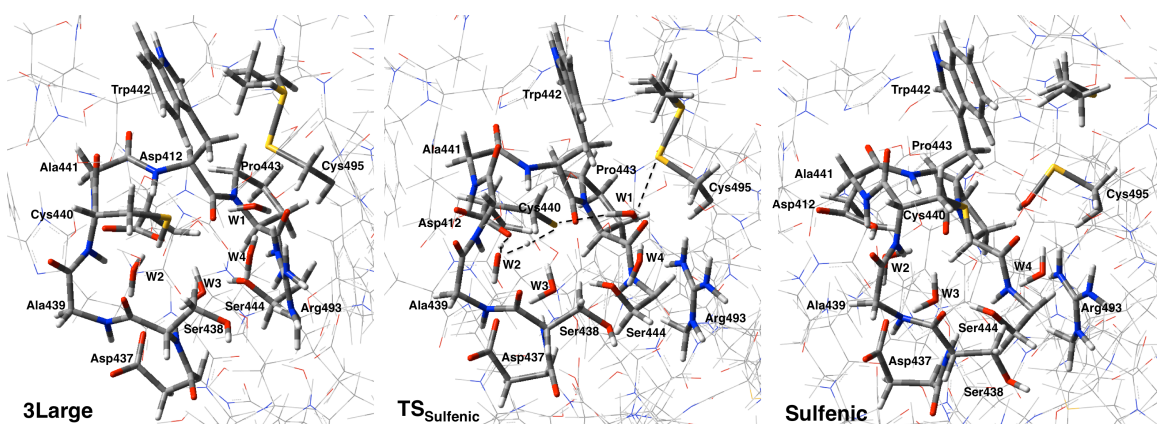


Figure 4.5. Optimized structures for the sulfenic acid formation pathway from the QM/MM calculations in which 3large were obtained from sulfonium intermediate MD simulation.

4.3.3 Molecular Dynamics Simulations Results

In order to confirm previous results and to understand why the S...S distance is different from crystal structure, we ran a series of MD simulations for several complexes starting from Apoenzyme to Michaelis complex, sulfonium and sulfenic acid, as described in the method section. We first examined the distance between the two Cys before substrate binding in wild type apoenzyme. A 16 ns MD simulation was enough to reach equilibrium as can be seen from the RMSD of the QM layer (Figure 4.6). The rmsd of the apoenzyme-generated structures fluctuate between approximately 0.75 Å and 1.5 Å indicating the stability of the system. This higher fluctuation range was investigated and found to be because that the recycling cysteine is located in a high flexible turn as previously suggested.

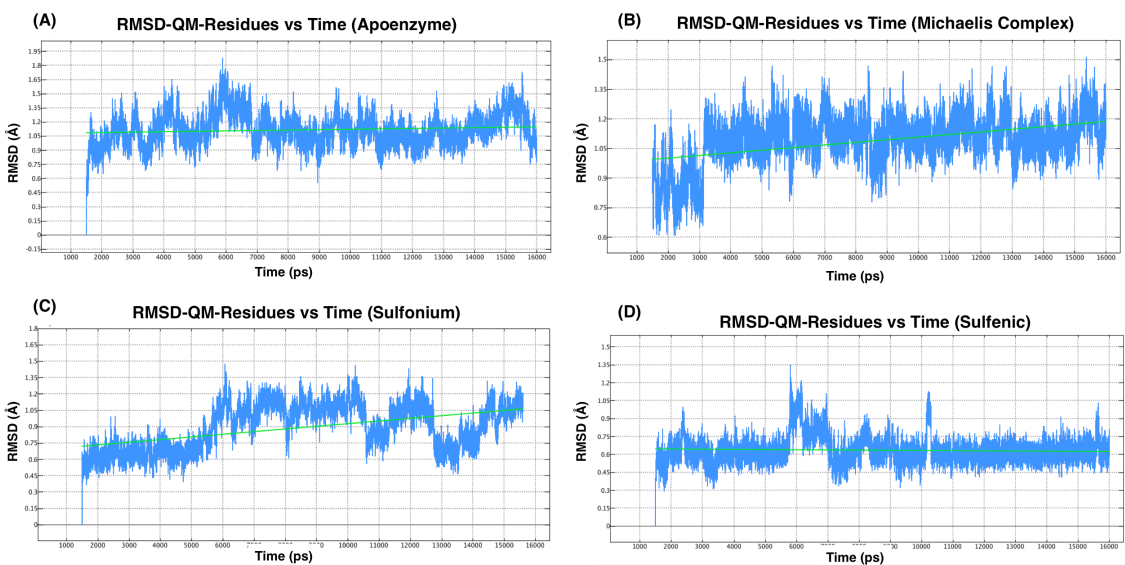


Figure 4.6. RMSD's of the QM layer vs time for the four models used to represent the catalytic cycle (A) Apoenzyme before substrate binding, (B) substrate bound active site, (C) sulfonium cation intermediate, and (D) sulfenic acid intermediate.

The main criterion in analyzing these simulations is the distance between the C_{α} 's of Cys440 and Cys495. As can be seen in Figure 4.7, this distance fluctuate approximately

between 7.0 Å and 9.5 Å with few exceptions where the distance increased near to 11.0 Å, as will be seen in the following discussion. Further, the mean distance is 8.36 Å with a maximum distance of 11.27 Å and minimum distance of 6.76 Å. Furthermore, all generated structures were clustered into five clusters and the average structure for each were determined. The highest populated cluster represents approximately 60% of the simulation in which the distance in the average structure is 8.36 Å, which is agreement with the crystal structure, ≈ 7.40 Å. Furthermore, the highest and lowest distance in this cluster are 8.88 Å and 7.83 Å, respectively. The second populated cluster represents approximately 20% of the simulation. The C_{α} distance significantly decreased in the average structure to 7.31 Å. In contrast, the distance in the 3rd populated cluster average structure has increased to 9.4 Å. All three clusters represent more than 95 % of the simulation. Structures with a distance between 9.93 Å and 11.27 Å in the 4th cluster represented less than 2.5 %. It is also worth to mention that the average S-S distance is 5.79 Å with a maximum distance of 11.08 and minimum distance of 3.08 Å. All these results show that the formation of disulfide bond is feasible with no need for conformational changes in the active site and the average distance is in agreement with crystal structure.

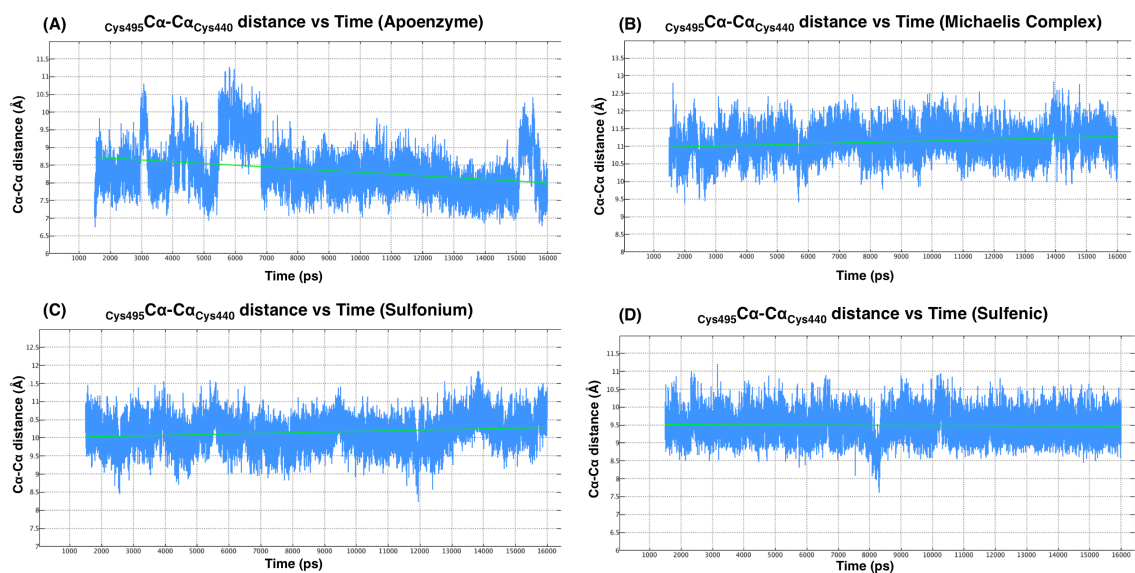


Figure 4.7. Distance change between C α of Cys495 and Cys440 vs time the four models used to represent the catalytic cycle (A) Apoenzyme before substrate binding, (B) substrate bound active site, (C) sulfonium cation intermediate, and (D) sulfenic acid intermediate. Distances are in Angstrom.

Since the Apoenzyme simulation shows a similar close distance, we also investigated the effect of the substrate binding. As shown in Figure 4.6, the rmsd of the QM layer are equilibrated with a fluctuation after the first 3 ns between approximately 1.0 Å and 1.4 Å. Similarly, the distance between the two Cys C α fluctuate between 10-12 Å with a mean distance 11.12 which is significantly higher than in the apoenzyme with a huge minimum distance of 9.39 Å and a maximum distance of 12.83 Å. As before, the generated structures were clustered into five clusters, which have been further analyzed. The most populated cluster represents approximately 59% of the simulation with a maximum and minimum distances of 11.45 Å and 10.79 Å, respectively. It is important to mention a distance more than 11 Å does not allow for disulfide bond formation. More importantly, the average structure of this cluster has 11.12 Å distance between the C α of the two cysteine. Furthermore, the second and third most populated clusters have a similar population of 19.5% with an average distance of 11.77 Å and 10.46 Å, respectively. All

three clusters constitute more than 97% of the simulation. Furthermore, the mean distance between the two sulfurs has significantly increased to 6.93 Å with a minimum of 4.72 Å and a maximum of 9.31 Å. Therefore, it's obvious that substrate binding leads to active site conformational changes. In fact Cys440 is located on the same loop as Trp442, which also changes its position upon substrate binding. The strong hydrophobic interaction between the substrate and Trp442 is very important for substrate binding and catalysis as have been examined experimentally. This suggest that, as can be seen in Figure 4.8, the substrate binding enforce Trp442 conformational change which in turn affects the conformation of the whole turn including Cys440 extending the distance between the two sulfurs.

In order to confirm prior conclusion, we also considered long MD simulation for the sulfonium intermediate in which bigger range of fluctuation in the QM rmsd between approximately 0.6 Å-1.3 Å; this actually an artifact of the turn flexibility and the sulfonium cation. It's important to mention that the positive sulfur of the sulfonium cation was not properly parameterized was the current force field, however its partial charges was correct, which might be a reason for this fluctuation. However, the current results are in agreement with the former one, as the distance of the C_α is equilibrated and fluctuate between 9.5 Å and 11.5 Å with a mean of 10.14 Å with a minimum of 8.24 Å and a maximum of 11.15 Å. This range is still significantly large compared to the apoenzyme simulation. Furthermore, the distance in the average structures of the three highest populated, 96%of the simulation, clusters are 10.14 Å, 9.48 Å and 10.80 Å, respectively. Furthermore, the mean distance between the two sulfurs is 5.80 Å with a minimum of 3.88 Å and a maximum of 8.49 Å. These results of the longer simulation confirm our QM/MM MD based results were the S··S distance are larger than the crystal structure allowing for formation of sulfenic acid intermediate in the mechanism as experimentally suggested.

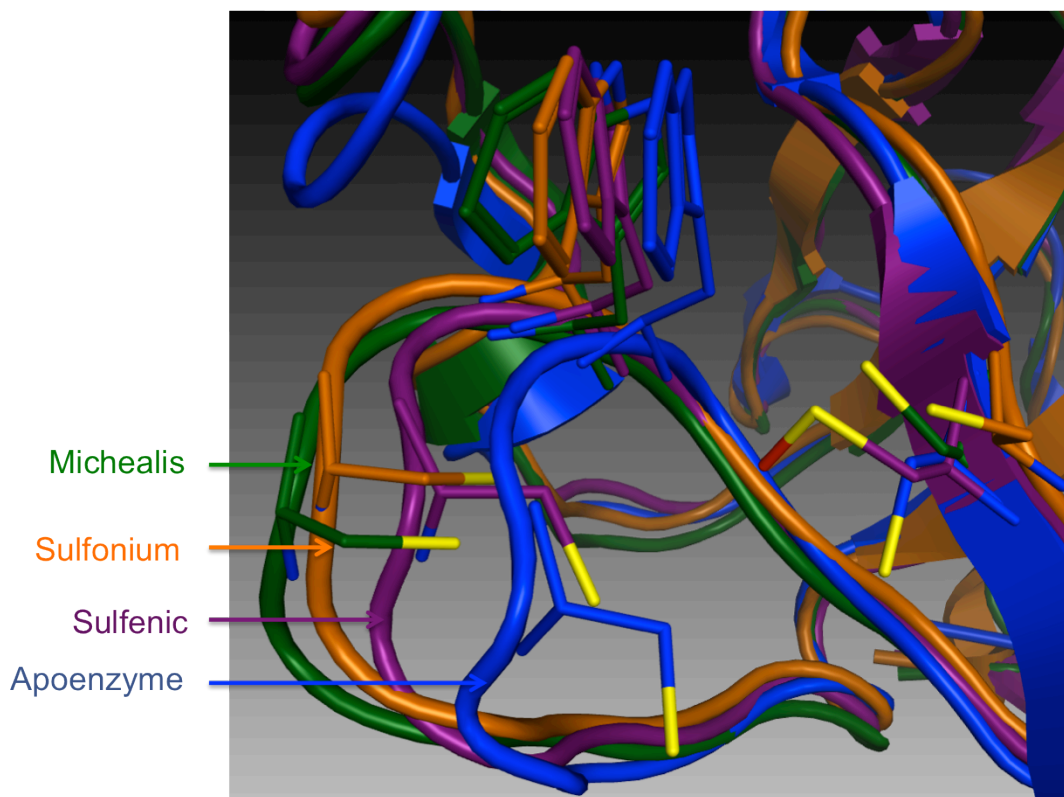


Figure 4.8. Superposition of the 4 MD's active site average structures in apo-enzyme, Michaelis complex, sulfonium and sulfenic acid intermediates.

Although earlier results explain the effect of substrate binding on the mechanism and its effect on the proceeding pathway, however there is still a question to be answered, which is how the disulfide bond is formed. Therefore we ran another 16ns MD simulation on the sulfenic acid structure investigating the effect of substrate removal; starting with a large S-S distance using the average structure of the 5ns sulfonium simulation. As shown in Figure 4.6, the RMSD of the QM layer is equilibrated fluctuating between 0.4 Å and 1.0 Å. This emphasized from the distance between the two C_{α} of the two cysteine which fluctuate approximately between 8.5 Å and 10.5 Å with a mean distance of 9.48 Å with a minimum of 7.63 Å and maximum of 11.19 Å. Furthermore, the highest populated cluster represents 59% of the simulation with an average structure distance of 9.48Å which would allow for a direct disulfide formation with no need to any conformational changes.

Similarly, the distance in the average structure of the second and third most populated cluster is 8.89 Å and 10.08 Å, respectively. All three clusters represent more than 97.5% of the simulation showing that upon substrate removal the Trp442 and Cys440 turn would move back toward its original conformation as in the wild type simulation and crystal structure.

Comparing all previous simulation results show that, before substrate binding the distance between the C_α of Cys440 and Cys495 in the average structure is 8.36 Å which then is elongated upon substrate binding to 11.12 Å showing that substrate induce conformational changes in the active site specially via its hydrophobic interaction with Trp442 which is located on the same turn as Cys440. These results were confirmed by the 10.14 Å distance in the average structure of the sulfonium cation allowing for sulfenic acid formation as have been shown in the previous QM/MM section. Finally, upon formation of sulfenic the substrate leaves the active site allowing for a shorter interaction distance, 9.48 Å, between the two Cys allowing for disulfide bond formation with no need for any conformational changes.

Our results emphasize the importance of starting structure preparation as well as question the applicability of QM cluster approach to system with flexible turns showing that misleading results might be obtained. Our previous QM-cluster study on MsrB showed that both mechanisms are energetically feasible however the direct disulfide bond formation is favorable. In contrast, in this study, running several MD simulations as well as investigating the effect of substrate binding in the active site showed that a sulfenic acid intermediate would occur in the mechanism.

4.4 Conclusions

In this study several computational approaches, QM cluster, QM/MM and MD simulations, have been employed comprehensively to reinvestigate that last step in the reductase step of the catalytic mechanism of MsrB. First, we investigated the effect of

level of theory by utilizing several basis sets for optimization and single point calculations. Increasing the basis set size confirmed our previous conclusion in which the direct disulfide bond formation is favorable over sulfenic intermediate formation. Furthermore, the HMGGA M062x was also used to examine the effect of functional on previous conclusion, however M062x emphasized previous conclusion. In addition, we also investigated the effect of model choice using QM/MM (ONIOM) approach including a large QM layer around Cys440. The QM/MM model was directly extracted from the crystal structure. Interestingly, deprotonating the second Cys440 leads to direct nucleophilic attack on Cys495 sulfur forming disulfide. Therefore we considered using a neutral cysteine in our model. Using neutral cysteine changed the nature of the optimized TS as it is now represent the activation of Cys440 leads to direct formation of disulfide emphasizing previous results.

In order to eliminate the possibility that all previous results are an artifact of the used same crystal structure in which the S...S distance is 3.29 Å, we investigated the effect of the starting distance by running a 5 ns MD simulation on the sulfonium cation intermediate. Surprisingly, the MD simulation results shows that the distance between the two sulfur is significantly increased to 4.88 Å. More importantly a water molecule is located in a proper position for attack forming sulfenic acid intermediate. To investigate the mechanism from the average structure of the MD simulation, we run QM/MM using a similar QM layer as in previous calculations. Surprisingly, the an anionic Cys440 leads to direct formation of sulfenic acid intermediate. We also found the activation of Cys440 occurs via a low barrier of 16.7 kJ mol⁻¹ proposing a catalytic role for Asp412.

Since Our MD results shows different distance from the crystal structure, we considered running long MD simulations for 16ns using several models representing active site changes during the catalytic mechanism. MD results show that, before substrate binding the distance between the C_α of Cys440 and Cys495 in the average structure is 8.36 Å which then is elongated upon substrate binding to 11.12 Å showing

that substrate induce conformational changes in the active site, especially via its hydrophobic interaction with Trp442 which is located on the same turn as Cys440. These results were confirmed by the 10.14 Å distance in the average structure of the sulfonium cation allowing for sulfenic acid formation as have been shown in the previous QM/MM section. Finally, upon formation of sulfenic the substrate leaves the active site allowing for a shorter interaction distance, 9.48 Å, between the two Cys allowing for disulfide

4.5 References

- (1) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.
- (2) Llano, J.; Gault, 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, USA. 2010; Vol. 2, p 643.
- (3) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2010**, *46*, 2354.
- (4) Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. *J. Chem. Theory Comput.* **2009**, *5*, 1400.
- (5) Ban, F. Q.; Rankin, K. N.; Gault, J. W.; Boyd, R. J. *Theor. Chem. Acc.* **2002**, *108*, 1.
- (6) Shaik, S.; Kumar, D.; de Visser, S. P. *J. Am. Chem. Soc.* **2008**, *130*, 10128.
- (7) Chung, L. W.; Li, X.; Sugimoto, H.; Shiro, Y.; Morokuma, K. *J. Am. Chem. Soc.* **2008**, *130*, 12299.
- (8) Lundberg, M.; Siegbahn, P. E. M.; Morokuma, K. *Biochemistry* **2008**, *47*, 1031.
- (9) Banas, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *Methods* **2009**, *49*, 202.

- (10) Masson, F.; Laino, T.; Rothlisberger, U.; Hutter, J. *ChemPhysChem* **2009**, *10*, 400.
- (11) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142.
- (12) Li, X.; Chung, L. W.; Paneth, P.; Morokuma, K. *J. Am. Chem. Soc.* **2009**, *131*, 5115.
- (13) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949.
- (14) Huang, W. J.; Gherib, R.; Gauld, J. W. *J. Phys. Chem. B* **2012**, *116*, 8916.
- (15) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645.
- (16) Robinet, J. J.; Dokainish, H. M.; Paterson, D. J.; Gauld, J. W. *J. Phys. Chem. B* **2010**, *115*, 9202.
- (17) Neiers, F.; Boschi-Muller, S.; Branlant, G. *J. Phys. Chem. B* **2011**, *115*, 10775.
- (18) Robinet, J. J.; Dokainish, H. M.; Paterson, D. J.; Gauld, J. W. *J. Phys. Chem. B* **2011**, *115*, 10776.
- (19) Weissbach, H.; Resnick, L.; Brot, N. *Biochim. Biophys. Acta.* **2005**, *1703*, 203.
- (20) Moskovitz, J. *Biochim. Biophys. Acta* **2005**, *1703*, 213.
- (21) Boschi-Muller, S.; Gand, A.; Branlant, G. *Arch. Biochem. Biophys.* **2008**, *474*, 266.
- (22) Olry, A.; Boschi-Muller, S.; Branlant, G. *Biochemistry* **2004**, *43*, 11616.
- (23) Antione, M.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2003**, *278*, 45352.
- (24) Lowther, W. T.; Brot, N.; Weissbach, H.; Honek, J. F.; Matthews, B. W. *Proc. Nat. Acad. Sci. U.S.A.* **2000**, *97*, 6463.

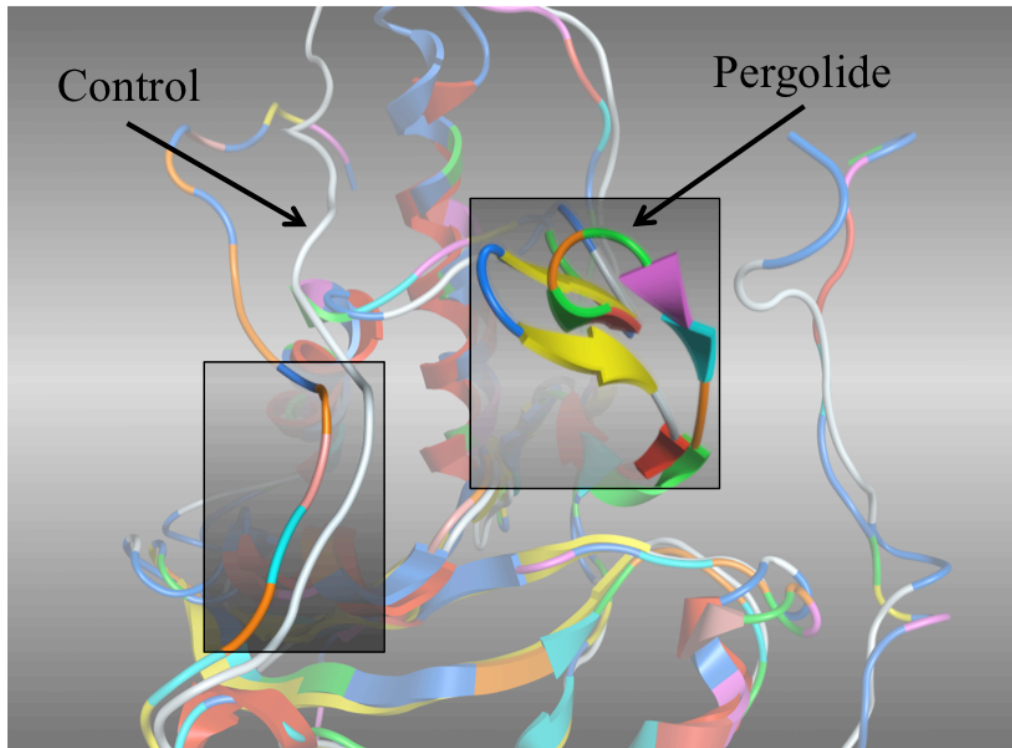
- (25) Boschi-Muller, S.; Azza, S.; Sanglier-Cianferani, S.; Talfournier, F.; Van Dorsselear, A.; Branlant, G. *J. Biol. Chem.* **2000**, *275*, 35908.
- (26) Lowther, W. T.; Brot, N.; Weissbach, H.; Matthews, B. W. *Biochemistry* **2000**, *39*, 13307.
- (27) Kettenhofen, N. J.; Wood, M. J. *Chem. Res. Toxicol.* **2010**, *23*, 1633.
- (28) McGrath, A. J.; Garrett, G. E.; Valgimigli, L.; Pratt, D. A. *J. Am. Chem. Soc.* **2010**, *132*, 16759.
- (29) Roos, G.; Messens, J. *Free Radic. Biol. Med.* **2011**, *51*, 314.
- (30) Rehder, D. S.; Borges, C. R. *Biochemistry* **2010**, *49*, 7748.
- (31) Alvarez, B.; Carballal, S.; Turell, L.; Radi, R. Formation and Reactions of Sulfenic Acid in Human Serum Albumin. In *Methods in Enzymology.*; Cadenas, E., Packer, L., Eds.; Elsevier Academic Press Inc: San Diego, USA. 2010; Vol. 473, p 117.
- (32) Poole, L. B.; Karplus, P. A.; Claiborne, A. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 325.
- (33) Nelson, K. J.; King, S. B.; Poole, L. B. *FASEB J.* **2005**, *19*, A258.
- (34) Hawkins, C. L.; Morgan, P. E.; Davies, M. J. *Free Radic. Biol. Med.* **2009**, *46*, 965.
- (35) Mansuy, D.; Dansette, P. M. *Arch. Biochem. Biophys.* **2011**, *507*, 174.
- (36) Saurin, A. T.; Neubert, H.; Brennan, J. P.; Eaton, P. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17982.
- (37) Carballal, S.; Radi, R.; Kirk, M. C.; Barnes, S.; Freeman, B. A.; Alvarez, B. *Biochemistry* **2003**, *42*, 9906.
- (38) Baker, L. M. S.; Poole, L. B. *J. Biol. Chem.* **2003**, *278*, 9203.
- (39) Roussel, X.; Kriznik, A.; Richard, C.; Rahuel-Clermont, S.; Branlant, G. *J. Biol. Chem.* **2009**, *284*, 33048.

- (40) Flohe, L.; Toppo, S.; Cozza, G.; Ursini, F. *Antioxid. Redox Signal.* **2011**, *15*, 763.
- (41) Blackinton, J.; Lakshminarasimhan, M.; Thomas, K. J.; Ahmad, R.; Greggio, E.; Raza, A. S.; Cookson, M. R.; Wilson, M. A. *J. Biol. Chem* **2009**, *284*, 6476.
- (42) Joensson, T. J.; Murray, M. S.; Johnson, L. C.; Lowther, W. T. *J. Biol. Chem* **2008**, *283*, 23846.
- (43) Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6238.
- (44) Fuangthong, M.; Helmann, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 6690.
- (45) Salmeen, A.; Andersen, J. N.; Myers, M. P.; Meng, T. C.; Hinks, J. A.; Tonks, N. K.; Barford, D. *Nature* **2003**, *423*, 769.
- (46) Hill, B. G.; Bhatnagar, A. *J. Mol. Cell. Cardiol.* **2012**, *52*, 559.
- (47) Mieyal, J. J.; Gallogly, M. M.; Qanungo, S.; Sabens, E. A.; Shelton, M. D. *Antioxid. Redox Signal.* **2008**, *10*, 1941.
- (48) Taylor, A. B.; Benglis, D. M.; Dhandayuthapani, S.; Hart, P. J. *J. Bacteriol.* **2003**, *185*, 4119.
- (49) Lowther, W. T.; Weissbach, H.; Etienne, F.; Brot, N.; Matthews, B. W. *Nat. Struct. Biol.* **2002**, *9*, 348.
- (50) Kim, Y. K.; Shin, Y. J.; Lee, W. H.; Kim, H. Y.; Hwang, K. Y. *Mol. Microbiol.* **2009**, *72*, 699.
- (51) Ranaivoson, F. M.; Neiers, F.; Kauffmann, B.; Boschi-Muller, S.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2009**, *394*, 83.
- (52) Lange, O. F.; Rossi, P.; Sgourakis, N. G.; Song, Y.; Lee, H.-W.; Aramini, J. M.; Ertekin, A.; Xiao, R.; Acton, T. B.; Montelione, G. T.; Baker, D. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 10873.

- (53) Gaussian 09, Revision D.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
- (54) Becke, A. D. *J. Phys. Chem.* **1993**, *98*, 1372.
- (55) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (56) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (57) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (58) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (59) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- (60) Neiers, F.; Sonkaria, S.; Olry, A.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2007**, *282*, 32397.

Chapter 5

Small Molecules Activating an Antioxidant Enzyme



5.2 Introduction

In drug design, enzymes are one of the primary therapeutic targets.¹ Indeed, enzyme inhibitors are widely used as drugs for numerous diseases, today.¹ Moreover, in the recent years, a new role of enzymes in drug development has become apparent, in which small molecules have been tested to enhance the activity of specific enzymes instead of inhibiting them as a new therapeutic strategy.²⁻⁴ Many enzymes have been shown to be activated via small molecules including glucokinase, sirtuin, AMPK and RNase.^{3,5-7} Furthermore, these activation mechanisms have been classified into two main groups,² in which small-molecules activators binds either directly to an allosteric site of the enzyme or to a regulatory subunit, leading to a conformational change and the stabilization of an open active sites enabling catalysis. Notably, the discovery and validation of possible target enzymes for activation is a very complicated process and there is no systematic approach to achieve this goal. In general, the activation of enzymes via small-molecules is actually favored over their inhibition due to: 1) allosteric sites are not conserved as active sites, so activators binding is considered more specific in general, 2) unlike inhibitors, much smaller concentration of the activators is needed to amplify the enzyme efficiency significantly.²

Methionine sulfoxide reductase (Msr) is a group of ubiquitous antioxidant enzymes that catalyze the repair of oxidized proteins, reduction of methionine sulfoxide (Met-O) to methionine (Met),⁸ with multiple health implications, from aging to several age related diseases such as Alzheimer's and cancer.^{9,10} It is mainly consisted of two stereospecific classes, MsrA and MsrB. Remarkably, the overexpression of MsrA in several species has been shown to increase their life span as well as knocking it out leads to oxidative stress.¹¹⁻¹³ Indeed, the defensive role of this enzyme against reactive oxygen species has been the subject of numerous studies.

The catalytic mechanism of MsrA can be divided into two step:^{14,15} 1) the reductase step, in which Met-O is reduced forming MsrA-sulfenic acid intermediate that

subsequently reduced forming an intramolecular disulfide bond (MsrA_{Ox}), 2) the regeneration step, in which the protein disulfide is reduced by thioredoxin (Trx) regenerating the active site (MsrA_{Red}). Recently, we have elucidated the mechanism for the reductase step in detail, showing that it happens via low barriers multistep.¹⁴ Experimentally, the rate-limiting step of the overall mechanism has been shown to be associated with the regeneration step.¹⁵ Outstandingly, the two steps of the mechanism require two different conformations, switching from a closed form active site in the reductase step to a more open fold for the regeneration step, allowing for Trx binding.^{16,17}

The overall fold of MsrA_{Ox} and MsrA_{Red} does not change much and the core structure of the enzyme is maintained. The main difference between the two conformations is the unfolding of the two-stranded antiparallel β -sheet (Gln122 to Thr132). Coudeville et al.¹⁶ using NMR studies have shown that in the oxidized form the Gln122-Thr132 segment is unfolded and has no secondary structure. Indeed, the unfolding of this segment is essential for Trx binding and subsequent reaction to regenerate the active site. In general, in either conformation, the core enzyme structure is lined with N and C-terminal ends with no secondary structure. The C-terminal end in *Escherichia coli* contains the resolving cysteinyl residue. It is important to mention that the Gln122-Thr132 segment in the reduced form surrounds the MsrA active site. Similar confirmations of the reduced form have been found in several species using X-ray crystallography.

Recently, Moskovitz et al.¹⁸ have shown that the activity of MsrA in rat neurons can be induced using a substrate mimic molecule, N-Acetyl Methionine Sulfoxide (Ac-Met-O). In vitro and in vivo studies have shown that Ac-Met-O protects the neuron from A β 42 toxicity suggesting a protective role in Alzheimer's. Based on their results, the authors suggested that the activity of MsrA might be provoked by small molecules containing a sulfoxide functional group, providing a novel way to develop antioxidant drugs in Alzheimer's and other oxidative stress related diseases.¹⁸ Consequently, more recently, Frankiln et al.¹⁹ have shown that Pergolide sulfoxide (PergSO) and Pergolide (Perg) as

well as S-adenosyl methionine (SAM) increases the activity of MsrA in neuronal cells suggesting similar role to Ac-Met-O. Perg is a dopamine receptor agonist that used to treat Parkinson's disease until 2007. Furthermore, the authors chose PergSO as it contains a methyl sulfoxide moiety suggesting that Perg will be first enzymatically converted to PergSO before activating MsrA.¹⁹ However, their results show that Perg is more potent than PergSO. At 10 μM concentration, Perg and PergSO increased the activity of MsrA by 158% and $\sim 130\%$, respectively. Interestingly, the upregulation of MsrA activity was partially correlated to its overexpression, suggesting that the activation occurs via an uncertain mechanism. In addition, SAM was found to have a less activation effect compared to Perg and PergSO.

Here we investigated the mechanism of MsrA activation by Perg, PergSO and SAM, considering the possibility of their direct interaction with the enzyme. The binding pockets were identified using docking techniques. The effects of small-molecules binding were elucidated using a 9.5 ns MD simulations. Furthermore, virtual-screening was used to identify other possible activator small-molecules.

5.2 Computational methods

In order to investigate the possibility of direct interaction, the *Escherichia coli* NMR structure (PDB: 2GT3) of the reduced form of MsrA was used as a starting structure.¹⁶ In addition, another starting structure was also considered in which the MsrA_{Red} was modified manually to form MsrA_{Ox}, by forming the Cys51-Cys198 disulfide bond. In both cases, the enzyme structure was first solvated, up to 15 Å away from any residue, and minimized using AMBER12:EHT forcefield. Subsequently, both MsrA_{Red} and MsrA_{Ox} were docked using the three ligands Perg, PergSO and SAM. The whole enzyme structure was used as a receptor for docking. The MOE suite of programs was used for docking.²⁰ First, the top 100 conformations for each ligand were generated using the

London dG scoring function. Second, using forcefield refinement method, only the top 50 conformations were retained and analyzed.

5.3 Results and Discussions

Docking the MsrA_{Red} form has shown that the three ligands mainly bind to a hydrophobic pocket at the interaction between the C-terminal and the α_4 secondary structure, in which the ligands interact with Leu207, Pro206, Ala159 and Phe156. Furthermore, very few conformation were found to occur at the N-terminal end along with β_5 .

Notably, upon forming the Cys51-Cys198 bond (MsrA_{Ox}) all three ligands were found to constantly bind to new binding pocket at the interaction of N-terminal and more importantly the two-stranded antiparallel β -sheet (Gln122-Thr132 segment). These interactions were found to include several residues such as His130, Arg125 from Gln122-Thr132 segment and Phe3, Lys5, Lys5 from the N-terminal. Furthermore, the binding score of all three ligands in MsrA_{Ox} is higher than the previously obtained for MsrA_{Red}. SAM was found to bind less regularly to this new pocket. Moreover, our docking structures show that the methyl sulfoxide moiety is not involved in these interactions; instead cation- π and hydrophobic interactions are the main reason for binding.

These results raised a main question if the binding of these ligands to the Gln122-Thr132 segment in the oxidized form might facilitate its unfolding and the subsequent regeneration reaction with Trx leading to MsrA direct activation, as observed experimentally. Therefore, using 9.5 ns molecular dynamic simulations, we investigated the dynamics of the Gln122-Thr132 segment upon ligand binding. Prior to MD's, we docked the three ligands using induced fit approach, in which the previously identified pocket residues were allowed to move and adjust during the simulation. The highest 10 conformations for each ligands were retained and top three poses were used as a starting point for MD simulations. Furthermore, a control MD simulation of MsrA_{Ox} was also

performed in order to compare the obtained results. As previous MOE suite of programs was used to generate and analyze MD's as well as NAMD was used to run the MD simulations.²¹

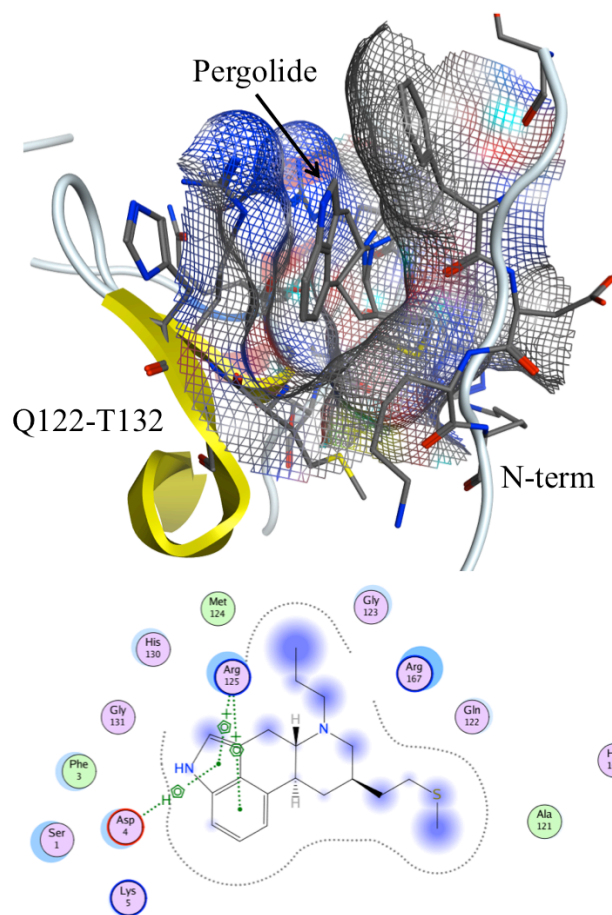


Figure 5.1. Top) docking predicted binding mode of Pergolide, bottom) schematic representation of Pergolide interactions in the MsrA allosteric site.

Experimentally, based on previous X-ray structures of MsrA_{Red}, a main hydrogen bond between Asp129 (from the Gln122-Thr132 segment) and Tyr197 was thought to play an important role in keeping the folded reduced form of the active site.¹⁷ However, in the used NMR structure, this hydrogen bond is broken, therefore in our analysis we also considered both scenarios as starting structures. Furthermore, in our analysis we

considered two main criteria, 1) the molecular surface of the Cys51-Cys198 disulfide bond and its surroundings residues as well as the disulfide exposure to the solvent, which reflects its subsequent reaction with Trx, 2) the root-mean-square deviation (RMSD) of the Gln122-Thr132 residues as well as the facing interacting residues, Gly196, Tyr197, Cys198 and Gly199 in the C-terminal. Although, all three ligands were investigated, in the following section we will mainly discuss the results obtained for Perg as it has the highest effect on MsrA activity.

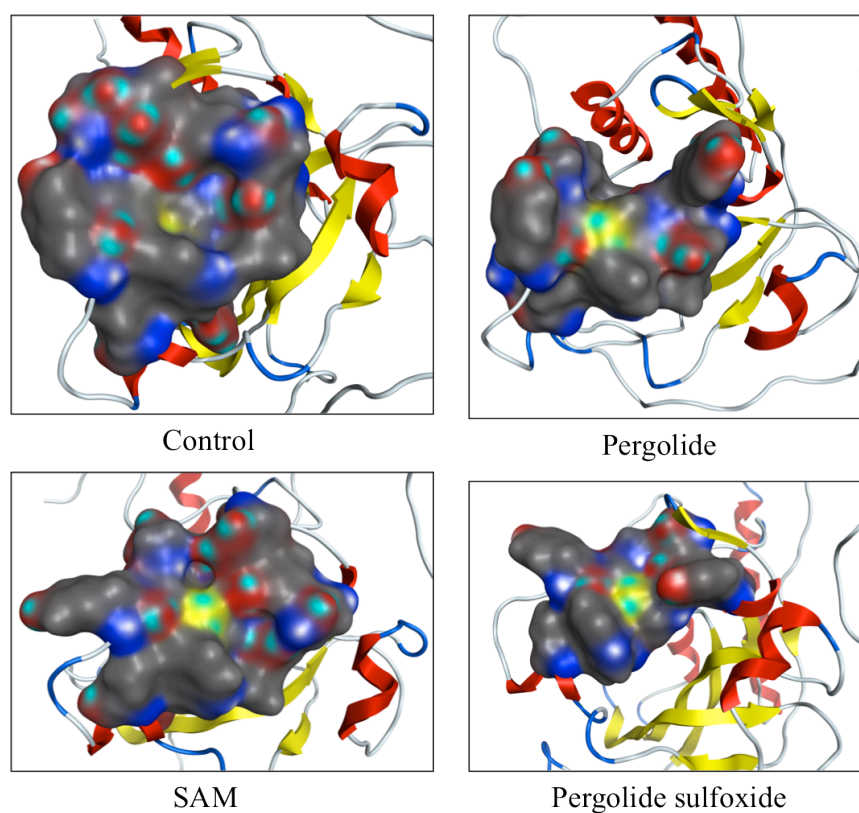


Figure 5.2. Molecular surface representation of the disulfide surrounding residues extracted from MD simulations. The surface exposure of the disulfide bond is represented by yellow colour.

As shown in Figure 5.2, from the control MD, the disulfide bond is buried in the active site with a very minimal solvent exposure. Furthermore, either upon forming the

Tyr197...Asp129 hydrogen bond or in its absence the Gln122-Thr132 segment is still folded completely, covering the active site. The rmsd of the involved residues shows that after significant change due to annealing the rmsd were kept constant between 2-3 Å reflecting a small increase upon time, see Figure 5.3. It is also important to mention that in our previous study on MsrA mechanism we ran a 25 ns simulation of the oxidized structure and no significant change in the Gln122-Thr132 segment was observed.¹⁴

Unlike the control MD, the introduction of Perg was found to increase the disulfide bond exposure to the solvent, dramatically, see Figure 5.2. More importantly, the Gln122-Thr132 segment residues are no longer interacting with the Gly196-Gly199 C-terminal residues. The superposition of the control versus Perg structure clearly emphasize this finding. Indeed, this is reflected in the obtained rmsd, in which after almost 7 ns of the simulation a huge jump was observed to occur breaking the original interactions between the Gln122-Thr132 residues and the Gly196-Gly199 C-terminal, see Figure 5.3. Similar dynamics were also observed for PergSO and SAM. Based on the MD generated structures, a cation- π between Arg125 and Perg arene might be the cause of this dynamics, see Figure 5.1. Furthermore, using different starting structure, another interaction was also found to occur with His130. All these results suggest that Perg binding disrupt the interactions between the Gln122-Thr132 and C-terminal residues, which might affect the segment folding. However, the generated structures still maintain the two-stranded antiparallel β -sheet.

In order to speed up the unfolding dynamics within a moderate simulation time, we ran a second MD (15 ns) starting with the last structure of the Perg MD upon removing Perg from the allosteric site. Indeed, the rmsd of the Gln122-Thr132 segment shows significant increase upon time. Notably, snapshots of the MD's conformation show the total uncovering of the active site and the fluctuation of the Gln122-Thr132 segment from two-stranded antiparallel β -sheet and no secondary structure. A superposition of the MD generated structure and the NMR unfolded form of the enzyme (PDB:2IEM)¹⁶

highlighting the exposure of the active site to subsequent reaction with Trx, see Figure 5.4. It is also important to mention that based on docking results that the binding of these small-molecules activators to the reduced form of the enzyme does not affect its folding and subsequently will not disturb the first step (the reductase step) of the mechanism which require a folded active site.

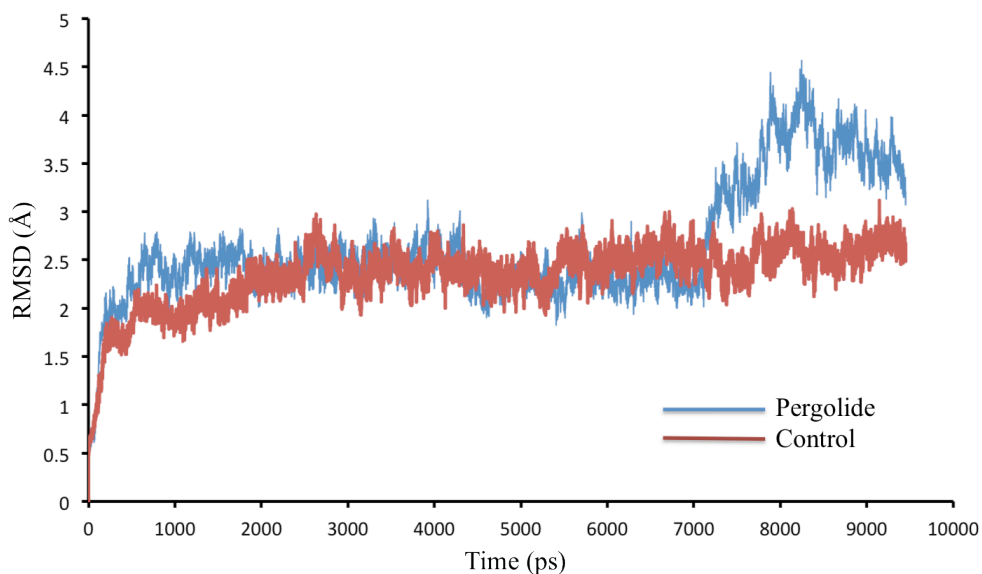


Figure 5.3. Time evaluation of root-mean-square deviation (Rmsd's) of Gln122-Thr132 segment and Gly196-Gly199 residues with and without Pergolide binding.

Previous results suggest that the activation mechanism of MsrA by Perg, PergSO and SAM occurs due to direct interaction with an allosteric site of the enzyme facilitating the unfolding of the Gln122-Thr132 segment. However, Perg cannot be longer used as a drug as it was withdrawn from the market due to undesirable side effects. Therefore, we also considered the search for other ligands, which can bind to MsrA producing similar effect. Using virtual-screening, we searched the subset ZIM from the ZINC drug database for binding in the previously identified allosteric site.²² ZIM includes all compounds have been in human including worldwide drugs. Using the triangle matcher placement and London dG scoring, we docked 11421 molecules, retaining the highest 10 poses for each

molecule. Based on the scoring function, the highest 292 molecules were selected for further analysis. In which an induced fit docking approach as well as a second refinement step, as described earlier, was used retaining the top highest 5 poses for each ligand. A full list of the 292 molecules with their docking score are included (See Appendix Table B2).

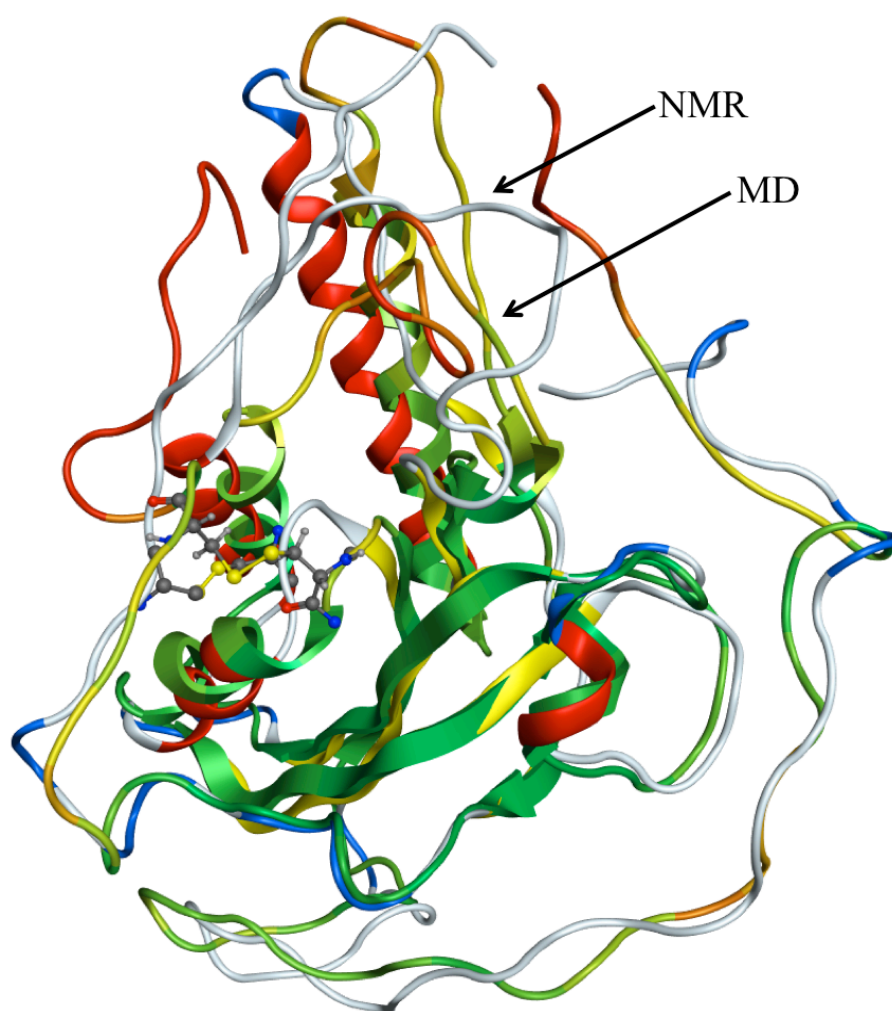


Figure 5.4. Superimposition of the MsrAOx NMR solution structure with that of MsrAOx MD generated structure upon Pergolide binding and removal.

In general, most of the identified systems were found to include an arene function group and some of them include the cation- π interaction as in Perg suggesting the possibility of discovering new small-molecules MsrA activators. Surprisingly, many

antibiotics such as Cefoperazone, Natamycin and Lymecycline were found to have high binding affinity to MsrA. Furthermore, many natural occurring molecules including NAD⁺, NADH and ATP were also identified as possible ligand. These results raise another main question, if MsrA is naturally regulated by biomolecules! Finally, part of the identified molecules has been found or suggested previously to have an antioxidant effects such as Nicardipine, folinic acid and Sildenafil.²³⁻²⁵ In addition, we run a short MD simulation (2 ns) on selected ligands including NADH, Natamycin and Cefoperazone. Although the short simulation time compared to control, our MD's show that the molecular surface of the disulfide was found to increase significantly in comparison to the longer 10 ns control simulation.

5.4 Conclusions

In summary, this report suggests the direct activation of MsrA via binding of small-molecules providing an explanation of the previously recognized experimental results. Furthermore, our results also suggest other possible activator ligands. Further experimental studies to test this direct interaction as well as high-throughout screening will provide tools to confirm our proposed mechanism of activation and develop new activators with an antioxidant activity.

5.5 References

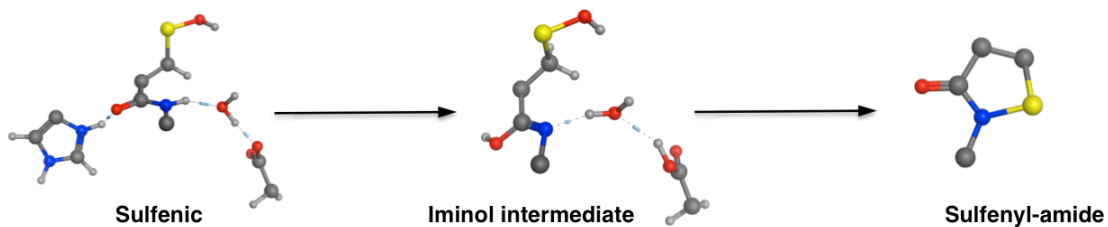
- (1) Copeland, R. A.; Harpel, M. R.; Tummino, P. J. *Expert Opin. Ther. Targets* **2007**, *11*, 967.
- (2) Zorn, J. A.; Wells, J. A. *Nat. Chem. Biol.* **2010**, *6*, 179.
- (3) Bishop, A. C.; Chen, V. L. *J. Chem. Biol.* **2009**, *2*, 1.
- (4) Wassman, C. D.; Baronio, R.; Demir, O.; Wallentine, B. D.; Chen, C. K.; Hall, L. V.; Salehi, F.; Lin, D. W.; Chung, B. P.; Hatfield, G. W.; Chamberlin, A. R.; Luecke, H.; Lathrop, R. H.; Kaiser, P.; Amaro, R. E. *Nat. Commun.* **2013**, *4*.

- (5) Sinclair, D. A.; Guarente, L. In *Annual Review of Pharmacology and Toxicology, Vol 54*; Insel, P. A., Ed.; Annual Reviews: Palo Alto, 2014; Vol. 54, p 363.
- (6) Merdanovic, M.; Monig, T.; Ehrmann, M.; Kaiser, M. *ACS Chem. Biol.* **2013**, *8*, 19.
- (7) Zorn, J. A.; Wille, H.; Wolan, D. W.; Wells, J. A. *J. Am. Chem. Soc.* **2011**, *133*, 19630.
- (8) Lim, J. C.; You, Z.; Kim, G.; Levine, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10472.
- (9) De Luca, A.; Sanna, F.; Sallese, M.; Ruggiero, C.; Grossi, M.; Sacchetta, P.; Rossi, C.; De Laurenzi, V.; Di Ilio, C.; Favalaro, B. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18628.
- (10) Koc, A.; Gladyshev, V. N. You have full text access to this contentMethionine Sulfoxide Reduction and the Aging Process. In *Biogerontology: Mechanisms and Interventions*; Rattan, S. I. S., Akman, S., Eds.; Blackwell Publishing: Oxford, UK. 2007; Vol. 1100, p 383.
- (11) Singh, V. K.; Moskovitz, J. *Microbiology-(UK)* **2003**, *149*, 2739.
- (12) Douglas, T.; Daniel, D. S.; Parida, B. K.; Jagannath, C.; Dhandayuthapani, S. *J. Bacteriol.* **2004**, *186*, 3590.
- (13) Nan, C. L.; Li, Y. J.; Jean-Charles, P. Y.; Chen, G. Z.; Kreymerman, A.; Prentice, H.; Weissbach, H.; Huang, X. P. *Biochem. Biophys. Res. Commun.* **2010**, *402*, 608.
- (14) Dokainish, H. M.; Gauld, J. W. *Biochemistry* **2013**, *52*, 1814.
- (15) Antoine, M.; Gand, A.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2006**, *281*, 39062.
- (16) Coudevylle, N.; Antoine, M.; Bouguet-Bonnet, S.; Mutzenhardt, P.; Boschi-Muller, S.; Branlant, G.; Cung, M. T. *J. Mol. Biol.* **2007**, *366*, 193.

- (17) Ranaivoson, F. M.; Antoine, M.; Kauffmann, B.; Bosch-Muller, S.; Aubry, A.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2008**, *377*, 268.
- (18) Moskovitz, J.; Maiti, P.; Lopes, D. H. J.; Oien, D. B.; Attar, A.; Liu, T. Y.; Mittal, S.; Hayes, J.; Bitan, G. *Biochemistry* **2011**, *50*, 10687.
- (19) Franklin, J. M.; Carrasco, G. A.; Moskovitz, J. *Neurosci. Lett.* **2013**, *533*, 86.
- (20) MOE, version 2012.10 (2012) Chemical Computing Group Inc., Montreal.
- (21) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (22) Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* **2005**, *45*, 177.
- (23) Kouoh, F.; Gressier, B.; Dine, T.; Luyckx, M.; Brunet, C.; Ballester, L.; Cazin, J. C. *Cardiovasc. Drugs Ther.* **2002**, *16*, 515.
- (24) Perk, H.; Armagan, A.; Naziroglu, M.; Soyupek, S.; Hoscan, M. B.; Sutcu, R.; Ozorak, A.; Delibas, N. *J. Clin. Pharm. Ther.* **2008**, *33*, 635.
- (25) Bayes, B.; Pastor, M. C.; Bonal, J.; Romero, R. *Kidney Int.* **2005**, *67*, S39.

Chapter 6

Formation of Stable Iminol Intermediate in the Redox Regulation Mechanism of PTP1B



6.1 Introduction

Protein tyrosine phosphatases (PTP's) are a large family of enzymes responsible for dephosphorylating phosphorylated tyrosyl residues in proteins,^{1,2} a physiologically important post-translational modification (PTM). In particular, the sub-class PTP1B, first characterized in 1988,^{1,3,4} plays a key role in inhibiting insulin and leptin signaling.⁵⁻⁷ Conversely, it has been shown to have a crucial positive role in signaling of, for instance, cSrc tyrosine kinases in breast cancer^{3,8,9} and the ubiquitous Ras proteins.¹⁰ Consequently, PTP1B is considered to be an outstanding drug target for the treatment of several diseases including diabetes, obesity and cancer.^{3,11-13}

PTP1B catalyzes the dephosphorylation of phosphotyrosine via a two step ping-pong mechanism (Scheme 1).¹⁴ In the first step, the sulfur of an active site cysteinyl (Cys215) nucleophilically attacks the phosphotyrosine ester to form a phosphoenzyme intermediate. This is accompanied by the concomitant release of the tyrosine. In the second step, hydrolysis of the phosphoenzyme intermediate occurs via nucleophilic attack of H₂O upon activation by Asp181. Tiago et al.¹⁴ have previously investigated both steps using X-Ray crystallography to characterize transition state analogs.

To date, four different mechanisms have been identified by which the function of PTP1B can be regulated: phosphorylation, sumoylation, proteolysis and oxidation.³ In particular, in the latter, Cys215 has been experimentally observed to be reversibly oxidized to a sulfenic acid (Cys251SOH) by reactive oxygen species' (ROS).¹⁵ This modification is in part facilitated by the unusually low pK_a (4.5-5.5) for the thiol of Cys215.¹ This oxidative PTM mediates several signaling pathways. For instance, with regards to insulin the stimulation of trans-membrane receptor kinase (RTK) leads to the activation of NADPH oxidase, producing ROS that oxidize Cys215 in PTP1B, thus transiently inhibiting its function.¹ The activity of PTP1B is restored upon reduction by an external thiol such as thioredoxin (Trx), dithiothreitol (DTT) or Glutathione (GSH).¹⁶

In general, sulfenic acid is susceptible to further and irreversible oxidation to sulfinic or sulfonic acid. X-ray crystallographic analysis, and pulse-chase labeling and mass spectrometry experiments suggested that the sulfenic acid can undergo an intramolecular reaction to give a seemingly unique cyclic sulfenyl-amide species; thus protecting it from further oxidation.^{15,17-19} To-date this mechanism has only been observed for PTP1B. But it should be noted that it has been suggested that such an intermediate may occur in other proteins including PTP1 α , and other protein families such as organic hydroperoxide resistance regulator (OhrR) in *Bacillus subtilis*.^{20,21} In other enzymes sulfenic acid is protected via, for example, formation of a disulfide (e.g., methionine sulfoxide reductase) or hypervalent sulfurane species (e.g., archaeal peroxiredoxin).^{22,23}

Initially it was proposed that the sulfenyl-amide forms via a direct S_N2 mechanism. Specifically, the backbone nitrogen of the neighboring serinyl (Ser216) nucleophilically attacks the sulfenic acid's S γ atom with concomitant release of H₂O (scheme 1).¹⁷ Furthermore, the hydrogen bonding interaction between the N δ 1 atom of the invariant histidyl (His214) and the carbonyl oxygen of Cys215 was suggested to play a key role in enhancing the nucleophilicity of the Ser216 backbone nitrogen.¹⁷ Indeed, it has been shown that mutation of His214 to Asn or Ala increases the pK_a of Cys215.²⁴ Generation of the sulfenyl-amide intermediate induces an active site conformational change. In particular, formation of the S–N bond disrupts a hydrogen bond between the R-groups of Ser216 and Tyr46, rendering the enzyme inactive.²⁵ Thus, there is interest in inducing or stabilizing this inactive oxidized form as a potential therapeutic approach for several diseases.²⁵ It is noted that experimentally, several external thiols have been shown to be able to reduce the sulfenyl-amide, regenerating the activity of the catalytic Cys215.¹⁶ This restoration mechanism has been confirmed via re-soaking crystals of sulfenyl-amide with dithiothreitol (DTT).¹⁷

Sarma et al.²⁶ have previously studied, both experimentally and computationally, the mechanism of sulfenyl-amide formation in PTP1B using model non-protein molecules. In

particular, formation of sulfenyl-amide via the above proposed direct mechanism was calculated at the B3LYP/6-31G(d) level of theory to have a barrier of 206.2 kJ mol⁻¹. In addition, they considered an alternate mechanism involving heterocyclic substitution of an oxazoline *ortho* to the sulfenic acid moiety. This modification significantly reduced the calculated barrier to 119.8 kJ mol⁻¹. However, they noted that the sulfenic acid model used did "not effectively mimic the cyclization of protein sulfenic acids".²⁶ Similarly, Sarma et al.²⁷ examined *ortho* substitution effects on a small amido thiophenol molecule and concluded that S···N/O interactions could influence the properties of the sulfenic acid. Furthermore, nearby residues may have a role in the sulfenyl-amide formation. More recently, as part of a review, Defelipe et al.²⁸ discussed results they had obtained for the mechanism of PTP1B using QM/MM. The reactive (QM) region of their chemical model consisted of Cys215 and Ser216. Similar to the results of Sarma et al.²⁶ they concluded that the direct formation of sulfenyl amide occurs with a high barrier of 205.0 kJ mol⁻¹. Experimentally, however, several studies have suggested that the rate-limiting step in sulfenyl-amide formation is generation of the sulfenic acid, not the sulfenyl-amide.^{15,20} Hence, many questions remain about the enzymatic mechanism.

In this present study, formation of the putative sulfenyl-amide intermediate from a Cys215 derived sulfenic acid, within the enzyme environment, is investigated via complementary application of molecular dynamics simulations and extensive quantum mechanical/molecular mechanical (QM/MM) modeling.

6.2 Computational Methods

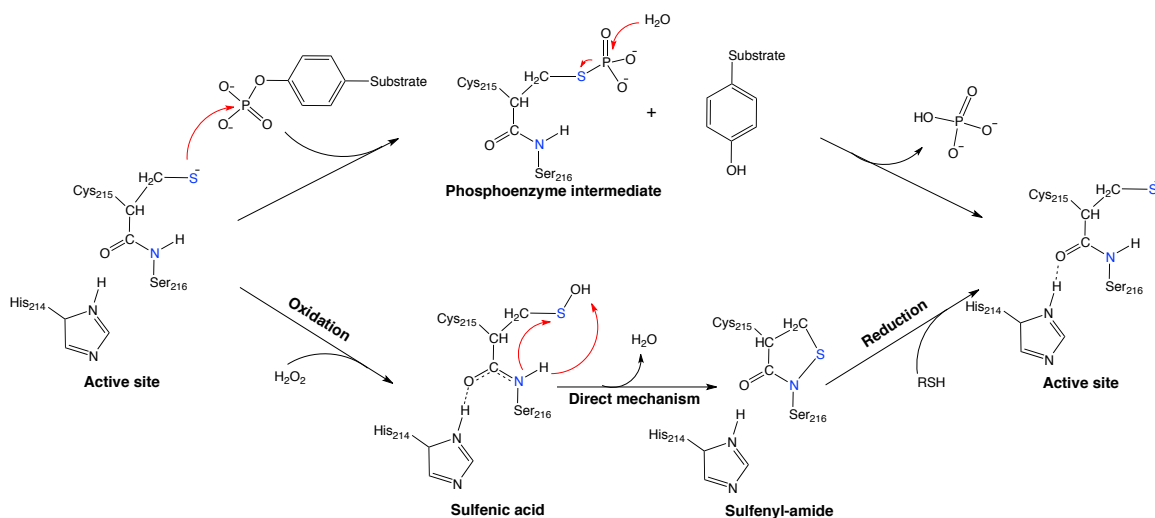
6.2.1 Protein Model Preparation and Molecular Dynamics Simulations

The X-ray crystal structure preparation and the MD analysis were performed using the Molecular Operating Environment (MOE) software package.²⁹ Molecular dynamics simulations were conducted using the NAMD Molecular Dynamics software.³⁰ The X-ray

crystal structure of the *Homo sapiens* PTP1B in its sulfenic acid oxidation state was used as starting structure and obtained from the PDB structure 1OET.¹⁷

First, problems encountered in the crystal structure were automatically corrected using the structural preparation applications in MOE. Second, the missing protons in the crystal structure were added using the protonate 3D application in MOE that assigns each residue ionization state by minimizing the total free energy of the system.³¹ Then, the enzyme was spherically solvated up to 15 Å beyond every protein atom. Later, the newly generated structure was minimized using AMBER12:EHT force field, which uses AMBER12 parameters for protein and Extended Hückel Theory for parameterizing small molecules.³²⁻³⁴ Finally, in order to allow for thermal relaxation and multiple conformers generation, the minimized structure was used as starting point for a 3 ns MD simulation.

Scheme 6.1. The proposed dephosphorylation mechanism of phosphotyrosine in PTP as well as the oxidative regulation mechanism in PTP1B via formation of sulfenyl-amide and the subsequent reactivation.^{14,17}



In the MD simulation, all atoms were set free to move and a time step of 2 fs was used. The Columbic interactions were calculated with the PME method and the short range van der Waals interactions were truncated at 8-10 Å. The simulation started with an annealing dynamics from 150 K to 300 K then to 400 K to finally 300 K at a constant volume for 550 ps, followed by a production run in the NPT ensemble at 300K and 1 bar for 2500 ps.

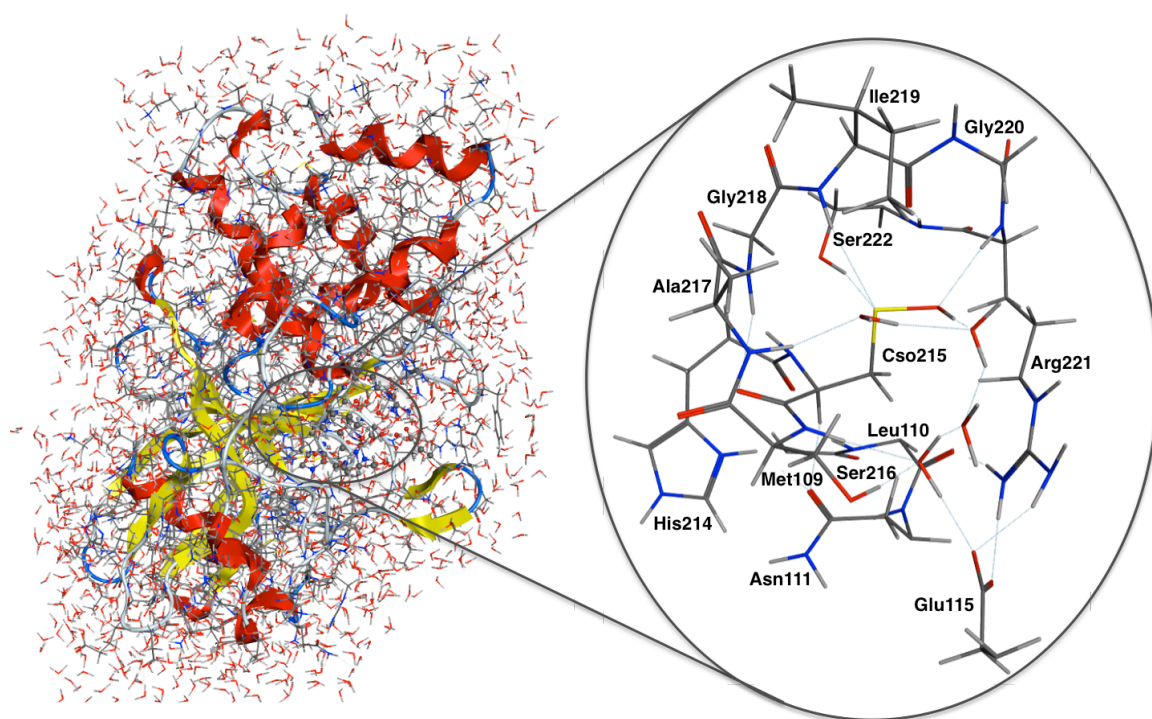


Figure 6.1. Illustration of the QM/MM model used (Model I) in this study obtained from the average structure. The QM layer atoms are highlighted and shown in sticks representation to the right.

All generated structures were analyzed using MOE software. Furthermore, based on the root mean square deviation (rmsd) of the active site residues, the generated structures were clustered into 5 main groups. The average structure of the highest populated group was selected and further minimized using AMBER12:EHT force field and the resulting

structure was used for further calculations using quantum mechanical/ molecular mechanical (QM/MM) approach. Finally, the rmsd for several active site interactions were also calculated to confirm the consistency of these interactions during the simulation.

6.2.2 QM/MM Models and Calculations

All calculations were performed within the ONIOM scheme using the Gaussian 09 suite of programs.³⁵ The QM/MM starting structures were acquired from the previous preparation steps using the whole protein. The QM high layer was described using the hybrid-meta-exchange-correlation functional M06-2X.³⁶ While the MM low layer was described by AMBER96 force field as implemented in Gaussian.³⁰ Optimized geometries were obtained using the 6-31G(d,p) basis set for the high layer. All atoms in the systems were kept free to move including solvent atoms. Relative energies were then obtained using single point energy calculations on the optimized structures at the ONIOM(M06-2X/6-311+G(2df,p):AMBER96) level of theory. The choice of functional and basis sets was based on our previous benchmarking study of several biological sulfur species in which M06-2X was found to be one of the best functionals in describing SN bonds with respect to QCISD and MP2 (manuscript in preparation). Due to the models size and computational cost, frequency calculations were conducted only on the high layer to confirm the nature of the optimized structures as previously performed.^{37,38} In addition, topological analysis of the electron densities as well as natural bond orbital (NBO) analysis for certain intermediates were performed at the M062-X/6-311+G(2df,p) level of theory using AIM2000 program.³⁹

In model I, as shown in Figure 6.1, the QM layer is formed of the catalytic Cys215 in its sulfenic acid form, Ser216, His214 in protonated form, Ala217, Gly218, Ile219, Gly220, Arg221, Ser222 and Asn111. In addition the R group of Glu115 as well as the backbone of Met109 and Leu110 were included. Furthermore, 4 active site water

molecules were included in the high layer. Besides the previous model, we also considered modifying model I in which we used a neutral His214 (Model II) to confirm the role of the histidyl residue in sulfenyl-amide formation.

6.3 Results and Discussion

6.3.1 Insights into Reactive Complex Structures

The MD average structure was first optimized using ONIOM approach as previously discussed. In model 1, the optimized structure shows that (see RC in Figure 6.4) Ser216 amide hydrogen is strongly hydrogen-bonded to a water molecule with a distance of 2.00 Å; the former is also found to be strongly hydrogen-bonded to Glu115, 1.67 Å. Furthermore, Arg221 is also found to be doubly hydrogen bonded to Glu115. The distance between the sulfenic acid sulfur and the amide nitrogen is quite large, 3.83 Å, which explains the unfeasibility of the direct mechanism.

Since previous proposed mechanism has highlighted the role of His214 hydrogen-bonding to Cys215 carbonyl oxygen, we investigated this hydrogen-bond in the optimized RC. As can be seen in Figure 6.4, His214 is weakly hydrogen bonded to Cys215; instead it is strongly hydrogen bonded to Asn111. Therefore, we considered optimizing new RC' structure upon decreasing the distance between the two residues. It is important to mention that mutating His214 was found to increase the pK_a of Cys216.²⁴ In model I, see Figure 6.4, the RC' structure was found to lie just 1.8 kJ mol⁻¹ lower in energy than RC. More importantly, the His214H⁺...O_Cys216 distance is now 2.09 Å showing strong interaction. Furthermore, His214 is also still hydrogen bonded to Asn111 with a distance of 2.10 Å. More importantly, other active site interaction remain similar to RC with minute differences such as the hydrogen-bond between Ser216 amide and the water molecule as well as the water...Glu115 have become slightly stronger with a distance of 1.94 Å and 1.54 Å, respectively.

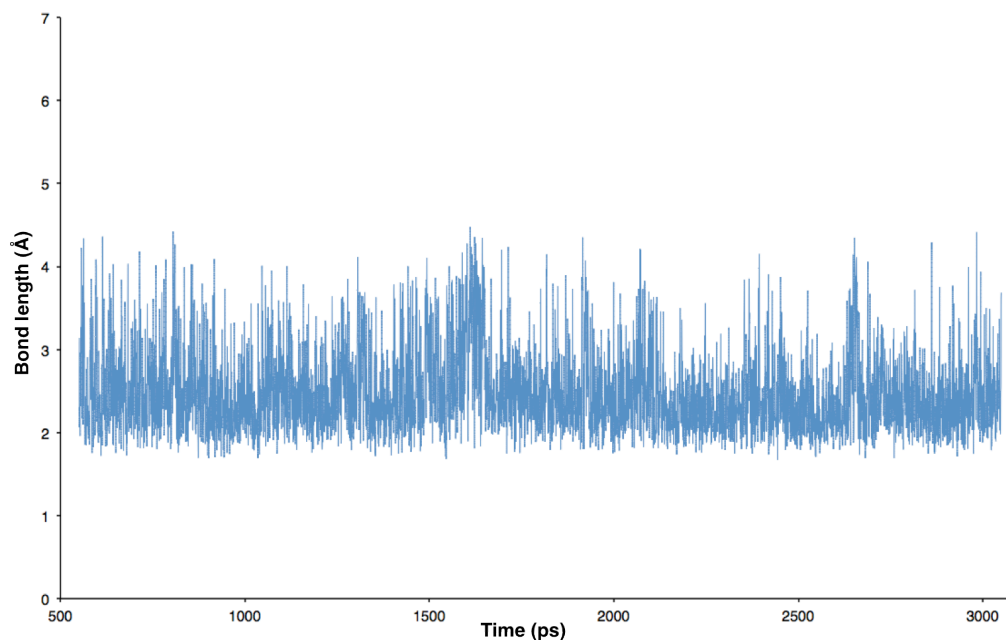


Figure 6.2. Plot of the hydrogen-bond distance between His214 and Cys215 with respect to time in the MD simulation.

In order to rationalize our choice of RC', we analyzed the consistency of this hydrogen bond interaction during the MD simulation. As can be seen in figure 6.2, the distance between the His214 hydrogen and the carbonyl oxygen is mainly fluctuating between 2 Å to 3 Å with less fluctuation around 3 Å to 4 Å. This clearly shows the strength of this hydrogen bond, fluctuating between strong to weak but it always exists.

6.3.2 Formation of Sulfenyl-Amide via Direct Mechanism

Although as indicated in the introduction that Sarma et al.²⁶ as well as Defelipe et al.²⁸ have studied the formation of sulfenyl-amide via direct mechanism showing an unfeasible high barrier of over 200 kJ mol⁻¹, we first considered reexamining this possibility in protein environment using both models. Especially, our MD simulations and QM/MM optimized structure suggest the presence of a water molecule that might mediate this reaction via a 6 member ring TS. However, such direct reaction was not found to occur, as we could not optimize a transition state that correspond to this mechanism. More

importantly, all attempted calculations have shown that the direct proton transfer to the sulfenic acid oxygen did not lead to sulfenyl-amide formation instead all calculations collapse back to form the reactive complex. It is important to mention that the obtained TS in Defelipe et al. QM/MM study shows a quite large distance of 3.30 Å and 2.36 Å for S \cdots N and S \cdots O, respectively.²⁸

All previous results suggest the need for investigating other possible mechanisms. Furthermore, for the nitrogen atom of the S216 to act as nucleophile it has to be first deprotonated, otherwise it is considered as a poor nucleophile as the π conjugation with the carbonyl group lead to resonance structure with positive charged nitrogen atom.

By investigating the MD generated structures, we found that a water molecule always bridge a hydrogen-bond interaction between the Ser216 amide hydrogen and Glu115 oxygen suggesting alternative mechanism. Indeed, this interaction was found to be consistent during the simulation.

6.3.3 Formation of Sulfenyl-Amide via Stepwise Mechanism

As mentioned earlier, RC and RC' active site interactions as well as MD results suggest the possibility of activating Ser216 amide nitrogen via Glu115. This step was found to be enzymatically feasible in with a barrier of 46.8 kJ mol⁻¹. Surprisingly, the TS1 structure shows triple proton transfers in which Ser216 amide hydrogen is shared with the water molecule oxygen with a distance of 1.27 Å and 1.23 Å, respectively. Furthermore, the water molecule proton is already transferred to Glu115 that has broken one of its hydrogen bonds to Arg221. The third proton transfer is between His214 and the carbonyl oxygen. Indeed, the resulting structure, I1, shows a complete proton transfer from His214 to Cys215 carbonyl oxygen, shifting the double bond to be between the amide nitrogen and carbonyl carbon forming a novel iminol intermediate. Definitely the nucleophilicity of the iminol nitrogen is much higher compared to the amide form. Notably, this would facilitate the nucleophilic attack on the sulfenic acid sulfur and the subsequent formation

of sulfenyl-amide intermediate. A similar mechanism of amide-iminol tautomerization has been recently proposed to occur in the case of succinimide formation in the C-terminal peptide group of aspartic acid residue.^{40,41}

These results suggest a new role for both His214 and Glu115 in sulfenyl-amide formation. The optimized structure of I1 was found to be slightly less stable than RC' with an energy difference of 21.3 kJ mol⁻¹.

The same step was reexamined starting from RC in which a weak hydrogen-bond interaction exists between His216 and Cys215. However, no stable I1 intermediate was found to exist where in the optimization the proton is transferred back to the amide nitrogen.

Although the iminol N is nucleophilic in I1, the distance between the Cys215S and Ser216 backbone N is quite large, 3.56 Å. This implies the need for structural rearrangements in the active site prior to subsequent reactions. Therefore, we considered optimizing another intermediate (I1') which was found to be more stable than I1, 22.7 kJ mol⁻¹ lower in energy with respect to I1. More importantly, the S...N distance has largely decreased to 2.78 Å allowing for following reaction. This rearrangement occurs upon changing the CCCS dihedral angle from 99.62° to 130.20°. Furthermore, the Glu115 is still hydrogen-bonded to the water molecule, which is now quite closer to the sulfenic acid oxygen allowing for proton transfer.

Recently the non-bonded interactions in protein sulfur species are gaining more attention and they are implicated to play a crucial role in sulfenic acid chemistry.⁴² Therefore we considered investigating the S...N interaction in I1', using Quantum Theory Atoms-in-Molecules (QTAIM) to gain further insights. In Bader's QTAIM theory the presence of critical point between two atoms reflect the existence of chemical bonding between them.⁴³ Therefore, we performed AIM analysis on I1 and I1'. Indeed, the generated molecular graphs for I1' indicated the presence of critical bond between S...N with electron density ρ of 0.016. Furthermore, we also considered the Laplacian $\nabla\rho^2$ of

this critical point to indicate the strength of the interaction, which was found to be 0.017, indicating the presence of a weak interaction between the two atoms. In general, the non-bonded electrostatic interaction between S and N atoms have been previously reported.^{44,45} The nature of this interaction has been suggested to occur due to the shift of electron density from the N lone pair to the S antibonding orbital (σ^*). This shows the importance of iminol intermediate formation to allow for such interaction to occur.

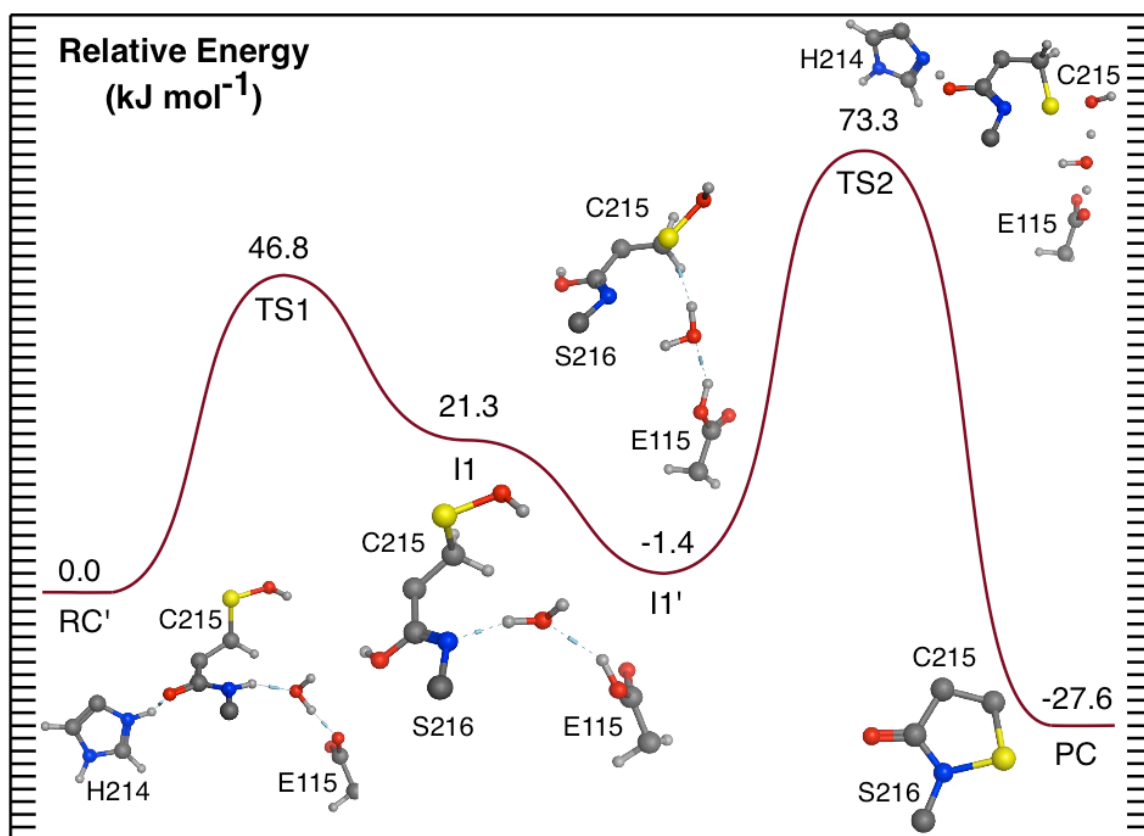


Figure 6.3. Potential energy surface obtained for the formation of sulfenyl-amide from sulfenic acid via iminol intermediate, see computational methods.

In order to investigate this shift in electron density between the S \cdots N, we used NBO analysis on I1 and I1'. The negative charge on the sulfur atom has increased upon I1' formation from +0.52 to +0.48 in I1 and I1', respectively. Furthermore the electron density on the sulfenic acid oxygen has also been increased from -0.91 to -0.94.

Consequently, the electron density on the N atom has decreased from -0.67 to -0.64 in I1 and I1', respectively. This clearly indicates the charge transfer and the interaction between the N lone pair and the sulfur σ^* .

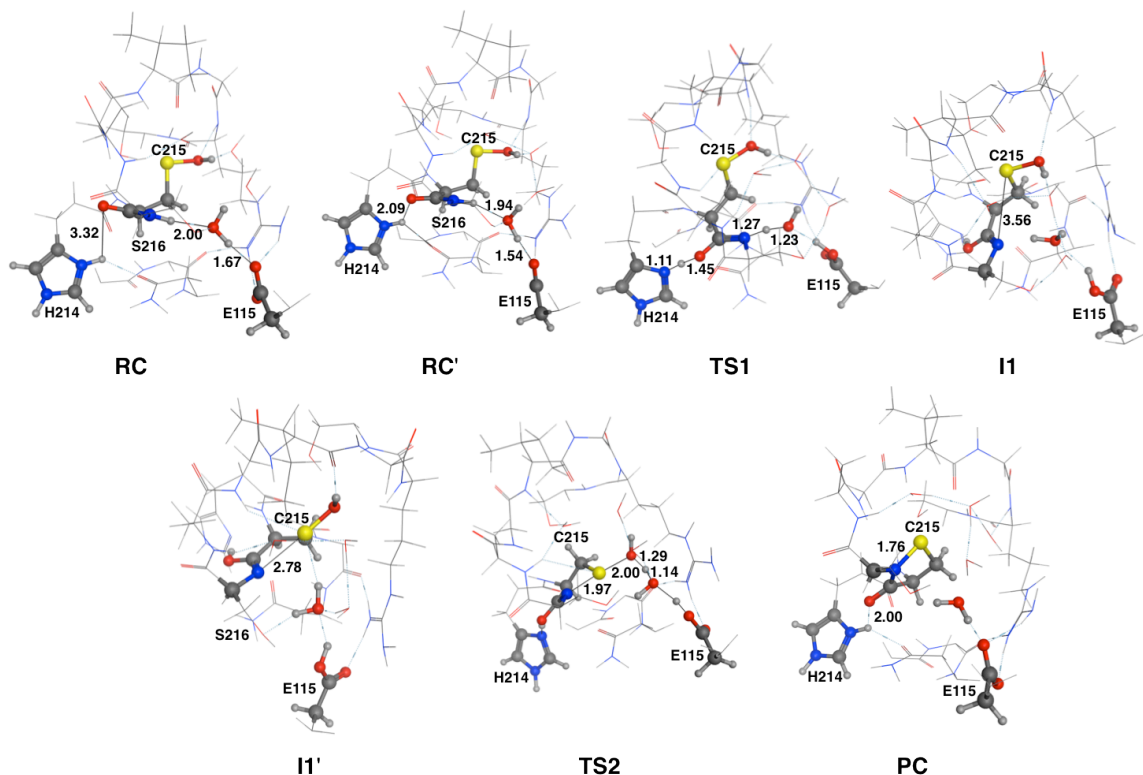


Figure 6.4. Stationary points obtained for model I for the sulfenyl-amide formation mechanism from sulfenic acid. All atoms in the high QM layer are included in the representation. However, only the highlighted residues are the ones involved in the reaction.

The last step of the mechanism was found to occur in a concerted fashion as shown in TS2, in which the S \cdots N bond is formed concomitantly with S \cdots O dissociation upon a proton transfer from Glu115 via water molecule. In TS2, the S \cdots N distance is reduced to 1.97 Å while the S \cdots O bond is elongated to 2.00 Å. Furthermore, the water molecule proton is partially transferred to the sulfenic acid oxygen with a distance of 1.29 Å and

1.14 Å with respect to molecular oxygen and sulfenic acid oxygen, respectively. The barrier for this step was found to be 73.3 kJ mol⁻¹ from RC.

Finally, the sulfenyl-amide species was found to lie -27.6 kJ mol⁻¹ lower in energy with respect to RC forming a stable product as experimentally proposed. In PC, the S...N distance is 1.76 Å in agreement with the X-ray crystal structure, 1.7 Å.¹⁷ Furthermore, His214 has restored its proton from the carbonyl carbon.

6.3.4 Formation of Sulfenyl-Amide using Neutral His214

In order to confirm previous results and especially the proposed role of His214 as an acid/base, we also considered modifying model I in which we included a neutral His. However, no stable II' intermediate was found to occur emphasizing the newly proposed role. It is important to mention as indicated in the introduction that the mutation of the invariant His increased the pK_a of the catalytic Cys which suggest the presence of a protonated His in the active site.

For completeness, we also considered the proton transfer from a neutral His214 with a concomitant abstraction of another proton from the surrounding residues. Investigating the surroundings of His214 suggested the possibility of hydrogen-bond network including the tetrad His214, Tyr124, His173 and Arg156. Therefore, we considered expanding the modified model I by including previous residues in the QM. However, based on our current model, there is no stable II intermediate.

Although PTP1B has been examined extensively especially regarding to the dephosphorylation mechanism, less studies have been performed to understand the regulatory oxidation mechanism. Therefore, based on our obtained results several experimental results can be performed to confirm our suggested mechanism, for instance the effect of His214 mutation on sulfenyl-amide, sulfinic and sulfonic acid formation. Furthermore, the role of His214 as an acid could be confirmed by mutating His to Asn to

preserve hydrogen-bonding interaction with the carbonyl oxygen. In addition, pH and kinetic studies during the oxidation mechanism would provide useful insights.

Our results show that the amide-iminol tautomerization is a more energetically favorable pathway over the unfeasible direct cyclization, which could be transferable to other amide nitrogen cyclization mechanisms in other PTP's as well as in proteins in general. Furthermore, the importance of the non-bonded interaction of cysteine sulfur in proteins is highlighted as a precursor to sulfenyl-amide formation. The rate-limiting step for this reaction is the formation of sulfenyl-amide from a stable iminol intermediate.

6.4 Conclusions

In this study the atomistic description of sulfenyl-amide formation from sulfenic acid in PTP1B has been elucidated. Several computational modeling techniques including molecular dynamics (MD), quantum mechanical/molecular mechanical (ONIOM) calculations and quantum theory atoms-in molecules (QTAIM) have been employed in cooperative fashion. More specifically, MD simulation on the solvated PTP1B sulfenic acid crystal structure was performed to generate potential Michaelis complexes. The MD generated structures were grouped into 5 main groups based on the root mean square deviation (rmsd) of the active site residue. The average structures of the highest populated clusters was selected for subsequent ONIOM calculations. Furthermore, several active site interactions including the indirect hydrogen bonding between Ser216 amide nitrogen and Glu115 via water molecule as well as the Cys215 carbonyl oxygen His214 hydrogen-bonding were investigated during the course of the simulation to confirm the stability of these interactions.

Three different models were generated from the MD average structure including model I which has protonated His214. For completeness, modified QM/MM structures of model I was also considered including a neutral His214 as well as upon increasing the size of QM layer.

The optimized structure of RC indicates the presence of a weak hydrogen bond between Cys215 carbonyl oxygen and His214. Therefore, another reactive complexes (RC') was optimized in which this hydrogen-bond become stronger. This has been justified by investigating the consistency of this interaction during the MD simulation. Furthermore, RC' was found to be slightly more stable than RC. Using both RC' and RC as starting structures, our calculation indicated the unfeasibility of the sulfenyl-amide direct formation. Instead ONIOM calculations as well as the MD results suggested the stepwise mechanism. In which the Ser216 amide N is first deprotonated by Glu115 with an energy barrier of 46.8 kJ mol⁻¹. The formed iminol intermediates were found to lay slightly higher in energy with respect to RC', 21.3 kJ mol⁻¹. The nucleophilicity of the generated iminol nitrogen is much higher compared to the amide form facilitating the nucleophilic attack on the sulfenic acid sulfur and subsequent formation of sulfenyl-amide intermediate. The same step was also examined starting from RC. However, no stable I1 intermediate was found to occur proposing a new role for His214 as an acid.

A second iminol intermediate (I1') was also considered where the S...N distance has largely decreased to 2.78 Å. The nature of the S...N interaction in I1' was analyzed using QTAIM and NBO indicating the presence of a weak interaction between the two atoms. Furthermore, NBO analysis confirmed the nature of this interaction in which there is a charge transfer between the N lone pair and the sulfur σ^* .

Finally, the S...N bond is formed concomitantly with S...O dissociation forming a stable sulfenyl-amide intermediate. Notably, the two modified forms of model I including a neutral His214 showed that no stable iminol intermediate is formed.

Our results suggest the formation a cyclic sulfenyl-amide from sulfenic acid in stepwise fashion using an amide-iminol tautomerization reaction in which His214 and Glu115 act as an acid and base catalysts, respectively. Finally, we think this work is important to the field of PTP1B drug discovery, as one of the main ways to inhibit PTP1B is to stabilize the sulfenyl-amide inactive oxidized form. Therefore, knowing the exact

mechanism and the role of the active site residues would be crucial for future drug development.

6.5 References

- (1) Feldhammer, M.; Uetani, N.; Miranda-Saavedra, D.; Tremblay, M. L. *Crit. Rev. Biochem. Mol. Biol.* **2013**, *48*, 430.
- (2) Tonks, N. K. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 833.
- (3) Yip, S. C.; Saha, S.; Chernoff, J. *Trends Biochem. Sci.* **2010**, *35*, 442.
- (4) Tonks, N. K. *FEBS J.* **2013**, *280*, 346.
- (5) Bence, K. K.; Delibegovic, M.; Xue, B.; Gorgun, C. Z.; Hotamisligil, G. S.; Neel, B. G.; Kahn, B. B. *Nat. Med.* **2006**, *12*, 917.
- (6) Gonzalez-Rodriguez, A.; Gutierrez, J. A. M.; Sanz-Gonzalez, S.; Ros, M.; Burks, D. J.; Valverde, A. M. *Diabetes* **2010**, *59*, 588.
- (7) Ali, M. I.; Ketsawatsomkron, P.; de Chantemele, E. J. B.; Mintz, J. D.; Muta, K.; Salet, C.; Black, S. M.; Tremblay, M. L.; Fulton, D. J.; Marrero, M. B.; Stepp, D. W. *Circ. Res.* **2009**, *105*, 1013.
- (8) Chatelain, E. H.; Dupuy, J. W.; Letellier, T.; Dachary-Prigent, J. *Cell. Mol. Life Sci.* **2011**, *68*, 2603.
- (9) Zhu, S.; Bjorge, J. D.; Fujita, D. J. *Cancer Res.* **2007**, *67*, 10129.
- (10) Dube, N.; Cheng, A.; Tremblay, M. L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 1834.
- (11) Combs, A. P. *J. Med. Chem.* **2010**, *53*, 2333.
- (12) Lessard, L.; Stuiblé, M.; Tremblay, M. L. *BBA-Proteins Proteomics* **2010**, *1804*, 613.
- (13) Sobhia, M. E.; Paul, S.; Shinde, R.; Potluri, M.; Gundam, V.; Kaur, A.; Haokip, T. *Expert Opin. Ther. Patents* **2012**, *22*, 125.

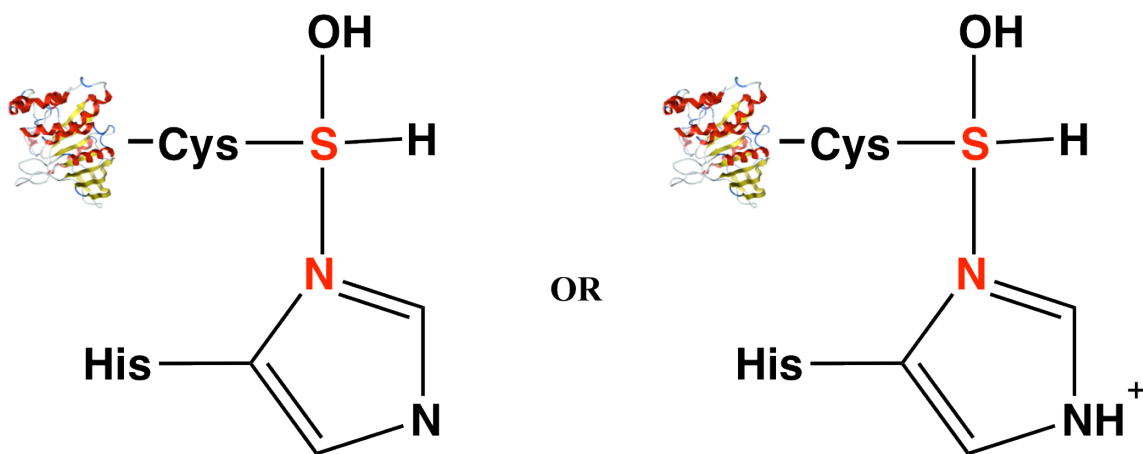
- (14) Brandao, T. A. S.; Hengge, A. C.; Johnson, S. J. *J. Biol. Chem.* **2010**, *285*, 15874.
- (15) Salmeen, A.; Andersen, J. N.; Myers, M. P.; Meng, T. C.; Hinks, J. A.; Tonks, N. K.; Barford, D. *Nature* **2003**, *423*, 769.
- (16) Parsons, Z. D.; Gates, K. S. *Biochemistry* **2013**, *52*, 6412.
- (17) van Montfort, R. L. M.; Congreve, M.; Tisi, D.; Carr, R.; Jhoti, H. *Nature* **2003**, *423*, 773.
- (18) Sivaramakrishnan, S.; Cummings, A. H.; Gates, K. S. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 444.
- (19) Shetty, V.; Neubert, T. A. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1540.
- (20) Yang, J.; Groen, A.; Lemeer, S.; Jans, A.; Slijper, M.; Roe, S. M.; den Hertog, J.; Barford, D. *Biochemistry* **2007**, *46*, 709.
- (21) Lee, J. W.; Soonsanga, S.; Helmann, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 8743.
- (22) Dokainish, H. M.; Gauld, J. W. *Biochemistry* **2013**, *52*, 1814.
- (23) Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6238.
- (24) Zhang, Z. Y.; Dixon, J. E. *Biochemistry* **1993**, *32*, 9340.
- (25) Haque, A.; Andersen, J. N.; Salmeen, A.; Barford, D.; Tonks, N. K. *Cell* **2011**, *147*, 185.
- (26) Sarma, B. K.; Mugesh, G. *J. Am. Chem. Soc.* **2007**, *129*, 8872.
- (27) Sarma, B. K. *J. Mol. Struct.* **2013**, *1048*, 410.
- (28) Zeida, A.; Guardia, C.; Lichtig, P.; Perissinotti, L.; Defelipe, L.; Turjanski, A.; Radi, R.; Trujillo, M.; Estrin, D. *Biophys. Rev.* **2014**, *1*.
- (29) MOE, version 2012.10 (2012) Chemical Computing Group Inc., Montreal.

- (30) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (31) Labute, P. *Proteins* **2009**, *75*, 187.
- (32) 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.* **1996**, *118*, 2309.
- (33) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2013**, *3*, 198.
- (34) Gerber, P. R.; Muller, K. *J. Comput. Aided Mol. Des.* **1995**, *9*, 251.
- (35) Gaussian 09, Revision D.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
- (36) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (37) Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. *J. Am. Chem. Soc.* **2012**, *134*, 4743.

- (38) Polyak, I.; Reetz, M. T.; Thiel, W. *J. Phys. Chem. B* **2013**, *117*, 4993.
- (39) Biegler-König F. AIM2000 Version 2.0. University of Applied Science: Bielefeld, Germany: 2002.
- (40) Takahashi, O., Intramolecular cyclization of aspartic acid residues assisted by three water molecules: a density functional theory study. In *Tyrosine and Aspartic Acid: Properties, Sources and Health Benefits*; Jones, J. E. M., D.M., Ed.; Nova Science Publishers: New York, USA. 2012, p 131.
- (41) Takahashi, O. *Health* **2013**, *5*, 2018.
- (42) Iwaoka, M.; Isozumi, N. *Molecules* **2012**, *17*, 7266.
- (43) Nakanishi, W.; Hayashi, S.; Narahara, K. *J. Phys. Chem. A* **2008**, *112*, 13593.
- (44) Adhikari, U.; Scheiner, S. *Chem. Phys. Lett.* **2011**, *514*, 36.
- (45) Scheiner, S. *J. Chem. Phys.* **2011**, *134*.

Chapter 7

Pseudo-hypervalent Sulfur Intermediate as a Protective Mechanism in Peroxiredoxin Enzymes



Sulfurane

7.1 Introduction

Peroxiredoxins (Prxs), also known as thiol peroxidases, are a class of ubiquitous enzymes with central roles in both the regulation of signaling pathways as well as being a potent antioxidant.¹⁻³ Their main function, foremost, is to catalyze the reduction of peroxide substrates such as hydrogen peroxide (H_2O_2) and alkyl-hydro-peroxide (ROOH).⁴ Notably, along with their ubiquity, these enzymes are vastly abundant in cells, up to 1% of the total soluble protein and reach catalytic rates in the order of $\sim 10^7 \text{ M}^{-1} \text{ S}^{-1}$.¹ In fact, these remarkable catalytic efficiencies have credited Prxs to reduce $\sim 90\%$ of the mitochondrial and approaching 100% of the cytoplasmic H_2O_2 .^{3,5}

Prxs are classified into 6 main groups based on sequence similarity: Prx1, Prx6, Prx5, Tpx, BCP and AhpE (in *Mycobacterium tuberculosis*).^{1,4} All groups contain a reducing thioredoxin (Trx) fold with differences in their oligomeric states, from monomer to dodecamer.^{1,6,7} In addition, they also possess high similarities in their active site structures including the catalytic cysteine (C_p), proline (Pro), threonine (Thr)/serine (Ser) and arginine (Arg).^{1,8,9} Furthermore, based on mechanistic differences, Prxs are further divided into three main groups: 1-Cys, typical 2-Cys and atypical 2-Cys.⁶ This classification is dependent mainly on the existence and location of a second mechanistic cysteinyl residue, known as the resolving cysteine (C_R).¹ For instance, 1-Cys Prxs does not contain C_R .^{1,6} In typical 2-Cys Prxs (the largest class of Prxs), however, the enzyme function is dependent on the presence of an intact homodimer where C_p and C_R are located on different monomers.^{4,10} In contrast, in atypical 2-Cys Prxs, the two mechanistic cysteinyl residues are located on the same monomer.^{4,11}

All classes of Prxs have a common catalytic cycle involving three main steps:^{1,4} 1) the peroxidation step, in which C_p nucleophilically attacks the peroxide substrate forming reactive Prx-sulfenic acid ($\text{S}_p\text{-OH}$) intermediate, 2) the resolution step, where either the resolving C_R or external thiol in 2-Cys Prx or the 1-Cys Prx, respectively, reduce ($-\text{S}_p\text{-OH}$) to form an intramolecular or intermolecular disulfide bond, 3) the recycling step that

leads to the regeneration of the active site via the reduction of the disulfide by external thiol such as Trx.¹²

Several X-ray structures of the Prx family reveal similarities of H₂O₂ (HO^A–O^BH where O^A is in close proximity to C_P) binding in their active sites, preserving several key interactions.^{9,13-15} More specifically, the conserved Arg always is hydrogen bonding to C_P.^{6,9} This allows for the stabilization of the nucleophilic thiolate as well as it is hydrogen bonded to the H₂O₂ oxygen (O^A).^{6,9} The conserved Pro is thought to protect C_P from interacting with solvent. In addition, the Thr residue acts as a hydrogen bond acceptor with respect to O^A.^{6,9} Furthermore, the H₂O₂ molecule has been shown to be well positioned for attack via the previously described hydrogen bonds and two other H-bonds with the backbone amide of C_P and its neighboring residue.^{6,9} Although existing X-ray structures highlights the roles of the active site residues, their specific roles in catalysis remains vague.⁹ More importantly, the nature of the proton donor needed to neutralize the hydroxide-leaving group is unknown.⁹

In general, sulfenic acids are fundamental intermediates for numerous redox processes in proteins.¹⁶ It is highly reactive and can undergo several reactions as a nucleophile or electrophile.^{16,17} Thus it can be readily overoxidized to sulfinic and subsequently sulfonic acids in the presence of H₂O₂ or other oxidizing agents.^{16,17} This modification, in general, is considered irreversible and leads to protein deactivation.¹⁸ Notably, the second step of the mechanism that is characterized by the reduction of –S_P-OH by C_R, is competing with its overoxidation especially under oxidative stress.^{4,19} In Prx, the overoxidation of the typical 2-Cys occurs at a higher rate than the atypical 2-Cys or 1-Cys subclasses.²⁰ Notably, this occurs as a result of the location of the C_R on the adjacent monomer which requires large structural rearrangement before sulfenic reduction allowing for –S_P-OH interaction with a second peroxide.⁴ Fortunately, a unique enzyme (sulfiredoxin) was found to reduce typical 2-Cys Prx sulfinic acid back to sulfenic.²⁰⁻²² This enzyme was only found in eukaryotic organisms and was found to be highly specific to the typical 2-

Cys Prx implying the presence of alternative mechanisms to protect the typical 2-Cys in bacteria and archaea as well as in 1-Cys Prx from overoxidation and subsequently their inactivation.²⁰

In Archaea, the typical 2-Cys Prx (ApTPx) was recently proposed, based on X-ray crystallography, to be protected from overoxidation via a unique hypervalent sulfurane intermediate.²³ Nakamura *et al.*²³ characterized an X-ray structure at high resolution of 1.77 Å for a unique intermediate with an electron density representing a covalent interaction between S of C_P50 and the Nδ1 of His42 of approximately 2.2 Å. Small models calculations were performed at the B3LYP/6-31G* level in the absence of the active site residues in order to describe the chemical nature of this interaction.²³ In result, two possible sulfurane intermediates were suggested to occur where the C_P50 sulfur was covalently bound to His42, a hydrogen ligand and a hydroxyl group.²³ The two proposed structures differ only by the ionization state of His42. Their calculations show that the removal of hydrogen ligand from the C_P sulfur leads to the disappearance of S···N interaction.²³ However, recently we assessed the performance of several density functional theory methods in dealing with hypervalent sulfur species in biological systems using different size basis set; our calculations show that the removal of hydrogen ligand does not break the interaction using other DFT functionals (manuscript in preparation). Furthermore, after characterizing the chemical structure of the intermediate the authors also investigated the mechanism of its formation. They suggested that sulfurane could not be formed from the direct interaction of sulfenic acid with an imidazole as the obtained barrier was found to be 215.9 kJ mol⁻¹.²³ Thus, an alternative mechanism where a thiol, a poor nucleophile, reduces the peroxide directly was proposed forming the sulfurane intermediate in concomitant step.²³ It is important to mention that a similar interaction with a histidyl residue was previously reported in the human 1-Cys Prx (hORF6) however the S···N distance was suggested to reflect hydrogen-bonding interaction.²⁴ In addition, this histidyl residue is not conserved in all Prxs.²³

Recently, there has been several hypervalent sulfur species reported in biological systems. For instance, in methionine sulfoxide reductase (Msr), a different sulfurane intermediate has been suggested computationally to occur.²⁵ In addition, in Srx mechanism, two-hypervalent sulfur-based intermediate have been also suggested experimentally.²⁶ More recently, a pseudo-hypervalent divalent sulfur species has been shown to play a critical role in protein architecture and functions.²⁷ This type of weak interaction ($S \cdots X$) between divalent sulfur and oxygen, nitrogen or sulfur atoms is quite similar to halogen-bonds.²⁸ This σ hole bonding originates from a positive electrostatic potential on the sulfur center allowing for orbital interaction between the X lone pair and the anti-bonding orbital of the sulfur atom, $n_X \rightarrow \sigma^*_S$.²⁹ Iwaoka et al. have characterized this weak interactions in four proteins: phospholipase, ribonuclease A, insulin and lysozyme, suggesting their importance in proteins in general.²⁷

In this study we first used the hybrid quantum mechanical/molecular mechanical calculations (QM/MM) calculations to elucidate the first step of the mechanism. This revealed thermodynamics and the roles of active site residues during the reduction of peroxide and sulfenic acid formation. Second, we investigated the pathway of formation of the currently proposed hypervalent sulfurane in ApTPx. Third, using QTAIM and NBO analysis we identified the degree of covalency of the $S \cdots N$ interaction in the enzyme's active site. Fourth, we also considered the formation of this intermediate in hORF6 to determine if this is feasible to occur in similar enzymes as a general intrinsic precaution to over oxidation. Finally, details into the overoxidation mechanism of Prx-sulfenic acid to sulfinic acid have been determined.

7.2 Computational Methods

7.2.1 Structural Preparation and Molecular Dynamic Simulations.

The Molecular Operating Environment (MOE) software package was used to prepare the starting X-ray structures for MD simulations as well as to analyze the simulations

results.³⁰ The NAMD molecular dynamic software was used to run the MD simulations.³¹ For ApTPx, the H₂O₂ bound X-ray of the C207S mutated crystal structure was obtained from the protein data bank (PDB: 3A2V)¹³ and used as starting configuration for simulation using one monomer of the decameric form. In addition, one monomer of the human 1-Cys Prx (hORF6) Prx-sulfenic acid X-ray structure (PDB:1PRX)²⁴ was used as starting structure.

Prior to MD simulations, errors in both crystal structures were automatically corrected using the structural preparation application in MOE. In ApTPX, the position of His42 was modified manually allowing for His42⁺·Cys50 interaction. The protonated 3D application in MOE was used to add the missing protons in the crystal structures. For the active site residues especially Cys50 and His42 the protonation states were determined based on their pK_a values obtained via PROPKA program.^{32,33} The generated structures were solvated up to 10 Å beyond every protein atom. Then, were minimized using AMBER12:EHT force field.^{34,35} Later, the minimized structures were used as starting points for 500 ps equilibration MD simulations to generate thermally relaxed structures in similar protocol as previously used.^{36,37} In the MD, 2 fs time step was used. The PME method was used to calculate columbic interactions and the van der Waals interactions were truncated at 8-10 Å.

In hORF6, the S⁺·O bond in the sulfenic acid intermediate was broken to generate the reduced form of the enzyme. The peroxide substrate was manually docked in the active site prior to simulation. In addition, a third MD simulation was performed to simulate the overoxidation mechanism. First, the QM/MM optimized structure of ApTPx sulfenic intermediate was used as starting structure. Second, peroxide substrate was also manually docked in the active site. The Cys50⁺·H₂O₂ distance was restrained during the minimization to 3 Å allowing for substrate binding adjustment in the active site. Later, similar solvation, minimization and MD simulation protocols were used as described above.

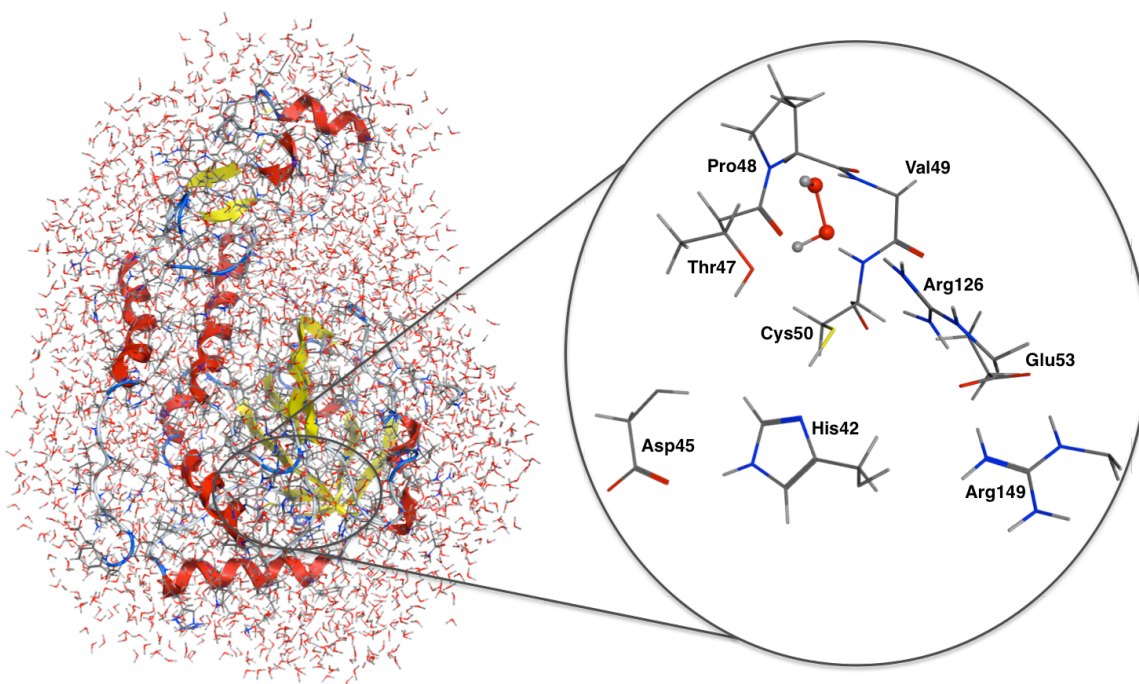


Figure 7.1. Illustration of the ApTPx QM/MM model used obtained from MD average structure. The QM layer atoms are highlighted and shown in sticks representation.

7.2.2 QM/MM Models and Calculations.

All calculations were performed within the ONIOM formalism using the Gaussian 09 suite of programs.³⁸ The last structures of the previous MD simulations were used as starting structures for QM/MM calculations.³⁹ The QM atoms were optimized at the M06-2X/6-31G(d,p) which was selected based on previous assessment of several level of theories to optimize the proposed intermediates in ApTPx (manuscript in preparation).⁴⁰ The rest of the monomer was calculated using the AMBER96 forcefield as implemented in Gaussian09.³⁵ Frequency calculations were used to confirm the nature of the optimized stationary points using the QM layer only, due to model size as previously used.^{36,37} Relative energies were obtained using the ONIOM electronic embedding formalism using single points calculations at the ONIOM(M06/6-311+G(2d,p):AMBER96) In addition, other basis sets were also tested such as 6-311G(2d,p), 6-311+G(2df,p) and 6-

311++G(3df,3pd) using the M06 functional as well as ONIOM(MP2/6-31G(d,p):AMBER), (See Appendix Figure A7).

Three QM/MM models structures were generated from the last conformer of the equilibrated MD simulations. All MM atoms 15 Å away from Cys50 sulfur were kept fixed in the optimization. For ApTPx, the QM layer included H₂O₂, Cys50, Pro48, Thr47, the backbone of Val40 and the R groups of His42, Asp45, Arg126 and Arg149, see Figure 7.1. For ApTPx overoxidation state of sulfenic acid, the previously described residues were included in the QM layer as well as H₂O₂, Pro43, Ala44 and a water molecule (See Appendix Figure A8). Finally for hORF6, the QM was formed of H₂O₂, Cys47, backbone of Val46, Pro45, Thr44, His39, Pro40, part of Ser38 backbone and the R groups of Arg132, Arg155, Glu50 and 3 water molecules (See Appendix Figure A8).

In order to investigate the nature of the S···N interaction, the Quantum Theory Atoms-in-Molecules (QTAIM) was employed using AIM2000 program.⁴¹ Furthermore, natural bond analysis (NBO) was also used to determine the atomic partial charge of the optimized structures.

7.3 Results and discussion

7.3.1 Sulfurane intermediate formation in ApTPx.

In ApTPx, the QM/MM optimized reactive complex (RC) structure revealed similar interactions to X-ray structure⁹ as discussed in the introduction. Based on estimated pK_a's, our MD starting structure contained a neutral His42 and Cys50. However, in the QM/MM optimization Cys50 was found to be deprotonated, by transferring a proton to His42 that in turn transferred its proton to Asp45. Thus, the RC is formed of anionic Cys50, neutral His42 and Asp45, see Figure 7.2.

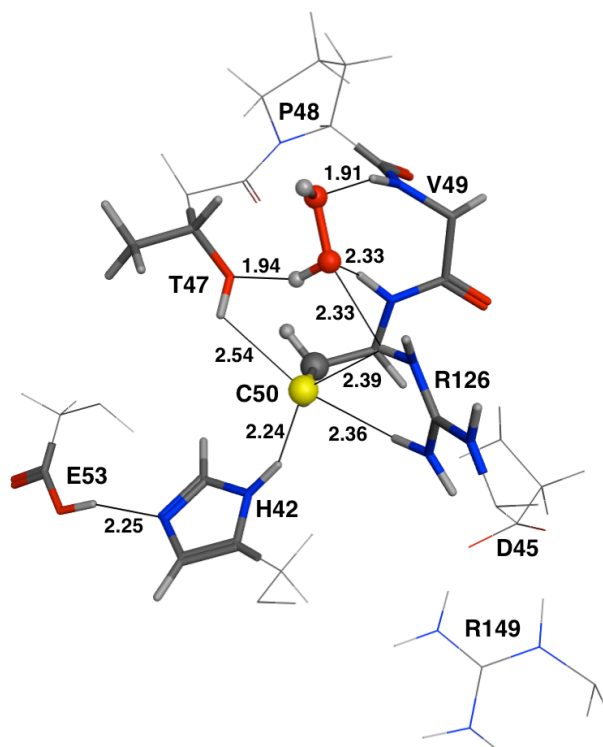


Figure 7.2. Illustration of the QM/MM optimized reactive complex of ApTPx. Only the QM layer is shown.

The nucleophilic Cys50 anionic sulfur is stabilized via four hydrogen bonds with Arg126, His42 and Thr49. Furthermore, the H₂O₂ binding in the active site was found to involve several strong hydrogen bonds including O_B···HN_{Val49}, O_A···HN_{Cys50}, O_A···Arg126 and O_AH···Thr49. Notably, these hydrogen bonds destabilize the H₂O₂ in the active site promoting catalysis as experimentally suggested. NBO analysis was used to compare the substrate's charges in the active site and in the absence of the active site using both in gas phase and protein environment (dielectric= 4). In absence of active site, the negative charge on H₂O₂ two oxygen's were found to be symmetric with values of -0.47 and -0.48 in gas phase and protein environment, respectively. In the active site the O–O bond is polarized with charges of -0.49 and -0.52 for O_B and O_A, respectively. The ∠H-OB-OA-H dihedral angle is reduced from 179.9° to 112.0°. In addition, Cys50 anionic sulfur has a charge of -0.63 and the _{Cys50}S···O_A distance is 3.2 Å.

The first step in the proposed mechanism is the reduction of the peroxide. By investigating this step, our calculations show that this can readily occur via a barrier of 82.1 kJ mol^{-1} , see Figure 7.3. In TS1, the $\text{H}_2\text{O}_2 \text{ O}_A \cdots \text{O}_B$ bond is elongated to 1.80 \AA . Notably, the $_{\text{Cys50}}\text{S} \cdots \text{O}_A$ is reduced to 2.21 \AA . Two intermediates were optimized from TS1, sulfenic acid (I1) and sulfenate intermediates (I2). However, I2 was found to be lower in energy than I1 by 15.0 kJ mol^{-1} . In I1, see Figure 7.4, a proton is being transferred from Arg126 to O_A . In addition, Thr47 is now hydrogen bonded to sulfenic acid oxygen with a distance of 1.71 \AA . I1 is 88.9 kJ mol^{-1} more stable than RC. In I2, a sulfenate, the $\text{S} \cdots \text{O}$ bond is slightly reduced with respect to I1 from 1.66 \AA to 1.64 \AA . Additionally, the charge on the oxygen has increased from -0.93 in I1 to -1.09 in I2. Notably, the oxyanion is stabilized with three strong hydrogen bonds with Thr47, Arg126 and the formed water molecule with distances of 1.56 \AA , 1.43 \AA and 1.93 \AA , respectively. I2 is $103.9 \text{ kJ mol}^{-1}$ more stable than RC. Since TS1 leads to the formation of either sulfenic or sulfenate, we considered optimizing TS2 to represent the interchange between the two possible intermediates. From I1, I2 can be readily formed via a low barrier of 4.8 kJ mol^{-1} . In TS2, a proton is being shared between Arg126 and sulfenic acid with distances of 1.27 \AA and 1.20 \AA , respectively.

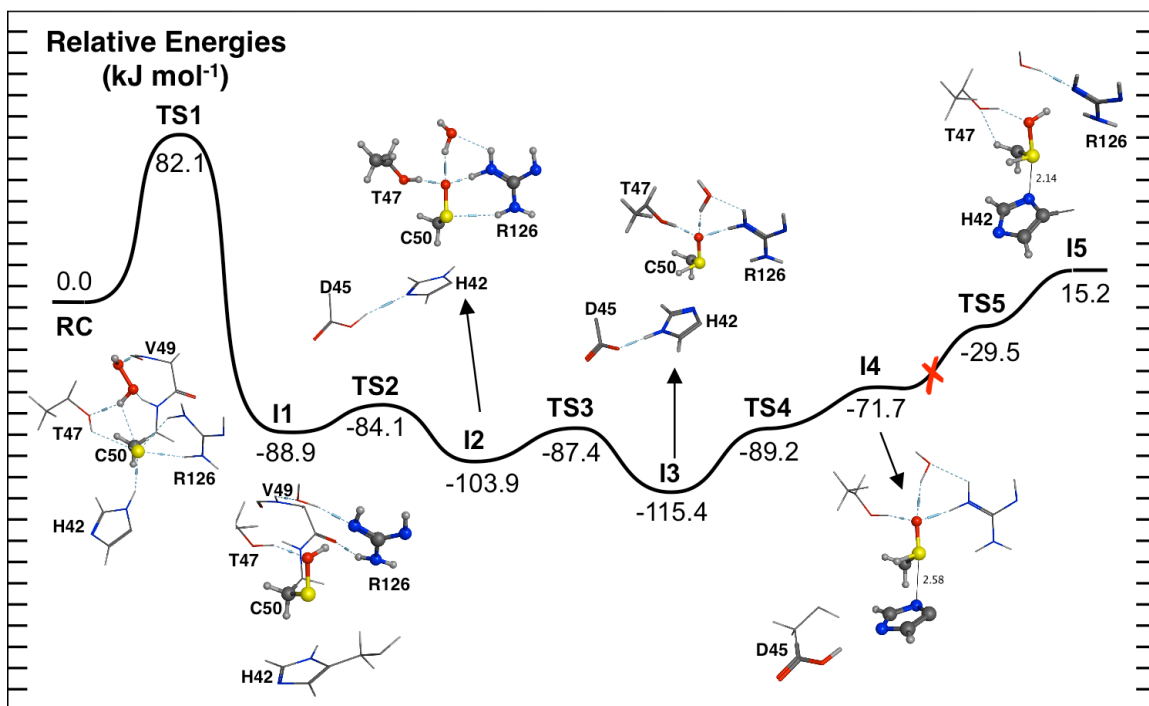


Figure 7.3. Potential energy surface obtained for the formation of pseudo-sulfurane intermediate in ApTPx.

In order to form the proposed sulfurane intermediate and allow for $\text{Cys50S}^{\cdots}\text{N}_{\text{His42}}$ interaction, His42 needs to lose its proton. Thus, based on His42 interactions in I2, this can only be achieved by transferring a proton to sulfenate sulfur to form sulfenic acid tautomer. In I2, the $\text{His42H}^{\cdots}\text{S}_{\text{Cys50}}$ distance is 2.70 Å which allows for direct transfer. This step was found to occur via TS3 with a low barrier of 16.5 kJ mol⁻¹. In TS3, a two proton transfer was found to occur in which the His42H is transferred to S_{Cys50} and Asp45 to His42. The optimized structure shows that the latter proton is 1.14 Å and 1.47 Å from His42 and Asp45, respectively. From TS3, a sulfenic acid tautomer (I3) was found to occur and be more stable than I2 by 11.5 kJ mol⁻¹ and lower than RC by -115.4 kJ mol⁻¹. In I3 optimized structure, the $\text{S}^{\cdots}\text{O}$ bond is reduced to 1.58 Å. As in I2, the sulfenic acid tautomer oxygen is stabilized by three hydrogen bonds. Notably, there is a weak interaction between $\text{His42N}^{\cdots}\text{S}_{\text{Cys50}}$ with a distance of 2.76 Å. This weak interaction ($n_{\text{N}} \rightarrow \sigma^*_{\text{S}}$) was analyzed using QTAIM in which a critical point was found to connect

His42 and Cys50 with electron density ρ of 0.020 and Laplacian $\nabla\rho^2$ of 0.017. NBO analysis of I3 shows that there is a positive charge on the Cys50 sulfur of 1.05 as well as a negative charge on the His42 nitrogen with a value of -0.63. The \angle N–S–O angle is 164.1° , which is close to X-ray structure.¹³

Although I3 represents a pseudo hypervalent structure similar to the X-ray obtained,¹³ the $\text{His42N}\cdots\text{S}_{\text{Cys50}}$ distance is still large. Therefore, we considered a subsequent reaction that may strengthen this interaction. This was found to occur via TS4 with a barrier of 25.2 kJ mol^{-1} . In TS3, a proton is transferred from His42 to Asp45 in which the $\text{H}\cdots\text{His42}$ and $\text{H}\cdots\text{Asp45}$ distances are 1.20 \AA and 1.36 \AA , respectively. For this TS, the $\text{His42N}\cdots\text{S}_{\text{Cys50}}$ was found to be slightly reduced to 2.75 \AA . The following intermediate (I4) was found to contain anionic His42 and neutral Asp45. Indeed, in I4, $\text{His42N}\cdots\text{S}_{\text{Cys50}}$ distance is reduced to 2.58 \AA which is in agreement with crystal structure. In addition, the sulfenic $\text{S}\cdots\text{O}$ distance was slightly increased to 1.58 \AA . The \angle N–S–O angle is increased to 166.1° . As previous, the nature of $\text{His42}\cdots\text{Cys50}$ interaction was investigated further using QTAIM. Our analysis was found to still reflect $\text{His42}\cdots\text{Cys50}$ weak interaction. However, the calculated density at critical point was found to be strengthened with respect to I3, $\rho=0.030$, as well as the Laplacian $\nabla\rho^2=0.023$. NBO analysis shows the increase of the negative charge on the sulfurane oxygen from -1.09 in I3 to -1.10. The charge on the sulfurane N was also found to increase to -0.68. In general I4 was found to be less stable than I3 by 43.7 kJ mol^{-1} . However, I4 is still more stable than RC by $-71.7 \text{ kJ mol}^{-1}$.

We also considered investigating a final intermediate (I5) upon transferring a proton from Arg126 to sulfurane oxygen via TS5, however this intermediate was found to be high in energy compared to I3 with a value of 15.2 kJ mol^{-1} . The energy difference between I3 and I5 was found to be $130.6 \text{ kJ mol}^{-1}$, which is unfeasible in enzymatic systems. More importantly, our QTAIM analysis shows that the $\text{His42}\cdots\text{Cys50}$ interaction is still noncovalent ($\rho=0.086$ and $\nabla\rho^2=0.083$) indicating that the X-ray obtained hypervalent sulfurane in ApTPx is formed due to noncovalent interactions.

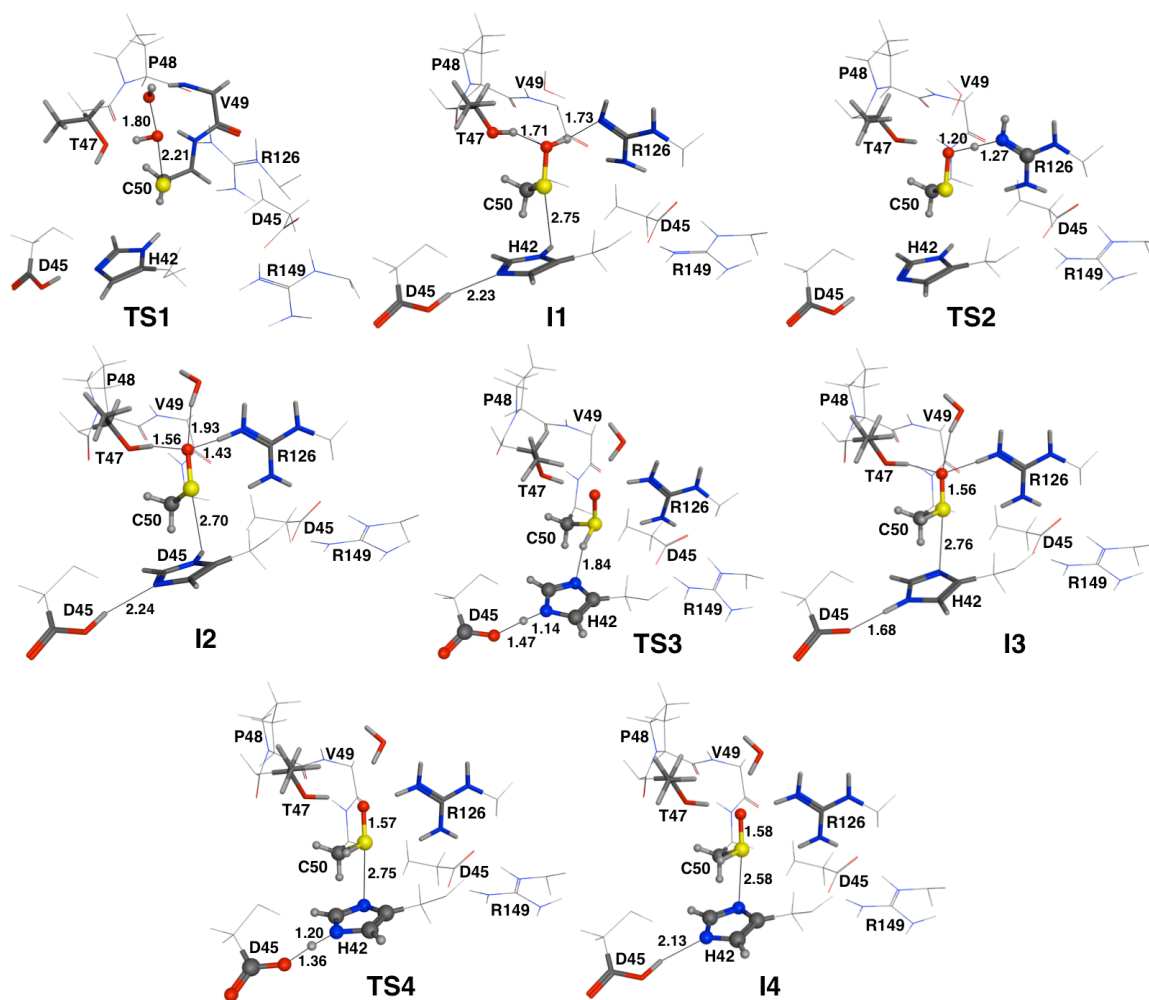


Figure 7.4. Stationary points obtained for the pseudo-sulfurane formation mechanism. Shown atoms represent the QM layer. However, only the highlighted are directly involved in the reaction.

The previous described mechanism provides a detailed mechanism to the formation of hypervalent sulfurane (I4) from sulfenic acid via low barriers multistep. Furthermore, in order for the last step in ApTPx to occur, which is the formation of disulfide bond, the formation of sulfurane has to be reversible allowing for sulfenic acid reduction. Our obtained potential energy surface shows that the sulfurane (I4) can be readily reduced

back to sulfenate (I2) or sulfenic acid (I1) via maximum barriers of 28.0 kJ mol⁻¹ and 31.3 kJ mol⁻¹, respectively.

7.3.2 Sulfenic acid oxidation mechanism in *ApTPx*.

Since the formation of sulfurane intermediate is proposed to protect sulfenic acid from oxidation, we also considered investigating the overoxidation mechanism in the active site. As indicated in the methods section, the sulfenic acid QM/MM optimized structure (I1) was used as starting point for equilibrated MD simulation. Two models were used in which His42 was neutral/protonated. However in our discussion, only the neutral His42 based model will be discussed in details as the obtained barrier for the protonated form shows unfeasible barrier of 183.3 kJ mol⁻¹.

In the MD simulation, the binding of H₂O₂ was found to only occur upon the rotation of the ∠Cα–Cβ–S–O dihedral angle, as experimentally suggested.⁴² In fact this might explain the role of sulfurane formation as the presence of His42···Cys50 noncovalent interaction would prevent such rotation and subsequently overoxidation. In the optimized QM/MM structure (RC_{Ox}), similar H₂O₂ interactions were found to occur including O_B···HN_{Val49}, O_A···HN_{Cys50}, O_A···Arg126 and O_AH···Thr49 with the following distances 1.99 Å, 2.06 Å, 2.57 Å and 1.68 Å, respectively. More importantly, the sulfenic acid S···O_A distance is 3.09 Å. Furthermore, the sulfenic acid is hydrogen bonded to His42, see Figure 7.5.

The overoxidation mechanism was found to occur via TS_{Ox} in which similar interactions to TS1 were observed. In TS_{Ox} the O_B···O_A distance is increased to 1.87 Å as well as the C_{ys50}S···O_A distance is reduced to 2.05 Å. The obtained barrier for this step was found to be 121.6 kJ mol⁻¹. The optimized sulfinic acid product (PC_{Ox}) was found to be much more stable than sulfenic acid with energy difference of -179.4 kJ mol⁻¹. This is in agreement with experiment and explains the irreversibility of the overoxidation.^{16,17}

Indeed, comparing the energetics of the overoxidation mechanism with the sulfurane formation shows that sulfurane formation is extremely favorable over sulfinic acid.

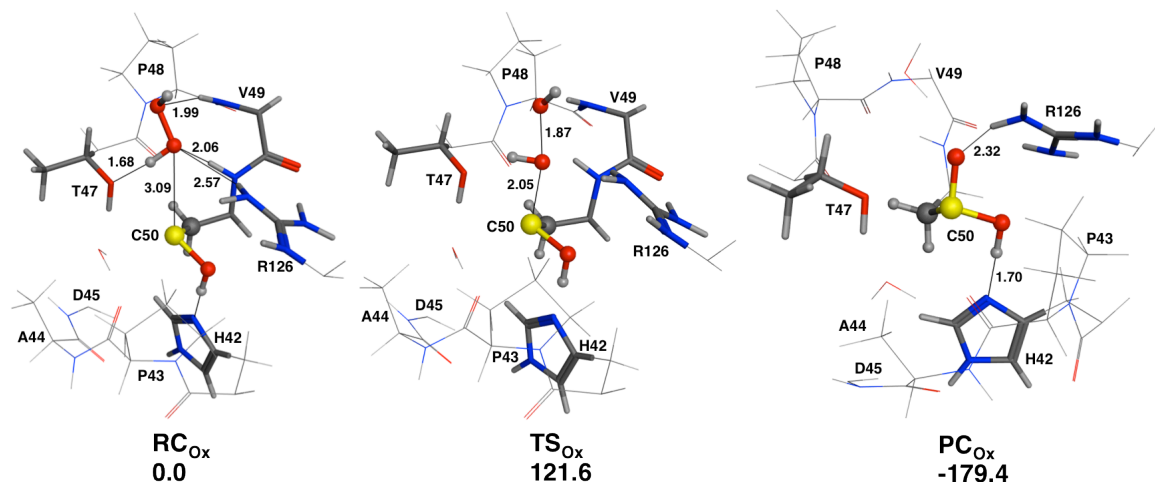


Figure 7.5. Illustration of the optimized stationary points for sulfenic acid overoxidation to sulfinic acid including their relative energies. Shown atoms represent the QM layer. However, only the highlighted are the ones involved in the reaction.

7.3.3 Sulfurane formation in human hORF6.

Finally, we investigated the possibility of pseudo-sulfurane formation in other enzymes in the Prx family. As discussed in the introduction, 1-Cys Prxs lack the presence of a second cysteine,¹ which suggests a sulfinic acid protective mechanism. In human 1-Cys hORF6, an X-ray structure shows that His39 nitrogen is 3.0 Å away from Cys47 sulfur indicating hydrogen bonding. Thus we elucidated the possibility of sulfurane formation in hORF6. In order to form this intermediate, as in ApTPx, a base is needed to deprotonate His39. From MD simulation, a water molecule was always found to connect His39 to Glu50 implying the possibility of sulfurane formation. Thus we optimized several intermediates along the reaction pathway.

In hORF6, the QM/MM optimized reactive complex (RC_{hORF6}) was found to show similar H_2O_2 binding interactions including $O_B \cdots HN_{Val46}$, $O_A \cdots HN_{Cys47}$, $O_A \cdots Arg132$ and

$\text{O}_A\text{H}\cdots\text{Thr44}$ with distances of 2.08 Å, 2.02 Å, 2.29 Å and 1.59 Å, respectively. The $\text{Cys47S}\cdots\text{O}_A$ is 3.42 Å as well as the anionic Cys47 sulfur is stabilized by four hydrogen bonds with His39, a water molecule and two hydrogen bonds with Arg132 with distances of 2.02 Å, 2.02 Å, 2.39 Å and 2.27 Å. Based on approximated pK_a 's, our model contained a neutral His39 and Cys47, however during the optimization an anionic Cys47 and protonated His39 was formed. The protonated His39 was found to be strongly hydrogen bonded to a water molecule, 1.52 Å. In turn this water is strongly hydrogen bonded to Glu50 with a distance of 1.48 Å.

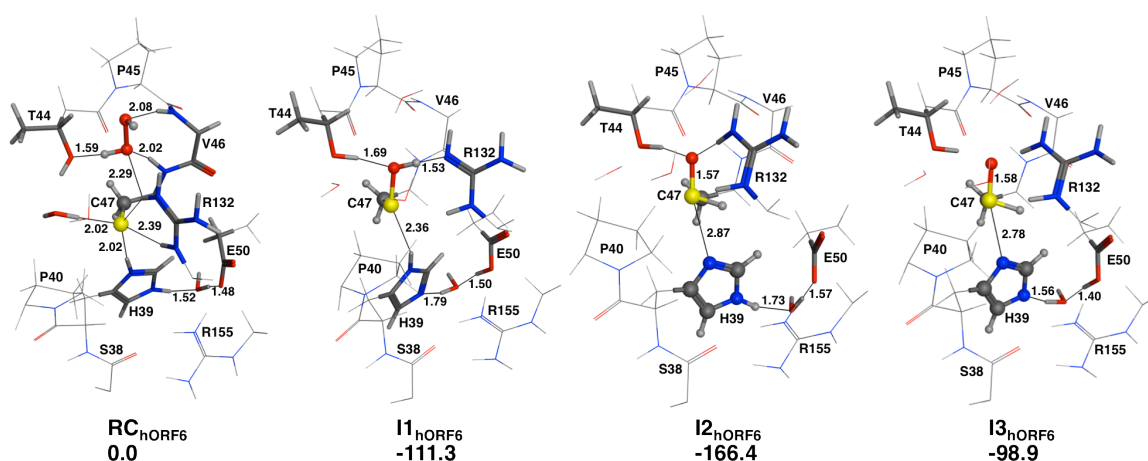


Figure 7.6. Illustration of the optimized stationary points for pseudo-sulfurane formation mechanism in hORF6, including their relative energies. Shown atoms represent the QM layer. However, the only highlighted are the ones involved in the reaction.

Subsequently, we optimized the sulfenic acid intermediate (I1_{hORF6}) which was found to be more stable than RC_{hORF6} by $-111.3 \text{ kJ mol}^{-1}$. Upon sulfenic acid formation His39 was found to donate its proton to Glu50 via the water molecule. Furthermore, as in ApTPx, sulfenic acid is formed with proton transfer from the neighboring arginine residue Arg132. The $\text{His42}\cdots\text{Cys50}$ distance has increased to 2.36 Å. In addition, His39 is

still strongly hydrogen bonded to a water molecule, 1.79 Å, which subsequently is hydrogen bonded to the now neutral Glu50, 1.50 Å.

As previous, in order for His39 to noncovalently bound to Cys47, His39 has to lose its proton. Based on X-ray and optimized QM/MM structure, this can be only achieved via sulfenic acid tautomer formation similar to ApTPx. Thus, we optimized the tautomer structure in which it was found to be lower in energy than RC_{hORF6} and I1_{hORF6} by -166.4 kJ mol⁻¹ and -55.1 kJ mol⁻¹, respectively. Notably, $_{\text{His39}}\text{N}\cdots\text{S}_{\text{Cys47}}$ is now reduced to 2.87 Å and the QTAIM calculation indicate the presence of noncovalent interaction with a $\rho=0.017$ and $\nabla\rho^2=0.015$. Furthermore, the optimized structure shows that His39 is now neutral upon proton transfer from Glu50. Similarly, sulfenic acid donates its proton to Arg132.

Finally, we also considered forming an anionic His39 to strengthen the S⁻⋯N interaction as in ApTPx. Indeed in I3_{hORF6}, the $_{\text{His39}}\text{N}\cdots\text{S}_{\text{Cys47}}$ is reduced to 2.78 Å and the electron density and the Laplacian at the critical point was increased to $\rho=0.022$ and $\nabla\rho^2=0.017$, respectively. Notably, our calculations show that a pseudo-sulfurane intermediate could form in 1-Cys Prx protecting sulfenic acid from overoxidation. The presence these noncovalent interaction between His nitrogen lone pair and Cys sulfur σ^* orbital prevent the $\angle\text{C}\alpha\text{-C}\beta\text{-S-O}$ dihedral angle rotation and subsequently H₂O₂ binding.

Previous results might suggest the generality of this protective mechanism in Prxs. Therefore, we also considered investigating the current X-ray structures in the protein data bank. As in hORF6, another 1-Cys Prx in *Arenicola Marina* shows that a histidynyl residue is 3.96 Å away from the Cp in the C45S mutated X-ray structure (PDB:2V32).⁴³ Surprisingly, several structures of human PrxV were found to include histadinyll residues in close proximity to the catalytic Cys with an approximate distance of 5 Å.⁴⁴⁻⁴⁶ Although PrxV is an atypical 2-Cys Prx, the distance between the two cysteinyl residues are over 13 Å and require large conformational change before disulfide bond formation which might imply the need for sulfenic acid protective mechanism.⁴⁷

In human PrxIV, typical 2-Cys Prx similar to AtPTx, a threonyl residue was found to be located beneath the catalytic cysteine (PDB: 3TKS) similar to previous discussed histidinyl with a distance of approximately 3 Å.⁴⁸ This suggests the occurrence of noncovalent interaction between Thr···Cys to form a novel sulfurane intermediate. Indeed, similar to nitrogen atom in His, the oxygen lone pair of Thr could form a noncovalent interaction with cysteinyl sulfur atom. As indicated in the introduction similar O···S interactions have been characterized in several proteins.²⁷

Finally, we also investigated the possibility of S···X noncovalent interaction formation in other enzymes. By investigating several X-ray structures including sulfenic acid, we found similar electrostatic interactions to occur. For instance, SUMO protease crystal structure indicate the presence of His residue 3.61 Å from sulfenic acid sulfur (PDB:2HKP).⁴⁹ Earlier, it was suggested that SUMO proteases is regulated by H₂O₂.⁴⁹ Similar interaction was also found in *E coli* Gsp amidase with His···Cys distance of 3.7 Å.⁵⁰ These structures as well as our results might suggest the generality of noncovalent interactions to stabilize sulfenic acid in proteins and their protection from overoxidation.

7.4 Conclusion

In this study the mechanistic details of pseudo-hypervalent sulfurane species in ApTPx and hORF6 have been investigated using molecular dynamics simulations (MD) and quantum mechanical/ molecular mechanical (ONIOM) calculations. The chemical nature of the obtained intermediates was also revealed using quantum theory atoms in molecules (QTAIM) and natural bond orbital analysis (NBO). Furthermore, the atomistic details of hydrogen-peroxide reduction mechanism in ApTPx have been elucidated.

First, MD and ONIOM calculations have shown that the binding of the substrate in the active site is similar to previous X-ray structures. Structural and NBO analysis of the reactive complex suggest the polarization of the O_A···O_B bond allowing for catalysis. The catalytic C_P was also found to exist as anionic sulfur upon substrate binding and stabilized

with four hydrogen bonds. The substrate was found to be reduced forming sulfenic/sulfenate and water molecule via a barrier of 82.1 kJ mol⁻¹. This step was found to be the rate-limiting step of the mechanism. Arg126 was found to act as an acid in case of sulfenic acid formation. However, the sulfenate intermediate was found to be more stable in the active site in which the oxyanion is stabilized by three strong hydrogen bonds.

The experimentally proposed hypervalent species was found to occur via a low barrier multistep reaction in which Asp45 was found to play a crucial role in alternating the ionization state of His42. The first step in this reaction was found to occur via the formation of sulfenic acid tautomer from sulfenate upon proton transfer from His42. Indeed, the neutral His42 nitrogen is now 2.76 Å from Cys50 sulfur. QTAIM analysis confirmed the presence of weak interaction between the two residues. This interaction was found to be strengthening upon the formation of anionic His42 upon proton transfer to Asp45. The S⁻N distance in the formed intermediate is in agreement with the X-ray structure. Another step was also considered in which the sulfenic acid tautomer oxygen abstract a proton from Arg126. However, this was found to be energetically unfeasible.

Second, we also considered investigating the overoxidation mechanism of sulfenic to sulfinic acid. This was found to occur upon another peroxide binding in the active site upon sulfenic acid dihedral rotation as suggested by previous experiment. The mechanism was found to occur via a high barrier, still enzymatically feasible, of 121.6 kJ mol⁻¹. Indeed the formation of pseudo hypervalent sulfurane is much more favorable over sulfinic acid formation emphasizing the role of sulfurane formation in protecting ApTPx sulfenic acid.

Third, we also considered investigating the possibility of similar hypervalent species in human 1-Cys hORF6. MD and ONIOM results suggested that Glu50 could act as an acid/base catalyst as in ApTPx. In order to investigate the mechanism, we optimized several intermediate along the reaction pathway. Notably, a similar reaction was found to

occur as in ApTPx. Although the S[⋯]N distance in hORF6 pseudo hypervalent species is larger than ApTPx, QTAIM suggest there is a presence of a weak interaction between the two residues.

Finally, we also considered investigating the protein data bank X-ray structures for similar possible interactions. Similarly to hORF6, the 1-Cys Prx in *Arenicola Marina* have a histidinyl residue. Remarkably, several structures of human PrxV were found to include histidinyl residues in close proximity to the catalytic Cys with an approximate distance of 5 Å. Furthermore, in the typical 2-Cys human PrxIV, a threonyl residue was found to 3 Å away from C_P suggesting the possibility of a novel sulfurane formation. We also considered investigating sulfenic acid X-ray structures in other proteins. In SUMO protease a His residue was found to be located 3.61 Å from sulfenic acid sulfur. Similarly, in *E coli* Gsp amidase a His residue was found to be in close proximity.

Our MD, ONIOM, QTAIM, NBO and X-ray structure analysis suggest the generality of sulfenic acid sulfur noncovalent interactions forming pseudo hypervalent as protective mechanism against overoxidation in Prxs and other proteins. It also emphasizes the role of weak interaction in restricting dihedral rotation and subsequently in catalysis. Several studies can be conducted in order to confirm the generality of these mechanisms in protein sulfenic acid chemistry.

7.5 References

- (1) Hall, A.; Nelson, K.; Poole, L. B.; Karplus, P. A. *Antioxid. Redox Signal.* **2011**, *15*, 795.
- (2) Zhu, H.; Santo, A.; Li, Y. B. *Exp. Biol. Med.* **2012**, *237*, 143.
- (3) Cox, A. G.; Winterbourn, C. C.; Hampton, M. B. *Biochem. J.* **2010**, *425*, 313.
- (4) Flohe, L.; Toppo, S.; Cozza, G.; Ursini, F. *Antioxid. Redox Signal.* **2011**, *15*, 763.

- (5) Winterbourn, C. C. *Nat. Chem Biol.* **2008**, *4*, 278.
- (6) Wood, Z. A.; Schroder, E.; Harris, J. R.; Poole, L. B. *Trends Biochem. Sci.* **2003**, *28*, 32.
- (7) Cao, Z. B.; Tavender, T. J.; Roszak, A. W.; Cogdell, R. J.; Bulleid, N. J. *J. Biol. Chem.* **2011**, *286*, 42257.
- (8) Karplus, P. A.; Hall, A. *Sub-cellular biochemistry* **2007**, *44*, 41.
- (9) Hall, A.; Parsonage, D.; Poole, L. B.; Karplus, P. A. *J. Mol. Biol.* **2010**, *402*, 194.
- (10) Deponete, M.; Becker, K. *Mol. Biochem. Parasitol.* **2005**, *140*, 87.
- (11) Knoop, B.; Goemaere, J.; Van der Eecken, V.; Declercq, J. P. *Antioxid. Redox Signal.* **2011**, *15*, 817.
- (12) Cejudo, F. J.; Ferrandez, J.; Cano, B.; Puerto-Galan, L.; Guinea, M. *FEBS Lett.* **2012**, *586*, 2974.
- (13) Nakamura, T.; Kado, Y.; Yamaguchi, T.; Matsumura, H.; Ishikawa, K.; Inoue, T. *J. Biochem.* **2010**, *147*, 109.
- (14) Choi, J. K.; Choi, S.; Choi, J. W.; Cha, M. K.; Kim, I. H.; Shin, W. *J. Biol. Chem.* **2003**, *278*, 49478.
- (15) Liao, S. J.; Yang, C. Y.; Chin, K. H.; Wang, A. H. J.; Chou, S. H. *J. Mol. Biol.* **2009**, *390*, 951.
- (16) Roos, G.; Messens, J. *Free Radical Biol. Med.* **2011**, *51*, 314.
- (17) Kettenhofen, N. J.; Wood, M. J. *Chem. Res. Toxicol.* **2010**, *23*, 1633.
- (18) Jeong, J. H.; Jung, Y. S.; Na, S. J.; Jeong, J. H.; Lee, E. S.; Kim, M. S.; Choi, S.; Shin, D. H.; Paek, E.; Lee, H. Y.; Lee, K. J. *Mol. Cell. Proteomics* **2011**, *10*.
- (19) Lim, J. C.; Choi, H. I.; Park, Y. S.; Nam, H. W.; Woo, H. A.; Kwon, K. S.; Kim, Y. S.; Rhee, S. G.; Kim, K.; Chae, H. Z. *J. Biol. Chem.* **2008**, *283*, 28873.

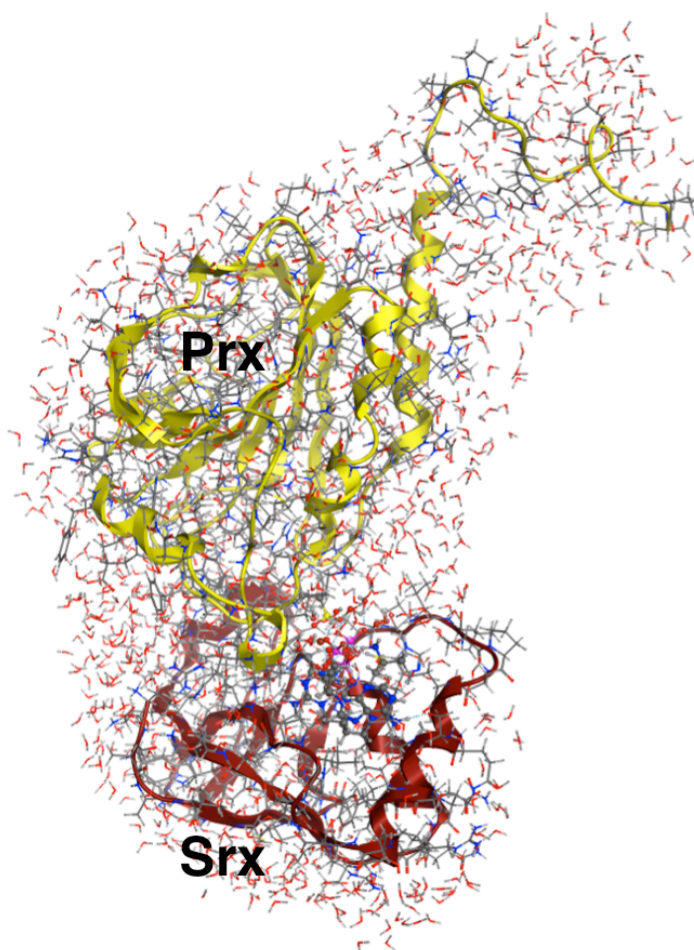
- (20) Jeong, W.; Bae, S. H.; Toledano, M. B.; Rhee, S. G. *Free Radical Biol. Med.* **2012**, *53*, 447.
- (21) Lowther, W. T.; Haynes, A. C. *Antioxid. Redox Signal.* **2011**, *15*, 99.
- (22) Biteau, B.; Labarre, J.; Toledano, M. B. *Nature* **2003**, *425*, 980.
- (23) Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6238.
- (24) Choi, H. J.; Kang, S. W.; Yang, C. H.; Rhee, S. G.; Ryu, S. E. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **1998**, *54*, 436.
- (25) Dokainish, H. M.; Gauld, J. W. *Biochemistry* **2013**, *52*, 1814.
- (26) Joensson, T. J.; Murray, M. S.; Johnson, L. C.; Lowther, W. T. *J. Biol. Chem.* **2008**, *283*, 23846.
- (27) Iwaoka, M.; Isozumi, N. *Molecules* **2012**, *17*, 7266.
- (28) Adhikari, U.; Scheiner, S. *Chem. Phys. Lett.* **2011**, *514*, 36.
- (29) Scheiner, S. *J. Chem. Phys.* **2011**, *134*.
- (30) MOE, version 2012.10 (2012) Chemical Computing Group Inc., Montreal.
- (31) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (32) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525.
- (33) Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 2284.
- (34) Gerber, P. R.; Muller, K. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251.
- (35) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198.
- (36) Polyak, I.; Reetz, M. T.; Thiel, W. *J. Phys. Chem. B* **2013**, *117*, 4993.

- (37) Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. *J. Am. Chem. Soc.* **2012**, *134*, 4743.
- (38) Gaussian 09, Revision D.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
- (39) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815.
- (40) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (41) Nakanishi, W.; Hayashi, S.; Narahara, K. *J. Phys. Chem. A* **2008**, *112*, 13593.
- (42) Sarma, G. N.; Nickel, C.; Rahlfs, S.; Fischer, M.; Becker, K.; Karplus, P. *A. J. Mol. Biol.* **2005**, *346*, 1021.
- (43) Smeets, A.; Loumaye, E.; Clippe, A.; Rees, J. F.; Knoops, B.; Declercq, J. P. *Protein Sci.* **2008**, *17*, 700.
- (44) Declercq, J. P.; Evrard, C.; Clippe, A.; Vander Stricht, D.; Bernard, A.; Knoops, B. *J. Mol. Biol.* **2001**, *311*, 751.

- (45) Evrard, C.; Capron, A.; Marchand, C.; Clippe, A.; Wattiez, R.; Soumillion, P.; Knoops, B.; Declercq, J. P. *J. Mol. Biol.* **2004**, *337*, 1079.
- (46) Evrard, C.; Smeets, A.; Knoops, B.; Declercq, J. P. *J. Chem. Crystallogr.* **2004**, *34*, 553.
- (47) Smeets, A.; Marchand, C.; Linard, D.; Knoops, B.; Declercq, J. P. *Arch. Biochem. Biophys.* **2008**, *477*, 98.
- (48) Wang, X.; Wang, L. K.; Wang, X.; Sun, F.; Wang, C. C. *Biochem. J.* **2012**, *441*, 113.
- (49) Xu, Z.; Lam, L. S. M.; Lam, L. H.; Chau, S. F.; Ng, T. B.; Au, S. W. N. *FASEB J.* **2008**, *22*, 127.
- (50) Chiang, B. Y.; Chen, T. C.; Pai, C. H.; Chou, C. C.; Chen, H. H.; Ko, T. P.; Hsu, W. H.; Chang, C. Y.; Wu, W. F.; Wang, A. H. J.; Lin, C. H. *J. Biol. Chem.* **2010**, *285*, 25345.

Chapter 8

QM/MM Investigation of the Reduction Mechanism of Cysteine Sulfinic Acid in Peroxiyredoxin via Sulfiredoxin



8.1 Introduction

Elevated levels of reactive oxygen species (ROS), such as hydrogen peroxide (H_2O_2), have been known for their abilities of modifying several cellular components including proteins, DNA and lipids leading to cell damage.^{1,2} Such oxidative stress in cells has been highly related to aging and age related diseases including cancer and Alzheimer's.²⁻⁵ Dissimilarly, low levels of H_2O_2 , under subtoxic concentration, are essential for a variety of cellular functions including proliferation, differentiation, adhesion and migration by stimulating signal transduction.⁶⁻⁸ Indeed, H_2O_2 fulfills the requirement to act as a second messenger as it is rapidly produced and controlled via antioxidant enzymes.⁹ However, for H_2O_2 to act as a signaling molecule its concentration has to reach and to maintain a certain threshold level.¹⁰

In cells, H_2O_2 is reduced via two main mechanisms;⁹ 1) It is being reduced by one electron reduction mechanism using transition metal such as in catalases or,¹¹ 2) it is reduced by the two electron nucleophilic substitution reaction by protein cysteinyl or seleno-cysteinyl residues, forming sulfenic/selenenic acid and water molecule such as in thiol and glutathione-peroxidases.¹²⁻¹⁴

Peroxiredoxin (Prxs), also known as thiol peroxidases, are ubiquitous thiol-dependent enzymes that reduce H_2O_2 as well as other ROS species.^{12,15} It is highly abundant within cells and it represents the main enzyme for reducing H_2O_2 with a catalytic rate constant of $10^7 \text{ M}^{-1} \text{ s}^{-1}$.¹⁵ Previous studies have shown that Prxs can readily reduce all cytoplasmic and nearly 90% of the mitochondrial H_2O_2 emphasizing its role in regulating low levels of H_2O_2 and its subsequent signaling.¹⁰

In general, Prxs are categorized into three main classes, 1-Cys, typical 2-Cys and atypical 2-Cys, based on the number and location of the catalytic cysteinyl residues as well as the formation of inter or intra disulfide bond in the mechanism.¹⁶ For example, in 1-Cys there is only one catalytic cysteinyl residue. In both 2-Cys classes two catalytic cysteinyl residues are involved in the mechanism.¹² The first step of the H_2O_2 reduction

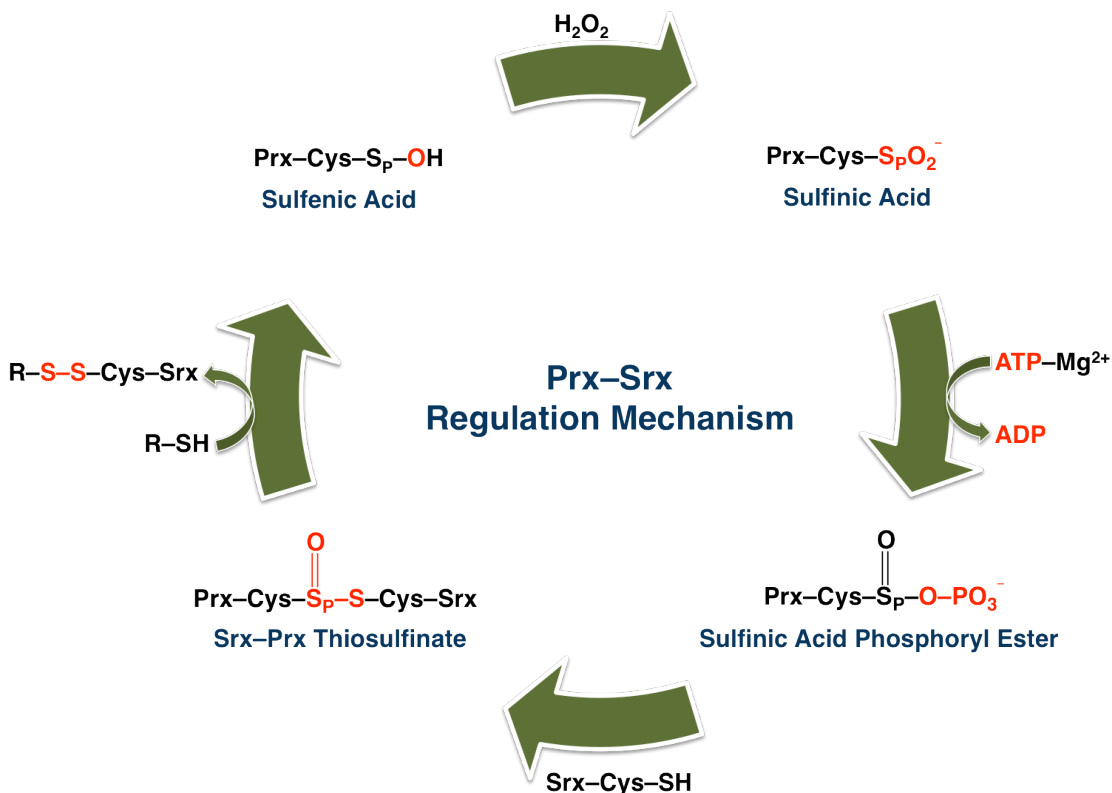
via Prxs is similar in which a conserved catalytic cysteinyl residue, known as peroxidatic cysteine, (Cys-S_PH) attacks the H₂O₂ molecule forming a highly reactive sulfenic acid intermediate (Cys-S_POH).¹² Subsequently, the later molecule is reduced using a resolving cysteinyl residue (Cys-S_RH) forming a disulfide bond. In the atypical 2-Cys, the resolving Cys is located on the same monomer.¹⁷ Dissimilarly, in the typical 2-Cys, Cys-S_RH is located on the adjacent monomer of the Prxs dimer. Ultimately, an external thiol reduces the sulfenic acid intermediate in 1-Cys or the disulfide bond in 2-Cys restoring the catalytic activity of the enzyme as the formation of sulfenic acid or disulfide transiently stops the peroxidatic activity.¹²

In eukaryotic cells, the reduction of sulfenic acid, in the typical 2-Cys, by the adjacent monomer Cys-S_RH requires large structural rearrangements decreasing the rate of disulfide formation.^{18,19} Notably, this can lead to the overoxidation of the highly reactive sulfenic acid by a second H₂O₂ molecule forming a sulfinic acid (–SO₂H) intermediate.¹⁹ Subsequently, under oxidative stress sulfinic acid is overoxidized to sulfonic acid (–SO₃H).²⁰ The formation of sulfinic acid in Prxs was found to not only inactivate the peroxidatic activity allowing for the increase in the peroxide level and signal stimulation but also it leads to the formation of high molecular mass Prxs aggregates with a chaperone activity.^{21,22} In addition to typical 2-Cys, the oxidation of sulfenic acid has been found to occur in 1-Cys and in the atypical 2-Cys but at slower rates.²³ It is important to mention that this overoxidation mechanism has been found to occur in other proteins such as protein tyrosine phosphatases 1B (PTP1B) and carbonic anhydrases.^{24,25} In general, these overoxidation states of sulfenic acid are considered irreversible, leading to the inactivation of the enzyme.^{26,27} Furthermore, the formation of sulfinic and sulfonic acid has been found to be associated with multiple disorders including cancer and age related diseases.^{22,28-30} Fortunately, the overoxidation of the typical 2-Cys Prxs to sulfinic acid is reversible by sulfiredoxin (Srx) via a unique mechanism.²²

Srx was first identified in 2003 in *Saccharomyces cerevisiae* in which a disulfide bond was formed between Srx and 2-Cys Prx.³¹ The reduction mechanism was found to be dependent on the presence of ATP and magnesium.³¹ Later several studies have indicated the conservation of Srx in eukaryotic organisms.²² In human, Srx is mainly localized in the cytosol allowing for the restoration of the peroxidatic activity of PrxI and PrxII as well as the mitochondrial PrxIII and the ER PrxIV.²² Therefore, Srx can bind to all human typical 2-Cys Prxs and reverse the oxidation to sulfinic acid. However, previous studies have also shown that the Srx-Prx reduction mechanism is inefficient, with catalytic rate constant of 0.1 to 1.8 min⁻¹, emphasizing the role of Srx in regulating 2-Cys Prxs and subsequently H₂O₂ signaling.^{22,32} Wood et al.³¹ have proposed a floodgate hypothesis in which transient intracellular peroxide burst increases the level of peroxide leading to the rapid inactivation of 2-Cys Prx allowing for H₂O₂ to act as messenger interacting with other proteins. However, in the presence of high concentration of Prx peroxide signaling would be stopped. Hence this overoxidation/reduction mechanism is essential for peroxide signaling as well as Prxs chaperon activity.^{21,23}

The molecular basis of this unique interaction of Srx-Prx has been investigated using several techniques including structural, mutational and mass spectrometry studies.²² However, the exact mechanism is debatable. Originally, Biteau et al.³³ proposed that the mechanism would occur in two steps in which Srx act as a phosphotransferase and thioltransferase. First, the sulfinic acid is phosphorylated forming a novel sulfinic acid phosphoryl ester intermediate, see scheme 8.1. This reaction is reminiscent and is known to occur in several enzymes for the activation of a carboxylic group by phosphorylation such as in glutamine synthetases.³⁴ However, this reaction is unique for sulfur chemistry. Then, the activated phosphoryl ester intermediate might reacts with a conserved Srx cysteinyl residue (Cys99) forming Srx-Prx thiosulfinate intermediate,³³ see scheme 8.1. Finally, an external thiol, such as Trx, reduces the thiosulfinate intermediate restoring the Prx activity.

Scheme 8.1. The proposed regulation mechanism of typical 2-Cys Prx via Srx upon the hyperoxidation of Prx sulfenic acid by hydrogen peroxide.



On the contrary, Jeong et al.¹⁹ proposed different mechanism in which the Srx conserved Cys is first phosphorylated similar to the phosphorylation mechanism of PTP1B active site cysteinyl residue.³⁵ Followed by the transfer of the phosphate group to the sulfenic acid oxygen forming the sulfenic acid phosphoryl ester. Later the intermediate is directly reduced via thiol molecule (RSH). Furthermore, their mutation studies indicated that Arg50, Asp57 and Asp79 are essential for substrate binding.¹⁹ It is important to mention that the mutation of Cys99 was found to terminate the reductase activity of Srx.³³

Jönsson et al.^{32,36,37} have performed several crystallographic studies to investigate the nature of the formed intermediates in the mechanism. Their results support the original mechanism, see scheme 8.1, in which Srx performs bifunctional roles. Since their

obtained crystal structure of Srx-Prx complex as well as the quaternary structure of Srx-Prx in the presence of ATP and Mg^{2+} shows that the distance between the γ -phosphate and Cys99 sulfur is nearly 5 Å.³⁶ Notably, the position of Cys99 does not allow for in line nucleophilic attack. On the contrary, the Prx Cys52Asp oxygen is in a perfect position for in line nucleophilic attack with a distance of 4.3 Å from the phosphate group.³⁶ In addition, in their Srx, ATP and Mg^{2+} crystal structure the γ -phosphate was found to point toward the solvent.³⁷ Furthermore, the formation of the thiosulfinate intermediate has been proven in both human and yeast in previous studies.^{38,39} The role of the Mg ion was also investigated in which it was found to ligate to the three phosphate groups of ATP resulting in directing the γ -phosphate toward the Prx Cys-S_p.²² Finally crystal structure indicated the unfolding of the Prx to allow for its interaction with Srx.³²

Until now, this unique sulfur reaction has not been investigated neither using small molecules nor within the Srx-Prx active site. Therefore in this study the reduction mechanism of Prx sulfinic acid via Srx is investigated in details using quantum mechanical/ molecular mechanical (QM/MM) modeling approach based on the Srx-Prx ATP and Mg^{2+} bound protein crystal structure. Computational chemistry is now established as a main tool for investigating catalytic mechanisms and has been successfully used to gain atomistic details of the enzymatic mechanisms including intermediates, transition states and energetics.⁴⁰⁻⁴³

8.2 Computational Methods

8.2.1 Protein Model Preparation

The X-ray crystal structure preparation and the MD analysis were accomplished using the Molecular Operating Environment (MOE) software package.⁴⁴ Molecular dynamic simulation was performed using the NAMD Molecular Dynamics software.⁴⁵ The initial coordinates were taken from the X-ray crystal structure of human Srx in complex with typical 2-Cys Prx (PrxI), ATP and Mg^{2+} , PDB: 3HY2.³⁶

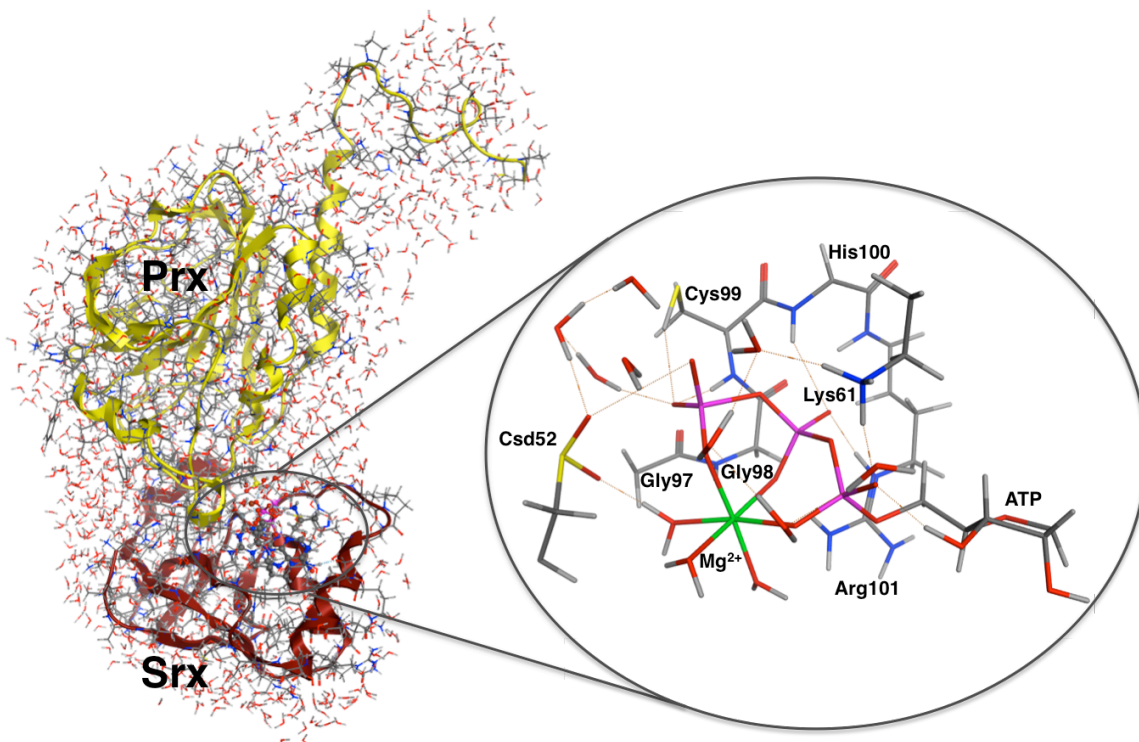


Figure 8.1. The QM/MM model for the Srx-Prx Michaelis complex used to investigate the first step of the mechanism. The QM layer atoms are highlighted in sticks to the right of the Figure.

First, the PrxI Cys52Asp was mutated to sulfinic acid. Second, the missing protons in the crystal structure were added using the protonate 3D application in MOE that assigns each residue ionization state by minimizing the total free energy of the system.⁴⁶ Then, the two enzymes complex structure was spherically solvated up to 15 Å beyond every protein atom. The distance between the γ -phosphate and one of the sulfinic oxygen's was restrained to 3–3.5 Å. Subsequently, the protein complex was minimized using AMBER12:EHT force field in which AMBER 12 parameters are used for protein and Extended Hückel Theory for parameterizing small molecules.^{47,48} Finally, in order to allow for thermal relaxation especially after introducing sulfinic acid in the structure, the minimized structure was used as starting point for a short 500 ps MD simulation at 300 K

similar to previous studies.⁴⁹ Finally, the final structure of this short simulation was chosen as the starting structure for further calculations.

In order to investigate the second step of the reaction, a second 500 ps MD simulation was also performed to model possible arrangements in the active site upon sulfinic acid phosphoryl ester intermediate formation. The starting structure for this simulation was obtained from the QM/MM optimized structure of the intermediate. The distance between sulfinic acid phosphoryl ester sulfur and the Srx Cys99 sulfur was restrained to 2.5–3 Å in order to allow for subsequent reaction. The same protocol was used as in the first simulation and the last structure of the simulation was chosen as the starting point for QM/MM calculations.

8.2.2 QM/MM Models and Calculations

All calculations were performed within the ONIOM scheme using the Gaussian 09 suite of programs.⁵⁰ The QM/MM starting structures were taken from previous steps using the whole protein Srx-Prx complex. The QM high layer was described using the hybrid-meta-exchange-correlation functional M06-2X.⁵¹ While the MM low layer was described by AMBER96 force field as implemented in Gaussian. Optimized geometries were obtained using the 6-31G(d,p) basis set for the high layer. Due to computational cost as well as to keep the integrity of the protein-protein interaction all atoms 15 Å away from the γ and β phosphate linking oxygen were kept frozen in the calculations. Relative energies were then obtained using a single point energy calculations on the optimized structures at the ONIOM(M06-2X/6-311+G(2df,p):AMBER96) level of theory within the electrostatic embedding (EE) formalism.⁵² The EE formalism accounts for the polarization of the core layer by the low layer at the QM level of theory. The choice of functional and basis sets was based on our previous benchmarking study of several biological sulfur species in which M06-2X was found to be one of the best functionals with respect to a benchmarking QCISD and MP2 calculations (manuscript in

preparation). Frequency calculations were performed on the high layer only to confirm the nature of the optimized structures as previously conducted.^{53,54}

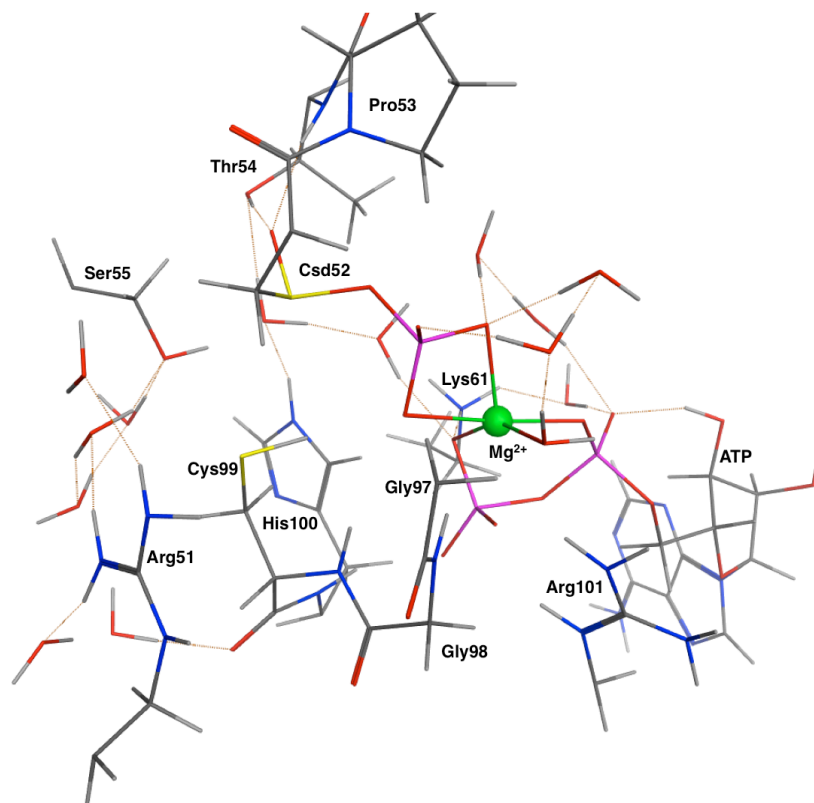


Figure 8.2. The QM layer residues included in model II to study the second step of the reduction mechanism converting sulfinic acid phosphoryl ester into thiosulfinate.

As can be seen in Figure 8.1, the Michaelis complex structure included only the cysteinyl sulfinic acid residue (Csd52) from PrxI. Besides several residues from the Srx protein as the ATP molecule is mainly bonded to the Srx active site, similar to experimental results, including the conserved Cys99, His100, Lys61, Arg101, Gly97 and Gly98. Furthermore, in addition to ATP and Mg^{2+} , several active site water molecules were also added to the high layer.

For the second step of the reaction, the QM layer of the sulfinic acid phosphoryl intermediate was slightly increased allowing for better description of the sulfinic acid

phosphoryl ester surroundings interaction. As can be seen in Figure 8.2, the used model included Csd52, Pro53 and Thr54 from PrxI. In addition, the same Srx residues as in model I were included beside Ser55 and Arg51 as well as the R group of His100.

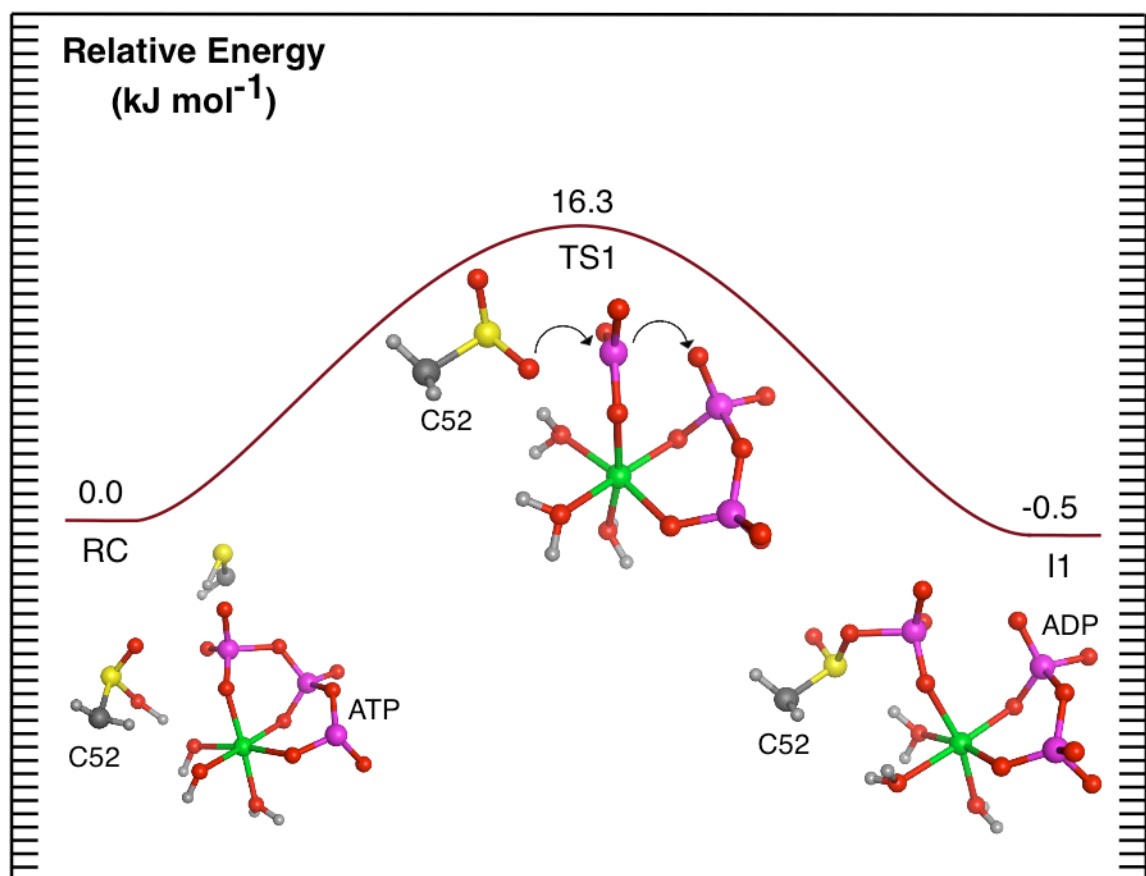


Figure 8.3. Potential energy surface obtained for the formation of sulfinic acid phosphoryl intermediate and ADP from Prx-sulfinic acid and ATP via a concerted mechanism.

8.3 Results and Discussion

8.3.1 Formation of sulfinic acid phosphoryl ester

The total active site interactions of the optimized structure of the reactive complex (RC) are similar to the Srx-Prx X-ray structure. The phosphate groups of the ATP molecule is stabilized and bound to the Srx proteins by several hydrogen bonding's with

Srx residues including Lys62, Arg101, Cys99 and the backbone amide of His100, Cys99 and Gly98. Previous mutational studies have confirmed the role of these residues in ATP binding. In addition, as in the crystal structure the Mg^{2+} ion is coordinated to the three phosphate of the ATP molecule. Besides Mg^{2+} ion coordination to the phosphate groups, it is also coordinated to three water molecules with an average distance of 2.06 Å. this octahedral interactions of the Mg^{2+} ion direct the γ -phosphate toward the Prx sulfinic acid (Csd52).

The Srx-Cys99 sulfur is 3.85 Å from the γ -phosphate atom. More importantly, the distance between the Prx-Csd52 oxygen and the γ -phosphate atom is 3.08 Å, see Figure 8.4. Therefore, the position of Prx-Csd52 is in a better position for direct in line nucleophilic attack as previously proposed. Furthermore, the Prx-Csd52 second oxygen is sharing a proton with one of the Mg ion bound water molecule in which the proton is 1.39 Å and 1.09 Å from the water oxygen and the sulfinic acid oxygen, respectively. In addition, the Mulliken population analysis shows that the sulfinic acid oxygen has a negative charge of -0.54 and the γ -phosphate atom has a positive charge of 1.28.

Based on structural arrangement and charge distribution, we investigated the first step of the mechanism in which a direct in line nucleophilic attack of Prx-Csd52 oxygen to form sulfinic acid phosphoryl ester. This step was found to occur via a low barrier of 16.3 kJ mol⁻¹. In TS1, the Prx-csd52 oxygen is 2.13 Å form the γ -phosphate atom. Similarly, the γ -phosphate...O bond is elongated to 2.14 Å showing that phosphoryl transfer is happened via a concerted mechanism. Other interactions in the active site remain similar to the reactive complex.

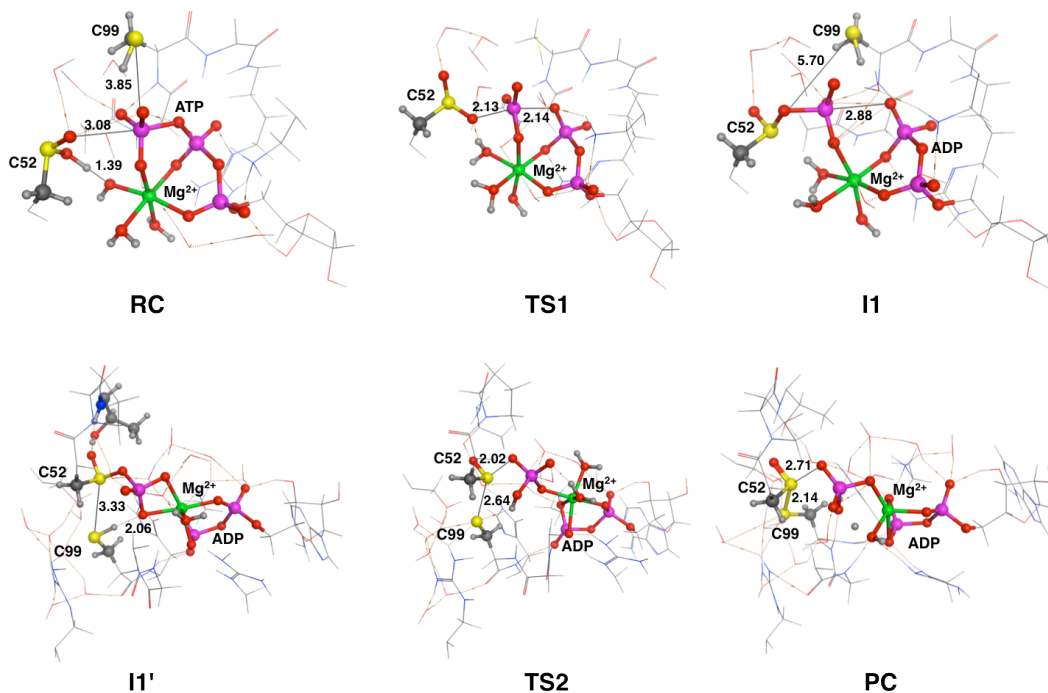


Figure 8.4. Schematic representation of the optimized stationary points for the overall reaction, from sulfinic acid to thiosulfinate product formation. All atoms in the QM layer are included however, we only highlighted the atoms included in the reaction as balls and sticks.

The energy of the sulfinic acid phosphoryl intermediate (I1) is quite similar to RC with a difference of -0.5 kJ mol^{-1} . Indeed in the I1 optimized structure, see Figure 8.3, the phosphoryl group is already transferred to the sulfinic acid in which the Prx-Csd52 oxygen...phosphorus distance is 1.7 \AA . Furthermore, the phosphoryl group is totally dissociated from the ATP forming an ADP molecule with a $\text{P}\cdots\text{O}$ distance of 2.88 \AA . Furthermore, the dissociated ADP oxygen is now hydrogen bonded to Srx-Cys99 and the backbone amide of His100. The Mg^{2+} ion is still coordinated to the three phosphate groups. In I1, the distance between the Srx-Cys99 sulfur and the sulfinic acid phosphoryl intermediate sulfur is quite long for subsequent reaction, 5.70 \AA . This indicates the need for structural rearrangements of the active site upon formation of I1. Therefore, as

indicated in the computational method section, we performed a second MD on the I1 optimized structure allowing for the following interaction to happen. It is important to mention that previous study indicated the change of the two sulfur's distance to approximately 3 Å upon sulfinic acid phosphoryl ester formation.

8.3.2 Formation of thiosulfinate intermediate

The obtained MD structure of I1' was optimized as in the previous step however, a larger QM layer was included to better describe new interactions in the intermediate, see computational section. Several new interactions were obtained in the QM/MM optimized I1' upon allowing for structural rearrangements such as Mg^{2+} ion is now coordinated to two oxygen of the sulfinic acid phosphoryl ester as well as three oxygen from the ADP molecule in addition to a water molecule keeping its octahedral structure. The second oxygen of the Csd52 is double hydrogen bonded to Prx-Thr54. In addition several water molecules are hydrogen bonded to the phosphoryl moiety of the sulfinic acid ester. Furthermore, the Srx-Cys99 is strongly hydrogen bonded to the sulfinic acid phosphoryl ester, 2.06 Å, see Figure 8.4. More importantly, the distance between the Srx-Cys99 sulfur and the sulfinic acid sulfur is now 3.33 Å allowing for subsequent reaction. Furthermore, the Mulliken population analysis shows that the sulfinic acid phosphoryl ester has a positive sulfur center of 0.152. In contrary the Srx-Cys99 sulfur has a negative charge of -0.312.

The second step of the reaction involving thiosulfinate formation was also found to occur in a concerted mechanism with an energy barrier of 81.6 kJ mol⁻¹ with respect I1', see Figure 8.5. The optimized transition state for this reaction (TS2) shows that the Srx-Cys99 is already activated forming a better nucleophile by the phosphoryl sulfinic acid ester oxygen. Notably the S···S distance is reduced to 2.64 Å with a concomitant elongation of the S···OPO₃H bond to 2.02 Å. This indicates that the reaction would occur via an S_N2 mechanism as in the first step. Furthermore, in TS2, the Mg^{2+} ion losses one of

its coordination with the now protonated oxygen of the phosphoryl moiety, replacing it with a second water molecule. The now nucleophilic Srx-Cys99 sulfur is currently stabilized by hydrogen bonding with Srx-Arg51 as well as a water molecule.

The formation of the disulfide bond forming the thiosulfinate product complex and phosphoryl group (PC) is found to lay 66.2 kJ mol^{-1} higher in energy than I1'. Indeed, the PC optimized structure shows the cleavage of the $\text{S}\cdots\text{OPO}_3\text{H}$ bond with a distance of 2.71 \AA . More importantly, the Srx-Cys99 sulfur is bounded to Prx-Cys52 with a distance of 2.14 \AA . Furthermore, the Mg^{2+} ion has restored its double coordination with the released phosphoryl group.

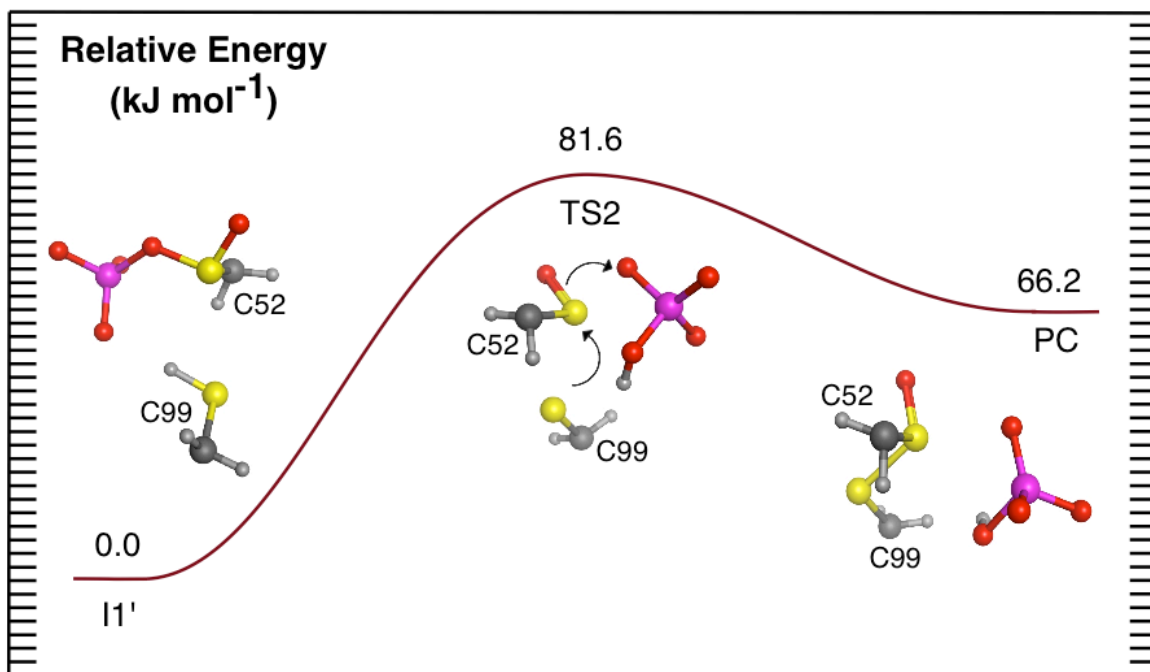


Figure 8.5. Potential energy surface obtained for the formation of thiosulfinate product complex via the interaction of Prx-sulfenic acid phosphoryl ester and Srx-Cys99 via a concerted mechanism.

In general, our results support the previously suggested reduction of Prx-sulfenic acid via ATP and Srx. Additionally, our obtained energies suggest that the rate-limiting step of

this part of the reduction mechanism is the formation of the thiosulfinate intermediate. Our obtained energies also suggest that the inefficiency and the slow catalytic rate for this reaction is mainly due to the final step in which an external thiol reduces the thiosulfinate intermediate. In fact this is in agreement with several enzymatic mechanisms including external thiol such as Trx. For example in methionine sulfoxide reductase Msr's where the reduction of the disulfide by Trx is the rate-limiting step.⁴³ Finally, our results provide a detailed mechanism for sulfinic acid reduction, which could open the door to design new antioxidants capable of reducing such irreversible modifications of proteins.

8.4 Conclusions

The atomic description of the reduction mechanism of Prx-sulfinic acid via ATP and Srx has been elucidated in details using the hybrid quantum mechanical/ molecular mechanical (QM/MM) approach, within the ONIOM formalism. The Srx-Prx protein-protein starting model was first obtained from the X-ray structure of the Prx-Cys52Asp mutated structure with Srx, ATP and Mg^{2+} . Prior to QM/MM calculation, first, the Prx-Asp52 was modified to sulfinic acid. Subsequently, the whole structure was solvated and minimized using AMBER12:EHT force field. Then the minimized structure was thermally relaxed using short molecular dynamics simulation for 500 ps allowing for equilibration of the active site upon modifications. Finally, the last MD structure was chosen as a starting structure for the QM/MM calculations.

The previously proposed mechanism for sulfinic acid reduction via direct formation of sulfinic phosphoryl ester and finally thiosulfinate was examined. The Srx-Prx complex was divided into two layers in which the high layer included only the active site residues and their calculations were performed using the M06-2X, DFT functional. The rest of the complex was optimized using AMBER96 force field.

The first step of the mechanism was found to occur via a direct concerted mechanism forming sulfinic phosphoryl ester with a small activation energy of 16.3 kJ mol^{-1} . In TS1,

the Prx-csd52 oxygen perform an in line nucleophilic attack on the γ -phosphate as previously suggested forming a slightly stable sulfinic acid phosphoryl intermediate with respect to RC. The I1 optimized structure shows that the phosphoryl group is already transferred to the sulfinic is dissociated from the ATP forming an ADP molecule. In I1, the distance between the Srx-Cys99 sulfur and the sulfinic acid phosphoryl intermediate sulfur indicated the need for structural rearrangements of the active site upon the intermediate formation allowing for subsequent reaction.

A second MD on the sulfinic acid phosphoryl ester was performed allowing for structural rearrangements. The last MD structure was chosen for subsequent QM/MM calculations as in the previous step. A larger QM layer was included in the starting structure of the second step allowing for better description of the active site. The QM/MM optimized of I1' shows several structural rearrangements in the active site, such as the Mg^{2+} ion double coordination the sulfinic acid phosphoryl moiety. Notably, the distance between the Srx-Cys99 sulfur and the sulfinic acid sulfur is 3.33 Å allowing for subsequent reaction.

The second step of the mechanism was also found to occur in a concerted mechanism with an activation energy of 81.6 kJ mol⁻¹. In TS2, the Srx-Cys99 is first activated forming a better nucleophile followed by nucleophilic attack on the sulfinic acid sulfur with a concomitant cleavage of the S-O bond via an S_N2 mechanism. The formed thiosulfinate product complex was found to lay 66.2 kJ mol⁻¹ higher in energy than I1'. Our results suggest that the rate-limiting step of this part of the reduction mechanism is the formation of the thiosulfinate intermediate indicating that the rate limiting-step of the overall reaction is related to the reduction of the thiosulfinate intermediate via an external thiol as in other proteins.

Finally, this study set up the basis for understanding the chemistry behind this fascinating sulfur reaction in Srx which could open the door to design Srx ATP mimic antioxidants as new hope to fight oxidative stress which is hallmark for many diseases.

8.5 References

- (1) Dickinson, B. C.; Chang, C. J. *Nat. Chem. Biol.* **2011**, *7*, 504.
- (2) Chen, X. P.; Guo, C. Y.; Kong, J. M. *Neural Regen. Res.* **2012**, *7*, 376.
- (3) Emerit, J.; Edeas, M.; Bricaire, F. *Biomed. Pharmacother.* **2004**, *58*, 39.
- (4) Finkel, T.; Serrano, M.; Blasco, M. A. *Nature* **2007**, *448*, 767.
- (5) Barnham, K. J.; Masters, C. L.; Bush, A. I. *Nat. Rev. Drug Discov.* **2004**, *3*, 205.
- (6) Sauer, H.; Wartenberg, M.; Hescheler, J. *Cell. Physiol. Biochem.* **2001**, *11*, 173.
- (7) Miki, H.; Funato, Y. *J. Biochem.* **2012**, *151*, 255.
- (8) Chiarugi, P.; Fiaschi, T. *Cell. Signal.* **2007**, *19*, 672.
- (9) Forman, H. J.; Maiorino, M.; Ursini, F. *Biochemistry* **2010**, *49*, 835.
- (10) Winterbourn, C. C. *Nat. Chem. Biol.* **2008**, *4*, 278.
- (11) Vidossich, P.; Alfonso-Prieto, M.; Rovira, C. *J. Inorg. Biochem.* **2012**, *117*, 292.
- (12) Flohe, L.; Toppo, S.; Cozza, G.; Ursini, F. *Antioxid. Redox Signal.* **2011**, *15*, 763.
- (13) Rahmanto, A. S.; Davies, M. J. *IUBMB Life* **2012**, *64*, 863.
- (14) Lubos, E.; Loscalzo, J.; Handy, D. E. *Antioxid. Redox Signal.* **2011**, *15*, 1957.
- (15) Hall, A.; Nelson, K.; Poole, L. B.; Karplus, P. A. *Antioxid. Redox Signal.* **2011**, *15*, 795.
- (16) Wood, Z. A.; Schröder, E.; Robin Harris, J.; Poole, L. B. *Trends Biochem. Sci.* **2003**, *28*, 32.
- (17) Knoop, B.; Goemaere, J.; Van der Eecken, V.; Declercq, J. P. *Antioxid. Redox Signal.* **2011**, *15*, 817.

- (18) Ogasawara, Y.; Ohminato, T.; Nakamura, Y.; Ishii, K. *Int. J. Biochem. Cell Biol.* **2012**, *44*, 1072.
- (19) Jeong, W.; Park, S. J.; Chang, T. S.; Lee, D. Y.; Rhee, S. G. *J. Biol. Chem.* **2006**, *281*, 14400.
- (20) Lim, J. C.; Choi, H. I.; Park, Y. S.; Nam, H. W.; Woo, H. A.; Kwon, K. S.; Kim, Y. S.; Rhee, S. G.; Kim, K.; Chae, H. Z. *J. Biol. Chem.* **2008**, *283*, 28873.
- (21) Neumann, C. A.; Cao, J. X.; Manevich, Y. *Cell Cycle* **2009**, *8*, 4072.
- (22) Lowther, W. T.; Haynes, A. C. *Antioxid. Redox Signal.* **2011**, *15*, 99.
- (23) Jeong, W.; Bae, S. H.; Toledano, M. B.; Rhee, S. G. *Free Radical Biol. Med.* **2012**, *53*, 447.
- (24) van Montfort, R. L. M.; Congreve, M.; Tisi, D.; Carr, R.; Jhoti, H. *Nature* **2003**, *423*, 773.
- (25) Mallis, R. J.; Hamann, M. J.; Zhao, W.; Zhang, T. Q.; Hendrich, S.; Thomas, J. A. *Biol. Chem.* **2002**, *383*, 649.
- (26) Roos, G.; Messens, J. *Free Radical Biol. Med.* **2011**, *51*, 314.
- (27) Brandes, N.; Schmitt, S.; Jakob, U. *Antioxid. Redox Signal.* **2009**, *11*, 997.
- (28) Yoshida, Y.; Yoshikawa, A.; Kinumi, T.; Ogawa, Y.; Saito, Y.; Ohara, K.; Yamamoto, H.; Imai, Y.; Niki, E. *Neurobiol. Aging* **2009**, *30*, 174.
- (29) Musicco, C.; Capelli, V.; Pesce, V.; Timperio, A. M.; Calvani, M.; Mosconi, L.; Zolla, L.; Cantatore, P.; Gadaleta, M. N. *Biochim. Biophys. Acta-Bioenerg.* **2009**, *1787*, 890.
- (30) Joshi, G.; Aluise, C. D.; Cole, M. P.; Sultana, R.; Pierce, W. M.; Vore, M.; St Clair, D. K.; Butterfield, D. A. *Neuroscience* **2010**, *166*, 796.
- (31) Wood, Z. A.; Poole, L. B.; Karplus, P. A. *Science* **2003**, *300*, 650.
- (32) Jonsson, T. J.; Johnson, L. C.; Lowther, W. T. *Nature* **2008**, *451*, 98.
- (33) Biteau, B.; Labarre, J.; Toledano, M. B. *Nature* **2003**, *425*, 980.

- (34) Eisenberg, D.; Gill, H. S.; Pfluegl, G. M. U.; Rotstein, S. H. *Biochim. Biophys. Acta-Protein Struct. Molec. Enzym.* **2000**, *1477*, 122.
- (35) Yip, S. C.; Saha, S.; Chernoff, J. *Trends Biochem. Sci.* **2010**, *35*, 442.
- (36) Jonsson, T. J.; Johnson, L. C.; Lowther, W. T. *J. Biol. Chem.* **2009**, *284*, 33305.
- (37) Joensson, T. J.; Murray, M. S.; Johnson, L. C.; Lowther, W. T. *J. Biol. Chem.* **2008**, *283*, 23846.
- (38) Jonsson, T. J.; Tsang, A. W.; Lowther, W. T.; Furdui, C. M. *J. Biol. Chem.* **2008**, *283*, 22890.
- (39) Roussel, X.; Bechade, G.; Kriznik, A.; Van Dorsselaer, A.; Sanglier-Cianferani, S.; Branlant, G.; Rahuel-Clermont, S. *J. Biol. Chem.* **2008**, *283*, 22371.
- (40) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2010**, *46*, 2354.
- (41) Hou, G. H.; Cui, Q. *J. Am. Chem. Soc.* **2013**, *135*, 10457.
- (42) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708.
- (43) Dokainish, H. M.; Gault, J. W. *Biochemistry* **2013**, *52*, 1814.
- (44) MOE, version 2012.10 (2012) Chemical Computing Group Inc., Montreal.
- (45) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (46) Labute, P. *Proteins* **2009**, *75*, 187.
- (47) Gerber, P. R.; Muller, K. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251.
- (48) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2013**, *3*, 198.
- (49) Liao, R. Z.; Thiel, W. *J. Phys. Chem. B* **2012**, *116*, 9396.
- (50) Gaussian 09, Revision D.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.

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

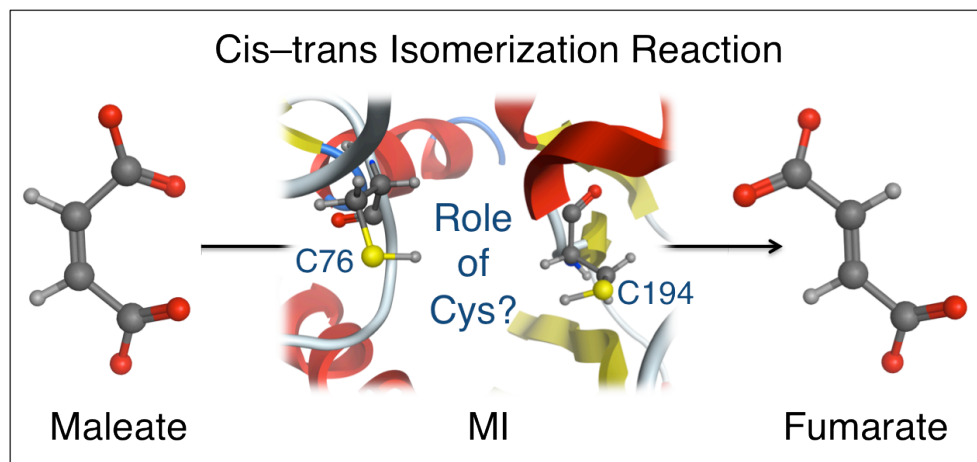
(52) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815.

(53) Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. *J. Am. Chem. Soc.* **2012**, *134*, 4743.

(54) Polyak, I.; Reetz, M. T.; Thiel, W. *J. Phys. Chem. B* **2013**, *117*, 4993.

Chapter 9

Computational Investigations on the Catalytic Mechanism of Maleate Isomerase: the Role of the Active Site Cysteine Residues



9.1 Introduction

For biomolecules, structure and function are often intimately inter-related. Consequently, their chemical and physical diversity is an essential factor for life as it enables them to exhibit numerous differentiated and highly specific functionalities.¹ This range of functionality arises not only from the use of dissimilar molecules but also from geometrical isomers of the same molecule.² For example, L-glutamate is a building block of proteins and has a role as an essential neurotransmitter in all complex living organisms.³ Its stereoisomer D-glutamate, however, does not share these functionalities; for example, it is instead a key component in bacterial cell wall synthesis.⁴ Cells can also differentiate between isomers of exogenous molecules such as therapeutic drugs, e.g., the enantiomers of thalidomide in which one is an effective treatment for morning sickness while the other causes birth defects.^{5,6}

Cis-trans isomerization (CTI) is an important approach for generating geometrical isomers,⁷ and is involved in many biochemical phenomena such as protein folding.⁸ Such reactions require the breaking of a double bond, which typically has a high activation barrier. However, they can be chemically-facilitated by, for example, the use of metal ions, nucleophilic attack or acid/base catalysis.^{8,9} Alternatively, photo-excitation can lead to isomerization via π - π^* singlet and triplet excited states.¹⁰⁻¹²

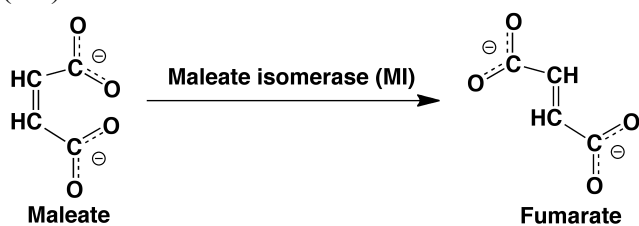
A number of enzymes that catalyze CTI have been examined experimentally and theoretically, revealing a wide range of mechanisms. For example, the glutathione (GSH)-dependent enzyme maleylacetoacetate isomerase, converts maleylacetoacetate to fumarylacetoacetate via transient covalent modification.¹³ The related enzyme maleylpyruvate isomerase utilizes a similar mechanism.¹⁴ In contrast, the mechanism by which retinal isomerases converts 11-trans-retinal to 11-cis-retinal proceeds via a radical intermediate.¹⁵ The peptidyl-prolyl cis-trans isomerase family catalyzes the interconversion of cis/trans peptide bonds that involve prolyl. Notably, they have been

proposed to use a variety of different mechanistic approaches including nucleophilic and proton addition.^{16,17}

Asp/Glu racemase superfamily members share several structural features including a pseudosymmetrical active site-containing domain, carboxylate-containing substrate, dioxyanion hole to help stabilize the carboxylate,¹⁸ and in most members, two catalytic cysteinyls.⁴ The catalytic role of the latter two residues has been experimentally confirmed using site directed mutagenesis,^{9,19,20} and two mechanisms have been proposed for their activation. In particular, in some members such as glutamate racemase (GluRs) the catalytic Cys may be activated by a conserved His and Asp residues.²¹ In contrast, a previous study on proline racemase (ProRs), involving molecular dynamics (MD) simulations suggested that activation may occur via a water or the substrate itself.¹⁸ Regardless of such differences, however, the catalytic mechanism of all members is believed to involve formation of an enediolate intermediate.^{18,22}

Maleate cis-trans Isomerase (MI) is a member of the Asp/Glu racemase superfamily found in bacteria that collectively exhibit a diverse range of functions including catalyzing the conversion of L-amino acids to D-amino acids during cell wall biosynthesis.^{19,23,24} MI is a key enzyme in the metabolic degradation pathway of nicotinic acid.²⁵ Notably, it is utilized by many microorganisms such as *Pseudomonas*, *Alcaligenes*, *Serratia* and *Proteus* to catalyze the geometric isomerization of maleate to fumarate (Scheme 1).^{6,26} The latter is an essential intermediate in the citric acid cycle.²⁷ In addition, it is also an important industrial target as it is involved in aspartic and L-maleic acid production.^{28,29} Recently, there has also been increasing attention in using MI for degradation of tobacco waste.³⁰ Hence, there is great interest in gaining a better understanding of the mechanism and properties of MI and its related enzymes.^{23,29,31,32}

Scheme 9.1. Illustration of the overall isomerization reaction catalyzed by Maleate cis-trans Isomerase (MI).

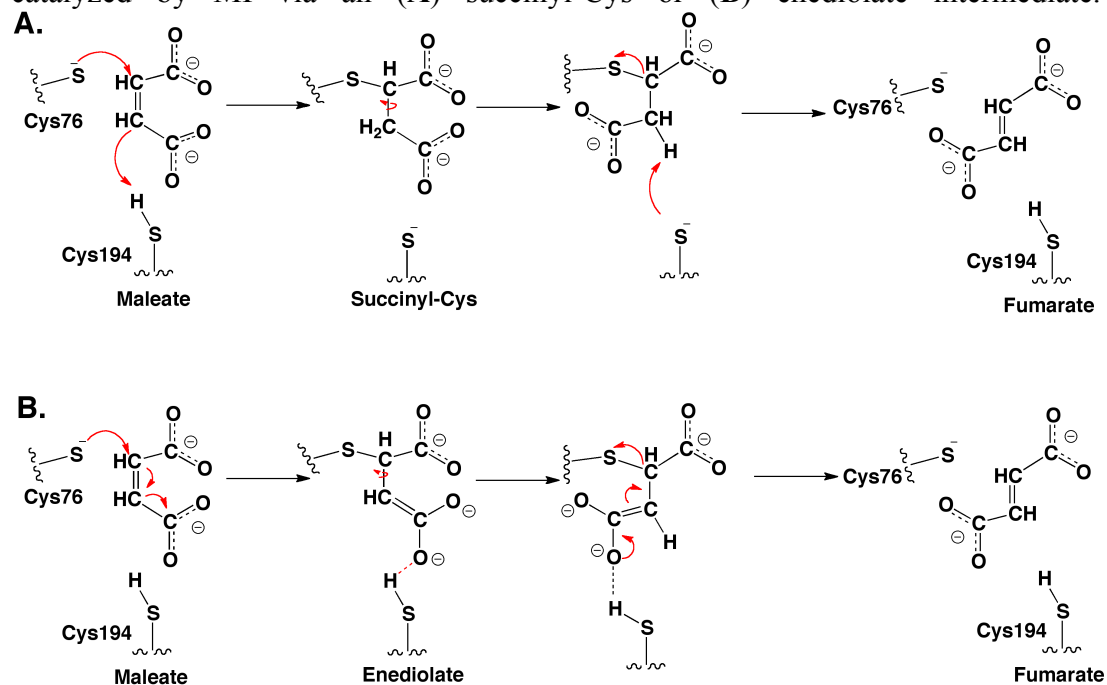


Maleate cis-trans Isomerase (MI) is a cofactor-independent member of the Asp/Glu racemase superfamily.^{26,33} Recently, X-ray crystal structures of both wild-type MI from *Nocardia farcinia* and the corresponding C194A mutant with a succinyl-cysteine intermediate trapped within the active site has been determined.³⁴ In the same study, site directed mutagenesis studies showed that mutation of Cys194 to Ala results in enzyme inactivation.³⁴ Furthermore, substitution of either active site cysteinyl Cys194 or Cys76 by serine reduces the rate of reaction of MI by 8000 and 1474-fold, respectively.³⁴ In addition, it was noted that as with other Asp/Glu racemase members several residues form a dioxyanion hole to help stabilize reaction intermediates.^{9,34} However, the more hydrophobic nature of MI's active site likely results in a less effective stabilization.³⁴

Based in part on these studies, two possible isomerization mechanisms have been proposed as shown in Scheme 2. Both involve an initial direct nucleophilic attack of deprotonated Cys76 (i.e., Cys76S⁻) at the maleate substrate's C2 carbon centre. Notably, it is as yet still unclear how Cys76 is deprotonated (i.e., activated) so that it can more readily act as a nucleophile.³⁴ Furthermore, this nucleophilic role is distinctly different to that observed in other superfamily members such as GluR and arylmalonate decarboxylase (AMD) in which the active site cysteinyls act as acids and/or bases and do not form a covalent enzyme-substrate complex.^{18,35} Importantly, in one pathway (A) this occurs with concomitant transfer of the Cys194 thiol proton

onto the substrate's C3 carbon centre to form a succinyl-cysteine-type intermediate (Scheme 2A). It is important to mention that this intermediate has been observed using X-ray crystallography in the C194A mutant structure at high resolution and it was also confirmed using mass spectrometry.³⁴ In the alternate pathway, however, Cys194 acts simply as a hydrogen bond donor to one of the substrate carboxylates throughout the mechanism, stabilizing its anionic charge (Scheme 2B). That is, pathway **B** proceeds via an enediolate-type intermediate. Despite these differences, the next step in both proposed mechanisms is rotation around the newly formed C2–C3 single bond to give a fumarate-like structure. In the succinyl-Cys pathway (Scheme 2A), the Cys76S–C2 bond dissociates concomitantly with deprotonation of -C3H₂- by Cys194S⁻, thus forming fumarate with regeneration of a neutral Cys194SH. In contrast, in the enediolate pathway (Scheme 2B), cleavage of the Cys76S–C2 bond leads directly to formation of fumarate.

Scheme 9.2. Proposed mechanisms for the maleate/fumarate isomerization reaction catalyzed by MI via an (A) succinyl-Cys or (B) enediolate intermediate.³⁴



Computational chemistry has been shown to provide detailed insights into biological systems, and in particular, enzymatic mechanisms.³⁶⁻³⁹ In this present study, density functional theory-based QM-cluster and ONIOM QM/MM methods have been used to investigate the initial substrate-bound active site complex and protonation states of key residues, as well as the catalytic mechanism of Maleate cis-trans Isomerase (MI).

9.2 Computational Methods

All docking and molecular dynamics simulations were performed using the Molecular Operating Environment (MOE) program,⁴⁰ while all QM-cluster and ONIOM QM/MM calculations were performed using the Gaussian 03⁴¹ and 09⁴² suite of programs. The density functional theory method B3LYP, a combination of Becke's three parameter exchange functional⁴³ and Lee, Yang and Parr's correlation functional⁴⁴ as implemented in the Gaussian programs, was the QM method used in the present calculations.

9.2.1 DFT-Small Model Studies

A series of initial studies were done in order to help determine an appropriate basis set to use in the larger studies (see below), and to examine the effects of the environment's polarity on the protonation state and properties of the substrate, mechanistic intermediates and product. More specifically, optimized geometries were obtained of maleic acid, its mono- and di-anionic deprotonated derivatives, succinate and succinyl-methylthiol using the B3LYP method in combination with basis sets ranging from 6-31G(d) to 6-311+G(2df,p). Effects of a polar environment were included by use of the integral equation formalism polarizable continuum model (IEF-PCM) as implemented in Gaussian.^{41,42} In particular, dielectric constants (ϵ) of 4 and

10 were used to model a protein environment as previously suggested,^{45,46} while $\epsilon=78.39$ was used to model an aqueous environment.

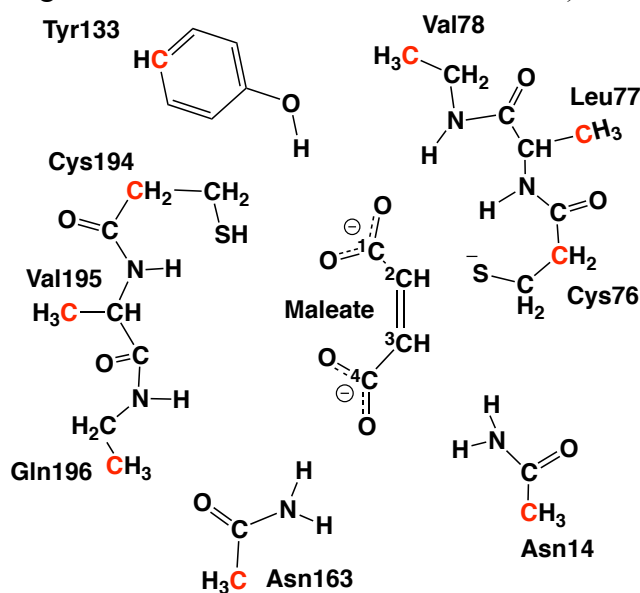
9.2.2 QM-Cluster Studies

All structures were optimized at the B3LYP/6-31G(d,p) level of theory. Relative energies were obtained via single points at the B3LYP/6-311+G(2df,p) level of theory on the optimized structures with the inclusion of the corresponding solvation correction obtained at the IEF-PCM ($\epsilon=4.0$)-B3LYP/6-31G(d,p) level of theory. Frequency calculations were used to characterize transition structures as first-order saddle-points.

A suitable chemical model was derived using the X-ray crystal structure of *Nocardia farcinia*, NfMI (PDB ID: 2XEC).³⁴ Specifically, dianionic maleate was docked into the active site of NfMI; all residues within 10 Å of the catalytic cysteine (Cys76) being considered as the active site. In the above crystal structure the R-group of Cys76 points away from the active site pocket and therefore was manually reoriented prior to docking and MD simulations. Docking was performed using the London dG scoring function followed by optimization of the top 100 generated structures using a force field refinement method using AMBER99. The best 30 scoring structures were then examined visually to choose the most suitable starting structure for further calculations. The active site of the chosen structure was then solvated up to 10 Å from the substrate. The solvated enzyme-substrate complex was then allowed to thermally relax by performing an MD simulation for 1 ns with a time step of 2 fs as has been previously used.^{47,48} A cluster analysis was then performed based on the distance between the sulfur of Cys76 and maleate's -C2H₂- carbon in order to obtain an average structure, which was then optimized using the AMBER99 force field.⁴⁹ From this optimized structure, the active site-bound substrate model

shown in Scheme 3 was obtained for use in the QM-cluster studies. In particular, it included appropriately truncated models of Cys76, Cys194, Val78 and Gln196 as they are known or have been suggested to be catalytically important.³⁴ In addition, the R-groups of Tyr133, Asn14, and Asn163 were included as they directly interact with the substrate, e.g., via hydrogen bonding. Finally, Leu77 and Val195, both modelled as alanyls, were also included. As is common practice⁵⁰ when using QM-cluster models, in order to maintain the integrity of the model, a minimum number of atoms, remote from the reactive region, were kept fixed at their MM optimized coordinates and are highlighted in red in Scheme 3.

Scheme 9.3. The active site-bound substrate chemical model of NfMI used for the QM-cluster studies (atoms fixed at their MM optimized coordinates are highlighted in red and atom numbering used for Maleate carbons is also shown).

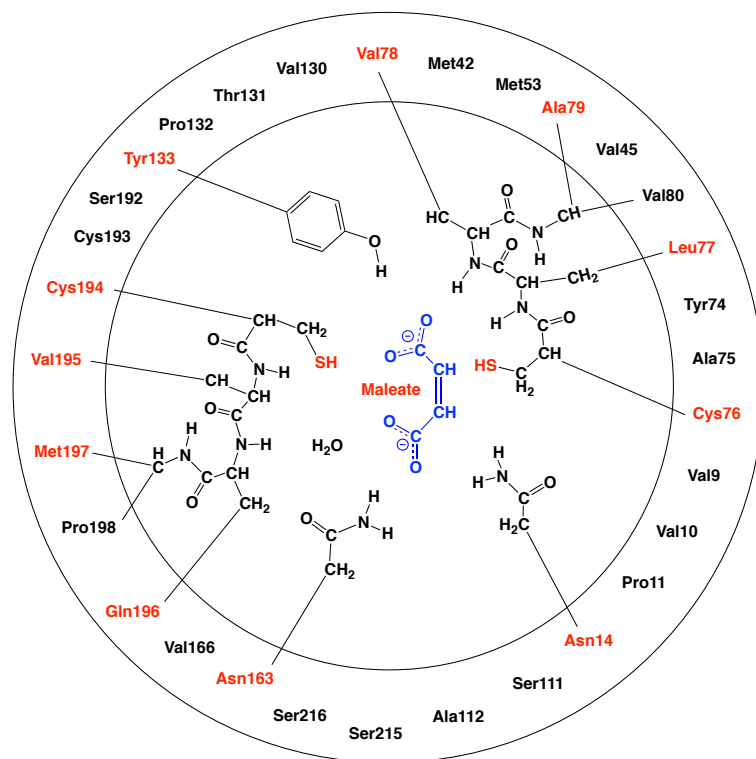


9.2.3 QM/MM Studies

The hybrid quantum mechanics/molecular mechanics (QM/MM) method as implemented in the ONIOM formalism in Gaussian 09⁴² was used for all QM/MM

calculations.⁵¹ The reactive region, high (QM)-layer, was described using the same level of theory and basis set size as per the above QM-cluster approach, B3LYP/6-31G(d,p). The rest of the chemical model, the low (MM)-layer, was described using the AMBER96 MM force field. Relative energies were obtained via single points at the ONIOM(B3LYP/6-311+G(2df,p):AMBER96) level of theory on the above optimized structures.

Scheme 9.4. Schematic illustration of the chemical model, derived from the X-ray crystal structure PDB ID: 2XED, used in the ONIOM QM/MM calculations. The inner circle represents the high (QM)-layer while the outer represents the low (MM)-layer.



A suitable chemical model for use in the QM/MM calculations was obtained from the X-Ray crystal structure of the Cys194Ala NfMI mutant enzyme with a covalently active site-bound succinyl-cysteine intermediate (PDB ID: 2XED).³⁴ A wild-type

Michaelis complex was generated by mutating Ala194 to Cys and cleaving the enzyme-substrate covalent bond. The structure was then minimized using the AMBER99 force field. The resulting minimized structure was then truncated to include all residues within 15 Å of the substrate and is shown in Scheme 4. Within this, the QM-layer was chosen to contain all residues previously used in the QM-cluster chemical model as well as the Gln196–Met197 and Val78–Ala79 peptide bonds. All other residues were placed in the MM-layer. In order to help maintain the model's integrity, and since a large QM-layer was selected, most MM-layer atoms were held fixed at their minimized (see above) positions.⁵² The QM-cluster and QM/MM optimized reactive complexes (RCs) were compared to verify consistency in their structures. Notably, their RMSDs were determined to have only quite negligible differences. In addition, the QM/MM optimized succinyl-Cys intermediate was compared to the corresponding chemical region of the crystallized intermediate (i.e., PDB ID: 2XED)³⁴ and found to have minor RMSDs of just 0.29 Å.

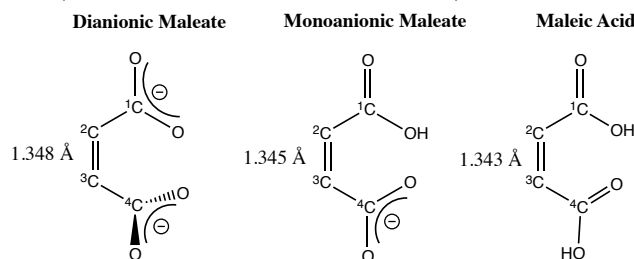
9.3 Results and Discussion

9.3.1 DFT-small model studies on isomerization

As noted in the Introduction, both proposed mechanisms involve nucleophilic attack of a cysteinyl thiolate at the C2/C3 position of dianionic maleate.³⁴ However, one proceeds via a succinyl-Cys and the other an enediolate intermediate. The former occurs with protonation of the adjacent =CH- while the latter does not (see Scheme 2). To help obtain additional insights into factors that may influence maleate to fumarate isomerization, a series of DFT-small model studies were performed. Specifically, we considered formation of these intermediates for all 3 possible ionization states of the substrate; di- (the most common form in aqueous solution)⁵³ and monoanionic maleate, and neutral maleic acid (Scheme 5). The HOMO and LUMO of each species

was also determined, where those of maleate are shown in Figure 9.1.

Scheme 9.5. Schematic illustration of the gas-phase optimized structures of the 3 possible substrate states ($\angle\text{C1-C2-C3-C4} = 0.0^\circ$ in each).



The gas-phase optimized structure of dianionic maleate has a $\angle\text{C1-C2-C3-C4}$ dihedral angle (ϕ) of 0.0° with the two carboxylates almost perpendicular to each other (Supporting Information: Table S1). The C3–C4 bond (1.543 Å) is slightly elongated with respect to C1–C2 (1.537 Å) as the HOMO lies mainly on the C1 carboxylate which lies more in the plane of the carbon backbone. Notably, C2 has only a small contribution to the HOMO orbital while to the LUMO it makes the largest contribution (Figure 9.1). Thus, it would be expected to be the carbon centre most susceptible to nucleophilic attack as has been proposed.³⁴ Increasing the polarity of the environment to 4.0 and 10.0, values commonly used to model the internal environment of a protein, the HOMO and LUMO of maleate had negligible change. It is noted that decreasing the charge in maleate via sequential protonation of the carboxylates has only quite minor effects (-0.003 and -0.002 Å respectively) on the C2=C3 bond length. Notably, however, the carboxylate/carboxylic groups now lie more in the plane of the carbon backbone (see Scheme 5).

For maleate and monoanionic maleate, nucleophilic attack of a methylthiolate (CH_3S^-) at C2 without concomitant protonation of C3 does not give a stable species for all environmental polarities (ϵ values) considered ($\epsilon = 1.0, 4.0, 10.0$ and 78.39

(water)). Specifically, formation of an enediolate is unstable with respect to dissociation of the S–C2 bond. The same occurs for monoanionic maleate when CH_3S^- attacks at C3 (C3C4COOH) instead of C2.

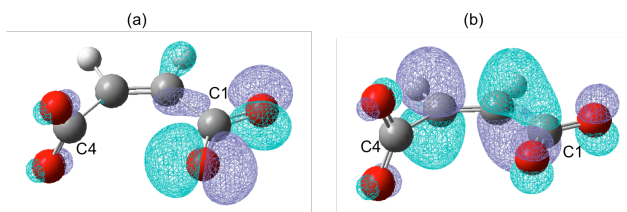


Figure 9.1. The (a) HOMO and (b) LUMO of dianionic maleate in the gas-phase ($\epsilon = 1$).

In contrast, for maleic acid nucleophilic attack of CH_3S^- at C2 gives a stable enediolate structure. This occurs with marked lengthening in the C2–C3 bond from 1.343 to 1.471 Å, respectively. Simultaneously, a significant increase in the $\angle\text{C1-C2-C3-C4}$ angle (ϕ) occurs from 0.0° to 57.6°, respectively. Furthermore, such changes are observed for all values of ϵ ; i.e., regardless of the environments polarity.

A stable succinyl-cysteinyl type intermediate was obtained for all 3 ionization states of the substrate. However, its nature was sensitive to the environment and the ionization state of the substrate. For example, for dianionic maleate, gas-phase formation of such an intermediate occurred with significant increases in both the C2–C3 bond length from 1.35 to 1.53 Å, and the dihedral ϕ from 0.0° to 159.7°. However, as e was increased to 4.0, 10.0 and higher, the magnitude of rotation decreased. In contrast, for maleic acid ϕ in the resulting intermediate was reasonably consistent at approximately 60.0° for all values of e .

Thus, it appears that both the ionization state of the substrate and the polarity of the environment can have significant effects on the stability and nature of possible mechanistic intermediates.

9.3.2 The ionization states of Cys76, Cys194 and the substrate

As described in the Introduction, in both proposed mechanisms, Cys76 acts as the nucleophile while Cys194 acts as a proton or hydrogen bond donor.³⁴ In order for Cys76 to act as a more effective nucleophile it must be deprotonated, as assumed in the mechanisms.³⁴ Knowing the likely initial ionization state of both active site cysteinyl residues is central for understanding their roles and the preferred reaction pathway. Consequently, the proton affinities (PAs) of the Cys76S⁻ and Cys194S⁻ thiolates within the various possible ionization states of the apo-enzyme and substrate-bound active site were examined using the present QM/MM models. The results obtained are shown in Figure 9.2.

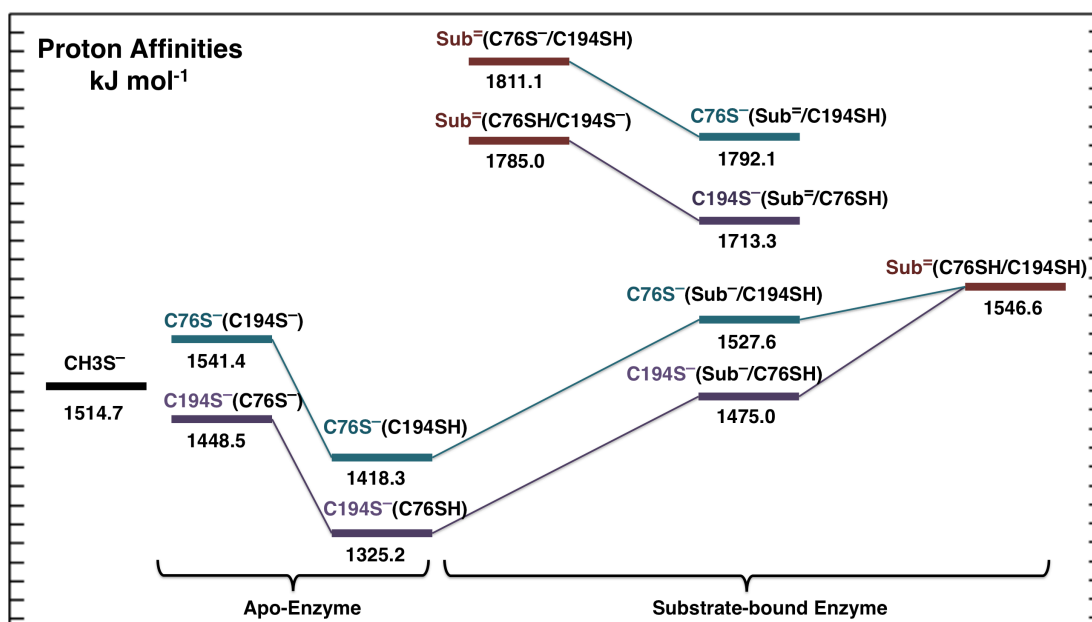


Figure 9.2. The calculated proton affinities (see Computational Methods) of C76S⁻ and C194S⁻ before and after substrate binding, and the PA of the substrate in the active site.

The proton affinity (PA) of methylthiolate (model for deprotonated cysteine) in

aqueous solution is calculated to be $1514.7 \text{ kJ mol}^{-1}$ at the present level of theory (see Computational Methods). For the active site cysteinyl thiolates (i.e., C76S^- and C194S^-), their PAs were first calculated within the apo-enzyme active site for both possible scenarios; where the other cysteinyl is (i) anionic or (ii) neutral. For the first case the PA of C76S^- is $1541.4 \text{ kJ mol}^{-1}$ while that of C194S^- is $1448.5 \text{ kJ mol}^{-1}$. That is, the PA of C76S^- has increased while that of C194S^- has decreased compared to CH_3S^- in aqueous solution. In contrast, in the case where the other cysteinyl is kept neutral the PAs of both C76S^- and C194S^- decrease significantly to 1418.3 and $1325.2 \text{ kJ mol}^{-1}$, respectively. Interestingly, in both scenarios the PA of C194S^- is lowest. This may reflect the fact that as seen in the X-ray crystal structure (PDB ID: 2XED),³⁴ C194 is surrounded by more possible hydrogen bond donors than C76, thus any anionic charge on the former is likely to be better stabilized.

From Figure 9.2 it can be seen that upon binding the dianionic substrate ($\text{Sub}^{=}$), the proton affinity of both C76 and C194 has increased markedly from their corresponding values in all possible apo-enzyme active sites. In contrast, in all substrate-bound active sites containing one or more neutral cysteinyls, the PA of $\text{Sub}^{=}$ has decreased from its calculated value in aqueous solution ($1877.3 \text{ kJ mol}^{-1}$; not shown). Importantly, however, despite these decreases, in those cases where only one of the cysteinyls is neutral, the PA of $\text{Sub}^{=}$ remains higher than that of the ionized cysteinyl. For instance, when C194 is neutral the PA of $\text{Sub}^{=}$ is $1811.1 \text{ kJ mol}^{-1}$ while that of C76S^- is 19.0 kJ mol^{-1} lower at $1792.1 \text{ kJ mol}^{-1}$. Similarly, when C76 is neutral, the decreased PA of $\text{Sub}^{=}$ ($1785.0 \text{ kJ mol}^{-1}$) is still 71.7 kJ mol^{-1} higher than that of C194S^- ($1713.3 \text{ kJ mol}^{-1}$). Thus, in either system $\text{Sub}^{=}$ will be preferentially protonated over the ionized active site cysteinyl.

Indeed, the proton affinity of $\text{Sub}^{=}$ when both cysteinyls are neutral is $1546.6 \text{ kJ mol}^{-1}$, while that of the ionized cysteinyl in $\text{C76S}^-/\text{Sub}^-/\text{C194SH}$ and $\text{C76SH}/\text{Sub}^-$

/C194S⁻ is 1527.7 and 1475.0 kJ mol⁻¹, respectively. It is noted that in the latter system the PA of C194S⁻ is also lower than that of methylthiolate in aqueous solution.

These results thus suggest that the substrate-bound active site prefers to exist as having a monoanionic maleate and ionized Cys194 (i.e., C194S⁻), but a neutral Cys76 (i.e., C76SH). This further suggests that the substrate itself may be able to play a role in activating an active site cysteinyl, specifically Cys194, to be the required nucleophile, while Cys76 may instead act as the proton or hydrogen bond donor.

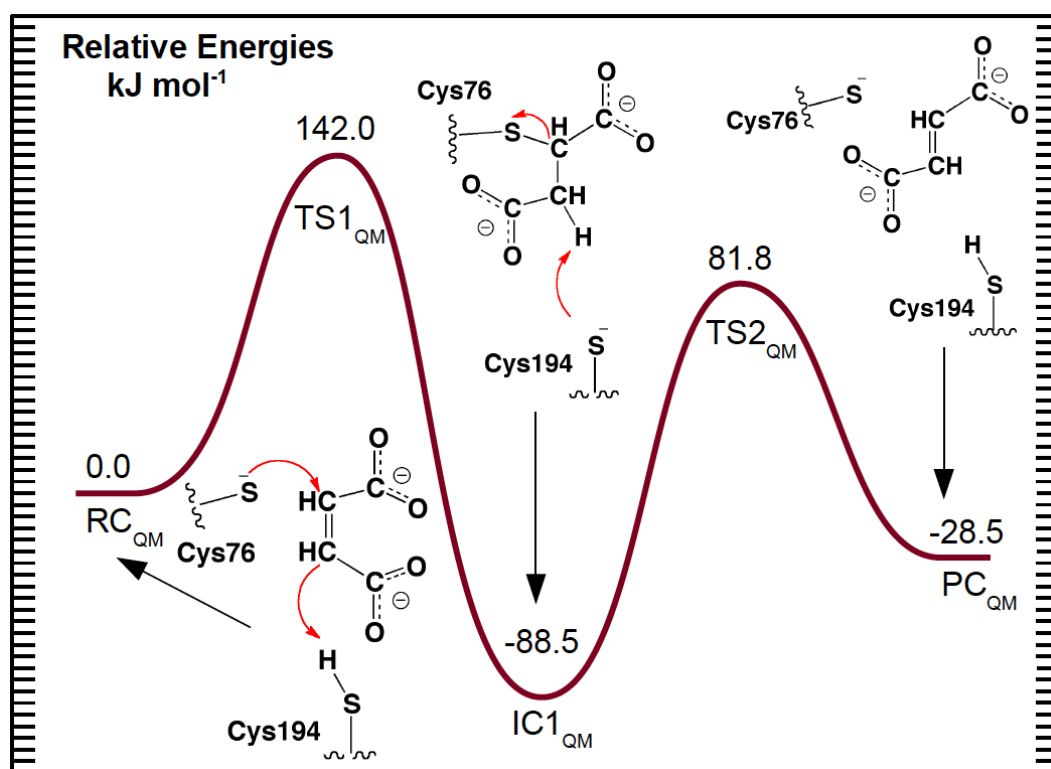


Figure 9.3. PES obtained using a QM-cluster approach (see Computational Methods) of the catalytic mechanism of MI in which Cys76 acts as a nucleophile.

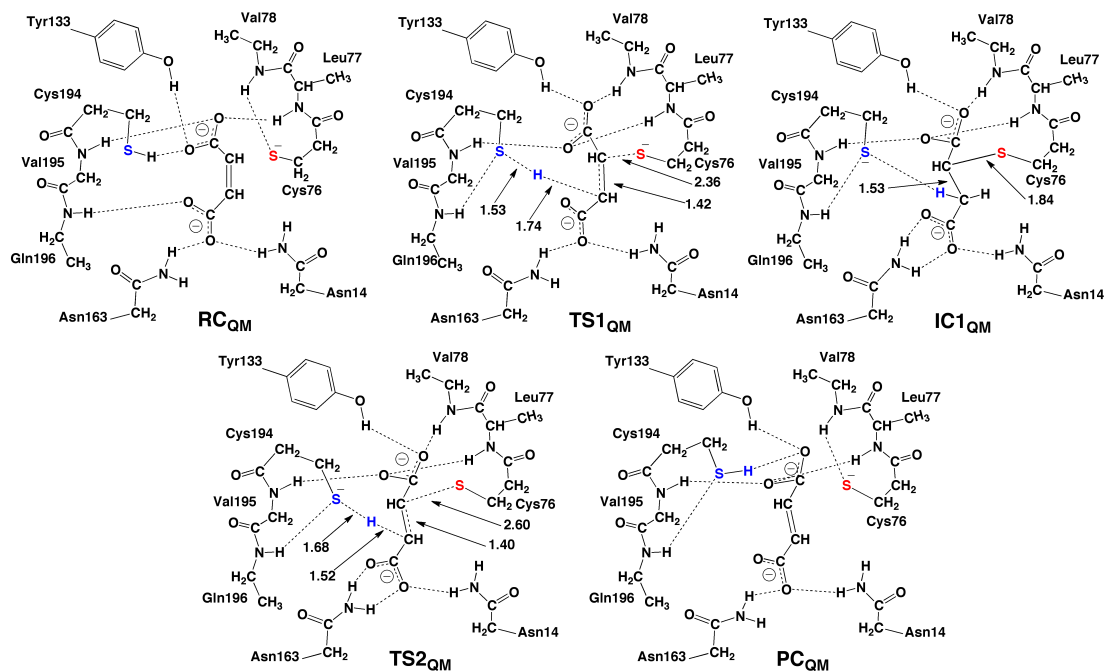
9.3.3 Cys76-pathway: mechanism with Cys76 as the nucleophile

In both of the experimentally proposed mechanisms the thiol of Cys76 is deprotonated and acts as a nucleophile to attack the C2 position of dianionic maleate,

while Cys194 is neutral. Such possible catalytic mechanisms were thus considered using a QM-cluster approach. The potential energy surface (PES) obtained is shown in Figure 9.3 while the corresponding optimized stationary point structures, with selected distances, are shown in Scheme 9.6.

In the optimized structure of the reactant complex, \mathbf{RC}_{QM} , the substrate's carboxylates are each stabilized by multiple hydrogen bonds. More specifically, $-\text{C1OO}^-$ forms hydrogen bonds with the R-groups of Cys194 and Tyr133, and the backbone $-\text{NH}-$ moieties of Leu77 and Val195. The $-\text{C4OO}^-$ carboxylate, in contrast, is stabilized by only three hydrogen bonds formed with the R-groups of Asn14 and Asn163, and the backbone $-\text{NH}-$ of Gln196. Meanwhile, the Cys76S⁻ forms just a single hydrogen bond with the backbone $-\text{NH}-$ of Val78.

Scheme 9.6. Schematic illustration of the optimized structures obtained using a QM-cluster approach (see Computational Methods) for the mechanism in which Cys76 acts as nucleophile



Using the QM-cluster approach no mechanism involving an enediolate-intermediate could be characterized. However, an alternate possible pathway involving a succinyl-Cys type intermediate was obtained. The latter begins with nucleophilic attack of the Cys76S⁻ at the substrate's C2 center with a concomitant proton transfer from the thiol of Cys194 onto C3. This step occurs via **TS1_{QM}** with a markedly high relative energy barrier of 142.0 kJ mol⁻¹. In the optimized structure of **TS1_{QM}** (Scheme 6) the Cys76S⁻⋯C distance has shortened considerably to 2.36 Å while the C3⋯H⋯SCys194 distances are 1.74 and 1.53 Å, respectively. These distances further illustrate the concomitant formation of the Cys76S—C2 bond and proton transfer from Cys194. Notably, during formation of the succinyl-Cys intermediate the ∠C1-C2-C3-C4 (ϕ) increases from 1.1° to 68.0° while the C2—C3 bond has lengthened to 1.42 Å; it now has significantly reduced double bond character. That is, **IC1_{QM}** resembles more a fumarate-like structure.

The resulting intermediate (**IC1_{QM}**) formed lies 88.5 kJ mol⁻¹ lower in energy than **RC_{QM}**. The dihedral angle ϕ has significantly increased to 210.8 (-149.2)° while the C2—C3 bond has lengthened to 1.53 Å, i.e., is now essentially a single bond with a *trans*-like orientation of the substrate's carbon backbone. The hydrogen bond network between the substrate and active site residues is generally retained, with only some minor differences. For example the Gln196 –NH– backbone now hydrogen bonds to the thiolate of Cys76 instead of the substrate's –C4OO⁻ group.

In the next and final step the Cys76S—C2 bond is cleaved while concomitantly the –C3H₂– moiety transfers a proton onto the Cys194S⁻ thiolate. This concerted step proceeds via **TS2_{QM}** with a barrier of 81.8 kJ mol⁻¹ with respect to **RC_{QM}**; 160.3 kJ mol⁻¹ with respect to **IC1_{QM}**. That is neither step 1 or 2 are likely to be enzymatically feasible.^{54,55} In **TS2_{QM}**, the Cys76S⁻⋯C2 bond has elongated to 2.60 Å; while the Cys194S⁻⋯H⋯C3 distances are 1.68 and 1.52 Å, respectively. As a result the C2—C3

bond now has regained some double bond character. Furthermore, the dihedral angle ϕ is now 182.9°. The product complex (**PC_{QM}**) lies 28.5 kJ mol⁻¹ lower in energy than **RC_{QM}**. Notably, the C2—C3 bond is now formally a double bond with a distance of 1.34 Å while ϕ has increased slightly to 191.9°.

It is noted that the corresponding "succinyl-Cys" mechanism in which the proton transfers involving Cys194 occurred via a H₂O moiety was also examined. However, the relative energies with respect to **RC_{QM}** of **TS1_{QM}** and **TS2_{QM}** increased significantly to 123.0 and 193.9 kJ mol⁻¹, respectively.

An ONIOM(QM/MM) approach (see Computational Methods) was then used to further examine possible mechanisms in which Cys76 may act as the nucleophile. In particular, mechanisms in which Cys76 and Cys194 may initially be neutral, as suggested by the above PA calculations, were considered. The PES obtained is shown in Figure 9.4 while the corresponding optimized structures, with selected distances, are shown in Scheme 9.7.

Again, in the reactant complex the -C1OO⁻ carboxylate is stabilized via multiple strong hydrogen bonds with the side chain hydroxyl of Tyr133 and the backbone -NH- functionalities of Ala79, Leu77 and Val195. Similarly, the -C4OO⁻ group is again stabilized via three strong hydrogen bonds with the side chains of Asn14 and Asn163, and the backbone-NH- of Gln196. It is noted that the thiol of Cys194 forms a weak hydrogen bond with -C1OO⁻; $r(\text{SH}\cdots\text{O}) = 2.52 \text{ \AA}$.

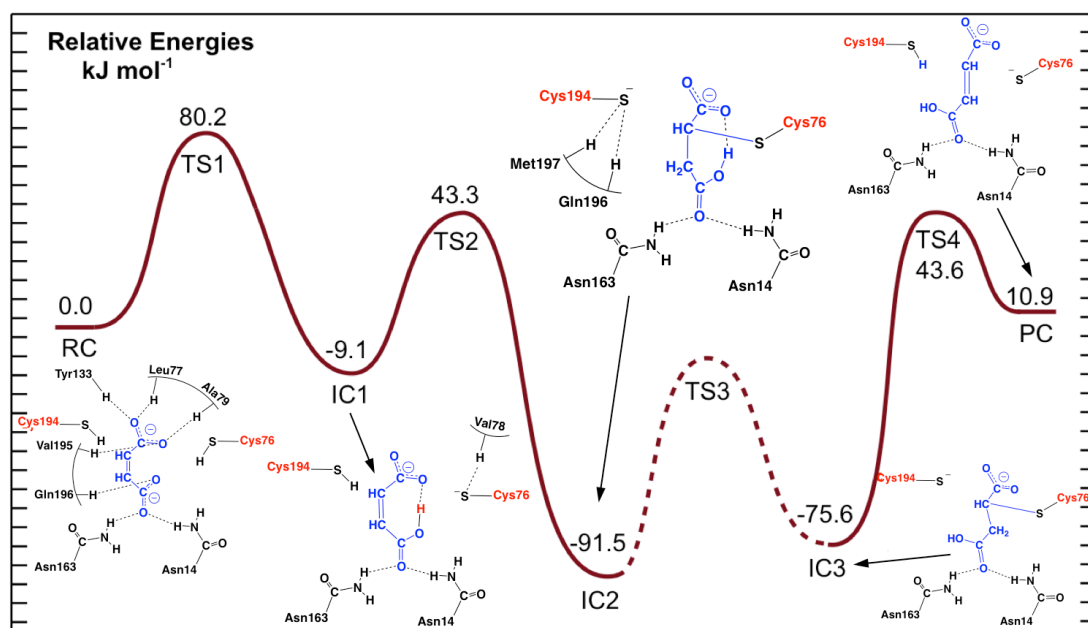


Figure 9.4. PES obtained using an ONIOM(QM/MM) approach (see Computational Methods) of the catalytic mechanism of MI in which Cys76 acts as a nucleophile.

As noted, Cys76SH is proposed to act the nucleophile after it has been activated, i.e., deprotonated. Unfortunately, Cys76 is situated in a hydrophobic region with no suitable candidate base residue to cause its activation. However, an alternate possibility is that the substrate may be able to facilitate such a process due to its carboxylates. Indeed, the PA calculations described above suggest that the substrate may have a suitably high-enough PA compared to Cys76SH.

In **RC**, the Cys76SH proton is 2.85 and 3.70 Å from the nearest oxygen of the $-C1OO^-$ and $-C4OO^-$ groups, respectively (Table S2). The first step of the overall mechanism is transfer of the thiol proton of Cys76 via the $-C1OO^-$ moiety onto the $-C4OO^-$ group. This occurs via **TS1** with a barrier of 80.2 kJ mol⁻¹ with respect to **RC** (Figure 9.4). In **TS1** (Scheme 7) the proton being transferred is essentially on the $-C1OO^-$ group; $r(C1OO^- \cdots H^+) = 1.05$ Å. Simultaneously, however, it lies about midway between the Cys76 thiolate sulfur and the nearest oxygen of the $-C4OO^-$

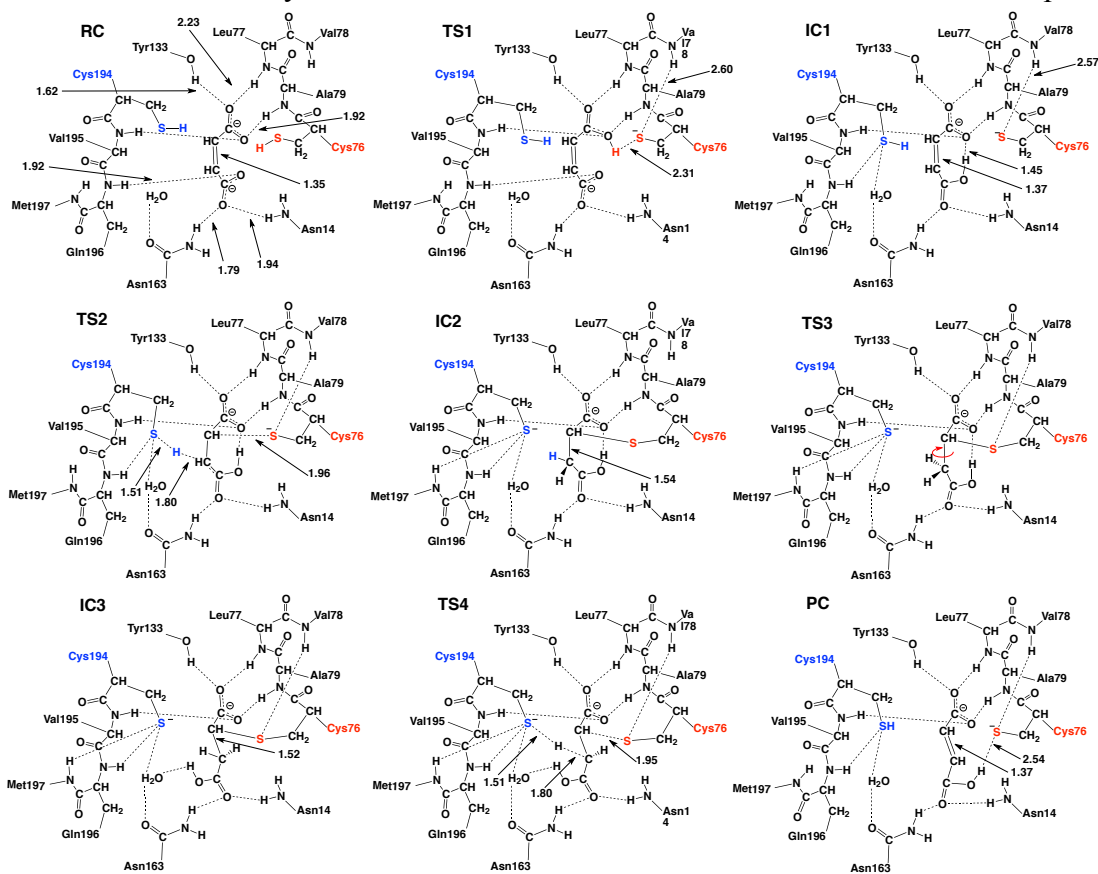
moiety with distances of 2.31 and 2.15 Å, respectively (see Scheme 7). It is noted that no intermediate was obtained with the Cys76SH proton on the $-C1OO^-$ group. This may reflect that this carboxylate already makes stronger and more hydrogen bonds with active site residues than the $-C4OO^-$ group. Thus, the $-C1OO^-$ group's anionic character is better stabilized than that of the $-C4OO^-$ group. The Cys76S⁻ thiolate now appears suitably positioned to attack C2 (Scheme 7).

In the resulting intermediate formed, **IC1**, the proton from Cys76SH has been transferred fully to the $-C4OO^-$ group. That is, the substrate is now a monoanionic maleate with a neutral Cys194. Meanwhile, the anionic Cys76S⁻ forms a single weak hydrogen bond (2.57 Å) with the backbone $-NH-$ of Val78. More importantly the Cys76S⁻⋯C2 distance has shortened to 2.75 Å while the C2–C3 bond has elongated from 1.35 to 1.37 Å. **IC1** lies lower in energy than **RC** by just 9.1 kJ mol⁻¹.

As suggested by the DFT-small model studies (see above), a stable enediolate intermediate could not be obtained within the active site using the QM/MM model. Rather, the next step is nucleophilic attack of the thiolate of Cys76S⁻ at the C2 carbon centre of the substrate. Concomitantly, the thiol of Cys194 transfers its proton onto the substrate's adjacent C3 center. This concerted step proceeds via **TS2** at a cost of 52.4 kJ mol⁻¹ with respect to **IC1**. This is illustrated by the fact that in the optimized structure of **TS2** the Cys76S⁻⋯C2 distance has shortened significantly to 1.96 Å. Meanwhile, the Cys194SH proton now lies between the Cys194 sulfur and C3 with distances of 1.51 and 1.80 Å, respectively.

Scheme 9.7. Schematic illustration of optimized structures obtained using an ONIOM(QM/MM) approach (see Computational Methods) for the mechanism in

which **Cys76** acts as nucleophile.



The resulting succinyl-Cys intermediate formed, **IC2**, lies significantly lower in energy than **RC** by 91.5 kJ mol^{-1} . Also, the C2–C3 bond distance is now 1.54 \AA . Notably, it corresponds to the crystallographically obtained "intermediate" in a Cys194Ala mutant MI enzyme.³⁴

The next step is likely rotation about the C2–C3 bond (1.52 \AA) to give an enzyme-bound fumarate-like intermediate **IC3**. Such a species similarly lies considerably lower in energy than **RC** by 75.6 kJ mol^{-1} . Unfortunately, no TS (**TS3**) for such a rotation could be optimized at the present level of theory. However, the barrier is

expected to be feasible as observed in the Cys194-pathway discussed below (see Figure 9.5).

The subsequent and final step is then formation of the product complex **PC**; an active site-bound fumarate. This concerted step, involving both cleavage of the C2—SCys76 bond and a proton transfer from the substrate's -C3H2- group to Cys194, occurs via **TS4** at a cost of $119.2 \text{ kJ mol}^{-1}$ with respect to **IC3**. Notably, the energy difference between **IC2** and **TS4** is $135.1 \text{ kJ mol}^{-1}$; which has been suggested to be thermodynamically greater than that which is enzymatically feasible.^{54,55} This also represents the largest difference between any minimum and TS along this possible mechanism, and is likely the rate limiting step for such a mechanism. The apparent considerable energy required for **IC2** to either proceed to product or back to reactant (requires $171.7 \text{ kJ mol}^{-1}$) may provide insights into the ability of experimentalists to obtain an X-ray crystal structure of such an intermediate in a Cys194Ala mutant enzyme. Notably, the detection of the succinyl-Cys in C194A suggests the presence of an alternate proton-transferring agent that helps stabilize the intermediate. However, the absence of a suitable mechanistic base (i.e., loss of Cys194S⁻) does not allow for product formation.

The final product complex, **PC**, lies slightly higher in energy than **RC** by 10.9 kJ mol^{-1} , indicating that the overall mechanism is endothermic. The C2–C3 double bond is 1.37 \AA , further indicating the formation of the fumarate product.

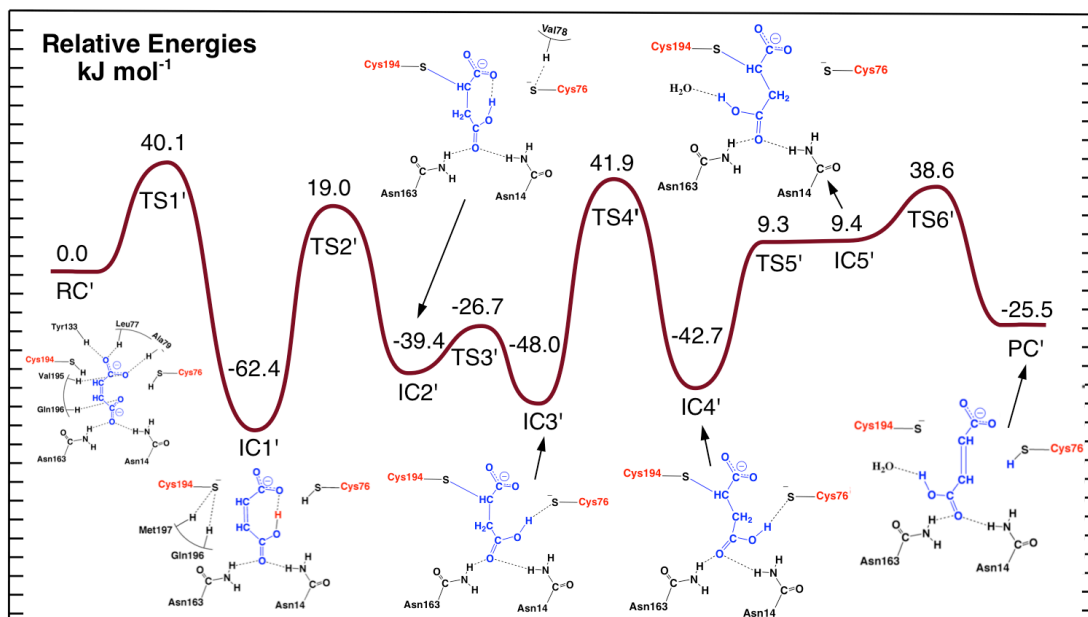


Figure 9.5. PES obtained using an ONIOM(QM/MM) approach (see Computational Methods) of the catalytic mechanism of MI in which Cys194 acts as a nucleophile.

9.3.4 Cys194-pathway: mechanism with Cys194 as the nucleophile

As noted above, the PA calculations suggested that rather than Cys76, Cys194 may in fact be preferably ionized within the substrate-bound active site. Hence, possible catalytic mechanisms in which Cys194 may act as the nucleophile were investigated. The resulting PES obtained is shown in Figure 9.5 while the corresponding optimized stationary point structures are illustrated in Scheme 8.

Beginning from the same **RC** as for the above mechanism in which Cys76 acts as the nucleophile, the first step is the analogous substrate facilitated activation of Cys194. This step, however, proceeds via **TS1'** at a cost relative to RC of only 40.1 kJ mol⁻¹ (Figure 9.5). This is half that required to activate Cys76 via an analogous pathway (cf. Figure 9.4). Furthermore, unlike that observed for activation of Cys76, the Cys194 thiol proton is transferred directly onto an oxygen of the -C4OO⁻ group

with $r(\text{Cys194S}\cdots\text{H})$ and $r(\text{H}\cdots\text{OOC4})$ distances in **TS1'** of 1.63 and 1.24 Å, respectively (Scheme 7). The resulting intermediate formed **IC1'**, lies lower in energy than **RC** by 62.4 kJ mol⁻¹. Notably, this is 53.3 kJ mol⁻¹ lower in energy than **IC1** on the Cys76-pathway (cf. Figure 9.4). This is likely due in part to the fact that in contrast to the single weak stabilizing hydrogen bond observed in **IC1**, in **IC1'** the Cys194S⁻ thiolate forms three hydrogen bonds. Specifically, it forms two with the backbone -NH-'s of Gln196 and Met197, and one with the amide side chain of Asn163 via a water molecule. Meanwhile, as in **IC1**, the proton from Cys194 is wholly transferred onto the substrate's -C4OO⁻ group and now forms an intermolecular hydrogen bond with the -C1OO⁻ moiety. Also, the C2-C3 bond in **IC1'** (1.35 Å) remains little changed from that obtained for **RC**, in contrast to that observed for **IC1** in the alternate Cys76-pathway.

As for the Cys76-pathway the subsequent step is formation of a succinyl-Cys type intermediate (**IC2'**). Again this involves nucleophilic attack of the thiolate, though now it is Cys194S⁻, at the substrate's C2 center, with concomitant transfer of the thiol proton from the second active site cysteine (now Cys76) onto the adjacent C3 center. No stable enediolate intermediate could be found. This step occurs via **TS2'** with a barrier of 19.0 kJ mol⁻¹ with respect to **RC**, or 81.4 kJ mol⁻¹ relative to **IC1'**. The resulting succinyl-Cys intermediate **IC2'** lies lower in energy than **RC** by -39.4 kJ mol⁻¹. However, notably, this is in fact 52.1 kJ mol⁻¹ *higher* in relative energy than the same corresponding intermediate **IC2** of the Cys76-pathway (cf. Figure 9.4). The C2-C3 bond has now elongated to 1.55 Å; that is, it is now a single bond.

At the heart of the isomerization mechanism is rotation about the C2-C3 bond, i.e., the *cis-trans* isomerization. In contrast to the seemingly one-step isomerization upon formation of a succinyl-Cys intermediate for the Cys76-pathway, a multi-step process was obtained on the Cys194-pathway (Figure 9.5). This process

itself can be thought to occur in 3-stages. In the first, the intramolecular -C4OOH...⁻OOC1- hydrogen bond is broken. Instead, the -C4OOH group now forms a strong hydrogen bond with the thiolate of Cys76S⁻; $r(\text{C4OOH}\cdots\text{SCys76}) = 1.84 \text{ \AA}$. This step occurs via **TS3'** with a low barrier of just 12.7 kJ mol^{-1} to give the alternate succinyl-Cys type intermediate **IC3'**. The latter in fact lies slightly lower in energy than **IC2'** by 8.6 kJ mol^{-1} . More importantly, the $\angle\text{C2-C3-C4-O}$ increases from 44.8° to 71.8° . Notably, the twist about the C2—C3 bond, i.e., $\angle\text{C1-C2-C3-C4}$ (f), has slightly changed from that observed in **IC2'**, -294.3° to 306.0° in **IC3'**.

The next step is essentially the twist from a cis conformation to trans. This occurs via **TS4'** with a barrier of 41.9 kJ mol^{-1} relative to **RC** or 89.9 kJ mol^{-1} with respect to **IC3'**. In the resulting alternate succinyl-Cys type intermediate **IC4'**, lying just slightly higher in energy than **IC3'** by 5.3 kJ mol^{-1} , $\angle\text{C1-C2-C3-C4}$ (f), has increased significantly to 137.2° . All active site-substrate interactions observed for **IC3'** are maintained (and the C2—C3 bond length stays the same as in **IC3'**). Importantly, this represents the rate-limiting step along the Cys194-pathway having both the highest barrier for a single reaction step and relative to **RC**. Furthermore, it is in good agreement with the barrier of $\sim 70 \text{ kJ mol}^{-1}$ calculated using experimental kinetics measurements.³⁴ In addition, it is significantly lower than the $135.1 \text{ kJ mol}^{-1}$ required for the rate-limiting step of the alternate Cys76-pathway: going from **IC2** to **PC**.

The third stage is cleavage of the Cys76S⁻...HOOC4 hydrogen bond. This step occurs via **TS5'** at a cost of 52.0 kJ mol^{-1} with respect to **IC4'**, or 9.3 kJ mol^{-1} relative to **RC**. In the resulting alternate succinyl-Cys type intermediate formed, **IC5'**, the C2—C3 bond has shortened by 0.02 \AA to 1.53 \AA , while the dihedral angle $\angle\text{C1-C2-C3-C4}$ is now 160.4° . More importantly, the C4OH group now forms a quite strong hydrogen bond of length 1.80 \AA with an active site water molecule (Scheme 8). It should be noted that this latter water simultaneously forms a weak hydrogen bond (2.45 \AA) with

the sulfur of Cys194 (Scheme 8). In addition, the distance between the thiolate sulfur of Cys76 and the nearest -C3H₂- proton has now decreased from 3.05 Å in **IC4'** to 2.45 Å in **IC5'**. Thus, the Cys76 now seems well-positioned relative to the substrate to abstract a proton. Thermodynamically, **IC5'** has an energy relative to **RC** of 9.4 kJ mol⁻¹, 0.1 kJ mol⁻¹ above that of **TS5'**. This is a common artefact of single-point energy calculations on a flat PES and indicates that the reverse reaction, **IC5'** to **IC4'**, essentially occurs without a barrier.

The final step is proton abstraction by the thiolate of Cys76 from the substrates -C3H₂- moiety, with concomitant cleavage of the Cys194S—C2 bond. However, unlike the Cys76-pathway where the analogous step occurred with a high barrier, on the Cys194-pathway this step proceeds via **TS6'** at a very low cost of only 29.9 kJ mol⁻¹ relative to **IC5'**. The concerted nature of this step is highlighted by the fact that in **TS6'** the Cys194S...C2 bond has elongated to 1.97 Å, while the C3H...SCys76 and C3—HSCys76 distances are now 1.65 and 1.50 Å, respectively. The final product complex **PC**, in which fumarate is now non-covalently bound within the active site, is 25.5 kJ mol⁻¹ lower in energy than **RC** with a C2—C3 bond of 1.35 Å. Thus, unlike that calculated above for the Cys76-pathway, this alternate mechanism in which Cys194 acts as the nucleophile is thermodynamically favoured.

9.4 Conclusions

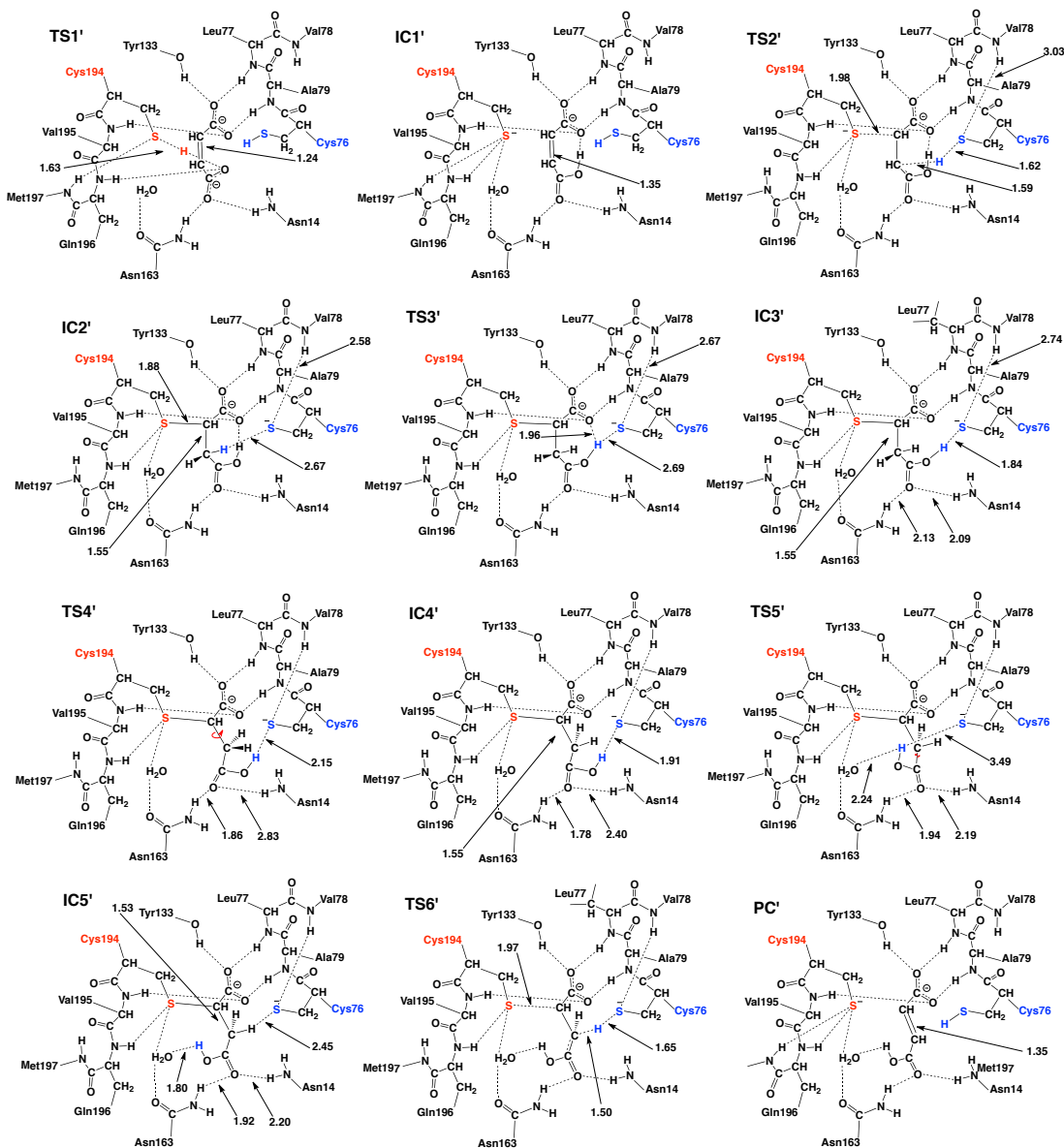
In this study the mechanism by which the enzyme maleate isomerase catalyzes the cis-trans interconversion of maleate and fumarate has been computationally investigated. Specifically, DFT methods in combination with small chemical models were used to gain fundamental insights into the nature of possible mechanistic intermediates, while QM-cluster and ONIOM(QM/MM) methods have been used to examine the nature of the substrate-bound active site and the catalytic mechanism.

The feasibility of the formation of proposed enediolate and succinyl-Cys type intermediates from a neutral, monoanionic or dianionic maleate substrate was initially examined using DFT-small chemical model studies. For both neutral and monoanionic maleate an enediolate intermediate, formed by nucleophilic attack of methyl thiolate (CH_3S^-) at their C2 center, was unstable, i.e., dissociated back to the original substrates. In contrast, the succinyl-Cys type intermediate, formed by concomitant thiolate addition to C2 and protonation of C3, was stable for all 3 ionization states of the initial maleate. Furthermore, for dianionic maleate the formation of succinyl-Cys leads to a barrierless rotation about the substrates central C2—C3 bond. Increasing the polarization of the environment was observed to decrease the angle of rotation. In order to understand the driving force for rotation, the LUMO and HOMO were also examined.

The ionization state of the active site cysteines (Cys76 and Cys194) and the substrate were examined using an ONIOM(QM/MM) based approach for both the apoenzyme and the substrate-bound active site. The results suggest that the proton affinity (PA) of Cys76S^- is higher than that of Cys194S^- in both the apo-enzyme and when the dianionic maleate substrate is bound within the active. Furthermore, the combined results suggest that upon substrate binding the preferred initial state of the substrate-bound active site contains a monoanionic maleate substrate, an ionized Cys194 (i.e., Cys194S^-) and a neutral Cys76 (i.e., Cys76SH).

Two possible mechanistic pathways were investigated using QM-cluster and/or an ONIOM(QM/MM) approach. The first pathway corresponds to that previously experimentally proposed in which the active site residue Cys76 acts as the mechanistic nucleophile that attacks the substrates sp^2 C2 centre; the Cys76-pathway. In the alternate pathway Cys194 acts as the required nucleophile; the Cys194-pathway.

Scheme 9.8. Schematic illustration of optimized structures obtained using an ONIOM(QM/MM) approach (see Computational Methods) for the mechanism in which Cys194 acts as nucleophile.



In both pathways the Cys76 or Cys194 thiol, can be deprotonated via proton transfer onto the maleate substrate's carboxylate positioned nearest Asn14 and Asn163. However, the barrier for this step on the Cys194-pathway (40.1 kJ mol^{-1}) is only half that of the analogous step on the Cys76-pathway (80.2 kJ mol^{-1}). This is due

in part to greater stabilization of Cys194S⁻ than Cys76S⁻ by hydrogen bonding within the active site.

The subsequent step in these pathways is the nucleophilic attack of the thiolate at the maleate substrate's C2 centre. Additionally, this occurs with concomitant proton transfer from the R-group thiol of the second active site cysteinyl to give a succinyl-Cys intermediate (**IC2** and **IC2'**, respectively). Importantly, in the case of the Cys76-pathway such an intermediate (**IC2**) lies significantly lower in energy than the initial reactant complex (**RC**) by 91.5 kJ mol⁻¹. Furthermore, subsequent rearrangement and reaction of **IC2** to give the final product complex (**PC**) requires a considerable amount of energy; 135.1 kJ mol⁻¹. This is in fact the rate-limiting process of the Cys76-pathway. In contrast, for the Cys194-pathway the resulting succinyl-Cys intermediate (**IC2'**), while lower in energy than the **RC**, is not as significantly stabilized, lying just 39.4 kJ mol⁻¹ lower in energy than **RC**. It is able to undergo a series of relatively low-energy rearrangements and reactions to give the final product complex (**PC'**). The rate-limiting step along the Cys194-pathway is the step in which a twist from a cis conformation to trans occurs. The calculated barrier for this step is 89.9 kJ mol⁻¹ is in good agreement with the barrier calculated using experimentally determined kinetics; ~70 kJ mol⁻¹.³⁴

Thus, the present results suggest that the overall catalytic mechanism of Maleate Isomerase is initiated by a substrate-assisted activation of the active site cysteinyl Cys194. This is followed by the concerted formation of a succinyl-Cys intermediate in which Cys76 acts as an acid. Thus, Cys194 is the mechanistic nucleophile while Cys76 instead acts as an acid/base along the mechanism. The desired rotation about the C2—C3 bond occur via multiple relatively low-barrier steps with assistance of the anionic Cys76S⁻, generated during formation of a succinyl-Cys intermediate. Finally, cleavage of the Cys194S—C2 bond concomitant with proton abstraction by Cys76S⁻

from the intermediates C3H2 group leads to the formation of the fumarate product. The conserved nature of the active site in the racemase superfamily suggests possible transferability of the mechanism outlined to other species in the family.

9.5 References

- (1) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13.
- (2) Fischer, G. In *Cis-Trans Isomerization in Biochemistry*; Dugave, C., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006.
- (3) Hrabovszky, E.; Liposits, Z. *J. Neuroendocrinol.* **2008**, *20*, 743.
- (4) Mixcoha, E.; Garcia-Viloca, M.; Lluch, J. M.; Gonzalez-Lafont, A. *J. Phys. Chem. B* **2012**, *116*, 12406.
- (5) Trapnell, C. B. *AIDS clinical care* **1998**, *10*, 3.
- (6) Fabro, S.; Smith, R. L.; Williams, R. T. *Nature* **1967**, *215*, 296.
- (7) Dugave, C. In *Cis-Trans Isomerization in Biochemistry*; Dugave, C., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006.
- (8) Koyama, Y.; Kakitani, Y.; Nagae, H. In *Cis-Trans Isomerization in Biochemistry*; Dugave, C., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006.
- (9) Fisch, F. A., PhD thesis, University of York, 2009.
- (10) Wang, L. X.; Yi, C. H.; Zou, H. T.; Xu, J.; Xu, W. L. *J. Phys. Org. Chem.* **2009**, *22*, 888.
- (11) Hoffman, D. P.; Mathies, R. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 6298.
- (12) Mizuno, K.; Nire, K.; Sugita, H.; Maeda, H. *Tetrahedron Lett.* **2001**, *42*, 2689.
- (13) Polekhina, G.; Board, P. G.; Blackburn, A. C.; Parker, M. W. *Biochemistry* **2001**, *40*, 1567.

- (14) Marsh, M.; Shoemark, D. K.; Jacob, A.; Robinson, C.; Cahill, B.; Zhou, N. Y.; Williams, P. A.; Hadfield, A. T. *J. Mol. Biol.* **2008**, *384*, 165.
- (15) Redmond, T. M.; Poliakov, E.; Kuo, S.; Chander, P.; Gentleman, S. *J. Biol. Chem.* **2010**, *285*, 1919.
- (16) Fanghanel, J.; Fischer, G. *Front. Biosci.* **2004**, *9*, 3453.
- (17) Agarwal, P. K. *Proteins* **2004**, *56*, 449.
- (18) Puig, E.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M.; Field, M. J. *J. Phys. Chem. B* **2007**, *111*, 2385.
- (19) Lundqvist, T.; Fisher, S. L.; Kern, G.; Folmer, R. H. A.; Xue, Y. F.; Newton, D. T.; Keating, T. A.; Alm, R. A.; de Jonge, B. L. M. *Nature* **2007**, *447*, 817.
- (20) Hatakeyama, K.; Goto, M.; Uchida, Y.; Kobayashi, M.; Terasawa, M.; Yukawa, H. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 569.
- (21) Puig, E.; Mixcoha, E.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Am. Chem. Soc.* **2009**, *131*, 3509.
- (22) Okrasa, K.; Levy, C.; Hauer, B.; Baudendistel, N.; Leys, D.; Micklefield, J. *Chem.-Eur. J.* **2008**, *14*, 6609.
- (23) de Jonge, B. L. M.; Kutschke, A.; Uria-Nickelsen, M.; Kamp, H. D.; Mills, S. D. *Antimicrob. Agents Chemother.* **2009**, *53*, 3331.
- (24) Zeida, A.; Guardia, C.; Lichtig, P.; Perissinotti, L.; Defelipe, L.; Turjanski, A.; Radi, R.; Trujillo, M.; Estrin, D. *Biophys Rev* **2014**, *1*.
- (25) Scher, W.; Jakoby, W. B. *J. Biol. Chem.* **1969**, *244*, 1878.
- (26) Hatakeyama, K.; Asai, Y.; Uchida, Y.; Kobayashi, M.; Terasawa, M.; Yukawa, H. *Biochem. Biophys. Res. Commun.* **1997**, *239*, 74.
- (27) Tanaka, K.; Kobayashi, K.; Ogasawara, N. *Microbiology-(UK)* **2003**, *149*, 2317.
- (28) Kato, Y.; Yamagishi, J.; Asano, Y. *J. Ferment. Bioeng.* **1995**, *80*, 610.

- (29) Engel, C. A. R.; Straathof, A. J. J.; Zijlmans, T. W.; van Gulik, W. M.; van der Wielen, L. A. M. *Appl. Microbiol. Biotechnol.* **2008**, *78*, 379.
- (30) Chen, D. D.; Tang, H. Z.; Lv, Y.; Zhang, Z. Y.; Shen, K. L.; Lin, K.; Zhao, Y. L.; Wu, G.; Xu, P. *Mol. Microbiol.* **2013**, *87*, 1237.
- (31) Bugg, T. D. H.; Braddick, D.; Dowson, C. G.; Roper, D. I. *Trends Biotechnol.* **2011**, *29*, 167.
- (32) Ichikawa, S.; Iino, T.; Sato, S.; Nakahara, T.; Mukataka, S. *Biochem. Eng. J.* **2003**, *13*, 7.
- (33) Takamura, Y.; Takamura, T.; Soejima, M.; Uemura, T. *Agric. Biol. Chem.* **1969**, *33*, 718.
- (34) Fisch, F.; Fleites, C. M.; Delenne, M.; Baudendistel, N.; Hauer, B.; Turkenburg, J. P.; Hart, S.; Bruce, N. C.; Grogan, G. *J. Am. Chem. Soc.* **2010**, *132*, 11455.
- (35) Obata, R.; Nakasako, M. *Biochemistry* **2010**, *49*, 1963.
- (36) Borowski, T.; Georgiev, V.; Siegbahn, P. *J. Mol. Model.* **2010**, *16*, 1673.
- (37) Stenta, M.; Calvaresi, M.; Altoe, P.; Spinelli, D.; Garavelli, M.; Galeazzi, R.; Bottoni, A. *J. Chem. Theory Comput.* **2009**, *5*, 1915.
- (38) Liao, R.-Z.; Himo, F.; Yu, J.-G.; Liu, R.-Z. *J. Inorg. Biochem.* **2010**, *104*, 37.
- (39) Llano, J. G., J. W. In *Quantum Biochemistry: Electronic Structure and Biological ActiVity*; Matta, C. F., Ed.; , Ed.; Wiley-VCH: Weinheim: 2010; Vol. 2, p 643.
- (40) MOE, version 2012.10 (2012) Chemical Computing Group Inc., Montreal.
- (41) Gaussian 03, Revision C.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; 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.;

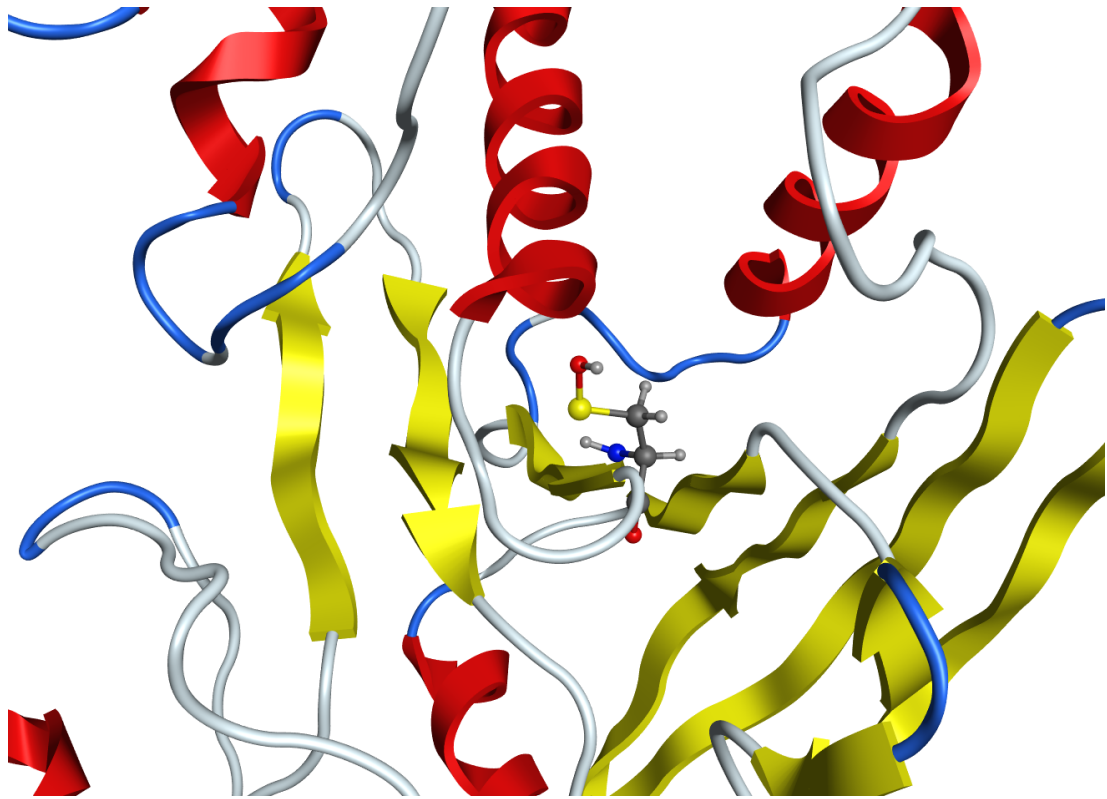
Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, 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.; and Pople, J. A.; Gaussian, Inc., Wallingford CT, 2004.

(42) Gaussian 09, Revision D.01, 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. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; 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, Ö.;

- Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009. (43) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- (44) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (45) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421
- (46) Noodleman, L.; Lovell, T.; Han, W.-G.; Li, J.; Himo, F. *Chem. Rev.* **2004**, *104*, 459.
- (47) Bushnell, E. A. C.; Erdtman, E.; Llano, J.; Eriksson, L. A.; Gault, J. W. *J. Comput. Chem.* **2010**, *32*, 822.
- (48) Bushnell, E. A. C.; Huang, W.; Llano, J.; Gault, J. W. *J. Phys. Chem. B* **2012**, *116*, 5205.
- (49) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- (50) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.
- (51) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (52) Przybylski, J. L.; Wetmore, S. D. *Biochemistry* **2011**, *50*, 4218.
- (53) Razmisleviciene, I.; Baltuskonyte, R.; Padarauskas, A.; Naujalis, E. *Chemija* **2008**, *19*, 33.
- (54) Lim, V. I.; Curran, J. F.; Garber, M. B. *J. Mol. Biol.* **2005**, *351*, 470.
- (55) Lim, V. I.; Kljashtorny, V. G. *Mol. Biol.* **2006**, *40*, 572.

Chapter 10

Conclusions



In this thesis the redox chemistry of sulfur in several enzymes has been investigated using various computational approaches. Particularly, previously proposed mechanisms for sulfenic acids formation and reduction/overoxidation pathways have been examined.

In Chapter 3, the catalytic mechanism of MsrA in *Mycobacterium tuberculosis*, in which S-methionine sulfoxide (Met-O) is reduced to Methionine (Met), has been investigated using docking, molecular dynamics (MD) simulations and ONIOM (QM/MM) methods. In addition, the roles of specific active site residues including an aspartyl (Asp87) near the recycling cysteine, tyrosyl's (Tyr44 and Tyr92) and glutamyl (Glu52) have been examined, as well as the general effects of the protein and active site on the nature and properties of mechanistic intermediates. The mechanism is initiated by proton transfer from the catalytic cysteine's thiol (Cys13SH) via a bridging water to the R-group carboxylate of Glu52. The now anionic sulfur of Cys13 nucleophilically attacks the substrate's sulfur with concomitant proton transfer from Glu52 to the sulfoxide oxygen, generating a sulfurane. The active site enhances the proton affinity of the sulfurane oxygen which, can readily accept a proton from the phenolic hydroxyls of Tyr44 or Tyr92 to give a sulfonium cation. Subsequently, Asp87 and the recycling cysteine (Cys154) can facilitate nucleophilic attack of a solvent water at the Cys13S center of the sulfonium to give a sulfenic acid (Cys13SOH) and Met. For the subsequent reduction of Cys13SOH with intramolecular disulfide bond formation Asp87 can help facilitate nucleophilic attack of Cys154S at the sulfur of Cys13SOH by deprotonating its thiol. This reduction is found to likely occur readily upon suitable positioning of the active site hydrogen bond network and the sulfur centers of both Cys13 and Cys154. The calculated rate-limiting barrier is in good agreement with experiment.

In Chapter 4, the last step in the reductase step of the catalytic mechanism of MsrB (MsrA sister enzyme) was re-investigated using several computational approaches. Our previous DFT cluster paper showed that two possible mechanisms could occur, however

the direct formation of disulfide from sulfonium cation was favored over sulfenic acid formation. In contrary, experimental studies suggest sulfenic acid formation. Therefore, first, we investigated the effect of level of theory, which confirmed previous conclusion. In addition, the effect of model choice was also investigated using ONIOM including a large QM layer around Cys440. Interestingly, deprotonating Cys440 leads to direct nucleophilic attack on Cys495 forming disulfide. Second, to eliminate the possibility that all previous results are an artifact of the used crystal structure in which the S...S distance is 3.29 Å, we ran a 5 ns MD simulation on the sulfonium cation intermediate. Surprisingly, our results show that the distance between the two sulfur is significantly increased to 4.88 Å. More importantly a water molecule is located in a proper position for nucleophilic attack. QM/MM calculations shows that sulfenic acid is formed via low barrier of 16.7 kJ mol⁻¹. Finally, the effect of substrate binding on the two Cys's distance were investigated via running several MD simulations of possible intermediates, showing that substrate binding induces conformational changes increasing the sulfur's distance which is decreased upon substrate removal upon sulfenic acid formation. These results question the applicability of QM cluster approach in systems including flexible turns. It also emphasizes the importance of proper preparation of the starting structure.

Both Chapters 3 and 4, show that the same mechanism occurs in MsrA and MsrB despite their active sites structural differences. As future direction, it would be interested to investigate the mechanism in the 1-Cys Msr, which contains only one Cys residue in the active site, particularly the stabilization mechanism of sulfenic acid in the active site against overoxidation. Furthermore, in human, MsrB reduction mechanism uses an active site selenocysteine (Sec) residue. Therefore, mechanistic difference between Cys and Sec in MsrB could be revealed.

In Chapter 5, Docking and Molecular dynamic simulations were synergistically used to elucidate the mechanism of MsrA activation via small molecules. Experimentally, Pergolide (Perg) and Pergolide-sulfoxide (PergSO) have been shown to increase the

activity of MsrA via an unknown mechanism. In this study we investigated the possibility of MsrA induction via direct molecular interaction. Our docking calculations show that Perg and PergSO as well as S-adenosyl methionine (SAM) bind to a novel allosteric site of the oxidized form of the enzyme. MD simulations have also shown that the binding of these molecules disrupts the interaction between the Gln122-Thr132 segment and C-terminal residues facilitating the binding of thioredoxin and the regeneration step in the mechanism. These findings represent for the first time a direct mechanism to activate MsrA providing a novel therapeutic pathway. Furthermore, virtual screening was used to screen the subset ZIM from the drug database ZINC against our identified allosteric site. Several ligands were identified to bind with high affinity to MsrA. Several following studies are needed to verify the proposed activation mechanism. For instance, the direct interaction hypothesis could be tested experimentally. Furthermore, identified possible ligands can also be tested to select lead compounds. Subsequently, quantitative structural relationship studies could be used to optimize the lead ligands.

In Chapter 6, the formation mechanism of sulfenyl-amide intermediate from sulfenic acid in protein tyrosine phosphatase 1B (PTP1B) has been investigated. Using a wide array of computational modeling techniques including MD simulations and high-level ONIOM calculations possible catalytic pathways has been investigated. Several potential reactive complexes were considered in the calculations. The only feasible mechanism was found to occur in a stepwise fashion, in which a stable iminol intermediate is formed. This step has an activation energy of 48.6 kJ mol^{-1} . Later, a much more stable iminol intermediate is formed in which a noncovalent electrostatic interaction of the sulfenic acid sulfur anti-bonding orbital with the iminol nitrogen lone pair was found to occur. Subsequently, a cyclic sulfenyl-amide is formed with a concomitant proton transfer from Glu115 to the sulfenic acid oxygen. Our results suggest that Glu115 and His214 play a crucial role in the mechanism. These results could contribute to the discovery of PTP1B inhibitors and the stabilization of the enzyme oxidized form.

In Chapter 7, mechanistic details for the reduction of hydrogen peroxide and the formation of a proposed hypervalent sulfurane species in ApTPx have been elucidated via the synergistic use of MD simulations and ONIOM. Our results show that the reduction mechanism is driven by the destabilization of the peroxide substrate in the active site, facilitating the nucleophilic attack from the peroxidative cysteine. Following the reduction of hydrogen peroxide, a pseudo hypervalent sulfurane intermediate is suggested to form with a weak interaction between His42 and the oxidized Cys50. This interaction was confirmed using QTAIM and NBO analysis. Indeed, the formation of the sulfurane was found to be energetically favorable over the overoxidized sulfinic acid species and thus would provide a mechanism protecting sulfenic acid from overoxidation. A similar intermediate was discovered for the first time in the human 1-Cys hORF6. Moreover, by comparing X-ray structures with several different Prxs and other proteins we suggest this intermediate may be generally formed when sulfenic acid is involved in protein chemistry. This can be tested via synergistic use of computational and experimental techniques to elucidate sulfenic acid interactions in PrxV and PrxIV as well as other proteins such as SUMO proteases and Gsp amidase.

In Chapter 8, the reduction of typical 2-Cys peroxiredoxins (Prxs) sulfinic acid to sulfenic acid via a repair enzyme known as sulfiredoxin (Srx) has been investigated. Molecular modeling techniques including MD simulations (MD) and ONIOM approach were used to elucidate the atomistic details of this unique reaction in sulfur chemistry. Our results support the previous experimentally proposed mechanism in which the sulfinic acid oxygen perform an in line direct nucleophilic attack on the γ -phosphate of ATP forming sulfinic acid phosphoryl ester intermediate and ADP, via a low barrier of 16.3 kJ mol^{-1} . Subsequently, the formed intermediate is directly reduced via an S_N2 mechanism by the Srx-Cys99 forming thiosulfinate. Our results suggest that the rate-limiting step of the reduction mechanism is associated with the reduction step of the thiosulfinate intermediate. Notably, this work improves the current knowledge of this

unique reaction, which could contribute to the discovery of new groups of antioxidants capable of reducing this irreversible overoxidized state in other proteins.

In Chapter 9, the maleate isomerase (MI) catalyzed isomerization of maleate to fumarate has been investigated using a wide range of computational modeling techniques, including small model DFT calculations, QM-cluster approach, ONIOM and MD simulations. Several fundamental questions regarding the mechanism were answered in detail, such as the activation and stabilization of the catalytic Cys in a rather hydrophobic active site. The two previously proposed mechanisms were considered, where either enediolate or succinyl-Cys intermediate forms. Small model calculations as well as an ONIOM-based approach suggest that an enediolate intermediate is too unstable. Furthermore, the formation of succinyl-Cys intermediate via the nucleophilic attack of Cys76⁻ on the substrate C2 (as proposed experimentally) was found to be energetically unfeasible in both QM-cluster and ONIOM approaches. Instead, our results show that Cys194, upon activation via the substrate, acts as a nucleophile and Cys76 acts as an acid/base catalyst, forming a succinyl-Cys intermediate in a concerted fashion. Indeed, the calculated PA of Cys76 is always higher than that of Cys194 before or upon substrate binding in the active site. Furthermore, the mechanism proceeds via multiple steps by substrate rotation around C2–C3 with the assistance of the now negatively charged Cys76, leading to the formation of fumarate. Finally, our calculated barrier is in good agreement with experiment. These findings represent a novel mechanism in the racemase superfamily.

With the fast growing number of X-ray identified sulfenic acids in proteins, more computational studies are needed to further elucidate sulfenic acid chemistry specially their protective mechanism and interactions. In addition, the role of non-covalent interactions in stabilizing sulfur species' needs further investigation. In addition the transferability of any of our studied mechanism to other proteins would improve our understanding of these reactions. Although our studies do not discuss

all aspect of sulfur/sulfenic chemistry, it gave some important insights that enrich our knowledge about sulfur chemistry.

Appendix C



RightsLink®

[Home](#)
[Create Account](#)
[Help](#)


ACS Publications
Most Trusted. Most Cited. Most Read.

Title: A Molecular Dynamics and Quantum Mechanics/Molecular Mechanics Study of the Catalytic Reductase Mechanism of Methionine Sulfoxide Reductase A: Formation and Reduction of a Sulfenic Acid

Author: Hisham M. Dokalnish, James W. Gauld

Publication: Biochemistry

Publisher: American Chemical Society

Date: Mar 1, 2013

Copyright © 2013, American Chemical Society

User ID
<input type="text"/>
Password
<input type="text"/>
<input type="checkbox"/> Enable Auto Login
LOGIN
Forgot Password/User ID?
If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)
[CLOSE WINDOW](#)

Copyright © 2014 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#).
Comments? We would like to hear from you. E-mail us at customer@copyright.com

Computational investigations on the catalytic mechanism of maleate isomerase: the role of the active site cysteine residues

H. M. Dokainish, B. F. Ion and J. W. Gauld, *Phys. Chem. Chem. Phys.*, 2014, **16**, 12462
DOI: 10.1039/C4CP01342E

If you are not the author of this article and you wish to reproduce material from it in a third party non-RSC publication you must [formally request permission](#) using RightsLink. Go to our [Instructions for using RightsLink page](#) for details.

Authors contributing to RSC publications (journal articles, books or book chapters) do not need to formally request permission to reproduce material contained in this article provided that the correct acknowledgement is given with the reproduced material.

Reproduced material should be attributed as follows:

- For reproduction of material from NJC:
Reproduced from Ref. XX with permission from the Centre National de la Recherche Scientifique (CNRS) and The Royal Society of Chemistry.
- For reproduction of material from PCCP:
Reproduced from Ref. XX with permission from the PCCP Owner Societies.
- For reproduction of material from PPS:
Reproduced from Ref. XX with permission from the European Society for Photobiology, the European Photochemistry Association, and The Royal Society of Chemistry.
- For reproduction of material from all other RSC journals and books:
Reproduced from Ref. XX with permission from The Royal Society of Chemistry.

If the material has been adapted instead of reproduced from the original RSC publication "Reproduced from" can be substituted with "Adapted from".

In all cases the Ref. XX is the XXth reference in the list of references.

If you are the author of this article you do not need to formally request permission to reproduce figures, diagrams etc. contained in this article in third party publications or in a thesis or dissertation provided that the correct acknowledgement is given with the reproduced material.

Reproduced material should be attributed as follows:

- For reproduction of material from NJC:
[Original citation] - Reproduced by permission of The Royal Society of Chemistry (RSC) on behalf of the Centre National de la Recherche Scientifique (CNRS) and the RSC
- For reproduction of material from PCCP:

- [Original citation] - Reproduced by permission of the PCCP Owner Societies
- For reproduction of material from PPS:
[Original citation] - Reproduced by permission of The Royal Society of Chemistry (RSC) on behalf of the European Society for Photobiology, the European Photochemistry Association, and RSC
 - For reproduction of material from all other RSC journals:
[Original citation] - Reproduced by permission of The Royal Society of Chemistry

If you are the author of this article you still need to obtain permission to reproduce the whole article in a third party publication with the exception of reproduction of the whole article in a thesis or dissertation.

Information about reproducing material from RSC articles with different licences is available on our [Permission Requests page](#).

MDPI Journals A-Z For Authors For Editors For Librarians About Open Access Policy Submit to IJMS Login Register

International Journal of Molecular Sciences

Title / Keyword: Journal: Volume:
 Author: Section: Issue:
 Article Type: Special Issue: Page:

IMPACT FACTOR 2.339

IJMS Volume 15, Issue 1 *Int. J. Mol. Sci.* **2014**, *15*(1), 401-422; doi:10.3390/ijms15010401 [Open Access](#)

- Article Versions**
- Abstract
- Full-Text HTML
- Full-Text PDF [858 KB]
- Full-Text XML
- Article Versions Notes
- Related Info**
- Article Statistics
- PubMed/Medline
- Google Scholar
- Order Reprints
- More by Authors**
- [+] on DOAJ
- [+] on Google Scholar
- [+] on PubMed
- Share This Article**
- CiteULike
- Facebook
- Mendeley
- Twitter
- More services

Review
Multi-Scale Computational Enzymology: Enhancing Our Understanding of Enzymatic Catalysis

Rami Gherib^{1,2}, Hisham M. Dokainish^{1,2} and James W. Gauld^{1,2}

Authors' affiliation

Received: 11 November 2013; in revised form: 5 December 2013 / Accepted: 24 December 2013 / Published: 31 December 2013

(This article belongs to the collection Computational, Structural and Spectroscopic Studies of Enzyme Mechanisms, Inhibition and Dynamics)

This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
PDF Full-Text
Citations to this Article
Article Metrics
Cite This Article

Abstract: Elucidating the origin of enzymatic catalysis stands as one of the great challenges of contemporary biochemistry and biophysics. The recent emergence of computational enzymology has enhanced our atomistic-level description of biocatalysis as well as the kinetic and thermodynamic properties of their mechanisms. There exists a diversity of computational methods allowing the investigation of specific enzymatic properties. Small or large density functional theory models allow the comparison of a plethora of mechanistic reactive species and divergent catalytic pathways. Molecular docking can model different substrate conformations embedded within enzyme active sites and determine those with optimal binding affinities. Molecular dynamics simulations provide insights into the dynamics and roles of active site components as well as the interactions between substrate and enzymes. Hybrid quantum mechanical/molecular mechanical (QM/MM) can model reactions in active sites while considering steric and electrostatic contributions provided by the surrounding environment. Using previous studies done within our group, on OvoA, EglB, ThrRS, LuxS and MsrA enzymatic systems, we will review how these methods can be used either independently or cooperatively to get insights into enzymatic catalysis.

Keywords: enzyme catalysis; density functional theory (DFT) cluster method; quantum mechanics/molecular mechanics (QM/MM); molecular dynamics (MD) simulations; molecular docking

Int. J. Mol. Sci. EISSN 1422-0067 Published by MDPI AG, Basel, Switzerland [RSS](#) [E-Mail Table of Contents Alert](#)

[Terms & Conditions](#)
[Privacy Policy](#)
[Contact MDPI](#)
[Payment Information](#)
[Jobs at MDPI](#)
 © 1996-2014 MDPI AG (Basel, Switzerland) unless otherwise stated

The applicable mediation rules will be designated in the copyright notice published with the work, or if none then in the request for mediation. Unless otherwise designated in a copyright notice attached to the work, the UNCITRAL Arbitration Rules apply to any arbitration.

[More info.](#)

You may also use a license listed as compatible at <https://creativecommons.org/compatiblelicenses>

[More info.](#)

A commercial use is one primarily intended for commercial advantage or monetary compensation.

[More info.](#)

Merely changing the format never creates a derivative.

[More info.](#)

VITA AUCTORIS

Name: Hisham Mohammed Mohammed Dokainish

Place of Birth: Al Basrah, Iraq

Year of birth: 1981

Education :

PhD Chemistry October 2014

University of Windsor Windsor, Ontario, Canada

Master's in Chemistry April 2010

University of New Brunswick Fredericton, New Brunswick, Canada

Bachelor of Pharmacy May 2004

Helwan University Cairo, Egypt